

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ



Proyecto “Fortalecimiento de la Infraestructura Tecnológica para Procesos de Investigación del IIAP”

Programa BIOINFO - Instituto de Investigaciones de la Amazonía Peruana - IIAP



2017

Isaac Ocampo Yahuarcani
Luis Exequiel Campos Baca



PERÚ

Ministerio
del Ambiente



CONCYTEC
CONSEJO NACIONAL DE CIENCIA
TECNOLOGÍA E INNOVACIÓN TECNOLÓGICA



CIENCIAACTIVA

Scholarships and Grants by Concytec

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Isaac Ocampo Yahuarcani

Luis Exequiel Campos Baca

**Proyecto "Fortalecimiento de la Infraestructura
Tecnológica para Procesos de Investigación del IIAP"**

Instituto de Investigaciones de la Amazonía Peruana - IIAP
Programa de Investigación en Información de la Biodiversidad
Amazónica - BIOINFO

IQUITOS, 2017

Introducción a la supercomputación en el Perú.

ISBN: 978-9972-667-99-2

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2017-08981

Primera edición, julio de 2017.

© Instituto de Investigaciones de la Amazonía Peruana - IIAP

Programa de Investigación en Información de la Biodiversidad Amazónica - BIOINFO

Av. Abelardo Quiñones KM 2,5 – Iquitos

Telef: +51 65 265515

<http://www.iiap.org.pe>

Obra realizada con el financiamiento de Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC)

Fondo Nacional de Desarrollo Científico, Tecnológico y de Innovación Tecnológica - CIENCIACTIVA

Calle Schell 459, Miraflores - Lima

Telef: +51 1 6440004

<http://www.cienciaactiva.gob.pe>

Autores:

Isaac Ocampo Yahuarcani

Luis Exequiel Campos Baca

Colaboradores:

Daniel Díaz Ataucuri (INICTEL-UNI), Milton Ríos Julcapoma (INICTEL - UNI), Hugo Enrique Hernández Figueroa (UNICAMP-Brasil), Maicon Farias (Empresa ABAX - Brasil), Alex Jesús Cuadros Vargas (Universidad Católica San Pablo), Alexander Víctor Ocsa Mamani (UNSA), Jose Alfredo Herrera Quispe (UNSA), Hernán Tello Fernandez (MINAM), Rosanna Patrógic Rengifo (Empresa Tecnosys), Andrés Yong Hurtado (Empresa Tecnosys), Rodolfo Cárdenas Vigo (IIAP), Pedro Roberto García Vela (IIAP), Carlos García Cortegano (FISI-UNAP), Lelis Antony Saravia Llaja (FISI-UNAP), Franz Ghusept Chuje Panaifo (FISI-UNAP), Aurora Soledad Guerra Ugas (FISI-UNAP), Deivis Deksther Fernandez Meza (FISI-UNAP), Junior Antonio Cerna Ayapi (FISI-UNAP).

Cuidado de la Edición y de los Textos

Juan José Bellido Collahuacho, Manuel Martín Brañas, Isaac Ocampo Yahuarcani, Lelis Antony Saravia Llaja.

Carátula: Isaac Ocampo Yahuarcani, Lelis Antony Saravia Llaja, Franz Ghusept Chuje Panaifo.

Foto de Portada: Nodos de la Supercomputadora MANATI del IIAP.

Se imprimieron 500 ejemplares en :

FABRIGRAF S.A.C., Av. Argentina 144, Primer piso Int. Bk-28, Lima, Tel. 01 478 5584.

Julio de 2017

Impreso en Perú.

CONTENIDO

Presentación	9
Introducción	11
CAPÍTULO 1: Introducción a la supercomputación	15
1.1. Supercomputación o HPC	15
1.2. Supercomputadora o computador de alto rendimiento - HPC	17
1.3. Los FLOPS como medida de rendimiento de un computador	17
1.4. Conceptos básicos para entender la supercomputación	17
1.5. ¿En qué casos no es recomendable comprar una supercomputadora?	19
1.6. Algunas interrogantes para trabajar con supercomputadoras	20
1.7. Sistemas de alto rendimiento computacional	21
1.8. Capacidades a tener en cuenta en una supercomputadora	22
1.9. Historia de la supercomputación	23
1.10. International Business Machines (IBM)	29
1.11. El genio de Seymour Cray	34
1.12. Computadoras más veloces según el Top500	37
1.13. ¿Para qué usar una supercomputadora?	43
CAPÍTULO 2: Elementos de un sistema de alto rendimiento computacional: computadoras, sistema eléctrico, sistema de enfriamiento y otros complementos.	45
2.1. Equipamiento para alto rendimiento computacional	46
2.2. Sistema de refrigeración	77
2.3. Sistema eléctrico	81
2.4. Adecuaciones de ambiente	89
CAPÍTULO 3: Elementos de un sistema de alto rendimiento computacional: Software	91
3.1. Sistemas operativos	91
3.2. Software para gestión de supercomputadoras	95
3.3. Software para desarrollo de aplicaciones y cálculo numérico en supercomputadoras	100
3.4. APIs y Librerías	106
3.5. Software especializado	109

CAPÍTULO 4: La Supercomputación en el mundo	115
4.1. La Supercomputación a noviembre del 2016	115
4.2. Supercomputación en Latinoamérica	123
4.3. Centros de alto rendimiento computacional en Latinoamérica	125
4.4. Expertos referentes de computación de alto rendimiento	139
CAPÍTULO 5: La Supercomputación en el Perú	144
5.1. Centros de alto rendimiento computacional en Perú	145
5.2. Centros de formación de recursos humanos en el Perú	153
5.3. Proyectos de investigación sobre computación de alto rendimiento	160
5.4. Expertos, investigadores y usuarios de supercomputación en el Perú	165
CAPÍTULO 6: Primeros pasos en programación para computación de alto rendimiento	177
6.1. Ejemplos para sistema operativo Windows	177
6.2. Ejemplos para sistema operativo Linux	192
CAPÍTULO 7: Agenda de proyectos del Instituto de Investigaciones de la Amazonía Peruana – IIAP relacionadas a computación de alto rendimiento	207
7.1 Fortalecimiento de infraestructura tecnológica para procesos de investigación del Instituto de Investigaciones de la Amazonía Peruana	207
7.2 Circulo de Investigación en computación de alto rendimiento, para el desarrollo de métodos y técnicas de minería de datos de gran escala como soporte en investigaciones en cambio climático	213
7.3 Evaluación del estado ecosistémico de los aguajales en Iquitos empleando imágenes de alta resolución, computación de alto rendimiento y procesamiento digital de imágenes, que facilite la conservación y uso sostenible de la biodiversidad en la Amazonía	215
7.4 Reconocimiento de patrones de hojas de plantas de importancia económica en imágenes digitales utilizando redes neuronales	218
7.5 Procesamiento de imágenes satelitales para deforestación, monitoreo, tala ilegal, inundaciones, dinámica de ríos, sedimentación, cambio climático	218
7.6 Alcance del programa BIOINFO en proyectos relacionados con HPC	219
REFERENCIAS BIBLIOGRÁFICAS	221
ANEXOS	229
ANEXO 01: Directorio de centros de alto rendimiento computacional de Sudamérica	230
ANEXO 02: Listado de las 100 primeras supercomputadoras del Top500 a noviembre del 2016	235
ANEXO 03: Acceso de usuarios al MANATI IIAP	241
ANEXO 04: Cálculo de FLOPS por computadora	243

ÍNDICE DE TABLAS

Tabla 01: Capacidades en GIGAFLOPS de diversos dispositivos	23
Tabla 02: Equivalencias del Nodo supercomputador IIAP frente a diversos dispositivos	23
Tabla 03: Cuadro de resumen de la historia de la supercomputación	33
Tabla 04: Lista de las Supercomputadoras más veloces según el Top500 (desde junio de 1993 hasta noviembre de 2016)	42
Tabla 05: Especificaciones de NVIDIA Tesla K80	59
Tabla 06: Especificaciones de NVIDIA Titan Xp	60
Tabla 07: Especificaciones de GeForce GTX	60
Tabla 08: Especificaciones del Quadro GP100	61
Tabla 09: Especificaciones del procesador Tegra X1	61
Tabla 10: Especificaciones del procesador Tegra K1	62
Tabla 11: Especificaciones de Tegra 4	63
Tabla 12: Especificaciones Tesla P100	63
Tabla 13: Especificaciones AMD Embedded Radeon E9260 MXM Module	65
Tabla 14: Especificaciones AMD Embedded Radeon E8870 MXM Module	65
Tabla 15: Especificaciones AMD Embedded Radeon E8860 GPU MCM Module	65
Tabla 16: Características de Acelerador Pezy-1	67
Tabla 17: Ancho de banda teórico de INFINIBAND en GB/s	68
Tabla 18: Lista de Sistemas Operativos para computadoras de alto rendimiento	91
Tabla 19: Software para gestión de supercomputadoras	95
Tabla 20: Software para desarrollo de aplicaciones y cálculo numérico en HPC	100
Tabla 21: Lista de APIs y Librerías	106
Tabla 22: Lista de Software especializados	109
Tabla 23: Las 25 Supercomputadoras de mayor capacidad del mundo según el Top500 a noviembre de 2016	117
Tabla 24: Las 3 Supercomputadoras Latinoamericanas en el Top500 de noviembre del 2016	124
Tabla 25: Supercomputadoras latinoamericanas en el Top500 en noviembre de 2001	125
Tabla 26: Cuadro de universidades cuyas carreras de informática incluyen algunas asignaturas relacionadas a computación de alto rendimiento	154
Tabla 27: Listado de banco de Tesis de la Maestría en Informática de la UNSA Arequipa	156
Tabla 28: Proyectos de Tesis de la Maestría de Ciencia de la Computación de la UCSP de Arequipa	158
Tabla 29: Cantidad de proyectos identificados por institución	161
Tabla 30: Proyectos identificados por región	162
Tabla 31: Proyectos identificados por líneas de investigación	163
Tabla 32: Proyectos de computación de alto rendimiento en Perú	163
Tabla 33: Lista de usuarios del MANATI IIAP hasta marzo del 2017	175
Tabla 34: Instrucciones por ciclo (IPC) de algunas familias de procesadores	248

PRESENTACIÓN

Desde el año 1982, el Instituto de Investigaciones de la Amazonía Peruana (IIAP) viene trabajando para lograr el desarrollo y sostenibilidad de la región amazónica en el Perú. Durante todos estos años ha liderado una serie de estudios e investigaciones dirigidos a lograr este propósito, entre los que destacan aquellos relacionados con los procesos de Zonificación Ecológica Económica (ZEE) en la mayor parte de las regiones amazónicas, estudios de biodiversidad de fauna y flora, investigaciones sobre recursos hídricos, estudios ecosistémicos sobre los bosques, entre otros.

El año 2014, el IIAP con la valiosa colaboración de Instituto Nacional de Investigación y Capacitación de Telecomunicaciones del Perú de la Universidad Nacional de Ingeniería (INICTEL UNI), plantearon un proyecto cuyo objetivo fue la mejora de las capacidades relacionadas a infraestructura computacional orientada a potencializar los estudios científicos asociados al monitoreo del bosque (deforestación y uso de territorio), modelamiento y simulaciones de eventos climáticos, bioinformática, química computacional, entre otros. En este contexto fue clave, el apoyo del Fondo CIENCIACTIVA de CONCYTEC, que financió el mencionado proyecto. Uno de los principales resultados fue la creación del Centro de Alto Rendimiento Computacional de la Amazonía Peruana; un laboratorio científico que promete revolucionar la investigación en la Amazonía y, a la vez, generar servicios en tiempo real que permitan el monitoreo del bosque, los recursos hídricos y la biodiversidad.

Siendo el IIAP una entidad pública, adscrita al Ministerio del Ambiente, cuyo máximo órgano de Gobierno es su Consejo Superior, conformado por 25 instituciones donde convergen las universidades amazónicas, los gobiernos regionales, representantes de las comunidades indígenas, iglesia católica y organizaciones como el Instituto Nacional de Innovación Agraria (INIA) e Instituto Veterinario de Investigaciones Tropicales y de Altura de la Universidad Nacional Mayor de San Marcos de Lima (IVITA-UNMSM). El Consejo Superior es un espacio interinstitucional apropiado para promover la socialización, expansión y uso de equipamiento de alto rendimiento computacional, lo que permitirá llegar con estas tecnologías hacia diversos actores y beneficiarios. Por otro lado, a la fecha se han logrado convenios con la Universidad Nacional de la Amazonía Peruana (UNAP) y se viene preparando acuerdos con las universidades amazónicas como Universidad Nacional de Ucayali (UNU), Universidad Nacional de San Martín (UNSM), Universidad Nacional Intercultural de la Amazonía (UNIA) y Universidad Nacional Autónoma de Alto Amazonas (UNAAA). De la misma manera se vienen desarrollando actividades conjuntas y facilitando acceso a usuarios de la Universidad Nacional de San Agustín de Arequipa (UNSA), Universidad Católica San Pablo de Arequipa (UCSP), Universidad Nacional de Ingeniería (UNI), Pontificia Universidad Católica del Perú (PUCP), la Universidad de Ingeniería y Tecnología de Lima (UPEC), etc.

Es en este sentido que la presente publicación cobra especial interés, ya que se convierte en un material didáctico de primer orden para comprender los conceptos de supercomputación a partir de un proyecto concreto, los alcances que tiene y los retos que se nos presentan a la hora de incursionar en este campo del saber. Esperamos que este libro sea de su interés y sirva para adentrarlos en el desafiante mundo de la supercomputación.

Luis Exequiel Campos Baca
Presidente del IIAP

INTRODUCCIÓN

Desde el año 2001 el IIAP viene trabajando en la implementación de sistemas de información y otros mecanismos tecnológicos orientados a la reducción de las diversas brechas existentes en la Amazonía peruana. Estos instrumentos fueron creados para apoyar procesos de acopio, sistematización y acceso a la información de la biodiversidad amazónica y el ambiente entre ellos destacan: Sistema de Información de la Diversidad Biológica y Ambiental de la Amazonía Peruana (SIAMAZONIA), Sistema de Información Forestal de la Amazonía Peruana (SIFORESTAL), Sistema de Información Turística de Loreto (SITURISMO), Centro de Promoción de la Biodiversidad Amazónica (PROMAMAZONIA), entre otros.

El año 2009 el IIAP a partir de la implementación de su Plan Estratégico 2009-2018 contempló la creación del Programa de Investigación en Información de la Biodiversidad Amazónica (BIOINFO), creado como un espacio de generación investigaciones e innovaciones utilizando las tecnologías de la información y comunicación para el desarrollo de la Amazonía Peruana.

Los siguientes pasos necesariamente requieren la expansión de los sistemas de información hacia el campo, directamente al bosque, dejar las salas de servidores y llegar directamente a ríos, árboles, animales y todos los ecosistemas. Las interrogantes siguen siendo las mismas, sin embargo, las estrategias utilizadas en el Perú siempre fueron inapropiadas e insuficientes. Ello a pesar de que se conoce que el Perú es uno de los países más vulnerables frente al cambio climático en el mundo.

Nuestro país no cuenta con suficientes redes electrónicas de monitoreo (aire, agua, bosque, biodiversidad), aún estamos dando los pasos iniciales en el uso de satélites, radares y drones para el estudio de nuestra geografía peruana. Actualmente las consecuencias de no haber masificado estas tecnologías significan un alto precio para todos los peruanos, luego de haber sufrido los embates del fenómeno del Niño Costero, actualmente ciudades como Piura, Chiclayo o Trujillo están considerablemente destruidas y recién se habla de proyectos de monitoreo, prevención y pronósticos. Los efectos y consecuencias de las inundaciones demorarán muchos años en desaparecer.

En este camino el IIAP estableció una fructífera alianza con INICTEL-UNI, para trabajar en temas de redes de sensores de monitoreo de agua (pequeñas estaciones de bajo costo), además de proyectos de procesamiento de imágenes de teledetección en el bosque amazónico. A partir de estos proyectos y con toda la experiencia del IIAP en sistemas de información surgieron necesidades de capacidad procesamiento computacional o supercomputación.

En este contexto destaca la computación de alto rendimiento o supercomputación (también llamado HPC), como un instrumento capaz de generar respuestas para la apropiada prevención y gestión de los desastres naturales. Una supercomputadora es una sinergia o conjunto de procesadores conectados entre sí para procesar grandes volúmenes de información. La idea central es la realización de la mayor cantidad de operaciones computacionales en el menor tiempo posible usando la menor cantidad de recursos, es decir se busca la realización de procesos lo más eficientemente posibles.

Según el último ranking Top500 de noviembre del 2016 [1], China lidera el ranking mundial de Supercomputadoras con el equipo denominado Tianhe-2 (MilkyWay-2) ubicado en el National University of Defense Technology (Changsha, China) que supera 93 petaflops. Históricamente en la supercomputación ha existido hegemonía de los Estados Unidos (la última vez que lideró el Top500 fue con el Supercomputador Cray en el año 2012), sin embargo desde el año 2013 China viene consolidándose como líder mundial. Así mismo a nivel de Latinoamérica Brasil lleva el liderazgo con el supercomputador Santos Dumont GPU (en el puesto 364 del Top500), que funciona en el Laboratório Nacional de Computação Científica de Río de Janeiro.

En enero del 2017 el IAP, en su sede central de Iquitos, puso en servicio el Centro de Alto Rendimiento Computacional de la Amazonía Peruana, que es un laboratorio para procesamiento de grandes volúmenes de información (procesamiento paralelo gráfico y numérico), que cuenta con una clúster de supercomputación conformado por 3 nodos para cálculo numérico (Procesadores Intel Xeon E2680) y 6 nodos para procesamiento gráfico (aceleradores NVIDIA K80). Este sistema está disponible y con libre acceso para todos los peruanos interesados.

Actualmente los centros de alto rendimiento computacional del Perú se cuentan con los dedos de las manos, e incluso la mayor parte de ellos vienen implementándose desde mediados del 2015. En este punto es destacable el apoyo de entidades como el Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC) y el Programa Nacional de Innovación para la Competitividad y Productividad (INNOVATE Perú) que facilitaron el financiamiento para su implementación a través de fondos concursables. Sin embargo aún falta mucho camino que recorrer, pues en países como Brasil o México implementar infraestructura de alto rendimiento computacional es clave y muy necesario en el desarrollo de la ciencia y de la propia industria. Brasil cuenta con al menos 57 centros de alto rendimiento computacional, financiados por las propias universidades, entidades del estado, empresa privada, incluso con fondos de cooperación. En el Perú apenas logramos identificar 7 centros de alto rendimiento (4 en Lima, 1 en Arequipa, 1 en Iquitos y el séptimo se viene instalando en Ica) y nuestras instituciones aún no valoran las capacidades de procesamiento computacional, ello por una serie de motivos, a explicarse en las próximas páginas de este documento.

Emprender un proyecto de supercomputación en el Perú fue verdaderamente un reto muy desafiante y más complicado aún fue el hecho de instalarlo en una ciudad como Iquitos. Ello teniendo en cuenta el pequeño y débil ecosistema tecnológico nacional. Hasta ahora todos los elementos del mencionado ecosistema no favorecen este tipo de proyectos. Una de los principales limitantes es la escasez de empresas proveedoras especializadas en computación de alto rendimiento. La mayoría de las empresas consultadas ofrecían equipos de menor capacidad a altos precios, incluso equipamiento sin el sistema eléctrico o a la vez escaso soporte y capacitación. Otro factor limitante fue el trámite relacionado a las licitaciones públicas. Por otro lado la compra directa a un proveedor internacional, incluso incrementaba la complejidad, pues cualquier empresa extranjera no está muy interesado en someterse a los requisitos del sistema de adquisiciones peruana. En estos últimos puntos es destacable el apoyo de CONCYTEC que dispuso la modificación de la ley de adjudicaciones para que los equipamientos científicos puedan hacerse por adjudicación directa.

Otro factor crucial en este tipo de proyectos es la formación de recursos humanos. La mayoría de investigadores peruanos que usa supercomputadoras ha recibido formación posgrado en el extranjero, principalmente Brasil, México, Chile y Estados Unidos. En el Perú a pesar de existir más de 100 universidades con carreras de informática o sistemas, en menos del 5% reciben enseñanzas con

contenidos de computación de alto rendimiento. En este mismo punto, pero en relación a carreras de ciencias básicas como biología, física, química y matemática, prácticamente no se desarrollan capacidades para procesamiento de grandes volúmenes de datos y ello constituye una debilidad para el futuro de la ciencia en el Perú.

Otro aspecto que consideramos importante comentar, es que los profesionales que han aprendido a usar Supercomputadoras, no necesariamente tienen las capacidades para formular proyectos de implementación de centros de alto rendimiento. Este punto es crítico y se debe tener en cuenta, pues se han visto casos de proyectos en los que no participaron expertos en electricidad, ni se incluyó el presupuesto suficiente para instalaciones complementarias. Ello a la larga acarreará problemas en la implementación y sostenibilidad del supercomputador. Por eso debería ser obligatorio impartir contenidos de diseño de datacenters en las carreras de informática y sistemas.

El IIAP en los inicios del proyecto centró sus esfuerzos en procesamiento de información de simulaciones sobre cambio climático, deforestación, nichos ecológicos, bioinformática, inundaciones y dinámica fluvial. Sin embargo debido a las crecientes necesidades de procesamiento y a la escasa oferta nacional de capacidades de procesamiento ha puesto a disposición de diversos investigadores peruanos el acceso libre al MANATI IIAP. Actualmente el IIAP viene trabajando en un sistema de monitoreo en tiempo real para estimar deforestación y detectar actividades de tala ilegal, identificación de tipos de bosques, métodos de predicción de la temperatura a través de Deep learning (con el Círculo de Investigación en Cambio Climático a través de Computación de Alto Desempeño liderado por la UNSA), monitoreo de aguajales a través de imágenes hiperespectrales (en alianza con INICTEL-UNI). Además del convenio con la Universidad Nacional de la Amazonía Peruana (UNAP) para el desarrollo de capacidades de docentes y alumnos en computación de alto desempeño. Y finalmente se está trabajando en la firma de un convenio con la Universidad Católica San Pablo de Arequipa (UCSP) para desarrollar trabajos conjuntos en procesamiento de imágenes de teledetección en el bosque (inundaciones, cambio climático) a través de sus estudiantes de maestría y docentes.

La intención principal de este documento es proporcionar un punto de partida para facilitar información básica o introductoria, que permita generar en los lectores criterios mínimos para entender o involucrarse en procesos relacionados a computación de alto rendimiento. Teniendo en cuenta que en el Perú no existe material de consulta disponible que permita conocer el estado actual, experiencias, actores y primeros pasos para adentrarse en el mundo de la computación de alto desempeño.

Los pasos siguientes de la supercomputación en el Perú son desafiantes. Es necesaria la conformación de una red nacional de investigadores que promueva el tema e impulse proyectos de investigación útiles para el desarrollo de la ciencia e industria en el Perú. Además es urgente utilizar la computación de alto desempeño en la predicción y sistemas de alerta temprana de desastres ambientales y climáticos. Será importante involucrar a organizaciones como la Sociedad Peruana de Computación y las diversas universidades peruanas que ofrecen carreras de informática, sistemas y computación.

CAPÍTULO 1

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN

1.1. Supercomputación o HPC

A lo largo de la historia y evolución de las computadoras han surgido diversos conceptos para definir la supercomputación, también es llamada Computación de Alto Rendimiento o incluso Computación de Alto Desempeño, sus siglas para identificarla en inglés es HPC (High Performance Computing). Al ser este un documento introductorio no se pretende profundizar ni hacer comparaciones milimétricas sobre los significados y conceptos utilizados en diversas partes del mundo, sino que básicamente se pretende contribuir con los elementos básicos para entender y comenzar a trabajar en HPC.

Como definición resumida del concepto de “supercomputación” podemos catalogarla como una “sinergia de los microprocesadores”. La idea central se basa en la realización de la mayor cantidad de operaciones computacionales en el menor tiempo posible, usando la menor cantidad de recursos (tiempo, mano de obra, energía, etc.), es decir, se busca procesar grandes volúmenes de datos de la forma más eficiente posible.



Imagen 1: Nodos de procesamiento de la supercomputadora MANATI del IIAP - Iquitos.

La supercomputación es una rama de la ciencia de la computación que busca la generación de métodos y procedimientos para que los microprocesadores, junto a otros recursos computacionales, trabajen de forma articulada y lo más eficientemente posible, de manera que puedan satisfacer diversos requerimientos de cálculo, cómputo o procesamiento de información. Para ello se combinan una serie de subdisciplinas del conocimiento de las ciencias de la computación, entre las que destacan algoritmos, computación paralela, arquitectura de computadores y computación distribuida.

Entre los principales recursos computacionales que debemos tener en cuenta a la hora de hablar de supercomputación están los microprocesadores con sus núcleos (CPU), los aceleradores gráficos (GPU), memorias, discos de almacenamiento, bus de comunicaciones de alta velocidad, sistemas operativos y las aplicaciones informáticas con funcionalidades de procesamiento paralelo.

El portal especializado InsideHPC [2] indica que “High Performance Computing generalmente se refiere a la práctica de agregar potencia computacional de una manera que ofrezca un rendimiento mucho mayor que el que podría proporcionar una típica computadora de escritorio o estación de trabajo para resolver grandes problemas en la ciencia, la ingeniería o el negocio”.

Diego Darío Guerrero Fernández en su tesis Doctoral en la Universidad de Málaga en España [3], sostiene que la supercomputación “consiste básicamente en coordinar la utilización de grandes cantidades de recursos de cómputo para resolver problemas específicos”.

Según la Red Universitaria Nacional de Chile la Computación de Alto Rendimiento [4] “es una disciplina que se dedica a resolver problemas que, dada su alta complejidad computacional, así como por la gran abundancia de datos disponibles para ser analizados, no pueden ser resueltos en un computador de escritorio. Para lograr este objetivo la computación de alto rendimiento se apoya en equipos que poseen una mayor potencia y capacidad de cálculo, ya sea a través de computadores con múltiples CPU's y/o por grandes cantidades de computadores de múltiples CPU's conectados por redes de alta velocidad”.

El Dr. Fabián García Nocetti, de la UNAM de México, indica que la computación de alto rendimiento (HPC) consiste en el uso de procesamiento paralelo para ejecutar aplicaciones avanzadas de manera eficiente, confiable y rápida. El término se aplica en especial a los sistemas que superan el teraflops de procesamiento (10^{12}) [5].

En el portal web del Latin American Symposium on High Performance Computing [6] se menciona que la Computación de Alto Rendimiento (HPC) “constituye un área muy dinámica en la adopción y uso de recursos y nuevas tecnologías computacionales. Corresponde a un campo inherentemente interdisciplinario ya que las disciplinas en las que se utilizan estas herramientas se amplían día a día, teniendo que nutrirse tanto del *know-how* de los expertos en la rama puntual de la ciencia como también de las ciencias de la computación e ingenierías relacionadas”.

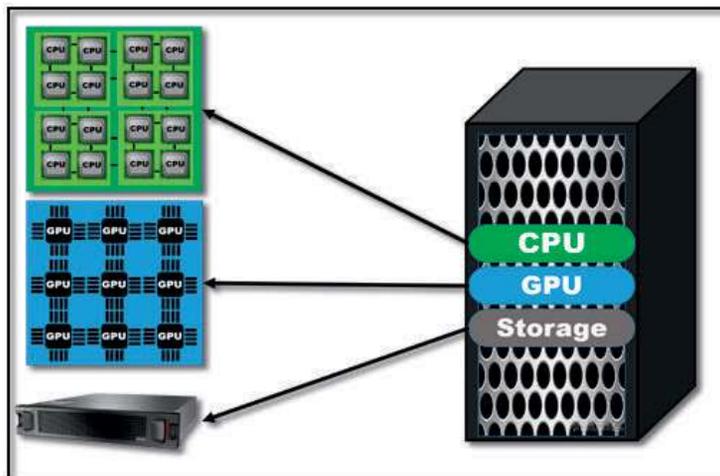


Imagen 2: Representación gráfica que integra una Unidad Central de Procesamiento (CPU), una Unidad de Procesamiento Gráfico (GPU) y una unidad de almacenamiento. Créditos: Elaboración propia.

1.2. Supercomputadora o computador de alto rendimiento - HPC

También llamado (High Performance Computer - HPC), es un computador o un conjunto de computadores conectados por comunicaciones de alta velocidad (clúster), cuyos procesadores funcionan de manera integrada y articulada para satisfacer grandes necesidades de cálculos numéricos, procesamientos de gráficos y otras diversas operaciones para las que son utilizadas las computadoras. Un HPC ofrece grandes capacidades computacionales para procesar grandes volúmenes de datos o contenidos.

La capacidad de procesamiento de un sistema computacional se mide en FLOPS, *floating point operations per second*, que significa “operaciones en coma flotante por segundo”. 1 gigaflops comprende 1000 millones de flops (10^9 flops) y es la unidad de medida del rendimiento de los celulares y computadores personales actuales.

Diversos autores indican que actualmente la HPC comienza a partir de 1 teraflops (10^{12} flops o 1000 gigaflops). Aunque en este caso, a finales del 2016, según el Top500 [1], la HPC de mayor rendimiento es la “Sunway TaihuLight” ubicada en China, que supera los 125 petaflops, unos 125 mil teraflops.

1.3. Los FLOPS como medida de rendimiento de un computador

El FLOPS es una medida de rendimiento de las computadoras [7]. Es uno de los términos que se usa para determinar que tan veloz es un computador para el procesamiento de datos. Es un acrónimo de *floating point operations per second*, se traduce como “operaciones o cálculos de coma flotante por segundo”. Este es un término acuñado en informática para designar a las operaciones matemáticas donde intervienen números tan grandes que es necesario expresarlas con notación científica o notación exponencial. En palabras sencillas son números que se expresan como un número real con un solo número entero (el resto son decimales) multiplicado por un exponente cuya base generalmente es 10 pero en informática, debido al sistema binario, se suele usar una base.

Fórmula para calcular los flops en un CPU:

1 Gigaflops = (Velocidad de la CPU en GHz) x (Nº núcleos) x (Instrucciones por ciclo).

Donde GHz es gigahercio.

1.4. Conceptos básicos para entender la supercomputación

Algoritmos: Formas de resolución de un problema mediante una serie de pasos definidos, precisos y finitos. En este caso se resuelven en computadores.

Algoritmo paralelo: Algoritmo diseñado para que sus tareas o instrucciones se ejecuten por partes y a la vez. Sus salidas serán unidas para generar el resultado final del algoritmo.

Algoritmo secuencial: Algoritmo cuyas tareas o instrucciones se ejecutan una detrás de otra, es decir, siguen una secuencia de pasos. La salida de una es la entrada de la siguiente y así sucesivamente hasta el fin del algoritmo.

Clúster: Un clúster computacional (o de cálculo) es una red de computadoras que trabajan conjuntamente y en paralelo con el objeto de resolver problemas complejos.

Ciencia de la Computación: Comprende el estudio de las bases teóricas de la información y la computación, así como su aplicación en sistemas computacionales. Existen diversos campos o disciplinas dentro de las Ciencias de la Computación o Ciencias Computacionales; algunos enfatizan los resultados específicos del cómputo (como los gráficos por computadora), mientras que otros (como la teoría de la complejidad computacional) se relacionan con propiedades de los algoritmos usados al realizar cómputos.

Computación paralela: Es una técnica de programación en la que muchas instrucciones se ejecutan simultáneamente en diversos núcleos e incluso en diversos procesadores y computadoras. Se basa en el principio de que los problemas grandes se pueden dividir en partes más pequeñas que pueden resolverse de forma concurrente (“en paralelo”). Es la principal técnica utilizada en supercomputación.

Computación científica: Comprende la construcción de modelos matemáticos y técnicos numéricos para resolver problemas científicos, sociales y de ingeniería.

CPU: “Unidad Central de Procesamiento”, es la encargada de procesar absolutamente todas las operaciones de la computadora. Realizan sus operaciones de forma secuencial.

GPU: “Unidad de Procesamiento Gráfico”, por lo general, se combina con una CPU para acelerar aplicaciones con necesidades de procesamiento gráfico (y otras que demanden considerables capacidades de procesamiento).

Mainframe: Son también conocidas como macro computadoras, son sistemas grandes, rápidos, con capacidad de controlar cientos de usuarios en forma simultánea, así como manejar cientos de dispositivos de entrada y salida (manejo de grandes cantidades de entrada, salida y almacenamiento.). Por ejemplo, para el procesamiento de transacciones bancarias.

Núcleo o Core: Una unidad de procesamiento que lee las instrucciones y ejecuta las acciones específicas. Actualmente los procesadores vienen con más de un núcleo. Ejemplos:

Los procesadores Intel Core i5, de acuerdo a la presentación cuentan con entre 2 y 4 núcleos.

Los procesadores que usa el supercomputador MANATI del IIAP cuentan con 14 núcleos cada uno.

Procesador: Es un dispositivo electrónico que incluye varios bloques, entre ellos se encuentran los que se denominan núcleos, los cuales se encargan de ejecutar las instrucciones y pueden ser vistos como unos micros en miniatura.

Procesamiento de imágenes: Conjunto de técnicas para la adquisición, transmisión, procesamiento y representación de imágenes digitales.

Red neuronal artificial: Modelos matemáticos inspirados en sistemas biológicos adaptados y simulados en computadoras.

Rpeak: Rendimiento máximo teórico del supercomputador (lo que podría alcanzarse en el mejor de los casos).

Sistema de Alto Rendimiento Computacional: Conjunto de elementos asociados para facilitar servicios de procesamiento de grandes volúmenes de información. Entre los elementos destacan:

- Equipos de grandes capacidades de procesamiento.

- Sistema de refrigeración.
- Sistema eléctrico para acceso y respaldo de energía.
- Sistema eléctrico de soporte.
- Software de gestión.
- Personal encargado de la gestión y acceso.

Teraflops: Número de operaciones de punto o coma flotante por segundo que es capaz de realizar un procesador o núcleo.

- 1 gigaflops comprende 1000 millones de flops (10^9 flops).
- 1 teraflops comprende 1000 gigaflops.
- 1 petaflops comprende 1000 teraflops.

Visión computacional: Conjunto de técnicas que permiten la emulación de la capacidad visual humana en las computadoras, mediante la interpretación de imágenes adquiridas, de tal forma se puedan reconocer los diversos objetos en el ambiente y su posición en el espacio.

1.5. ¿En qué casos no es recomendable comprar una supercomputadora?

Las supercomputadoras son equipos sofisticados que demandan altos costos de adquisición y operación. Nadie sabe a ciencia cierta cuantas supercomputadoras existen en el mundo, ni en qué se utilizan, sin embargo, la decisión de adquirirlas siempre está ligada a una serie de necesidades de procesamiento de grandes volúmenes de datos en el menor tiempo posible. Estos procesos se dan mayormente en el sector de la investigación científica. En otros casos, es poco probable que se la necesite por lo que se debe adquirir equipamiento acorde a las necesidades.

En algunos casos se confunde lo que es, o en lo que se debe usar una supercomputadora. Es pertinente, por lo tanto, dar las siguientes recomendaciones:

No es un gran almacén de información, repositorio o una gran base de datos: La principal fortaleza de la supercomputadora no es su capacidad de almacenamiento. Sus funcionalidades y altos costos están relacionados a la cantidad de microprocesadores. Es un gran error adquirir un supercomputador para almacenar los grandes volúmenes de información. Existen diferentes soluciones de menor costo para almacenar información (Storages, servidores de almacenamiento o computación en la nube), incluso existen una serie de servicios gratuitos y de pago que pueden satisfacer las necesidades de almacenamiento.

No es un servidor para publicar sitios web: No es apropiado adquirir una supercomputadora para usarla como alojamiento de los sitios web institucionales, pues se subutilizarían los procesadores. Para publicar sitios web se recomienda utilizar diversas familias de servidores disponibles en el mercado para tal fin. Además se pueden arrendar servicios de alojamiento web.

No es un servidor de aplicaciones para los sistemas de las instituciones: Tampoco es recomendable en la actualidad adquirir una supercomputadora para servir de plataforma a las diversas aplicaciones organizacionales (sistemas administrativo, sistemas de recursos humanos, sistemas de toma de decisiones, y otras muy comunes). En todo caso lo que se deba poner en el supercomputador debería aprovechar al máximo su cantidad de procesadores (los programas deben usar procesamiento paralelo).

No es una gran computadora para servicios transaccionales tipo operaciones bancarias. Otro error muy común es pensar que las supercomputadoras pueden servir para dar servicio permanente de operaciones, por ejemplo, a los clientes de la empresa (tipo entidades bancarias). En realidad en este esquema también se subutilizaría las capacidades de procesamiento. Para ello se recomienda utilizar equipos tipo computadoras mainframe con menos procesadores y mejor preparadas para soportar ese tipo de operaciones (las Mainframe están configuradas para trabajar con gran cantidad de usuarios, entrada y salida de datos).

No es un equipo para almacenamiento en la nube. Tal como pasa en el caso de los sitios web, una supercomputadora no puede ser utilizada para ofrecer servicios empresariales de almacenamiento por internet.

No es acelerador de operaciones de los aplicativos o sistemas informáticos: El hecho de que se instale un aplicativo o programa en la supercomputadora no garantiza que este se ejecute de forma más rápida. Si el programa no está hecho (o no es adaptable) para ejecutarse de forma paralela, de nada servirá que esté instalado en el supercomputador, pues no podrá aprovechar eficientemente las capacidades de procesamiento.

1.6. Algunas interrogantes para trabajar con supercomputadoras

Es necesario realizar una serie de aclaraciones o precisiones sobre el uso de los supercomputadores:

¿Es necesario estar conectado dentro de la misma red (intranet) para acceder a trabajar en una supercomputadora?

No, incluso en la mayoría de supercomputadoras los usuarios acceden a través de internet desde lugares geográficos distintos (ciudades, países).

¿La conexión a internet es crítica para su uso?

No necesariamente. En el Perú se viene utilizando, por ejemplo, la supercomputadora del IIAP ubicado en Iquitos, sin mayores dificultades (teniéndose en cuenta las limitaciones del internet) desde Lima, Arequipa y otras ciudades extranjeras.

El internet podría ser un factor crítico en caso de que se quiera enviar grandes volúmenes de información (por ejemplo, miles de archivos de imágenes satelitales) o en el caso de que se haga procesamiento paralelo y distribuido (distribución del procesamiento en supercomputadoras ubicadas geográficamente).

¿Cualquier programa, software o sistema informático puede ser instalado en una supercomputadora?

No, en realidad depende de que exista una versión orientada al sistema operativo instalado en la supercomputadora. No olvidar que la finalidad es aprovechar al máximo las capacidades de procesamiento (sino es para eso no tiene sentido ponerlo en la supercomputadora). Actualmente la mayoría de supercomputadores del mundo utiliza diversas versiones del sistema operativo Linux.

¿Un programa fácilmente puede aprovechar las capacidades de la supercomputadora?

Respuesta: No necesariamente. Actualmente existe software especializado que se comercializa con versiones para procesamiento paralelo. Si se va elaborar un software propio el reto es adaptarlo para que funcione en procesamiento paralelo.

¿Qué se debe hacer para que un programa aproveche las capacidades?

Respuesta: En el caso de que se decida elaborar programas propios, se debe incluir código que pueda ejecutarse de forma paralela. Es un trabajo básicamente de programadores con conocimientos de programación paralela.

¿Cuál es la diferencia entre procesadores numéricos (CPU) y procesadores gráficos (GPU)?

Respuesta: En primer lugar, cada uno tendrá mejor desempeño frente al tipo de información que se desee procesar. Los procesadores numéricos (CPU) son procesadores tradicionales mayormente utilizados para cálculos matemáticos, y vale la pena mencionar que estos procesadores numéricos trabajan de forma secuencial (paso a paso y operación tras operación), estas son las que comúnmente se usan en computadoras personales (PC) y computadoras portátiles.

En el caso de los procesadores gráficos (GPU), son relativamente nuevos en la computación, incorporan miles de núcleos GPU y funcionan netamente de forma paralela, esto hace que sean considerables las velocidades alcanzadas (son también llamadas tarjetas gráficas o aceleradores gráficos). También es cierto que cada núcleo GPU puede tener menor capacidad que un núcleo CPU, sin embargo las tarjetas gráficas cuentan con mayor cantidad de núcleos GPU por lo que alcanza velocidades muy superiores.

1.7. Sistemas de alto rendimiento computacional

Definimos un sistema de alto rendimiento computacional como la articulación sinérgica de un conjunto de elementos que permiten la generación de servicios de procesamiento de grandes volúmenes de información para resolver problemas específicos. En este caso, la cohesión de los elementos es crítica para el cumplimiento de los objetivos del sistema.

Esta definición va más allá del concepto que utilizamos para un supercomputador o HPC, pues además comprende una serie de elementos necesarios para que este pueda operar de la forma más eficiente posible.

Entre los elementos a tener en cuenta dentro de un sistema de alto rendimiento computacional destacan:

- Equipamiento de alto desempeño computacional.
- Sistema de refrigeración.
- Software.
- Sistema eléctrico.
- Adecuaciones del espacio físico donde será ubicado.



Imagen 3: Esquema de un sistema de alto rendimiento computacional. Créditos: Elaboración propia.

Cualquier proyecto de HPC debe tener en cuenta los elementos mencionados. Este tema será tratado con mayor amplitud en el capítulo 2 y 3.

1.8. Capacidades a tener en cuenta en una supercomputadora.

No existe conversación sobre supercomputadoras en la que no se realicen comentarios sobre las capacidades de estos sistemas, generándose incluso discusiones sin sentido sobre las equivalencias de los equipos de estas computadoras respecto a las computadoras convencionales o computadoras personales (PC), computadoras portátiles, dispositivos móviles, tabletas, etc.

Si bien es cierto que la unidad de medida mayormente utilizada para representar la capacidad de procesamiento de una supercomputadora son los flops, debemos enfatizar que serán de igual importancia las capacidades de memoria, velocidad de comunicación entre núcleos, procesadores e incluso entre nodos. En el caso del uso de tarjetas gráficas será también determinante la frecuencia de shaders (velocidad con las que transfieren instrucciones gráficas). Por ello no debe ser raro ver supercomputadoras que, a pesar de contar con mayor capacidad de núcleos de procesamiento, logren menores rendimientos que supercomputadoras con menor cantidad de núcleos.

No vamos a profundizar en este tema, pero sí, a partir de información de algunas fuentes bibliográficas, vamos a introducir algunas ideas que nos permitirán alcanzar un mayor grado de comprensión.

Para ilustrar el ejemplo, vamos a tomar como punto de partida las capacidades de procesamiento gráfico teórico de uno de los 6 Nodos para procesamiento gráfico del Clúster MANATI IIAP, comparándolo con diversos computadores e incluso dispositivos (Tabla 01).

Tabla 01: Capacidades en GIGAFLOPS de diversos dispositivos.

Equipo/Dispositivo	Hardware	Gigaflops
Nodo MANATI IIAP	K80 NVIDIA	8500
Consola de Videojuego	XBOX 360	240
Computar Portátil	Intel(R) Core(TM) i7	112
Computar Portátil	Intel(R) Core(TM) i5	105
Computar Portátil	Intel(R) Core(TM) i3	40
Teléfono Celular	Samsung Galaxy S6	34.8
Consola de Videojuego	PlayStation 2S	6.2
Computador Portátil	Intel(R) Celeron(R) CPU N2830	5
Teléfono Celular	Apple iPhone 4	1.6
Computador Personal	Intel Pentium 4	1

Fuente: Elaboración propia.

Con la finalidad de ilustrar mejor las equivalencias, presentamos la Tabla 02, con las diversas equivalencias respecto al Nodo del MANATI IIAP.

Tabla 02: Equivalencias del Nodo supercomputador IIAP frente a diversos dispositivos.

Equivalencias Nodo del MANATI IIAP
35 Consolas de Videojuegos XBOX 360
76 Computadoras Portátiles con procesador Intel(R) Core(TM) i7
81 Computadoras Portátiles con procesador Intel(R) Core(TM) i5
213 Computadoras Portátiles con procesador Intel(R) Core(TM) i3
244 Teléfonos Celulares Samsung Galaxy S6
1371 Consolas de Videojuegos PlayStation 2S
1700 Computadoras Portátiles con procesador Intel(R) Celeron(R) CPU N2830
5213 Teléfonos Celulares Apple Iphone 4.
8500 Computadoras personales con procesador Intel Pentium 4

Fuente: Elaboración propia.

1.9. Historia de la supercomputación

Diversos autores coinciden que el equipo denominado CDC 6600 fue el primer supercomputador de la historia. Este equipo fue lanzado por la empresa Control Data Corporation – CDC en Estados Unidos, en 1961, teniendo como creador al legendario experto en hardware y “padre de la supercomputación”, Seymour Cray. Con la finalidad de mejorar la ilustración de este texto les presentamos una lista de computadoras previas al CDC 6600 que en su momento fueron consideradas como los más grandes avances de la computación.

Hitos de la computación antes del CDC 6600

Enigma: Fue creada por Arthur-Scherbius en 1918 y luego fue utilizada principalmente por el ejército alemán. Tuvo un papel importante en el envío de mensajes encriptados durante la II Guerra Mundial. Incluso algunas personalidades indican que tuvo mucho que ver en la derrota de Alemania.



Imagen 4: Máquinas Enigma en el Museo Criptológico Nacional, Maryland, USA. Grafico (GPU) y una unidad de almacenamiento. Créditos: Robert Malmgren

Z1, Z2 y Z3: Fueron creadas por Konrad Zuse desde 1936 hasta 1941 (al principio con fondos propios y luego por el contexto de la II Guerra Mundial con fondos del gobierno Alemán). La Z3 que es la evolución de las primeras 2, es considerada la primera computadora electrónica digital totalmente funcional, además en ella se introdujo el sistema de numeración binario a las computadoras.



Imagen 5: Réplica del Z1 en el Museo Alemán de Tecnología en Berlín. Créditos: ComputerGeek

ABC Atanasoff-Berry Computer: Fue creado por John Vincent Atanasoff en Iowa State University (Estados Unidos) entre 1937 y 1943. Fue creado para resolver ecuaciones diferenciales lineales de forma simultánea. Podía resolver veintinueve ecuaciones al mismo tiempo.



Imagen 6: Atanasoff-Berry en el Durham Center, Universidad Estatal de Iowa. Créditos: Marcin Wichary

Colossus: Fue creado por Tommy Flowers para el ejército inglés en plena II Guerra Mundial en 1943, su objetivo fue descifrar los mensajes encriptados del ejército alemán. Fue uno de los primeros computadores digitales.



Imagen 7: Réplica de Colossus en el Museo Nacional de Computación, Bletchley Park, Reino Unido. Créditos: User:Manop

ENIAC: Acrónimo de Electronic Numerical Integrator And Computer. Fue creado por John Presper Eckert y John William Mauchly en 1946, en la Universidad de Pensilvania (Estados Unidos) con la finalidad de resolver una serie de problemas del ejército de Estados Unidos. Su objetivo fue la ejecución de procesos y operaciones numéricas utilizadas para la artillería.

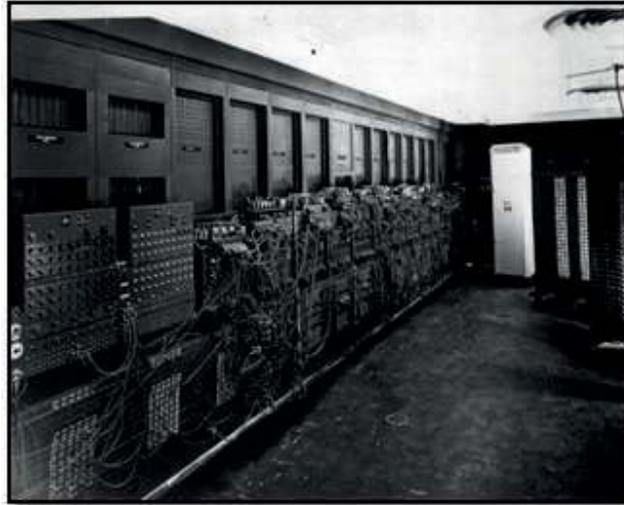


Imagen 8: Foto del lado izquierdo de ENIAC instalada en BRL Edificio 328. Créditos: U.S. Army Photo

EDVAC (Electronic Discrete Variable Automatic Computer): También fue creado por John Presper Eckert y John William Mauchly (aunque también trabajó en el proyecto John von Neumann) en la Universidad de Pensilvania (Estados Unidos) en 1949. Este computador permitía tener programas almacenados, además funcionaba en código binario.

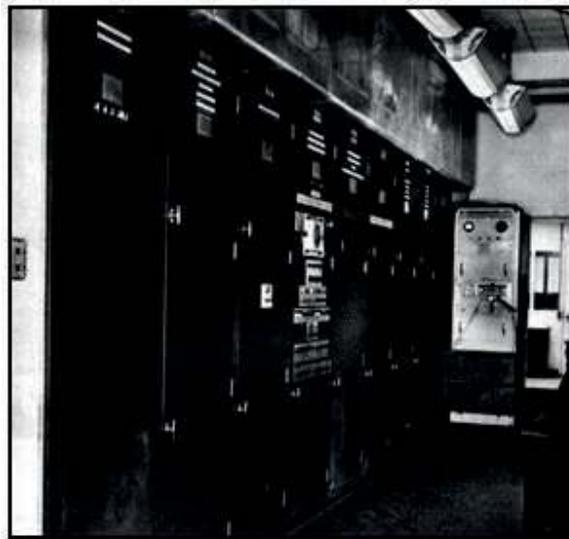


Imagen 9: Foto de la EDVAC del Ejército de EE.UU." Créditos: U.S. Army Photo

Manchester MARK1: Creada en la Universidad de Manchester (Inglaterra) en 1949, es la primera computadora en ejecutar un programa almacenado (la primera en utilizar memoria RAM). Alan Turing realizó un lenguaje para este computador.



Imagen 10: réplica del Manchester MARK 1 exhibida en el Museo de Ciencia e Industria. Créditos: Genio

CSIRAC: Fue creado por Maston Beard y Trevor Pearcey en el Laboratorio de Radiofísica en Sidney (Australia) entre 1947 y 1949. Fue el primer computador que reprodujo música y sonido.

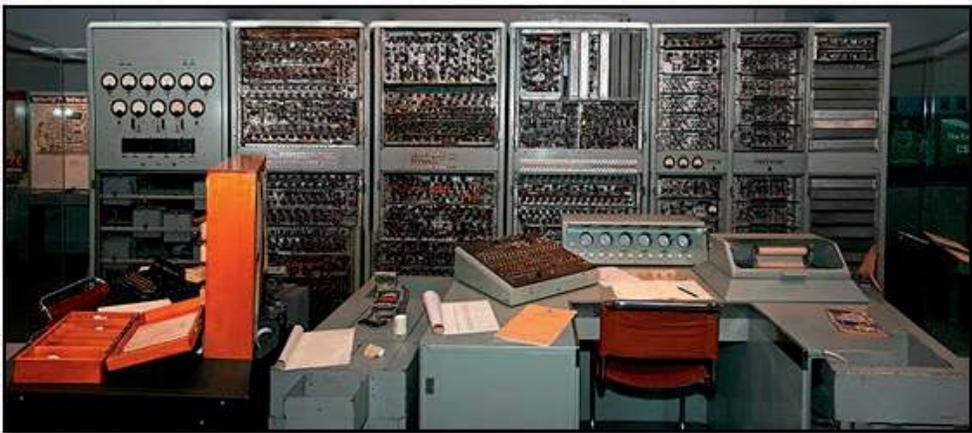


Imagen 11: Réplica del Computador CSIRAC exhibido en el Museo Victoria de Australia. Créditos: jjron

EDSAC (Electronic Delay Automatic Calculator): Creada en la Universidad de Cambridge (Inglaterra) en 1949, fue la primera computadora creada específicamente para resolver problemas prácticos de ingenieros, matemáticos y científicos. Según la web de la BBC [8] ayudó a 3 científicos de Cambridge a ganar el Premio Nobel, funcionó 10 años. Su creador fue Sir Maurice Wilkes.

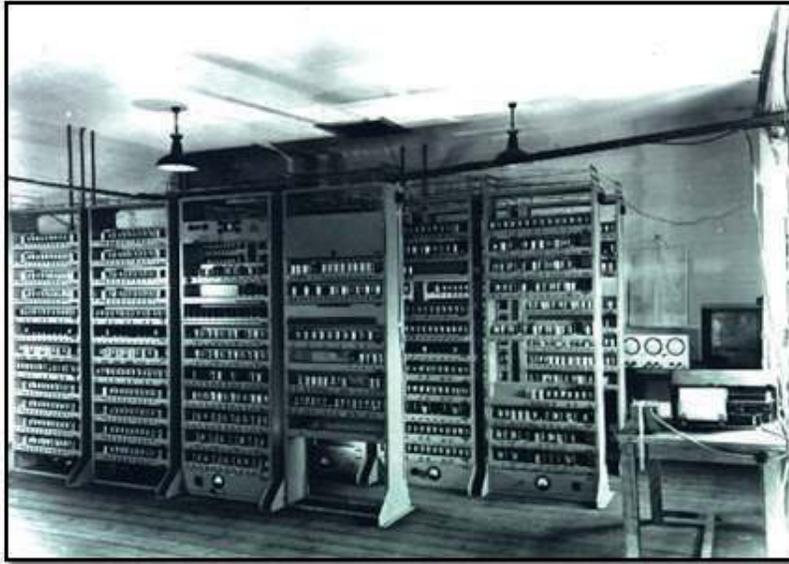


Imagen 12: Foto de la Computadora EDSAC, creada en 1949. Créditos: Thorpe

MONIAC (Monetary National Income Analogue Computer): Fue creado por William Phillips el año de 1949. Fue la primera computadora hidráulica capaz de predecir los vaivenes de la economía de un país en base a las decisiones de sus gobernantes.



Imagen 13: MONIAC exhibido en el Banco de la Reserva de Nueva Zelanda. Créditos: Kaihsu Tai

IAS Computer o Computador del Instituto de Estudios Avanzados: Fue creado en Princeton (Estados Unidos) entre 1946 y 1951. Fue diseñado por John Von Neumann. Esta máquina utilizó dispositivos de almacenamiento electrostático como la memoria principal (tubos de rayos catódicos). Cada tubo podría almacenar 1024 bits.

UNIVAC (UNIVersal Automatic Computer): También fue creado por John Presper Eckert y John William Mauchly entre 1946 y 1951 (algunas fuentes indican que Seymour Cray también participó en el proyecto). Fue el primer computador creado para procesar información numérica y alfabética. Fue utilizado para el censo de 1951 en Estados Unidos.



Imagen 14: foto de UNIVAC 9400 instalado en el Museo Alemán. Créditos: H. Müller

LINC: Fue creado por Wesley A. Clark y Charles Molnar en el Lincoln Laboratorio del MIT (Estados Unidos), en 1963. Es considerado el primer minicomputador y un precursor del ordenador personal (pensado para ser usado por un único usuario), además, estaba diseñado para usos médicos.



Imagen 15: LINC-8 equipo en exhibición en la universidad de Uppsala, Suecia Bjami Juliusson. Créditos: Bjami Juliusson

1.10. International Business Machines (IBM)

IBM 701: Computadora binaria creada en 1952 en el Watson Lab de la Universidad de Columbia. Fue creada para realizar cálculo científico y fue utilizada para cálculo de predicciones meteorológicas e incluso en fábricas de aviones. La primera unidad fue instalada en la sede de IBM en Nueva York.



Imagen 16: Consola de Operador de IBM 701. Créditos: Dan

IBM 650: Computador decimal de dos direcciones. Fue creada en 1954, siendo la primera computadora producida en masa del mundo. Su relativa accesibilidad permitió que sea empleada para cálculo científico, ingeniería, aplicaciones militares e incluso para enseñanza de programación en la universidad y en escuelas.



Imagen 17: Ordenador IBM 650. Créditos: MUNCYT2

IBM 704: Creada en 1956, fue la primera computadora producida en serie con hardware de aritmética de punto flotante, por lo que fue por buen tiempo la única computadora que podía manejar matemáticas complejas.

La creación del FORTRAN y LISP favoreció la expansión de esta computadora. Fue utilizada principalmente en el sector científico, para la observación astrofísica del Smithsonian Operation Moonwatch, además para diversos trabajos del MIT.



Imagen 18: Un IBM 704 unidad central. Créditos: Lawrence Livermore National Laboratory

IBM 709: Creada en 1958 para cálculos científicos, como el IBM 704. Alcanzó velocidades de hasta 1 megabit por segundo. Sirvió de apoyo a la Organización Europea para la Investigación Nuclear – CERN.



Imagen 19: Panel frontal del IBM 709 en el Museo de Historia de la Computadora. Créditos: Arno Reinhold

IBM 360: Creada en 1964, es una de las primeras computadoras comerciales que usó circuitos integrados. Podía realizar análisis numéricos, así como administración o procesamiento de archivos (además escalabilidad). Según diversos autores, a partir del IBM 360 empieza la tercera generación de computadoras. Su uso tenía orientación científica y empresarial.



Imagen 20: IBM 360 en el Museo Alemán. Créditos: Wolfgang Manousek de Dormagen, Alemania

Tabla 03: Cuadro de resumen de la historia de la supercomputación

N°	Nombre	Año	País	Creador(es)	Características y/o usos
1	Enigma	febrero de 1918	Alemania	Arthur-Scherbius	Envío de mensajes encriptados en la segunda guerra mundial.
2	Z1, Z2 y Z3	1936 hasta 1941	Alemania	Konrad Zuse	Introdujo el sistema de numeración binario a las supercomputadoras.
3	ABC – Atanasoff Berry Computer	1937 hasta 1943	Estados Unidos	John Vincent	Resoluciones de ecuaciones diferenciales lineales.
4	Colossus	1943	Inglaterra	Tommy Flowers	Descifrar mensajes encriptados.
5	ENIAC	1946	Estados Unidos	John Presper Eckert y John William Mauchlyen	Resolución, ejecución de procesos y operaciones numéricas.
6	EDVAC	1949	Estados Unidos	John Presper Eckert y John William Mauchlyen	Almacenamiento de programas, utilización de código binaria.
7	Manchester MARK1	1949	Inglaterra	Universidad de Manchester	Primera computadora en utilizar una memoria RAM, Alan Turing realizó un lenguaje para este computador.
8	CSIRAC	1947 hasta 1949	Australia	Maston Beard , Trevor Pearcey	Primera computadora en reproducir música y sonidos.
9	EDSAC	1949	Inglaterra	Universidad de Cambridge	Primera computadora en resolver prácticos de Ingenieros, científicos y matemáticos.
10	MONIAC	1949	Nueva Zelanda	William Phillips	Primera computadora Hidráulica.
11	IAS Computer	1946 hasta 1951	Estados Unidos	John Von Neumann	Almacenamiento electrostático como la memoria principal (tubos de rayos catódicos). Cada tubo podría almacenar 1024 bits.
12	UNIVAC	1946 hasta 1951	Estados Unidos	John Presper Eckert y John William Mauchlyen	Procesar información numérica y alfabética.
13	LINC	1963	Estados Unidos	Wesley A. Clark y Charles Molnar	Considerado el primer minicomputador y un precursor del ordenador personal.
14	IBM 701	1952	Estados Unidos	Watson Lab de la Universidad de Columbia	Realizar cálculo científico y predicciones meteorológicas.
15	IBM 650	1954	Estados Unidos	Watson Lab de la Universidad de Columbia	Su relativa accesibilidad permitió que sea empleada para cálculo científico, ingeniería, aplicaciones militares e incluso para enseñanza de programación en la universidad y en escuelas.
16	IBM 704	1956	Estados Unidos	Watson Lab de la Universidad de Columbia	Utilizado principalmente en el sector científico, como para la observación Astrofísica del Smithsonian Operation Moonwatch.
17	IBM 709	1958	Estados Unidos	Watson Lab de la Universidad de Columbia	Para cálculos científicos, Sirvió de apoyo a la Organización Europea para la Investigación Nuclear.
18	IBM 360	1964	Estados Unidos	Watson Lab de la Universidad de Columbia	Primeras computadoras comerciales que usaron circuitos integrados.

1.11. El genio de Seymour Cray

En 1957 un grupo de ingenieros y especialistas estadounidenses de mucha experiencia en proyectos de computación, asociados a su gobierno, decidió crear la empresa Control Data Corporation (CDC). Esta empresa surge con la visión de fabricar y comercializar computadoras de gran potencia, en el mismo rubro que ya destacaba IBM, con la que luego tendrá una serie de disputas.



Créditos: Michael Hicks

En 1958, Seymour Cray (1925 - 1996) se incorpora al CDC y lidera los principales proyectos de computadoras de grandes capacidades. Este ingeniero nacido en el poblado de Chippewa Falls, Wisconsin (Estados Unidos), se graduó en ingeniería eléctrica y matemáticas en la Universidad de Minnesota. Durante el periodo de Cray en el CDC se destaca la creación de los siguientes computadores:

- **CDC 1604:** Creado en 1960 por el equipo dirigido por Seymour Cray, fue uno de los primeros computadores que utilizó transistores en vez de tubo de vacío (diseñado para trabajar en 48 bits), y tuvo un interesante éxito comercial.

El primer CDC 1604 fue fabricado y entregado a la Marina de Guerra de los Estados Unidos en 1960, donde fue utilizado para problemas de los centros de control de las operaciones de la flota en Hawaii, Londres, y Norfolk, Virginia. En 1964 se construyeron más de 50 CDC 1604.

- **CDC 6600:** Creado en 1964 también por Seymour Cray y su equipo. La CDC 6600 es una computadora legendaria, considerada la primera supercomputadora del mundo, que en plenos años 60s alcanzó hasta 9 megaflops, tres veces más rápido que el mejor equipo de IBM. La CDC 6600 empleó transistores de Silicio en lugar del entonces habitual Germanio. Para alcanzar grandes velocidades se utilizaron 10 computadoras pequeñas, conocidas como unidades de procesamiento periférico (estas operaban en paralelo).

La CDC 6600 logró vender más de 100 equipos (cifra exorbitante para la época),

Se usó principalmente en laboratorios que fabricaban bombas nucleares. Algunos ejemplares llegaron a los laboratorios de computación universitarios.



Imagen 21: Computador CDC 6600. Créditos: Jitze Couperus

- **CDC 7600:** Fue lanzado en 1969, por el mismo equipo creador de los predecesores. Fue considerado el computador más veloz del mundo desde 1968 hasta 1975. Entre sus mejoras destaca el uso de una técnica llamada *pipelining*, que consiste en comenzar a leer una instrucción antes de que termine de ejecutarse la instrucción anterior, lo que incluso le permitió alcanzar rendimientos pico de 40 megaflops.

Su crecimiento computacional además generó problemas de refrigeración (generaba problemas de mantenimiento). Entre sus principales usuarios destacó el Laboratorio Nacional Lawrence Livermore (LLNL) del Departamento de Energía de los Estados Unidos de América. El sucesor del CDC 7600 nunca se llegó a fabricar, por una serie de discrepancias que originaron, incluso, la retirada de Cray de CDC.

En 1972, Cray decidió crear su propia empresa denominada Cray Research. Fue apoyado por seis miembros de su equipo de CDC. En 1989 creó otra empresa propia denominada Cray Computer Corporation (CCC). Desde el inicio busco fabricar computadoras para resolver problemas complejos y simulaciones para fines científicos y militares. Numeramos a continuación las creaciones de Cray en esta etapa de su vida:



Imagen 22: Computador CDC 7600. Créditos: Jitze Couperus

- **Cray-1:** Fue lanzado en 1976, siendo el computador más veloz del mundo (tardó en construirse cuatro años). Utilizaba procesadores vectoriales, 64-bits (pesaba 5,5 toneladas, incluyendo el sistema de refrigeración por Freón), alcanzando picos de 160 megaflops.

Fue utilizado principalmente para fines científicos. El Centro Nacional de Investigaciones Atmosféricas (NCAR) fue su primer cliente oficial (1977). Cray proyectaba vender al menos una docena de Cray-1, sin embargo, al final se vendieron más de 80, lo que hizo prosperar a la empresa.



Imagen 23: Cray-1, Museo de Historia de la Computación, Mountain View, California.
Créditos: Joi Ito de Inbamura, Japón

- **Cray-2:** Fue lanzado en 1985. Superando con creces la velocidad de Cray-1, alcanzó picos de 1.9 gigaflops. Utilizó procesadores vectoriales.

Su construcción inicial fue a pedido de los sectores de defensa (armas nucleares), energía, investigación oceanográfica de los Estados Unidos. Luego fue utilizada en universidades y empresas de todo el mundo.



Imagen 24: Computador Cray-2. Créditos: David.Monnaux

- **Cray-3 y Cray-4:** Ambas supercomputadoras no repitieron el éxito de los predecesores, debido a la baja demanda de supercomputadores por el crecimiento del mercado de los microprocesadores orientados a las computadoras personales.

El Cray-3 se empezó a construir en 1988 y estuvo listo en 1991. Utilizaba semiconductores de arseniuro de galio (GaAs) y cientos de circuitos integrados personalizados construidos en una CPU de 1 pie cúbico. El objetivo del diseño fue un rendimiento de alrededor de 16 gigaflops, aproximadamente 12 veces el de la Cray-2.

El Cray-4 era esencialmente una versión encogida y acelerada del Cray-3. Consistía de un número de procesadores vectoriales anclados a una memoria rápida. Nunca vio la luz. Cray se declaró en bancarrota en 1995. Falleció en 1996 en un accidente automovilístico.

1.12. Computadoras más veloces según el Top500.

En junio de 1993 fue lanzada la iniciativa Top500 [9], orientada a conocer las computadoras de mayor rendimiento del mundo (las primeras 500). El grupo que lidera este ranking está conformado por:

- Universidad de Mannheim (Alemania).
- Universidad de Tennessee (USA).
- NERSC/Lawrence Berkeley National Laboratory (USA).
- NERSC/Lawrence Berkeley National Laboratory (USA).

Según el Top500 de noviembre del 2016 [1] la supercomputadora más potente es la TaihuLight desarrollada por el NRCPC (Centro de Investigación Nacional de Tecnología e Ingeniería de Computación en Paralelo) de China.

A continuación presentamos las principales computadoras creadas o en funcionamiento desde 1993 hasta la actualidad.

1.Numerical Wind Tunnel: Fue instalada en la National Aerospace Laboratory of Japan, fabricada por la empresa, Fujitsu. Comprendía procesadores vectoriales conectados en paralelo. Alcanzó velocidades de 170 gigaflops. Lideró el Top500 en 4 oportunidades en el periodo de 1993 hasta 1995.

2.ASCI Red (supercomputador rojo): Fue instalada en la Sandia National Laboratories en Estados Unidos, fue fabricada por la empresa Intel, constaba de un procesador masivo paralelo (MPP), con una arquitectura múltiple de memoria distribuida de múltiples datos (MIMD). Alcanzó velocidades superiores a 2 teraflops (2000 gigaflops). Lideró el Top500 en 7 oportunidades en el periodo de 1997 hasta 2000. Su uso estuvo orientado a investigaciones científicas sobre energía nuclear.



Imagen 25: Supercomputadora ASCI Red. Créditos: Gobierno de los Estados Unidos

3.ASCI White (supercomputador blanco): Fue instalada en la Lawrence Livermore National Laboratory en Estados Unidos. Fue fabricada por la empresa IBM. Constaba de un clúster de ordenadores, basados en el modelo comercial RS/6000 SP de IBM. Alcanzó velocidades superiores a 7 teraflops (7000 gigaflops). Lideró el Top500 en 3 oportunidades en el periodo de 2000 hasta 2001. Su uso también estuvo orientado a investigaciones científicas sobre energía nuclear.

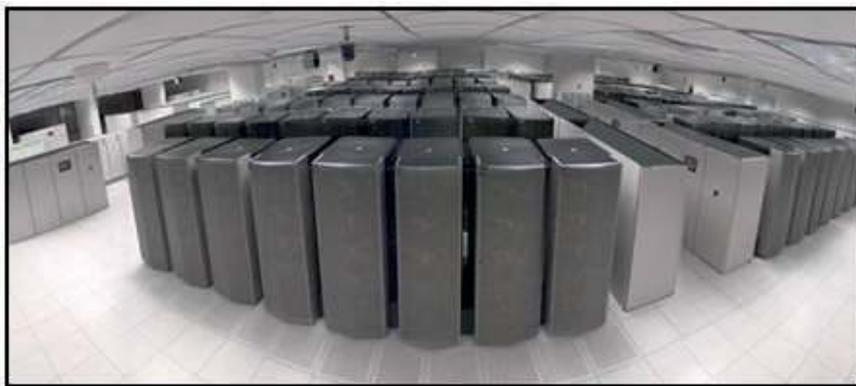


Imagen 26: Supercomputador ASCI White. Créditos: Gobierno de Estados Unidos: Lawrence Livermore National Laboratory

4.Earth Simulator: Fue instalada en el Earth Simulator Center en Japón. Fue fabricada por la empresa NEC. Constaba de 5120 CPUs especiales de 500 MHz. Alcanzó velocidades superiores a 40 teraflops. Lideró el Top500 en 5 oportunidades en el periodo de 2002 hasta 2004. Su uso estuvo orientado a investigaciones científicas, geológicas y ambientales.



Imagen 27: Supercomputador Earth Simulator. Créditos: GenGen (げんげん)

5.eServer Blue Gene Solution: Fue instalada en la Lawrence Livermore National Laboratory en Estados Unidos. Fue fabricada por la empresa IBM, constaba de un total de 212 992 núcleos de procesamiento. Alcanzó velocidades superiores a 478 teraflops. Lideró el Top500 durante 3 años, en 6 oportunidades en el periodo de 2005 hasta 2007. Su uso también estuvo orientado a investigaciones científicas sobre energía nuclear.



Imagen 28: Supercomputador eServer Blue Gene Solution. Créditos: Argonne National Laboratory

6.Jaguar (Cray XT5): Fue instalada en el Oak Ridge National Laboratory en Estados Unidos. Fabricada por la empresa Cray Inc. Constaba de 264 122 núcleos de procesamiento. Alcanzó velocidades superiores a 1.7 petaflops (1700 teraflops). Lideró el Top500 en 2 oportunidades en el periodo de 2009 hasta 2010. Su uso también estuvo orientado a investigaciones científicas sobre energía.



Imagen 29: Supercomputador JAGUAR CRAY XT5. Créditos: Laboratorio Nacional de Oak Ridge, EE.UU. Departamento de Energía

7. Titán (Cray XK7): Reemplazó al Jaguar en el Oak Ridge National Laboratory (Estados Unidos). También fue fabricada por la empresa Cray Inc. Su arquitectura fue híbrida, basada en una combinación de unidades centrales de procesamiento (CPUs) tradicionales con unidades de procesamiento gráfico (GPUs). Constaba de un total de 560 640 núcleos de procesamiento. Alcanzó velocidades superiores a 17 petaflops (17 000 teraflops). Lideró el Top500 en una oportunidad en el año 2011. Su uso también fue similar al del Jaguar.

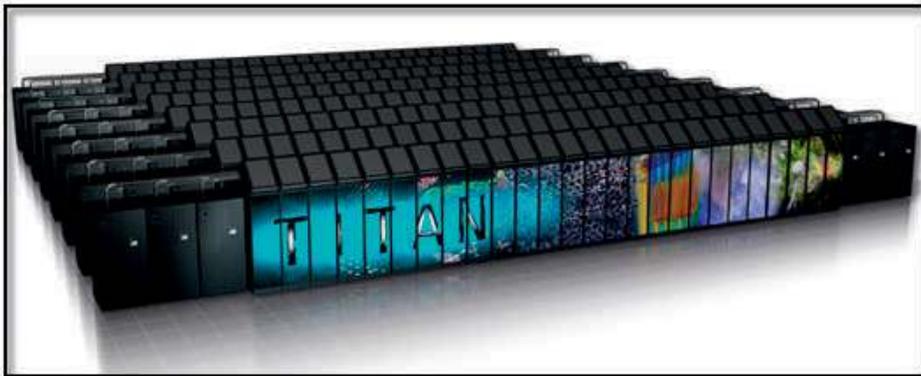


Imagen 30: Supercomputador TITAN CRAY XK7. Créditos: Laboratorio Nacional de Oak Ridge, EE.UU. Departamento de Energía

8. Tianhe-2: Instalada el año 2013 inicialmente en el National University of Defense Technology y posteriormente el mismo año en el National super Computer Center in Guangzhou, ambos de China. Fue construida por la National University of Defense Technology, con patrocinio del gobierno chino. Cuenta con más de 3 millones de núcleos de procesamiento (3 120 000), alcanzando rendimientos superiores a 33 petaflops (33 000 teraflops). Lideró el Top500 en 6 oportunidades en el periodo de 2013 hasta noviembre del 2015. Su uso estuvo orientado a investigaciones científicas sobre modelos climáticos atmosféricos, comportamiento del océano, etc.



Imagen 31: Supercomputadora Tianhe-2. Créditos: 001326

9.Sunway TaihuLight: Instalada el año 2016 en el National Supercomputing Center in Wuxi en China. Fue construida por el Centro Nacional de Investigación de Ingeniería y Tecnología de Computadores Paralelos (NRCPC). La principal innovación de esta supercomputadora lo constituye el procesador denominado SW26010, de fabricación china, que consta de 260 núcleos por cada procesador. Cuenta con más de 10 millones de núcleos de procesamiento (exactamente 10 649 600). alcanzado rendimientos superiores a 93 petaflop (93 000 teraflops). Actualmente viene liderando el Top500 desde junio del 2016. Su uso está orientado a modelos climáticos, investigación biológica y geológica avanzada, procesos industriales, ingeniería avanzada y análisis de datos.

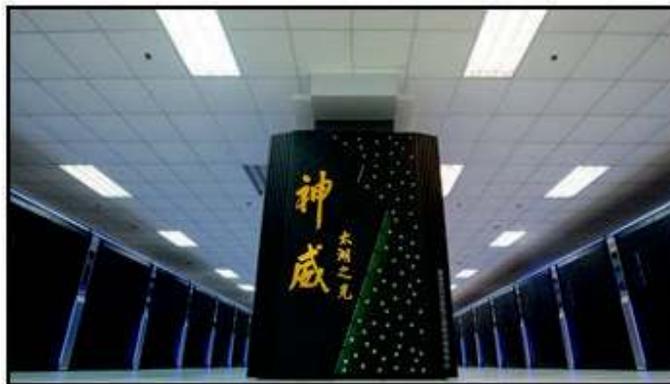


Imagen 32: Supercomputadora Sunway TaihuLight. Créditos: laboratorio linux

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Tabla 04: Lista de las Supercomputadoras más veloces según el Top500 (desde junio de 1993 hasta noviembre de 2016)

Año/mes	Computador	Institución	Fabricante	Núcleos	RMax
Jun-93	CM-5/1024	Los Alamos National Laboratory	Thinking Machines Corporation (USA)	1024	59.7
Nov-93	Numerical Wind Tunnel	National Aerospace Laboratory of Japan	Fujitsu (Japón)	140	124
Jun-94	XP/S140	Sandia National Laboratories	Intel (USA)	3680	143.4
Nov-94	Numerical Wind Tunnel	National Aerospace Laboratory of Japan	Fujitsu (Japón)	140	170
Jun-95	Numerical Wind Tunnel	National Aerospace Laboratory of Japan	Fujitsu (Japón)	140	170
Nov-95	Numerical Wind Tunnel	National Aerospace Laboratory of Japan	Fujitsu (Japón)	140	170
Jun-96	SR2201/1024	University of Tokyo	Hitachi (Japón)	1024	220.4
Nov-96	CP-PACS/2048	Center for Computational Science, University of Tsukuba	Hitachi (Japón)	2048	368.2
Jun-97	ASCI Red	Sandia National Laboratories	Intel (USA)	7264	1068
Nov-97	ASCI Red	Sandia National Laboratories	Intel (USA)	9152	1338
Jun-98	ASCI Red	Sandia National Laboratories	Intel (USA)	9152	1338
Nov-98	ASCI Red	Sandia National Laboratories	Intel (USA)	9152	1338
Jun-99	ASCI Red	Sandia National Laboratories	Intel (USA)	9472	2121
Nov-99	ASCI Red	Sandia National Laboratories	Intel (USA)	9632	2379
Jun-00	ASCI Red	Sandia National Laboratories	Intel (USA)	9632	2379
Nov-00	ASCI White	Lawrence Livermore National Laboratory	IBM (USA)	8192	4938
Jun-01	ASCI White	Lawrence Livermore National Laboratory	IBM (USA)	8192	7226
Nov-01	ASCI White	Lawrence Livermore National Laboratory	IBM (USA)	8192	7226
Jun-02	Earth-Simulator	The Earth Simulator Center	NEC (Japón)	5120	35860
Nov-02	Earth-Simulator	The Earth Simulator Center	NEC (Japón)	5120	35860
Jun-03	Earth-Simulator	The Earth Simulator Center	NEC (Japón)	5120	35860
Nov-03	Earth-Simulator	The Earth Simulator Center	NEC (Japón)	5120	35860
Jun-04	Earth-Simulator	The Earth Simulator Center	NEC (Japón)	5120	35860
Dec-04	BlueGene/L	IBM/DOE	IBM (USA)	32768	70720
Jun-05	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	65536	136800
Nov-05	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	131072	280600
Jun-06	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	131072	280600
Nov-06	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	131072	280600
Jun-07	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	131072	280600
Nov-07	eServer Blue Gene Solution	DOE/NNSA/LLNL	IBM (USA)	212992	478200
Jun-08	BladeCenter QS22/LS21 Cluster	DOE/NNSA/LANL	IBM (USA)	122400	1026000
Nov-08	BladeCenter QS22/LS21 Cluster	DOE/NNSA/LANL	IBM (USA)	129600	1105000
Jun-09	BladeCenter QS22/LS21 Cluster	DOE/NNSA/LANL	IBM (USA)	129600	1105000
Nov-09	Cray XT5-HE	Oak Ridge National Laboratory	Cray Inc.(USA)	224162	1759000
Jun-10	Cray XT5-HE	Oak Ridge National Laboratory	Cray Inc.(USA)	224162	1759000
Nov-10	NUDT TH	National supercomputing Center in Tianjin	NUDT (China)	186368	2566000
Jun-11	K computer	RIKEN Advanced Institute for Computational Science (AICS)	Fujitsu (Japón)	548352	8162000
Nov-11	K computer	RIKEN Advanced Institute for Computational Science (AICS)	Fujitsu (Japón)	705024	10510000
Jun-12	Sequoia	DOE/NNSA/LLNL	IBM (USA)	1572864	16324751
Nov-12	Titan	DOE/SC/Oak Ridge National Laboratory	Cray Inc.(USA)	560640	17590000
Jun-13	Tianhe-2 (MilkyWay-2)	National University of Defense Technology	NUDT(China)	3120000	33862700
Nov-13	Tianhe-2 (MilkyWay-2)	National super Computer Center in Guangzhou	NUDT(China)	3120000	33862700

Jun-14	Tianhe-2 (MilkyWay-2)	National super Computer Center in Guangzhou	NUDT(China)	3120000	33862700
Nov-14	Tianhe-2 (MilkyWay-2)	National super Computer Center in Guangzhou	NUDT(China)	3120000	33862700
Jun-15	Tianhe-2 (MilkyWay-2)	National super Computer Center in Guangzhou	NUDT(China)	3120000	33862700
Nov-15	Tianhe-2 (MilkyWay-2)	National super Computer Center in Guangzhou	NUDT(China)	3120000	33862700
Jun-16	Sunway TaihuLight	National supercomputing Center in Wuxi	NRCPC (China)	10649600	93014593.9
Nov-16	Sunway TaihuLight	National supercomputing Center in Wuxi	NRCPC (China)	10649600	93014593.9

Fuente: Top500 (www.top500.org). Créditos: Elaboración propia.

1.13. ¿Para qué usar una supercomputadora?

Según el profesor Felix García Merayo de la Universidad-UPM, “la investigación científica y el desarrollo técnico están soportados, hoy en día, en un alto grado, por técnicas de simulación computacional o simulación numérica; con ello es posible la experimentación, el contraste de hipótesis y, como consecuencia, la verificación de las teorías, todo lo cual evita el experimento físico directo que consume mucho tiempo e incluso, en muchos casos, es imposible de realizar, por razones económicas, políticas o medioambientales” [10]. En un mercado que evoluciona muy rápidamente, es necesario crear una fabricación ágil de todo tipo de productos. Esto conlleva la realización de simulaciones grandes, rápidas y complejas.

En realidad, son infinitos los usos que pueden darse a una supercomputadora. Si revisamos los usos que actualmente se vienen dando a las 500 principales supercomputadoras del mundo (www.top500.org), nos encontramos con la sorpresa de que la mayor cantidad de entidades de esta lista la constituyen una serie de industrias o empresas dedicadas a la fabricación de automóviles, aviones, exploración petrolera, gas, producción de energía atómica, servicios informáticos, fabricación de computadoras o servicios financieros. No siempre fue así, por ejemplo, en junio del año 1993, la mayor cantidad de supercomputadoras de este ranking estaban dedicadas a actividades de investigación científica.

Quince de las principales supercomputadoras del mundo están dedicadas hoy en día a la investigación científica. Entre los usos dados para el sector investigación destacan los estudios meteorológicos, el comportamiento del océano, los modelos climáticos atmosféricos, la bioinformática, la investigación biológica, la geología avanzada, la investigación en energía nuclear, los análisis de datos, etc. En este caso podemos mencionar a las supercomputadoras más potentes del mundo que son: Sunway TaihuLight (China), Tianhe-2(China), Titán (USA), Jaguar (USA), Cori (USA). Otros usos están relacionados con el trabajo académico de las universidades (que también incluye investigación científica) y labores propias de las agencias de diversos gobiernos.

A nivel de Latinoamérica, sólo 3 supercomputadoras brasileñas aparecen en el ranking. Dos de ellas están ubicadas en el Laboratorio Nacional de Computação Científica (Rio de Janeiro) y apoyan proyectos de investigación relacionados con la biología computacional, el petróleo, agua, gas, medicina asistida por computación, modelamiento computacional, métodos numéricos, etc. La otra supercomputadora se encuentra ubicada en el campo de SENAI CIMATEC (Bahia), donde apoya diversos procesos de manufactura (automatización y control de sistemas industriales).

En el caso del Perú se ha priorizado el uso para diversas aplicaciones relacionadas al estudio del clima y ambiente. El sector ambiente, a través del Instituto Geofísico del Perú, el Instituto de Investigaciones de la Amazonía Peruana - IIAP (ambas con el apoyo del CONCYTEC a través del programa CIENCIACTIVA) y el SENAMHI, han liderado interesantes proyectos que han incluido recientemente la adquisición de pequeños equipamientos de alto desempeño computacional. Así mismo, es pertinente destacar que el uso de la supercomputación en el Perú también se ha dirigido al sector académico, a través del supercomputador de

la Universidad San Agustín de Arequipa y la Universidad San Luis Gonzaga de Ica, con los que se desarrollan trabajos de investigación de postgrado y pregrado sobre modelamiento sobre el clima, *big data*, minería de datos, *deep learnig*, procesamiento de imágenes 2D y 3D, entre otros.

A finales de enero del año 2017 se puso en servicio la supercomputadora MANATI, ubicada en el Centro de Alto Rendimiento Computacional de la Amazonía Peruana del IIAP (con sede en Iquitos). Este sistema avanzado será utilizado para el procesamiento de imágenes de teledetección, estimación de la deforestación, inundaciones, desplazamientos hidrográficos, tipos de suelos, cambios en el clima, así como el modelamiento del clima, alertas de eventos climáticos extremos, nichos ecológicos, etc. El IIAP está trabajando en proyectos relacionados a estudios de plagas en especies agroforestales, reconocimiento de especies forestales, procesamiento de imágenes para agricultura de precisión, traductores de lenguas indígenas, etc. Incluso el IIAP viene iniciando trabajos relacionados a bioinformática y química computacional.



Imagen 33: Supercomputadora MANATI - IIAP. Créditos: IIAP Iquitos

Así mismo, vale la pena mencionar que en el Perú la computación de alto rendimiento aún está en los inicios y es necesario que las universidades prioricen su enseñanza en todas las carreras científicas o que involucren creación de tecnologías. Incluso consideramos importante que el uso y la creación de aplicativos software para supercomputadoras deberían ser incluidos como contenidos obligatorios en las carreras de pregrado, por ejemplo, en cursos de matemáticas, física, análisis matemáticos, estadísticas, simulaciones y programación. Los centros de alto rendimiento computacional deberían ser laboratorios que apoyen el proceso educativo del Perú, deberían estar al servicio de las asignaturas y cursos dictados en las aulas. Esta estrategia debería incluir a centros de educación técnica, tecnológica e instituciones educativas secundarias.

CAPÍTULO 2

ELEMENTOS DE UN SISTEMA DE ALTO RENDIMIENTO COMPUTACIONAL: COMPUTADORAS, SISTEMA ELÉCTRICO, SISTEMA DE ENFRIAMIENTO Y OTROS COMPLEMENTOS

En el contexto de esta publicación definimos un sistema de alto rendimiento computacional a la integración de una serie de recursos de gran capacidad (procesadores, memoria, almacenamiento, sistemas de comunicación, etc.), instalados y en funcionamiento en un ambiente acondicionado con suficiente energía, con apropiada capacidad de refrigeración y con diversas condiciones de seguridad física, bioseguridad, sistemas anti-desastres (incendios, eventos climáticos extremos, etc.). En su conjunto deben poder facilitar a sus usuarios acceso a servicios de procesamiento de datos y contenidos.

Implementar un sistema de alto rendimiento computacional no es fácil y siempre será costoso. Mantenerlo en funcionamiento es trabajoso y caro, considerando los costos de energía para su funcionamiento. Si seleccionar la arquitectura y los mismos procesadores que compondrán una Supercomputadora ya es complicado, definir todo el sistema será más complejo aún e incluso peligroso. Lo primero que se debe definir son los usos que se le espera dar, luego es crítico conocer las capacidades eléctricas del sitio donde se espera hacerlo funcionar, igual de importante es seleccionar los recursos humanos que se encargarán de su gestión y uso.

Es importante indicar que en el Perú existe muy poco personal especializado en instalación y mantenimiento de ambientes para sistemas de alto rendimiento computacional (a cualquier técnico o ingeniero le tomará tiempo adecuarse a la gestión de este tipo de infraestructura). Y de todas maneras instalar un sistema como éste requerirá contratar nuevo personal exclusivo para la operación, monitoreo y mantenimiento del mismo.

Por otro lado es erróneo formular un proyecto de adquisición o instalación de un sistema de alto rendimiento computacional sin planificar apropiadamente las condiciones ambientales de electricidad, refrigeración y seguridad para el acceso físico a los equipos. Se han visto proyectos en los que más del 90% del presupuesto se destina únicamente a los recursos computacionales. En una visita realizada al Centro Nacional de Processamento de Alto Desempenho de la Universidad Estatal de Campinas (Campinas, Brasil), el equipo de responsables de ese centro indicaban que es normal que en los proyectos de este tipo, el costo de las adecuaciones y ambientación resulte mayor que las propias supercomputadoras. E incluso ello no acaba ahí, pues en el mismo Brasil, cuyos expertos son los líderes de la supercomputación en Latinoamérica, igual han tenido problemas relacionados a la mala planificación del sistema y consumo eléctrico.

En una publicación de la BBC de Londres en julio de 2016 [11], indican que el Laboratorio Nacional de Computación Científica (LNCC) ubicado en el Estado de Río de Janeiro, adquirió la Supercomputadora Santos Dumont con más de 1.1 petaflops de capacidad computacional (el más potente de Latinoamérica). Sin embargo este no puede ser usado por completo debido al enorme costo de electricidad que genera. Según la BBC, el monto del costo mensual de la electricidad que necesita para funcionar la Supercomputadora asciende a 150 mil dólares. La lección aprendida es que en la planificación es importante poner a consideración el costo de operación que demandará el funcionamiento.

Y en ese sentido, vuelve a ser importante la definición de equipamiento de bajo consumo de energía o con algún mecanismo o tecnología que demande bajo consumo eléctrico.

El sistema de alto rendimiento computacional requiere mínimamente lo siguiente:

- Equipamiento de alto rendimiento computacional.
- Sistema de refrigeración.
- Sistema eléctrico.
- Adecuaciones del ambiente físico.
- Software (hemos decidido tratar los aspectos relacionados al software en el Capítulo 3).



Imagen 34: Supercomputadora Santos Dumont en Laboratorio Nacional de Computación Científica, Río de Janeiro, Brasil. Créditos: Paulo Faria. Fuente: <http://www.bbc.com>.

2.1. Equipamiento para alto rendimiento computacional

2.1.1. Arquitectura de una Supercomputadora: Es el esquema, orden, relacionamiento o conexión que deberán seguir los recursos computacionales (procesador, memoria, bus de datos, computadoras, etc.) en la supercomputadora. Podemos resumir como el orden y configuración que seguirán las partes de la supercomputadora para realizar el procesamiento lo más eficiente posible.

Muchos autores mencionan que la mayoría de las supercomputadoras actuales están basadas en la arquitectura Von Neumann o arquitectura Princeton (otros autores desvirtúan el trabajo de Von Neumann incluso se menciona a Alan Turing como uno de los padres de la misma arquitectura). Para el caso de sistemas de alto desempeño computacional, según el Top500 [12] de noviembre del 2016, se mencionan apenas 2 tipos de arquitecturas. La primera, la denominada Clúster, que es un conjunto de computadoras o nodos de gran capacidad, cada uno con su propia memoria, y todos los nodos conectados a grandes velocidades. La segunda denominada MPP (Massively Parallel Processor - Procesadores Masivamente Paralelos), que son equipos cuyos procesadores se encuentran conectados por mecanismos tecnológicos de alta velocidad, cada procesador tiene su propia memoria.

Con la finalidad de explicar de la mejor manera posible lo relacionado a las arquitecturas utilizadas en los sistemas de alto desempeño computacional se ha explorado el ranking Top500 [1] desde junio del 1993, priorizando las siguientes arquitecturas: AMP, Clúster, Constellations, SMP y MPP. Por lo que seguidamente daremos más detalles de cada una:

AMP o Multiprocesamiento Asimétrico (Asymmetric Multiprocessing en inglés): En este sistema hay un procesador (maestro) donde se ejecuta el sistema operativo y los demás procesadores (esclavos) esperan que el maestro les dé instrucciones o tienen asignadas tareas predefinidas. Es una relación maestro-esclavo.

En esta arquitectura el procesador maestro planifica el trabajo de los procesadores esclavos asignándoles tareas específicas. Tiene el control sobre toda la memoria y de todos los recursos de entrada y salida. Los procesadores esclavos solo ejecutan aplicaciones de usuario. Una vez que un proceso está activo y el esclavo necesita un servicio de E/S, le envía una solicitud al maestro y espera a que el servicio se lleve a cabo.

La principal ventaja es que al aumentar más procesadores se tiene que hacer un cambio mínimo y fácil para el manejo de estos y en general se eliminan muchos problemas de integridad de datos. La gran desventaja es que al haber solo una copia del sistema operativo en un solo procesador (maestro) cuando este procesador falla todo el sistema falla porque todos los recursos que son manejados por el sistema operativo no pueden ser accedidos. Esto quiere decir que el procesador maestro puede ser un cuello de botella en el rendimiento. Los sistemas en clúster son sistemas con múltiples CPU para llevar a cabo el trabajo.

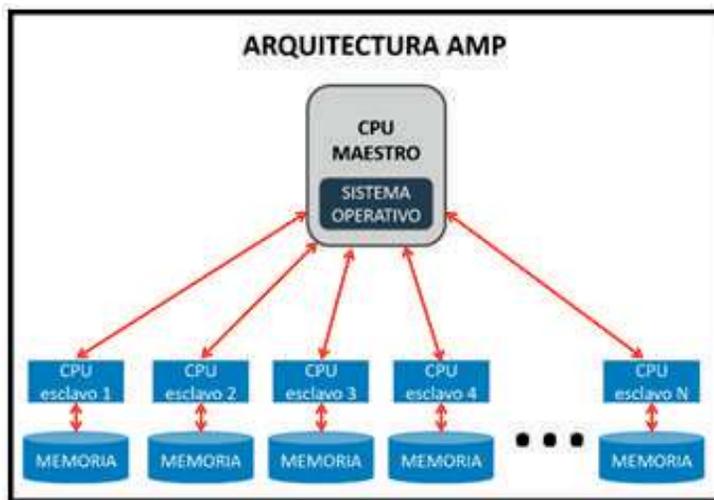


Imagen 35: Ilustración de la arquitectura AMP. Créditos: Elaboración propia.

Clúster (del inglés Cluster, grupo o racimo): conjunto o conglomerado de computadoras que funcionan de forma integrada. Cada computador en el contexto del clúster se denomina “nodo”. Normalmente los nodos están unidos por una red de alta velocidad, la finalidad de la integración es que el grupo de nodos funcionen como si fuese una única computadora.

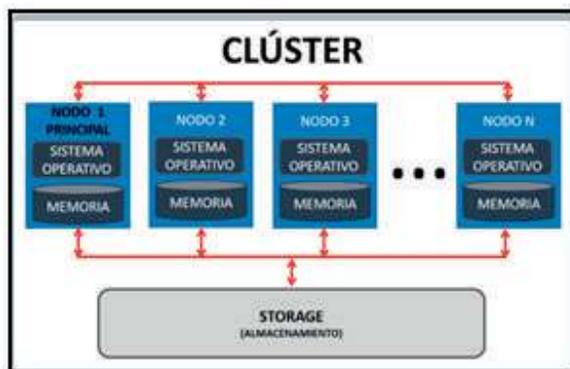


Imagen 36: Ilustración de la arquitectura Clúster. Créditos: Elaboración propia.

Los nodos del clúster comparten el almacenamiento y un sistema de conexión de alta velocidad. De acuerdo a como se estructuren las clúster puede clasificarse de la siguiente manera:

- **Clúster asimétrico:** Un nodo está en modo de espera mientras otro ejecuta las aplicaciones, monitorizándola, si el segundo falla, el primero en espera pasa a ser el nodo activo.
- **Clúster simétrico:** Dos o más nodos ejecutan aplicaciones y se monitorizan entre sí, obviamente es más eficiente.
- **Clúster en paralelo:** Permite que varios nodos procesen los mismos datos.
- **Clúster conectados a una red de área extensa:** Permiten que los nodos acceden a los mismos datos a lo largo de una red extensa (procesamiento distribuido).

SMP o Multiprocesamiento Simétrico (siglas de Symmetric Multiprocessing): Es una arquitectura donde dos o más procesadores idénticos están conectados a una única memoria principal común y son controlados por una sola instancia del sistema operativo.

Los sistemas SMP permiten que cualquier procesador trabaje en cualquier tarea sin importar donde los datos de esa tarea están ubicados en la memoria, a condición de que cada tarea en el sistema no está en ejecución en dos o más procesadores al mismo tiempo, con el apoyo adecuado del sistema operativo, los sistemas SMP puede mover fácilmente tareas entre los procesadores para equilibrar la carga de trabajo de manera eficiente.

La limitación más importante de la arquitectura SMP, en términos del rendimiento máximo que puede alcanzarse, es la posibilidad de saturación de los buses del sistema. Ya que todos los procesadores tienen acceso al mismo bus de memoria, y en general a todos los periféricos que se encuentran comunicados comúnmente por buses, conforme se incrementa el número de procesadores se incrementa también el tráfico en dichos buses. Esto causa una saturación que finalmente termina por negar el incremento de rendimiento obtenido teniendo varios procesadores. Por esta razón una arquitectura SMP difícilmente puede escalar más allá de algunas decenas de procesadores.

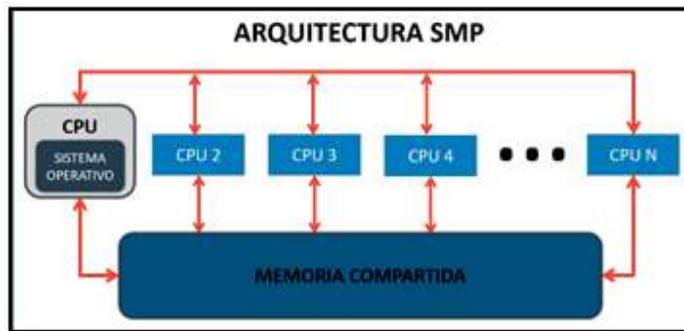


Imagen 37: Ilustración de la arquitectura SMP. Créditos: Elaboración propia.

Constellations (también llamados CLUMP o máquinas jerárquicas): Es un tipo particular de clúster. Es una arquitectura que funciona como si fuera la combinación de clúster y SMP, cada nodo de este clúster está compuesto por un SMP.

Tiene microprocesadores independientes y usan un bus compartido para comunicarse con la memoria. En esta arquitectura se tiende a gastar un tiempo considerable esperando a obtener los datos desde la memoria ya que no es el único en espera.

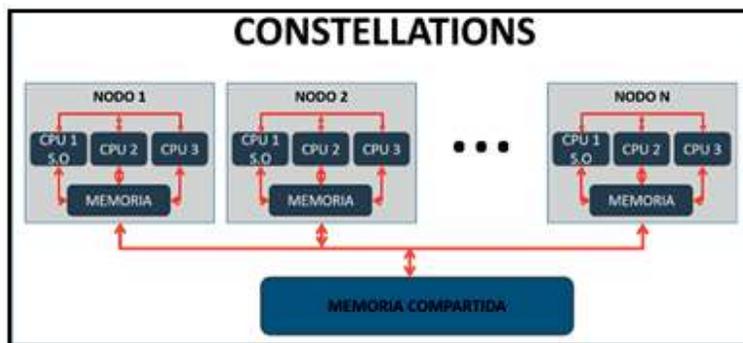


Imagen 38: Ilustración de la arquitectura Constellations. Créditos: Elaboración propia.

MPP o Procesamiento Paralelo Masivo (siglas de massively parallel processing): Esta arquitectura comprende una máquina paralela que consta de varias unidades de procesamiento básicamente independientes. Cada una de estas unidades, conocida como "nodo", es prácticamente una computadora en sí misma, contando con su propio procesador, memoria no compartida, y que se comunica con las demás unidades de procesamiento a través de un canal provisto exclusivamente para este propósito.

Su característica más importante es el hecho de que, en cada nodo, cada procesador opera básicamente como una computadora independiente, ejecutando su propio código independiente de los demás procesadores, y teniendo un área de memoria con datos también independientes. Para obtener un mayor desempeño, se requiere colaboración entre los nodos. Una máquina MPP debe contar con un canal que permita a los nodos comunicarse entre sí, a fin de intercambiar datos y coordinar sus operaciones. Ya que el objetivo principal es obtener alto rendimiento, para esto se busca

que este canal de comunicaciones sea lo más eficiente posible. En términos tanto de ancho de banda como de tiempo de latencia. En la mayoría de los casos este canal será un bus propietario, diseñado por el fabricante del equipo MPP. De esta manera se tiene una arquitectura que puede escalarse a varios cientos o miles de procesadores (las máquinas MPP más grandes en la actualidad tienen alrededor de 10 mil procesadores). Sin embargo el tener varias secciones de memoria independientes complica la programación en este tipo de arquitecturas.

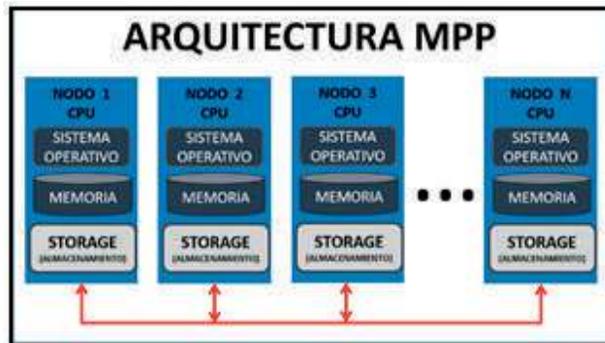


Imagen 39: Ilustración de la arquitectura MPP. Créditos: Elaboración propia.

2.1.2. Familias de Procesadores

El procesador es el elemento principal de un sistema de alto rendimiento computacional, en esencia es un conjunto de componentes electrónicos que forman un circuito integrado que se encarga de la ejecución de múltiples instrucciones.

A continuación se presentan las familias de procesadores más importantes, la lista se elabora teniendo en cuenta el Top500 [1].

Microprocesadores INTEL

Intel Nehalem (2008): Familia de procesadores de primera generación, fueron fabricados a 45 nanómetros. Esta familia proporcionaba una serie de mejoras respecto a sus predecesoras de la propia Intel. En este caso por ejemplo hasta 2,4 veces el ancho de banda de memoria, así mismo es capaz de ejecutar dos hilos en cada núcleo haciendo que el computador asigne al nuevo quad (realmente un procesador i7 de cuatro núcleos) ocho núcleos virtuales y al 8-core (con dos procesadores Xeon de cuatro núcleos).

Otra característica interesante de los Nehalem es el Turbo Boost que básicamente consiste en una tecnología que modifica dinámicamente el rendimiento del procesador variando su velocidad de reloj en función de la carga de trabajo. Es un overclocking pero realizado de forma automática cuando es necesario y la temperatura del procesador no está por las nubes.

Procesadores que componen esta familia son:

- Xeon
- Core i7 Extreme
- Core i7
- Core i5



Imagen 40: Intel Core i5 750. de la arquitectura Nehalem. Créditos: Curren.

Intel Westmere (2010) Familia de procesadores de primera generación, fueron fabricados a 32 nanómetros. Respecto a los Nehalem se incrementan dos núcleos más (seis en total), Westmere es igual a Nehalem en cargas de trabajo de un solo hilo, pero mucho más escalable gracias a los dos núcleos adicionales por dado.

De hecho, el aumento del rendimiento del 400 por ciento mostrado con las instrucciones AES-NI hace que el cifrado de disco completo sea casi imperceptible.

El Westmere se basa en las mismas directrices básicas que el controlador de memoria integrado Nehalem, caché L3 compartido por socket y QPI (Interconexión QuickPath), pero se basa en un proceso de 32nm en lugar de 45nm de Nehalem. Se ejecuta hasta 3.33GHz por núcleo, y dos hilos por núcleo con Hyper-Threading. Las CPUs Nehalem que funcionan a la misma velocidad de reloj.

Al igual que Nehalem, Westmere implementa el modo Turbo para incrementar la velocidad del reloj en ciertos núcleos dependiendo de la carga. El modo Turbo beneficia a las aplicaciones de rosca sencilla y ligeramente roscada al aumentar el rendimiento de algunos núcleos cuando sea necesario.

Procesadores que componen esta familia son:

- Gulftown (6C / 12T)
- Clarkdale (2C / 4T + iGFX)
- Arrandale (2C / 4T + iGFX)
- Clarkdale (2C / 4T + iGFX)



Imagen 41: Procesador Intel Core i7-980x de la familia Gulftown.

Intel Sandy Bridge (2011): Familia de procesadores de segunda generación, también fueron fabricados a 32 nanómetros. Es la familia sucesora del Intel Nehalem e Intel Westmere.

Entre las principales mejoras tenemos:

- Es Multinúcleo: procesadores de hasta 6 núcleos. En algunos modelos Intel incorpora la tecnología Hyperthread que permite procesar 2 tareas en cada núcleo
- Soporte de instrucciones: soporta instrucciones que trabajan a 256 bits, como las AVX. Ello permiten trabajar con cálculos complejos.
- Tarjeta gráfica integrada: Por primera vez integración real CPU GPU.
- Turbo (Turbo Boost.): Permite realizar un mayor número de tareas de manera simultánea, sin poner en riesgo que se queme el procesador.
- Quick Sync: Sirve para acelerar la creación, edición y conversión de vídeos.
- Overclock: Sirve para manejar la frecuencia de funcionamiento.

Procesadores que componen esta familia son:

- Core i7 Extreme
- Core i7 Extreme Edition
- Core i7
- Core i5
- Core i3
- Pentium
- Celeron
- Xeon E5
- Xeon E3

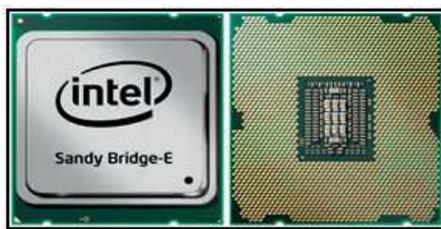


Imagen 42: Procesador Intel Sandy Bridge-E. Créditos: The Mr Friki

Intel IvyBridge (2012): Familia de procesadores de tercera generación, también fueron fabricados a 22 nanómetros. Es la familia sucesora del Intel Sandy Bridge. La cantidad de núcleos va entre 2 y 4 según sea la orientación de uso del procesador (escritorio, portátil).

Entre las principales mejoras tenemos:

Mayor cache: Al ocupar menor espacio, permite que se pueda aumentar la memoria cache y mejorar por tanto la velocidad de acceso a la memoria RAM. En algunos artículos comentan mejoras del 10 al 15 % de rendimiento a la misma frecuencia de funcionamiento.

Consumo: Al ser los transistores más pequeños consumen menos energía. De esta forma este cambio de tecnología es perfecto para aquellos aparatos portátiles que dependen de una batería para funcionar.

Mayor cantidad de transistores: Al ocupar menor espacio permite también disponer de más transistores que la mayoría de procesadores de la familia Intel Sandy Bridge.

Tarjeta gráfica integrada: Se añade soporte para Directx 11. Sus prestaciones son mayores. Y con ello mejor el rendimiento gráfico.

Procesadores que componen esta familia son:

- Core i3
- Core i5
- Core i7



Imagen 43: Procesador Intel de la arquitectura Ivy Bridge. Créditos: monomouth.com

Intel Haswell (2013): Familia de procesadores de cuarta generación, también fueron fabricados a 22 nanómetros. Es la familia sucesora del Intel IvyBridge. La cantidad de núcleos va hasta 8 núcleos. Así mismo es válido comentar que mantiene las características de sus predecesores (Sandy e Ivy Bridge) en cuanto a HyperThreading y Turbo Boost.

Entre las principales mejoras tenemos:

- **Menor consumo de energía:** Reducen el consumo de energía tanto en estado de actividad —manteniendo, o incluso mejorando el rendimiento— como en estado de reposo, en el que además se reduce el tiempo de espera al cambiar a un estado activo.
- **Mejoran la modularidad:** Orientado a posibilitar la implantación en dispositivos de menor tamaño, como tablets y smartphones.
- **Chipset dentro del procesador:** Lo que permite optimizar los mecanismos de uso de energía por parte del hardware, de tal manera, que equipos de bajos recursos como un Tablet o un Netbook.
- **Mejora de soporte gráfico en hardware:** Soporte para OpenGL 4.0 y Direct3D 11.1.

Los procesadores que componen esta familia son:

- Intel Core i7-4770K, Intel Core i7-4770, Intel Core i7-4770S, Intel Core i7-4770T, Intel Core i7-4770R, Intel Core i7-4765T.
- Intel Core i5-4670K, Intel Core i5-4670, Intel Core i5-4670S, Intel Core i5-4670T, Intel Core i5-46570, Intel Core i5-4570S, Intel Core i5-4570T
- Intel Core i7-4930MX, Intel Core i7-4900MQ, Intel Core i7-4800MQ, Intel Core i7-4950HQ, Intel Core i7-4850HQ, Intel Core i7-4650U, Intel Core i5-4350U



Imagen 44: Intel Haswell i7-4771. Créditos: Dsimic

Intel Broadwell (2014): Familia de procesadores de quinta generación, fabricados a 14 nanómetros. Creada para ser la sucesora de la Haswell, Alcanza hasta 22 núcleos (aunque la cantidad de núcleos en función a la presentación: Mainstream va de 2 a 4 núcleos, Enthusiast de 6 a 10 núcleos y Xeon de 4 a 22 núcleos).

Entre las principales mejoras tenemos:

- **Procesador más eficiente:** Al haber reducido a 14 nanómetros cada transistor, se incrementa el número de transistores, se reduce el consumo de energía y se incrementa la capacidad de procesamiento.
- **Facilidad de adaptación para trabajar en dispositivos móviles:** Intel confirma su interés por los dispositivos móviles, sus chips fueron fabricados pensando en ser utilizadas en tablets y computadoras portátiles.
- **Mejora funcionalidades gráficas:** El codificador de video de hardware Intel Quick Sync de Broadwell agrega la decodificación de hardware VP8 y el soporte de codificación.
- **Mejora de rendimiento gráfico:** La GPU integrada de Broadwell admite Windows Direct3D 11.2, OpenGL 4.4 (OpenGL 4.5 en Linux) y OpenCL 2.0.

Procesadores que componen esta familia son:

- Broadwell-Y
- Broadwell-U
- Broadwell-H
- Broadwell-K
- Broadwell EP y EX

Microprocesadores AMD (siglas en inglés de Advanced Micro Devices, Inc.)

Es el segundo productor más importante de procesadores, con sede en California (Estados Unidos). Sus productos principales incluyen microprocesadores, chipsets para placas base, circuitos integrados auxiliares, procesadores embebidos y procesadores gráficos para servidores, estaciones de trabajo, entre otros.

Se ha identificado el uso del procesador AMD x86-64 en algunos sistemas de alto rendimiento computacional. Si bien es cierto la arquitectura x86 es originaria de Intel sin embargo AMD ha desarrollado mejoras. En el año 2003 AMD lanzó los AMD64, un sistema basado en x86 pero para manejar estructuras de 64-bits.

Seguidamente presentamos un listado de la familia de AMD, y es importante indicar que muy pocas fueron utilizadas para computación de alto desempeño.

- AMD Phenom 64 Quad
- AMD Athlon 64
- AMD Athlon 64 X2
- AMD Athlon 64 FX
- AMD Opteron
- AMD Turion 64 X2
- AMD Turion 64
- AMD Sempron (todos los que utilizan sockets 754, 939 y AM2)
- AMD Phenom (K10)
- AMD Phenom II
- AMD Athlon II
- AMD FX (Bulldozer, Piledriver, Steamroller, Excavator)



Imagen 45: Procesador AMD Phenom II N950 advantages. Créditos: CPU world

Microprocesadores SW de China(siglas en inglés de SUNWAY)

Esta es una familia de procesadores chinos creados por el Centro Nacional de Diseño de Circuitos Integrados de Alto Rendimiento en Shanghai. Los procesadores Sunway es una propia creación china que ha permitido a este país actualmente liderar la Supercomputación en el mundo, supera ampliamente la cantidad de núcleos de CPU de los otros fabricantes.

Procesadores SW1600: Lanzadas el año 2010, estos procesadores funcionan con 16 núcleos CPU, alcanzando 140,8 gigaflops.

Son parte de esta familia los procesadores:

- Sunway SW-1
- Sunway SW-2
- Sunway SW-3
- Sunway SW26010

Procesadores SW26010: El SW26010 es un procesador de 260 núcleos que ha permitido a China a través de Sunway TaihuLight el supercoputador de mayor capacidad del mundo.

Son parte de esta familia los procesadores:

- Sunway SW-1
- Sunway SW-2
- Sunway SW-3
- Sunway SW1600

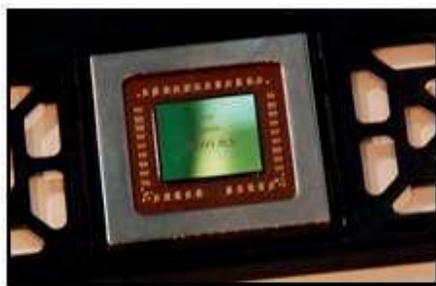


Imagen 46: Procesador Sunway SW1600. Créditos: Teleobjetivo.org

Procesadores de IBM (sigla en inglés de International Business Machines)

IBM es una empresa multinacional fabricante de tecnologías de la información con sede en Nueva York (Estados Unidos). La familia más importante de procesadores que fabrica IBM son las POWER (ello viene desde los inicios de los 90s). Pueden llegar a utilizar 18 núcleos.

En el Top500 del noviembre del 2016 [1] se identificó al menos una supercomputadora con procesadores POWER7 de la familia de los POWER de IBM.

Este POWER7 cuenta con ocho núcleos que corren hasta con cuatro hilos cada uno, transformado virtualmente cada procesador en un chip de 32 núcleos y dándole una clara ventaja sobre cualquier de los productos Intel o AMD para servidores.

Otros procesadores de la familia de los POWER son los siguientes:

- POWER1+ (1991)
- POWER1++ (1992)
- RSC (1992)
- POWER2 (1993)
- POWER2+ (1994)

- POWER2 (1997)
- RAD6000 (1997)
- POWER3 (1998)
- POWER3-II (1999)
- POWER4 (2001)
- POWER4+ (2002)
- POWER5 (2004)
- POWER5+ (2005)
- POWER6 (2007)
- POWER6+ (2009)
- POWER7 (2010)
- POWER7+ (2012)
- POWER8 (2014)

Por otro lado también se identificaron unos 20 procesadores Power BQC, perteneciente a la familia PowerPC (Cuya abreviación es PPC). En este caso estos procesadores están basados en la arquitectura RISC (del inglés Reduced Instruction Set Computer, en español Computador con Conjunto de Instrucciones Reducidas).

Otros Procesadores PowerPC son los siguientes:

- PowerPC 601
- PowerPC 7447
- PowerPC 970



Imagen 47: Procesador IBM PowerPC 601. Créditos: www.cideacme.uaeh.edu.mx

SPARC (en inglés significa Scalable Processor Architecture).

Inicialmente fueron desarrollados por SUN Microsystems (esta empresa fue adquirida por ORACLE). Previo a su venta Sun Microsystems licenció a fabricantes como Texas Instruments, Cypress Semiconductor, Fujitsu, LSI Logic entre otros. Esta familia de procesadores utiliza arquitectura RISC (del inglés Reduced Instruction Set Computer, en español Computador con Conjunto de Instrucciones Reducidas).

Características:

- Utiliza ventanas de registro.
- 32 registros de enteros de 32 bits.

- 16 registros de punto flotante de 64 bits (para doble precisión), que se pueden utilizar como registro de 32 bits (para precisión simple).
- Modos de direccionamiento
- Inmediato
- Directo
- Indirecto
- Utiliza instrucciones retardadas.
- Manejo de Memoria
 - Espacio virtual de 4 Gigabytes.
 - Unidad de manejo de memoria (MMU) que trabaja con páginas de tamaño configurable.

En el Top500 de noviembre de 2016 [1] se incluye hasta 7 Supercomputadoras utilizando SPARC64 VIIIfx, esta arquitectura fue desarrollada por Fujitsu, y justamente las 7 vienen funcionando en Japón. Esta tecnología posee transistores de 45 nanómetros y 8 núcleos.

Otros procesadores de la familia SPARC

- Micro SPARC I
- SuperSPARC I
- SPARCite
- hyperSparc
- TurboSPARC
- UltraSparc
- SPARC64
- Galaxy FT-1500
- LEON4
- R1000
- SPARC T4
- SPARC M5
- SPARC M6
- SPARC M7
- SPARC S7

2.1.3. Coprocesadores matemáticos y aceleradores

Se podría indicar que los aceleradores son la evolución de los antiguos coprocesadores matemáticos, ambos son básicamente apoyo o complemento del procesador central.

Según Alan Clements: “un coprocesador es una unidad de procesamiento auxiliar utilizada para complementar el poder de cómputo de un microprocesador principal. En síntesis, un coprocesador es un accesorio que provee funcionalidad que no puede incorporarse en el procesador principal, ya sea por razones económicas o técnicas” [13].

Según el portal SOFTDOIT [14] “un coprocesador es un microprocesador de un ordenador utilizado como suplemento de las funciones del procesador principal (la CPU). Las operaciones ejecutadas por uno de estos coprocesadores pueden ser operaciones de aritmética, procesamiento gráfico, de señales, procesamiento de texto o criptografía, etc”. Se usa como complemento o apoyo al procesador principal, esto reduce su carga y mejora el rendimiento del procesamiento.

El uso de los coprocesadores ha sido clave para el crecimiento de la computación de alto rendimiento. En la actualidad es de masivo y va en constante crecimiento. En el caso del MANATI IIAP se utilizan aceleradores NVIDIA K80. Los principales coprocesadores de aceleradores son los siguientes:

Aceleradores NVIDIA (creado en 1993, California, Estados Unidos)

Es un fabricante de tarjetas de procesamiento gráfico para diversos dispositivos y en especial para computación de alto rendimiento. Si bien es cierto al inicio empezó con la fabricación de hardware para juegos en la actualidad es el principal fabricante de procesadores gráficos GPU para las plataformas del Top500 (a noviembre del 2016).

Incluso NVIDIA en 1999 revolucionó la computación a través de la creación de las GPU (anteriormente las tarjetas de video fueron lo más parecido), en la actualidad sus competidores vienen manejando el mismo concepto. Los principales aceleradores nvidia son:

NVIDIA Tesla K80

Es el acelerador de aplicaciones más potente del mundo de la alta computación y diseñado para una amplia variedad de aplicaciones de aprendizaje automático, análisis de datos, cálculo científico y alta computación (HPC). La Supercomputadora MANATI del IIAP - Iquitos está implementada con ésta tecnología.

Nucleos Cuda	4992 núcleos CUDA en dos GPUs
Rendimiento Precisión doble	Hasta 2,91 Teraflops de rendimiento en operaciones de precisión doble con NVIDIA GPU Boost.
Rendimiento Precisión simple	Hasta 8,73 Teraflops de rendimiento en operaciones de precisión simple con NVIDIA GPU Boost.
Memoria	24 GB de memoria GDDR5.
Ancho de banda	480 GB/s de ancho de banda de memoria agregado.

Tabla 05: Especificaciones de NVIDIA Tesla K80.

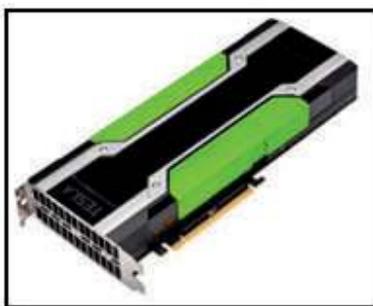


Imagen 48: Tarjeta NVIDIA Tesla K80. Créditos: www.nvidia.es

NVIDIA Titan Xp

Tabla 06: Especificaciones de NVIDIA Titan Xp

Arquitectura de GPU	Pascal
Memoria de vídeo	12 GB G5X
Frecuencia de la memoria	11.4 Gbps
Frecuencia acelerada	1582 MHz
Núcleos GPU	3840



Imagen 49: Tarjeta gráfica NVIDIA Titan Xp. Créditos: www.nvidia.es

GeForce GTX

Tabla 07: Especificaciones de GeForce GTX

	GeForce GTX 1080	GeForce GTX 980
Arquitectura de GPU	Pascal	Maxwell
Memoria de vídeo	8 GB GDDR5X	4 GB GDDR5
Frecuencia de la memoria	10 Gbps	7 Gbps
Frecuencia acelerada relativa	1x	1.4x
Frecuencia acelerada real	1733 MHz	1216 MHz
Núcleos GPU	2560	2048
GPU	GP104-400	GM204-400



Imagen 50: Tarjeta gráfica GeForce GTX 1080. Créditos: www.nvidia.es

Quadro GP100

Tabla 08: Especificaciones del Quadro GP100.

	Quadro GP100
Núcleos GPU	3,584
Interfaz de memoria	4096-bit.
Ancho de Banda de Memoria	Hasta 717 GB/s
Memoria de GPU	16 GB HBM2.
NVIDIA NVLINK™	Compatible con 2 tarjetas GP100
Interfaz del sistema	PCI Express 3.0 x16.
Consumo máximo de energía	235 W
Solución térmica	Activa
Factor de forma	4.4" de alto x 10.5" de largo, doble Ranura, altura completa.
Conectores de pantalla	4x DP 1.4, 1x DVI-D DL
Cantidad máxima de pantallas simultáneas	4 directas, 4 DP 1.4 Multi- Stream.
Resolución de pantalla	4x 4096 x 2160 a 120Hz 4x 5120 x 2880 a 60Hz.
API gráficas	Shader Model 5.1, OpenGL 4.56, DirectX 12.07, Vulkan 1.06.
API de cómputo	CUDA, DirectCompute, OpenCL

Fuente: www.nvidia.com



Imagen 51: Tarjeta gráfica Quadro GP100. Créditos: www.nvidia.es

NVIDIA Tegra X1

Tabla 09: Especificaciones del procesador Tegra X1

	Tegra X1
Núcleos GPU	256 núcleos de GPU NVIDIA.
CPU	CPU ARM® de 64 bits y 4 núcleos 4 A57 con 2 MB de L2.
VÍDEO	Video 4K a 60 fps con H.265 y VP9 4k con compresión H.265, VP9 y H.264
POTENCIA	SOC de 20 nm – TSMC Raíles aislados, alternancia de grupos de núcleos (cluster switching) de cuarta generación.
PANTALLA	4K x 2K a 60 Hz, 1080p a 120 Hz HDMI 2.0 a 60 fps, HDCP 2.2

Fuente: www.nvidia.com



Imagen 52: Procesador NVIDIA Tegra X1. Créditos: www.nvidia.es

NVIDIA Tegra K1

El procesador Tegra K1 está diseñado con la misma arquitectura NVIDIA Kepler™ en la que se basan los PCs de juego más avanzados del planeta y el superordenador más rápido de los EE.UU. Es la primera tecnología jamás creada para móviles que posee la capacidad de las consolas, lo que permite introducir tecnologías de juego de PC tales como DirectX 11, OpenGL 4.4, y la función de teselado (recubrimiento de superficies planas con figuras geométricas sin dejar espacio ni superposiciones).

Tabla 10: Especificaciones del procesador Tegra K1

	Tegra K1
Núcleos GPU	192 NVIDIA CUDA Cores.
CPU Cores and Architecture	NVIDIA 4-Plus-1™ Quad-Core ARM Cortex-A15 "r3".
Memory Type	DDR3L and LPDDR3.
Max Memory Size	8 GB (with 40-bit address extension).

Fuente: www.nvidia.com

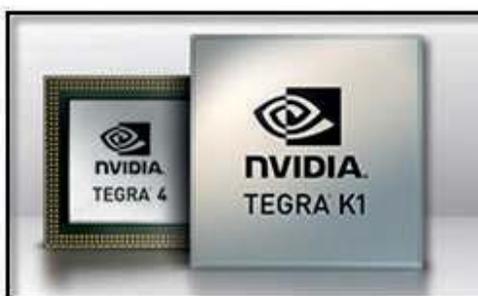


Imagen 53: Procesador NVIDIA Tegra K1. Créditos: www.nvidia.es.

NVIDIA Tegra 4

GPU NVIDIA con hasta 72 núcleos personalizados, innovaciones únicas para dispositivos móviles en fotografía, multimedia, juegos y la web, incluyendo imágenes con High Dynamic Range (HDR), WebGL y HTML5. Procesador ARM con cuatro núcleos - El procesador NVIDIA Tegra 4 aprovecha los núcleos más avanzados de la CPU que ARM ha desarrollado, además de un núcleo de segunda generación para ahorro de la batería con el fin de brindar niveles excepcionales de rendimiento y duración de la batería.

La CPU ARM Cortex-A15 es el motor que respalda Tegra 4, mientras que Tegra 4i está potenciado por la nueva CPU ARM Cortex-A9 r4 (definida por ARM con la ayuda de NVIDIA) y el núcleo de CPU más eficiente de su clase.

Tabla 11: Especificaciones de Tegra 4

	Tegra 4	Tegra 4i (procesador móvil con módem LTE 4G integrado)
Núcleos de CPU	4 + 1	4 + 1
Arquitectura de CPU	ARM Cortex-A15	ARM Cortex-A9 r4
Máxima frecuencia de reloj	1,9 GHz	Hasta 2 GHz
Núcleos de GPU	72	60
Arquitectura de fotografía computacional	✓	✓
Frecuencia de la memoria	DDR3L y LPDDR3	LPDDR3
Cantidad de memoria	4 GB	2 GB

Fuente: www.nvidia.com



Imagen 54: Procesador NVIDIA Tegra 4. Créditos: www.nvidia.es

NVIDIA Tesla P100

Tabla 12: Especificaciones Tesla P100

Arquitectura GPU	NVIDIA Pascal
Núcleos GPU	3584
Double Precision Performance	4.7 teraflops
Single-Precision Performance	9.3 teraflops
Half-Precision Performance	18.7 teraflops
GPU Memory	16GB CoWoS HBM2 at 732 GB/s or 12GB CoWoS HBM2 at 549 GB/s
System Interface	PCIe Gen3
Max Power Consumption	250 W

Fuente: www.nvidia.com

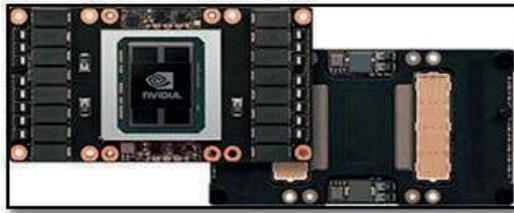


Imagen 55: Tarjeta NVIDIA TESLA P100 para HPC de escalabilidad fuerte. Créditos: www.nvidia.es

Aceleradores INTEL

Si bien cierto en líneas anteriores se mencionó las familias de procesadores INTEL. En este caso la presentaremos la familia de aceleradores de Intel Xeon Phi, cuya descripción es la siguiente:

Acelerador Intel Xeon Phi (2012). Esta es una familia de INTEL creada para trabajar como coprocesador (asistente del CPU central) para procesamiento numérico en coma flotante, incluso está basado en el modelo de las GPU (uso de hardware gráfico para solucionar problemas en general). Intel en su sitio web sobre el Intel Xeon Phi [14], dice lo siguiente: “Es un procesador host de inicio automático que proporciona enorme paralelismo y vectorización para admitir las aplicaciones de informática de alto desempeño más exigentes. La arquitectura integrada con consumo eficiente de energía proporciona significativamente mayor cantidad de procesos por unidad de energía consumida, en comparación con plataformas semejantes, para ofrecerle un costo total de propiedad mejor. La integración de memoria y estructura está por encima del máximo de memoria y reduce costos para ayudarlo a resolver sus desafíos más grandes con mayor rapidez.”

En junio del 2016 Intel anunció su procesador Intel® Xeon Phi™ 7290, que cuenta con 72 núcleos que alcanzan 3 teraflops

Procesadores que componen esta familia son:

- Procesador Intel® Xeon Phi™ 7250
- Procesador Intel® Xeon Phi™ 7230
- Procesador Intel® Xeon Phi™ 7210
- Procesador Intel® Xeon Phi™ 7290
- Procesador Intel® Xeon Phi™ 7290F
- Procesador Intel® Xeon Phi™ 7250F
- Procesador Intel® Xeon Phi™ 7230F
- Procesador Intel® Xeon Phi™ 7210F

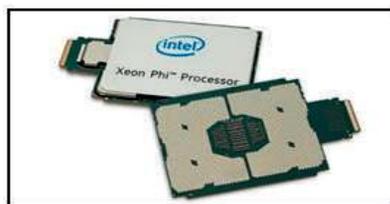


Imagen 56: Procesador Intel Xeon Phi. Créditos: www.computerhoy.com

Acelerador Advanced Micro Devices, Inc. (creada en 1969, California, Estados Unidos)

Junto a Intel es una de las compañías más grandes del mundo en producción de microprocesadores compatibles x86 y uno de los más importantes fabricantes de CPUs, GPUs, chipsets y otros dispositivos semiconductores.

Las GPU discretas incorporadas de alto rendimiento de AMD proporcionan un rendimiento gráfico excepcional Para sistemas de visualización únicos y múltiples sin comprometer el consumo de energía o el coste. Idealmente adaptado para dispositivos portátiles de imágenes médicas, instalaciones de señalización digital y juegos de casino y arcade.

AMD Embedded Radeon™ E9260 MXM Module

Tabla 13: Especificaciones AMD Embedded Radeon E9260 MXM Module.

Rendimiento de punto flotante	2,150.4 gigaflops
CORE	896
Unidades de cómputo:	14
Tamaño de la memoria	4096 megabytes
Tipo de memoria	GDDR5
Bus de Memoria	128 bits
GPU	Baffin

Fuente: www.amd.com

AMD Embedded Radeon™ E8870 MXM Module

Tabla 14: Especificaciones AMD Embedded Radeon E8870 MXM Module

Rendimiento de punto flotante	1,536 gigaflops
CORE	768
Unidades de cómputo:	12
Tamaño de la memoria	4,096 megabytes
Tipo de memoria	GDDR5
GPU	Saturn

Fuente: www.amd.com

AMD Embedded Radeon™ E8860 GPU MCM Module

Tabla 15: Especificaciones de AMD Embedded Radeon E8860 GPU MCM Module.

Rendimiento de punto flotante	768 gigaflops
CORE	640
Unidades de cómputo	10
Tamaño de la memoria	2048 megabytes
Tipo de memoria	GDDR5

Fuente: www.amd.com

GPU de servidores AMD FirePro™ S9300 x2

Acelera las cargas de trabajo HPC más complejas relacionadas con el análisis de datos o el procesamiento sísmico con el acelerador de GPU de cómputo con precisión simple más rápido del mundo, la GPU de servidores AMD FirePro™ S9300 x 2.24.



Imagen 57: GPU AMD FirePro S9300 x2. Créditos: www.amd.com

GPU del servidor AMD FirePro™ S9150

Hasta 5,07 teraflops de rendimiento pico de precisión simple y 2,53 teraflops de rendimiento de cómputo pico de precisión doble a un consumo máximo de energía de 235 vatios. Incluso las cargas de trabajo con cálculos más intensivos no representan un desafío para la GPU del servidor más potente del mundo.



Imagen 58: GPU AMD FirePro™ S9150. Créditos: www.amd.com

GPU del servidor AMD FirePro™ S9170

Acelera las cargas de trabajo más complejas en cálculos científicos, análisis de datos o procesamiento sísmico, y presencia la potencia de la primera GPU con servidor de 32 GB del mundo, la AMD FirePro™ S9170. Mantente a la vanguardia aprovechando OpenCL™ 2.0, el estándar de programación abierta para cálculos de propósitos generales en sistemas heterogéneos.



Imagen 59: GPU AMD FirePro™ S9170. Créditos: www.amd.com

Aceleradores PEZY Computing K.K (creada en 2010, Tokyo, Japón)

Es una empresa japonesa fabricante de coprocesadores para ser utilizadas en computación de alto desempeño. Incluso su tecnología figura en equipos del Top500 de noviembre del 2016 [1].

Pézy-SC

Pézy-SC tiene 1024 núcleos y ser fabricada por nodo de proceso de 28nm estado de la técnica. Además, Pézy pronto estará listo para DRAM-Logic 3-D apilamiento paquete por el principio mejorada y implimented 2,048bit ultra ancha IO SI SDRAM.

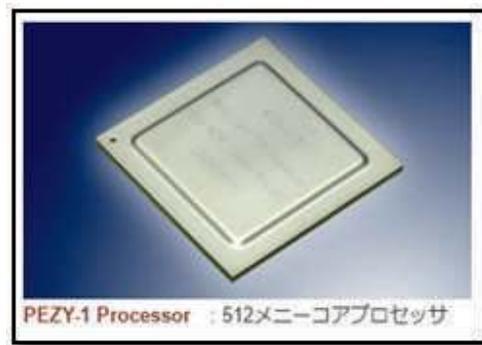


Imagen 60: Acelerador PEZY-1. Créditos: <http://pezy.co.jp>

Tabla 16: Características de Acelerador Pezy-1

Logic Cores(PE)	1,024
Core Frequency	733 MHz
Peak Performance	Floating Point Single 3.0 teraflops / Double 1.5 teraflops.
Host Interface	PCI Express GEN 3.0 x8 Lane x 4 Port (x16 bifurcation available) JESD204B Protocol support.
DRAM Interface	DDR4, DDR3 combo 64 bit x 8 Port Max B/W 1533.6 gigabytes/s +Ultra WIDE IO SDRAM (2 048 bit) x 2 Port Max B/W 102.4 gigabytes/s
Control CPU	ARM926 2 core
Process Node	28 nm
Package	FCBGA 47.5mm x 47.5mm, Ball Pitch 1mm, 2,112 pin.

Fuente: <http://pezy.co.jp/>

2.1.4. Comunicaciones de alta velocidad

Protocolo INFINIBAND (año 2001): Llamado anteriormente System I/O (Sistema de entrada/salida). INFINIBAND es un sistema de comunicaciones que facilita la conexión de computadoras de alto desempeño (por ejemplo en la conexión de nodos de clúster). INFINIBAND es una interconexión generalizada, de baja latencia y alto ancho de banda, que requiere una sobrecarga de procesamiento baja y es ideal para transportar múltiples tipos de tráfico (agrupación, comunicaciones, almacenamiento, administración) a través de una sola conexión.

INFINIBAND fue desarrollado por la IBTA (Infiniband Trade Association). A partir de 2002 en un número de vendedores ha comenzado a ofrecer sus productos basados en INFINIBAND. Su tecnología es el resultado de la unión de dos diseños en competencia (el Future I/O, desarrollado por Compaq, IBM y Hewlett-Packard, y el Next Generation I/O, desarrollado por Intel, Microsoft y Sun Microsystems). Además es importante comentar que en diversas fuentes se indica que el año 2012 el fabricante INTEL le compró a QLOGIC sus patentes y trabajos relacionado a INFINIBAND, la idea de INTEL es mejorar su modelo de negocio respecto al soporte para computación de alto desempeño.

A la fecha es una de las tecnologías de mayor aceptación incluso en el Top500 de noviembre del 2016, donde el 37% de los sistemas de alto rendimiento utiliza INFINIBAND. En el mencionado ranking además se presentan implementaciones como: INFINIBAND EDR/FDR, INFINIBAND FDR, INFINIBAND FDR14, INFINIBAND QDR. Por último en la tabla 17 presentamos el ancho de banda que logra INFINIBAND según los enlaces que logra integrar.

Tabla 17: Ancho de banda teórico de INFINIBAND en GB/s.

	SDR	DDR	QDR	FDR-10	FDR	EDR
1 Enlace	0.3	0.5	1	1.25	1.71	3.13
4 Enlaces	1	2	4	5	6.82	12.5

Gigabit Ethernet (IEEE 802.3, 1998)

Es una tecnología de alta velocidad utilizada para conectar equipos de alto rendimiento computacional, fue la evolución del típico Ethernet (de gran aceptación en redes locales). Gigabit Ethernet surge dado al interés que el comité del IEEE 802.3 tomó ante el paso de la Ethernet (10 Mbps) a FastEthernet (100 Mbps). Gigabit Ethernet también suele emplearse como troncal para conectar diversas redes Ethernet de menor velocidad. En el Top500 encontramos las 2 siguientes variantes de Gigabit Ethernet:

10G Ethernet: En junio de 2002 se ratificó el estándar IEEE 802.3ae de Ethernet de 10 Gigabit, que continúa con lo ya establecido por anteriores estándares de Ethernet de separación entre la capa MAC (Media Access Control) y la capa Phy (Physical Media Layer). En un principio, los estándares de 10 Gigabits fueron especificados para interfaces de fibra óptica exclusivamente. Pero, más recientemente, se ha acordado el estándar 10GBase-CX (IEEE 802.3ak) para utilizar cableado de cobre de alta calidad desarrollado inicialmente para el estándar de cableado de centros de datos INFINIBAND. también se está desarrollando un estándar a utilizar con cable de par trenzado sin apantallar (Cat. 5 o Cat. 7).

40G Ethernet: En el año 2010 a través de la norma IEEE 802.ba se estableció el ethernet de 40/100 Gigabit a través de fibra, que un estándar que permite la transferencia de Ethernet enmarca a velocidades de hasta 40 gigabits por segundo (Gbps). El estándar 40 GbE está destinado a un servidor de conectividad local, ejecutándose en Quad Small Form Factor Pluggable (QSFP) cableado, conector de fibra de alta densidad con 12 hebras de fibra.

2.1.5. Sistemas de alta velocidad: Cray Gemini Interconnect y Aries Interconnect

Estos 2 sistemas fueron adquiridos por INTEL a Cray en el año 2012 (hasta ese entonces ambas tecnologías de menor latencia y mayor ancho de banda).

Cray Gemini Interconnect, puede transferir 64 bytes de datos en cada dirección cada 5 ciclos, lo que corresponde a un ancho de banda máximo por dirección de $64 \times 650/5 = 8.3$ GBytes / s. Cray Gemini Interconnect tiene 10 conexiones de red, de 12 carriles en cada dirección que funcionan a 3.125 a 6.25 GHz, dando el enlace. Fiabilidad a nivel de enlace y enrutamiento adaptativo. El ancho de banda de inyección máximo por nodo es de 20 GB / s.

Aries Interconnect (sucesos de Cray Gemini Interconnect) proporciona mejoras sustanciales en las métricas de rendimiento de red estándar para la computación de alto rendimiento: ancho de banda, latencia, velocidad de mensajes y escalabilidad. Además, el diseño de Aries aborda un desafío clave en redes de alto rendimiento: cómo proporcionar un ancho de banda global rentable y escalable. Un solo dispositivo Aries Interconnect proporciona la conectividad de red para los cuatro nodos del modelo Cray XC. la Supercomputadora Titan (tercero en el Top500 de noviembre del 2017).

BULL de ATOS (1997, Francia)

Atos SE (Societas Europaea) es un líder global en servicios digitales con oficinas en 72 países. Proporciona a su cartera global de clientes servicios TI de Consultoría e Integración de Sistemas, Operaciones Gestionadas y BPO, operaciones Cloud, soluciones de *Big Data* y Ciberseguridad, abarca sectores como Defensa, Servicios Financieros, Salud, Manufacturing, Media, Utilities, Sector Público, Retail, Telecomunicaciones y Transporte. Ejerce sus actividades bajo las marcas Atos, Atos Consulting, Atos Worldgrid, Bull, Canopy, Unify y Worldline.

El año 2014 ATOS SE adquirió la empresa BULL, que hasta ese entonces era una de las empresas europeas líder en comunicaciones de alta velocidad en dicho contiene (tecnologías que incluso son utilizadas en sistemas de alto rendimiento que actualmente forman parte del TOP500). BULL previamente lanzó al mercado tecnologías como la Bull ExaScale o BXI, que es una tecnología de conexión de alta velocidad orientada a computación de alto desempeño. Introduce un cambio de paradigma en términos de rendimiento, escalabilidad, eficiencia, fiabilidad y calidad de servicio para cargas de trabajo extremas.

INTEL TRUE SCALE Y INTEL OMNI PATH

Desde hace décadas la empresa INTEL también busca liderar las tecnologías de altas velocidades para soporte de computadores de alto desempeño. De esta forma en el Top500 de noviembre del 2016 se mencionan los siguientes:

Intel True Scale Fabric está diseñado específicamente para ejecutar aplicaciones HPC / MPI, así como para aprovechar al máximo la última tecnología de procesador con implementaciones multi-core. Utiliza una biblioteca optimizada del HPC llamada PSM (Performance Scaled Messaging) para comunicaciones MPI. PSM es una biblioteca "ligera" que es específicamente construida para optimizar los requisitos de rendimiento MPI. Las principales medidas del rendimiento de HPC son el rendimiento de la tasa de mensajes MPI, la latencia de extremo a extremo, rendimiento colectivo y rendimiento de la aplicación.

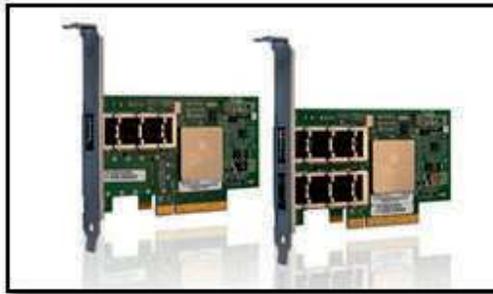


Imagen 61: Adaptador Intel True Scale Fabric. Créditos: Intel

Intel Omni-Path Architecture (Intel OPA), ofrece el desempeño para cargas de trabajo del futuro en HPC y la capacidad de escalar a decenas de miles de nodos y más con el tiempo, a un precio competitivo. Soporta una variedad de métodos de enrutamiento, en lugar de enviar todos los paquetes desde un origen a un destino a través de un único camino, por dispersión de enrutamiento distribuye el tráfico en varias rutas. La principal ventaja es que reduce la optimización de los flujos de la variación en la latencia visto a través de la red por el tráfico de alta prioridad en la presencia de tráfico de menor prioridad.

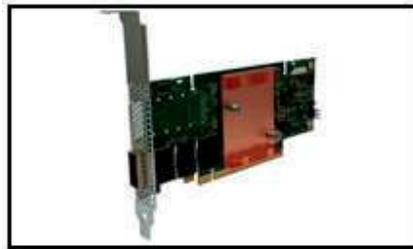


Imagen 62: Adaptador Intel Omni-Path Host Fabric Interface. Créditos: Intel

TOFU INTERCONNECT DE FUJITSU (JAPON)

FUJITSU es una empresa japonesa líder en fabricación de tecnologías de la información y comunicación, ofreciendo Soluciones y servicios de TI, Servidores, Almacenamiento, Equipos de comunicaciones y dispositivos electrónicos.

FUJITSU desarrolló Tofu Interconnect, que fue impulsado por FUJITSU como un Proyecto nacional del Ministry of Education, Culture, Sports, Science and Technology de Japón.

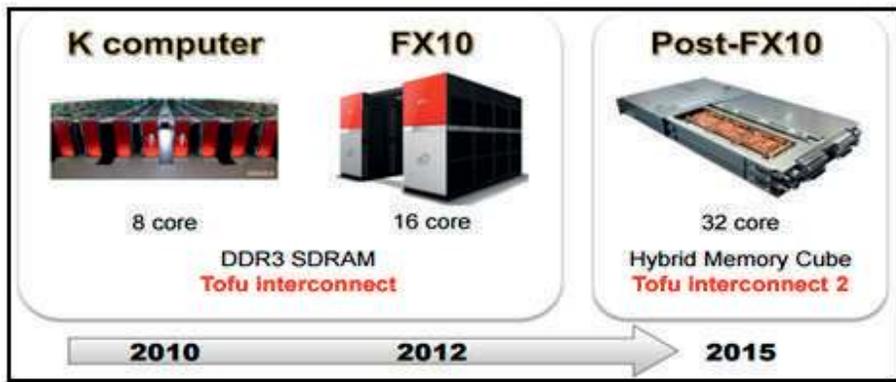


Imagen 63: Línea de tiempo de Tofu Interconnect de Fujitsu. Créditos: Fujitsu

La interconexión Tofu proporciona la escalabilidad para implementar el sistema de 10 computadoras petaflops que tiene más de 80 000 nodos. La topología de red es una malla 6D / toro. Las interfaces de red cuádruples proporcionan un alto rendimiento. La interfaz de barrera está dedicada a la descarga de comunicaciones colectivas. Este artículo describe la arquitectura de interconexión de Tofu, el enrutador de red Tofu, la interfaz de red Tofu y la interfaz de barrera Tofu, además de presentar los resultados preliminares de la evaluación. Destacan los sistemas TOFU INTERCONNECT Y TOFU INTERCONNECT 2.

2.1.6. Storages

Es un componente importante en los sistemas de alto rendimiento computacional, caracterizado por contar de grandes capacidades de almacenamiento (es un computador de gran capacidad de almacenamiento), su función es netamente complementaria al equipamiento encargado del procesamiento (nodos de procesamiento y otros).

El Storage funciona como un gran almacén o repositorio de datos, ficheros y otros contenidos a gestionarse desde la Supercomputadora. Los Storages en otro contexto como también son utilizados como servidor de datos.

Existen diversos fabricantes que ofrecen storages, sin embargo para efectos ilustrativos de este documento vamos a priorizar algunos Storages de sistemas contenidos de algunas marcas priorizando las principales marcas del Top500 de noviembre del 2016.

Hewlett Packard Enterprise

HPE StoreEasy 1000

El almacenamiento HPE StoreEasy 1000 se basa en HPE ProLiant Server DNA. Este equipo facilita incremento de capacidad hasta 280 TB. El HPE StoreEasy Pool Manager facilita las configuraciones. A la vez facilita entre el 50-60 por ciento de ahorro de espacio. Las capacidades de gestión avanzada de datos mejoran automáticamente la utilización de los recursos, cumplen las políticas de retención de datos y mejoran la protección de los archivos más importantes.

El almacenamiento HPE StoreEasy 1000 protege los datos en reposo con cifrado de sistema de archivos. El almacenamiento HPE StoreEasy 1000 aumenta el tiempo de actividad con características estándar que incluyen supervisión casi continua del estado con HPE Active Health System, sistema operativo (SO) en espejo, unidades de estado sólido (SSD), ventiladores redundantes (no disponibles en todos los modelos) y alimentación, y protección de memoria avanzada.



Imagen 64: HPE StoreEasy 1000. Créditos: www.hpe.com

HPE StoreEasy 3000

Según la web de HPE [15], este equipo permite la conexión con un gran número de usuarios y dispositivos, a la vez ahorro de espacio (del 50-60 por ciento con reduplicación de datos pormenorizada). Cuenta con gestión de datos avanzada para mejorar automáticamente el uso de los recursos, retención de datos y protección de archivos sensibles, así mismo almacenamiento automático en el caché de forma local (archivos de acceso frecuente del centro de datos corporativo).

StoreEasy protege los datos inactivos mediante cifrado de sistemas de archivos, y mientras se transfieren con el cifrado y la firma de SMB. El soporte nativo para Microsoft Azure le permite almacenar sin problemas las copias de seguridad de datos StoreEasy en la nube, utilizando la suscripción Azure, para protegerse contra desastres y errores graves.

Permite la autorreparación en línea de sistemas de archivos. Esta autorreparación reduce la necesidad de desconectar los sistemas, verificar en línea y reparar de forma precisa y rápida, lo que reduce el tiempo de inactividad resultante de la corrupción de archivos de días u horas a tan solo unos minutos.



Imagen 65: HPE StoreEasy 3000. Créditos: hpe.com

CRAY

• **Cray Tiered Adaptive Storage (TAS):** Esta solución proporciona un sistema abierto y optimizado de gestión de datos diseñado para reducir el costo de gestión del almacenamiento a largo plazo. Posee las siguientes ventajas:

- Bajo coste y la complejidad de la gestión de datos durante la vida útil de los datos.
- Basado en sistemas abiertos, Linux® y formatos abiertos para mantener iniciativas de conservación y acceso a datos a largo plazo.
- Capacidad pequeña y de escala sin interrupciones, facilita que los datos sean siempre disponibles y accesibles.
- Cray TAS incluye todo lo necesario para migrar, entregar y mantener archivos de ahora y en el futuro.
- Basado en estándares abiertos, TAS está diseñado para servir a grandes empresas y conjuntos de datos HPC con la necesidad de un acceso unificado a los datos. Como tal, ofrece una de las soluciones de almacenamiento escalonadas.
- Facilita la virtualización de todos los niveles de almacenamiento con una interfaz de sistema de archivos. Esta característica permite a los administradores de aplicaciones implementar y administrar aplicaciones fácilmente.



Imagen 66: Cray Tiered Adaptive Storage. Créditos: Cray.com

Cray® Sonexion® 3000: Es la mejor solución creada por CRAY para el almacenamiento para computación de alto desempeño. Este almacenador incorpora todo lo necesario para desplegar, escalar y administrar el almacenamiento en paralelo a gran escala para grandes datos y supercomputación. Cada módulo combina software y hardware en un diseño similar a un dispositivo para escalar de manera eficiente, simplificar el despliegue y la gestión y reducir el número de puntos de gestión física y la huella del centro de datos.

Cray Sonexion 3000 está diseñado específicamente para alto rendimiento de datos grandes y cargas de trabajo de supercomputación - para cualquier entorno de computación x86 Linux®. Además es un sistema de almacenamiento escalable, combina potentes servidores con el sistema de archivos paralelo Lustre® y el software de gestión en un producto de almacenamiento modular que escala, finalmente ofrece un rendimiento optimizado extremo a extremo por Cray. El diseño modular, pre-integrado y compacto de este sistema mantiene bajos los costos al tiempo que ofrece el rendimiento adecuado para análisis y computación de clusters y superordenadores de todo tipo.

Cuenta con herramientas de gestión y diagnóstico de infraestructura como propios para hardware Cray como Cray Systems Snapshot Analyzer (SSA) y Cray Sonexion Storage Manager (CSSM).



Imagen 67: Cray Sonexion 3000. Créditos: Cray.com

SGI (Silicon Graphics International)

- **SGI® InfiniteStorage™ 17500:** Este es el último y mejor almacenador del fabricante SGI. Respecto a equipos SGI ofrece un rendimiento sin precedentes de IOPS (operaciones de entrada/salida por segundo) y ancho de banda con una gestión de capacidad altamente eficiente para proporcionar la base para la próxima generación de conjuntos de aplicaciones de datos intensivos.

Este equipo fue diseñado para trabajar con conjuntos multi-core de múltiples aplicaciones actuales. El SGI IS17500 con Storage Fusion Architecture utiliza avanzada tecnología de procesadores, buses y memoria con un motor RAID optimizado y sofisticados algoritmos de gestión de datos. IS17500 está diseñado para simplificar y dominar el gran crecimiento de datos, lo que le permite diseñar y escalar su entorno de manera más inteligente, eficiente y rentable que nunca.

Reduce el consumo de energía en un 50%, eliminando gabinetes de almacenamiento innecesarios, interruptores, ventiladores y fuentes de alimentación. Está orientado a sacar ventajas respecto a espacio, potencia y refrigeración

SGI® InfiniteStorage™ 7700x: Es un dispositivo de almacenamiento flash híbrido, diseñado especialmente para los requisitos de Big Data. En el núcleo, el IS7700x está ejecutando el sistema operativo SFA OS. La unidad base es una carcasa de rack de 4U con capacidad para 60 unidades. El IS7700x es un sistema totalmente redundante y modular que puede expandir hasta 396 unidades de disco en un gabinete de 20U. Las opciones de conectividad de host de almacenamiento son cuatro Fibre Infiniti de 56 GB FDR o cuatro Fibre Channel de 16 GB.

SGI® InfiniteStorage™ 5600i: Es un storage ajustable a requerimientos de rendimiento y capacidad de almacenamiento. Ideal para gestión de datos de industrias (la fabricación, el petróleo y el gas, las ciencias de la vida y las ciencias de la tierra).

El software SGI DMF, el IS5600i es una solución de almacenamiento de nivel primario o secundario ideal para implementaciones de virtualización en niveles. Además de muchas opciones de protección de datos, por ejemplo posee sistemas de recuperación de las unidades que fallan, además pueden ser más eficientes que el RAID tradicional. (La serie IS5600i viene con tres opciones de recinto, 2U 12, 2U 24 y 4U 60).



Imagen 68: SGI InfiniteStorage 5600i. Créditos: Uv hpc

International Business Machines (IBM)

La familia DS8880.

DS8880 es un sistema de almacenamiento diseñado para proporcionar infraestructura para procesos de alto rendimiento y que requieran alta capacidad y almacenamiento. Está basada en la plataforma de software empresarial IBM DS8000 y ofrece aceleración de misión crítica, máxima disponibilidad, integración incomparable con servidores IBM y eficiencia transformacional mediante diversas funcionalidades.

Entre las principales características podemos mencionar:

- Alta capacidad de memoria cache (almacenamiento inteligente en caché aceleran el rendimiento).
- Almacenamiento avanzado en capas ayuda a optimizar de forma automática el rendimiento y el costo de los sistemas de una y varias capas.
- Entre las opciones flexibles de unidades de disco se incluyen los discos magnéticos tradicionales de alto rendimiento, discos magnéticos de alta capacidad, unidades flash y tarjetas flash ultrarrápidas (con HPFE).

- Una interfaz gráfica de usuario (GUI) intuitiva simplifica y centraliza la gestión y permite una consolidación más eficaz.
- Las unidades de cifrado automático y el gestor de claves basado en estándares ofrecen protección avanzada para los datos pasivos, prácticamente sin impacto en el rendimiento.
- Almacenamiento no volátil y en caché escalable hasta 2 terabytes (TB).
- Escalabilidad desde 3 TB hasta más de 4,6 petabytes (PB) de capacidad de almacenamiento físico.

IBM DeepFlash 150

Esta storage está basado en un sistema de discos de estado sólido que contiene varias unidades de memoria flash en lugar de unidades de disco duro. La memoria flash, que no tiene partes móviles, es un tipo de memoria no volátil que se puede borrar y reprogramar en unidades de memoria llamadas bloques. Es una variación de la memoria borrable programable de lectura (EEPROM) que tiene su nombre porque los bloques de memoria se pueden borrar en una sola acción o "flash". Una matriz de flash puede transferir datos desde y hacia unidades de estado sólido (SSD) mucho más rápido que las unidades de disco electromecánicas.

Según la propia IBM, IBM Flash Storage está diseñado para satisfacer los requisitos modernos de almacenamiento de alto rendimiento: latencia ultra baja, rentabilidad, eficiencia operacional y confiabilidad en aspectos críticos para la misión. Además es una opción ideal para acelerar los sistemas de interacción, datos no estructurados, datos grandes y otras cargas de trabajo que requieren latencia muy baja, alto rendimiento y datos sostenidos.

Este innovador sistema de almacenamiento se basa en un enfoque a nivel de sistemas que permite a las organizaciones gestionar conjuntos de datos mucho más grandes sin tener que gestionar unidades de disco o SSD individuales.

DeepFlash 150 viene completo con todo el hardware necesario para la empresa y el almacenaje de la hiper-escala, incluyendo hasta 64 tarjetas de destello diseñadas propósito en un chasis 3U y los conectadores del SAS 12-Gbps para hasta ocho servidores del anfitrión.

Algunas otras ventajas:

- Proporciona un bloque de construcción para las infraestructuras de almacenamiento definido por software, diseñadas para entornos de big data, streaming de medios de comunicación y ocio, escritorio virtual, base de datos de alta velocidad e hiperescalados.
- Rompe la barrera de 1 dólar por GB; incluso con el software de pila de almacenamiento el coste es inferior a 2 dólares por GB
- Proporciona gestión y escalado optimizados para flash, junto con interfaces de almacenamiento de objetos, archivos y bloques de gran capacidad diseñadas para facilitar la administración y ahorrar tiempo.

2.2. Sistema de refrigeración

El siguiente componente fundamental y complementario para un computador de alto rendimiento es su sistema de refrigeración. Las temperaturas que alcanzan estas computadoras son exageradamente altas y es muy probable que el calor afecte el funcionamiento e incluso deteriore y malogre el equipo. Por ello es crítico mantener adecuadamente refrigerado el lugar donde se aloja el equipo. Si bien es cierto en todos los manuales de datacenters se menciona la importancia de mantenerlos refrigerados, sin embargo se debe ser doblemente cuidadosos albergará un sistema de alto rendimiento computacional (es recomendable dimensionar de forma exclusiva las necesidades de refrigeración).

Vale la pena resaltar que en muy pocos casos sería recomendable utilizar aires acondicionados de uso común o también llamados “de confort”. Actualmente existen 2 sistemas recomendables, en el primer caso los sistemas de aire acondicionados de precisión, en el segundo los sistemas de enfriamiento adiabático. Seguidamente vamos a explicar de qué trata cada uno, incluso incluiremos ejemplos:

2.2.1. Sistema de enfriamiento por aires acondicionados de precisión: Se puede definir como aquellos equipos diseñados para lograr un ambiente, donde, en forma simultánea y continua, se controlen la temperatura, la humedad, la circulación y la limpieza del aire, a la vez que se mantiene una presión positiva en la sala, en relación con otros ambientes, para una exigencia de trabajo de 24 horas al día durante los 365 días del año, por un tiempo de vida útil entre 15 y 20 años. Cuanto mayor sea la cantidad de procesadores o nodos de procesamiento mayor será la necesidad de enfriamiento, y ello a decir verdad incrementará el costo del sistema de refrigeración. En este caso los altos costos del sistema de refrigeración dificultan el proyecto.

Seguidamente presentaremos 3 casos de sistema de refrigeración por aire acondicionado de precisión:

- **Aire acondicionado de precisión Liebert DS, de la marca Vertiv (USA):** El sistema de enfriamiento Liebert DS brinda un control confiable, preciso y eficiente de la temperatura, de la humedad y del flujo de aire de la sala para un funcionamiento apropiado del equipo electrónico crítico. El sistema también ofrece una alta eficiencia energética, un control Liebert iCOM amigable con el usuario, marco modular, acceso frontal para mantenimiento y opciones de compresor y ventilador. Funciona con potencia entre 28 y 103 Kw.

Fuente: <https://www.vertivco.com>

Características:

1	Cuenta con componentes redundantes como compresores, motores y ventiladores.
2	Se puede instalar múltiples configuraciones y tiene varios tamaños para permitir una aplicación flexible.
3	Monitorea y controla el funcionamiento del Liebert iCOM.
4	Permite una instalación en una configuración ascendente o descendente y modelos con economizador aire, agua, glicol, GLYCOOL.
5	Ofrece un economizador GLYCOOL energéticamente eficiente y modelos híbridos Dual Cool. Los modelos descendentes son compatibles con un economizador opcional de aire.

• **CyberAir 3 marca STULZ (Alemania):** Sistema independiente de aire acondicionado de precisión para aplicaciones de altas prestaciones. El sistema de circuito cerrado de aire CyberAir 3 de STULZ controla las condiciones del centro de datos con máxima precisión y fiabilidad, así como un uso eficiente de la energía eléctrica.

Diseñado para el funcionamiento fiable continuó durante largos años, el CyberAir 3 de STLZ es extremadamente preciso, silencioso y excepcionalmente económico. Mantiene su equipo IT disponible en todo momento. Funciona con potencia que va desde 20 a 82 Kw.

Características:

- Ventilador EC con funcionamiento silencioso, duradero y fácil de mantenimiento.
- Compresor con tecnología EC y válvulas de expansión con control electrónico.
- Control infinito del compresor para una potencia frigorífica de máxima eficiencia.
- Gestión de control del filtro.
- Siete sistemas de refrigeración con flujo ascendente y descendente.
- Tamaño compacto.
- Controlador C7000 para el control y monitorización.
- Mantenimiento accesible solo por la parte frontal.

Fuente: <https://www.vertivco.com/>



Imagen 69: Aire acondicionado de precisión Liebert DS. Créditos: www.vertivco.com

• **HDCV marca Uniflair (Italia):** Diseñado específicamente para el centro de datos modernos con un rendimiento optimizado a temperaturas elevadas en línea con las últimas guías ASHRAE, y combinado con la tecnología de ventiladores compuesto de vanguardia, la gama HDCV ofrece hasta 50% de ahorro de energía del ventilador en comparación con las unidades de flujo descendente convencionales. 30-250Kw.

Características:

- Control del microprocesador
- Advanced microprocessor control UG50

- Integración con Uniflair Chillers para una lógica de gestión optimizada
- Libre contacto para alarma general y 2 para alarmas direccionables
- Interruptor ON-OFF remoto
- Tarjeta serie RS485 integrada para conexión directa a BMS externo (Modbus). Segunda ranura para tarjeta serie adicional para conexión BMS (opcional).
- Tarjeta de reloj integrada en la unidad de medición de la temperatura de entrada en agua fría

2.2.2. Sistema de enfriamiento adiabático: El enfriamiento adiabático consiste en enfriar la temperatura ambiente del aire utilizando agua. Es un enfriador o condensador de diseño avanzado que utiliza la última tecnología de transferencia de calor y controles. Utilizando agua solamente en los días pico, el enfriador/condensador ACS frecuentemente puede reducir el consumo de agua al 20% de un sistema convencional enfriado por agua y reducir la demanda de energía en días pico en comparación con un sistema seco. Este sistema utiliza un sistema de gestión del agua y paneles de enfriamiento adiabático para maximizar la eficiencia, administrar de forma segura el uso del agua y reducir al mínimo su consumo. Además de su facilidad de operación, la eliminación de los sistemas de tratamiento de agua es típicamente aplicable. Presentamos 2 ejemplos:

• **Acondicionadores de aire FREE COOLING con enfriamiento adiabático marca techline (Italia):** Utilizan el sistema FREE COOLING indirecto con enfriamiento adiabático para garantizar un alto ahorro energético del sistema. Las unidades de la serie F son la solución ideal para el enfriamiento del aire de instalaciones donde el eco sostenibilidad y el ahorro energético son prioritarios, como los grandes centros de datos de última generación, permitiendo conseguir prestaciones similares a las de los sistemas de FREE COOLING directo. Las unidades están diseñadas para ser instaladas en exteriores, principalmente sobre techos, y constan de dos secciones de tratamiento; una del aire interior y otra del aire exterior, físicamente separadas y con dos secciones de filtrado y ventilación totalmente independientes.

Características

- Intercambiador de calor de placas certificadas EUROVENT
- Panel evaporador OXYVAP
- Estructura metálica de color blanco RAL 9010
- Paneles con aislamiento termoacústico de 50 mm
- Filtros de aire de eficiencia G4 con indicación de filtro sucio
- Ventiladores electrónicos EC FANS
- Cuadro eléctrico con dispositivos de regulación de seguridad
- Microprocesador de control SURVEYEVO con pantalla gráfica
- Sistema de apagado de la unidad por presencia de incendio
- Placa RS485 Modbus RTU slave
- Placa RJ45 Ethernet.

Fuente: <http://www.tecnairlv.it>



Imagen 70: serie F - acondicionadores de aire free cooling con enfriamiento adiabático
Créditos: www.tecnairlv.it

• **SMART COOLING™ de la marca Blue Energy Technologies (USA):** Denominados paneles de pre-enfriamiento adiabático. Este sistema utiliza las propiedades de enfriamiento del agua al evaporar agua irregularmente en frente de los paneles que están adheridos al frente del condensador, mientras que previenen que el agua alcance la bobina. Como resultado, la temperatura del aire disminuye antes de alcanzar al condensador y se vuelve a enfriar mientras pasa a través de la membrana saturada del panel.

Este panel se coloca delante del condensador de la unidad de aire acondicionado y refrigeración. La temperatura del aire que pasa a través del panel se reduce debido que el aire caliente se enfría mediante agua nebulizada generada por las instalaciones de aire acondicionado y refrigeración (una tecnología especial de pulverización de líquidos, la Smart Release technology es utilizada, que permite pulverizar agua en minúsculas gotas).

El panel de preenfriamiento adiabático "Smart Cooling™" consiste de los siguientes componentes:

1. Estación de bombeo.
2. Paneles con membranas reflectantes.
3. Sistema de pulverización de agua.
4. Sistema de tratamiento de agua.
5. Panel de control, proporciona un arranque automático del sistema a cierta temperatura.

El sistema aspersor de agua funciona bajo alta presión de agua la cual es suministrada a través de la estación de agua de alta presión especialmente diseñada combinada con un tratamiento de anti-sarro de agua para prevenir cualquier acumulación de sarro. Un procesador dedicado asegura un mínimo consumo de agua. Consume menos aceite y genera menor temperatura del viento del motor, y disminuye la obstrucción por polvo del condensador. Estos factores contribuyen grandemente a la confiabilidad del compresor y de todo el sistema.

Fuente: <http://www.blueenergytechnologies.com/>

2.3. Sistema eléctrico

Una de las cosas más difíciles en proyectos de alto rendimiento computacional es el dimensionamiento de los requerimientos de cargas eléctricas necesarias para su funcionamiento. Es muy importante hacer una estimación lo más precisa posible y ello debe ser incluido en el diseño del sistema eléctrico. Este tema es sumamente crítico, incluso son frecuentes los problemas o fracasos de este tipo de proyectos por no tomar las medidas respecto a este tema. El sistema eléctrico deberá proporcionar la energía suficiente y estable para el funcionamiento apropiado de la Supercomputadora.

Los proyectos de Supercomputación deben incluir el presupuesto necesario para adquirir un sistema eléctrico propio o reforzar el sistema eléctrico ya existente (de acuerdo a la capacidad instalada de la organización). Se recomienda no embarcarse en proyectos de Supercomputación, si no se dispone de los suficientes fondos o si no se cuenta con un presupuesto exclusivo para instalar el sistema eléctrico, de lo contrario todo lo que se invirtió en la Supercomputadora será infructuoso, no podrá funcionar o incluso correrá peligro de malograrse.

En el caso del Centro de Alto Rendimiento de la Amazonía Peruana del IIAP, se instaló un sistema eléctrico exclusivo para el MANATI IIAP, que tuvo un costo aproximado del 40% del costo de la Supercomputadora, y no es que haya sido la solución completa, sin embargo se implementó un sistema que facilite las condiciones básicas de funcionamiento. Por otro lado, una solución eléctrica ideal no hubiera bajado del 90% del costo de la Supercomputadora.

Hay que tener en cuenta que el MANATI IIAP es un clúster de apenas 11 equipos de alta velocidad, digamos que en número de nodos es pequeño respecto a las Supercomputadoras extranjeras que tienen más de 100 nodos, e incluso la Supercomputadora de la Universidad Nacional San Luis Gonzaga de Ica (UNICA) que cuenta con 70 equipos. Definitivamente el sistema eléctrico del MANATI IIAP no alcanzaría para dar soporte a este tipo de equipamiento. Cada nodo tiene su fuente de poder o alimentación eléctrica, que es desde donde se consume energía. Es importante diseñar el sistema eléctrico conociendo la cantidad de fuentes de poder y el consumo de Watts por cada núcleo. Se debe tener en cuenta que también el sistema de refrigeración deberá alimentarse de este mismo sistema eléctrico (según el diseño o sistema de la organización).

Finalmente se recomienda contratar a un experto en sistemas eléctricos (ingeniero eléctrico con mediana experiencia en instalación de sistemas eléctricos para datacenter o equipamientos con grandes requerimientos de energía) para liderar el proceso relacionado al sistema eléctrico. Este profesional incluso deberá ver aspectos de la alimentación eléctrica externa, sistemas de proyección complementarios e incluso la proyección de los gastos de electricidad durante el funcionamiento de la Supercomputadora. Se han visto casos en el Perú en que no se consultó a expertos eléctricos lo que generó graves problemas que luego fueron complicados de corregir.

Seguidamente presentamos los componentes básicos de un sistema eléctrico, basado en la implementación del Centro de Alto Rendimiento de la Amazonía Peruana

- Transformador de distribución.
- Tablero de transferencia automática (TTA).
- Tablero general eléctrico.
- Transformador de aislamiento.
- Estabilizador de tensión.
- UPS (Uninterruptible Power Supply).

- Tablero de Bypass.
- Tablero electrónico de distribución estabilizada.
- Pararrayos.
- Sistema de puesta a tierra.
- Grupo Electrónico.

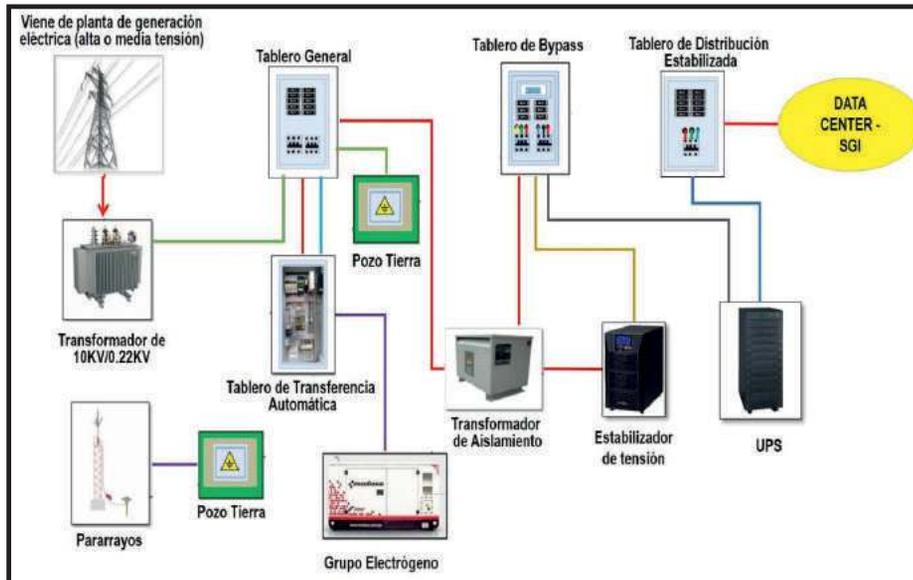


Imagen 71: Diseño del sistema eléctrico. Créditos: Elaboración propia

2.3.1. Transformador de Distribución: [16] Es una máquina eléctrica que permite aumentar o disminuir el voltaje o tensión en un circuito eléctrico de corriente alterna, manteniendo la frecuencia. Son utilizados para reducir la tensión trifásica en redes de distribución eléctrica a tensión de consumo domiciliario o empresarial, principalmente en áreas metropolitanas y para aplicaciones industriales.

En el caso del sistema eléctrico del MANATI IIAP, se utilizó uno de los transformadores que la empresa de electricidad local de Iquitos (Electro Oriente) pone a disposición del IIAP para su acceso al servicio eléctrico. Al inicio del diseño del sistema eléctrico se pensó en la adquisición de un propio transformador de distribución, sin embargo, luego se hizo un estudio del consumo eléctrico del IIAP, donde se determinó que se podía utilizar uno de los 2 que ya inicialmente se estaban utilizando para las necesidades institucionales. El adquirir uno nuevo hubiera demandado altos costos y varios meses de trámites, fabricación e instalación.

2.3.2. Tablero de transferencia automática (TTA): [17] Un tablero de transferencia es un interruptor eléctrico que cambia una carga entre dos fuentes, son automáticas y pueden cambiar cuando detectan que una de las fuentes ha perdido o ganado el poder. Un interruptor de transfe-

rencia automática (ATS) se instala a menudo donde se encuentra un generador de respaldo, para que el generador pueda proporcionar energía eléctrica temporal si la fuente de energía falla.

Por ejemplo, en una casa equipada con un generador de respaldo y un ATS y se produce un corte de energía eléctrica, el ATS le dirá al generador de respaldo cuando comenciar, una vez que el ATS ve que el generador está listo para proveer de energía eléctrica, el ATS rompe la conexión del hogar de la compañía eléctrica y se conecta el generador al panel eléctrico principal de la casa y el generador suministra energía a la carga eléctrica de la casa, pero no está conectado a las líneas de servicios eléctricos.



Imagen 72: Tablero de transferencia automática. Créditos: Elaboración propia

2.3.3. Tablero General Eléctrico: Es el tablero principal en toda instalación, en ellos van montados los dispositivos de protección y maniobra que protegen los alimentadores además permiten operar sobre toda la instalación de consumo en forma conjunta o fraccionada.

Características: [18]

- En Baja Tensión (BT), se pueden clasificar de diferentes maneras, tales como:
- Por su diseño interior: abierto, frente muerto, cerrado.
- Por su lugar de instalación: uso interior y exterior.
- Por el modo de montaje: mural y auto-soportado.
- Por su ejecución: Fijo y extraíble.
- Por la posibilidad de desplazamiento: estático y móvil.
- Por la magnitud de los interruptores: bajo, mediano y bajo.

Especificaciones técnicas:

Parámetros Eléctricos	Bajas Capacidades
Capacidad de Corriente (A)	63 – 100 – 160 – 200 – 250
Tensión	220 – 380 – 440
Capacidad de Ruptura (kA)	85 – 50 – 42

2.3.4. Transformador de Aislamiento: [19] Se encarga de aislar la entrada de energía de la salida, independizando completamente el suministro eléctrico externo del interno a través de una malla electrostática. El traspaso de energía es vía inducción, esto significa que la mayoría de perturbaciones eléctricas externas no son pasadas al circuito eléctrico interno protegiendo todos los equipos conectados de interferencias, pequeñas distorsiones en la frecuencia, armónicas, entre otros.

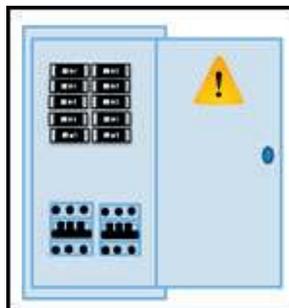


Imagen 73: Tablero General Eléctrico Créditos: Elaboración propia

Ejemplo de equipos de transformador de aislamiento:

Marca Prolite: HP9116T SERIES (0.5KVA - 75KVA) de la marca DLMRED (DLMRED UPS SYSTEM SAC)



Imagen 74: Transformador de aislamiento marca Prolite: HP9116T. Créditos www.dlmred-ups.com.pe

Especificaciones Técnicas:

Capacidad (VA/Kva)	Desde 500VA a 75Kva
Voltaje de Entrada (Vac)	110vac, 120vac, 220vac, 230vac, 240vac
Voltaje de Salida (Vac)	110vac, 120vac, 220vac, 230vac, 240vac
Fase	Monofásico (1Ø). + N +G
Estándar de Fabricación	IEC-76 / ITINTEC 370.002 / IEC600-76-11
Grado de Protección	IP21
Factor	K-20 ; K-13 ; K-1
Factor de Potencia	Cos Phi = 0.8

2.3.5. Estabilizador de Tension: [86] es un equipo electrónico o electromecánico, destinado a dar una tensión estabilizada en su salida (230 V ó 380 V), aunque en su entrada la tensión eléctrica sea más baja o más alta del valor de utilización (nominalmente 230 V en monofásica; 400 V en trifásica)

Acontinuación se detallan los tipos habituales de estabilizadores de tensión [87]:

- **Estabilizadores electrónicos o AVR** (Automatic Voltage Regulator): Se basan en el empleo de un autotransformador con varios puntos de trabajo entre los que se conmuta según sea la tensión de entrada. Su principio es el mismo que el del estabilizador que monta un SAI in-line. Habitualmente tienen tres modos de trabajo: normal (la tensión de salida es la misma que la de la entrada), boost (cuando la tensión de entrada es baja, se aumenta la tensión de salida) y buck (cuando la tensión de entrada es alta, se reduce la tensión de salida).

- **Estabilizadores electromecánicos:** Se basan en el empleo de un autotransformador de variación continua sobre el que se posiciona una escobilla accionada por un servomotor. Su precisión es mucho mayor que la de un estabilizador electrónico, siendo su valor típico inferior al 4%. Su velocidad de respuesta es más lenta y depende de la velocidad del servomotor, que responderá más rápido a variaciones pequeñas y con mayor lentitud a variaciones grandes en la tensión de entrada. Al igual que los estabilizadores electrónicos, filtran ruidos y recortan picos transitorios.

- **Estabilizadores de estado sólido:** A diferencia de los estabilizadores electrónicos, emplean TRIAC (Triodo para Corriente Alterna) para realizar las conmutaciones y suelen emplear un número mayor de puntos de trabajo. Al no emplear partes móviles son más fiables, aunque menos robustos cuando se excede su potencia máxima.

- **Estabilizadores ferromesónicos:** Emplean un transformador de saturación como medio para regular el voltaje. Gracias a la propiedad de saturación magnética, estos estabilizadores se basan en un circuito resonante que produce un voltaje medio de salida muy constante. Este tipo de estabilizadores son muy interesantes gracias a que no disponen de componentes activos ni partes móviles. Son los más robustos y precisos. Proporcionan además un aislamiento galvánico entre la entrada y la salida. Además de un mayor coste, sus principales inconvenientes radican en su tamaño, emisión de un zumbido audible y gran disipación de calor. Su eficiencia oscila en torno al 90% cuando se emplean con cargas altas, pero puede llegar a ser inferior al 60% con cargas bajas. Por las citadas características, son más propios de entornos industriales que de hogares u oficinas.



Imagen 75: Estabilizador de Tension Sólido Créditos: www.gamatec.com.com

2.3.6. UPS (Uninterruptible Power Supply): [20] Un UPS (Sistema de alimentación ininterrumpida). Es una fuente de suministro eléctrico que posee una batería con el fin de seguir dando energía a un dispositivo en el caso de interrupción eléctrica. Algunos UPS más avanzados también ofrecen aplicaciones que se encargan de realizar ciertos procedimientos automáticamente para los casos en que el usuario no esté presente y se corte el suministro eléctrico.

Tipos o tecnologías de UPS

- **UPS en línea:** Provee alimentación constante desde su batería y no de forma directa. Esta tecnología es la más cara de todas pero es la que ofrece el mayor nivel de protección. Por lo tanto es ideal para equipos sensibles o importantes como servidores centrales. El UPS on-line tiene un modo llamado by-pass, que es un paso paralelo que deriva la corriente eléctrica directamente desde la entrada hacia la salida del UPS. Este tipo de UPS es el que se tiene en el sistema eléctrico del MANATI IIAP.

- **UPS, Sistema de Energía Reservada:** Se encarga de monitorear la entrada de energía, cambiando a la batería apenas detecta problemas en el suministro eléctrico. Ese pequeño cambio de origen de la energía puede tomar algunos milisegundos (tiempo de conmutación), lo cual puede afectar a algunos aparatos sensibles.

- **UPS línea interactiva:** Los UPS de línea interactiva funcionan de forma similar al UPS offline-standby, pero con un autotransformador multivoltaje multitap. Este es un tipo especial de transformador que puede sumar o restar bobinas de cable, aumentando o disminuyendo el campo magnético y el voltaje de salida del transformador de ese modo.

Ejemplo de UPS en línea:

SU60K Sistema UPS modular trifásico SmartOnline de 120/208V, UPS en línea de doble conversión, wye, 60kVA SU60K de Tripp Lite de 60kVA. Modular, Trifásico, inteligente, proporciona disponibilidad del sistema al 100% con arquitectura modular N+1 y capacidad en paralelo 1+1. En una configuración N+1, el SU60K cuenta con tres módulos de potencia redundantes de 20kVA autónomos Hot-Swap que pueden intercambiarse aun operando con la carga energizada si se requiere mantenimiento.



Imagen 76: UPS SU60K. Créditos: www.tripplite.com

2.3.7. Tablero de Bypass: [21] Permite la selección del suministro eléctrico normal o regulado dado por el UPS. Funciona como un bypass (Bypass: en el lenguaje técnico indica una desviación), colocada en paralelo en un circuito eléctrico, que permite poner en comunicación directa dos puntos

Cummins Power Generation recomienda que el sistema de control del interruptor de bypass-aislamiento permanezca activo, aun cuando el interruptor de bypass esté integrado, de manera que en una falla de energía, el grupo electrógeno esté en funcionamiento y listo para aceptar la carga cuando se cierre manualmente el interruptor de bypass hacia la carga [22].



Imagen 77: Tablero de Bypass. Créditos: Elaboración propia

2.3.8. Tablero eléctrico de distribución estabilizada: [23] Los tableros de distribución tienen como función principal distribuir la energía eléctrica desde la acometida o medidores de la compañía eléctrica, hasta los equipos que transforman dicha energía en trabajo útil. Esta distribución de energía eléctrica tiene que hacerse de manera eficiente y segura.

Especificaciones Técnicas:

Parámetros Eléctricos	Bajas Capacidades
Capacidad de Corriente (A)	63 – 100 – 160 – 200 – 250
Tensión	220 – 380 – 440
Capacidad de Ruptura (KA)	85 – 36 – 35
Usos	Interior o Exterior
Normas Estándares	IEC
Tipo	Metal enclosed (No compartimiento)



Imagen 78: Tablero eléctrico de distribución estabilizada Créditos: Elaboración propia

2.3.9. Pararrayos: Es un instrumento cuyo objetivo es ionizar el aire circundante para atraer un rayo. Siendo Iquitos una ciudad propensa a las tormentas eléctricas al incidir el rayo sobre el pararrayos, este dirige la energía hacia el sistema de puesta a tierra (SPT) por medio de las bajadas.

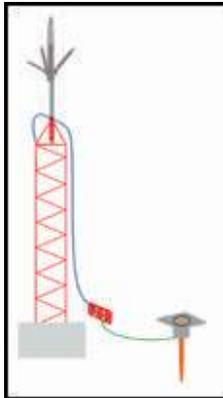


Imagen 79: Pararrayos. Créditos: Elaboración propia

2.3.10. Sistema de puesta a tierra: El sistema de puesta a tierra cumple normalmente dos funciones de seguridad. La primera es establecer conexiones equipotenciales. Si una persona está en contacto simultáneamente con dos piezas diferentes de una estructura metálica expuesta, el conductor de conexión eléctrica debe garantizar que la persona no reciba un choque eléctrico, haciendo que la diferencia de potencial entre los equipos sea insuficiente para que esto ocurra.

Otra función de un sistema de puesta a tierra es garantizar que, en el evento de una falla a tierra, toda corriente de falla que se origine, pueda retornar a la fuente de una forma controlada. Por una forma controlada se entiende que la trayectoria de retorno está predeterminada, de tal modo que no ocurra daño al equipo o lesión a las personas [25].

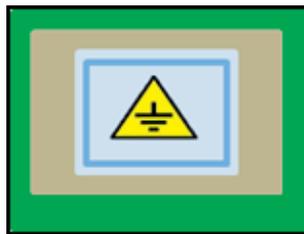


Imagen 80: Sistema de puesta a tierra. Créditos: Elaboración propia

2.3.11. Grupo Electrónico: [26] Es una máquina que mueve un generador de electricidad a través de un motor de combustión interna. Son comúnmente utilizados cuando hay déficit en la generación de energía eléctrica de algún lugar, o cuando son frecuentes los cortes en el suministro eléctrico.

Ejemplo de grupo electrógeno

GRUPOS ELECTRÓGENOS FIAT POWER TRAIN – FPT



Créditos: modasa.com.pe

Grupo Electrónico MF-55

Modelo	Potencia		Voltaje	Frecuencia	Factor de potencia	Amperaje
	Prime	Stand by				
Mf-55	50 Kw/63 KVA	53 Kw/67 KVA	220V	60Hz	0.8	174 A

Datos técnicos

Modelo	MF-55	
Motor	FIAT N45 SM1A	
Alternador	STAMFORD UCI 224D	
Módulo de control	Electrónico	
Fases	Trifásico	
Tanque combustión	abierto/insonoro 55 Galones / 68 Galones	
Sistema Eléctrico	12V	
Frecuencia	60Hz	50Hz
Radiador flujo aire	138 m3/min	112 m3/min
Gases de escape flujo	5.50 m3/min	4.50 m3/min
Temperatura gases escape	385°C	483°C

2.4. Adecuaciones de ambiente

Finalmente es de carácter obligatorio realizar diversas adecuaciones al ambiente o lugar donde será instalado la supercomputadora. Las adecuaciones deben garantizar la apropiada climatización, monitoreo y seguridad del equipamiento. Ello estará en función de las dimensiones y necesidades de la supercomputadora. No olvidar que el sistema de alto rendimiento computacional incluye el supercomputador, el sistema de refrigeración y todos los sistemas que garantizan su apropiado funcionamiento. En el caso del MANATI IAP, se hicieron dos acondicionamientos importantes:

- El primero de ellos fue la construcción de una propia casa de fuerza para albergar el sistema eléctrico, con las mejores condiciones de seguridad y temperatura posible.
- El segundo de ellos fue la adecuación de un espacio exclusivo para la Supercomputadora en el Data Center del IIAP (instalación de una mampara de vidrio templado) sistema de control de ingreso, al que le agregaron acceso de energía procedente de la casa de fuerza, sistema de aire acondicionado y otras adecuaciones de conectividad.

Como parte complementaria se agregaron los siguientes componentes:

- Sistema anti-incendio o detector de humo (con sensores y alarmas conectadas a las casetas de la vigilancia).
- Sistema de seguridad de acceso físico al MANATI IIAP (sistema de huellas dactilares para los usuarios que acceder directamente al equipo).
- Sistema de refrigeración de la casa de fuerza (aire acondicionado para evitar sobrecalentamiento de los componentes del sistema eléctrico).
- Sistema de cámaras de seguridad (para controlar el acceso al DataCenter y la casa de fuerza).
- Protocolo de acceso y mantenimiento del sistema de alto rendimiento computacional (política institucional del IIAP).
- Sistema de seguimiento a la temperatura dentro de espacio del supercomputador.
- Conectividad con usuarios, existen conexión directa a los usuarios institucionales y a usuarios externos a través de internet.

CAPÍTULO 3

ELEMENTOS DE UN SISTEMA DE ALTO RENDIMIENTO COMPUTACIONAL: SOFTWARE

El software es un elemento fundamental y complejo para computación de alto rendimiento. Al contrario de lo que se piensa en la informática convencional, en este contexto cobra importancia que el software tenga capacidad o esté preparado para aprovechar al máximo la cantidad de núcleos y otros recursos computacionales. Hablar de software para estos fines, sus usos y formas de uso merecería una publicación exclusiva de cientos de páginas.

Como éste es un documento introductorio hemos considerado un listado básico y breve descripción de software que el lector podría utilizar para empezar a usar Supercomputadoras.

3.1. Sistemas operativos:

Se han priorizado los sistemas operativos del TOP500 de noviembre del 2016 [28].

Tabla 18: Lista de Sistemas Operativos para computadoras de alto rendimiento.

N°	Sistema Operativo	Link de descarga
1	CentOS	https://www.centos.org/download/
2	Cray Linux Environment	Sin datos
3	Red Hat Enterprise Linux	https://access.redhat.com/downloads
4	SUSE Linux Enterprise	https://www.suse.com/es-es/download-linux/
5	TOSS	Sin datos

Fuente: Elaboración propia

3.1.1. CentOS (CentOS Development Team – 2004 – Varios países)

Community Enterprise Operating System (CentOs) [27], es una plataforma estable, predecible, manejable y reproducible derivado de las fuentes de Red Hat Enterprise Linux (RHEL).



CentOS es de código abierto. Desarrollado y diseñado para ser fundamentalmente similar a la de Red Hat, CentOS ejecuta únicamente las versiones más estables de software empaquetado. Esto reduce considerablemente el riesgo de accidentes y errores.

Los usuarios que instalan CentOS también acceder a un máximo de una década de actualizaciones de seguridad de nivel empresarial - debido a su estrecha relación con Red Hat. Lo que esto significa es que CentOS está equipado con una amplia gama de características de seguridad impresionantes, incluyendo un increíblemente potente firewall y el mecanismo de política de SELinux. Factor en el aumento de la estabilidad disfrutado por la distribución, y se convierte en una opción aún más atractiva [30].

Link de descarga:

<https://www.centos.org/download/>



Imagen 81: Interfaz principal Sistema Operativo Centos

3.1.2. Cray Linux Environment (Cray Inc. – 1985- EE.UU) [31] Es un sistema operativo, ofrece actualmente una función denominada Core Specialization, que permite al usuario solicitar los núcleos de los nodos para su trabajo, se dividan en dos grupos, uno para ejecutar la aplicación y otro para ejecutar los servicios del sistema. Esta característica ha mostrado beneficios de rendimiento para algunas aplicaciones.

El entorno Linux Cray fue explícitamente diseñado para escalar a más de 500 000 núcleos en un entorno que son sólo para HPC [32].

Los comandos CLE permiten a los administradores realizar funciones administrativas en los nodos de servicio para controlar el procesamiento. La mayoría de los comandos CLE se ejecutan desde el nodo de arranque, haciendo que el nodo de arranque sea el punto focal para la administración de CLE [33].

3.1.3. Red Hat Enterprise Linux (Red Hat – 2003 – EE.UU):

Es una plataforma operativa de la empresa Red Hat bien adaptado a una amplia gama de aplicaciones a través de la infraestructura de TI. Red Hat Enterprise Linux incluye soporte para las principales arquitecturas de hardware, hipervisores, y los proveedores de nubes, haciendo despliegues a través de entornos virtuales y físicos diferentes, predecibles y seguras. Herramientas mejoradas y nuevas capacidades de esta versión permiten a los administradores personalizar el entorno de aplicación para controlar y gestionar los recursos informáticos y de seguridad de manera eficiente. Red Hat Enterprise Linux es adecuado para equipos de sobremesa, servidores y mainframes [34].



Características

- Libertad a través de la estabilidad.
- Un ecosistema de soluciones y soporte.
- Confianza a través de la flexibilidad.
- Seguridad.
- Almacenamiento.
- Administración.
- Interoperabilidad.

Link de descarga:

<https://access.redhat.com/downloads>

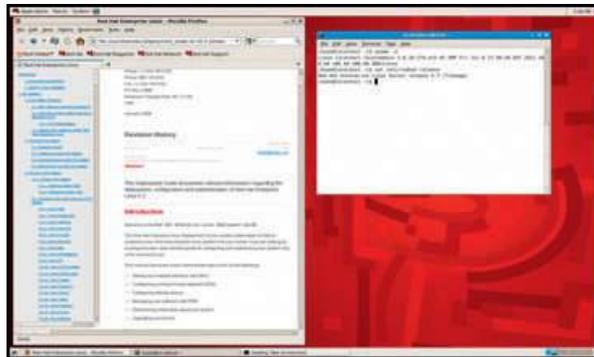


Imagen 82: Interfaz del sistema operativo Red Hat Enterprise Linux

3.1.4. SUSE Linux Enterprise (SUSE – 2000 - Alemania)

SUSE Linux Enterprise Server es un sistema operativo de servidor de prestigio mundial, seguro y de código abierto; concebido para procesar cargas de trabajo físicas, virtuales o basadas en nube que sean de misión crítica.



Entre sus funcionalidades se encuentran: Transformación de la conectividad definida por software y virtualización de funciones de red, cumplimiento de los estrictos requisitos de seguridad del gobierno federal (FISMA), adaptación a las nuevas tecnologías (ARM AArch64, Raspberry Pi y Omni-Path), Facilidad de acceso a los paquetes y las tecnologías más recientes mediante la integración de SUSE Package Hub con el Centro de servicios al cliente (SCC) [33].

Link de descarga:

<https://www.suse.com/es-es/download-linux/>



Imagen 83: Interfaz principal Sistema Operativo SUSE.

3.1.5. TOSS (Thiagarajar Facultad de Ingeniería 2009- India)

TOSS (TCE Operating System Services) [36] es una distribución Linux basada en Ubuntu y como afirma el desarrollador TOSS fue desarrollado para orientarse a la comunidad de los estudiantes desarrolladores con una variedad de software de ingeniería.

TOSS se centra en la facilidad de uso de herramientas de ingeniería y algún tipo de seguridad. TOSS también hace hincapié en la accesibilidad y la internacionalización de llegar a tantas personas como sea posible. Ya que es un sistema operativo de código completamente abierto, no hay ninguna restricción en la realización de tareas administrativas, lo que permite que la cuenta raíz permanezca abierta, y permite incluso que los nuevos desarrolladores puedan hacer cambios en el sistema.

Características:

- Facilidad de instalación y uso.
- Velocidad y estabilidad.
- Soporte y documentación.
- Rendimiento.



Imagen 84: Interfaz principal de TOSS.

3.2. Software para gestión de supercomputadoras

Tabla 19: Software para gestión de supercomputadoras.

Id	Nombre	Año de creación	Usos principales
1	Ganglia	1999	Sistema de control distribuido escalable para los sistemas de computación de alto rendimiento como los clusters y Grids
2	Nagios	1999	La monitorización de servicios de red (SMTP, POP3, HTTP, SNMP...), de los recursos de sistemas hardware (carga del procesador, uso de los discos, memoria, estado de los puertos...), independencia de sistemas operativos, posibilidad de monitorización remota mediante túneles SSL cifrados o SSH, y la posibilidad de programar plugins específicos para nuevos sistemas.
4	Portable Batch System (PBS)	1991	Optimiza la programación de tareas y la gestión de la carga de trabajo en entornos de computación de alto rendimiento (HPC).
5	PUTTY		Con PuTTY conseguimos abrir una sesión de línea de comandos en el servidor remoto para administrarlo.
6	Sun Grid Engine (SGE)	2000	La gestión de un sistema manejador de recursos computacionales o procesos distribuidos en ambientes heterogéneos.
7	Terascale Open-source Resource and QUEUE Manager TORQUE	2003	Proporciona control sobre los trabajos por lotes y nodos de computación distribuida.
8	WinSCP	2000	WinSCP permite conectarse a un servidor SSH (Secure Shell) empleando el protocolo SFTP (SSH File Transfer Protocol) o el servicio SCP (Secure Copy Protocol). SFTP es un estándar del paquete SSH-2.

3.2.1. Ganglia (Universidad de California, EEE.UU)

Ganglia [37] es un software que provee monitoreo en tiempo real y ejecución de ambientes usado por cientos de universidades tanto privadas como gubernamentales e implementaciones de clúster comerciales alrededor del mundo.



Ganglia es tan fácil de usar que al igual como puede correr en 16 nodos de un clúster puede correr en 512 o más nodos en un sistema de ambiente clúster.

Inicialmente Ganglia fue desarrollado como manera de enlazar clúster entre los campus de manera lógica. Ganglia fue desarrollado en un ambiente universitario y es completamente open-source y no tiene componentes adicionales que pertenezcan a un propietario. Todos los datos son intercambiados definitivamente por XML y XDR para un máximo de extensibilidad y portabilidad. Ganglia no enlaza nodos de un clúster con otro, la manera lógica es que enlaza un clúster con otros clúster. Ganglia enlaza líneas de clúster y computación distribuida por lo que llamamos “clúster to clúster” (C2C).

Link de descarga:

http://ganglia.info/?page_id=66

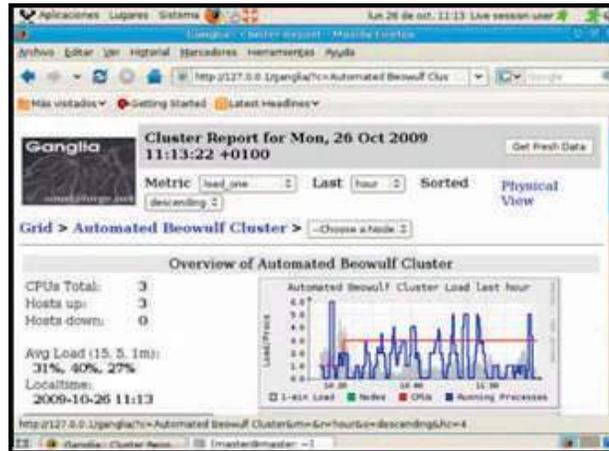


Imagen 85: Interfaz Ganglia.

3.2.2. Nagios (Ethan Galstad – 1999 – Nagios)

Nagios [38] es un sistema de monitorización de redes de código abierto ampliamente utilizado, que vigila los equipos (hardware) y servicios (software) que se especifiquen, alertando cuando el comportamiento de los mismos no sea el deseado.



Es utilizado para la monitorización de servicios de red (SMTP, POP3, HTTP, SNMP), la monitorización de los recursos de sistemas hardware (carga del procesador, uso de los discos, memoria, estado de los puertos), independencia de sistemas operativos, posibilidad de monitorización remota mediante túneles SSL cifrados o SSH, y la posibilidad de programar plugins específicos para nuevos sistemas.

Link de descarga:

<https://www.nagios.com/products/nagios-core/>



Imagen 86: Interfaz de trabajo de Nagios.

3.2.3. Open PBS (NASA Ames Research Center/Veridian Information Solution Inc – mediados del 2003, EE.UU)

OpenPBS [39] consta de tres componentes principales: un servidor de tareas, un ejecutor de trabajo, y un planificador de tareas, y un conjunto de comandos y herramientas para la presentación de trabajos y el seguimiento de las colas. El servidor de tareas (pbs_server) se encarga de los servicios básicos de puesta en cola como la creación y modificación de un trabajo por lotes y la colocación de un puesto de trabajo en ejecución cuando está previsto que se ejecute. El ejecutor de trabajo (pbs_mom) es el que ejecuta los trabajos. El planificador de tareas (pbs_sched). En la implementación más sencilla, pbs_server y pbs_sched se ejecutan sólo en el nodo front-end, mientras que pbs_mom se ejecuta en todos los nodos del clúster que pueden ejecutar los trabajos, incluyendo el nodo front-end.

3.2.4. Portable Batch System (NAS/NERSC - 1991 – EE.UU)

PBS Professional es un software que optimiza la planificación de tareas y la gestión de carga de trabajo en computación (HPC) de alto rendimiento, mejora la eficiencia del sistema y la productividad de las personas. Construido por la gente de HPC para la gente HPC.

PBS es rápido, escalable, segura y resistente, y es compatible con todas las modernas infraestructuras, middleware y aplicaciones.

Link de descarga:

<http://www.pbspro.org/Download.aspx#download>

3.2.5. Putty (Simon Tatham – 1999 -EE.UU)

Es cliente SSH y Telnet con el que podemos conectarnos a servidores remotos iniciando una sesión con ellos que nos permite ejecutar comandos [40].

Es un emulador gratuito de terminal que soporta SSH y muchos otros protocolos.



Putty ofrece una interfaz gráfica de configuración muy sencilla e integra múltiples opciones [41]:

- Guarda las preferencias de conexión para establecerla rápidamente en el futuro.
- Respuestas de puertos.
- Soporte Ipv6.
- Soporte SCP y SFTP.

Link de descarga:

<http://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>.

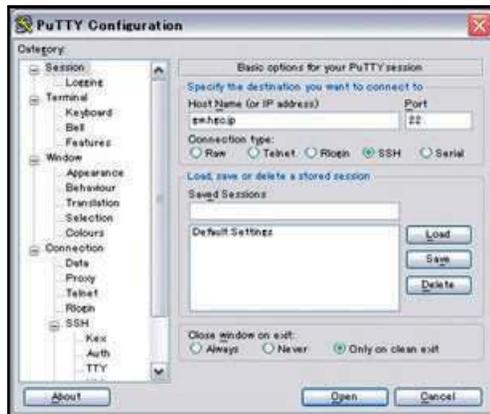


Imagen 87: Interfaz de trabajo de Putty.

3.2.6. Sun Grid Engine (Sun Microsystems 2000 – EE.UU)

Sun Grid Engine (SGE) [42] es un sistema de gestor de colas, es útil cuando se tiene una gran cantidad de tareas a ejecutar y se desea distribuir las tareas sobre un grupo de máquinas.



SGE se encarga de la planificación de tareas, esto significa que puede presentar todos sus puestos de trabajo y SGE ponerlos en cola y ejecutarlas cuando los recursos requeridos estén disponibles. SGE también va a logra equilibrio de carga en la que se distribuyen los puestos de trabajo de manera que los nodos específicos no se sobrecargan. Además, permitirá hacer el monitoreo y contabilidad de trabajos que será útil cuando se desea comprobar si su trabajo está en marcha, y si no que le ayudará a entender lo que salió mal [43].

Link de descarga:

http://www.softpanorama.org/HPC/Grid_engine/Implementations/SGE62u5/index.shtml



Imagen 88: Interfaz de trabajo de GridEngine.

3.2.7. Torque (PBS - 2003 – EE.UU)

Torque [44] es un producto de código abierto avanzado basado en el original proyecto de PBS e incorpora lo mejor de la comunidad y el desarrollo profesional. Incorpora avances significativos en las áreas de escalabilidad, fiabilidad y funcionalidad y se encuentra actualmente en uso en decenas de miles de líderes gubernamentales, académicas y sitios comerciales en todo el mundo.

Torque proporciona control sobre los trabajos por lotes y recursos de computación distribuida.

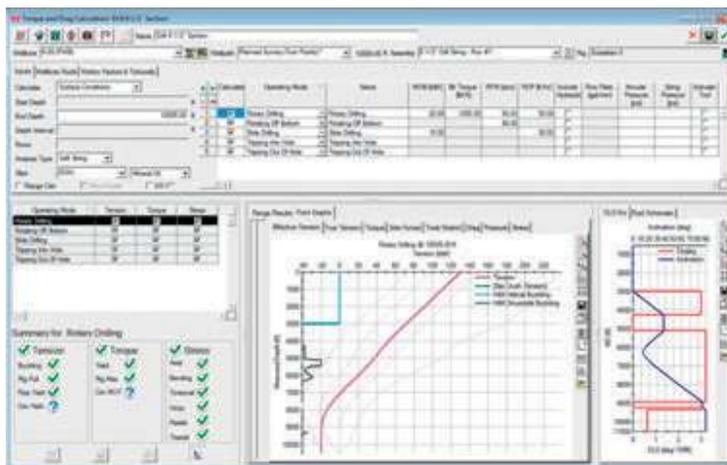


Imagen 89: Interfaz de trabajo de Torque

3.2.8. WinSCP (Kateřina Šrámková 2000 - EE.UU)

WinSCP es un cliente SFTP gráfico para Windows que emplea SSH. Su función principal es facilitar la transferencia segura de archivos entre dos sistemas informáticos, el local y uno remota que ofrezca servicios SSH.

WinSCP [45] es una aplicación de Software Libre.



WinSCP permite conectarse a un servidor SSH (Secure Shell) empleando el protocolo SFTP (SSH File Transfer Protocol) o el servicio SCP (Secure Copy Protocol).

Link de descarga:

<https://winscp.net/eng/download.php>.

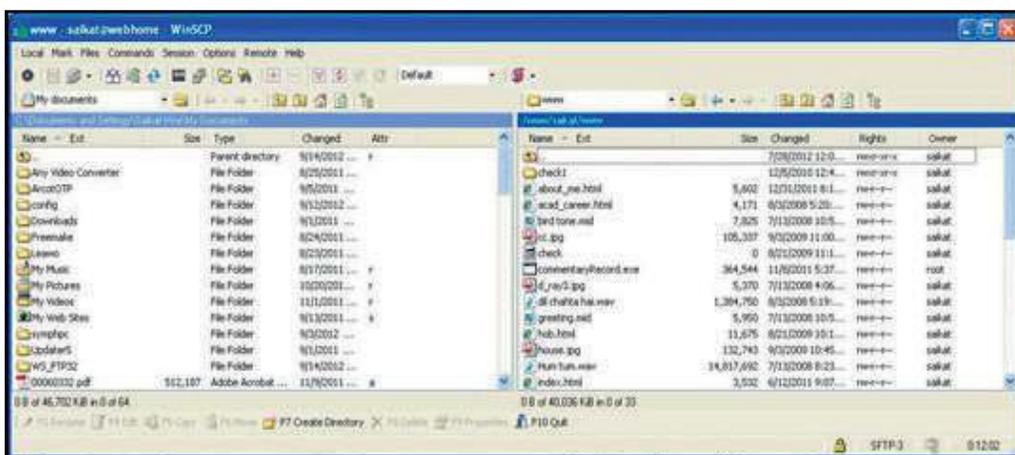


Imagen 90: Interfaz de trabajo de WinSCP

3.3. Software para desarrollo de aplicaciones y cálculo numérico en supercomputadoras

Tabla 20: Software para desarrollo de aplicaciones y cálculo numérico en HPC.

Id	Nombre	Año de creación	Usos principales
1	C	1970	Se ha utilizado para el desarrollo de muy diversas aplicaciones: sistemas operativos, hojas de cálculo, gestores de bases de datos.
2	C++	1983	Sistemas Operativos, drivers de dispositivos, gestores de paquetes y programas de configuración, aplicaciones gráficas de usuario, entre otros.
3	Fortran	1957	Se utiliza principalmente para aplicaciones científicas y el análisis numérico.
	Java	1995	Aplicaciones de cliente-servidor de web.
4	Python	1991	Desarrollar software para aplicaciones científicas, para comunicaciones de red, para aplicaciones de escritorio con interfaz gráfica de usuario (GUI), para crear juegos, para smartphones y para aplicaciones web.
5	R	1993	Investigación por la comunidad estadística, siendo además muy popular en el campo de la minería de datos, la investigación biomédica, la bioinformática y las matemáticas financieras.
6	Matlab	1984	Análisis de datos, modelado y diseño de sistemas DSP, modelado y control de sistemas de potencia y guiado.
7	Octave	1988	Está orientado al análisis y cálculo numérico.
8	Scilab	1994	Matemáticas y simulación, funciones de gráficos para visualizar 2D y 3D, herramientas para realizar análisis de datos y modelización, Xcos: simulador por diagramas en bloque de sistemas dinámicos híbridos, simulador de sistemas de modelado mecánico, circuitos hidráulicos, entre otros.

3.3.1. C (Dennis Ritchie-1972-Laboratorios Bell - EE.UU)

C [46] es un lenguaje de programación de propósito general que se utiliza para una amplia gama de aplicaciones, desde sistemas operativos como Windows y iOS a un software que se utiliza para la creación de películas en 3D.

Programación C es altamente eficiente, esa es la razón principal por la que es muy popular a pesar de tener más de 40 años de edad.



Características

- Es un lenguaje de procedimientos.
- Los programas en C son rápidos.
- Los programas estándar en C son portátiles-
- Modularidad (método que divide un problema, en problemas más pequeños, llamados subproblemas).
- Lenguaje tipo estático.
- Es de propósito general.

3.3.2. C++ (Bjarne Stroustrup – 1979 - Laboratorios Bell – EE.UU)

C++ [47] es un lenguaje de programación de propósito general sofisticado, eficiente y basado en C. Fue desarrollado por Bjarne Stroustrup en 1979.

Muchos de los sistemas operativos, controladores del sistema, navegadores y juegos de hoy en día utilizan C++ como su lenguaje principal. Esto hace que C++ uno de los lenguajes más populares hoy en día.



Características de C++

- Es rápido.
- Tipo estático.
- Lenguaje de programación multi-paradigma.
- Poder de la biblioteca estándar (Standard Template Library - STL).

3.3.3. FORTRAN (John Backus – 1956 – IBM – EE.UU)

El nombre de FORTRAN [48] es un acrónimo de FORmula TRANslation, es un lenguaje orientado al cálculo numérico, diseñado en sus inicios para las computadoras IBM y usado en aplicaciones científicas y de ingeniería y es el más antiguo de los lenguajes de alto nivel.



FORTRAN desde sus inicios dominó el área de la programación y se ha mantenido su desarrollo y aplicación por más de 50 años en distintos segmentos de la ciencia y técnica. FORTRAN uno de los lenguajes más usados en la Computación de alto rendimiento, además se utiliza para el desarrollo de programas que evalúan el desempeño y el posicionamiento de las Supercomputadoras. Este lenguaje fue propiedad absoluta de IBM hasta 1961, al pasar el tiempo se fue extendiendo a computadoras de otros fabricantes.

Características:

- Potencia en los cálculos matemáticos.
- Diseñado teniendo en cuenta el uso de la tarjeta perforada.
- No es muy usado fuera de los campos de la informática y el análisis numérico, pero permanece como el lenguaje a escoger para desempeñar tareas de computación numérica de alto rendimiento.

3.3.4. Java (Sun Microsystems – 1995 - EE.UU)

Como cualquier lenguaje de programación, Java [49] tiene su propia estructura, reglas de sintaxis y paradigma de programación. El paradigma de programación del lenguaje Java se basa en el concepto de programación orientada a objetos (POO), que las funciones del lenguaje soportan.



El lenguaje Java es un derivado del lenguaje C, por lo que sus reglas de sintaxis se parecen mucho a C.

Hay muchas aplicaciones y sitios web que no funcionarán a menos que tenga Java instalado y cada día se crean más. Java es rápido, seguro y fiable. Desde portátiles hasta centros de datos, desde consolas para juegos hasta súper computadoras, desde teléfonos móviles hasta Internet, Java está en todas partes [50].

Link de descarga:

<https://java.com/es/download/>

3.3.5. Python (Guido van Rossum – 1991 – Países Bajos)

Python [51] es un lenguaje de propósito general. Tiene una amplia gama de aplicaciones de desarrollo Web (como: Django y Bottle), para computación científica y matemática (Orange, Sympy, NumPy) y para escritorio con interfaces gráficas de usuario (Pygame, Panda3D)..



La sintaxis del lenguaje está limpia y la longitud del código es relativamente corto. Es divertido trabajar en Python, ya que le permite pensar sobre el problema en lugar de centrarse en la sintaxis

Características:

- Fácil de aprender.
- Libre y de código abierto.
- Portabilidad.
- Extensible e insertable (se puede combinar fácilmente piezas de C/C++ u otros lenguajes).
- Alto nivel, lenguaje interpretado.
- Bibliotecas para resolver las tareas comunes.
- Orientado a objetos.

Link de descarga:

<https://www.python.org/downloads/>

3.3.6. R (John Chambers– 1993 - Laboratorios Bell)

R es un lenguaje de programación [52] especialmente orientado al análisis estadístico y a la representación gráfica de los resultados obtenidos. Es un proyecto GNU. Por lo tanto, los usuarios son libres de modificarlo y extenderlo. Se trata de un lenguaje basado en comandos, en lugar de pinchar y arrastrar íconos o menús con el ratón se escriben comandos o instrucciones que son ejecutados. Una sucesión de instrucciones o comandos de R que implementa un flujo de trabajo para realizar una tarea se denomina script o guion R. Existe una amplia variedad de entornos de desarrollo para R que facilitan escribir scripts de R tales como R commander, RKWard y RStudio.



Características

- Un manejo eficaz de datos y la instalación de almacenamiento.
- Un conjunto de operadores para los cálculos en matrices.
- Una gran coherente colección integrada de herramientas intermedias para análisis de datos.
- Instalaciones gráficas para análisis de datos y de visualización.
- Un simple y eficaz lenguaje de programación bien desarrollado que incluye condicionales, bucles, funciones recursivas definidos por el usuario y las instalaciones de entrada y salida.

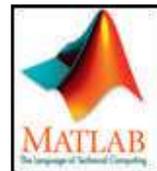
Link de descarga:

<https://cran.r-project.org/mirrors.html>

3.3.7. MATLAB (MathWorks- 1984 – EE.UU)

MATLAB es un lenguaje de alto rendimiento para la computación técnica. Se integra cómputo, visualización y programación en un entorno fácil de usar, donde los problemas y las soluciones se expresan en notación matemática conocida.

MATLAB es un sistema interactivo cuyo elemento básico de datos es una matriz que no requiere dimensionamiento. Esto le permite resolver muchos problemas de computación técnica, especialmente aquellos con formulaciones de matriz y el vector, en una fracción del tiempo que se tardaría en escribir un programa en un lenguaje no interactivo escalar como C o Fortran [53].



Link de descarga:

<https://www.mathworks.com/downloads>

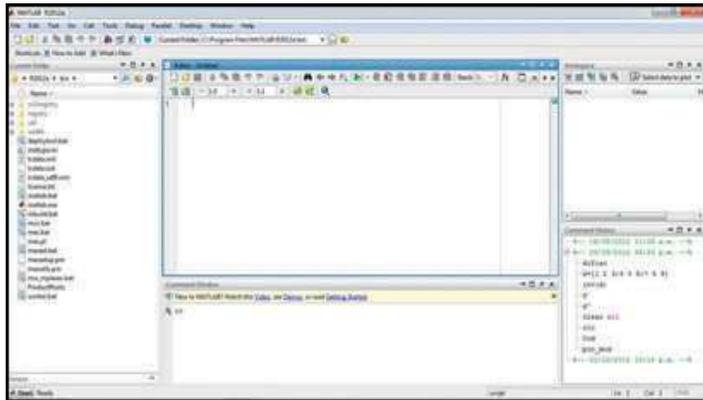


Imagen 91 Entorno de trabajo en Matlab

3.3.8. Octave (James B. Rawlings/ John G. Ekerdt – 1988 – EE.UU):

Octave es un proyecto de software libre maduro y consolidado. Es una herramienta de gran utilidad en el campo del cálculo numérico en todas sus vertientes.



Aunque no fue uno de sus objetivos iniciales, en su madurez ha ido convergiendo hacia una de las estrellas del Software en ingeniería, Matlab. Octave se ha convertido en una alternativa en dos casos: cuando la adquisición de Matlab no es viable por cualquier motivo y cuando la pericia de un buen programador en C++ puede extender su funcionalidad [54].

Octave tiene amplias herramientas para resolver problemas de álgebra lineal numérica comunes, encontrar las raíces de ecuaciones no lineales, la integración de las funciones ordinarias, manipular polinomios, e integrando diferenciales ordinarias y ecuaciones diferenciales algebraicas. Es fácilmente extensible y adaptable a través de funciones definidas por el usuario escritas en el idioma del Octave, o el uso de módulos cargados dinámicamente escritas en C++, C, Fortran, u otros idiomas [55].

Link de descarga:

<https://www.gnu.org/software/octave/download.html>

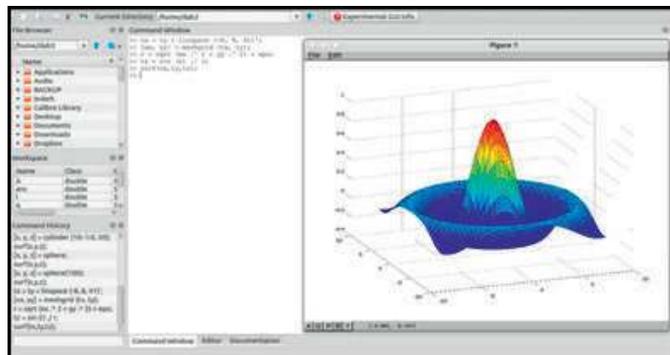


Imagen 92: Entorno de trabajo en Octave

3.3.9. Scilab (INRIA- 1994 - Francia)

Se ha desarrollado para el control del sistema y las aplicaciones de procesamiento de señales. Se distribuye libremente en formato de código fuente.



Una característica clave de la sintaxis de Scilab es su capacidad de manejar matrices: se llevan a cabo inmediatamente manipulaciones básicas de matriz tales como concatenación, extracción o de transposición, así como las operaciones básicas tales como suma o multiplicación [56].

Características:

- Entorno de programación libre.
- Las funciones y librerías de funciones están completamente en las manos del usuario. Además, sus funciones son conocidos como objetos de datos en Scilab.

Link de descarga:

<http://www.scilab.org/download/latest>

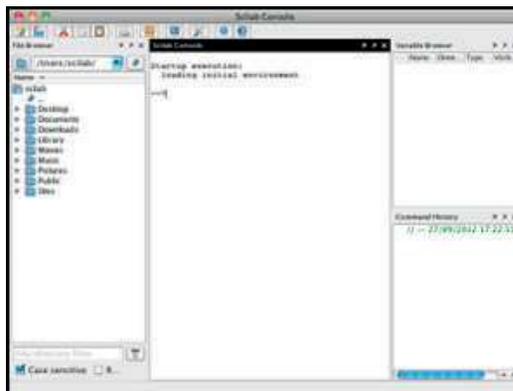


Imagen 93: Entorno de trabajo de Scilab.

3.4. APIs y Librerías

Tabla 21: Lista de APIs y Librerías.

N°	Nombre	Año de creación	Usos principales
1	Apache Hadoop	2011	Procesamiento distribuido de grandes conjuntos de datos a través de grupos de ordenadores.
2	Cuda	2007	Procesamiento de vídeo e imágenes, datos de biología y química computacional, simulación de la dinámica de fluidos, reconstrucción de imágenes de TC, análisis sísmico o trazado de rayos, entre otras.
3	MPI	1994	Es un mecanismo para la programación paralela por paso de mensajes. Por ello, no es necesario que todos los procesos MPI ejecuten en el mismo nodo.
4	OpenCL	2008	Permite implicar en los cálculos de una tarea a todos los núcleos del procesador central o toda la capacidad de computación de la unidad de procesamiento gráfico, cosa que, a fin de cuentas, reduce significativamente el tiempo de ejecución de un programa. Por eso, el uso de OpenCL es muy útil en las tareas relacionadas con computaciones muy laboriosas o que gastan muchos recursos.
5	OpenMP	1997	Permite añadir concurrencia a los programas escritos en C, C++ y Fortran sobre la base del modelo de ejecución fork-join.
6	TBB	2006	Usado en predicción meteorológica numérica, oceanografía y astrofísica, análisis de elementos finitos, IA y automatización, ingeniería genética.

3.4.1. Apache Hadoop (Apache Software Foundation -2011 - EE.UU):

La biblioteca de software Apache Hadoop [57] es un marco que permite el procesamiento distribuido de grandes conjuntos de datos a través de grupos de ordenadores que utilizan modelos de programación simples. Está diseñado para escalar desde los servidores individuales a miles de máquinas, cada una ofreciendo computación y almacenamiento local.



En lugar de depender de hardware para ofrecer alta disponibilidad, la biblioteca en sí está diseñada para detectar y manejar los fallos en la capa de aplicación, por lo que la entrega de un servicio de alta disponibilidad en la parte superior de un grupo de ordenadores, cada uno de los cuales puede ser propenso a fallos.

Una amplia variedad de empresas y organizaciones utilizan Hadoop tanto para la investigación y la producción.

Link de descarga:

<http://hadoop.apache.org/releases.html>

3.4.2. CUDA (NVIDIA – 2007 – EE.UU):

Compute Unified Device Architecture (CUDA) [58] es una arquitectura de cálculo paralelo de NVIDIA que aprovecha la gran potencia de la GPU (Unidad de Procesamiento Gráfico) para proporcionar un incremento extraordinario del rendimiento del sistema.



La plataforma de cálculo paralelo CUDA, proporciona unas cuantas extensiones de C y C++ que permiten implementar el paralelismo en el procesamiento de tareas y datos con diferentes niveles de granularidad. El programador puede expresar ese paralelismo mediante diferentes lenguajes de alto nivel como C, C++ y Fortran o mediante estándares abiertos como las directivas de OpenACC. En la actualidad, la plataforma CUDA se utiliza en miles de aplicaciones aceleradas en la GPU y en miles de artículos de investigación publicados.

Una amplia variedad de empresas y organizaciones utilizan Hadoop tanto para la investigación y la producción.

Link de descarga:

<https://developer.nvidia.com/cuda-downloads>

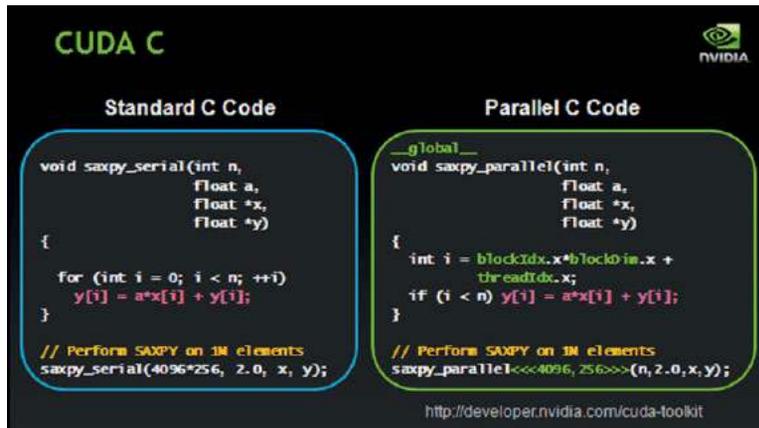


Imagen 94: Comparación de código C y código C paralelo.

3.4.3. MPI (Jack Dongarra/Tony Hey/David W. Walker – 1994 - EE.UU)

Message Passing Interfaz (MPI) [57] es una especificación para programación de paso de mensajes, que proporciona una librería de funciones para C, C++ o Fortran que son empleadas en los programas para comunicar datos entre procesos.



MPI es la primera librería de paso de mensajes estándar y portable, especificada por consenso por el MPI Forum (<http://www.mpi-forum.org>), con unas 40 organizaciones participantes, como modelo que permita desarrollar programas que puedan ser migrados a diferentes computadores paralelos.

Su ventaja más inmediata es su implementación y portabilidad a una gran variedad de sistemas: desde máquinas con memoria compartida (Origin2000) hasta una red de estaciones de trabajo [58].

Características:

- Estandarización.
- Portabilidad.
- Buenas prestaciones.
- Amplia funcionalidad.
- Existencia de implementaciones libres.

3.4.4. OpenCL (APPLE -2008 – EE.UU)

OpenCL (marca comercial de Apple Computers Inc.) [61] Es un estándar, que hace posible cálculos mucho más rápidos. El estándar es controlado por la organización de estándares sin fines de lucro Khronos. Mediante el uso de esta técnica y las tarjetas gráficas (GPU) o extensiones de los procesadores modernos se puede, por ejemplo, convertir un video en 20 minutos en lugar de 2 horas.



OpenCL [62] es útil en las tareas relacionadas con computaciones muy laboriosas o que gastan muchos recursos. Al contrario de CUDA o Stream, OpenCL fue creado originalmente por Apple quien luego la propuso al Grupo Khronos para convertirlo en un estándar abierto y libre que no dependa de un hardware de un determinado fabricante (CUDA sólo está disponible en gráficas NVidia y Stream en gráficas de ATI).

Link de descarga:

<https://developer.nvidia.com/ocl>

3.4.5. OpenMP (OpenMP Architecture Review Board - 1997 - EE.UU):

OpenMP [63] es una interfaz de programación de aplicaciones (API) para la programación multiproceso de memoria compartida en múltiples plataformas. También podría definirse como un modelo de programación portable y escalable que proporciona a los programadores una interfaz simple y flexible para el desarrollo de aplicaciones paralelas para las plataformas que van desde las computadoras de escritorio hasta las supercomputadoras.



Esta API ofrece todo el potencial de los nodos de cómputo, ya que se caracterizan por su gran volumen de memoria compartida.

OpenMP proporciona un modelo portátil, escalable para los desarrolladores de aplicaciones paralelas de memoria compartida. La API es compatible con C/C++ y Fortran [64]

Características:

- Facilidad de uso.
- Capacidad de poner en paralelo de forma incremental un programa en serie.

- Portabilidad.

Link de descarga:

<http://www.openmp.org/resources/openmp-compilers/>

3.4.6. TBB (Intel – 2006 – EE.UU)

Intel Threading Building Blocks (Intel TBB) es una biblioteca de C++ para el paralelismo de tareas, el cual permite escribir fácilmente programas paralelos de C++ que aprovechan al máximo el rendimiento de múltiples núcleos, que son portátiles y componibles, y escalables [65].



Características:

- Algoritmos paralelos y estructura de datos.
- Asignación de memoria escalable y programación de tareas.

Link de descarga:

<https://www.threadingbuildingblocks.org/download>

3.5. Software especializado

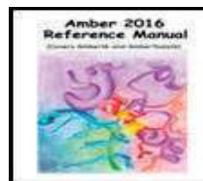
Tabla 22: Lista de Software especializado.

N°	Nombre	Año de Creación	Usos Principales
1	Ambertools 16	2002	Simulación de campos de fuerza de dinámica molecular de biomoléculas.
2	Gaussian	1970	Química Computacional.
3	ORCA	1999	Química cuántica con énfasis específico en las propiedades espectroscópicas de las moléculas open-shell.
4	AutoDock	1990	Simulación de modelado molecular, especialmente para acoplamiento de proteína-ligando.
5	Dock 6	1980	Predecir modos de unión de los pequeños complejos de molécula de proteínas, bases de datos de búsqueda de ligandos para los compuestos que se unen a una proteína particular, bases de datos de búsqueda de ligandos para compuestos que se unen dianas de ácido nucleico, examinar las posibles orientaciones de unión de la proteína-proteína y proteína de ADN, ayudar a guiar los esfuerzos sintéticos mediante el examen de pequeñas moléculas que se derivatizan computacionalmente.

3.5.1. Software para química computacional

Ambertools 16 (Universidad de California – 2016 – San Francisco, EE.UU):

AmberTools [66] consiste en varios paquetes desarrollados independientemente que funcionan bien por sí mismos. La suite AmberTools es gratis, y sus componentes son liberados en su mayoría bajo la Licencia Pública General de GNU (GPL). Unos componentes se incluyen los que están en el dominio público o que tienen otros, de código abierto, licencias.



Se distribuye en formato de código fuente, y debe ser compilado para poder ser utilizado. En compiladores de C, C++ y Fortran90.

Se puede utilizar para llevar a cabo completas simulaciones de dinámica molecular.

Link de descarga:

<http://ambermd.org/AmberTools16-get.html>

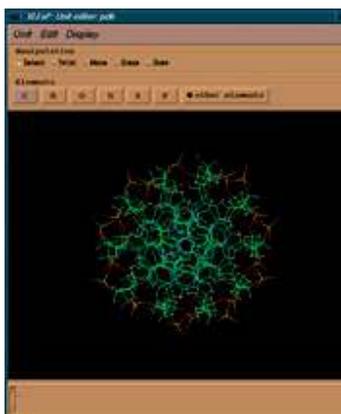


Imagen 95: Ejemplo de gráfica de una molécula usando AmberTools.

Gaussian (Gauss, Inc. – 1970 - EE.UU)

Es un software que predice las energías, estructuras moleculares y frecuencias de vibración de sistemas moleculares, junto con las diversas propiedades moleculares que se derivan.

Gaussian permite crear un modelo de reacciones químicas complejas y sustancias al más alto nivel de precisión disponibles [67].



Para adquirir el software es necesario comunicarse con la empresa <http://gaussian.com/>

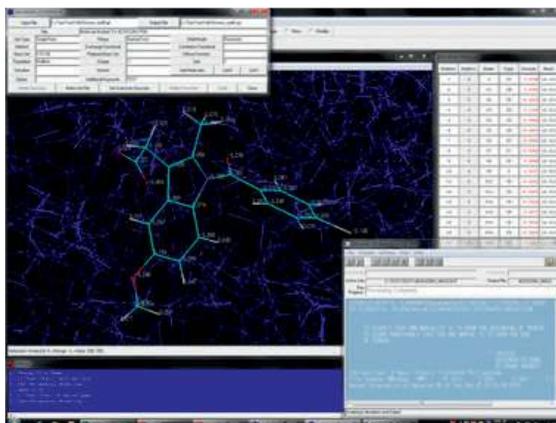


Imagen 96: Interfaz de Gaussian

Orca (F. Neese – 1999 – Universidad de Stanford, EE.UU):

ORCA es una herramienta de propósito general flexible, eficiente y fácil de usar para la química cuántica con énfasis específico en propiedades espectroscópicas de moléculas de capa abierta.



Cuenta con una amplia variedad de métodos de química cuántica estándar que van desde métodos semiempíricos para DFT (Teoría funcional de la densidad) a un solo y multireferencia correlacionado métodos ab initio. También puede tratar los efectos ambientales y relativistas.

Debido al estilo fácil de usar, ORCA se considera que es una herramienta útil no sólo para los químicos computacionales, sino también para los químicos, físicos y biólogos que están interesados en el desarrollo del contenido de la información completa de sus datos experimentales con ayuda de cálculos.

ORCA es capaz de llevar a cabo optimizaciones de geometría y de predecir un gran número de parámetros espectroscópicos en diferentes niveles de teoría. Además del uso de la teoría Hartree Fock, la teoría funcional de la densidad (DFT) y los métodos semiempíricos, alto nivel ab initio métodos químicos cuánticos, basado en la interacción de configuración y métodos de racimo acoplados, se incluyen en ORCA a un grado creciente [68].

Link de descarga:

<https://orcaforum.cec.mpg.de/downloads.php>

3.5.2. Software para Bioinformática

AutoDock (The Scripps Research Institute – 1990 –EE.UU)

AutoDock [69] es un conjunto de programas diseñados para predecir la forma en que moléculas pequeñas, como sustratos o candidatas a fármacos, se unen a un receptor biológico de estructura tridimensional conocida.



Constan de dos generaciones de software: AutoDock 4 y AutoDock Vina.

AutoDock 4 consta en realidad de dos programas principales: AutoDock realiza el acoplamiento del ligando a un conjunto de rejillas que describen la proteína Diana; AutoGrid pre-calcula estas rejillas.

Además de su uso para el docking, las rejillas de afinidad atómicas pueden ser visualizadas. Esto puede ayudar, por ejemplo, para guiar a los químicos sintéticos orgánicos diseñar mejores aglutinantes.

AutoDock Vina no requiere la elección de tipos de átomos y mapas de la red de pre-cálculo para ellos. En lugar de ello, calcula las rejillas internamente, para los tipos de átomos que se necesitan, y lo hace prácticamente al instante.

AutoDock es muy rápido, proporciona predicciones de alta calidad y buenas correlaciones entre las constantes de inhibiciones predichas y los experimentales. AutoDock también ha demostrado ser útil para el acoplamiento ciego, cuando no se conoce la ubicación del sitio de unión.

Link de descarga:

<http://autodock.scripps.edu/downloads>

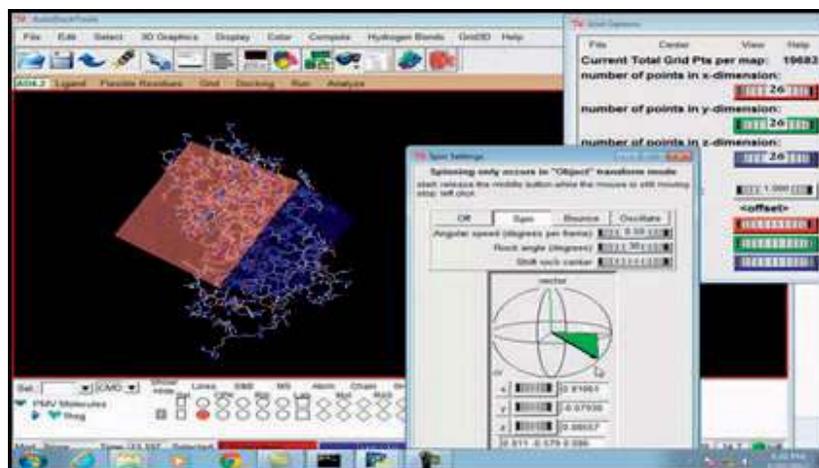


Imagen 97: Interfaz de trabajo de Autodock

Dock 6(Kuntz's Group – 1980 – Universidad California, San Francisco, EE.UU)

Está escrito en C++ y se separa funcionalmente en componentes independientes, lo que permite un alto grado de flexibilidad del programa. Los programas y accesorios están escritas en C y Fortran 77. Se proporcionan código fuente de todos los programas.



La suite de programas DOCK requiere del orden de 100 MB de espacio en disco y 48 MB de RAM.

Las nuevas características de DOCK 6 incluyen: opciones de puntuación adicionales durante la reducción al mínimo [70].

Permite: Examinar las posibles orientaciones de unión de la proteína-proteína y proteína de ADN, ayudar a guiar los esfuerzos sintéticos mediante el examen de pequeñas moléculas que se derivatizan computacionalmente.

Están disponibles de forma gratuita para las instituciones académicas, pero hay una cuota de licencia para las organizaciones industriales [71].

Para los académicos:

http://dock.compbio.ucsf.edu/Online_Licensing/dock_license_application.html

Para las organizaciones industriales: e-mail al administrador de licencia:

dock_industry@dock.compbio.ucsf.edu.

3.5.3. Software para medir rendimiento de PC

SiSoft Sandra (SiSoftware – 1997 – Inglaterra) Es un sistema de información y una herramienta de diagnóstico que puede proporcionar todo tipo de información útil acerca del hardware PC.



El programa ofrece una amplia gama de puntos de referencia para poner a prueba la velocidad de las unidades de disco duro, almacenamiento extraíble, sistema de archivos, transferencia de archivos del dispositivo móvil, la velocidad del procesador, procesamiento de vídeo, el ancho de banda de memoria y rendimiento de la red, entre otros. Estos controles son particularmente útiles si usted está tratando de optimizar su sistema con ajustes de hardware o software [72].

Link de descarga:

<http://www.sisoftware.eu/download-lite/#>

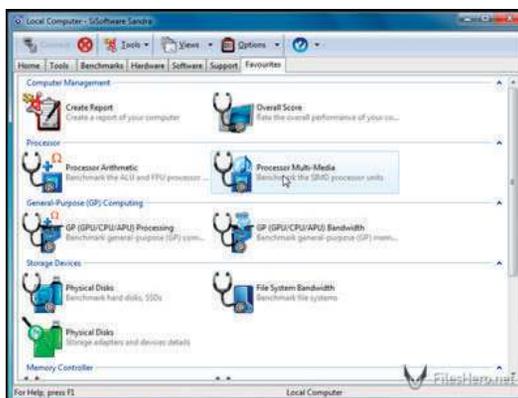


Imagen 98: Interfaz de Trabajo SiSoftware Sandra.

3DMark (Corporación Futuremark - 1998 - Finlandia)

3DMark incluye todo lo necesario para comparar su PC y dispositivos móviles en una sola aplicación. Ya sea que esté de juego en un teléfono inteligente, tableta, portátil o un PC de escritorio del juego, 3DMark incluye un punto de referencia diseñado específicamente para su hardware.



Una nueva interfaz que es más rápida, más flexible y más fácil de usar. Lo que es más, ahora puede obtener descargas más rápidas y ahorrar espacio de almacenamiento al optar por instalar sólo las pruebas que necesita [73].

Link de descarga:

<https://www.futuremark.com/benchmarks/3dmark>

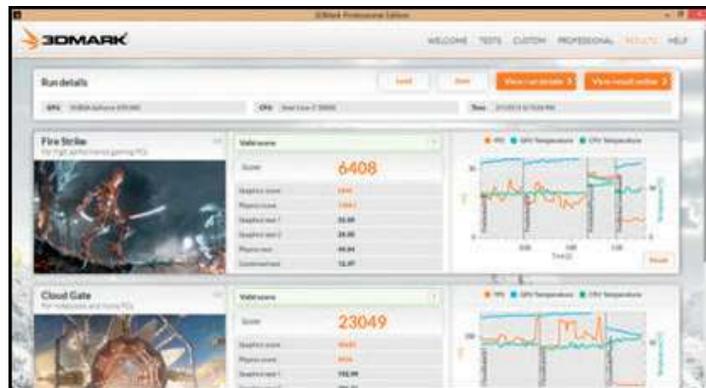


Imagen 99: Interfaz de trabajo en 3DMark

PCMark (Corporación Futuremark - 2002 - Finlandia)

PCMark 8 [74] es la versión más reciente de nuestra serie de herramientas de evaluación comparativa de PC estándar de la industria. Con PCMark 8 se puede probar el funcionamiento de todo tipo de PC, desde tabletas a los escritorios. Con cinco pruebas de referencia separadas, además de las pruebas de la vida de la batería, PCMark 8 ayuda a encontrar los dispositivos que ofrecen la combinación perfecta de eficiencia y rendimiento. PCMark 8 es la referencia completa de PC para el hogar y la empresa.

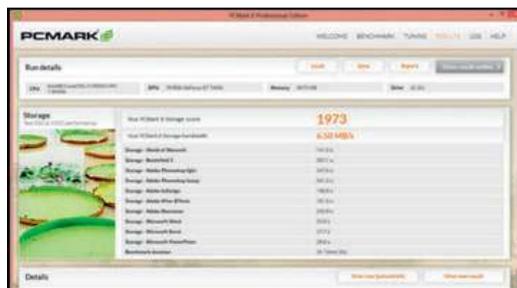


Imagen 100: Interfaz de PCMARK 8

CAPÍTULO 4

LA SUPERCOMPUTACIÓN EN EL MUNDO

Entre las principales interrogantes que surgen frecuentemente al hablar de Supercomputación en el mundo destacan preguntas como ¿cuál es la Supercomputadora más poderosa del mundo?, ¿dónde funciona?, ¿qué marca es?, ¿qué sistema operativo utiliza?. Hablar sobre supercomputación en realidad va mucho más allá de estas interrogantes, considerando lo cambiante que es la tecnología computacional. Lo primero que podemos comentar es que actualmente existe una fuerte disputa o carrera entre Estados Unidos y China, por el liderazgo mundial de construir y alojar la supercomputadora más potente del mundo, a la vez compiten en la cantidad de supercomputadoras más potentes en el ranking Top500 (estos 2 países han sacado considerable ventaja frente al resto de naciones).

Si bien es cierto en los inicios de la supercomputación existió una frecuente hegemonía de Estados Unidos, aunque recién desde junio de 1993 existe un registro apropiado de este tipo de tecnologías a través de ranking Top500. Aunque en aquellos años la disputa era entre Estado Unidos y Japón (el Top500 fue encabezado por un equipo marca Thinking Machines Corporation ubicado en Los Alamos National Laboratory en Nuevo Mexico, USA). Paradójicamente no existía ninguna supercomputadora china en el ranking, en los años posteriores China se preparó hasta llegar a ser lo que es hoy.

Según el último ranking Top500 de noviembre del 2016, Estados Unidos y China están actualmente empatados en cantidad de supercomputadoras, cada uno con 171 (entre ambos suman 342, que equivale el 68.4% de las 500 supercomputadoras más poderosas de mundo). Lo que se debe destacar es que China lidera el ranking desde junio del 2013, primero con el Tianhe-2 (MilkyWay-2) del National University of Defense Technology (Changsha, CHINA) y en la actualidad con Sunway TaihuLight ubicado en el National Supercomputing Center in Wuxi (Wuxi, Jiangsu, China), incluso los procesadores Sunway SW26010 son de fabricación china, lo que prácticamente consolida la hegemonía China en estos últimos años.

Respecto a Latinoamérica, según el propio ranking, es contundente el liderazgo de Brasil y México. En el caso brasileño es destacable la considerable inversión en laboratorios de alto rendimiento y desarrollo de capacidades. Respecto a Sudamérica aparecen al menos una vez en el ranking Colombia, Perú y Venezuela.

4.1. La Supercomputación a noviembre del 2016

La fuente más importante del mundo relacionada al estudio de las capacidades de las supercomputadoras es el Top500 (www.top500.org). Este Ranking es publicado desde 1993 cada junio y noviembre. En noviembre del 2016 se publicó la última versión (ver anexo 02), que evidencia ya un ligero liderazgo de China, lastimosamente no aparece ninguna supercomputadora peruana, incluso a nivel de Latinoamérica apenas aparecen 3 equipos brasileiros en funcionamiento. Seguidamente presentamos información sobre las computadoras con mayores capacidades de procesamiento, los países donde funcionan, fabricantes, arquitecturas, sistemas operativos, familias de procesadores más usados y los principales usos que se dan.

4.1.1. Las Supercomputadoras más poderosas del mundo

Según el Top500, desde junio de 2016 la Supercomputadora de mayor capacidad de procesamiento del mundo es la llamada Sunway TaihuLight que viene funcionando en el National Supercomputing Center in Wuxi en China, y que superó los 93 petaflops de operaciones de coma flotante (que es equivalente a unos 2 millones 325 mil Computadoras Portátiles con procesador Intel Core i3 o más de 2 millones 600 mil Teléfonos Celulares Samsung Galaxy S6). Además esta Supercomputadora cuenta con una capacidad teórica superior a 125 teraflops. China además también alberga a la segunda Supercomputadora del mundo, que es la denominada Tianhe-2 (MilkyWay-2) que vienen funcionando en el National Super Computer Center in Guangzhou que alcanza 33.8 petaflops (lideró el ranking hasta noviembre del 2015).

Hasta el año 2013 Estados Unidos lideró el Ranking Top500 con la Supercomputadora Titan, que alcanza los 17 petaflops, en la actualidad ésta Supercomputadora es la tercera del Top500. De las 10 Supercomputadoras más potentes del mundo 5 de ellas funcionan en los Estados Unidos, además destacan entre los primeros 10, 2 Supercomputadoras de Japón y una Supercomputadora de Suecia. Además es interesante mencionar que 9 de las primeras 10 del Top son utilizados en actividades de investigación y las restantes en actividades académicas. Así mismo es destacable comentar que en el actual Top500, 117 supercomputadoras logran superar 1 petaflops, equivalente a 106 gigaflops.

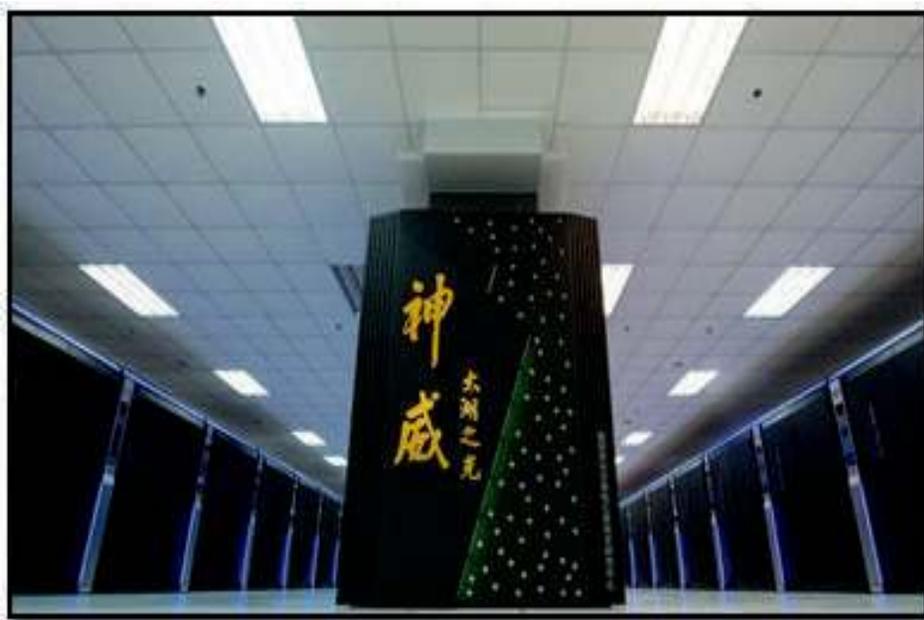


Imagen 101: Supercomputadora Sunway TaihuLight instalada en el National Supercomputing Center in Wuxi en China, Actualmente líder del Top500. Créditos: laboratorio linux

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Tabla 23: Las 25 Supercomputadoras de mayor capacidad del mundo según el Top500 a noviembre de 2016.

Rank	Nombre	Ubicación	Marca	País	Rmax (GIGAFLOPS)	Uso
1	Sunway TaihuLight	National Supercomputing Center in Wuxi	NRCPC	China	93014593.9	Investigación
2	Tianhe-2 (MilkyWay-2)	National Super Computer Center in Guangzhou	NUDT	China	33862700	Investigación
3	Titan	DOE/SC/Oak Ridge National Laboratory	Cray Inc.	Estados Unidos	17590000	Investigación
4	Sequoia	DOE/NNSA/LLNL	IBM	Estados Unidos	17173224	Investigación
5	Cori	DOE/SC/LBNL/NERSC	Cray Inc.	Estados Unidos	14014700	Investigación
6	Oakforest-PACS	Joint Center for Advanced High Performance Computing	Fujitsu	Japón	13554600	Académico
7	K computer, SPARC64 VIIifx 2.0GHz, Tofu interconnect	RIKEN Advanced Institute for Computational Science (AICS)	Fujitsu	Japón	10510000	Investigación
8	Piz Daint	Swiss National Supercomputing Centre (CSCS)	Cray Inc.	Suecia	9779000	Investigación
9	Mira	DOE/SC/Argonne National Laboratory	IBM	Estados Unidos	8586612	Investigación
10	Trinity	DOE/NNSA/LANL/SNL	Cray Inc.	Estados Unidos	8100900	Investigación
11	Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	United Kingdom Meteorological Office	Cray Inc.	Reino Unido	6765250	Investigación
12	Marconi Intel Xeon Phi	CINECA	Lenovo	Italia	6223040	Académico
13	Pleiades	NASA/Ames Research Center/NAS	HPE/SGI	Estados Unidos	5951550	Investigación
14	Hazel Hen	HLRS - Höchstleistungsrechenzentrum Stuttgart	Cray Inc.	Germany	5640170	Investigación
15	Shaheen II	King Abdullah University of Science and Technology	Cray Inc.	Saudi Arabia	5536990	Academico
16	Pangea	Total Exploration Production	HPE/SGI	France	5283110	Industria
17	Stampede	Texas Advanced Computing Center/Univ. of Texas	Dell	United States	5168110	Academico
18	Theta	DOE/SC/Argonne National Laboratory	Cray Inc.	United States	5095840	Investigación
19	JUQUEEN	Forschungszentrum Juelich (FZJ)	IBM	Germany	5008857	Investigación
20	Cheyenne	National Center for Atmospheric Research (NCAR)	HPE/SGI	Estados Unidos	4788190	Investigación
21	Vulcan	DOE/NNSA/LLNL	IBM	Estados Unidos	4293306	Investigación
22	Abel	Petroleum Geo-Services	Cray Inc.	Estados Unidos	4042460	Industria
23	Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	ECMWF	Cray Inc.	United Kingdom	3944680	Investigación
24	Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	ECMWF	Cray Inc.	United Kingdom	3944680	Investigación
25	Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K39	Government	Cray Inc.	United States	3577000	Gobierno

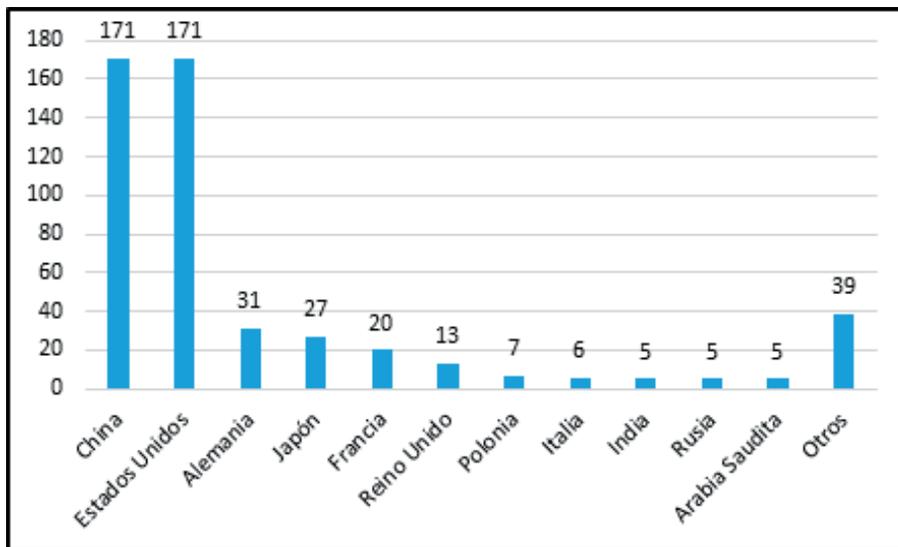
Fuente: Top500 - Elaboración propia.

4.1.2. Supercomputadoras por países

Históricamente Estados Unidos ha liderado registros en cantidad de Supercomputadoras en el ranking de supercomputadoras del Top500 desde junio de 1993. Sorpresivamente en junio del 2016 China lideró por primera vez el Top500 con un registro de 168 Supercomputadoras versus 165 de Estados Unidos. En el último Top500 de noviembre del 2016 cada uno registró 171 estableciéndose un empate, sin embargo si se tiene en cuenta que las 2 primeras Supercomputadoras actuales son de China, ya se habla de un liderazgo Chino en la Supercomputación a nivel mundial. Cabe indicar que en el caso peruano la Supercomputadora adquirida por la Universidad San Luis Gonzaga de Ica fue comprada a la marca china INSPUR.

La disputa titánica entre Estados Unidos y China viene minimizando el desarrollo y uso de estas tecnologías en los demás países. Muestra de ello por ejemplo es que el tercer país del ranking es Alemania con 31 Supercomputadoras, 140 menos que Estados Unidos y China. Seguidamente aparece Japón (antigua competencia de Estados Unidos) con 27, Francia 20, Reino Unido con 13, además aparecen en el Top500 Polonia (7), Italia (6), India (5), Rusia (5), Arabia Saudita (5). En el gráfico n° 1, Brasil está considerado entre los países denominados Otros, con apenas 3 Supercomputadoras, 3 menos de lo que tuvo en los ranking del 2015.

Gráfico N° 1: Cantidad de supercomputadoras por país según Top500 de noviembre del 2016.



Fuente: Elaboración propia según estadísticas del Top500.

4.1.3. Las marcas de las Supercomputadoras en el Top500

Respecto a las empresas fabricantes de Supercomputadoras en 23 años que tiene el Top500, esto ha sido igual de dinámico que el resto del sector informático, sobretodo en la creación, desaparición o fusión de marcas, aunque en este caso, con algunas empresas netamente especializadas en soluciones de alto rendimiento.

Empresas como Thinking Machines Corporation, cuyo equipo CM-5/1024 que funcionó en Los Alamos National Laboratory (California, Estados Unidos) fue la Supercomputadora de mayor capacidad de procesamiento del mundo en junio de 1993, sin embargo esta marca desapareció en los noventa (lo que quedó de ella después de su quiebra fue adquirida por Sun Microsystem).



Imagen 102: Computador CM-5. Créditos: mark Pellegrini

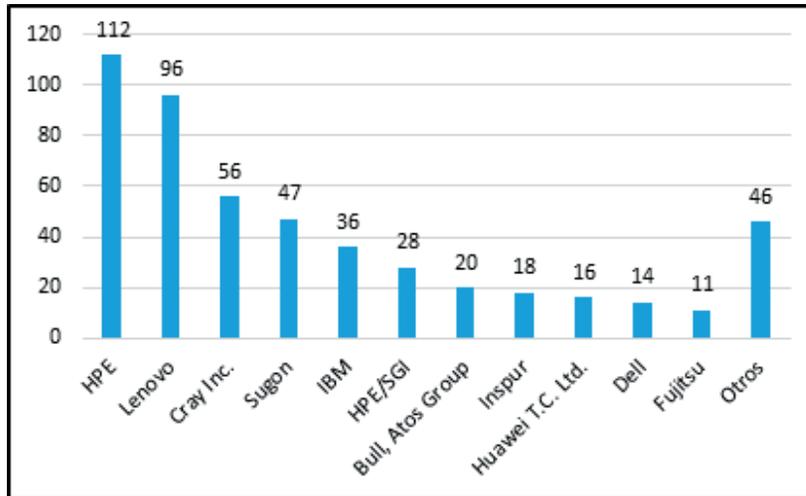
A la vez han surgido nuevas marcas, incluso en países como China que en el año 1993 prácticamente no existía en el mundo de la Supercomputación.

En estos últimos años es muy destacable el crecimiento de la industrias de la computación de China, que incluso a través de su gobierno viene impulsando la creación de propia tecnología para Supercomputación (como por ejemplo procesadores como Sunway SW26010 utilizados en el Sunway TaihuLight), y con ello el surgimiento y crecimiento de empresas o marcas fabricantes chinos de Supercomputadoras. En el último Top500 aparecen empresas chinas como Lenovo (segunda ubicación), Sugon (cuarta ubicación) e Inspur, y se prevé que en los próximos años esta participación incluso crezca.

Por otro lado en junio de 1993 el 42% (205) supercomputadoras eran de la marca norteamericana Cray Inc., en el Top500 de noviembre del 2016 apenas alcanzan el 11% (56%). Además en la actualidad la empresa norteamericana Hewlett Packard Enterprise HPE es la que lidera el ranking con el 22% (112), aunque es importante indicar que fines del 2016 incluso HPE compró la empresa Silicon Graphics International – SGI que aparece en el mismo ranking con el 5.6%, lo que daría ya a HPE una cifra superior al 27%. Es pertinente además comentar la histórica lucha entre las marcas Hewlett Packard Enterprise e IBM. Si bien es cierto IBM en el Top500 de noviembre del 2016 aparece con apenas 7.2%(36), sin embargo hace una década lideraba los Top500 con registros tan altos como 47.2% (236) en junio del 2006 y 48% en noviembre del 2006 (240), es evidente el desinterés de IBM en los últimos años por este mercado.

Para el caso del Perú tenemos al menos 4 Supercomputadoras de la marca SGI (IIAP, UNSA-AREQUIPA, SENAMHI y CORPAC), el equipo del Instituto Geofísico del Perú y del Instituto del Mar del Perú son de la marca DELL y de la Universidad San Luis Gonzaga de Ica es de la marca China INSPUR.

Gráfico N° 2: Cantidad de Supercomputadoras por marca fabricante según Top500 de noviembre del 2016.



Fuente: Elaboración propia según estadísticas del Top500

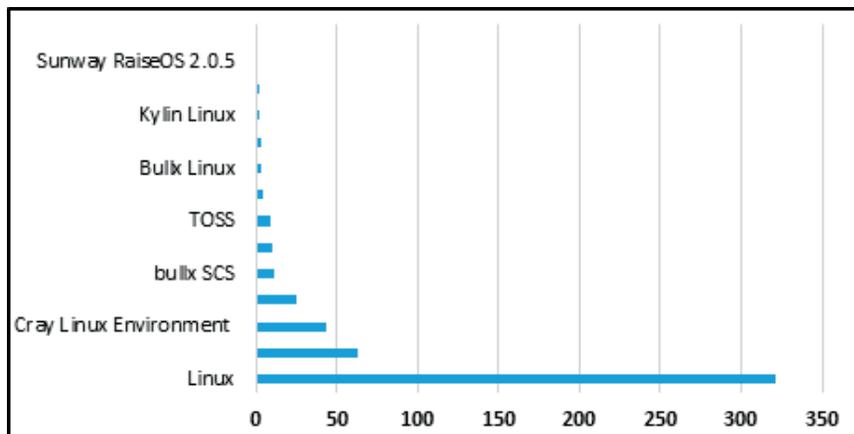
4.1.4. Sistemas Operativos utilizados en el Top500

En la actualidad la familia Linux (Linux y una serie de distribuciones basadas en el kernel Linux), son los sistemas operativos más utilizados mundialmente en Supercomputación. Según el último reporte del Top 500 de noviembre del 2016, el 99.6% (498 supercomputadoras) usan algún sistema operativo de la familia Linux (entre los que destacan Centos, Cray Linux, Suse Linux, entre otras), el restante 0.4% (2) es IAX que está familiarizado con UNIX.

Sin embargo no siempre fue así, por ejemplo en junio del 1993 el sistema operativo más utilizado era el UNICOS, que era una versión de UNIX para las Supercomputadoras CRAY. Linux apareció por primera vez en el Top500 de junio de 1998 en una Supercomputadora de Los Alamos National Laboratory (California, Estados Unidos). En junio de 2004 por primera vez en el Top500 el sistema operativo Linux estuvo instalado en más de 250 Supercomputadoras. Hasta ahí históricamente los sistemas operativos para Supercomputadoras estaban basadas en UNIX.

Por otro lado es interesante mencionar que también en la base de datos histórica del Top500 se identificó Supercomputadoras con Sistemas Operativos de la familia del Windows, aunque no pasaron de 6 por cada ranking (en junio del 2008), incluso el Top500 de noviembre del 2014 aparece por última vez en el Shanghai Supercomputer Center (China). Entre las versiones Windows que se usaron en Supercomputadoras aparecen Windows 2000, Windows 2003 Server, Windows Compute Cluster Server 2003, Windows HPC 2008, Windows Azure.

Gráfico N° 3: Cantidad de supercomputadoras por sistema operativo según Top500 de noviembre de 2016



Fuente: Elaboración propia según estadísticas del Top500

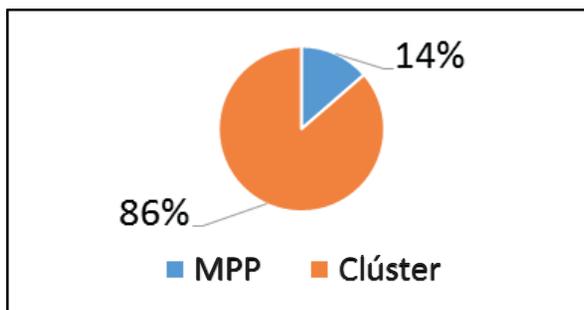
4.1.5. Arquitecturas más utilizadas por las Supercomputadoras

En la última versión del Top500 sólo aparecen 2 tipos de arquitecturas de supercomputadoras. La primera, la denominada Clúster, que es un conjunto de computadoras o nodos de gran capacidad, cada uno con una propia memoria, y todos ellos conectados a grandes velocidades. La segunda denominada MPP (Massively Parallel Processor - Procesadores Masivamente Paralelos), que son equipos cuyos procesadores se encuentran conectados por mecanismos tecnológicos de alta velocidad, en este caso cada procesador tiene su propia memoria.

Actualmente el 86% (430) de supercomputadoras tiene arquitectura clúster y el restante 14% (70) son MPP. Así mismo la Sunway TaihuLight – China (la más rápida del mundo) y el Titan (tercer más rápido del mundo) poseen arquitectura MPP. Así mismo la Tianhe-2 (MilkyWay-2) de China (segundo a nivel mundial) posee arquitectura clúster. Algo a tener en cuenta es que de los primeros 10 Supercomputadoras del Top500, sólo 3 de ellos tienen como arquitectura el esquema Cluster y las otras 7 son MPP.

Por otro lado para el caso peruana, las supercomputadoras del IAP, UNSA-Arequipa y del Instituto Geofísico del son las que poseen arquitectura clúster.

Gráfico N° 4: Porcentajes de Supercomputadoras del Top500 de noviembre del 2016, según arquitectura de computadora.

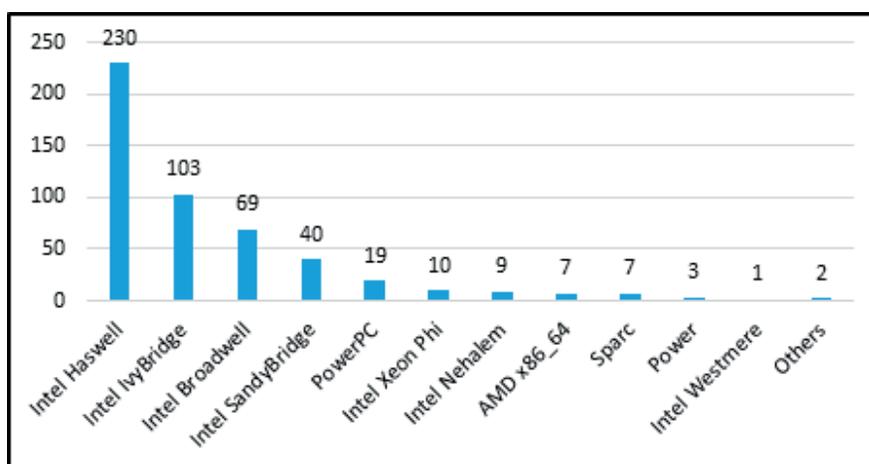


Fuente: Elaboración propia según estadísticas del Top500

4.1.6. Procesadores más utilizados en Supercomputadoras del Top500

Intel es una de las empresas líderes en la fabricación de procesadores de computo, incluso este liderazgo es muy evidente en el sector Supercomputación. Según el reporte del Top500 de noviembre del 2016, 462 (92.4%) supercomputadoras utilizan procesadores de las diversas familias Intel, siendo las Intel Haswell la más expandida en 230 (46% del total) supercomputadoras. Otros fabricantes que aparecen son el PowerPC, AMD, SPARC entre otras. Cabe comentar que en junio del 1993 la marca de procesador más expandido era el Cray con 193 (38.6%) y la marca Intel apenas llegaba a 65 (13%).

Gráfico N° 5: Cantidad de Supercomputadoras por familia de microprocesador del Top500 de noviembre del 2016.



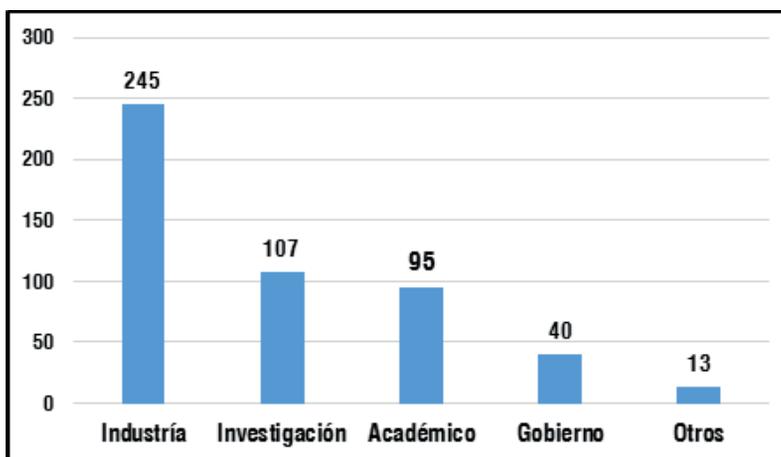
Fuente: Elaboración propia según estadísticas del Top500

4.1.7. Áreas que utilizan Supercomputadoras según el Top500

En los años 60, las supercomputadoras fueron creadas para apoyar los problemas de la ciencia siendo muy utilizadas en centros de investigación y universidades, hoy ambas áreas de trabajo han sido desplazadas. Según el reporte del Top500 de noviembre del 2016, el sector industrial es el sector que más utiliza supercomputadoras, siendo utilizadas para apoyar diversos procesos productivos, actualmente 49% (245 supercomputadoras) del Top500 son utilizados en el sector industrial. Este liderazgo del sector industrial en el ranking data de junio de 1997 y es impulsado por corporaciones globales dedicadas a fabricación de automóviles, aviones, exploración petrolera, de gas, producción de energía atómica, servicios informáticos, fabricación de computadoras, servicios financieros. Si bien es cierto la Supercomputación nació para apoyar diversos problemas de la ciencia, sin embargo por lo visto diversas corporaciones mundiales vienen sacándole provecho.

Las demás áreas de trabajo son Investigación 21.2% (107), Académico 19% (95), Gobierno 8% (40) y otros (para Benchmarking y clasificados). Vale indicar que el caso del uso de Supercomputadoras en el sector del gobierno, se viene registrando desde junio de 2006.

Gráfico N° 6: Cantidad de supercomputadoras por área de uso según Top500 de noviembre de 2016



Fuente: Elaboración propia según estadísticas del Top500

4.2. Supercomputación en Latinoamérica

La computación de alto rendimiento en Latinoamérica se encuentra considerablemente atrasada, muy lejos y desentonada de los avances de Norteamérica y algunos países asiáticos. Entre las enormes limitaciones de los países latinoamericanos está la inexistencia de producción propia de tecnologías para computación, pues no existe ningún productor de componentes electrónicos conectado a la cadena de valor de diversos dispositivos computacionales, a pesar de tener a países como Perú y Chile que figuran entre los productores mundiales de materia prima o metales que son utilizados en la cadena de valor de la computación.

El ecosistema empresarial latinoamericano aún es muy pequeño. No se alcanzaría cubrir la palma de la mano para contar los emprendimientos en esta área, sin embargo vale la pena citar algunos emprendimientos que con mayor apoyo podrían introducir a Latinoamérica como productor de equipos para supercomputación.

- **Plataformas de alto desempeño marca ABAX (Brasil):** Abax es una empresa brasileña con sede en Campinas (Sao Paulo), cuyo gerente es el experto Maicon Farias, que se dedica al ensamblaje, comercialización y soporte de plataformas de alto desempeño, y que vienen siendo utilizadas principalmente en entidades académicas brasileñas.
- **Supercomputadora portátil BIFE (Argentina):** BIFE, es un interesante proyecto de emprendimiento liderado por los jóvenes argentinos Demian Slobinsky y Javier Haboba. Es una supercomputadora de bajo consumo de energía, de pequeñas dimensiones, y que según ellos, puede lograr velocidades de 500 Pcs. Incluso este proyecto estuvo entre los ganadores del concurso Intel Global Challenge del 2011.
- **Supercomputadora para las Bolsas de Valores (México):** El experto Alberto Alonso de México en el junio del 2016, presentó una supercomputadora ensamblada en su propia casa con la ayuda de sus amigos y que está orientada a apoyar actividades de las bolsas de valores. Lo más desarrollado que tiene Latinoamérica respecto a la computación de alto desempeño se encuentra en los trabajos realizados por el sector académico y de investigación, expresado a

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

través de los centros de alto rendimiento o desempeño computacional. En este aspecto a nivel Latinoamérica Brasil y México son los más desarrollados. En el Top500 de noviembre del 2016 sólo aparecen 3 Supercomputadoras funcionando en Latinoamérica.

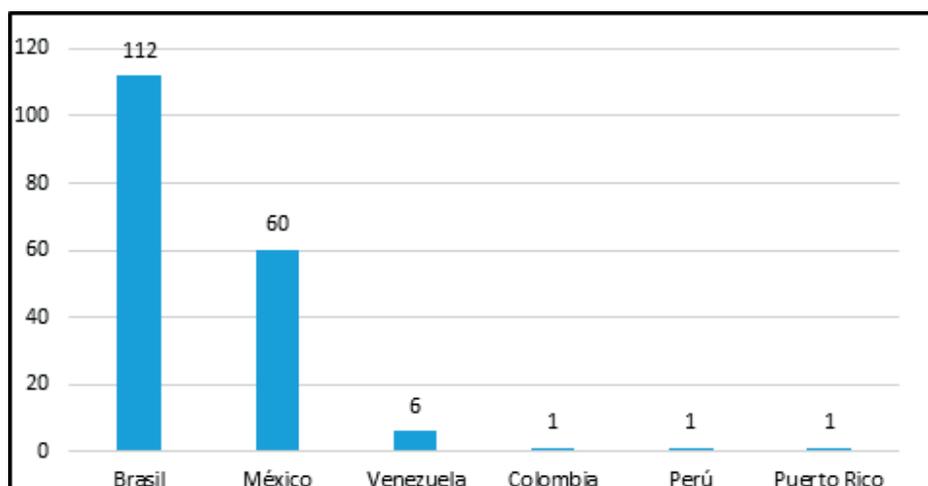
Tabla 24: Las 3 Supercomputadoras Latinoamericanas en el Top500 de noviembre de 2016.

Nro.	Top500	Supercomputadora	Ubicación	Fabricante	Arquitectura	Procesadores	TERAFLOPS (RMAX)
1	364	Santos Dumont GPU	Laboratório Nacional de Computação Científica	Bull, Atos Group	Clúster	Intel Xeon E5 - 2695v2 12C 2.4GHz	456.8
2	433	CIMATEC Yemoja	SENAI CIMATEC	HPE/SGI	Clúster	Intel Xeon E5-2690v2 10C 3GHz	405.365
3	476	Santos Dumont Hybrid	Laboratório Nacional de Computação Científica	Bull, Atos Group	Clúster	Intel Xeon E5-2695v2 12C 2.4GHz	363.225

Fuente: Top 500 - Elaboración propia.

Así mismo si se hiciera un ranking con los datos del número de veces de las apariciones de los países latinoamericanos en el Top500, tendríamos el enorme liderazgo de Brasil y México, ambos desde los inicios del ranking han invertido y adquirido equipamiento de gran capacidad. Los países que los acompañan son Venezuela (6 apariciones), Colombia (1 aparición en junio del 2004, con un equipo HP de Columbia Movil), Perú (1 aparición en junio del 2000, con un equipo Sun Microsystems HPC) y Puerto Rico (1 aparición en noviembre del 2001 con un equipo HP Superdome).

Gráfico N° 7: Cantidad de apariciones de países latinoamericanos en el Top500 desde junio de 1993 hasta noviembre del 2016.



Fuente: Elaboración propia según estadísticas del Top500

Así mismo la edición en que mejor le fue a Latinoamérica fue la de noviembre del 2001, donde aparecieron 8 Supercomputadoras, curiosamente liderados por 3 de Venezuela, 2 de Brasil, 2 de México y una de Puerto Rico. 7 de estos equipos eran la Supercomputadora SuperDome de HP.

Tabla 25: Supercomputadoras latinoamericanas en el Top500 en Noviembre de 2001.

Top500	Ubicación	Marca	Modelo	País	Sistema Operativo	GIGAFLOPS
153	Brasil Telecom	Hewlett-Packard	SuperDome/HyperPlex	Brasil	HP Unix (HP-UX)	281.6
169	Brasil Telecom	Hewlett-Packard	SuperDome/HyperPlex	Brasil	HP Unix (HP-UX)	281.6
335	Telcel Radiomovil Dipsa	Sun Microsystems	HPC 10000 400 MHz Cluster	México	Solaris	153.6
408	Alestra De LR	Hewlett-Packard	SuperDome	Venezuela	HP Unix (HP-UX)	140.8
412	Banco Popular	Hewlett-Packard	SuperDome	Puerto Rico	HP Unix (HP-UX)	140.8
413	Corporacion Digitel CA	Hewlett-Packard	SuperDome	Venezuela	HP Unix (HP-UX)	140.8
445	Telmex	Hewlett-Packard	SuperDome	México	HP Unix (HP-UX)	140.8
460	Cerveceria Polar	Hewlett-Packard	SuperDome/HyperPlex	Venezuela	HP Unix (HP-UX)	140.8

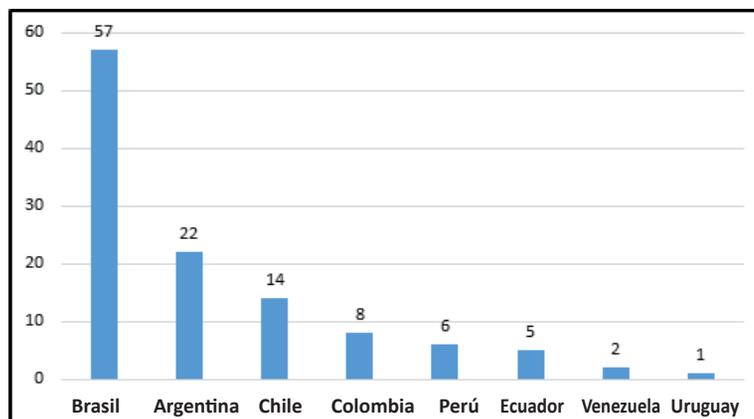
Fuente: Top500

4.3. Centros de alto rendimiento computacional en Latinoamérica

Los centros de alto rendimiento computacional son laboratorios de computo, data centers o centros de servicios que ofrecen servicios públicos y/o privados de procesamiento de grandes volúmenes de datos a través de equipamiento de computación paralela. Estos servicios pueden ser de diversa naturaleza y aplicaciones como es el caso de procesamiento numérico para investigación científica, big data y minería de datos, bioinformática, procesamiento de imágenes de teledetección, imágenes médicas, reconstrucción de escenarios, simulaciones gráficos para química computacional, etc.

En Latinoamérica son los principales espacios que vienen apoyando la expansión de la computación de alto desempeño. Se ha elaborado un listado de los centros de alto rendimiento de los países sudamericanos, habiéndose identificado 115 de ellos, encontrándose el liderazgo desbordado del Brasil y a la vez ningún centro procedente de Bolivia ni Paraguay.

Gráfico N° 8: Número de centros de alto rendimiento por países de Sudamérica en marzo de 2017.



Fuente: Elaboración propia según estadísticas del Top500.

Detrás de Brasil, destacan Argentina y Chile, un tanto más atrás Colombia y la misma Venezuela que incluso en el año 2000 logró tener 3 Supercomputadoras en el Top500. En el caso de Perú y Ecuador los centros de alto rendimiento computacional son recientes y dependen considerablemente del apoyo de entidades gubernamentales. Vale mencionar que entre los 7 centros de Perú ya se considera a la supercomputadora adquirida por la Universidad San Luis Gonzaga de Ica, cuyas autoridades anunciaron que brindará servicios en agosto de 2017. Puede ver el directorio completo en el Anexo 01.

a. Centro de Simulación Computacional para Aplicaciones Tecnológicas del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

Ubicación: Buenos Aires, Argentina.

El Centro de Simulación Computacional (CSC) es un instituto dedicado a buscar soluciones por iniciativa propia o a solicitud de otras instituciones públicas o privadas a problemas que puedan ser abordados con técnicas, procesos o metodologías derivadas de sus líneas de investigación.

CSC-CONICET hospeda a la máquina TUPAC, que es una máquina de 48 teraflops que cuenta con 4096 cores AMD Opteron y 16384 cores NVIDIA Cuda. Posee 72 terabytes de espacio para storage y 32 terabytes de espacio local en disco. La comunicación se establece mediante 5 redes independientes 2INFINIBAND QDR (muy baja latencia) y 2 redes Ethernet. Adicionalmente, la infraestructura de almacenamiento está implementada sobre una red independiente de fibra óptica montada sobre servidores en alta disponibilidad.

Modelado Avanzado: Modelado dentro de diferentes áreas que engloba la Mecánica Computacional. Modelos y desarrollos en computación de alto rendimiento. Paralelización bajo paradigmas MPI, OPENMP y GPU.

Ingeniería de la Información: Procesamiento de la información sobre redes inteligentes de potencia. Análisis y diseño de redes inalámbricas de detección distribuida. Diseño de RF y procesamiento de señales en radares ultrawideband. Técnicas distribuidas e interactivas de compresión de fuentes.

Generación de Energía: Se modelan diferentes procesos de generación de energía entre las que se destacan energías renovables (eólica y solar), energía producida por fisión y fusión nuclear y diferentes aspectos que tiene que ver con las obtenidas por combustibles fósiles.

Procesos en la Nanoescala (de angstroms a nanómetros): Analiza procesos de reactividad superficial, Fluidos confinados en matrices nanoporosas: propiedades estructurales, dinámicas y termodinámicas a través de Mecánica Cuántica (QM), Mecánica Clásica (MM), métodos híbridos (QM/MM)

Sitio web: <http://www.csc-conicet.gob.ar/>



Imagen 103: Interfaz del portal web del CSC - CONICET.

b. Laboratório Nacional de Computação Científica

Ubicación: Rio de Janeiro, Brasil

Creada en 1980, es un centro de investigación científica y desarrollo tecnológico perteneciente al Ministerio de Ciencia, Tecnología e Innovación (MCTI) de Brasil, especializada en la computación científica. Cuenta con la primera y tercera supercomputadoras más poderosas de Latinoamérica: Santos Dumont GPU y Santos Dumont Hybrid (según Top500 de noviembre del 2016).

Tiene como propósito llevar a cabo la investigación en el campo de la informática científica. Las líneas de investigación de laboratorio se fijan en áreas interdisciplinarias como biosistemas, la bioinformática, la biología computacional, la atmósfera y los océanos, medio ambiente y ciencia de múltiples escalas.



Imagen 104: Santos Dumont GPU I. Créditos: Laboratório Nacional de Computação Científica - LNCC

Las líneas de investigación que desarrolla son:

- **Modelado por computador:** Investigación y desarrollo de métodos para la comprensión y solución del problema en términos de la fenomenología, matemáticas e informática en áreas tan amplias como la ingeniería y las ciencias exactas, económicas, biológicas, humanidades y del medio ambiente, entre otros.
- **Métodos numéricos:** Comprende técnicas de desarrollo y análisis matemáticos y algoritmos numéricos para la simulación de fenómenos complejos.
- **Sistemas, control y señal:** Tiene como objetivo estudiar el comportamiento de sistemas dinámicos, con el objetivo de lograr ciertos estándares de referencias estatales o salida del sistema.
- **Computación:** Comprende retos y paradigmas que surgen en la informática en su conjunto y específicamente en la computación masivamente paralela y distribuida, computación cuántica, visualización científica y entornos de colaboración a virtual y la realidad aumentada, con el fin de afectar los modelos de investigación y desarrollo, métodos y algoritmos robustos y altamente eficientes, el punto de vista computacional.
- **Biología computacional:** Tiene como objetivo integrar el conocimiento de Ciencias de la Computación, Matemática Aplicada y Estadística con la biología para el desarrollo de modelos y herramientas computacionales para analizar los datos y fenómenos biológicos.
- **Petróleo, gas y agua:** En el cumplimiento de su misión, se ha centrado en unas pocas áreas estratégicas, incluidos los recursos hídricos, la calidad de las aguas subterráneas, el petróleo y el gas y el almacenamiento de dióxido de carbono.
- **Medicina asistida por computación científica:** Comprende el diagnóstico por imágenes que, acoplado al modelado y simulación, permite el desarrollo de nuevas técnicas terapéuticas en tiempo real para mejorar los procedimientos y tratamientos médicos:

Sitio web: <http://www.lncc.br/estrutura/default.php>

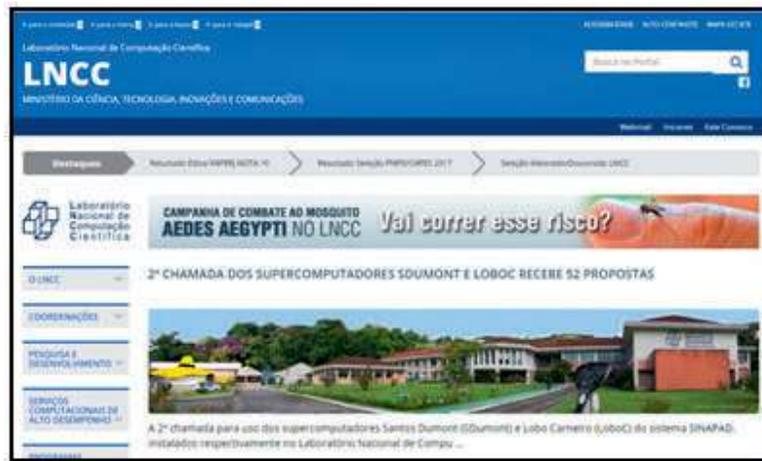


Imagen 105: Interfaz del portal web LNCC

c. CENAPAD-SP de la Universidad de Campinas

Ubicación: Campinas, Brasil

Creada en marzo de 1994, es uno de los nueve centros que componen el programa brasileño denominado SINAPAD implementado en Brasil por el Ministerio de Ciencia y Tecnología. Su misión es apoyar las actividades de investigación y desarrollo en las áreas de Ciencia y Tecnología del país, proporcionando un entorno de alto rendimiento y similares, que actúa como un centro de excelencia en las nuevas tendencias de la computación, lo que contribuye a la formación de la comunidad científica y empresarial.

Ofrece formación especializada en ciencias de la computación de alto rendimiento, abierta a la comunidad científica y empresarial.

Cuenta con 4 plataformas: SGI Altix PBS, SGI ICE PBS, IBM P750 LoadL, GPU / Phi Torque.



Imagen 106: Reunión de los autores en las instalaciones de CENAPAD-UNICAMP en Campinas

Se trabajan proyectos en las líneas siguientes:

- Agricultura.
- Biofísica.
- Biología.
- Computación.
- Ingeniería.
- Estadística.
- Farmacia.
- Física.
- Ciencias de la Tierra.
- Medicina.
- Meteorología.
- Química.
- Zootecnia.

Sitio web: <https://www.cenapad.unicamp.br/>



Imagen 107: Interfaz del portal web de CENAPAD

d. Universidade Federal de ABC

Ubicación: Sao Paulo, Brasil.

Esta universidad fue creada en el año 2005, a pesar de ser una universidad nueva viene alcanzando interesantes indicadores académicos a nivel de las universidades brasileñas. Entre las fortalezas de su infraestructura de investigación destacan hasta 3 plataformas de alto desempeño:

- Una Supercomputadora Altix 4700 de la marca SGI (actual HPE/SGI), con gran capacidad de memoria compartida.
- El Cluster Chrome UFABC.
- Bull Nova Scale Cluster UFABC.

Sitio web: <http://www.ufabc.edu.br/>



Imagen 108: Laboratorio de supercomputación de la Universidad Federal ABC.

Los equipos dan soporte a investigaciones relacionadas a:

- Neurobiología del comportamiento.
- Aplicaciones en tiempo real.
- Física matemática.
- Metodologías y technical computing.
- Computación gráfica.
- Informática Aplicada / Intera.
- Formación para enseñanza de la matemática.
- Biología computacional y bioinformática.
- Dinámica no lineal.
- Computación paralela y distribuida.
- Computación científica.
- Sistemas de software y de educación a distancia.



Imagen 109: Interfaz del portal web de la Universal Federal de ABC – Sao Paulo

e. Laboratorio Nacional de Computación de Alto Rendimiento (NLHPC) de la Universidad de Chile

Ubicación: Santiago de Chile, Chile.

Establecida el año 2010, el NLHPC (del inglés National Laboratory of High Performance Computing) es una iniciativa que tiene por objetivo instalar en Chile una red de infraestructura de computación de alto rendimiento, para el procesamiento de datos científicos y sus aplicaciones industriales.

Este proyecto es impulsado a través el Primer Concurso de Centros de Servicios de Equipamiento Mayor de uso Compartido del Programa de Investigación Asociativa (PIA) de CONICYT y liderado por el Centro de Modelamiento Matemático (CMM) de la Universidad de Chile en una modalidad colaborativa con Red Universitaria Nacional (REUNA) y otras seis universidades, lo cual permite que los principales centros científicos y académicos del país accedan a los nodos de súper-cómputo que instalará el proyecto, a través de una red de conectividad de alta capacidad.

El NLHPC está formado por la Universidad de Chile (Institución Patrocinadora), las universidades Pontificia Universidad Católica de Chile (PUC), Pontificia Universidad Católica de Valparaíso (PUCV), Universidad Austral de Chile (UACH), Universidad Católica del Norte (UCN), Universidad Católica del Maule (UCM), Universidad de Antofagasta (UA), Universidad del Bío Bío (UBB), Universidad de Concepción (UDEC), Universidad de la Frontera (UFRO), Universidad de La Serena (ULS), Universidad de Santiago (USACH), Universidad de Talca (UTalca), y la Universidad Técnica Federico Santa María (UTFSM), en asociación con Aura observatory en Chile, Centro Interdisciplinario de Neurociencia de Valparaíso (CINV), Fundación Ciencia & Vida (FCV), INRIA Chile y REUNA como instituciones asociadas (AI).

Este centro cuenta con la supercomputadora de mayor capacidad al que han denominado “Leftrarú” cuyo significado es “Halcon veloz” en mapuche.

Sus capacidades son las siguientes:

- 132 Nodos de cómputo HP (128 nodos HP SL230 y 4 nodos HP SL250), cada uno con procesadores de 10 cores Intel Xeon Ivy Bridge E5-2660 V2.
- 2640 cores (núcleos).
- 12 Co-procesadores Intel Xeon Phi 5110p de 2 teraflops.
- 6.25 terabytes de RAM
- Conexión Infiniband FDR a 56Gbps.
- 274TB de almacenamiento Lustre (DDN EXAScaler).
- Capacidad de cómputo: 70 teraflops.

Entre las principales líneas de investigación trabajadas destacan:

- **Biomedicina:** Análisis computarizado de imágenes, genómica, bioinformática en el contexto de cáncer y enfermedades neurodegenerativas.
- **Ingeniería y Diseño Industrial:** Simulaciones para resistencia de materiales.
- **Astrofísica y Astronomía:** Desarrollo de trabajos relacionados a agujeros negros, estudios sobre la gravedad.
- **Otras líneas:** Procesamiento de señales en tiempo real, minería de datos, meteorología, física de estado sólido, entre otros.

Sitio web: <http://ingenieria.uchile.cl/>



Imagen 110: Laboratorio Nacional de Computación de Alto Rendimiento de Chile



Imagen 111: Interfaz del portal web del Laboratorio Nacional de computación de Alto Rendimiento.

f. Centro de Bioinformática y Biología Computacional (CBBC)

Ubicación: Manizales, Colombia.

Inició sus actividades el año 2013. Es un centro especializado en bioinformática, líder en supercomputación dedicado a la prestación de servicios al gobierno, la academia y la industria interesados en la investigación y desarrollo de la biotecnología y la bioprospección de los recursos de la gran biodiversidad del país; es una corporación mixta, sin ánimo de lucro, que ofrece infraestructura de supercomputación robusta y personal altamente capacitado para desarrollar actividades de investigación e innovación.

Busca proporcionar herramientas computacionales en relación con las ciencias de la vida;

desarrollar software, almacenar y procesar datos suministrados por sus investigadores, por miembros de otros grupos de investigación, o por terceros; adelantar actividades académicas y educativas, asesorar y orientar investigaciones intensivas en el uso de datos y su interpretación y participar en el desarrollo de investigaciones, que requieran el uso de sus facilidades de procesamiento de datos en otras áreas del conocimiento, de importancia para el desarrollo de la ciencia, la tecnología y la innovación.

Este centro surge de la convergencia de esfuerzos de una serie de instituciones colombianas como el Ministerio de Tecnologías de la Información y las Comunicaciones- MINTIC, la empresa transnacional Microsoft, Colciencias, en asociación a entidades académicas como la Universidad Autónoma de Manizales, Universidad Católica de Manizales, Universidad de Caldas, Universidad de Manizales, Universidad del Quindío, Universidad Tecnológica de Pereira, Universidad del Tolima, etc.

En sus líneas de trabajo destacan:

- **Bioinformática y bioprospección:** Genómica, metabolómica, metagenómica, proteómica, Transcriptómica, etc.
- **Bioingeniería.**



Imagen 112: Clúster del Centro de Bioinformática y Biología Computacional

Capacidades de cómputo

- Almacenamiento: 274 TB.
- Núcleos CPU: 312.
- Núcleos GPU: 9984.
- 5 Tarjetas Xeon Phi 5110P.
- Memoria RAM: 4.8 TB.
- Infiniband.

Sitio web: <http://bios.org.com>



Imagen 113: Interfaz del portal web del Centro de Bioinformática y Biología Computacional

g. Universidad de Investigación de Tecnología Experimental Yachay

Ubicación: Imbabura, Ecuador.

Esta universidad fue creada en el año 2013, es parte de un auspicioso proyecto del gobierno ecuatoriano de crear una ciudad tecnológica denominada Yachay, en Ecuador (la primera ciudad planificada de Ecuador, creada para convertirse en un polo de desarrollo científico y tecnológico del país y de Sudamérica).

Entre sus laboratorios cuenta con equipamiento de alto rendimiento computacional según sus autoridades que alcanzan 350 teraflops. Se tiene planificado que ésta universidad albergará hasta 3 supercomputadoras.



Imagen 114: Equipamiento de la Supercomputadora de Yachay (Ecuador)

Según una noticia publicada en el propio sitio web de Yachay [75] La solución de computación de alta performance de IBM (High Performance Computing – HPC) está basada en IBM Power Systems S822LC, e incorpora innovaciones del ecosistema de OpenPOWER Foundation, incluyendo NVIDIA y MELLANOX. Tiene 84 nodos S822LC y 4 nodos de gestión, además de Mellanox Infiniband Fabric (8828-ED0), IBM Rackswitch fabricado por Lenovo, IBM Storwize V7000 y Storwize V5000, IBM Spectrum Compute Suite, IBM Spectrum Scale, IBM Spectrum Protect, IBM Spectrum Virtualize, IBM Spectrum Control, IBM Protectier además de servicios de nuestras áreas de Servicios Tecnológicos y Mantenimiento.

Entre las líneas de investigación priorizadas destacan: Ciencias de la vida, nanociencias, tecnologías de la información y comunicación, petroquímica y energía. Así como temas cruciales relacionados a cambio climático a través del estudio de los vientos, el agua, biodiversidad, etc.

Sitio web: <http://www.yachaytech.edu.ec/home/>



Imagen 115: Interfaz del portal web de la Universidad de Investigación de Tecnología Experimental Yachay

h. Clúster FING de la Universidad de la República

Ubicación: Montevideo, Uruguay.

Inició sus actividades el año 2008. Es una infraestructura de cómputo de alto desempeño adquirida por la Facultad de Ingeniería de la Universidad de la República de Uruguay. Su principal objetivo consiste en proveer soporte para la resolución de problemas complejos que demanden un gran poder de cómputo. El cluster FING fue adquirido con fondos del llamado de Fortalecimiento de Equipamientos para la Investigación de la Comisión Sectorial de Investigación Científica.

Sirve de soporte al Núcleo Interdisciplinario de Computación Científica de Alto Desempeño (NICCAD) que está integrado por investigadores interesados en las temáticas de métodos matemáticos y computacionales aplicados a la resolución de problemas científicos y de computación de alto desempeño. Está organizado de acuerdo a los siguientes grupos de trabajo:

- Centro de Cálculo.
- Computational chemistry and biology group.

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

- Física Computacional.
- Geofísica Planetaria.
- Grupo de Mecánica de Sólidos Computacional.
- Grupo de Simulaciones Biomoleculares.
- Grupo de Trabajo en Energías Renovables.
- Ingeniería del Viento.
- Investigación Operativa.
- Laboratorio de Biología de Sistemas.
- Laboratorio de Probabilidad y Estadística (LPE).
- Laboratorio de Química Teórica y Computacional.
- Mecánica de los Fluidos Computacional.
- MINA (Network Management - Artificial Intelligence).
- Modelación numérica del Río de la Plata y el Frente Marítimo.
- Simulación numérica de convertidores energéticos.
- Sistemas de Energía Eléctrica.

Capacidades de Computo

tipo	nodos	modelo	#	procesador	proc/ nodo	cores/ proc	cores / nodo	cores / tipo	mem. / nodo	mem. total	disco / nodo
head	node0	Dell Power Edge 2950	1	Intel Xeon E5430 2.66GHz	2	4	8	8	8GB	8GB	80GB
class0	node02-09	Dell Power Edge 2950	8	Intel Xeon E5430 2.66GHz	2	4	8	64	8GB	64GB	80GB
class1	node20-23	HP Proliant DL180 G6	4	Intel Xeon E5520 2.26GHz	2	4	8	32	8GB	32GB	-
class2	node30-32	HP Proliant DL385 G7	3	AMD Opteron 6172 2.10GHz	2	12	24	72	24GB	72GB	-
class3	node33-38	HP Proliant DL385 G7	5	AMD Opteron 6172 2.10GHz	2	12	24	120	72GB	360GB	250GB
class4	node40	HP Proliant DL585	1	AMD Opteron 6272 2.09GHz	4	16	64	64	48GB	48GB	-
class5	node50-53	DELL Power Edge R720	2	Intel Xeon E52650 2.00GHz	2	8	16	64	64GB	256GB	600GB
class6	node60-61	HP Proliant DL385p G8	2	AMD Opteron Processor 6376 2.30GHz	2	32	1	64	32GB	64GB	1TB
class7	node70-71	HP Proliant DL380 G9	2	Intel Xeon CPU E5-2680 v3 2.50GHz	24	1	24	48	128GB	256GB	1TB
xeon phi	node50- 51-phi	-	2	Xeon Phi 31S1P	1	1	57	114	8GB	16iGB	-
tesla	tesla1	-	1	Intel Xeon E5530 2.40GHz	1	16	16	16	8GB	8GB	-
Totales			31					609		848	

Imagen 116: capacidad de Computo del Clúster FING de la Universidad de la República, Montevideo, Uruguay.

Sitio web: <https://www.fing.edu.uy/cluster/index.php>

Cluster FING

El clúster FING es una infraestructura de cómputo de alto desempeño perteneciente a la Facultad de Ingeniería. Su principal objetivo consiste en proveer soporte para la resolución de problemas complejos que demandan un gran poder de cómputo. El clúster FING fue adquirido con fondos del Ramado de Fortalecimiento de Equipamientos para la Investigación de la Comisión Sectorial de Investigación Científica (2003).

Si utiliza la plataforma en su trabajo por favor reconocerlo con un texto similar al siguiente: "Los experimentos computacionales fueron desarrollados en el Clúster FING, la plataforma de Computación Científica de alto Desempeño de la Universidad de la República, Uruguay (Nesmachnov2010). Detalles disponibles en www.fing.edu.uy/cluster/". La referencia bibliográfica correspondiente es: "S. Nesmachnov. Computación científica de alto desempeño en la Facultad de Ingeniería, Universidad de la República. Revista de la Asociación de Ingenieros del Uruguay 61 (1), 12-15, 2016."

Cumplimos 8.000.000 de horas de cómputo
El proyecto cluster FING cumplió 8.000.000 de horas de cómputo en settembre de 2016. Nuestros agradecimientos a todos los que colaboraron para alcanzar este hito.
Total de horas de cómputo al día de hoy:

Investigación
Núcleo Interdisciplinario de Computación Científica de Alto Desempeño (NICCAD)

Infraestructura

Imagen 117: Interfaz del Portal web del Clúster FING.

i. Centro Nacional de Cálculo Científico - CeCalCULA de la Universidad de Los Andes

Ubicación: Mérida, Venezuela

Es el primer centro de supercomputación de Venezuela. CeCalCULA es un proyecto surgido de la asociación entre la Universidad de los Andes y el Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICIT) de Venezuela. Y fue desarrollado dentro de la Corporación Parque Tecnológico de Mérida (CPTM). Permite la afiliación de institutos, centros, grupos de I&D e investigadores independientes venezolanos y facilita el acceso a fondos de cooperación internacional.

Así mismo este centro desarrolla servicios y entrenamiento de personal en el área de computación de alto rendimiento y provee a la comunidad académica e industrial de herramientas y técnicas computacionales altamente competitivas.

Áreas de Trabajo e Investigación:

- Bioinformática.
- Bioingeniería e Imagenología Médica.
- Computación de Alto Rendimiento.
- Estadísticas y Software Matemático.
- Física Computacional.
- Geociencias Computacionales.
- Información y Colaboración Electrónica.
- Ingeniería Computacional.
- Química Computacional.
- Visualización y Arquitectura Computacional.

Sitio web: www.ula.ve/



Imagen 118: Interfaz del portal web de la Universidad de los Andes de Venezuela.

4.4. Expertos Referentes de Computación de alto rendimiento

• Seymour Cray

Lugar y Fecha de Nacimiento: Chippewa Falls, Wisconsin (EE.UU.), nació el 28 de septiembre de 1925, falleció el 6 de octubre de 1996.

Principal aporte: Creador de la primera supercomputadora del mundo, el equipo modelo CDC 7600 con la empresa Control Data Corporation. Incluso es reconocido como padre de la supercomputadora.

Fuente: <https://www.quotationof.com/>

Resumen de su biografía:

Desde muy joven manifestó su pasión por la ciencia. Se unió al ejército de su país en 1943 al pelotón de comunicaciones de infantería, incluso participó de la Segunda Guerra Mundial. Durante los años 50 fue parte del equipo que diseñó el ordenador UNIVAC 1103.

En 1957, junto con otros ingenieros fundó la compañía denominada Control Data Corporation, en abreviatura CDC, para la cual construyó el CDC 1604, que fue uno de las primeras computadoras que utilizó transistores en lugar de tubos de vacío.



Créditos: Michael Hicks

Cray siempre buscaba crear las computadoras más rápidas del mundo. A finales de los 60s, Control Data, lanzó el CDC 7600, la primera supercomputadora. En el año 1972 fundó Cray Research, para fabricar y vender exclusivamente supercomputadoras, y lo hacía de uno en uno, a pedido. Su primer producto fue el CRAY-1, que fue pedido por el Laboratorio Nacional Los Álamos. Consolidó liderazgo hasta los 80s (llegó a tener el 70% del mercado mundial), luego se retiró de la empresa por problemas financieros, y en 1996 fundó llamada SRC Computers. Poco después luego de un accidente automovilístico falleció el 6 de octubre de 1996.

• Gordon Moore

Fecha de Nacimiento: 3 de enero de 1929.

Lugar de Nacimiento: San Francisco (California)

Principales aportes: Cofundador de la Empresa Intel, principal fabricante actual de procesadores en el mundo no solo para computación de alto rendimiento electrónicos. Además es autor de la Ley de Moore.



Créditos: Chemical Heritage Foundation

Resumen de su biografía

Graduado en química por la Universidad de California en Berkeley en 1950 y un doctorado en química y física en el Instituto de Tecnología de California (Caltech) en 1954. Previo a sus estudios en Berkeley, pasó sus años de bachiller en la Universidad Estatal de San José donde conoció a su esposa Betty.

sino para otros dispositivos

El 18 de julio de 1968 junto a Robert Noyce fundo la empresa Intel, cuyo nombre completo es Integrated Electronics Corporation (fue vicepresidente ejecutivo hasta 1975). Desde 1975 hasta 1987 fue Presidente Ejecutivo.

Es el autor de una ley predictiva sobre los procesadores denominada “Ley de Moore” indica que aproximadamente cada 18 meses se duplica el número de transistores en un circuito integrado. Gordon E. Moore la publicó el 19 de abril de 1965, cuyo cumplimiento se ha podido constatar hasta hoy. Sin los procesadores Intel la Supercomputación no se hubiera desarrollado y llegado hasta donde llega hoy [76].

• Jen-Hsun Huang

Fecha de nacimiento: 17 de febrero de 1963.

Lugar de nacimiento: Taipei (Taiwán)

Principales aportes: Su empresa NVIDIA el año 2009 creó las GPU (aceleradores para procesamiento gráficos). Las GPU vienen revolucionando la computación de alto rendimiento.



Créditos: Nvidia corporation

Resumen de su biografía

Graduado en Ingeniería Eléctrica en la Universidad Estatal de Oregón en 1984, su maestría la realizó en la Universidad de Stanford en 1992. Trabajó como Director de Coreware en LSI Logic, luego como diseñador de microprocesadores en AMD. En 1993, Jen Hsun Huang AMD fundó la empresa Nvidia, donde actualmente es CEO y director presidente de la compañía.

El año 1999 Nvidia creo las GPU (Unidad de Procesamiento Gráfico, que es un coprocesador dedicado al procesamiento de gráficos u operaciones de coma flotante), con lo cual sistemas

informáticos están pasando de realizar el “procesamiento central” en la CPU a realizar “coprocesamiento” repartido entre la CPU y la GPU. Desde entonces, ha establecido nuevos estándares en computación visual con impresionantes gráficos interactivos disponibles en dispositivos que van desde teléfonos inteligentes y tabletas hasta computadoras portátiles, estaciones de trabajo, computadoras de alto desempeño.

El año 2006, Nvidia lanza CUDA, que es una arquitectura de cálculo paralelo que aprovecha la gran potencia de la GPU para proporcionar un incremento extraordinario del rendimiento del sistema. El modelo de programación de CUDA está diseñado para que se creen aplicaciones que de forma transparente escalen su paralelismo para poder incrementar el número de núcleos computacionales. CUDA intenta explotar las ventajas de las GPU frente a las CPU utilizando el paralelismo que ofrecen sus múltiples núcleos, que a la vez permiten el lanzamiento de un altísimo número de hilos simultáneos.

• Mateo Valero

Fecha de nacimiento: 6 de agosto de 1952.

Lugar de nacimiento: Zaragoza(España).

Principales aportes: Impulsó y promovió los centros de alto rendimiento computacional en países iberoamericanos. Así mismo construyó en España un centro de investigación de Arquitectura de Computadores de talla mundial, y es pionero en nuevos enfoques básicos para el paralelismo a nivel de instrucción.



Créditos: Gemma ribas maspoch

Resumen de su biografía

Ingeniero de Telecomunicación por la Universidad Politécnica de Madrid en 1974 y Doctor. Ingeniero de Telecomunicación por la Universidad Politécnica de Catalunya en 1980. Es profesor de la UPC desde el año 1974 y Catedrático del Departamento de Arquitectura de Computadores desde 1983. Ha sido director del Departamento y Decano de la Facultad de Informática. Así mismo sus temas de Docencia e Investigación se centran en el campo de la Arquitectura de las Computadoras, con especial énfasis en algunos aspectos de las Supercomputadoras y Computadoras de Altas Prestaciones.

Fue el primer investigador hispanohablante que recibió el premio de supercomputación Seymour Cray, el mayor reconocimiento internacional en computación de alto rendimiento otorgado por la IEEE Computer Society (considerado el Premio Nobel de la Supercomputación).

Desde su creación en el 2004 es el director del Barcelona Supercomputing Center – Centro Nacional de Supercomputación Español (BSC - CNS) que alberga la supercomputadora MareNostrum www.bsc.es. Ha publicado más de 600 artículos sobre computación de alto rendimiento.

- **Jorge Titinger**

Lugar y Fecha de Nacimiento: Lima, Perú.

Principales aportes: Hasta fines del 2016 fue CEO de la empresa norteamericana SGI (nombre completo de Silicon Graphics International), que es un fabricante especializado en computación de alto rendimiento (uno de los pocos latinos con este tipo de posición). SGI es una de las marcas más expandidas en Latinoamérica, incluso en el Perú 4 de las 6 plataformas de alto rendimiento registradas son de la marca SGI.



Créditos: events.bizzabo.com

Realizó su licenciatura y maestría en ingeniería eléctrica y una maestría en administración en Stanford. Llegó a Stanford becado como jugador de fútbol (incluso jugó en el equipo nacional de interior de los EEUU a partir de 1988 a 1993).

Pasó 10 años trabajando en Hewlett-Packard Company, MIPS Computer Systems. Desde 27 de febrero de 2012 fue presidente y director ejecutivo de SGI.-Silicon Graphics International Corp. hasta fines del 2016, ya que esta empresa fue adquirida por Hewlett-Packard.

Recibió el Premio People (Silicon Valley Latino) Choice Award en la categoría de categoría de Trailblazer Latino.

CAPÍTULO 5

LA SUPERCOMPUTACIÓN EN EL PERU

Empezamos este capítulo indicando que la computación de alto rendimiento en el Perú aún está en pañales y sostenemos ello por el débil y reducido ecosistema institucional existente. A la fecha, en el Perú estas tecnologías no han sido difundidas de forma apropiada, lo que ha generado que nuestras instituciones y la mayoría de profesionales desconozcan su uso. Así mismo, en el Perú no existe una carrera, especialidad o postgrado que se ocupe de netamente de la supercomputación. No obstante, es válido destacar el rol de algunas universidades que impulsan diversas escuelas de Ciencias de la Computación que cubren ciertos aspectos de la supercomputación.

Consideramos que los principales factores que actualmente afectan negativamente el desarrollo de la supercomputación son los siguientes:

- Escasa oferta de capacitación sobre el procesamiento de grandes volúmenes de datos en los planes curriculares de las carreras o disciplinas científicas (biología, química, medio ambiente, forestal, etc). Esto genera desconocimiento de técnicas y métodos modernos de investigación en diversas ciencias.
- Reducido número de profesionales de informática capacitados en computación de alto rendimiento. A pesar que en Perú existe más de un centenar de universidades con escuelas de informática, sistemas o computación y más de 500 centros de educación técnica, el sistema educativo peruano no forma profesionales con capacidades para desarrollar proyectos o utilizar supercomputación. Los pocos profesionales residentes en Perú que manejan correctamente el tema, recibieron mayormente formación de posgrado en el extranjero. En los pocos casos encontrados en algunas universidades destacan asignaturas como procesamiento paralelo, uso de supercomputadores, computación científica, diseño y gestión de centros de informática avanzada, etc.
- Escasa presencia de empresas especializadas en comercialización, adaptación y soporte de equipamiento para computación de alto rendimiento. Este factor genera incertidumbre a la hora de elaborar proyectos de adquisición de este tipo de equipamiento. Ello está articulado con la nula presencia en Perú de marcas especializadas en este tipo de computadoras (como Cray, Fujitsu, IBM, etc.).
- Mínima inversión del sector estatal en proyectos de infraestructura computacional de alto rendimiento (incluye capacitación, uso del software, redes de alta velocidad). En este punto podemos destacar la inversión promovida por CONCYTEC, a través de CIENCIACTIVA y el mismo INNOVATE, que han facilitado fondos para la adquisición de equipamiento en el Instituto Geofísico del Perú, Instituto de Investigaciones de la Amazonía Peruana y la Universidad Nacional de San Agustín.
- Mínima inversión del sector empresarial-industrial en proyectos de infraestructura computacional de alto rendimiento. Actualmente, en el Top500, existen más supercomputadores en las corporaciones mundiales que en los centros de investigación. Lastimosamente en el sector empresarial peruano la presencia es nula, pues muy pocas empresas peruanas investigan, desarrollan productos o buscan

innovación (el Perú es básicamente un país proveedor de materia prima, prácticamente último en innovación a nivel mundial).

- Reducido número de laboratorios o centros de alto rendimiento computacional en el Perú. Esto coincide con la escasa inversión económica, escasos proyectos con necesidades de procesamiento e incluso por las equivocadas prioridades de las instituciones peruanas. Algo parecido pasa en la mayoría de países sudamericanos.

- Ausencia de proyectos de investigación y servicios de información de cobertura nacional que respondan a diversas problemáticas del país. En el Perú aún no se han priorizado la construcción de servicios de información que integren el territorio y que den respuestas a diversos problemas ambientales, riesgos de desastres, deforestación de bosques, recursos hídricos, problemas sociales, etc. No obstante, cada vez surgen mayores necesidades para medir fenómenos que afectan al ambiente y monitorear sus efectos. Por ejemplo, para el caso del niño costero es urgente la expansión y apropiada cobertura de redes de estaciones meteorológicas y de sensores (imágenes satelitales, radares y de drones) que produzcan los datos suficientes para prevención de fenómenos climáticos extremos.

- En el caso de entidades públicas, el sistema de adquisiciones con fondos públicos es un factor limitante, considerando la cantidad de trámites, el tiempo de cada proceso, tiempos de importación, las condiciones y los requisitos que deben cumplir las empresas proveedoras. Esto encarece los costos y desanima o limita a todos los actores del ecosistema. En este punto es necesario destacar los esfuerzos del CONCYTEC, que logró simplificar los procesos administrativos sin quebrar por ello el proceso de control e idoneidad de los equipos.

- Escasa producción científica e innovación peruana (a menor producción científica y/o innovación, menor necesidad de procesamiento).

Para hablar de la Supercomputación en el Perú hemos considerado relevante hablar de los centros de alto rendimiento existentes, hacer un breve análisis de la situación de los centros de educación superior y finalmente presentar información de profesionales y expertos que poseen trabajos y están utilizando computación de alto desempeño.

5.1. Centros de Alto Rendimiento Computacional en Perú

Estos centros son laboratorios de cómputo, data centers o centros de servicios que ofrecen servicios públicos y/o privados de procesamiento de grandes volúmenes de datos, a través de equipamiento de computación paralela. Estos servicios pueden ser de diversa naturaleza y aplicaciones, como es el caso de procesamiento numérico para investigación científica para el cambio climático, big data y minería de datos, bioinformática, procesamiento de imágenes de teledetección, imágenes médicas, reconstrucción de escenarios, simulaciones gráficas para química computacional, etc. En Latinoamérica son los principales espacios que vienen apoyando la expansión de la computación de alto desempeño. En el Perú encontramos registro de la existencia de hasta 7 centros de alto rendimiento computacional, 4 en Lima, 1 en Arequipa, 1 en Iquitos y el último en Ica (4 de ellos están ubicados en instituciones relacionadas al Ministerio del Ambiente del Perú, por lo que se podría decir que en Perú ya existe una preponderancia a los aspectos ambientales y climáticos).

Seguidamente hacemos una pequeña presentación:

Laboratorio de Dinámica de Fluidos Geofísicos Computacional del Instituto Geofísico del Perú (IGP)

Sede: Lima

Tiene como objetivo estudiar el comportamiento complejo de los fluidos (e.g., el océano, la atmósfera) desarrollados a diferentes escalas de movimiento, en diferentes sistemas y bajo diferentes formas, los cuales se rigen por una dinámica similar (i.e. los principios fundamentales de la dinámica de fluidos y la termodinámica).

Cuenta con el sistema computacional de alto rendimiento de última generación (HPC-Linux-Clúster), dedicado a suplir la “necesidad computacional” (capacidad de cálculo o procesamiento) de las actuales investigaciones científicas de alto nivel que se desarrollan en el Instituto Geofísico del Perú, las cuales están dedicadas a resolver y analizar los fenómenos geofísicos (naturales y/o de origen antropogénico) que afectan a la población peruana y sus actividades.



Imagen 119: Equipo del Laboratorio de Dinámica de Fluidos Geofísicos Computacional del Instituto Geofísico del Perú (IGP) con el equipo del Centro de Alto Rendimiento Computacional de la Amazonía Peruana (IIAP) en la sede del IGP en febrero del 2015.

Facilita acceso a universidades, instituciones y centros de investigación que requieran (siguiendo el protocolo establecido) la capacidad de cómputo necesaria, a fin de fortalecer las investigaciones científicas y el desarrollo tecnológico en beneficio del Perú.

El Sistema Clúster HPC está compuesto por dos Clúster:

Clúster El Niño: Se compone de 6 nodos que contienen un total de 72 núcleos CPU. Este equipo es marca DELL.

Clúster Cinnamon: Se compone de 20 nodos que contienen un total de 480 núcleos CPU (Capacidad teórica pico de 20.29 teraflops). Este equipo es marca DELL.

Contacto: Responsable de la Unidad de Oceanografía: Dra. Ivonne Montes, e-mails: dfgc@igp.gob.pe, imontes@igp.gob.pe, Telef. : (511) 317-2300 Anexo 135

Sitio web: www.igp.gob.pe



Imagen 120: Interfaz del sitio web del Laboratorio de Dinámica de Fluidos Geofísicos Computacional del Instituto Geofísico del Perú (IGP)

NASA LASER TRACKING STATION AREQUIPA de la Universidad Nacional San Agustín de Arequipa (UNSA)

Sede: Characato, Arequipa

Este es uno de los pocos observatorios de rastreo satelital que la NASA tiene en Sudamérica, su denominación oficial es Instituto Astronómico Aeroespacial Pedro Paulet - IAAPP de la UNSA y está ubicado en la ciudad de Characato.

Funciona bajo un convenio con UNSA, y tiene como objetivo el rastreo de Satélites por Laser para la determinación de Órbitas y de Posicionamiento. Así mismo, realiza una serie de proyectos de investigación relacionados a procesamiento de imágenes Satelitales con el Clúster y datos de GNSS, para lo cual tiene instalado un supercomputador adquirido por la UNSA el año 2015 con apoyo de Innóvate Perú.

Este clúster de supercomputación se compone de 36 nodos que contienen un total de 432 núcleos CPU, administrado conjuntamente con el Departamento de Ingeniería de Sistemas e Informática de la UNSA. El clúster además facilita el acceso de investigadores y estudiantes de la UNSA y pronto a otras universidades o instituciones públicas y privadas, para uso en diversas temáticas tales como Minería de Datos, Deep Learning, Big data, Simulaciones con Modelos y otros.

Contacto: Dr. Raúl Yanyachi, email: raulpab@unsa.edu.pe

Sitio web: www.unsa.edu.pe



Imagen 121: Dr. Raúl Yanyachi responsable del Observatorio de Characato UNSA, recibiendo al Mg. Isaac Ocampo Yahuarcani, coordinador de Centro de Alto Rendimiento Computacional de la Amazonía Peruana



Imagen 122: Clúster de UNSA en pleno funcionamiento

Centro de Alto Rendimiento Computacional de la Amazonía Peruana del Instituto de Investigaciones de la Amazonía Peruana (IIAP)

Sede: Iquitos

Es un laboratorio informático diseñado para apoyar investigaciones científicas y tecnológicas que tengan necesidades de procesamiento de grandes volúmenes de información. Fue instalada en la ciudad de Iquitos (Perú) a principios del 2017.

Cuenta con un clúster compuesto por 9 nodos: 3 nodos de procesamiento numérico (un total de 280 núcleos de procesador de gran capacidad) y por 6 nodos de procesamiento gráfico (tarjetas NVIDIA con un total de 29952 núcleos CUDA Core). Fue bautizado como MANATI.

Aplicaciones para la Amazonía Peruana: Los principales usos estos orientados a las investigaciones sobre deforestación, conservación de la biodiversidad, recursos hídricos, cambio climático, ordenamiento territorial, poblaciones indígenas, bioinformática y otros estudios.

Servicios

- Procesamiento de información numérica (ejecución de aplicaciones, algoritmos, modelos matemáticos y paralelización, etc.)
- Procesamiento gráfico (pre-procesamiento y procesamiento de imágenes espaciales).
- Capacitación en computación de alto rendimiento.

Contactos:

- Dr. Luis Campos Baca. Presidente del IIAP, presidencia@iiap.org.pe, lcamposba@gmail.com
- Isaac Ocampo Yahuarcani, Coordinador del Centro de Alto Rendimiento Computacional de la Amazonía Peruana, iocampo@iiap.org.pe

Sitio web: <http://www.iiap.org.pe/manati/>



Imagen 123: Sitio web del Centro de Alto Rendimiento Computacional de la Amazonía Peruana IIAP

Servicio Nacional de Meteorología e Hidrología de Perú (SENAMHI)

Sede: Lima

La Subdirección de Modelamiento Numérico de la Atmósfera de la Dirección de Redes de Observación y Datos del SENAMHI con la finalidad de generar información de pronósticos más detallados de la distribución de temperatura y de lluvias en todo el territorio peruano, impulsó la implementación de un sistema de alto desempeño computacional en su sede central en Lima.

Este equipo cuenta con 2 módulos de procesamiento con un total de 80 núcleos CPU, fue bautizado como “ANKA”, cuyo significado es águila andina en quechua.

Entre las prioridades de trabajo para el SENAMHI destacan los modelos operacionales ETA y WRF en los dominios Sudamérica (33 km de resolución espacial) y Perú (22 km de resolución espacial) para el pronóstico de 5 a 7 días.

Sitio web: <http://www.senamhi.gob.pe/>



Imagen 124: Plataforma de alto rendimiento de SENAMHI. Fuente: www.senamhi.gob.pe

Corporación Peruana de Aeropuertos y Aviación Comercial (CORPAC)

Sede: Lima.

Esta institución tiene entre sus funciones apoyar la gestión y control de navegación aérea comercial, esto incluye monitoreo de meteorología (incluso dispone de estaciones meteorológicas en diversos aeropuertos), para lo cual le es necesario el uso de plataformas de alto rendimiento computacional.

Cuenta con un equipo de alto rendimiento computacional cuyas principales capacidades comprenden 6 procesadores de 8 núcleos cada uno, que dan un total de 48 núcleos en total.

Contacto: Julio Quezada, Jefe Equipos de Pronósticos y Climatología CORPAC jquezada@corpac.gob.pe

Sitio web: <http://www.corpac.gob.pe/>



Imagen 125: Plataforma de alto rendimiento de CORPAC



Imagen 126: Interfaz principal de la web de CORPAC. Fuente: www.corpac.gob.pe

Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC) del Instituto del Mar del Perú (IMARPE)

Sede: Lima

En este laboratorio se realizan investigaciones de modelado y análisis de procesos, a diferentes escalas espaciales y temporales, para fines de predicción y manejo. Así como investigaciones sobre el impacto del cambio climático en los ecosistemas acuáticos y medidas de adaptación, de conformidad con los lineamientos de política y planes de investigación institucional. Así como evaluaciones de escenarios hipotéticos, tales como proyecciones de impactos del cambio climático, y evaluaciones de estrategias de manejo pesquero con enfoque ecosistémico mediante modelos atmosféricos (WeatherResearch and Forecasting, WRF), oceánicos (Regional OceanicModelSystem, ROMS), biogeoquímicos (PISCES), biológicos (Ichthyop), ecosistémicos (EwE) entre otros.

Cuenta con un clúster de supercomputación compuesto por 12 nodos que contienen un total de 288 núcleos CPU.

Contacto:

Dr. Jorge Tam jtam@imarpe.gob.pe
Jefe del área de modelado numérico de IMARPE

Fuente: www.imarpe.gob.pe



Imagen 127: Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC).

Universidad San Luis Gonzaga de Ica (UNICA)

Para agosto del 2017 se tiene prevista la puesta en funcionamiento del Supercomputador denominado STRATEGY, adquirido por esta universidad en marzo del 2016, con la finalidad de impulsar diversos proyectos de investigación y desarrollo de capacidades áreas de salud, farmacología, epidemiología, minería, ingeniería y meteorología.

Fuente: <https://www.facebook.com/UnicaSLG>

Este Supercomputador cuenta con 40 nodos CPU y 30 nodos de procesamiento gráfico (Intel Xeon Phi) entre lo que alcanzan los 70 nodos alcanzan 1680 núcleos CPU y 3660 núcleos gráficos (los datos fueron obtenidos del documento BASES CONTRATACIÓN INTERNACIONAL N° 001-2015-UNICA, UNIVERSIDAD NACIONAL SAN LUIS GONZAGA DE ICA. Los datos corren el riesgo de haberse modificado, sin embargo son una buena aproximación).

Sitio web: <http://www.unica.edu.pe/>



Imagen 128: Cajas de equipamiento de plataforma de alto desempeño de la UNICA



Imagen 129: Interfaz del sitio web de la UNICA

5.2. Centros de formación de recursos humanos en el Perú

5.2.1. Universidades peruanas

Al inicio del capítulo se hizo referencia a la escasez de recursos humanos con capacidades para desempeñarse en la Computación de Alto Rendimiento en el Perú. Esta escasez de presenta a dos niveles, el primero a nivel de profesionales usuarios de aplicaciones (investigadores de diversas disciplinas científicas diferentes a computación), el segundo a nivel de desarrolladores o creadores de tales aplicaciones.

En este contexto la universidad peruana tiene un rol importante y obligatorio. Después de consultar la página web de la SUNEDU [88], se encontraron un total de 143 universidades, de las cuales, 51 son universidades estatales y 92 son universidades privadas. Posteriormente se revisó el sitio web de cada una de las universidades, buscando identificar las carreras de computación, informática o sistemas, cuyo contenido curricular esté relacionado a la computación de alto desempeño (generación, desarrollo y adaptación de aplicaciones para computación paralela, así como la gestión y mantenimiento de equipamiento de alto desempeño). Los resultados no fueron los mejores y debe ser materia de análisis para las entidades educativas del nivel superior.

En el Perú pasa algo anecdótico, se logró identificar que en 109 universidades existe al menos una carrera relacionada a informática, mayormente ingeniería con sus variantes sistemas, informática, computación y software. En algunos casos se usa la combinación de los términos, así mismo, se encontraron algunas carreras denominadas “ciencia de computación” (el total incluye también los programas de posgrados maestrías y doctorados). Este conjunto de universidades ya el año 2014 habían superado los 25 mil estudiantes, 36% de ellos es universidades estatales (ANR, 2014). De todas esas universidades no llegan ni a 10 las que tienen contenido relacionable a computación de alto desempeño y estas se vinculan mayormente a las carreras de Ciencia de la Computación.

Tampoco no es que en el Perú exista o se espera que existan carreras especializadas (como existe en la UNAM en México). En el caso de este documento se han priorizado los planes curriculares donde se incluyen contenidos sobre Computación Paralela, Procesamiento de señales (imágenes, sonidos, etc.), sin excluir por ejemplo a las carreras no adecuadas a estándares internacionales como de ACM/IEEE (Association for Computing Machinery (ACM)/ Instituto de Ingeniería Eléctrica y Electrónica). Así mismo es importante indicar que no existe la más mínima intención de entrar en polémica sobre el cumplimiento de estándares internacionales en universidades peruanas, ello es materia de otra discusión.

Para efectos de este documento, se ha logrado identificar 8 universidades, que son presentadas en la Tabla 26. De estas 8, la mitad se encuentran en Lima y la otra mitad en Arequipa (2), La Libertad y Puno.

Tabla 26: Cuadro de universidades cuyas de carreras de informática incluyen algunas asignaturas relacionadas a computación de alto rendimiento.

Nro.	Universidad	Región	Régimen	Carreras y Programas relacionados a Informática
1	Pontificia Universidad Católica del Perú - PUCP	Lima	Privado	<p>Pregrado</p> <ul style="list-style-type: none"> • Ingeniería Informática. • Ingeniería Biomédica. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Informática. • Maestría en Integración e Innovación Educativa de las Tecnologías de la Información y la Comunicación (TIC). • Maestría en Ingeniería Biomédica. • Maestría en Procesamiento de Señales e Imágenes Digitales. <p>Doctorados</p> <ul style="list-style-type: none"> • Doctorado en Ingeniería.
2	Universidad Católica San Pablo - UCSP	Arequipa	Privado	<p>Pregrado</p> <ul style="list-style-type: none"> • Ciencia de la Computación. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Ciencia de la Computación.
3	Universidad de Ingeniería y Tecnología - UTEC	Lima	Privado	<p>Pregrado</p> <ul style="list-style-type: none"> • Ciencia de la computación.
4	Universidad Nacional de Ingeniería - UNI	Lima	Estatal	<p>Pregrado</p> <ul style="list-style-type: none"> • Ciencia de la Computación. • Ingeniería de Sistemas. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Ingeniería de Sistemas. <p>Doctorados</p> <ul style="list-style-type: none"> • Doctorado en Ingeniería de Sistemas.

5	Universidad Nacional de San Agustín - "UNSA"	Arequipa	Estatal	<p>Pregrado</p> <ul style="list-style-type: none"> • Esc. Ingeniería de Sistemas. • Esc. Ciencia de la Computación. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Informática. <p>Doctorados</p> <ul style="list-style-type: none"> • Doctorado en Ciencias: Computación.
6	Universidad Nacional de Trujillo - UNITRU	Trujillo	Estatal	<p>Pregrado</p> <ul style="list-style-type: none"> • Ingeniería de Sistemas. • Ciencias de la Computación. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Ingeniería de Sistemas, mención: Administración y Dirección de Tecnologías de la Información. <p>Doctorados</p> <ul style="list-style-type: none"> • Doctorado en Ciencias e Ingeniería.
7	Universidad Nacional del Altiplano - UNAP	Puno	Estatal	<p>Pregrado</p> <ul style="list-style-type: none"> • Escuela Profesional de Ingeniería Estadística e Informática. • Escuela Profesional de Ingeniería de Sistemas. <p>Maestrías</p> <ul style="list-style-type: none"> • Ingeniería de Sistemas. <p>Doctorados</p> <ul style="list-style-type: none"> • Ciencias de la Computación. • Estadística e Informática.
8	Universidad Peruana de Ciencias Aplicadas - UPC	Lima	Privado	<p>Pregrado</p> <ul style="list-style-type: none"> • Ciencias de la Computación. • Ingeniería de Sistemas de Información. • Ingeniería de Software. <p>Maestrías</p> <ul style="list-style-type: none"> • Maestría en Dirección de Sistemas y Tecnologías de la Información

Fuente: SUNEDU y webs institucionales - Elaboración propia

Las 8 instituciones presentadas en la Tabla 25, no representan ni el 10% del total de las universidades peruanas, y ello si constituye una situación desventajosa para el Perú, pues la computación de alto rendimiento debería ser un soporte permanente al desarrollo científico y tecnológico del país, sin embargo más del 90% de universidades no lo aborda ni ligeramente, con lo cual producen profesionales que años posteriores lideraran proyectos en el sector público y privado e incluso algunos serán importantes funcionarios de gobiernos o autoridades que difícilmente priorizaran y comprenderán la importancia de la computación de alto rendimiento en obras públicas, proyectos o procesos de investigación y desarrollo que deberían impulsarse en el Perú.

Es importante destacar la labor de las universidades como la Universidad Católica San Pablo y Universidad San Agustín de Arequipa, que son las que han promovido con mayor energía el enfoque científico de la computación que incluye obligatoriamente computación de alto rendimiento (basado en el Estándar ACM/IEEE-CS), con lo cual vienen sirviendo de referencia y apoyando a todo el sistema universitario peruano. Actualmente ambas universidades, con el apoyo de CIENCIACTIVA de CONCYTEC, vienen desarrollando programas de postgrado a tiempo completo, además lideran una serie de proyectos de investigación que utilizan computación de alto rendimiento. Así mismo, la UNSA fue la primera universidad peruana en adquirir un sistema de alto rendimiento computacional.

Así mismo, el apoyo de CIENCIACTIVA del CONCYTEC ha sido clave para el crecimiento y consolidación de la ciencia de la computación, el siguiente paso tendría que ser la expansión de Ciencia de la Computación en otros lugares del Perú.

5.2.2. Programas de maestría

Para efectos de este documento, consideramos relevante mencionar 4 programas de maestría peruanos que incluyen contenido formativo para computación de alto rendimiento. De ellos, 2 se dictan en Arequipa y 2 en Lima, cabe destacar que 3 de ellos actualmente cuentan con financiamiento de CIENCIACTIVA del CONCYTEC:

UNIVERSIDAD NACIONAL SAN AGUSTÍN (AREQUIPA):

Maestría en Informática. Iniciada en junio del 2015, con una duración de dos años. Este programa recibió financiamiento de CIENCIACTIVA de CONCYTEC y actualmente cuenta con 15 estudiantes becados a tiempo completo.

Tienen como objetivo contribuir en la formación y capacitación de capital humano, especialistas en informática calificados en resolver problemas de la industria, sector productivo y sector educación, a través de proyectos de investigación usando TICs.

La maestría es llevada a cabo por la Escuela de Posgrado CITEC, que tiene entre sus líneas de investigación Computación de Alto Desempeño y Procesamiento Multimedia (procesamiento de imágenes). La maestría incluye en sus asignaturas cursos de Procesamiento Multimedia, Minería de Datos, Tópicos de Programación Avanzada, entre otros.

Así mismo, la UNSA cuenta con alianzas con instituciones internacionales como Universidad Tecnológica de La Habana José Antonio Echeverría de (Cuba), École Nationale Supérieure d'Électronique, d'Électrotechnique, d'Informatique, d'Hydraulique, et des Télécommunications, de Francia, la Universidad de Chile, etc.

Tabla 27: Listado de banco de Tesis de la Maestría en Informática de la UNSA Arequipa.

Nro.	Título	Autor	Año
1	Modelo de búsqueda semántica con un enfoque de expansión de consultas, caso de estudio: búsqueda de información de videos sobre educación ambiental.	Jhon Edilberto Monroy Barrios	2016
2	Reconocimiento de defectos físicos de granos de café verde arábica en muestras usando visión artificial.	Juan Ramirez Ticona	2016
3	Exploración de artículos científicos con realidad virtual y procesamiento del lenguaje natural.	Ricardo Josué Linárez Juárez	2017
4	Un enfoque probabilístico de sistemas de recomendación basado en contenido.	Laura Vanessa Cruz Quispe	2016
5	Potenciando los dispositivos móviles mediante técnicas de Offloading en arquitecturas Cloudlet.	Anrique Arturo Soto Mendoza	2016

Fuente: UNSA - Arequipa. Elaboración propia

Datos de Contacto

Director: Dr. José Herrera
Email: citec@unsa.edu.pe

UNIVERSIDAD CATÓLICA SAN PABLO (AREQUIPA): Maestría en Ciencia de la Computación.

Este programa de maestría se inició en abril del 2016, al igual que la UNSA por promoción recibe 15 alumnos becados, también es financiado por CIENCIACTIVA del CONCYTEC y tiene una duración de 2 años

Su objetivo es formar recursos humanos de alto nivel, capaces de desarrollar ciencia básica y aplicada, transmitir el conocimiento, innovar y también aplicar tecnologías derivadas de ciencia de la computación, tanto tradicionales como nuevas, participando como líderes de grupos de trabajo en la solución de problemas del sector público y privado.

La maestría es llevada a cabo por el Centro de Investigación e Innovación en Ciencia de la Computación (RICS), que trabaja tres áreas troncales de investigación e innovación: Gráficos, Imágenes y Sistemas Inteligentes, y a través de ellas profundizar en áreas estratégicas e importantes de Ciencia de la Computación, considerando su estándar internacional.

Como contenido para computación de alto desempeño hemos identificado las siguientes asignaturas:

- **Estructura de Datos y Algoritmos:** que incluye: Memoria dinámica, Programación Genérica, Estructuras de datos espaciales, Análisis de complejidad de algoritmos.
- **Programación Paralela y Distribuida:** que incluye: Programación Paralela en CPU, Programación en Tarjetas Gráficas GPU, Programación Distribuida.
- **Imágenes: que incluye:** Procesamiento digital de imágenes y videos, Visión Computacional, Reconocimiento de patrones.
- **Gráficos:** Que incluye: Computación Gráfica, Geometría Computacional, Visualización de Datos, Humano Computador.
- **Seminarios Avanzados de Computación:** que incluye: Problemas relacionados a Big Data, Desarrollo de Videojuegos, Innovación y Emprendimientos Tecnológicos.

Así mismo, el Dr. Alex Cuadros, Director de la RICS, nos ha compartido la lista de los trabajos de Tesis que vienen trabajando sus estudiantes de maestría, que según sus propias palabras vienen siendo evaluadas y asesoradas por un grupo de científicos de universidades brasileras, norteamericanas, etc. Más de la mitad de estos trabajos están orientados directamente al uso de Computación de Alto desempeño o Supercomputación.

Tabla 28: Proyectos de Tesis de la Maestría de Ciencia de la Computación de la UCSP de Arequipa.

Nro.	TEMA	TESISTA	ORIENTADOR
1	Método de segmentación para contextos generalizados, con un enfoque paralelo.	Juan Carlos Isaías Dueñas García	Alex Jesús Cuadros Vargas
2	Desarrollo de un modelo híbrido usando modelos de aprendizaje profundo para la recuperación de información multi-modal en texto e imágenes.	Miler Díaz Zevallos	José Eduardo Ochoa Luna
3	δ -radius IVRL: Paradigma de integración de aprendizaje por refuerzo en sistemas multi-agente.	Jesús Alejandro Camargo Monroy	Dennis Barrios Aranibar
4	Implementación de un método para la detección de cambios orientados a objetos en zonas urbanas a partir de imágenes obtenidas desde un VANT.	Fredy Calsin Cari	Marc Antoine Le Guen
5	Reconocimiento de rostros con Elastic Bunch Graph Matching en aplicaciones de vídeo.	José Rodrigo Galdós Chávez	Juan Carlos Gutiérrez Cáceres
6	A practical succinct data structure for ordinal trees.	Joshimar Córdova Monroy	Gonzalo Navarro
7	Reconocimiento de actividades cotidianas.	Kelly Vizconde La Motta	Guillermo Cámara Chávez
8	kMesh: algoritmo paralelo para construir mallas adaptables a la forma de un objeto.	Ronald Ubel Adolfo Gonzales Vega	Alex Jesús Cuadros Vargas
9	Reconocimiento de escenas violentas en Video usando massive threading.	Jorge Thony Ramírez Ticona	Guillermo Cámara Chávez
10	Suavizado de Fronteras en Mallas Simpliciales Segmentadas Generadas a Partir de Imágenes.	Tony Liedyn Choque Ramos	Alex Jesús Cuadros Vargas
11	Generación de triangulaciones de delaunay persistentes.	Jainor Nestor Cárdenas Choque	Christian López del Álamo
12	Reconocimiento de entidades nombradas para el idioma español utilizando Conditional Random Fields con características no supervisadas.	Jenny Linet Copara Zea	José Eduardo Ochoa Luna
13	Un método radial interactivo para la exploración visual de datos en alta dimensión.	Germain García Zanabria	Erick Gómez Nieto
14	Una propuesta de algoritmo evolutivo de inspiración cuántica para representación real usando filtro de partículas.	Josimar Edison Chire Saire	Yván Tupac
15	Detección de eventos anómalos en vídeo.	Neptalí Menejes Palomino	Guillermo Cámara Chávez
16	Detección de placas vehiculares en un contexto masivamente paralelo.	David Ramirez Añasco	Alex Cuadros

Fuente: DINA CONCYTEC - Elaboración propia.

Datos de Contacto

Director: Dr. Alex Jesús Cuadros Vargas
 Teléfono: +51 (54) 605630 anexo 664
 Email: rics@ucsp.edu.pe

UNIVERSIDAD NACIONAL DE INGENIERÍA (LIMA): Programa de Maestría en Ciencias en Ingeniería Electrónica con Mención en Procesamiento Digital de Señales e Imágenes

Este programa de maestría se inició en mayo de 2016, tiene financiamiento de CIENCIACTIVA del CONCYTEC que ha permitido otorgar becas de tiempo completo a 10 estudiantes. en este caso, es valido indicar, que si bien es cierto la maestría proviene de ingeniería electrónica, sin embargo, su contenido incluye asignaturas orientadas a computación de alto rendimiento.

Entre los objetivos principales se encuentran “formar especialistas en la línea de procesamiento digital de señales e imágenes con énfasis en las áreas prioritarias de teledetección, biotecnología y medio ambiente.

La maestría es llevada a cabo por la Escuela Central de Posgrado de la Universidad Nacional de Ingeniería y las asignaturas relacionadas a Computación de Alto Desempeño que considera son: Procesamiento Digital de Señales, Procesamiento Digital de Imágenes, Tópicos Especiales en Procesamiento Digital de Señales e Imágenes, cabe indicar que en este caso, desde el inicio de la maestría se han priorizados áreas de aplicación como: Agricultura de Precisión, Monitoreo del medio ambiente y los recursos naturales, Arqueología, Minería, Urbanismo y Paisajismo

Datos de Contacto

Director: Dr. Guillermo Leopoldo Kemper Vásquez
Teléfono: 382-2163, 9510-35137
Email: postgrado_fiee@uni.edu.pe

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERU LIMA: Maestría en Procesamiento de Señales e Imágenes Digitales. Este programa es más antiguo que los anteriores, a la fecha no reciben financiamiento del estado, por lo que los estudiantes deben cubrir los costos.

Es llevada a cabo por la Escuela de Posgrado de la PUCP. Esta maestría es la única de las 4 maestrías relacionadas a tecnologías de la información que contiene asignaturas adaptables a Computación de Alto desempeño. Las asignaturas relacionadas son las siguientes:

- Procesamiento de Imágenes Digitales.
- Algoritmos paralelos
- Procesamiento de imágenes médicas
- Métodos computacionales para procesamiento digital de señales

Datos de Contacto

Director: Dr. Roberto Lavarello Montero
Teléfonos: (511) 626-2530 / (511) 626-2531
Email: <http://posgrado.pucp.edu.pe/maestria/procesamiento-de-senales-e-imagenes-digitales/admision/>

5.2.3. Supercomputación en Institutos técnicos

También se ha hecho una ligera revisión de los institutos de formación técnica del Perú, teniendo en cuenta la base de datos de la iniciativa “Ponte en Carrera” (<http://www.ponteencarrera.pe/>) complementada por búsqueda en Google. Se ha encontrado la existencia de al menos 390 instituciones de educación técnica con carreras de Computación e Informática, en algunos casos imparten clases de administración de redes, programación, gestión de sistemas operativos Linux, sin mayor contenido relacionado a computación de alto desempeño. Dentro de este listado instituciones destacan TELESUP, SENATI, UNITEK, CIBERTEC y el mismo Alas Peruanas con sedes en descentralizadas.

5.3. Proyectos de investigación sobre computación de alto rendimiento

En el Perú a pesar de las limitaciones institucionales, en diversas regiones se vienen formando grupos de investigadores que impulsan proyectos relacionados a computación de alto desempeño, entre los que podemos destacar el de la Universidad Nacional de Ingeniería UNI (incluyendo INICTEL UNI), Universidad San Marcos, Universidad San Agustín de Arequipa, Universidad Católica San Pablo, Universidad Nacional de Trujillo, etc. Al no existir un banco o registro oficial de proyectos, se consultó los proyectos de investigación registrados por los investigadores del Directorio Nacional de Investigadores – DINA de CONCYTEC. Se tuvo en cuenta los siguientes criterios:

- Se priorizaron proyectos liderados por peruanos (sin importar el lugar de sus estudios) para instituciones peruanas públicas y privadas.
- El siguiente paso fue priorizar proyectos de investigación directamente relacionados al empleo de computación paralela, así como proyectos de infraestructura para computación de alto rendimiento.
- También se priorizaron proyectos de investigación relacionados a las líneas de investigación de procesamiento de imágenes y visión computacional, computación gráfica, big data y minería de datos, redes neuronales, algoritmos que incluyan procesamiento de grandes volúmenes de datos. Es importante mencionar que estos proyectos no necesariamente utilizan procesamiento paralelo, sin embargo los resultados de estos pueden ser adaptados para aprovechar las capacidades de las plataformas de alto rendimiento.
- Para priorizar los proyectos de investigación en el listado se tuvieron en cuenta la existencia de información básica (Título, Autor, Fechas, Descripción). Teniendo en cuenta las limitaciones de la estructura de los datos del DINA.

Se lograron identificar 225 proyectos procedentes de las siguientes instituciones:

Tabla 29: Cantidad de proyectos identificados por institución

N°	INSTITUCIÓN	PROCEDENCIA	CANTIDAD DE PROYECTOS
1	UNIVERSIDAD NACIONAL DE INGENIERÍA - UNI	LIMA	26
2	INSTITUTO NACIONAL DE INVESTIGACIÓN Y CAPACITACIÓN DE TELECOMUNICACIONES - UNI	LIMA	19
3	UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS	LIMA	18
4	PONTIFICA UNIVERSIDAD CATOLICA DEL PERU	LIMA	15
5	UNIVERSIDAD PRIVADA CAYETANO HEREDIA	LIMA	13
6	UNIVERSIDAD NACIONAL SAN AGUSTÍN DE AREQUIPA	AREQUIPA	11
7	UNIVERSIDAD NACIONAL DE TRUJILLO	LA LIBERTAD	8
8	UNIVERSIDAD CATOLICA SAN PABLO	AREQUIPA	7
9	UNIVERSIDAD CATÓLICA SANTA MARÍA	AREQUIPA	7
10	UNIVERSIDAD NACIONAL DEL ALTIPLANO	PUNO	7
11	UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN	TACNA	7
12	UNIVERSIDAD NACIONAL SAN ANTONIO ABAD DEL CUZCO	CUZCO	7
13	UNIVERSIDAD NACIONAL DE SAN MARTIN	SAN MARTÍN	5
14	UNIVERSIDAD NACIONAL DEL CENTRO DEL PERÚ	JUNÍN	5
15	UNIVERSIDAD SENOR DE SIPAN	LAMBAYEQUE	5
16	UNIVERSIDAD DE CIENCIAS Y HUMANIDADES	LIMA	4
17	UNIVERSIDAD DE PIURA	PIURA	4
18	UNIVERSIDAD DEL PACIFICO	LIMA	4
19	INSTITUTO GEOFÍSICO DEL PERU	LIMA	3
20	UNIVERSIDAD ANDINA NESTOR CACERES VELÁSQUEZ	PUNO	3
21	UNIVERSIDAD NACIONAL DE MOQUEGUA	MOQUEGUA	3
22	UNIVERSIDAD NACIONAL MICAELA BASTIDAS DE APURIMAC	APURIMAC	3
	OTROS		41
	TOTAL		225

Fuente: DINA CONCYTEC - Elaboración propia

De los 225 proyectos identificados, 125 proceden de instituciones de Lima (55.5%), teniendo a la Universidad Nacional de Ingeniería – UNI, como la institución peruana líder en proyectos de investigación vinculables a computación de alto desempeño. Incluso si se tendría en cuenta también los proyectos liderados por INICTEL UNI (centro de investigación que desde el 2010 forma parte de la UNI), la UNI en total alcanzaría 45 proyectos. Se decidió separarlos por fines de mejor análisis de la información.

La Universidad Nacional de San Marcos y la Pontificia Universidad Católica del Perú cuentan con 18 y 15 proyectos. Posterior a ellos sigue en la lista un grupo creciente de universidades procedentes de Arequipa, La Libertad, Puno, Tacna y Cuzco, y el mismo Lima. En total se identificaron 53 instituciones peruanas pertenecientes a 18 regiones que trabajan o trabajaron en algún proyecto de computación de alto desempeño. Fuera de Lima, Arequipa es la región con mayor cantidad de proyectos,

y ellos por el esfuerzo de sus 3 universidades principales Universidad Nacional San Agustín de Arequipa, Universidad Católica San Pablo y Universidad Católica Santa María. Luego de ellos le acompañan en la lista La Libertad, Puno, Cuzco, Junín, Lambayeque y Tacna. El detalle puede ser visto en el siguiente cuadro:

Tabla 30: Proyectos identificados por región.

Nro.	REGIÓN	CANTIDAD DE PROYECTOS
1	LIMA	125
2	AREQUIPA	27
3	LA LIBERTAD	10
4	PUNO	10
5	CUZCO	7
6	JUNÍN	7
7	LAMBAYEQUE	7
8	TACNA	7
9	PIURA	6
10	SAN MARTÍN	5
11	APURIMAC	3
12	MOQUEGUA	3
13	ANCASH	2
14	AYACUCHO	2
15	LORETO	1
16	HUANCAVELICA	1
17	HUANUCO	1
18	ICA	1
	TOTAL	225

Fuente: DINA CONCYTEC - Elaboración propia

Respecto a las líneas de investigación de los proyectos, se ha registrado un importante número de ellos en la línea Procesamiento de imágenes y visión computacional, alcanzando los 62 proyectos (27.5%), teniéndose como principales aplicaciones lo relacionados teledetección para el estudio del bosque y espacios naturales, agricultura de precisión, agroindustria, así como aplicaciones médicas.

Así mismo es importante destacar, en los últimos años, el crecimiento del número de proyectos relacionados a redes neuronales (Deep Learning), Big data y Minería de Datos, en ambos casos es clave el uso de computación paralela en el procesamiento de los datos. Respecto a la línea propia de HPC, la cifra alcanzó apenas el 6% del total de los proyectos, una cifra que revela la escasa expansión del HPC

en el Perú (es válido mencionar que para seleccionar los proyectos básicamente se observó el Título y el Resumen, sin verificar los resultados o logros de los proyectos, por lo que existe un riesgo que no en todo los casos se haya aplicado realmente procesos de paralización de los programas, lo que en la práctica no es computación de alto rendimiento, con ello incluso la cifra podría ser menor al 6%).

Tabla 31: Proyectos identificados por línea de investigación

Nro.	Línea de investigación	Cantidad de Proyectos
1	Procesamiento de imágenes y visión computacional	62
2	Redes neuronales	32
3	Algoritmos genéticos y computación bioinspirada	25
4	Big Data y Minería de Datos	23
5	HPC	14
6	Bioinformática	13
7	Química Computacional	3
8	Astrofísica	2
9	Otros Algoritmos	51
	Total	225

Fuente: DINA CONCYTEC - Elaboración propia

Respecto a proyectos de computación de alto desempeño se han identificado 14 proyectos, la mitad de ellos en Lima y el resto en otras regiones (APURIMAC, AREQUIPA, CUZCO y LORETO). Es pertinente mencionar que 12 de estos proyectos han sido realizados en los últimos 5 años, lo que demuestra que es un tema de reciente penetración en el Perú. Así mismo en el listado se incluyen los proyectos del IIAP y del IGP que incluyen la adquisición de equipamiento.

Tabla 32: Proyectos de computación de alto rendimiento en Perú.

N°	Proyecto	Institución	Región	Inv. Principal	Descripción	Año
1	Implementación de sistemas HPC sobre plataformas de software libre para la resolución de problemas con gran demanda de procesamiento	Universidad Nacional San Antonio Abad del Cuzco	CUZCO	EDWIN CARRASCO POBLETE	Sin datos.	2005
2	Algoritmo Paralelo para Principal Component Pursuit en la Arquitectura CUDA con Casos de Estudio orientados a Seguridad Ciudadana y de Ayuda a Personas Discapacitadas	Pontificia Universidad Católica del Perú	LIMA	PAUL RODRIGUEZ VALDERRAMA	PCP es exitosamente usado para segmentar automáticamente objetos en movimiento del fondo (objetos estáticos) en aplicaciones médicas y en seguridad ciudadana, etc. Sin embargo, el costo computacional de los algoritmos existentes es muy elevado y no permite aplicaciones prácticas en tiempo real. Este proyecto propone desarrollar un nuevo algoritmo paralelo, computacionalmente eficiente, para el cómputo de PCP y aplicaciones, teniendo especial énfasis en el modelo de programación SIMT o CUDA.	2012

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

3	Computación de Alto Performance (HPC) basado en clúster heterogéneo usando hardware de bajo costo; diseño y construcción – UNAMBA	Universidad Nacional Micaela Bastidas de Apurímac	APURIMAC	NORA GLADYS ECHEGARAY PEÑA	Proyecto de investigación en proceso	2012
4	Laboratorio de Computación de Alto Desempeño	Universidad Nacional San Agustín de Arequipa	AREQUIPA	LUIS ALFARO CASAS	El proyecto tiene como objetivo la adquisición e implantación de un equipamiento de alto desempeño para la Universidad Nacional de San Agustín. Este equipamiento debe contemplar las propiedades para soportar Computación de Alto Desempeño. Con este equipamiento se proyecta ofrecer recursos computacionales para problemas que requieran: Alto Desempeño; alto procesamiento; alta disponibilidad; escalabilidad; entre otros.	2013
5	Desarrollo e implementación de algoritmos de convolución-2D para el procesamiento digital de imágenes satelitales en programación paralela	Instituto Nacional de Investigación y Capacitación de Telecomunicaciones - UNI	LIMA	VICTOR FLORES BENTES	Sin datos.	2013
6	Centro de Excelencia en Computación de Alto Desempeño para Centros Urbanos	Universidad Católica San Pablo	AREQUIPA	ALEX CUADROS	Sin datos.	2014
7	Validación de una máquina de micro gravedad midiendo los niveles de expresión genética provenientes de un microarrays utilizando algoritmos de programación paralela en una arquitectura GRID Computing.	Instituto Nacional de Investigación y Capacitación de Telecomunicaciones - UNI	LIMA	DR. DAVID LAVÁN QUIROZ	Tarjeta de control de velocidad de motores DC y tarjeta de adquisición de datos provenientes de sensores (Temperatura, humedad, potenciales de membrana, pH, campo magnético, giroscopio, acelerómetro, torque y video) ubicados en la plataforma móvil con comunicación por puerto USB 2.	2014
8	Desarrollo de algoritmos paralelos para la recuperación por contenido en imágenes médicas	Pontificia Universidad Católica del Perú	LIMA	CÉSAR BELTRÁN	Sin datos.	2014
9	Fortalecimiento de capacidades de infraestructura para procesamiento de información sobre recursos naturales y ambiente amazónico	INSTITUTO DE INVESTIGACIONES DE LA AMAZONÍA PERUANA	LORETO	ISAAC OCAMPO YAHUARCANI	Comprendió la adquisición de equipamiento de alto rendimiento computacional orientado al procesamiento de grandes volúmenes de información sobre deforestación, clima, hidrología, biodiversidad. Cambio climático de la Amazonía peruana.	2015
10	SISTEMA COMPUTACIONAL DE ALTO RENDIMIENTO PARA LA SIMULACION DE FLUIDOS	INSTITUTO GEOFÍSICO DEL PERU	LIMA	IVONNE KATHERINE, MONTES	El proyecto enfrenta la necesidad de ampliar el conocimiento sobre los fenómenos geofísicos naturales que generan pérdidas socioeconómicas en el Perú. Este conocimiento es adquirido mejorando el modelado numérico del océano y la atmósfera. El sistema de alto rendimiento será utilizado por investigadores en universidades, instituciones y centros de investigación en el Perú.	2015
11	Círculo de Investigación en Computación de Alto Desempeño con énfasis en el desarrollo de Métodos y Técnicas de Minería de Datos de Gran Escala para el apoyo en investigaciones de cambio climático.	Universidad Nacional San Agustín de Arequipa	AREQUIPA	CRISTIAN JOSE LOPEZ DEL ALAMO	El cambio climático duplicará la frecuencia de eventos climatológicos extremos. En este contexto, con los datos almacenados durante varios años por el Servicio Nacional de Meteorología e Hidrología (SENAMHI) y el Instituto de Investigación de la Amazonía Peruana (IIAP), las técnicas de aprendizaje profundo emergentes en la última época, permitirán la construcción de un modelo computacional para la detección de patrones ocultos para la comprensión de los fenómenos climáticos y meteorológicos.	2015

12	Estudio de Algoritmos de Balanceo de carga implementados en un clúster de sistemas embebidos	Universidad Nacional de Ingeniería	LIMA	YURI NUÑEZ MEDRANO	El proyecto tiene como objetivo construir un clúster de Alta disponibilidad de configuración activo/activo, en los que se instala paquetes para Balanceo de carga. Luego de configuración para su funcionamiento como clúster, se harán las pruebas de rendimiento en la calidad de servicios. Teniendo la certeza del óptimo funcionamiento del clúster se procederá a las pruebas de estudio de algoritmos como: Opportunistic Load Balancing (OLB) , Load Balance Min-Min (LBMM), Round-robin y Weighted	2016
13	Generación de modelos tridimensionales de huesos humanos para fines de exploración y cuantificación ósea	Universidad Católica San Pablo	AREQUIPA	ALEX CUADROS VARGAS	El proyecto comprende la creación de modelos geométricos tridimensionales, a partir de tomografías, donde se distinga cada hueso por separado, para calcular la densitometría, obteniendo resultados equivalentes a las de un densitómetro, pero de forma localizada.	2016
14	Computación de Alto Desempeño con Minería de Datos a gran escala en Cambio Climático	Universidad Nacional San Agustín de Arequipa	AREQUIPA	LUIS ALFARO CASAS	El objetivo de la propuesta es desarrollar capacidades en el procesamiento y análisis de datos masivos con indexación temporal, competencia que se requiere para consolidar y fortalecer el centro de investigación en "Computación de Alto Desempeño" ubicado en la UNSA, aplicado al estudio del Cambio Climático.	2016

Fuente: DINA CONCYTEC - Elaboración propia

5.4. Expertos, investigadores y usuarios de Supercomputación en el Perú

En el Perú no existe un registro que a ciencia cierta nos presente de usuarios e incluso creadores de aplicaciones para equipos de alto desempeño computacional. Sin embargo a partir de la recopilación de proyectos que hicimos (consultando el DINA de CONCYTEC [89]) y a partir de las primeras experiencias gestionando el MANATI IIAP en su primer mes, nos atrevimos a elaborar y presentar algunos pequeños listados, que seguidamente presentamos:

5.4.1. Expertos peruanos que utilizan plataformas de alto rendimiento computacional

El primer caso es un listado de al menos 5 investigadores peruanos que ya desarrollan proyectos de investigación cuyos procesos necesariamente necesitan trabajar con Supercomputadoras. En este caso son investigadores relacionados a Bioinformática, Astrofísica, Oceanografía y Química Computacional. Algo que es importante destacar es que todos los expertos del listado hicieron sus estudios de doctorado en el extranjero en países como: Brasil, Estados Unidos, España y Alemania, países donde justamente la supercomputación es un instrumento vital para la ciencia. Es muy importante indicar que los datos presentados de cada investigador fueron extraídos principalmente de la base de datos del DINA de CONCYTEC (sólo se consideró parte de los trabajos de cada investigador). Seguidamente presentamos un cuadro por cada investigador:

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Nombres y apellidos	David Lavan Quiróz	
Línea de investigación	Bioinformática	
Datos de contacto	Email: dlavan@inictel-uni.edu.pe	
Centro laboral actual	Instituto Nacional de Investigación y Capacitación de Telecomunicaciones de la Universidad Nacional de Ingeniería – INICTEL UNI.	
Registro REGINA:	5889	
Ejerce ó ejerció docencia en	Universidad del Callao Universidad Nacional de Ingeniería	
Estudios realizados	<ul style="list-style-type: none"> ● Doctorado en Bioquímica, Biología Molecular, Biomedicina y Biotecnología (Biociencias Moleculares). Se graduó como doctor en Ciencias defendiendo la tesis con el nombre de: “Análisis de la expresión Génica en pupas de <i>D. Melanogaster</i> expuestas a microgravedad (ug)”. En Marzo del 2012 en la Universidad Autónoma de Madrid. ● Master en Biofísica por la Universidad Autónoma de Madrid. ● Estudios concluidos en el Master de Física Teórica en la Pontificia Universidad Católica del Perú. ● Licenciado en Física por la Universidad Nacional de Trujillo. ● Bachiller en Física Matemática por la Universidad Nacional de Trujillo. 	
Experiencias y proyectos actuales en HPC	<ul style="list-style-type: none"> ● Validación de una máquina de microgravedad midiendo los niveles de expresión genética provenientes de un microarrays utilizando algoritmos de programación paralela en una arquitectura de GRID Computing. ● Búsqueda de MOTIF en zonas conservadas evolutivamente en <i>Drosophila melanogaster</i> utilizando super computación. 	
Proyectos	<ul style="list-style-type: none"> ● Experimentos biológicos en microgravedad real y simulada en el programa ELIPSE-2 de la Agencia Espacial Europea ● Análisis de la expresión genética de <i>Drosophila melanogaster</i> en condiciones de microgravedad real y simulada. Universidad Autonoma de Madrid 2006 – 2009. ● Validación de una máquina de microgravedad midiendo los niveles de expresión genética provenientes de un microarrays utilizando algoritmos de programación paralela en una arquitectura de GRID Computing ● “Validar la máquina de microgravedad mediante estudios de expresión genética mediante microarrays, los resultados son comparados con los publicados por la ESA.” Unidad ejecutora 002 - INICTEL-UNI noviembre 2013. 	

Nombres y apellidos	Fiestas Iqira José Antonio	
Línea de investigación	Astrofísica	
Datos de contacto	joseafiestasi@gmail.com	
Centro laboral actual	Universidad Nacional de Ingeniería	
Registro REGINA	4927	
Ejerce docencia en	Universidad Nacional de Ingeniería	
Estudios realizados	Estudios de Maestría y Doctorado en la Universidad de Heidelberg en Alemania.	
Trabajos en HPC	<ul style="list-style-type: none"> ● Cyberinfrastructure for Accelerating Physics & Astronomy Applications With Many-core and Accelerator-Based Systems. Desarrollo de Software en Supercomputadores híbridos (CPU-GPU). Inicio setiembre del 2011, final setiembre del 2013. Investigador Principal: Horst Simon. ● Silk-Road Project in computational astrophysics: Desarrollo de software con aplicación a la astrofísica en arquitecturas híbridas CPU-GPU. Inicio Octubre 2013 octubre 2014 Investigador Principal: Jose Antonio Fiestas Iqira. 	

Trabajos en HPC	<ul style="list-style-type: none"> ● Witnesses of Cosmic History: Formation and evolution of galaxies, black holes, and their environment. Modelos computacionales evolutivos de galaxias con agujeros negros. Inicio: Marzo 2006, Fin: Agosto 2011 Investigador Principal: Jose Antonio Fiestas Iquira.
Publicaciones	<ul style="list-style-type: none"> ● Accelerated Many-Core GPU Computing for Physics and Astrophysics on Three Continents (En SCOPUS, Autor principal: Spurzem R. Año 2012). ● Astrophysical particle simulations with large custom GPU clusters on three continents (En SCOPUS, Autor principal:Spurzem R.Año2011).

Nombres y apellidos	Gómez Valdéz Badhin	
Línea de Investigación	Química Computacional	
Datos de Contacto	Badhingomez@gmail.com	
Centro Laboral Actual	Universidad Católica Santa María	
Registro REGINA	102	
Ejerce Docencia en	Universidad Nacional de Ingeniería, Universidad de Lima	
Estudios Realizados:	<ul style="list-style-type: none"> ● Doctor en Química, Universidad de Chile, Chile. ● Licenciado en Química, Universidad Nacional de San Agustín de Arequipa. ● Bachiller en Química, Universidad Nacional de San Agustín de Arequipa, Perú. 	
Trabajos en HPC:	<ul style="list-style-type: none"> ● "Estudio de la reactividad de ciclotidos para el potencial uso en tratamientos del alzheimer, mediante técnicas mixtas de mecánica cuántica y molecular (qmmm)". Inicio enero del 2016, final enero del 2018. ● Identificación y estudio de los sitios activos de las lipoproteínas del m. Bovis y m. Caprae usando herramientas de bioinformática. Inicio julio del 2016, final julio del 2018. ● "Análisis computacional de la interacción de las proteínas de unión a lípidos (lbps) de <i>echinococcus granulosus</i> con variantes de aceites vegetales usando herramientas de Química Computacional, en busca de nuevos tipos de drogas para la terapia de hidatidosis". Inicio en enero 2017 a enero 2019 ciencias naturales. ● "Estudio mediante química computacional para la elucidación de las propiedades fisicoquímicas, cinéticas y de reactividad del 3-amino-1- hidroxí-3,4-dihidroquinolin-2(1h)- ona y sus derivados, orientado al diseño de fármacos para la esquizofrenia." inicio enero del 2016, final enero del 2018. 	

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Nombres y apellidos	Ivonne Katherine Montes Torres	
Línea de investigación	Oceanografía Física	
Datos de contacto	imontes@igp.gob.pe	
Centro laboral actual	Instituto Geofísico Del Perú	
Registro REGINA	961	
Ejerce docencia en	Universidad Peruana Cayetano Heredia	
Estudios realizados	Doctor en Oceanografía (Universidad de Concepción - Chile).	
Trabajos en HPC	Sistema Computacional de Alto Rendimiento para la Simulación de Fluidos Geofísicos: el objetivo del proyecto es disponer, a mediano plazo, de un sistema computacional de alto rendimiento para la simulación de Fluidos Geofísicos, que cubra el costo computacional de las diversas áreas científicas del Instituto Geofísico del Perú y las universidades, a fin de contribuir al entendimiento de los fenómenos geofísicos (naturales y antropogénicos) que afectan a la población peruana y sus actividades, con énfasis a problemas relacionados con el Fenómeno El Niño y Cambio Climático. (Inicio: Febrero 2015, Final junio 2017. Investigador Principal: Ivonne Katherine Montes Torres.	
Publicaciones	<ul style="list-style-type: none"> ● Oceanografía de la Cuenca de Yucatán, Mar Caribe: Implicaciones dinámicas, observaciones y modelos numéricos, Capítulo de Libro Editorial Academia Española. ● Evolución del Sistema Computacional de Alto Rendimiento en el IGP para un mejor pronóstico y estudio de los fenómenos climático. Artículo de Boletín. ● Modelado Numérico de la Circulación del Pacífico Tropical Este y su conexión subsuperficial con el Sistema de Corrientes de Humboldt: Condiciones Climatológicas. Resumen de Congreso. II Congreso de Ciencias del Mar del Perú High-resolution modeling study of the Peru Upwelling System. 	

Nombres y apellidos	Mirko Juan Zimic Peralta:	
Línea de investigación	Bioinformática	
Datos de contacto	mirko.zimic@upch.pe	
Centro laboral actual	Jefe del Laboratorio de Bioinformática y Biología Molecular de la Universidad Peruana Cayetano Heredia	
Registro REGINA	2058	
Ejerce docencia en	Universidad Peruana Cayetano Heredia	
Estudios realizados	<ul style="list-style-type: none"> ● Posee grado PhD en Control y Prevención de Enfermedades por la Universidad Johns Hopkins. ● MSc. Bioquímica por la Universidad Peruana Cayetano Heredia (UPCH). ● BSc. Física por la Universidad Nacional de Ingeniería. 	
Trabajos en HPC	<ul style="list-style-type: none"> ● Desarrollo de un sistema automático experto basado en un servidor WEB, para el diagnóstico automático y a distancia de tuberculosis mediante el análisis de imágenes usando el método MODS. Grant PIN FINCYT (Agosto 2008). Investigador Principal: Mirko Zimic Peralta. ● Creación de un microscopio invertido de bajo costo e implementación de un software de reconocimiento de patrones. Subvencionado por The Wellcome Trust. Investigador Principal: Mirko Zimic Peralta. 	

Trabajos en HPC	<ul style="list-style-type: none"> ● Implementación de un GRID computacional para el desarrollo de herramientas de telemedicina, bioinformática. Grant FINCYT para Equipamiento (2008). Investigador Principal: Mirko Zimic Peralta. ● Implementación de una plataforma WEB para facilitar el diagnóstico de lesiones malignas de piel en la región de Huancavelica de manera remota. Subvenciones Especiales-CONCYTEC 2012. Investigador Principal: Mirko Zimic Peralta. ● Development of an algorithm for the recognition and automatic count of intestinal parasites. PROCYT-CONCYTEC (2008-2009) & PIBAB-FINCYT (2009-2011). Investigador Principal: Mirko Zimic Peralta. ● Desarrollo de una vacuna multiepitópica contra influenza aviar empleando herramientas genómicas e inmuno-bioinformáticas. PIPEA FIDECOM FINCYT 2012. Investigador Principal: Mirko Zimic Peralta. ● Bioinformatics-driven Antigen Identification for the Development of an Immunologic Assay for the Diagnosis of Drug-Resistant Latent Mycobacterium tuberculosis (MTB) Infection. Proyecto de colaboración UPENN-UPCH. Co-Investigador principal: Mirko Zimic Peralta. Duración: desde marzo 2011 a diciembre 2012. ● Sistema Automático de Lectura de Láminas de Papanicolaou, vinculado a una plataforma web para Telediagnóstico de Cáncer de Cuello Uterino. Investigador Principal: Mirko Zimic Peralta. Duración: Marzo 2012 a Abril 2013. ● Desarrollo de un sistema integral de visualización de imágenes médicas microscópicas con capacidad de observación directa e inversa en modo manual/digital, y de reconocimiento automático de patrones en una plataforma web para telediagnóstico rápido y económico. Investigador Principal: Mirko Zimic Peralta. ● "Desarrollo de una nueva herramienta inmunoinformática de predicción de epítopes inmunogénicos específica para patógenos de pollo para el diseño más eficiente de vacunas aviares". FINCTY (2013 UPCH Noviembre 2013. Investigador Principal: Mirko Juan Zimic Peralta. ● Sistema automático de diagnóstico de cáncer de cuello uterino y vaginosis bacteriana basado en un algoritmo de reconocimiento de patrones embebido en un teléfono celular. Investigador Principal: Mirko Juan Zimic Peralta.
Publicaciones:	<ul style="list-style-type: none"> ● Molecular distribution of amino acid substitutions on neuraminidase from the 2009 (H1N1) human influenza pandemic virus. Artículo en revista científica (Autor Principal: Quiliano M., Año: 2013). ● Genome sequence and comparative analysis of Avibacterium paragallinarum. Artículo en revista científica (Autor Principal: Requena D. Año: 2013). ● Structure-Activity relationship in mutated pyrazinamidases from Mycobacterium tuberculosis. Artículo en revista científica, Año: 2011. ● Immunoinformatics prediction of linear epitopes from Taenia solium TSOL18. Artículo en revista científica, Año: 2011. ● Trypanocidal properties, structure-activity relationship and computational studies of Quinoxaline 1,4-di-n-oxide derivatives. Artículo en revista científica, año: 2011.

5.4.2. Expertos peruanos desarrolladores de aplicaciones para computación de alto desempeño

El segundo caso es un listado de al menos 5 investigadores peruanos que ya desarrollan aplicaciones informáticas que requieren de grandes capacidades computacionales (capacidades de supercomputación). En algunos casos estos desarrolladores tienen capacidades para crear aplicaciones informáticas para apoyar problemáticas de diversas ramas científicas (bioinformática, física, química computacional, hidrología, biología, geografía, meteorología etc.), en otros se han especializado en algoritmos computacionales para responder diversas inquietudes de la propia relacionadas a ciencia de la computación como: inteligencia artificial, redes neuronales, procesamiento de imágenes, visión computacional, minería de datos, etc. En este caso es importante destacar que estos científicos fueron formados en su mayoría en universidades brasileñas, por lo que se debe reconocer el aporte del sistema universitario brasileño en la computación peruano, pues incluso estos investigadores lideran la mayoría de los pocos proyectos existentes a nivel del Perú. Desde Universidades como la UNSA. Universidad Católica San Pablo (ambas de Arequipa) y a propia Universidad Nacional de Ingeniería- UNI. Como en el caso anterior los datos presentados fueron extraídos principalmente de la base de datos del DINA de CONCYTEC. Seguidamente presentamos un cuadro por cada investigador:

Nombres y apellidos	Cristian Jose Lopez Del Alamo	
Línea de investigación	Inteligencia Artificial	
Datos de contacto	criloal23@gmail.com	
Centro laboral actual	Universidad la Salle, Universidad Nacional de San Agustín	
Ejerce docencia en	Universidad la Salle, Universidad Nacional de San Agustín	
Estudios realizados	Dr. en ciencias de la computación, con Felicitaciones públicas, por la Universidad Nacional de San Agustín, y Magister en Ingeniería de Software con Máxima distinción por la Universidad de Tarapacá de Chile.	
Trabajos en HPC	<ul style="list-style-type: none"> ● Desarrollo de una plataforma de clasificación de imágenes por contenido como soporte computarizado al diagnóstico y telerradiología, para acelerar el proceso de pre-selección de normalidad en estudios de patologías cerebrales en imágenes médicas. Investigador Principal: Yvan Jesus Tupac Valdivia. Duración: Enero 2011 a febrero 2014. ● Plataforma de Cálculo Estocástico de Rutas Óptimas para Transporte Público y Privado. Éste proyecto tiene como objetivo el desarrollo de una plataforma de software para brindar información de rutas que conecten diferentes puntos de la ciudad determinados por el usuario. La ruta calculada deberá incluir tanto desplazamientos a pie como con vehículos motorizados -- de transporte público y privado. Investigador Principal: Ernesto Cuadros Vargas. Duración: Marzo 2009 a febrero 2011. ● Circulo de Investigación en Computación de Alto Desempeño con énfasis en el desarrollo de Métodos y Técnicas de Minería de Datos de Gran Escala para el apoyo en investigaciones de cambio climático. El círculo de investigación tiene como objetivo investigar en patrones de cambio climático en el sur del Perú mediante técnicas de Deep Learning y Visualización científica. Abril 2016 a abril 2018. Cristian Jose Lopez Del Alamo. 	
Publicaciones	<ul style="list-style-type: none"> ● Efficient approach for interest points detection in non-rigid shapes. Artículos científico. Autor principal: Del Alamo C. 2015. ● 3D mesh interest point detection using GISIFs and heat diffusion. Artículos científico. Autor principal: Hurtado J. 2015. ● A novel approach for image feature extraction using HSV model color and niters wavelets. Artículos científico. Autor principal: Del Alamo, C.J.L. 	

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Publicaciones	<ul style="list-style-type: none"> ● Discovery of patterns in software metrics using clustering techniques. Artículos científico. Autor principal: Del Alamo, C.J.L. ● M-CBIR: A medical content-based image retrieval system using metric data-structures. Artículos científico. Autor principal: Chuctaya H.
----------------------	--

Nombres y apellidos	Guillermo Leopoldo Kemper Vasquez	
Línea de investigación	Procesamiento y Codificación de Señales.	
Datos de contacto	guillermo.kemper@gmail.com	
Centro laboral actual	Instituto Nacional de Investigación y Capacitación de Telecomunicaciones.	
Registro REGINA	3522	
Ejerce docencia en	Universidad Nacional De Ingeniería – UNI, Universidad Peruana De Ciencias Aplicadas.	
Estudios realizados	<ul style="list-style-type: none"> ● Ingeniero Electrónico graduado en la Universidad Privada Antenor Orrego de Trujillo-Perú (1994). ● Grados de Doctor (2001) y Magister (1996) en Ingeniería Electrónica con mención en Telecomunicaciones y Telemática en la Universidad Estadual de Campinas - UNICAMP-Brasil. 	
Trabajos en HPC	<ul style="list-style-type: none"> ● Desarrollo de un Codificador Escalable de Imágenes Satelitales Basado en Descomposición de Sub-Bandas Wavelets - (INICTEL-UNI) - Participación: Asesor y parte del equipo de investigadores. Investigador Principal: Joel Telles. Duración: enero 2011 a diciembre 2012. ● Desarrollo de un Software de Procesamiento y Segmentación de Regiones en Imágenes Satelitales - (INICTEL-UNI) - Participación: Asesor y parte del equipo de investigadores. Participación: Asesor y parte del equipo de investigadores. Investigador Principal: Joel Telles. Duración: enero 2012 a Diciembre 2012. ● Sistema Portátil de Adquisición de Imágenes y Cuantificación de Fluorescencia, Orientado a Brindar Información Característica para La Detección del Plasmodium Spp y Diagnóstico Rápido de la Malaria Proyecto que es Parte del Proyecto Macro: "Smartphones feat Smartmolecules" marzo 2015 diciembre 2015 Pohl Milón. ● Desarrollo de un Software de Procesamiento y Extracción de Características de Señales de Sonar (Marina de Guerra del Perú - INICTEL-UNI) - Participación: Investigador Principal. Investigador Principal: Guillermo Leopoldo Kemper Vasquez. Año: 2013. ● Desarrollo de un aplicativo de software para la identificación automática de la vegetación, superficies de agua y zonas urbanizadas en los pantanos de villa basada en el procesamiento de imágenes satelitales. Investigador principal: Guillermo Leopoldo Kemper Vasquez. Duración: agosto 2014 a setiembre 2015. ● Desarrollo de un Aplicativo de Software para el Registro de la Información Catastral de la Cuenca del Rio Santa, basado en el Procesamiento de Imágenes Satelitales - (MINAGRI - ANA - CTIC - UNI). Investigador Principal: Guillermo Kemper. Duración: Enero 2013 setiembre 2015. ● Evaluación del estado eco sistémico de los aguajales en Iquitos empleando imágenes de alta resolución, computación de alto rendimiento y procesamiento digital de imágenes, que facilite la conservación y uso sostenible de la biodiversidad en la Amazonia Peruana. Investigador Principal: Joel Telles. Diciembre 2014 agosto 2016. 	

Trabajos en HPC	<ul style="list-style-type: none"> ● Desarrollo de una plataforma de teledetección basada en computación paralela para el procesamiento de imágenes satelitales, haciendo uso de firmas espectrales y análisis de suelos orientada a incrementar el rendimiento del maíz amarillo duro en Lambayeque. Investigador principal: Daniel Diaz. Duración: diciembre 2014 diciembre 2016.
Publicaciones	<ul style="list-style-type: none"> ● A detection method of ectocervical cell nuclei for pap test images, based on adaptive thresholds and local derivatives. Artículo científico. Autor principal: Oscanoa J. Año: 2015. International Journal of Multimedia and Ubiquitous Engineering. ● Identification and cadastral registration of water bodies through multispectral image processing with multi-layer Perceptron Neural Network. Artículo científico. Autor principal: Dianderas E. Año 2015. ● One Channel Subvocal Speech Phrases Recognition Using Cumulative Residual Entropy and Support Vector Machines. Artículo científico. Autor principal: Chau G. Año: 2015 IEEE Latin America Transactions. ● Landscape Units estimation in WorldView-2 images by using segmented urban areas, green areas and water bodies for monitoring variation/evolution of cities. Artículo científico. Autor principal: Ramirez A. Año: 2015. ● A novel fuzzy logic-based metric for audio quality assessment: Objective audio quality assessment. Artículo científico. Autor principal: León L. Año: 2011. ● An audio compression method based on wavelets subband coding. Artículo científico. Autor principal: Kemper G. Año: 2011. IEEE Latin America Transactions. ● Frequency conversion of digital image for application of SDVPI Conversion de frecuencia de imágenes digitalizadas para aplicaciones en simulación de SDVPI. Artículo científico. Autor Principal: do Nascimento A. Año: 1997. ● Un método automático de extracción de información de imágenes digitales basado en la detección de posicionamiento de marcas ópticas a partir de distancias euclidianas y redes neuronales. Autor principal: Kemper G. Año: 2016. ● Diseño e Implementación de una herramienta de software orientada a la extracción y medición de temperatura, superficies de agua, vegetación y nieve utilizando imágenes satelitales LANDSAT 7. Artículo en congreso. Año 2013. ● Un Algoritmo de Mejora de Imágenes Aéreas Adquiridas Vía UAV Orientado a Mejorar la Detección de Presencia de Aguajes Jóvenes en la Selva del Perú. Artículo en congreso. Año 2015.

Nombre	Alex Jesús Cuadros Vargas	
Línea de investigación	Computación Gráfica	
Datos de contacto	alex@ucsp.pe	
Centro laboral actual	Universidad Católica San Pablo.	
Ejerce docencia en	Universidad Nacional de San Agustín, Universidad Católica San Pablo.	
Estudios realizados	<ul style="list-style-type: none"> ● Doctor en Ciencias Matemáticas y de Computación en la Universidad de Sao Paulo, Brasil. ● Maestría en Ciencias Matemáticas y de Computación en la Universidad de Sao Paulo, Brasil. ● Realizó un Pos-Doctorado [2007-2009] investigando sobre generación de mallas a partir de imágenes, en la Universidad de Sao Paulo. 	

<p>Trabajos en HPC</p>	<ul style="list-style-type: none"> ● Maestría en Ciencia de la Computación, el Programa de Maestría en Ciencia de la Computación dirigido por el Centro de Investigación en Ciencias de la Computación (CICC) de la Universidad Católica San Pablo es el primer programa de esta área de conocimiento del país que se realiza a tiempo completo (stricto-sensu). Cuenta con un cuerpo de docente de primera línea académica, compuesto por profesionales formados en las mejores universidades de Latinoamérica y el Mundo Esta maestría cuenta con el apoyo y respaldo de CONCYTEC. Investigador Principal: Alex Jesús Cuadros Vargas. Duración: Abril 2014 a Marzo 2016. ● Generación de modelos tridimensionales de huesos humanos para fines de exploración y cuantificación ósea. Este proyecto propone la creación de modelos geométricos tridimensionales, a partir de tomografías, donde se distinguida cada hueso por separado. Siendo así, se propone calcular la densitometría de cada hueso por separado, obteniendo resultados equivalentes al de un Densitómetro, sin embargo, de forma focalizada. Como consecuencia, el proceso propuesto pretende evitar el uso de un densitómetro, por ende evitar su irradiación, y obtener resultados equivalentes a partir de imágenes Investigador Principal: Alex Jesús Cuadros Vargas. Duración: Diciembre 2015 a Diciembre 2017. ● Maestría en Ciencia de la Computación El Programa de Maestría en Ciencia de la Computación dirigido por el Centro de Investigación en Ciencias de la Computación (RICS) de la Universidad Católica San Pablo es el primer programa de esta área de conocimiento del país que se realiza a tiempo completo (stricto-sensu). Cuenta con un cuerpo de docente de primera línea académica, compuesto por profesionales formados en las mejores universidades de Latinoamérica y el Mundo Esta maestría cuenta con el apoyo y respaldo de CONCYTEC / FONDECYT. Investigador Principal: Alex Jesús Cuadros Vargas. Duración: Abril 2016 a Marzo 2020. ● Círculos de Investigación y Tecnología Es un proyecto financiado por FONDECYT/CONCYTEC. Este proyecto involucra a las universidades nacionales: Universidad Nacional de San Agustín (UNSA), la Universidad Nacional de Ucayali (UNU), la Universidad Católica San Pablo (UCSP), y a FONDECYT/CONCYTEC. Además cuenta con el apoyo y asesoramiento de investigadores de la Universidad de Sao Paulo (Brasil), Universidad de Minas Gerais (Brasil), Universidad de Nueva York (UNY) y INTEL-USA., con lo que se busca promover la interacción académica. Investigador Principal: ALEX JESÚS CUADROS VARGAS. Duración: Diciembre 2014 a Diciembre 2017.
<p>Publicaciones</p>	<ul style="list-style-type: none"> ● Aceleración en GPU del Proceso de Cálculo de Patrones de Color en Mallas Triangulares Generadas a Partir de Imágenes. Artículo en revista científica. Año: 2013. ● Combinatorial laplacian image cloning. In Proceedings of. Conference on Graphics, Patterns and Images. Artículo en revista científica. Año: 2011. ● Generating segmented meshes from textured color images. Artículo en revista científica. Año: 2009. ● Generating segmented quality meshes from non-processed images. Artículo en revista científica. Año: 2007. ● Generating segmented tetrahedral meshes from regular volume data for simulation and visualization applications. Artículo en revista científica. Año: 2006. ● Beta-connection: Generating a family of models from planar cross sections. Artículo en revista científica. Año: 2005. ● Imesh: An image based quality mesh generation technique. Artículo en revista científica. Año: 2005. ● Bidimensional Delaunay mesh generation from images. Artículo en revista científica. Año: 2004. ● Renderizado de Mallas Tetraedrales Segmentadas. Artículo en revista científica. Año: 2007.

Nombres y apellidos	Álvaro Henry Mamani Aliaga	
Línea de Investigación	Computación de alto desempeño.	
Datos de Contacto	Sin datos.	
Centro laboral actual	Universidad Nacional de San Agustín, Universidad Católica San Pablo.	
Registro REGINA	3886	
Ejerce docencia en	<ul style="list-style-type: none"> ● Universidad Nacional de San Agustín ● Universidad Católica San Pablo. 	
Estudios realizados	<ul style="list-style-type: none"> ● Magister en ciencias de la computación en Universidad De Sao Paulo. ● Ingeniero de sistemas de la Universidad Nacional De San Agustín. 	
Trabajos en HPC	<ul style="list-style-type: none"> ● Laboratorio de Computación de Alto Desempeño: El proyecto tiene como objetivo la adquisición e implantación de un equipamiento de alto desempeño para la Universidad Nacional de San Agustín. Este equipamiento debe contemplar las propiedades para soportar Computación de Alto Desempeño. Con este equipamiento se proyecta ofrecer recursos computacionales para problemas que requieran: Alto Desempeño; alto procesamiento; alta disponibilidad; escalabilidad; entre otros. Investigador Principal: Luis Alfaro Casas, Duración: noviembre 2013 a marzo 2015. ● Teledetección de humedales usando imágenes satelitales y computación de alto desempeño La idea fundamental de este proyecto es crear una interfaz para que aplicaciones del análisis de imágenes satelitales puedan acceder a un conjunto de funcionalidades de uso intensivo de cómputo y datos. Esta herramienta se evaluara con imágenes locales de la región de Arequipa, lo que permitirá analizar, monitorear, etc. Los humedales encontrados en dicha región, y pudiendo extenderse a diferentes regiones del país. Investigador Principal: Pablo Raul Yanyachi Aco-Cardenas, Duración: marzo 2015 noviembre 2016. ● A comparative study on Task Dependent Scheduling Algorithms for Grid Computing The large amount of data processing is growing fast. In many scenarios, both in industry and academia, high-performance computing (HPC) is a needed to process Big Data in short time. An important topic in HPC is task scheduling. In this project is presented a comparative study designed to discover the behavior of scheduling algorithms exhibit under different settings. This was done through a methodology with four criteria: performance, scalability, workload distribution and adaptability. Investigador Principal: Alvaro Henry Mamani Aliaga, Duración: marzo 2009 a Agosto 2011. 	
Publicaciones	<ul style="list-style-type: none"> ● Pytos: A Framework for Mobile Computation Offloading in Python. Autor Principal: Mendoza E. Año: 2016. ● A new approach to the massive processing of satellite images. Autor Principal: Nina W. Año: 2015. ● HIPI, as alternative for satellite images processing Autor Principal: Choquehuayta W. Año: 2015. ● Empowering mobile users: Create your own mobile application for data collection in the cloud. Autor Principal: Da Conceição A. Año: 2014. ● Open architecture for mobile data collection using cloud computing. Autor Principal: Da Conceicao, Autor Principal: A.F. Año: 2013. ● A comparative study on task dependent scheduling algorithms for grid computing. Autor Principal: Mamani-Aliaga, A.H. Año:2012 ● Estudio comparativo de algoritmos de escalonamiento para grades computacionais. Artículo en congreso. Año: 2011 ● Artículo en congreso A Comparative Study On Task Dependent Scheduling Algorithms For Grid Computing. Artículo en congreso. Año: 2012. 	

Nombre	Alexander Victor Ocsa Mamani	
Línea de Investigación	Redes Neuronales	
Datos de Contacto	alexaooca.cs@gmail.com	
Centro Laboral Actual	TVARE I+D	
Ejerce Docencia en	Universidad Nacional de San Agustín	
Estudios Realizados	<ul style="list-style-type: none"> ● Magister en ciencias de la computación en Universidad de Sao Paulo. ● Ingeniero de Sistemas de la Universidad Nacional de San Agustín. 	
Trabajos en HPC	<ul style="list-style-type: none"> ● Sistema automático de diagnóstico de parásitos intestinales a través de micrografías. El presente proyecto formula una propuesta de desarrollo de un sistema de diagnóstico automático de parásitos, cuyo propósito es implementar una nueva técnica que permita el diagnóstico de parásitos intestinales de humanos y animales, el mismo que será realizado mediante el análisis de imágenes de muestras fecales. Para ello será necesario desarrollar técnicas computacionales que permitan visualizar y procesar las imágenes microscópicas. Investigador Principal: Cesar Beltran Castañón. Año 2011. ● AgroDataMine: Desenvolvimento de Métodos e Técnicas de Mineração de Dados para Apoiar Pesquisas em Mudanças Climáticas com Ênfase em Agrometeorologia. The objective of the project is to understand the correlation between various climate and agriculture analysis parameters and to improve agro climate models, evaluating and cross-checking terabytes of data collected by sensors placed in the soil, by meteorological radars, and by satellites. The goal is to create models and mathematical algorithms that will enable us to identify tendencies and establish correlations in this data to help agro meteorologists make more accurate forecasts. Investigador Principal: Agma Juci Machado Traina. Duración: Enero 2009 a Diciembre 2012. ● Circulo de Investigación en Computación de Alto Desempeño con énfasis en el desarrollo de Métodos y Técnicas de Minería de Datos de Gran Escala para el apoyo en investigaciones de cambio climático. El creciente aumento de datos y su disponibilidad en diferentes áreas y ámbitos ha motivado el desarrollo de diversas técnicas para el descubrimiento de conocimiento sobre grandes volúmenes de datos complejos. A pesar de las recientes investigaciones en métodos y técnicas de minería de datos, su aplicación a gran escala aún es limitada debido a su alto costo computacional. Para el desarrollo de soluciones de alto desempeño en tareas que requieren enormes recursos computacionales. Investigador Principal: Cristian Lopez Del Alamo. Duración: Enero 2016 a enero 2019. 	
Publicaciones	<ul style="list-style-type: none"> ● Cloth simulation using AABB hierarchies and GPU parallelism. Autor Principal: Salazar F. Año 2011. ● Sublinear querying of realistic timeseries and its application to human motion. Autor Principal: Florez O. Año: 2010. ● DBM-Tree: An efficient metric acces method. Autor Principal: Ocsa A. Año: 2007. ● DB-GNG: A constructive Self-Organizing Map based on density. Autor Principal: Ocsa A. Año: 2007. 	

5.4.3. Usuarios del Centro de Alto Rendimiento Computacional de la Amazonía Peruana

El 27 de enero del 2017 se realizó el lanzamiento del Centro de Alto Rendimiento Computacional de la Amazonía Peruana con sede en el Instituto de Investigaciones de la Amazonía Peruana. El mismo que cuenta con un supercomputador con arquitectura clúster, denominado MANATI, compuesto por 10 equipos de alta velocidad. 6 de estos equipos poseen aceleradores gráficos (GPU). Esto fue posible gracias al apoyo financiero de CIENCIACTIVA del CONCYTEC.

El MANATI es un recurso público peruano, por lo cual es accesible a investigadores e incluso estudiantes peruanos. Teniendo en cuenta que uno de los objetivos de este centro es apoyar la penetración de la supercomputación en la ciencia peruana, y este tipo de recursos son escasos en las instituciones peruanas.

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

De los usuarios existentes, y de acuerdo a sus necesidades podemos mencionar que existen hasta 3 tipos:

- En el primer caso los usuarios Expertos que tienen proyectos y que necesitan las capacidades de cómputo para concluir sus estudios.
- En el segundo caso los usuarios En formación, que vienen trabajando en actividades académicas como los trabajos de tesis (principalmente posgrado) y propios trabajos de las asignaturas.
- En el tercer casos los usuarios Docentes, que utilizan las supercomputadoras como instrumento para hacer demostraciones en sus aulas.

Seguidamente presentamos un pequeño cuadro con 20 usuarios del MANATI IIAP (en el listado se ha priorizado investigadores externos a la Amazonía peruana).

Tabla 33: Lista de usuarios del MANATI IIAP hasta marzo de 2017.

N°	Nombres y Apellidos	Línea de Investigación	Institución	Resumen de trabajos
1	Alexander Ocsa	Redes neuronales HPC	Universidad Nacional de San Agustín Arequipa	Deep Learning y Base de Datos en GPU
2	Cesar Marchelo Bragagnini Mendizabal	Redes neuronales, Visión computacional.	Universidad Católica San Pablo	Deep Learning, Natural Language Processing, Image Caption, Sentiment Analysis, Computer Vision
3	César Martín Cruz Salazar	Matemática	Facultad de Ciencias - Universidad Nacional de Ingeniería	Ejecución de algoritmos matemáticos para cálculo de primos y valores propios de matrices de 400 x 400
4	Christian solis Calero	Química Computacional	Universidad Nacional Mayor de San Marcos	El trabajo que pretendo ejecutar está relacionado a mi tesis para obtener el grado académico de Doctor en Farmacia y Bioquímica de la Facultad de Farmacia y Bioquímica de la Universidad Nacional Mayor de San Marcos, que se encuentra registrada en la Unidad de Post Grado de la Institución con el título de "Diseño computacional de nuevos inhibidores de la isocitrato liasa del Mycobacterium tuberculosis". Los cálculos que realizaría en el clúster son de las energías libres de unión de la isocitrato liasa unida a inhibidores de su actividad bioquímica por simulaciones de dinámica molecular (DM).
5	Cristian José López Del Alamo	Inteligencia Artificial	Universidad (Arequipa) La Salle	Desarrollo de una estructura de datos persistente para el manejo temporal de Triangulaciones de Delaunay.
6	David Alfonso Laván Quiroz	Bioinformática	INICTEL-UNI	Estudios de sitios de unión a factores de transcripción
7	Disraeli Fausto Ari Mamani	Minería de datos	Universidad de San Arequipa Nacional Agustín	Trabajos en procesamiento de lenguaje natural, específicamente el procesamiento en "Spanish Billion Words Corpus and Embeddings" para la tarea específica de Análisis de sentimientos ya que operar esta gran cantidad de datos, necesita una arquitectura más grande para realizar las pruebas necesarias al algoritmo en la tesis "Twitter sentiment analysis with conditional random fields and word embeddings"
8	Disraeli Fausto Ari Mamani	Minería de datos	Universidad de San Arequipa Nacional Agustín	Investigación relacionada al Análisis de Sentimientos en cadenas de textos de Twitter, usando para ello word embeddings, que son representaciones de palabras en forma de vector, se requiere hacer clusters y binarización de los mismos para lo que se necesita un mayor procesamiento, y luego realizar el procesamiento en un clasificador probabilístico para obtener los resultados, proceso que también requiere un gran procesamiento.

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

9	Elisban Flores Quenaya	Bioinformática	Universidad del Altiplano/ Universidad Católica San Pablo	Evaluación feno-genotípica de alpacas
10	Gerardo Cristian Jacome Vergaray	Meteorología	SENAMHI	Regionalización dinámica del modelo HadGEM2-ES en alta resolución, utilizando el modelo WRF a 5 km de resolución espacial, simulado a nivel nacional, para evaluar los cambios de lluvia y temperatura del aire al 2065
11	Jorge Ibáñez Gijón	Inteligencia Artificial	Aix-Marseille Universite	Modelos biomecanicos de control del movimiento humano
12	Jose Antonio Fiestas Iquira	Astrofísica	Universidad de Ingeniería	Simulaciones de N-cuerpos
13	Josimar Edinson Chire Saire	Computación bioinspirada	Universidad Católica San Pablo	Experimentos de algoritmo evolutivo de inspiración cuántica con funciones benchmark y alta dimensionalidad
14	Julio Cesar Mendoza Bobadilla	Redes neuronales	University of Campinas	La clasificación de nódulos pulmonares es uno de los principales tópicos en sistemas de ayuda al diagnóstico (CAD), y se ha abordado el problema con muchos métodos. En este estudio abordamos el problema utilizando Deep Learning, específicamente redes neuronales convolucionales. El objetivo del estudio es analizar métodos para entrenamiento de Redes Neuronales Convolucionales utilizando funciones de pérdida que consideran tanto objetivos supervisados (clasificación) como no supervisados (reconstrucción) para clasificación de nódulos pulmonares en imágenes radiográficas. El análisis involucra la evaluación de métodos del estado del arte, así como la búsqueda de hiper-parámetros.
15	Julio Rubén Barrios Llacuchaqui	Química Computacional	Universidad Nacional de Ingeniería	Cálculos de Química Computacional
16	Junior Usca Huacasi	Minería de datos	Universidad Nacional de San Agustín	Análisis de datos geoespaciales distribuidos de tweets
17	Manuel Castillo- Cara	Inteligencia artificial	Universidad Nacional de Ingeniería	Para la localización en indoor tenemos 65000 posibles combinaciones de potencias y posiciones, estas combinaciones tardan aproximadamente 2 días.
18	Rayner Harold Montes Condori	Procesamiento de imágenes	Universidad de São Paulo	El proyecto se centra principalmente en el análisis de imágenes de la superficie de las hojas y en las secciones histológicas de los órganos de las hojas, tales como estomas y vasos. La principal motivación de este estudio es contribuir a las técnicas de reconocimiento de patrones del estado del arte, aplicados a la identificación y evaluación nutricional de algunas especies de plantas. Las imágenes biológicas presentan una no linealidad intrínseca que resulta de su naturaleza no trivial. Nuestro grupo de investigación ha estado implementando soluciones en problemas relacionados durante casi una década a través del uso de redes complejas y análisis fractal. Se utilizará los GPUs del cluster en el desarrollo y aplicación de técnicas de aprendizaje profundo en el análisis de imágenes de plantas microscópicas y macroscópicas.
19	Rel Guzmán Apaza	Redes neuronales	Universidad Nacional de San Agustín	Training & hyperparameter tuning de algoritmos de Machine Learning
20	Ricardo Rildo Coronado Pérez	Redes neuronales	Universidad Nacional de San Agustín	Deep Learning en imágenes médicas

Fuente: IIAP

CAPÍTULO 6

PRIMEROS PASOS EN PROGRAMACIÓN PARA COMPUTACIÓN DE ALTO RENDIMIENTO

El primer paso para empezar a desarrollar aplicaciones para supercomputadoras es atreverse a generar y ejecutar código. En el contexto de este documento hemos considerado interesante incluir ejemplos básicos de programación, basado en la experiencia del Centro de Alto Rendimiento Computacional de la Amazonía Peruana del IIAP.

Es importante resaltar que para crear y ejecutar ejemplos básicos de Supercomputación no es necesario disponer de un Supercomputador, pues incluso la mayoría de computadoras personales de escritorio o portátil como las que tienen procesadores Intel I3, I5 e I7 (e incluso otras que ya vienen con tarjetas gráficas NVIDIA) vienen con más de 2 núcleos y tienen capacidades para ejecutar programación paralela. En este caso es clave identificar el software a utilizar. Por experiencias propias antes que la codificación de algoritmos, muchas veces más complejo es la configuración del software o de las propias librerías.

Para hacer más fácil a los programadores interesados en comenzar a programar, se han incluido ejemplos de codificación tipo “Hola Mundo” para equipos con sistemas operativos Windows y Linux, utilizando el siguiente software (más información del software en el capítulo 3)

- TBB
- MPI
- OpenMP
- OpenCL
- CUDA

6.1. Ejemplos para sistema operativo Windows:

Los ejemplos tipo “Hola mundo” en paralelo desarrollados a continuación con las diferentes Apis y librerías para paralelizar, serán desarrollados para que se ejecuten en plataformas x86/x64.

Pre-requisitos para los ejemplos

- a) Tener Instalado Visual Studio 2012
- b) Conocimiento básico de programación C++

6.1.1. Hola Mundo en TBB (Threading Building Blocks) para Windows

a. Instalación de Apis y librerías

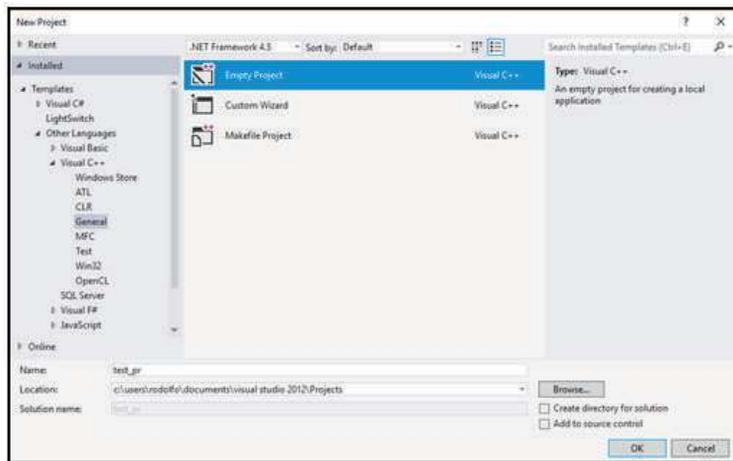
- i. Descargar Librería TBB desde la página oficial para Windows
<https://www.threadingbuildingblocks.org/download#stable-releases>

El archivo descargado ser como el siguiente tbb2017_20170226oss_win.

ii. Descomprimos el archivo en la carpeta de su preferencia, en este caso esta C:\TBB\tbb2017_20170226oss

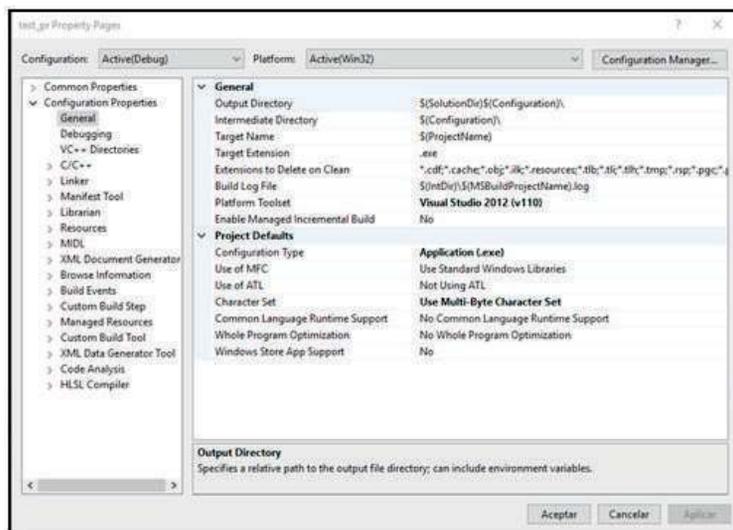
b. Creación y configuración del entorno para programar en paralelo

i. Crear un proyecto vacío de C++ en Visual Studio

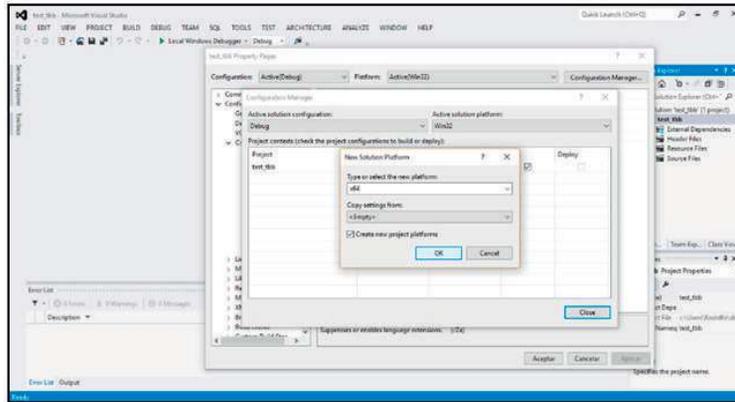


ii. Configurar el entorno del proyecto

1. Ingresar a propiedades del proyecto



2. Configurar la plataforma de 32 a 64



3. En la propiedad "C/C++ -> General -> Additional Include directorios", agregar "C:\TBB\tbb2017_20170226oss\include"
4. En la propiedad "Linker -> All Options -> Additional Dependencies", agregar tbb_debug.lib;tbb.lib
5. En la propiedad "Linker -> All Options -> Additional Library Directories", agregar C:\TBB\tbb2017_20170226oss\lib\intel64\vc11
6. En la propiedad "Build Events -> Post-Build Event " agregar en:

a. Command Line:

```
COPY "C:\TBB\tbb2017_20170226oss\bin\intel64\vc11\tbb.dll" "$(SolutionDir)\$(Platform)\$(Configuration)\\"
COPY "C:\TBB\tbb2017_20170226oss\bin\intel64\vc11\tbb_debug.dll" "$(SolutionDir)\$(Platform)\$(Configuration)\\"
```

Seleccionar Aceptar

b. Description: Copiar tbb.dll y tbb_debug.dll en Carpeta de salida

iii. Crear archivo hola_mundo.cpp en "Source Files", donde se escribirá el programa de hola mundo utilizando TBB.

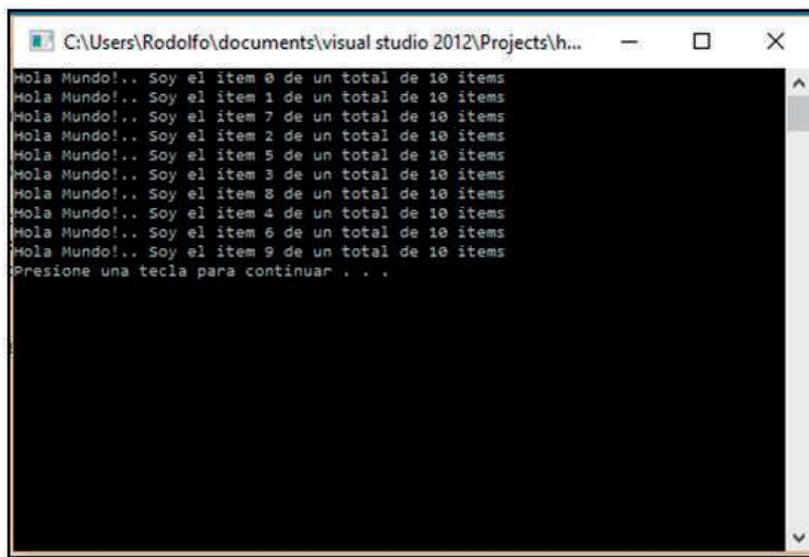
iv. Escribir el siguiente código.

```
#include "tbb/tbb.h"
#include <stdio.h>
using namespace tbb;
```

```
int main( )
{
    int dim=10;
    tbb::parallel_for(0,dim,[&](int i){
        printf("Hola Mundo!.. Soy el item %d de un total de %d items \n",i,dim);
    });

    system("PAUSE");
}
```

v. Compilar y ejecutar el programa



```
C:\Users\Rodolfo\documents\visual studio 2012\Projects\h...
Hola Mundo!.. Soy el item 0 de un total de 10 items
Hola Mundo!.. Soy el item 1 de un total de 10 items
Hola Mundo!.. Soy el item 7 de un total de 10 items
Hola Mundo!.. Soy el item 2 de un total de 10 items
Hola Mundo!.. Soy el item 5 de un total de 10 items
Hola Mundo!.. Soy el item 3 de un total de 10 items
Hola Mundo!.. Soy el item 8 de un total de 10 items
Hola Mundo!.. Soy el item 4 de un total de 10 items
Hola Mundo!.. Soy el item 6 de un total de 10 items
Hola Mundo!.. Soy el item 9 de un total de 10 items
Presione una tecla para continuar . . .
```

6.1.2. Hola Mundo en MPI (Message Passing Interface) [75]

En el ejemplo se utilizará el MS-MPI [76].

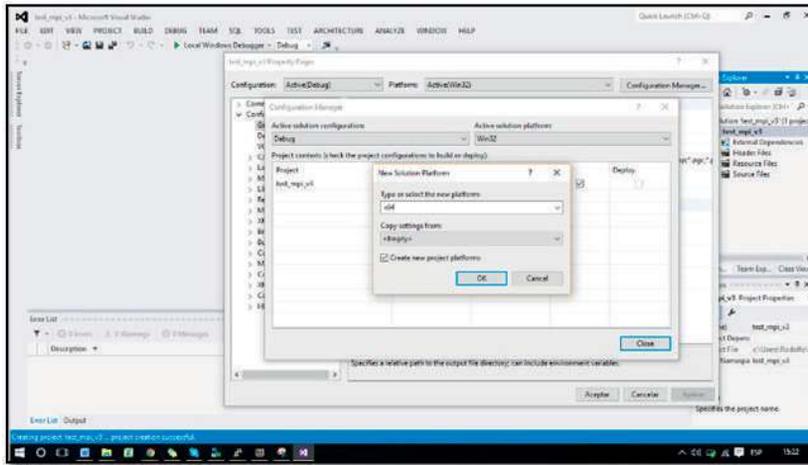
a. Instalación de Apis y librerías

- i. Descargar e instalar “msmpisdsk.msi” y “MSMpiSetup.exe”
<http://go.microsoft.com/fwlink/p/?LinkID=389556>
- ii. Comprobar si la instalación fue correcta escribiendo en consola lo siguiente “set MSMPI” y debe mostrar los siguiente:
 1. MSMPI_BIN=C:\Program Files\Microsoft MPI\Bin\
 2. MSMPI_INC=C:\Program Files (x86)\Microsoft SDKs\MPI\Include\
 3. MSMPI_LIB32=C:\Program Files (x86)\Microsoft SDKs\MPI\Lib\x86\
 4. MSMPI_LIB64=C:\Program Files (x86)\Microsoft SDKs\MPI\Lib\x64\

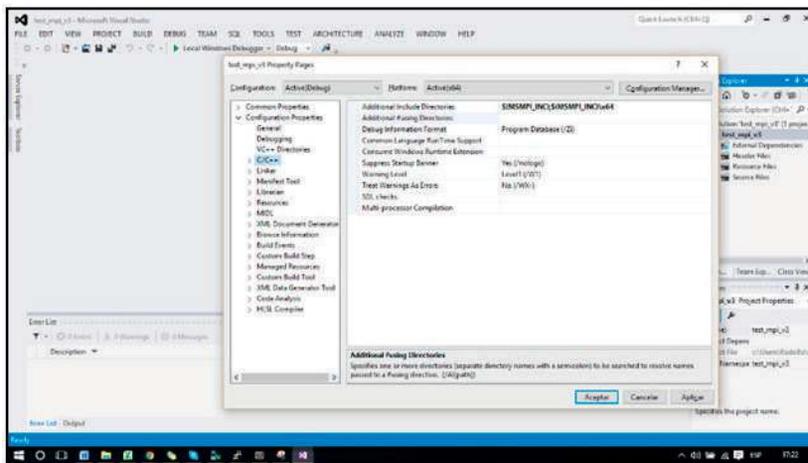
b. Creación y configuración del entorno para programar en paralelo

- i. Crear un proyecto vacío de c++ en Visual Studio
- ii. Configurar el entorno del proyecto:

- 1. Ingresar a propiedades del proyecto
- 2. Configurar la plataforma de 32 a 64



- 3. En la propiedad "C/C++" agregar en additional include directories $\$(MSMPI_INC); \$(MSMPI_INC)\x64$. Si está trabajando para un entorno de 32 bits, agregar $\$(MSMPI_INC); \$(MSMPI_INC)\x86$.



- 4. En la propiedad "Linker ->All Options", agregar en "Additional Dependencies" msmpl.lib, y en Additional Library Directories $\$(MSMPI_LIB64)$. Si está trabajando para un entorno de 32 bits, agregar $\$(MSMPI_LIB32)$

- 5. Seleccionar Aceptar


```
MPI_Init(&argc,&argv);
MPI_Comm_size( MPI_COMM_WORLD, &numprocs ); // devuelve el numero de procesos
en este COMM_WORLD
MPI_Comm_rank( MPI_COMM_WORLD, &myid ); // identificate por el myid asignado

printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos \n",myid,numprocs);

MPI_Finalize();

return 0;
}
```

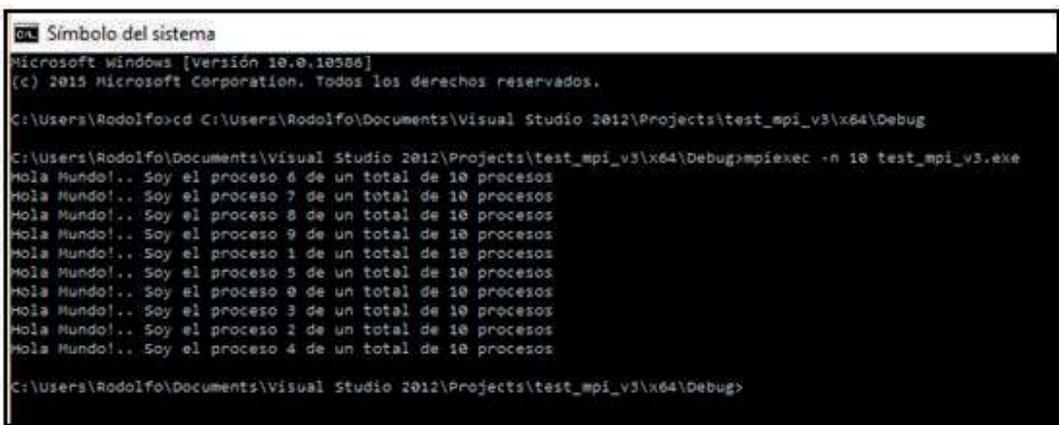
8. Compilar el programa para generar los ejecutables

9. Ejecutar el ejecutable generado desde línea de comandos

a. Abrir consola CMD

b. Ir a Tu_Ruta\Visual Studio 2012\Projects\test_mpi_v3\x64\Debug

c. Ejecutar `mpiexec -n 10 test_mpi_v3.exe`



```
Simbolo del sistema
Microsoft Windows [Versión 10.0.10586]
(c) 2015 Microsoft Corporation. Todos los derechos reservados.

C:\Users\Rodolfo>cd C:\Users\Rodolfo\Documents\Visual Studio 2012\Projects\test_mpi_v3\x64\Debug

C:\Users\Rodolfo\Documents\Visual Studio 2012\Projects\test_mpi_v3\x64\Debug>mpiexec -n 10 test_mpi_v3.exe
Hola Mundo!.. Soy el proceso 6 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 7 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 8 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 9 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 1 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 5 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 0 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 3 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 2 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 4 de un total de 10 procesos

C:\Users\Rodolfo\Documents\Visual Studio 2012\Projects\test_mpi_v3\x64\Debug>
```

10. Como se observa en la salida se tiene 10 hola mundos de diferentes procesos que invocamos.

6.1.3. Hola Mundo en OpenMP (Open Multi-Processing)

a. Instalación de Apis y librerías [77]

i. OpenMP viene instalado por defecto en Visual Estudio, y no requiere librerías adicionales

b. Creación y configuración del entorno para programar en paralelo

i. Crear un proyecto vacío de c++ en Visual Estudio

ii. Configurar el entorno del proyecto

1. Ingresar a propiedades del proyecto
2. En la propiedad "C/C++" habilitar la opción de Open MP Support (Yes/openmp) en "Language".
3. Seleccionar "Aceptar"
- iii. Crear archivo hola_mundo.cpp en "Source Files", donde se escribirá el programa de hola mundo en open MP.
- iv. Escribir el siguiente código

```
#include <stdio.h>
#include <stdlib.h>
#include <omp.h>

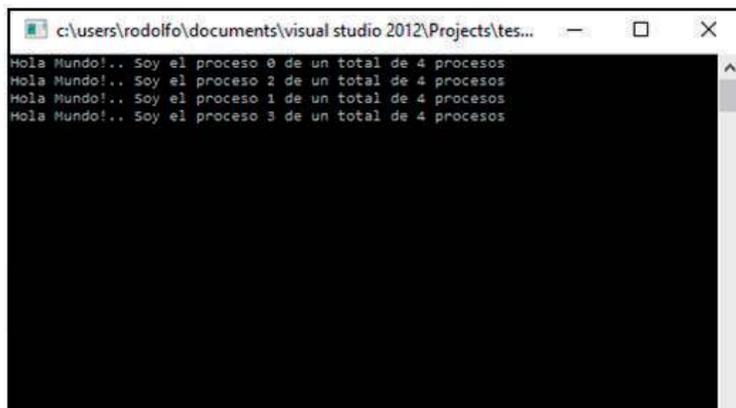
int main(int argc, char **argv)
{
    int p=0, id_hilo, id_proceso;

    p=omp_get_max_threads();
    omp_set_num_threads(p);
    #pragma omp parallel
    {
        id_hilo=omp_get_thread_num();
        id_proceso=omp_get_num_procs();
        printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos
\n",id_hilo,id_proceso);
    }

    system("PAUSE");

    return 0;
}
```

- v. Compilar y ejecutar el programa



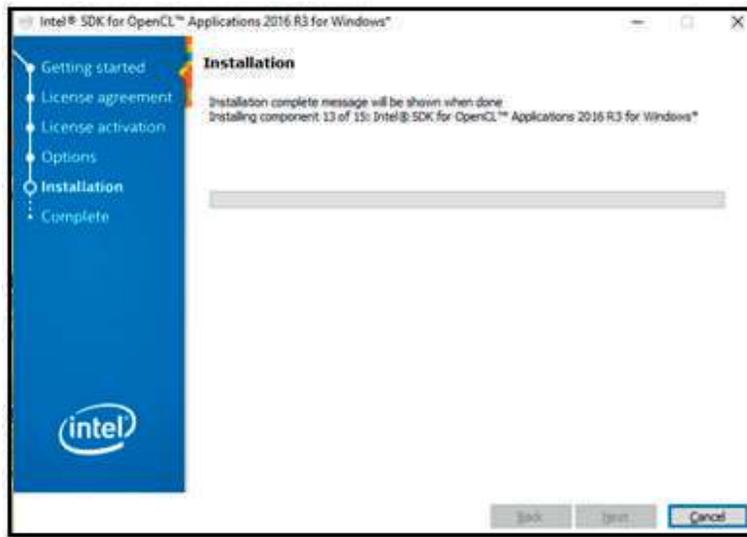
```
c:\users\rodolfo\documents\visual studio 2012\Projects\tes...
Hola Mundo!.. Soy el proceso 0 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 2 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 1 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 3 de un total de 4 procesos
```

6.1.4. Hola Mundo en OpenCL (Open Computing Language) [78]

a. Instalación de Apis y librerías [79]

i. Open Cl, permite realizar programación paralela en diversos tipos de procesadores [80] tanto en CPU (Intel, AMD) como en GPU (NVIDIA, AMD), para ello se debe contar con los SDKs de acuerdo a la marca del procesador. Y en esta oportunidad se trabajaremos con INTEL

ii. Descargar e instalar Intel® SDK for OpenCL™ Applications
<https://software.intel.com/en-us/intel-openc1>



iii. Configurar el entorno del proyecto

1. Ingresar a propiedades del proyecto
2. Configurar la plataforma de 32 a 64
3. En la propiedad "C/C++ -> General -> Additional Include directories", agregar I "C:\Program Files (x86)\Intel\OpenCL SDK\6.3\include"
4. En la propiedad "Linker -> All Options -> Additional Dependencies", agregar OpenCl.lib
5. En la propiedad "Linker -> All Options -> Additional Library Directories", agregar C:\Program Files (x86)\Intel\OpenCL SDK\6.3\lib\x64

iv. Creamos los archivos código fuente necesarios para Open CL:

1. Crear archivo hola_mundo.cpp en "Source Files", donde se escribirá el programa de hola mundo para ejecutar Hola Mundo de Open CL, escribir el siguiente código:

```
#include <stdio.h>
#include <stdlib.h>
#include <CL/cl.h>
```

//Para el tamaño de la matriz de string que contendra el codigo fuente del kernel

```

#define MAX_SOURCE_SIZE (0x100000)

int main(void) {

    // Carga del código fuente del kernel en la matriz cod_kernel
    FILE *fp;
    char *cod_kernel;
    size_t source_size;

    fp = fopen("kernel_hola_mundo.cl", "r");
    if (!fp) {
        fprintf(stderr, "Error al cargar el kernel.\n");
        exit(1);
    }
    cod_kernel = (char*)malloc(MAX_SOURCE_SIZE); //inicializamos la matriz de
string y reservamos un espacio de memoria
    cod_kernel[MAX_SOURCE_SIZE+1]='\0';
    source_size = fread( cod_kernel, 1, MAX_SOURCE_SIZE, fp);
    fclose( fp );

    cl_int ret;
    int j;

    //Para Imprimir la Plataforma seleccionada
    cl_platform_id *platform_id;
    char* info;
    size_t infoSize;
    cl_uint ret_num_platforms;

    ret = clGetPlatformIDs(5, NULL, &ret_num_platforms);

    printf("\n Numero de plataformas detectadas: %d\n",ret_num_platforms);

    platform_id = (cl_platform_id *)malloc(sizeof(cl_platform_id) * ret_num_platfor-
ms);

    ret = clGetPlatformIDs(ret_num_platforms, platform_id, NULL);

    // Para obtener el tamaño del nombre de la plataforma y reservar espacio en memo-
ria
    clGetPlatformInfo(platform_id[0], CL_PLATFORM_NAME, 0, NULL, &infoSize);
    info = (char*) malloc(infoSize);
    // Para obtener el nombre de la plataforma
    clGetPlatformInfo(platform_id[0], CL_PLATFORM_NAME, infoSize, info, NULL);
    printf("\n Plataforma seleccionada: %s", info);

    //Para Imprimir la Dispositivo seleccionado

```

```

cl_device_id device_id=NULL;
cl_uint ret_num_devices=1;
//Aca definimos el tipo de dispositivo habilitada para la plataforma.
//Para este ejemplo con CL_DEVICE_TYPE_CPU, devuelve Intel Intel(R) Core(TM)
i7-7500U si la plataforma es Intel.
//Y con CL_DEVICE_TYPE_GPU, devuelve GeForce GTX 950M si es NVIDIA
ret = clGetDeviceIDs( platform_id[0], CL_DEVICE_TYPE_CPU, 1,&device_id, NULL);

        char* value;
        size_t valueSize;
// Para obtener el tamaño del nombre del dispositivo y reservar espacio en memoria
clGetDeviceInfo(device_id, CL_DEVICE_NAME, 0, NULL, &valueSize);
value = (char*) malloc(valueSize);
// Para obtener el nombre del dispositivo
clGetDeviceInfo(device_id, CL_DEVICE_NAME, valueSize, value, NULL);
printf("\n Dispositivo seleccionado: %s \n\n",value);
free(value);

//Creamos el contexto para manejar los objetos y recursos de OPENCL
cl_context context = clCreateContext( NULL, 1, &device_id, NULL, NULL, &ret);

if (ret != CL_SUCCESS)
{
    printf( "\n->No se pudo crear el contexto de la CPU");

}
else
{
    printf("\n->Se creo el Contexto de la CPU con exito");

}

//Creamos una cola de comandos que nos permitira comunicarnos del HOST al Dispositivo
seleccionado
cl_command_queue command_queue = clCreateCommandQueue(context, device_id, 0,
&ret);
if (command_queue == NULL)
{
    printf("\n->Error al crear comandoQueue para dispositivo 0");
    exit(0);

}
else
{
    printf("\n->Se creo commandQueue con exito");

}
//Creamos el objeto programa que contiene el codigo fuente del Kernel
cl_program program = clCreateProgramWithSource(context, 1, (const char **)&co-
```

```

d_kernel, (const size_t *)&source_size, &ret);

    if (program == NULL)
    {
        printf("\n->No se pudo crear el programa de CL desde el origen");
    }
    else
    {
        printf("\n->Se creo el programa de CL con exito\n");
    }

//Compilamos y enlazamos el programa creado al Dispositivo seleccionado
ret = clBuildProgram(program, 1, &device_id, 0, 0, 0);

if (ret != CL_SUCCESS)
{
    // Determine the reason for the error
    char buildLog[32000];
    clGetProgramBuildInfo(program, device_id, CL_PROGRAM_BUILD_LOG, sizeof(
f(buildLog), buildLog, NULL);
    printf("\nError en el kernel BUILD PGM: %s\n ",buildLog);

}

//Instanciamos el kernel definido en al codigo fuente del kernel que se encuentra compila-
do en el programa CL
cl_kernel kernel = clCreateKernel(program, "hello", &ret);

if (ret != CL_SUCCESS)
{
    printf("Error en: clCreateKernel (falla en la función del kernel):\n");

}

//Definimos la cantidad de grupos y los hilos a utilizar del dispositivo seleccionado
size_t global_work_size = 1;
size_t global_item_size = 10;

//Ejecutamos el kernel con la cantidad de hilos definidos anteriormente
ret = clEnqueueNDRangeKernel(command_queue, kernel, global_work_size, NULL,
&global_item_size, NULL, 0, NULL, NULL);

if(ret== CL_SUCCESS) {
    //printf("\n\nNO Error in clEnqueueNDRangeKernel\n");
}

else {

```

```

        printf("\n->Error clEnqueueNDRangeKernel\n");
    }

    //Liberamos los objetos creados
    ret = clFlush(command_queue);
    ret = clFinish(command_queue);
    ret = clReleaseKernel(kernel);
    ret = clReleaseProgram(program);
    ret = clReleaseCommandQueue(command_queue);
    ret = clReleaseContext(context);

    system("PAUSE");

    return 0;
}

```

2. Crear archivo `kernel_hola_mundo.cl` el cual tendrá el código kernel de Open CL, escribir el siguiente código.

```

#pragma OPENCL EXTENSION cl_amd_printf : enable
__kernel void hello()
{
    int i = get_global_id(0);
        int dim = get_global_size(0);
        printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos \n",i+1,-
dim);
}

```

v. Compilar y ejecutar el programa

```

Seleccionar c:\users\r-cardenas-na\documents\visual studio 2012\Projects\test_openc1(v64)\Debug\test_openc1.exe
Numero de plataformas detectadas: 3
Plataforma seleccionada: Intel(R) OpenCL
Dispositivo seleccionado: Intel(R) Core(TM) i7-7500U CPU @ 2.70GHz

->Se creo el Contexto de la EPU con exito
->Se creo commandQueue con exito
->Se creo el programa de CL con exito
1 warning generated.
Hola Mundo!.. Soy el proceso 1 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 2 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 7 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 8 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 5 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 6 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 3 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 4 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 9 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 10 de un total de 10 procesos
Presione una tecla para continuar . . .

```

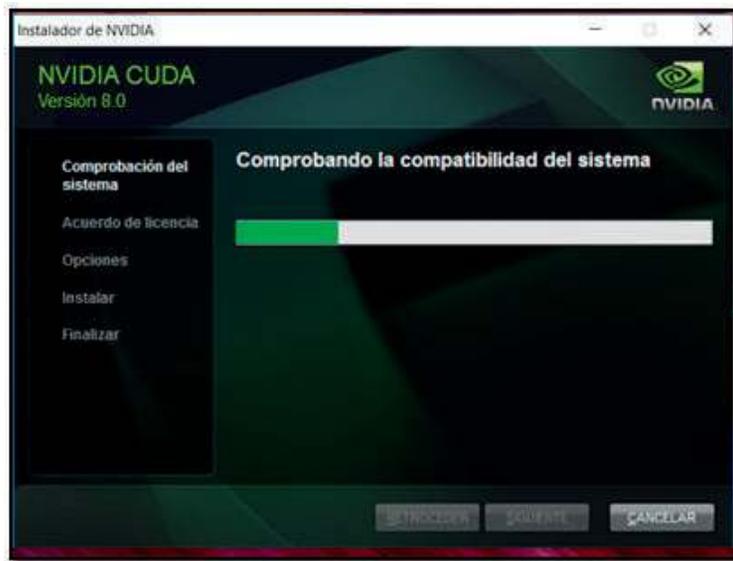
6.1.5. Hola Mundo en CUDA [81]

Para este ejemplo utilizaremos una laptop que cuente con tarjeta gráfica NVIDIA.

a. Instalación de Apis y librerías

i. Descargar e instalar CUDA Toolkit, para este ejemplo utilizaremos la versión 8.0 para Windows 10:

<https://developer.nvidia.com/cuda-toolkit>

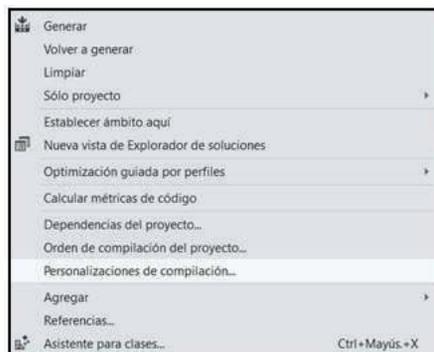


b. Creación y configuración del entorno para programar en paralelo

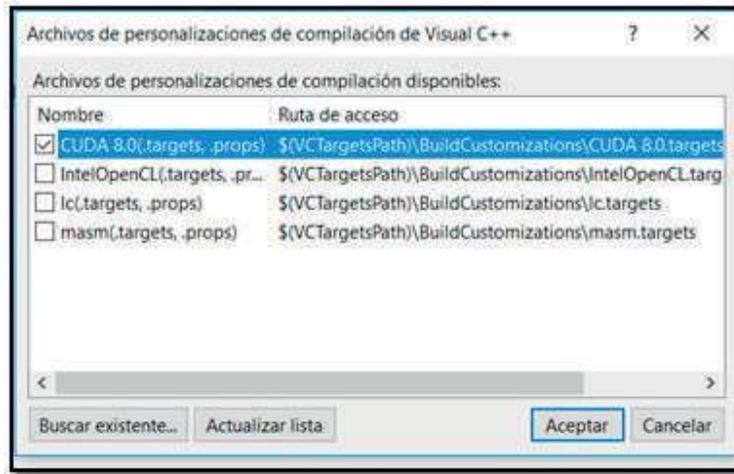
i. Crear un proyecto vacío de c++ en Visual Estudio
ii. Configurar el entorno del proyecto

1. Ingresar a personalizaciones de compilación

a. Clic derecho -> personalizaciones de compilación



- b. Activamos el compilador de CUDA, en este caso CUDA 8.0, y seleccionamos aceptar



2. Configuramos la propiedades del proyecto

- a. En la propiedad "Linker -> All Options -> Additional Dependencies", agregar cudart.lib;

- iii. Crear archivo hola_mundo.cu en "Source Files", donde se escribirá el programa de hola mundo utilizando CUDA.

```
/*Hola Mundo llamando una funcion kernel desde un host*/
//La funcion imprime los hola mundo en cada hilo.
#include "cuda_runtime.h"
#include <stdio.h>
#include <iostream>
```

```
/* Function executed on device (GPU) kernel */
__global__ void hello(void) {
```

```
    printf( "\t Hola mundo desde GPU: Soy el proceso %d de un total de %d procesos\n", threadIdx.x+1, 35);
```

```
}
```

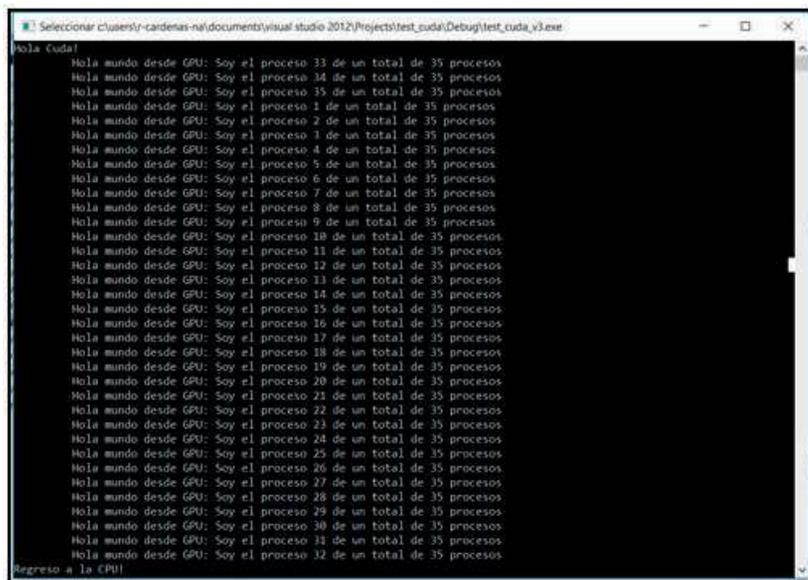
```
//Funcion principal HOST
```

```
int main(void){
    //Definimos el numero de hilos a utilizar
    unsigned int num_threads = 35;
    unsigned int num_block = 1;
```

```
    dim3 dimBlock (num_threads);
    dim3 dimGrid (num_block);
```

```
/* print message from CPU */  
printf( "Hola Cuda!\n" );  
  
/* execute function on device (GPU) */  
hello<<< dimGrid, dimBlock >>>();  
  
/* wait until all threads finish their job */  
cudaDeviceSynchronize();  
  
/* print message from CPU */  
printf( "Regreso a la CPU!\n" );  
system("PAUSE");  
return 0;  
}
```

iv. Compilar y ejecutar el programa



```
Seleccionar c:\users\cardenas-na\documents\visual studio 2012\Projects\test_cuda\Debug\test_cuda_v3.exe  
Hola Cuda!  
Hola mundo desde GPU: Soy el proceso 33 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 34 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 35 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 1 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 2 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 3 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 4 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 5 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 6 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 7 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 8 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 9 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 10 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 11 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 12 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 13 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 14 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 15 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 16 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 17 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 18 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 19 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 20 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 21 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 22 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 23 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 24 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 25 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 26 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 27 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 28 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 29 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 30 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 31 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 32 de un total de 35 procesos  
Regreso a la CPU!
```

6.2. Ejemplos para sistema operativo Linux

Los ejemplos tipo “Hola mundo” en paralelo desarrollados a continuación con las diferentes Apis y librerías, se requieren los siguientes requisitos:

Requisitos

a. Tener instalado GCC y compiladores C++ (gcc y gcc-c++)

b. Conocimiento básico de programación en C++

Seguidamente los ejemplos:

6.2.1. Hola Mundo en TBB (Threading Building Blocks) para Linux

a. Instalación de Apis y librerías

i. Descargar Librería TBB desde la página oficial para Linux

<https://www.threadingbuildingblocks.org/download#stable-releases>

El archivo descargado tendrá el siguiente nombre `tbb2017_20170226oss_lin.tgz`

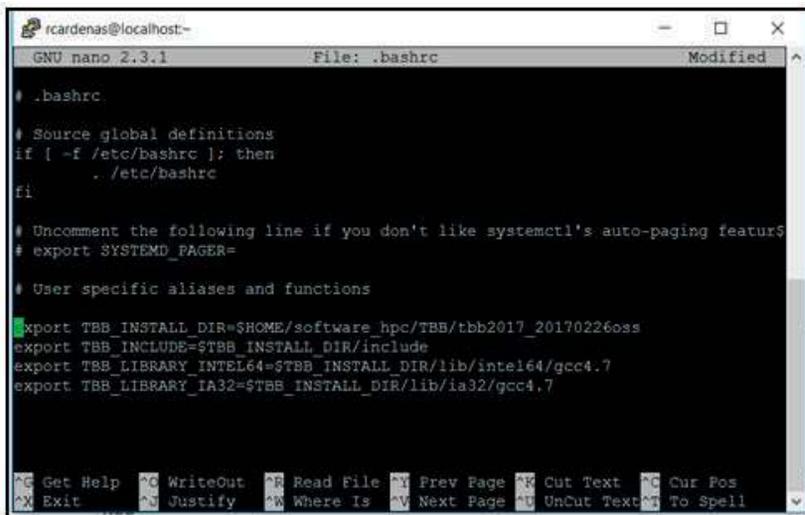
ii. Descomprimos el archivo en la carpeta de su preferencia, en este caso está en `/home/r-cardenas/software_hpc/TBB/tbb2017_20170226oss_lin`

b. Creación y configuración del entorno para programar en paralelo

i. Creación de variables de entorno que apunten a las carpetas de `include` y `lib` de TBB en la cuenta local, en este caso en el archivo `.bashrc` en la ruta `/home/[user]/`

ii. Dentro de `.bashrc` escribir el siguiente código (utilizar el editor `nano .bashrc`)

```
export TBB_INSTALL_DIR=$HOME/software_hpc/TBB/tbb2017_20170226oss
export TBB_INCLUDE=$TBB_INSTALL_DIR/include
export TBB_LIBRARY_INTEL64=$TBB_INSTALL_DIR/lib/intel64
export TBB_LIBRARY_IA32=$TBB_INSTALL_DIR/lib/ia32
```



```
rcardenas@localhost:~$ nano .bashrc
GNU nano 2.3.1 File: .bashrc Modified
# .bashrc
# Source global definitions
if [ -f /etc/bashrc ]; then
    . /etc/bashrc
fi
# Uncomment the following line if you don't like systemctl's auto-paging feature
# export SYSTEMD_PAGER=
# User specific aliases and functions

export TBB_INSTALL_DIR=$HOME/software_hpc/TBB/tbb2017_20170226oss
export TBB_INCLUDE=$TBB_INSTALL_DIR/include
export TBB_LIBRARY_INTEL64=$TBB_INSTALL_DIR/lib/intel64/gcc4.7
export TBB_LIBRARY_IA32=$TBB_INSTALL_DIR/lib/ia32/gcc4.7

Get Help WriteOut Read File Prev Page Cut Text Cur Pos
Exit Justify Where Is Next Page UnCut Text To Spell
```

iii. Recargar el archivo `bashrc` con la siguiente línea de comando:

```
source ~/.bashrc
```

iv. Crear archivo de código fuente con el siguiente nombre `“hola_mundo_tbb.cpp”` con el siguiente comando:

```
touch hola_mundo_tbb.cpp
```

v. Escribir el siguiente código

```
#include "tbb/parallel_for.h"
#include <stdio.h>

using namespace tbb;

int main()
{
    size_t dim=50;
    tbb::parallel_for(size_t(0),dim,[=](size_t i){
        printf("Hola Mundo!... Soy el item %d de un total de %d items \n",i+1,dim);
    });
}
```

vi. Compilar el código para generar el ejecutable, con el siguiente comando

```
g++ -std=c++11 -o hola_mundo_tbb.o hola_mundo_tbb.cpp -I$TBB_INCLUDE -L$TBB-  
B_LIBRARY_INTEL64 -ltbb
```

vii. Ejecutar el programa con el siguiente comando:

```
./hola_mundo_tbb.o
```

6.2.2. Hola Mundo con MPI usando OpenMPI para Linux [82]

a. Instalación de Apis y librerías

ii. Descargar las librerías de OpenMPI desde la página oficial de de OpenMPI.

<https://www.open-mpi.org/software/ompi/v2.1/>

El archivo descargado tiene el siguiente nombre "openmpi-2.1.0.tar"

iii. Descomprimos el archivo en la carpeta de su preferencia, en este caso está en /home/rcardenas/software_hpc/MPI/openmpi-2.1.0

iv. Configuramos el archivo de instalación con el siguiente comando:

```
./configure --prefix="$HOME/software_hpc/MPI/openmpi-2.1.0"
```

v. Dentro de la carpeta /home/rcardenas/software_hpc/MPI/openmpi-2.1.0, hacemos uso de la herramienta "make" para construir las librerías utilizando el siguiente comando:

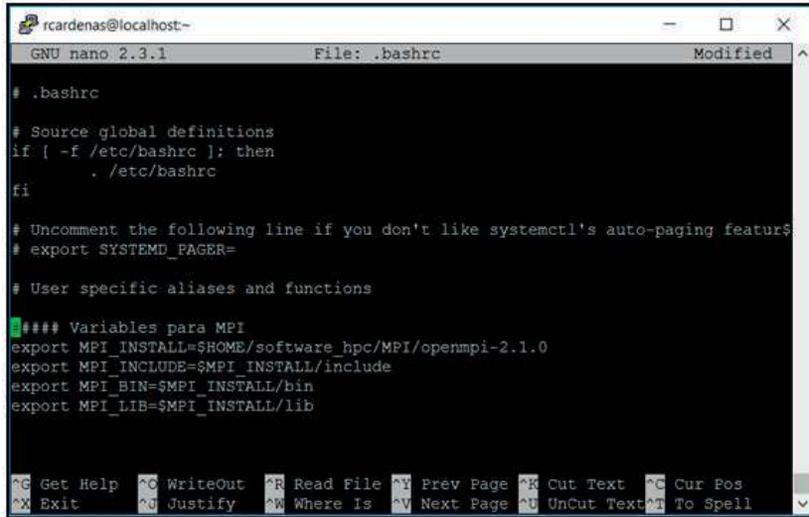
```
make install
```

b. Creación y configuración del entorno para programar en paralelo

i. Creación de variables de entorno que apunten a las carpetas de bin y lib de OpenMPI en la cuenta local, en este caso en el archivo .bashrc en la ruta /home/[user]/:

ii. Dentro de `.bashrc` escribir el siguiente código (utilizar el editor `nano .bashrc`)

```
export MPI_INSTALL=$HOME/software_hpc/MPI/openmpi-2.1.0
export MPI_INCLUDE=$MPI_INSTALL/include
export MPI_BIN=$MPI_INSTALL/bin
export MPI_LIB=$MPI_INSTALL/lib
```



```
rcardenas@localhost:~$ nano .bashrc
GNU nano 2.3.1 File: .bashrc Modified
# .bashrc
# Source global definitions
if [ -f /etc/bashrc ]; then
    . /etc/bashrc
fi

# Uncomment the following line if you don't like systemctl's auto-paging featur$
# export SYSTEMD_PAGER=

# User specific aliases and functions

#### Variables para MPI
export MPI_INSTALL=$HOME/software_hpc/MPI/openmpi-2.1.0
export MPI_INCLUDE=$MPI_INSTALL/include
export MPI_BIN=$MPI_INSTALL/bin
export MPI_LIB=$MPI_INSTALL/lib
```

iii. Recargar el archivo `bashrc` con la siguiente línea de comando:

```
source ~/.bashrc
```

iv. Crear archivo de código fuente “`hola_mundo_mp.cpp`” con el siguiente comando:

```
touch hola_mundo_mp.cpp
```

v. Escribir el siguiente código

```
/* Programa 'hola mundo' donde cada procesador requerido se identifica,
basado en ejemplos orginales de Tim Kaiser (http://www.sdsc.edu/~tkaiser),
del San Diego Supercomputer Center, en California */
```

```
#include <stdio.h>
#include <stdlib.h>
#include <mpi.h>
```

```
int main(int argc, char **argv)
{
```

```
    int myid, numprocs;
    FILE *arch;
```

```
    MPI_Init(&argc,&argv);
```

```
    MPI_Comm_size( MPI_COMM_WORLD, &numprocs ); // devuelve el numero de procesos
```

en este COMM_WORLD

```
MPI_Comm_rank( MPI_COMM_WORLD, &myid ); // identificate por el myid asignado
printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos \n",myid,numprocs);

MPI_Finalize();

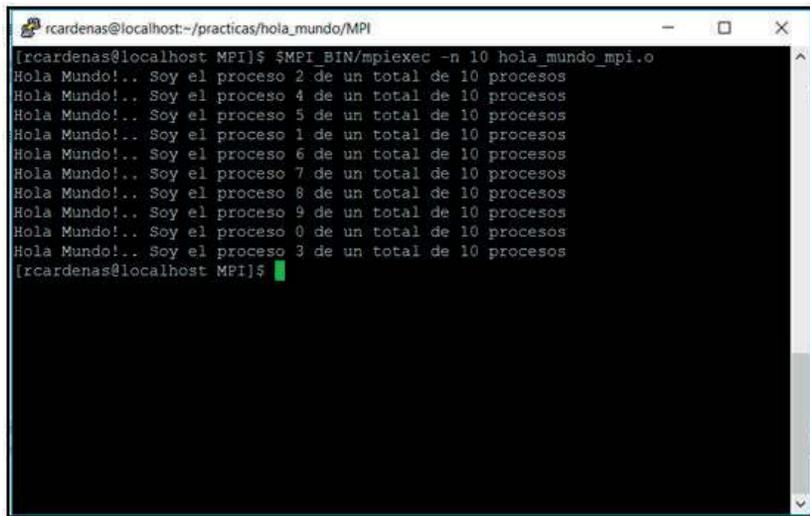
return 0;
}
```

vi. Compilar el código fuente con el siguiente comando:

```
$MPI_BIN/mpicc++ mpi_hola_mundo.cpp -o o_mpi_hola_mun
```

vii. Ejecutar el programa generado con el siguiente comando:

```
$MPI_BIN/mpiexec -n 10 hola_mundo_mpi.o
```



```
rcardenas@localhost:~/practicahola_mundo/MPI
[rcardenas@localhost MPI]$ $MPI_BIN/mpiexec -n 10 hola_mundo_mpi.o
Hola Mundo!.. Soy el proceso 2 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 4 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 5 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 1 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 6 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 7 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 8 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 9 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 0 de un total de 10 procesos
Hola Mundo!.. Soy el proceso 3 de un total de 10 procesos
[rcardenas@localhost MPI]$
```

viii. Como se observa en la salida se tiene 10 hola mundos de diferentes procesos que invocamos.

6.2.3. Hola Mundo con OpenMP en Linu

a. Instalación de Apis y librerías

Según la página oficial de OpenMP, las librerías vienen por defecto al instalar el compilador GCC <http://www.openmp.org/resources/openmp-compilers/>

b. Creación y configurando el entorno para programar en paralelo

i. Crear archivo de código fuente con el siguiente nombre "hola_mundo_tbb.cpp" con el

siguiente comando:

```
touch hola_mundo_openmp.cpp
```

ii. Escribir el siguiente código

```
#include <stdio.h>
#include <stdlib.h>
#include <omp.h>
```

```
int main(int argc, char **argv)
```

```
{
    int p=0, id_hilo, id_proceso;

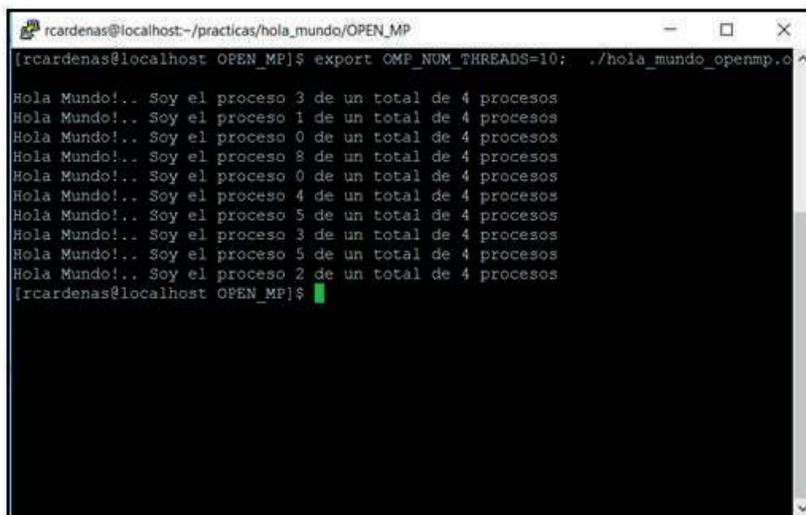
    p=omp_get_max_threads();
    omp_set_num_threads(p);
    #pragma omp parallel
    {
        id_hilo=omp_get_thread_num();
        id_proceso=omp_get_num_procs();
        printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos \n",id_hilo,i-
d_proceso);
    }
    return 0;
}
```

iii. Compilar el código para generar el ejecutable, con el siguiente comando

```
g++ -fopenmp hola_mundo_openmp.cpp -o hola_mundo_openmp.o
```

iv. Ejecutar el programa con

```
export OMP_NUM_THREADS=10; ./hola_mundo_openmp.o
```



```
rcardenas@localhost:~/practicas/hola_mundo/OPEN_MP
rcardenas@localhost OPEN_MP]$ export OMP_NUM_THREADS=10; ./hola_mundo_openmp.o
Hola Mundo!.. Soy el proceso 3 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 1 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 0 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 8 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 0 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 4 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 5 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 3 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 5 de un total de 4 procesos
Hola Mundo!.. Soy el proceso 2 de un total de 4 procesos
rcardenas@localhost OPEN_MP]$
```

6.2.4. Hola Mundo con OpenCL en Linux [83]

a.Requisitos

Tener instalado dkms,

b.Instalación de Apis y librerías [84]

i. Open Cl, permite realizar programación paralela en diversos tipos de procesadores tanto en CPU (Intel, AMD) como en GPU (NVIDIA, AMD), para ello se debe contar con los SDKs de acuerdo a la marca del procesador. Y en esta oportunidad se trabajaremos con INTEL

ii. Descargar e instalar Intel® SDK for OpenCL™ Applications, para este ejemplo se descargo intel_sdk_for_openc1_2016_6.3.0.1904_x64.tgz

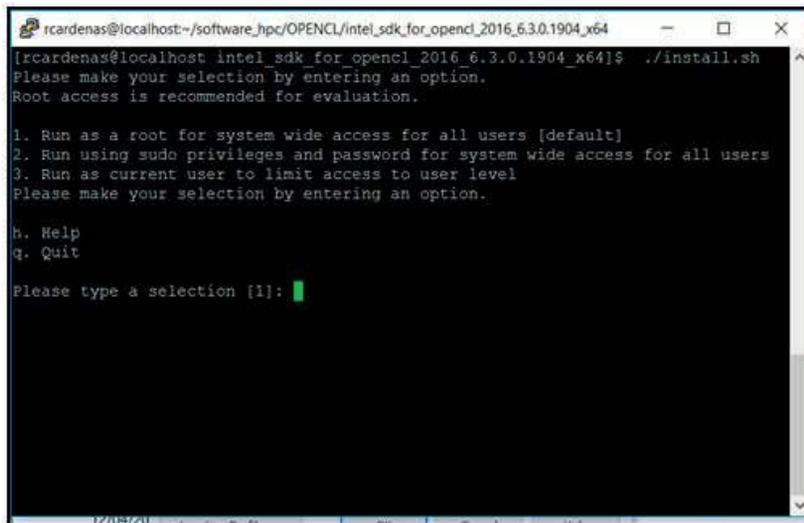
<https://software.intel.com/en-us/intel-openc1>

iii. Descomprimir en la carpeta de preferencia en este caso:

```
/home/rcardenas/software_hpc/OPENCL/intel_sdk_for_openc1_2016_6.3.0.1904_x64
```

iv. Dentro de la carpeta ejecute el siguiente comando (para la interfaz de línea de comandos) y siga las instrucciones del instalador:

```
./install.sh
```

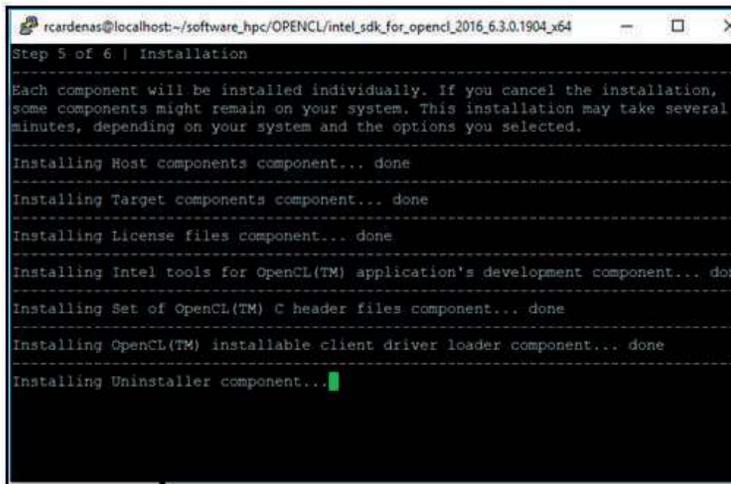


```
rcardenas@localhost:~/software_hpc/OPENCL/intel_sdk_for_openc1_2016_6.3.0.1904_x64
[rcardenas@localhost intel_sdk_for_openc1_2016_6.3.0.1904_x64]$ ./install.sh
Please make your selection by entering an option.
Root access is recommended for evaluation.

1. Run as a root for system wide access for all users [default]
2. Run using sudo privileges and password for system wide access for all users
3. Run as current user to limit access to user level
Please make your selection by entering an option.

h. Help
q. Quit

Please type a selection [1]: █
```



```
rcardenas@localhost:~/software_hpc/OPENCL/intel_sdk_for_opencl_2016_6.3.0.1904_x64
Step 5 of 6 | Installation
-----
Each component will be installed individually. If you cancel the installation,
some components might remain on your system. This installation may take several
minutes, depending on your system and the options you selected.
-----
Installing Host components component... done
-----
Installing Target components component... done
-----
Installing License files component... done
-----
Installing Intel tools for OpenCL(TM) application's development component... don
-----
Installing Set of OpenCL(TM) C header files component... done
-----
Installing OpenCL(TM) installable client driver loader component... done
-----
Installing Uninstaller component... █
```

v. Verificamos que en la ruta

/etc/alternatives/

Que existan las carpetas correspondientes de OpenCL

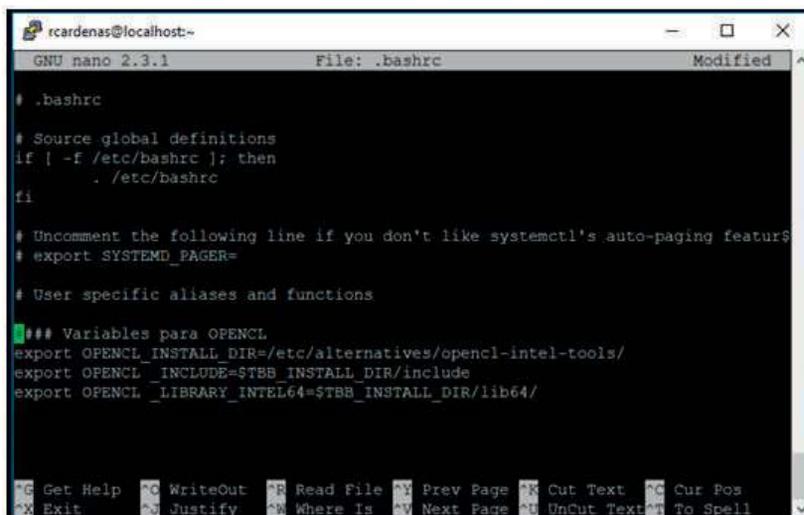
c. Creación y configuración del entorno para programar en paralelo

i. Creación de variables de entorno que apunten a las carpetas de include y lib de OPENCL en

la cuenta local, en este caso en el archivo .bashrc en la ruta /home/[user]/

ii. Dentro de .bashrc escribir el siguiente código (utilizar el editor nano .bashrc)

```
export OPENCL_INSTALL_DIR=/etc/alternatives/opencv-intel-tools/
export OPENCL_INCLUDE=$OPENCL_INSTALL_DIR/include
export OPENCL_LIBRARY_INTEL64=$OPENCL_INSTALL_DIR/lib64/
```



```
rcardenas@localhost:~
GNU nano 2.3.1 File: .bashrc Modified
# .bashrc
# Source global definitions
if [ -f /etc/bashrc ]; then
    . /etc/bashrc
fi
# Uncomment the following line if you don't like systemctl's auto-paging feature
# export SYSTEMD_PAGER=
# User specific aliases and functions

### Variables para OPENCL
export OPENCL_INSTALL_DIR=/etc/alternatives/opencv-intel-tools/
export OPENCL_INCLUDE=$OPENCL_INSTALL_DIR/include
export OPENCL_LIBRARY_INTEL64=$OPENCL_INSTALL_DIR/lib64/

Get Help WriteOut Read File Prev Page Cut Text Cur Pos
Exit Justify Where Is Next Page UnCut Text To Spell
```

iii. Recargar el archivo bashrc con la siguiente línea de comando:

```
source ~/.bashrc
```

iv. Creamos los archivos código fuente necesarios para Open CL:

1. Crear archivo de código fuente con el siguiente "hola_mundo_opencl.cpp" con el siguiente comando:

```
touch hola_mundo_opencl.cpp
```

Escribir el siguiente código:

```
#include <stdio.h>
#include <stdlib.h>
#include <CL/cl.h>

//Para el tamaño de la matriz de string que contendra el codigo fuente del kernel
#define MAX_SOURCE_SIZE (0x100000)

int main(void) {

    // Carga del código fuente del kernel en la matriz cod_kernel
    FILE *fp;
    char *cod_kernel;
    size_t source_size;

    fp = fopen("kernel_hola_mundo.cl", "r");
    if (!fp) {
        fprintf(stderr, "Error al cargar el kernel.\n");
        exit(1);
    }
    cod_kernel = (char*)malloc(MAX_SOURCE_SIZE); //inicializamos la matriz de string y reservamos un espacio de memoria
    cod_kernel[MAX_SOURCE_SIZE+1]='\0';
    source_size = fread( cod_kernel, 1, MAX_SOURCE_SIZE, fp);
    fclose( fp );

    cl_int ret;
    int j;

    //Para Imprimir la Plataforma seleccionada
    cl_platform_id *platform_id;
    char* info;
    size_t infoSize;
    cl_uint ret_num_platforms;

    ret = clGetPlatformIDs(5, NULL, &ret_num_platforms);

    printf("\n Numero de plataformas detectadas: %d\n",ret_num_platforms);
```

```

platform_id = (cl_platform_id *)malloc(sizeof(cl_platform_id) * ret_num_platforms);

ret = clGetPlatformIDs(ret_num_platforms, platform_id, NULL);

// Para obtener el tamaño del nombre de la plataforma y reservar espacio en memoria
clGetPlatformInfo(platform_id[0], CL_PLATFORM_NAME, 0, NULL, &infoSize);
info = (char*) malloc(infoSize);
// Para obtener el nombre de la plataforma
clGetPlatformInfo(platform_id[0], CL_PLATFORM_NAME, infoSize, info, NULL);
printf("\n Plataforma seleccionada: %s", info);

//Para Imprimir la Dispositivo seleccionado
cl_device_id device_id=NULL;
cl_uint ret_num_devices=1;
//Aca definimos el tipo de dispositivo habilitada para la plataforma.
//Para este ejemplo con CL_DEVICE_TYPE_CPU, devuelve Intel Intel(R) Core(TM) i7-7500U
si la plataforma es Intel.
//Y con CL_DEVICE_TYPE_GPU, devuelve GeForce GTX 950M si es NVIDIA
ret = clGetDeviceIDs( platform_id[0], CL_DEVICE_TYPE_CPU, 1,&device_id, NULL);

        char* value;
size_t valueSize;
// Para obtener el tamaño del nombre del dispositivo y reservar espacio en memoria
clGetDeviceInfo(device_id, CL_DEVICE_NAME, 0, NULL, &valueSize);
value = (char*) malloc(valueSize);
// Para obtener el nombre del dispositivo
clGetDeviceInfo(device_id, CL_DEVICE_NAME, valueSize, value, NULL);
printf("\n Dispositivo seleccionado: %s \n\n",value);
free(value);

//Creamos el contexto para manejar los objetos y recursos de OPENCL
cl_context context = clCreateContext( NULL, 1, &device_id, NULL, NULL, &ret);

if (ret != CL_SUCCESS)
{
    printf( "\n->No se pudo crear el contexto de la CPU");
}
else
{
    printf("\n->Se creo el Contexto de la GPU con exito");
}

//Creamos una cola de comandos que nos permitira comunicarnos del HOST al Dispositivo
seleccionado
cl_command_queue command_queue = clCreateCommandQueue(context, device_id, 0,
&ret);

```

```

if (command_queue == NULL)
{
    printf("\n->Error al crear comandoQueue para dispositivo 0");
    exit(0);
}
else
{
    printf("\n->Se creo commandQueue con exito");
}

//Creamos el objeto programa que contiene el codigo fuente del Kernel
cl_program program = clCreateProgramWithSource(context, 1, (const char **)&cod_kernel,
(const size_t *)&source_size, &ret);

if (program == NULL)
{
    printf("\n->No se pudo crear el programa de CL desde el origen");
}
else
{
    printf("\n->Se creo el programa de CL con exito\n");
}

//Compilamos y enlazamos el programa creado al Dispositivo seleccionado
ret = clBuildProgram(program, 1, &device_id, 0, 0, 0);

if (ret != CL_SUCCESS)
{
    // Determine the reason for the error
    char buildLog[32000];
    clGetProgramBuildInfo(program, device_id, CL_PROGRAM_BUILD_LOG, sizeof(build-
Log), buildLog, NULL);
    printf("\nError en el kernel BUILD PGM: %s\n ", buildLog);
}

//Instanciamos el kernel definido en al codigo fuente del kernel que se encuentra compilado en el
programa CL
cl_kernel kernel = clCreateKernel(program, "hello", &ret);
if (ret != CL_SUCCESS)
{
    printf("Error en: clCreateKernel (falla en la función del kernel):\n");
}

//Definimos la cantidad de grupos y los hilos a utilizar del dispositivo seleccionado
size_t global_work_size = 1;
size_t global_item_size = 10;

```

```
//Ejecutamos el kernel con la cantidad de hilos definidos anteriormente
ret = clEnqueueNDRangeKernel(command_queue, kernel, global_work_size, NULL,
&global_item_size, NULL, 0, NULL, NULL);

if(ret == CL_SUCCESS) {
    //printf("\n\nNO Error in clEnqueueNDRangeKernel\n");
}

else {
    printf("\n->Error clEnqueueNDRangeKernel\n");
}

//Liberamos los objetos creados
ret = clFlush(command_queue);
ret = clFinish(command_queue);
ret = clReleaseKernel(kernel);
ret = clReleaseProgram(program);
ret = clReleaseCommandQueue(command_queue);
ret = clReleaseContext(context);

system("PAUSE");

return 0;
}
```

2. Crear archivo kernel_hola_mundo.cl siguiente código

```
touch kernel_hola_mundo.cl
```

Escribir el siguiente código:

```
#pragma OPENCL EXTENSION cl_amd_printf : enable
__kernel void hello()
{
    int i = get_global_id(0);
    int dim = get_global_size(0);
    printf("Hola Mundo!.. Soy el proceso %d de un total de %d procesos \n",i+1,dim);
}
}
```

v. Compilar el código para generar el ejecutable, con el siguiente comando
g++ -o hola_mundo_opencl.o hola_mundo_opencl.cpp -I\$OPENCL_INC
LUDE -L\$OPENCL_LIBRARY_INTEL64 -lOpenCL

vi. Ejecutar el programa con el siguiente comando

```
./hola_mundo_opencl.o
/
```


c. Creación y configuración del entorno para programar en paralelo

i. Creación de variables de entorno que apunten a las carpetas de include y lib de CUDA en la cuenta local, en este caso en el archivo `.bashrc` en la ruta `/home/[user]/`

ii. Dentro de `.bashrc` escribir el siguiente código (utilizar el editor `nano .bashrc`)
`export CUDA_INSTALL=/usr/local/cuda/`
`export CUDA_BIN=$CUDA_INSTALL/bin`

iii. Recargar el archivo `bashrc` con la siguiente línea de comando:

```
source ~/.bashrc
```

iv. crear archivo de código fuente "hola_mundo_cuda.cu" con el siguiente comando:

```
touch hola_mundo_mp.cpp
```

v. Escribir el siguiente código

```
/*Hola Mundo llmando una funcion kernel desde un host*/
//La funcion imprime los hola mundo en cada hilo.
#include "cuda_runtime.h"
#include <stdio.h>
#include <iostream>

/* Function executed on device (GPU) kernel */
__global__ void hello(void) {

    printf( "\t Hola mundo desde GPU: Soy el proceso %d de un total de %d procesos\n",
threadIdx.x+1, 35);
}
//Funcion principal HOST
int main(void){
    //Definimos el numero de hilos a utilizar
    unsigned int num_threads = 35;
    unsigned int num_block = 1;

    dim3 dimBlock (num_threads);
    dim3 dimGrid (num_block);

    /* print message from CPU */
    printf( "Hola Cuda!\n" );

    /* execute function on device (GPU) */
    hello<<< dimGrid, dimBlock >>> ();

    /* wait until all threads finish their job */
```

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

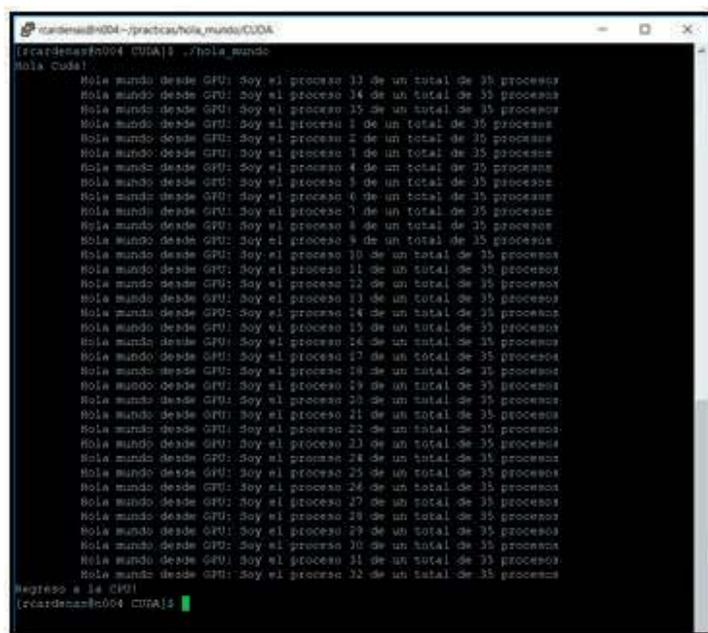
```
cudaDeviceSynchronize();  
  
/* print message from CPU */  
printf("Regreso a la CPU!\n");  
return 0;  
}
```

vi. Compilar el código fuente con el siguiente comando:

```
$CUDA_BIN/nvcc -o hola_mundo hola_mundo_cuda.cu
```

vii. Ejecutar el programa generado con el siguiente código fuente

```
./hola_mundo
```



```
rcardenas@M04:~/proyectos/hola_mundo/CUDA$ ./hola_mundo  
Hola Cuda!  
Hola mundo desde GPU: Soy el proceso 0 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 1 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 2 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 3 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 4 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 5 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 6 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 7 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 8 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 9 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 10 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 11 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 12 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 13 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 14 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 15 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 16 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 17 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 18 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 19 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 20 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 21 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 22 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 23 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 24 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 25 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 26 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 27 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 28 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 29 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 30 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 31 de un total de 35 procesos  
Hola mundo desde GPU: Soy el proceso 32 de un total de 35 procesos  
Regreso a la CPU!  
rcardenas@M04:~/proyectos/hola_mundo/CUDA$
```

CAPÍTULO 7

AGENDA DE PROYECTOS DEL INSTITUTO DE INVESTIGACIONES DE LA AMAZONIA PERUANA – IIAP RELACIONADAS A COMPUTACION DE ALTO RENDIMIENTO

El IIAP a lo largo de 35 años de labor, ha impulsado diversos trabajos de investigación relacionados a ordenamiento territorial, estudios de biodiversidad, inundaciones y otros eventos relacionados al bosque amazónico. Desde el año 2011 viene impulsando proyectos relacionados a computación de alto rendimiento. En la actualidad y en asociación con instituciones como INICTEL UNI, Universidad San Agustín de Arequipa, Universidad Nacional de la Amazonía Peruana, se vienen impulsando una serie de proyectos, entre los que destacan:

7.1 Fortalecimiento de infraestructura tecnológica para procesos de investigación del Instituto de Investigaciones de la Amazonía Peruana.

Desde el año 2015 el Instituto de Investigaciones de la Amazonia Peruana (IIAP) en asociación con el Instituto Nacional de Investigación y Capacitación de Telecomunicaciones de la UNI (INICTEL – UNI), ejecutó el proyecto “Fortalecimiento de Infraestructura Tecnológica para Procesos de Investigación del Instituto de Investigaciones de la Amazonia Peruana”, financiada con fondos de CIENCIACTIVA promovida por CONCYTEC. En el proyecto se planteó el fortalecimiento de la infraestructura de investigación del IIAP mediante la adquisición de una supercomputadora y la implementación de una red de sensores que permita capturar en tiempo real de datos sobre los recursos naturales amazónicos (agua, biodiversidad y otros elementos del ambiente), su posterior almacenamiento, sistematización y altos niveles de procesamiento (supercomputadora), para su posterior visualización, uso, accesibilidad y difusión de resultados, abarcando las regiones Loreto, Ucayali y Madre de Dios.



Imagen 130: Noticia sobre la aprobación del proyecto para la adquisición de una supercomputadora.

El proyecto finalizó en mayo de 2017, teniendo los siguientes principales resultados:

Un clúster de supercomputación.

Como parte de los hitos del proyecto el IAP instaló en su sede principal un Sistema de Alto Rendimiento Computacional del Instituto de Investigaciones de la Amazonía Peruana IAP. Este equipamiento fue conceptualizado como una solución “llave en mano” que comprende entre otras cosas: un clúster compuesto de 9 nodos de procesamiento, 6 de ellos con procesadores GPU (gráficos) y los otros 3 sólo con procesamiento numérico, 1 nodo líder, un sistema de almacenamiento (storage). Además se ha implementado un sistema de alimentación eléctrica, adecuaciones del Data Center, e incluso 40 horas de capacitación para gestión, uso y creación de aplicaciones para HPC.

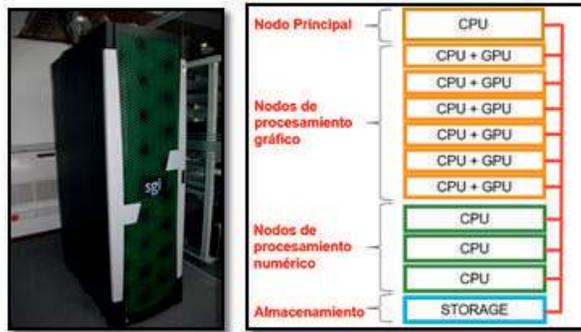


Imagen 131: A la derecha foto de la clúster de supercomputación, a la izquierda arquitectura de la clúster.

03 redes de sensores integrados.

Se instalaron en la ciudad Iquitos 3 nodos sensores, en Pucallpa 3 nodos sensores y 3 nodos sensores en Puerto Maldonado, conformando 3 redes de sensores en las principales ciudades de la Amazonía Peruana. Cada nodo sensor está compuesto por un sistema de flotación, un sistema electrónico que incorpora un módulo de comunicación inalámbrica (GSM o ZigBee), módulo de adquisición y procesamiento de datos, módulo de sensores electroquímicos y módulo de alimentación y carga de energía (batería y paneles solares) para su autonomía. Los nodos sensores permiten monitorear las variables de calidad de agua ph, temperatura del agua, oxígeno disuelto, temperatura del ambiente y humedad relativa.

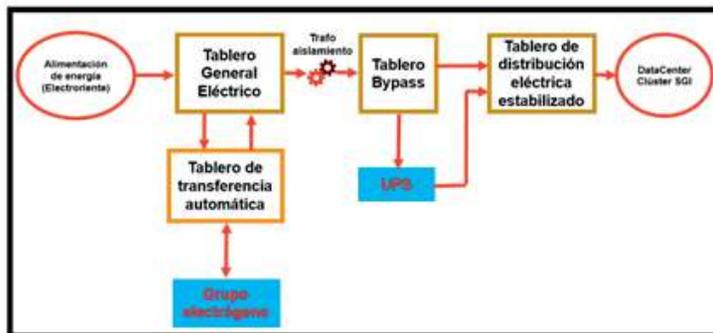


Imagen 132: Gráfico funcional del sistema eléctrico del Sistema de Alto Rendimiento Computacional.



Imagen 133: Nodo sensor transmitiendo datos en Loreto.



Imagen 134 Equipo de INICTEL – UNI, preparando los nodos sensores para ser instalados.



Imagen 135: Casa de fuerza del Centro de Alto Rendimiento Computacional de la Amazonía Peruana.



Imagen 136: Nodo sensor transmitiendo datos en Pucallpa.



Imagen 137: Nodo sensor transmitiendo datos en Pto. Maldonado.

Sistema de monitoreo en tiempo real de variables de calidad de agua

Con la finalidad de integrar y almacenar los datos generados por la red de nodos sensores, se desarrolló un sistema de información para almacenar y monitorear en tiempo real lo datos capturados por la red de nodos sensores.

El Sistema de monitoreo en tiempo real es una herramienta de monitoreo de la calidad del agua y el ambiente, que plasma, a través de una serie de gráficos, los datos primarios de fuentes propias (red de nodos sensores) y externas (otros sistema o equipos sensores) sobre el estado del agua y el ambiente en las distintas puntos de monitoreo en la amazonia peruana. Entre los componentes o módulos que se desarrollaron en este sistema, está la herramienta de Reporte en Tiempo Real, Reporte de Datos históricos y Reporte en Mapa, así como también, la Administración para la Gestión de Usuarios del Sistema, la Gestión de Nodos sensores, y Gestión de Estándares de Calidad Ambiental.



Imagen 138: Interfaz de la ubicación de los nodos sensores en google mapas.



Imagen 139: Interfaz principal del sistema de monitoreo.

Capacitaciones en temas de redes de sensores, procesamiento de imágenes, gestión y programación en supercomputación.

Unos de los hitos principales de las proyecto consistía en la formación de recursos humanos para la gestión y programación en la cluster de supercomputación, además de capacitación de redes de sensores y aplicaciones en HPC, a continuación se lista los principales talleres desarrollados durante el proyecto.

- Pasantía en Gestión de Clúster, Consistió en la visita al IGP en lima de 3 profesionales al IAP, para conocer sobre gestión y configuración de su sistema de clúster de supercomputación.
- Curso introductorio a Gestión de Clúster con LINUX, Consistió en capacitar en los conceptos básicos en el uso de sistema operativo LINUX.
- Curso de Introducción a la Computación paralela, consistió en conocer los conceptos de computación paralela.
- Capacitación en gestión de Clúster Beowulf SGI. Consistió en el desarrollo de capacidades de los gestores de la cluster de supercomputación instalado en el IAP.
- Introducción a la Simulación numérica, curso introductorio al uso de modelos matemáticos utilizando herramientas computacionales. Este curso está dirigido a investigadores de

las temáticas del medio ambiente e hidrología.

- Simulaciones climáticas, hidrológicas, curso de nivel intermedio de aplicaciones a problemáticas del ambiente, cambio climático o hidrología. Este curso está dirigido a investigadores de las temáticas de medio ambiente e hidrología.
- Introducción al FORTRAN, curso introductorio al uso del Fortran como herramienta para computación científica. Este curso está dirigido a investigadores de las temáticas del medio ambiente e hidrología, e informáticos.
- Introducción Machine learning, curso introductorio al uso del Machine Learning como herramienta para estudio del cambio climático. Este curso está dirigido a informáticos y especialistas en medio ambiente.
- Introducción Data mining, curso introductorio al uso del Data Mining como herramienta para aprovechamiento de datos ambientales. Este curso está dirigido a informáticos y especialistas en medio ambiente.
- Introducción al procesamiento de imágenes, comprendió un curso básico de procesamiento de imágenes, con casos básicos de estudio.
- Módulo Matlab y computación paralela con GPU, comprendió los fundamentos básicos de computación paralela utilizando Matlab y procesadores gráficos. Este curso está dirigido básicamente a informáticos.
- Módulo básico de CUDA, comprendió introducción a programación en tarjetas gráficas NVIDIA utilizando CUDA.
- Introducción a MPI, OpenMP, TBB y OpenCL, comprendió introducción a programación paralela en las librerías de programación paralela.
- Introducción a nodos sensores y programación de Arduino, consistió en introducción a los conceptos de nodos sensores y programación de placas arduino.



Imagen 140: Curso de capacitación de Introducción MPI, OpenMP, TBB, CUDA.



Imagen 141: Equipo capacitado en el curso básico de programación básica en MATLAB.

7.2 Circulo de Investigación en computación de alto desempeño, para el desarrollo de métodos y técnicas de minería de datos de gran escala como soporte en investigaciones en cambio climático.

En enero de 2016 Universidad Nacional de San Agustín por medio de la Cátedra CONCYTEC en TICs de UNSA, gana el financiamiento en CIENCIACTIVA de CONCYTEC el proyecto “Circulo de Investigación en computación de alto desempeño, para el desarrollo de métodos y técnicas de minería de datos de gran escala como soporte en investigaciones en cambio climático” el cual el IIAP forma parte como entidad asociada, así como grupo de Investigación en Procesamiento de Imágenes y minería de datos (IPRODAM) de la Universidad La Salle de Arequipa, el grupo de investigación en Reconocimiento de Patrones e Inteligencia Artificial Aplicada (GRPIAA) de la PUCP de Lima.

El proyecto tiene como objetivo el desarrollo de métodos y técnicas de computación de alto desempeño de algoritmos de filtrado, análisis y extracción de patrones de asociación y visualización entre los datos climáticos, para de este modo analizar los resultados de los modelos de pronóstico de cambio climático regional y compararlas con las mediciones reales recogidos por las estaciones de tierra con el fin de evaluar la calidad de los datos generados y tener en cuenta la calibración de los modelos de predicción.

Para esto los proyectos de investigación del círculo abarcan las áreas de HPC, Deep Learning y Visión Computacional, como a continuación se describe de manera general:

- Predicciones a partir de series temporales utilizando redes neuronales.
- Operaciones de Base de datos a nivel de GPU para mejora de tiempos de procesamiento de grandes volúmenes de datos.
- Caracterización de series temporales mediante autocorreladores.

Y por último uno de los resultados del proyecto es la mejora de las capacidades de las entidades Asociadas, en la cual el IIAP empleara las metodologías y técnicas desarrolladas en el proyecto para el análisis de series temporales para la predicción y caracterización del clima de Iquitos.



Imagen 142: Equipo IIAP exponiendo el aporte del IIAP en el proyecto, al equipo monitor de CONCYTEC.



Imagen 143: Coordinador del proyecto presentando los avances al equipo monitor de CONCYTEC.



Imagen 144: Equipo IIAP presentando los avances del IIAP en el proyecto, al equipo monitor de CONCYTEC.

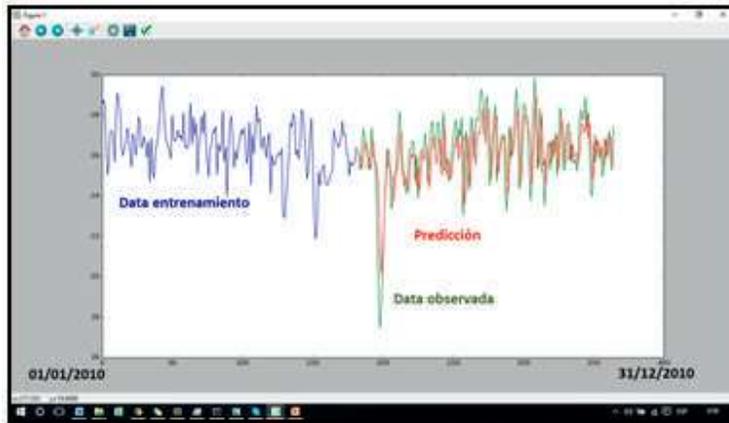


Imagen 145: Predicción de la temperatura de Iquitos con modelo autorregresivo integrado de promedio móvil (ARIMA). Créditos: Elaboración Propia

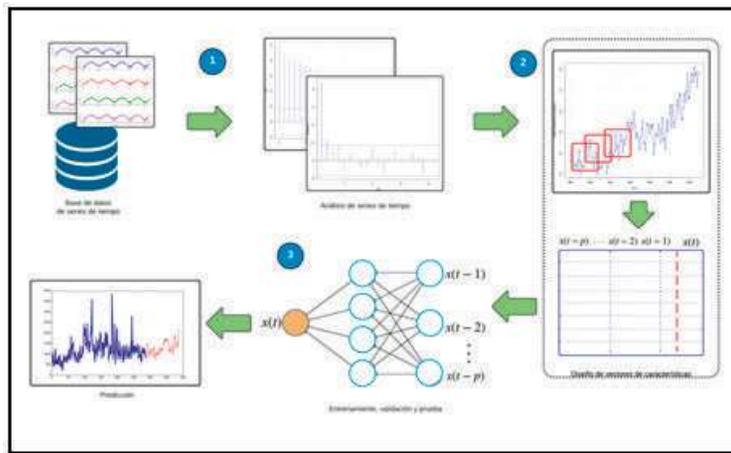


Imagen 146: Modelo de predicción de la temperatura con redes neuronales, fuente proyecto de tesis desarrollado durante el proyecto.

7.3 Evaluación del estado eco sistémico de los aguajales en Iquitos empleando imágenes de alta resolución, computación de alto rendimiento y procesamiento digital de imágenes, que facilite la conservación y uso sostenible de la biodiversidad en la Amazonía

En el 2014 el Instituto Nacional de Investigación y Capacitación de Telecomunicaciones de la UNI (INICTEL – UNI), ganó el financiamiento en Innovate Perú del proyecto “Evaluación del estado eco sistémico de los aguajales en Iquitos empleando imágenes de alta resolución, computación de alto rendimiento y procesamiento digital de imágenes, que facilite la conservación y uso sostenible de la biodiversidad en la Amazonía”, en la cual el IIAP participó como entidad asociada.

El proyecto tiene como objetivo desarrollar un aplicativo de software validado para la identificación, monitorización del estado de veredor y deforestación de la Mauritia Flexuosa mediante el

uso de UAV/ DRONE, cámaras multispectrales y la generación de índices de vegetación en aguajales cercanos a la ciudad de Iquitos.

A la fecha de la culminación del proyecto (Setiembre 2017), los principales productos del proyecto es un software para la detección de aguajes, un software para la reconstrucción y georeferenciación de mosaicos de las imágenes aéreas y un GEOPORTAL para la visualización de Imágenes Satelitales, DRONE y muestras de campo generadas en el proyecto.

Los resultados brindan al IIAP una herramienta automática para realizar el monitoreo del ecosistema conocido como Aguajales para hacer comparaciones de resultados multitemporales, y así, poder estudiar la pérdida de bosques determinando si fue a causa de degradación (natural) o deforestación

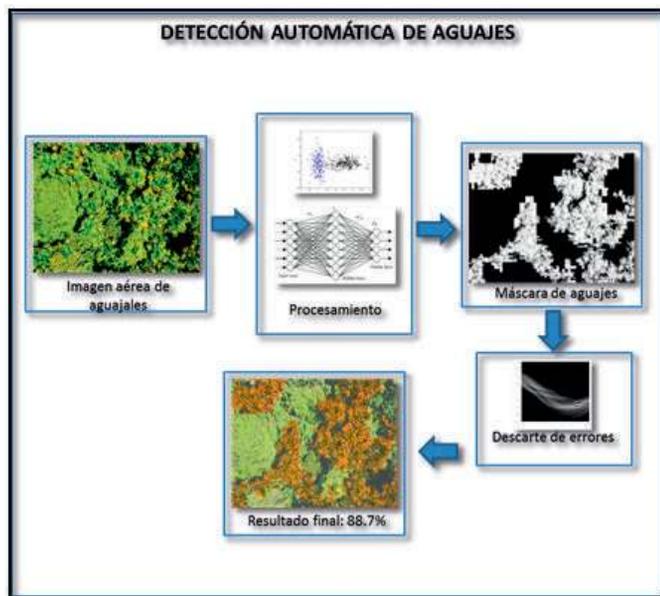


Imagen 147: Proceso de detección Automática de aguajes, fuente informe de proyecto.



Imagen 148: Ciclo de procesamiento de la información del proyecto, fuente informe del proyecto.

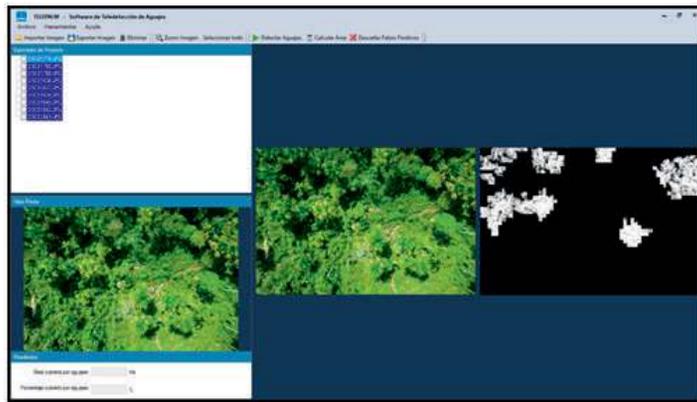


Imagen 149: Software de teledetección de Aguajes desarrollado por INICTEL UNI.



Imagen 150: Proceso de reconstrucción y georeferenciación de mosaicos de imágenes aéreas, fuente informe de proyecto.

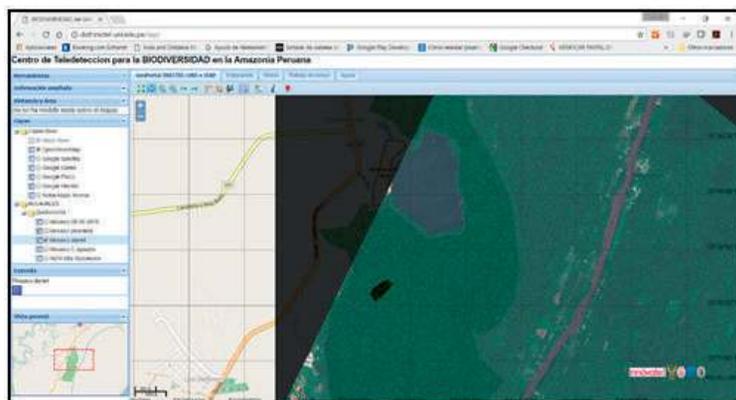


Imagen 151: Pantalla principal del GEOPORTAL desarrollado en el proyecto.

7.4 Reconocimiento de patrones de hojas de plantas de importancia económica en imágenes digitales utilizando redes neuronales

La identificación de plagas y/o enfermedades de las plantas de importancia económica constituyen un proceso crucial en las actividades de los agricultores amazónicos, por lo cual, aprovechando las capacidades de la Supercomputadora, se viene desarrollando un software para reconocimiento de patrones, plagas y/o enfermedades de las hojas de plantas de importancia económica.

Se priorizaron las hojas de las plantas de Camu Camu, Sacha Inchi y Cacao, con las cuales se generó un banco de imágenes, que contiene como mínimo 1000 imágenes digitales de las hojas por especie. Posteriormente se utilizó las redes neuronales convolucionales (CNNs), VGG16/VGG19/ResNet para el reconocimiento de patrones, plagas y/o enfermedades. La primera funcionalidad del software es identificar las hojas plantas y segundo funcionalidad en poder determinar la presencia de plagas y/o enfermedades.



Imagen 152: Procesos de identificación de patógenos en especies vegetales de importancia económica a través de procesamiento de imágenes.

7.5 Procesamiento de imágenes satelitales para deforestación, monitoreo, tala ilegal, inundaciones, dinámica de ríos, sedimentación, cambio climático.

Aprovechando las capacidades del procesamiento grafico que cuenta la Supercomputadora, se tiene como proyecto el desarrollo de un sistema de monitoreo de deforestación así también el desarrollo de herramientas que permitan producir información sobre problemas ambientales que afecta a la amazonia peruana como la tala ilegal, sedimentación y dinámica de los ríos y el cambio climático. Para esto mediante una colección de imágenes satelitales proveniente de diversas fuentes (CONIDA, Satélites Lansat, RapidEye y otros servicios), serán el principal insumo para los algoritmos de las herramientas que correrán en la Supercomputadora



Imagen 153: Imagen satelital Lansat 8, mostrando Iquitos y la zona deforestada de tamshiyacu.

7.6 Alcance del programa BIOINFO en proyectos relacionados con HPC

El 27 de enero de 2017 el IIAP inauguró el Centro de Alto Rendimiento de Computacional de la Amazonía Peruana, el cual posee un clúster de supercomputación con capacidades de procesamiento numérico (para trabajar con bases de datos de millones de registros) y gran capacidad de procesamiento gráfico (para el tratamiento de miles de imágenes de satélite, radar, DRONES y otros sensores)

Las actividades del Centro abarcan las áreas de Deep Learnig, Bioinformática y HPC, con el objetivo de realizar investigaciones sobre deforestación, conservación de la biodiversidad, recursos hídricos, cambio climático, ordenamiento territorial, poblaciones indígenas, bioinformática y otros estudios. Y desde el centro se promueve el desarrollo de la Ciencia de Computación en la rama de computación de alto rendimiento en las universidades de la Amazonía Peruana, mediante las orientaciones a los practicantes de pre-grado.



Imagen 154: Página web del Centro de Alto Rendimiento Computacional de la Amazonía Peruana.



REFERENCIAS BIBLIOGRAFICAS

- [1] Top500, Top500 List - November 2016 [En línea]. Waibstadt: Top500, 2016. Disponible en: <https://www.top500.org/lists/2016/11/>. [Último acceso: marzo 2017].
- [2] InsideHPC, What is high performance computing? [En línea]. Portland: InsideHPC, 2017. Disponible en: <http://insidehpc.com/hpc-basic-training/what-is-hpc>. [Último acceso: marzo 2017].
- [3] D. Guerrero Fernández, «Plataforma de supercomputación para bioinformática,» Tesis doctoral, Dep. de Biología Molecular y Bioquímica, Universidad de Málaga, Málaga, 2015.
- [4] REUNA, Laboratorio de computación de alto rendimiento del CMM operará a través de Red Académica Nacional [En línea]. Santiago: REUNA, 2014. Disponible en: <http://www.reuna.cl/component/content/article.html?id=3110centro-computacional-de-alto-rendimiento-del-cmm-operara-a-traves-de-red-academica-nacional>. [Último acceso: febrero 2017].
- [5] F. García Nocetti, Cómputo de alto rendimiento (HPC) & big data [En línea]. México: UNAM, 2014. Disponible en: <http://www.inegi.org.mx/eventos/2014/big-data/doc/P-DemetrioGareia.pdf>. [Último acceso: febrero 2017].
- [6] ECAR, Escuela de Computación de Alto Rendimiento (ECAR 2013) [En línea]. Mendoza: ECAR, 2013. Disponible en: <http://ecar2013.hpclatam.org>. [Último acceso: marzo 2017].
- [7] Misrespuestas.com, Qué son los flops? [En línea]. Santiago: Misrespuestas.com, 2016. Disponible en: <http://www.misrespuestas.com/que-son-los-flops.html>. [Último acceso: abril 2017].
- [8] BBC, Cómo se resucita una computadora de hace 65 años [En línea]. Londres: BBC, 2014. Disponible en: http://www.bbc.com/mundo/noticias/2014/11/141127_tecnologia_computadora_historica_recreacion_museo_ig. [Último acceso: marzo 2017].
- [9] Top500, About [En línea]. Waibstadt: Top500, 2016. Disponible en: <https://www.top500.org/project/>. [Último acceso: marzo 2017].
- [10] F. García Merayo, «Aplicaciones de la supercomputación,» In Acta: revista digital y manuales formativos, n° 32, pp. 73-85, 2004.
- [11] G. Lissardy, La frustración de tener el mayor supercomputador de América Latina y no poder aprovecharlo [En línea]. Londres: BBC, 2016. Disponible en: <http://www.bbc.com/mundo/noticias-america-latina-36771239>. [Último acceso: marzo 2017].
- [12] A. Clements, The Coprocesor [En línea]. Middlesbrough: University of Teesside, 2014. Disponible en: <http://alanclements.org/coprocesso.html>. [Último acceso: marzo 2017].
- [13] SofDoit, ¿Qué es un coprocesador? [En línea]. Barcelona: SofDoit, 2016. Disponible en: <https://www.softwaredoit.es/definicion/definicion-coprocesador.html>. [Último acceso: marzo 2017].

- [14] INTEL, Procesadores INTEL Xeon Phi [En línea]. Santa Clara CA: INTEL, 2017. Disponible en: <http://www.intel.la/content/www/xl/es/products/processors/xeon-phi/xeon-phi-processors.html?wapkw=intel+xeon+phi>. [Último acceso: marzo 2017].
- [15] HP Enterprise, HPE StoreEasy 3000 Storage Gateway [En línea]. Palo Alto CA: HP Enterprise, 2016. Disponible en: <https://www.hpe.com/us/en/product-catalog/storage/file-storage/pip.HPE-StoreEasy-3000-storage.5335826.html>. [Último acceso: marzo 2017].
- [16] DELCROSA, Transformador de distribución [En línea]. Lima: DELCROSA, 2013. Disponible en: <http://www.delcrosa.com.pe/ver-unidades-negocio/transformador-de-distribucion>. [Último acceso: marzo 2017].
- [17] DSG, ¿Qué es un tablero de transferencia? [En línea]. Santiago de Querétaro: DGS, 2015. Disponible en: <http://venta-deplantasdeluz.com.mx/que-es-tablero-de-transferencia.html>. [Último acceso: marzo 2017].
- [18] Tecniases, Tablero eléctrico general [En línea]. Lima: Tecniases, 2015. Disponible en: <http://tecniases.com/tablero-general-tablero-electrico/>. [Último acceso: abril 2017].
- [19] DLMRED, Transformadores de aislamiento en seco [En línea]. Lima: DLMRED, 2013. Disponible en: http://www.dlmred-ups.com.pe/espanol/transformadores_aislamiento.html. [Último acceso: abril 2017].
- [20] L. ALEGSA, Definición de UPS (uninterruptible power supply) [En línea]. Santa Fe: ALEGSA, 2016. Disponible en: <http://www.alegsa.com.ar/Dic/ups.php>. [Último acceso: abril 2017].
- [21] RedGiga, By-pass - definición - significado [En línea]. Madrid: RedGiga, 1998. Disponible en: <https://diccionario.motorgiga.com/diccionario/by-pass-definicion-significado/gmx-niv2015-con193299.htm>. [Último acceso: abril 2017].
- [22] G. Olson, Mecanismos de interruptores de transferencia de bypass. Columbus IN: Cummins, 2007.
- [23] Solinfra, Tableros eléctricos de distribución [En línea]. Quito: Solinfra, 2016. Disponible en: <http://www.solinfra.com/soluciones-de-energia-normal-y-regulada/>. [Último acceso: abril 2017].
- [24] Tecniases, Tablero eléctrico de distribución - especificaciones técnicas [En línea]. Lima: Tecniases, 2015. Disponible en: <http://tecniases.com/tablero-de-distribucion-tablero-electrico/>. [Último acceso: abril 2017].
- [25] Programa Casa Segura, La puesta a tierra en instalaciones eléctricas [En línea]. México : Programa Casa Segura, 2016. Disponible en: <http://programacasasegura.org/mx/la-puesta-a-tierra-en-instalaciones-electricas-2/>. [Último acceso: abril 2017].
- [26] EcuRed, Grupo electrógeno [En línea]. La Habana: EcuRed, 2016. Disponible en: https://www.ecured.cu/Grupo_Electr%C3%B3geno. [Último acceso: abril 2017].
- [27] CentOS Project, CentOS Linux [En línea]. Raleigh NC: CentOS Project, 2017. Disponible en: <https://www.centos.org/about/#centos-linux>. [Último acceso: abril 2017].

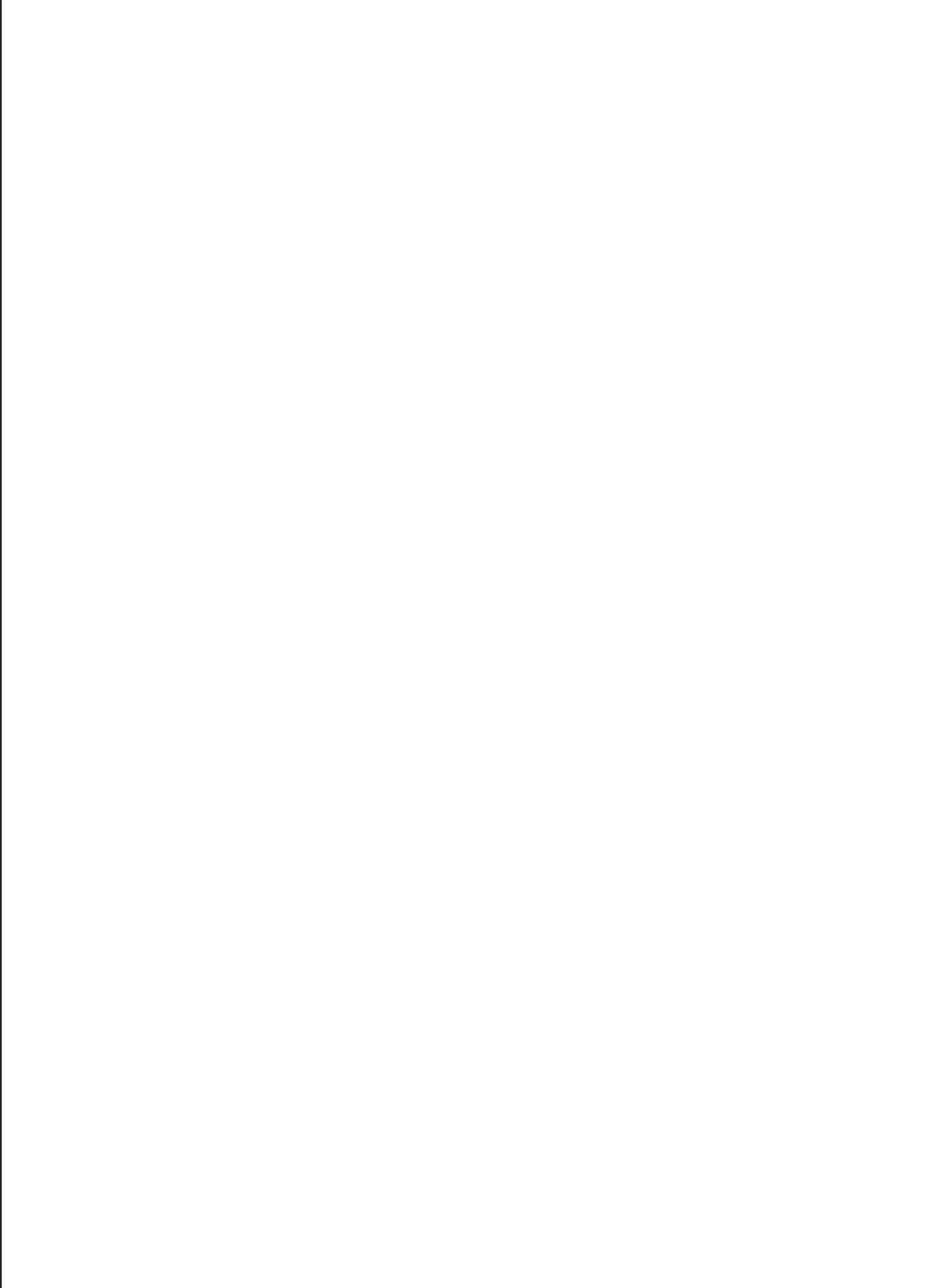
- [28] M. Davis, Las ventajas y desventajas de CentOS [En línea]. Southfield MI: Future Hosting, 2014. Disponible en: www.futurehosting.com/blog/the-advantages-and-disadvantages-of-centos/. [Último acceso: abril 2017].
- [29] L. Kaplan, Full operating system core specialization. Seattle WA: Cray, 2012.
- [30] Cray Inc. Cray releases new version of Cray Linux Environment [En línea]. Seattle WA: Cray, 2010. Disponible en: <http://insidehpc.com/2010/04/cray-releases-new-version-of-cray-linux-environment/>. [Último acceso: abril 2017].
- [31] Cray Inc. Manage system software for Cray Linux Environment. Seattle WA: Cray, 2015.
- [32] QBS Software, Red Hat Enterprise Linux [En línea]. Wembley: QBS Software, 2016. Disponible en: http://www.qbssoftware.com/products/Red_Hat_Enterprise_Linux/overview/_prodredhatad. [Último acceso: abril 2017].
- [33] SUSE, SUSE Linux Enterprise Server [En línea]. Austin TX: SUSE, 2017. Disponible en: <https://www.suse.com/es-es/products/server/>. [Último acceso: abril 2017].
- [34] DCI, TOSS: a Linux-distro from Madurai [En línea]. Tamilnadu: DCI, 2010. Disponible en: <http://www.dotcominfoway.com/blog/toss-a-linux-distro-from-madurai/>. [Último acceso: abril 2017].
- [35] Centro Nacional de Cálculo Científico. Universidad de los Andes. Ganglia toolkit [En línea]. Caracas: Universidad de Los Andes, 2003. Disponible en: http://www.cecalc.ula.ve/HPCLC/slides/day_06/Monitoring/Exercises_Monitoring/Practica_de_GANGLIA_TOOLKIT.pdf. [Último acceso: abril 2017].
- [36] A. Hernandez, J. C. Avila, A. Sardinha y E. Teran, ¿Qué es Nagios? [En línea]. México: s.n., 2016. Disponible en: <https://prezi.com/hpsapkbye5j5/que-es-nagios/>. [Último acceso: abril 2017].
- [37] N. Bezroukov, OpenPBS [En línea]. s.l.: Softpanorama, 2016. Disponible en: http://www.softpanorama.org/Admin/Job_schedulers/openpbs.shtml. [Último acceso: abril 2017].
- [38] V. Zeokat, ¿Qué es PuTTY y para qué sirve? [En línea]. Madrid: VozIdea, 2014. Disponible en: <http://www.vozidea.com/que-es-putty-y-para-que-sirve>. [Último acceso: abril 2017].
- [39] S. Filippi, ¿Qué es PuTTY y para qué sirve? [En línea]. Madrid: Internetlab, 2010. Disponible en: <http://www.internetlab.es/post/891/que-es-putty-y-para-que-sirve/>. [Último acceso: abril 2017].
- [40] Software Tools for Academics and Researchers, Sun grid engine (SGE) quickstart [En línea]. Cambridge MA: STAR, 2011. Disponible en: <http://star.mit.edu/cluster/docs/0.93.3/guides/sge.html>. [Último acceso: abril 2017].
- [41] Oracle Corp., Sun N1 grid engine 6.1 user's guide, Redwood Shores CA: Oracle, 2007.
- [42] Adaptive Computing, TORQUE resource manager [En línea]. Provo UT: Adaptive Computing, 2017. Disponible en: <http://www.adaptivecomputing.com/products/open-source/torque/>. [Último acceso: abril 2017].

- [43] J. Havlíček, Introducción WinSCP [En línea]. Boston: Free Software Foundation, 2014. Disponible en: <https://winscp.net/eng/docs/lang:es>. [Último acceso: abril 2017].
- [44] Programiz, Learn C programming, the definitive guide [En línea]. Lalitpur: Parewa Labs, 2017. Disponible en: <https://www.programiz.com/c-programming>. [Último acceso: abril 2017].
- [45] Parewa Labs., Learn C++ programming: the definitive guide [En línea]. Patan: Parewa Labs, 2017. Disponible en: <https://www.programiz.com/cpp-programming>. [Último acceso: abril 2017].
- [46] EcuRed, Fortran [En línea]. La Habana: EcuRed, 2016. Disponible en: <https://www.ecured.cu/Fortran>. [Último acceso: abril 2017].
- [47] J. S. Perry, Introducción a la programación Java, parte 1: Conceptos básicos del lenguaje Java [En línea]. New York: IBM, 2012. Disponible en: <https://www.ibm.com/developerworks/ssa/java/tutorials/-introtojava1/>. [Último acceso: abril 2017].
- [48] Oracle Co., ¿Qué es la tecnología Java y para qué la necesito? [En línea]. Redwood Shores CA: Oracle, 2017. Disponible en: https://www.java.com/es/download/faq/whatis_java.xml. [Último acceso: abril 2017].
- [49] Parewa Labs., Learn Python programming: the definitive guide [En línea]. Patan: Parewa Labs, 2017. Disponible en: <https://www.programiz.com/python-programming>. [Último acceso: abril 2017].
- [50] F. J. Romero-Campero, Introducción al lenguaje de programación R [En línea]. Sevilla: Universidad de Sevilla, 2013. Disponible en: https://www.cs.us.es/~fran/curso_unia/introduccion_R.html. [Último acceso: abril 2017].
- [51] M. Gerritsen, A brief introduction to Matlab [En línea]. Stanford: Stanford University, 2006. Disponible en: <https://web.stanford.edu/class/cme001/handouts/matlab.pdf>. [Último acceso: abril 2017].
- [52] G. Borrell i Nogueras, Octave: una alternativa real a Matlab a coste cero [En línea]. Madrid: Universidad Politécnica de Madrid, 2008. Disponible en: http://torroja.dmt.upm.es/media/files/paper_logrono.pdf. [Último acceso: abril 2017].
- [53] J. W. Eaton, GN U Octave [En línea]. London: Network Theory, 2002. Disponible en: <https://www.gnu.org/software/octave/about.html>. [Último acceso: abril 2017].
- [54] Scilab Group, Introduction to Scilab user's guide [En línea]. Marne: Scilab Group: 2000. Disponible en: http://cermics.enpc.fr/scilab_new/site/Liens/intro/node3.html. [Último acceso: abril 2017].
- [55] Apache Software Foundation, Apache Hadoop [En línea]. Wakefield, MA: ASF, 2017. Disponible en: <http://hadoop.apache.org>. [Último acceso: abril 2017].
- [56] NVidia Corporation, Procesamiento paralelo CUDA [En línea]. Disponible en: <http://www.nvidia.es/object/cuda-parallel-computing-es.html> [Último acceso: abril 2017].
- [57] V. Cerverón, Introducción a MPI [En línea]. Valencia: Universidad de Valencia, 2005. Disponible en: http://informatica.uv.es/iiiguia/ALP/materiales2005/2_2_introMPI.htm. [Último acceso: abril 2017].

- [58] Centro de Supercomputación Complutense, Curso de introducción a la programación con MPI [En línea]. Madrid: Universidad Complutense de Madrid, 2005. Disponible en: <http://pendientedemigracion.ucm.es/info/csc/Usuarios/Formacion/Conline/IntroMPI/Default.htm> [Último acceso: abril 2017].
- [59] StreamHPC, What is OPENCL? [En línea]. Amsterdam: StreamHPC, 2017. Disponible en: <https://streamcomputing.eu/knowledge/what-is/opencl/>. [Último acceso: abril 2017].
- [60] P. Santamaría, OpenCL, la alternativa libre a CUDA y el futuro de la computación GPG [En Línea] Madrid: Weblog SL, 2010. Disponible en: <https://www.applesfera.com/aplicaciones-os-x-1/opencl-la-alternativa-libre-a-cuda-y-el-futuro-de-la-computacion-gpgpu>. [Último acceso: abril 2017].
- [61] Fundación Computación y Tecnologías Avanzadas de Extremadura, ¿Qué es OpenMP? [En línea]. Cáceres: Fundación Computaex, 2017. Disponible en: <http://www.cenits.es/faq/preguntas-generales/que-es-openmp> [Último acceso: abril 2017].
- [62] B. Barney, OpenMP [En línea]. Livermore, CA: Lawrence Livermore National Laboratory, 2016. Disponible en: <https://computing.llnl.gov/tutorials/openMP/> [Último acceso: abril 2017].
- [63] INTEL, Threading Building Blocks - Why Use Intel® TBB? [En línea]. Santa Clara CA: INTEL, 2017. Disponible en: <https://www.threadingbuildingblocks.org/learn> [Último acceso: abril 2017].
- [64] D.A. Case, D.S. Cerutti, T.E. Cheatham, AMBER home page [En línea]. Athens: Ambermd, 2017. Disponible en: <http://ambermd.org>. [Último acceso: abril 2017].
- [65] Consorci de Serveis Científics i Acadèmics de Catalunya, Gaussian [En línea]. Barcelona: CSCAC, 2017. Disponible en: <http://www.csuc.cat/es/personal/gaussian> [Último acceso: abril 2017].
- [66] Max Planck Institute for Chemical Energy Conversion, ORCA [En línea]. Mülheim an der Ruhr: MPI, 2016. Disponible en: <https://orcaforum.cec.mpg.de> [Último acceso: abril 2017].
- [67] G.M. Morris, Autodock [En línea]. La Jolla CA: The Scripps Research Institute, 2013. Disponible en: <http://autodock.scripps.edu>. [Último acceso: abril 2017].
- [68] Kuntz Group, DOCK 6 [En línea]. San Francisco: UCSF Molecular Design Institute, 2003. Disponible en: http://dock.compbio.ucsf.edu/DOCK_6/index.htm [Último acceso: abril 2017].
- [69] Kuntz Group, Licensing information [En línea]. San Francisco: UCSF Molecular Design Institute, 2003. Disponible en: http://dock.compbio.ucsf.edu/Online_Licensing/index.htm. [Último acceso: abril 2017].
- [70] M. Williams, SiSoftware Sandra Lite [En línea]. Londres: IDG, 2016. Disponible en: <http://www.pcadvisor.co.uk/download/system-desktop-tools/sisoftware-sandra-lite-2016-v2220-605/> [Último acceso: abril 2017].
- [71] Futuremark Co., 3D Mark [En línea]. Espoo: Futuremark, 2017. Disponible en: <https://www.futuremark.com/benchmarks/3dmark>. [Último acceso: abril 2017].

- [72] Futuremark Co., PCMark [En línea]. Espoo: Futuremark, 2017. Disponible en: <https://www.futuremark.com/benchmarks/pcmark8> [Último acceso: abril 2017].
- [73] Yachay EP, Yachay EP e IBM consolidan acciones de alto desarrollo tecnológico para el país [En línea] Quito: Yachay, 2016. Disponible en: <http://www.yachay.gob.ec/yachay-ep-e-ibm-consolidan-acciones-de-alto-desarrollo-tecnologico-para-el-pais>. [Último acceso: abril 2017].
- [74] J. Velasco, Historia de la tecnología: Gordon Moore [En línea]. Madrid: Hipertextual, 2012. <https://www.futuremark.com/benchmarks/3dmark> . [Último acceso: abril 2017].
- [75] mavedfrank, Introducción a MPI: Hola mundo! [En línea] 2012. Disponible en: <https://mavedfrank.wordpress.com/2012/05/26/hola-mundo-con-mpi-en-c/> [Último acceso: marzo 2017].
- [76] A. Vo, How to compile and run a simple MS-MPI program [En línea]. Redmont WA: Microsoft Technet, 2015. Disponible en: <https://blogs.technet.microsoft.com/windowshpc/2015/02/02/how-to-compile-and-run-a-simple-ms-mpi-program/> [Último acceso: marzo 2017].
- [77] S. Sayyad, Installation Open MPI on Windows [En línea]. Jaipur: Tech Tutorials, 2014. Disponible en: <http://www.suhelsayyad.com/2013/08/installation-open-mpi-on-windows.html> [Último acceso: marzo 2017].
- [78] R. Ioffe, OpenCL tutorials [En línea]. Santa Clara CA: INTEL, 2015. Disponible en: <https://software.intel.com/en-us/articles/openc1-tutorials> [Último acceso: marzo 2017].
- [79] A. Hamurur, How to install OpenCL on Windows [En línea]. Amsterdam: StreamHPC, 2015. Disponible en: <https://streamcomputing.eu/blog/2015-03-16/how-to-install-openc1-on-windows/> [Último acceso: marzo 2017].
- [80] V. Kartoshkin y T. Mattson, Tutorial: OpenCL: introducción para programadores de HPC. [En línea]. Santa Clara CA: INTEL, 2011. Disponible en: <https://software.intel.com/en-us/articles/tutorial-openc1-introduction-for-hpc-programmers> [Último acceso: marzo 2017].
- [81] S. Ui Haq, How to run CUDA in Visual Studio 2010 [En línea]. Islamabad: TunaCode, 2011. Disponible en: <http://blog.cuvilib.com/2011/02/24/how-to-run-cuda-in-visual-studio-2010/> [Último acceso: marzo 2017].
- [82] D. Guerrero y S. Rodríguez, Installation of OpenMPI: parallel and distributed programming [En línea]. Granada: Universidad de Granada, 2014. Disponible en: http://lsi.ugr.es/jmantas/pdp/ayuda/datos/instalaciones/Install_OpenMPI_en.pdf [Último acceso: marzo 2017].
- [83] Fixstars Solutions, The OpenCL programming book [En línea]. Irvine CA: Fixstars Solutions, 2017. Disponible en: <https://www.fixstars.com/en/openc1/book/OpenCLProgrammingBook/first-openc1-program>. [Último acceso: marzo 2017].
- [84] R.T. McGibbon, Install Intel-CPU OpenCL on Ubuntu.md [En línea]. San Francisco: GitHub, 2012. Disponible en: <https://gist.github.com/rmcgibbo/6314452> [Último acceso: marzo 2017].

- [85] J. Prats, WMIC de Windows [En línea]. Barcelona: Systemadmin, 2017. Disponible en: <http://systemadmin.es/2011/07/wmic-de-windows> [Último acceso: abril 2017].
- [86] Hogar Digital e Inmótica Ingenieros, Qué es un estabilizador de tensión [En línea]. Madrid: Hogar Digital e Inmótica Ingenieros, 2017. Disponible en: <https://todosai.com/blog/que-es-un-estabilizador-de-tension-b49.html> [Último acceso: marzo 2017].
- [87] Hogar Digital e Inmótica Ingenieros, Estabilizadores de tension – tipos y usos [En línea]. Madrid: Hogar Digital e Inmótica Ingenieros, 2017. Disponible en: <https://todosai.com/blog/estabilizadores-de-tension-tipos-y-usos-b33.html> [Último acceso: abril 2017].
- [88] Superintendencia Nacional de Educación Superior Universitaria - SUNEDU. [En línea] Available: <https://www.sunedu.gob.pe/> [Último acceso: marzo 2017].
- [89] Directorio Nacional de Investigadores e Innovadores - DINA. [En línea] Available: <https://dina.concytec.gob.pe/appDirectorioCTI/> [Último acceso: marzo 2017].



ANEXOS

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

ANEXO 01: Directorio de centros de alto rendimiento computacional de Sudamérica

Nro.	Centro de Alto Rendimiento / Data Center	Institución	Ciudad	País	Website
1	Ceconea	Universidad Nacional del Nordeste	Corrientes	Argentina	http://gica.exa.unne.edu.ar/
2	Centro Atómico de Bariloche	Comisión Nacional de Energía Atómica	San Carlos de Bariloche	Argentina	http://www.cab.cnea.gov.ar/
3	Centro de Computación de Alto Desempeño de la Universidad Nacional de Córdoba	Universidad Nacional de Córdoba	Córdoba	Argentina	http://ccad.unc.edu.ar/
4	Centro de Computación de Alto Desempeño de la Universidad Nacional de Río Cuarto	Universidad Nacional de Río Cuarto	Río Cuarto-Córdoba	Argentina	http://www.unrc.edu.ar/
5	Centro de Investigación de Métodos Computacionales CIMEC	Centro de Investigación de Métodos Computacionales	Santa Fé	Argentina	http://www.cimec.santafe-conicet.gov.ar/
6	Centro de Investigaciones del Mar y la Atmosfera	Universidad de Buenos Aires	Buenos Aires	Argentina	http://www.cima.fcen.uba.ar/
7	Centro de Simulación Computacional para Aplicaciones Tecnológicas	Centro de Simulación Computacional para Aplicaciones Tecnológicas	Buenos Aires	Argentina	http://www.polo.mincyt.gov.ar
8	Cluster ICB-ITIC - Mendoza	Universidad Nacional de Cuyo	Mendoza	Argentina	http://www.icb.uncu.edu.ar/
9	Facultad de Ciencias Exactas – CIFICEN	Universidad Nacional del Centro	Tandil	Argentina	http://www.exa.unicen.edu.ar/
10	Facultad de Ciencias Exactas y Naturales (FCEyN)	Universidad de Buenos Aires	Buenos Aires	Argentina	http://exactas.uba.ar/
11	Facultad de Ciencias Físico Matemáticas y Naturales-LIDIC	Universidad Nacional de San Luis	San Luis	Argentina	http://www.lidic.unsl.edu.ar/
12	GFC / GEMA	Universidad Nacional de la Plata	La Plata	Argentina	http://www.gfc.ing.unlp.edu.ar /
13	GTIC	Comisión Nacional de Energía Atómica	Buenos Aires	Argentina	http://www.cnea.gov.ar/
14	HPC Cluster Rosario	Centro Científico Tecnológico Rosario	Rosario	Argentina	https://www.rosario-conicet.gov.ar/
15	IFISUR	Instituto de Física del Sur	Bahia Blanca	Argentina	http://www.ifisur-conicet.gob.ar/
16	IFLYSIB	Universidad Nacional de la Plata	La Plata	Argentina	http://www.iflysib.unlp.edu.ar/
17	INQUIMAE	Universidad de Buenos Aires	Buenos Aires	Argentina	http://www.inquimae.fcen.uba.ar/
18	Instituto de Astronomía y Física del Espacio	Universidad de Buenos Aires	Buenos Aires	Argentina	http://www.iafe.uba.ar/
19	Instituto de Genética Veterinaria (IGEVET)	Universidad Nacional de la Plata	La Plata	Argentina	http://www.fcv.unlp.edu.ar/
20	Instituto de Investigación en Informática LIDI (III-LIDI)	Universidad Nacional de la Plata	La Plata	Argentina	http://www.lidi.info.unlp.edu.ar/wp/
21	Nodo Bioinformático IMPaM CONICET UBA	Universidad de Buenos Aires	Buenos Aires	Argentina	http://www.brnhid.org.ar/
22	Unidad de Cálculo para Física y Química Teórica (UnCaFIQT)	Universidad Nacional de la Plata	La Plata	Argentina	calculo.inifta.unlp.edu.ar
23	A Coppe – Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	http://www.coppe.ufrj.br/
24	CENAPAD - UFPE	Universidade Federal de Pernambuco - UFPE	Recife	Brasil	http://cenapad-pe.ufpe.br/
25	CENAPAD-SP	Universidad de Campinas	Campinas	Brasil	https://www.cenapad.unicamp.br/

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

26	Centro de Ciências Matemáticas Aplicadas à Indústria (CeMEAI)	Universidade de São Paulo	Sao Carlos	Brasil	http://www.icmc.usp.br/
27	Centro de Computación de Instituto de Física	Universidade de São Paulo	Sao Paulo	Brasil	http://portal.if.usp.br/ccifusp/
28	Centro de Hidrografía da Marinha	Centro de Hidrografia da Marinha	Rio de Janeiro	Brasil	https://www.mar.mil.br/dhn/chm/meteo/
29	Centro de Pesquisa e Análise de São Paulo - SPRACE	Universidade Estadual Paulista "Júlio de Mesquita Filho"	Sao Paulo	Brasil	https://www.sprace.org.br
30	Centro Nacional de Processamento de alto Desempenho	Instituto Nacional de Pesquisas Amazônicas.	Manaus	Brasil	http://cenapad-am.inpa.gov.br/index.php
31	Centro Nacional de Processamento de Alto Desempenho Ambiental - CENAPAD	Instituto Nacional de Pesquisas Espaciais - INPE	Sao Paulo	Brasil	http://www.cptec.inpe.br/
32	Centro Nacional de Processamento de Alto Desempenho de Minas Gerais (CENAPADMG)	Universidade Federal de Minas Gerais	Belo Horizonte	Brasil	http://www.cenapad.ufmg.br/
33	Centro Nacional de Processamento de Alto Desempenho do Ceará – UFC	Universidade Federal de Ceará	Fortaleza	Brasil	http://www.cenapad.ufc.br/
34	Centro Nacional de Supercomputação (CESUP)	Universidade Federal do Rio Grande do Sul	Porto Alegre - Rio Grande do Sul	Brasil	http://www.cesup.ufrgs.br/
35	Dep. de Bioinformática, Engenharia Mecânica y Eléctrica - UFG	Universidade Federal de Goiás	Goias	Brasil	https://www.ufg.br/
36	Departamento de Computação Aplicada	Universidade Federal de Santa Maria	Santa Maria - Rio Grande del Sur	Brasil	http://www.inf.ufsm.br/index
37	Departamento de Engenharia Naval e Oceânica – Escola Politécnica – UFRJ	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	http://www.oceanica.ufrj.br/
38	Departamento de Física da Universidade Federal de Santa Maria	Universidade Federal de Santa Maria	Santa Maria - Rio Grande Del Sur	Brasil	http://coral.ufsm.br/dfisica/
39	Departamento de Física y Química de UFABC	Universidade Federal de ABC	Sao Paulo	Brasil	http://www.ufabc.edu.br/
40	Departamento de Informática	Universidade Federal de Santa Maria	Santa Maria - Rio Grande del Sur	Brasil	http://site.ufsm.br/
41	Departamento de Química	Universidade de São Paulo	Ribeirão Preto	Brasil	http://www.ffclrp.usp.br/departamentos/quimica/
42	Diretoria de hidrografia e navegação da marinha	Marinha do Brasil	Rio de Janeiro	Brasil	https://www.mar.mil.br/dhn/chm/meteo/
43	Engenharia Química	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	https://ufrj.br/
44	Facultad de Ingeniería Química – UNICAMP	Universidad de Campinas	Campinas	Brasil	www.feq.unicamp.br/
45	Facultade Unb Panaltina	Universidade de Brasilia	Brasilia	Brasil	http://fup.unb.br/
46	Fundação Oswaldo Cruz (Fiocruz)	Fundação Oswaldo Cruz (Fiocruz)	Rio de Janeiro	Brasil	https://portal.fiocruz.br/pt-br
47	INPE (National Institute for Space Research)	INPE (National Institute for Space Research)	Sao Paulo	Brasil	http://www.inpe.br/
48	Instituto de Astronomia, Geofísica e Ciências Atmosféricas (IAG)	Universidade de São Paulo	Sao Paulo	Brasil	http://www.iag.usp.br/
49	Instituto de Bioinformática de UCB	Universidade Católica de Brasília	Brasília	Brasil	http://www.ucb.br/

INTRODUÇÃO A LA SUPERCOMPUTACIÓN EN EL PERÚ

50	Instituto de Física "Gleb Wataghin"	Universidad de Campinas	Campinas	Brasil	https://portal.ifi.unicamp.br/
51	Instituto de Física Teórica (IFT) UNESP	Universidade Estadual Paulista "Júlio de Mesquita Filho"	Sao Paulo	Brasil	http://www.ift.unesp.br/
52	Instituto de Geociências	Universidad en Salvador de Bahía	Bahia	Brasil	https://www.ufba.br/
53	Instituto de Química	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	https://www.iq.ufrj.br/es/
54	Instituto Nacional de Meteorologia	Instituto Nacional de Meteorologia	Brasília	Brasil	http://www.inmet.gov.br/
55	Instituto Nacional de Metrologia Qualidade e Tecnologia - INMETRO	Instituto Nacional de Metrologia Qualidade e Tecnologia	Rio de Janeiro	Brasil	http://www.inmetro.gov.br/
56	Instituto Tecnológico de Aeronáutica	Instituto Tecnológico de Aeronáutica	Sao Paulo	Brasil	http://www.ita.br/
57	Instituto Tecnológico VALE	Instituto Tecnológico VALE	Belem do Pará	Brasil	http://www.itv.org/
58	IPT - Instituto de Pesquisas Tecnológicas do Estado de SP	IPT - Instituto de Pesquisas Tecnológicas do Estado de SP	Sao Paulo	Brasil	http://www.ipt.br/
59	Laboratório de Astroinformática	Universidade de São Paulo	Sao Paulo	Brasil	http://www.iag.usp.br/astrofotografia/grupo/laboratorio-de-astroinformatica
60	Laboratório de Computación de Alto Rendimiento (LCAD)	Universidade Federal da Integração Latino-Americana	Foz do Iguaçu – Paraná	Brasil	https://www.unila.edu.br/
61	Laboratório de Tecnologia Oceânica - LabOceano	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	http://www.laboceano.coppe.ufrj.br/
62	Laboratório Nacional de Computação Científica	Laboratório Nacional de Computação Científica	Rio de Janeiro	Brasil	http://www.lncc.br/estrutura/default.php
63	Laboratory of Environmental and Thermal Engineering	Universidade de São Paulo	Sao Paulo	Brasil	http://www.lete.poli.usp.br/
64	Núcleo Avançado de Computação de Alto Desempenho (NACAD) da COPPE/UFRJ	INTEL	Rio de Janeiro	Brasil	http://www.nacad.ufrj.br/
65	Núcleo de Dinâmica e Fluidos (NDF)	Universidade de São Paulo	Sao Paulo	Brasil	http://www.ndf.poli.usp.br/
66	Observatório Nacional	Ministério da Ciência, Tecnologia, Inovações e Comunicações	Rio de Janeiro	Brasil	http://www.on.br/
67	Petrobras	Petrobras	Rio de Janeiro	Brasil	http://www.petrobras.com.br/
68	Pontifícia Universidad Católica de Río	Pontifícia Universidad Católica de Río	Rio de Janeiro	Brasil	www.puc-rio.br/
69	Pontifícia Universidade Católica de São Paulo	Pontifícia Universidade Católica de São Paulo	Sao Paulo	Brasil	http://www.pucsp.br/
70	Programa De Engenharia Química - UFRJ	Universidade Federal do Rio de Janeiro	Rio de Janeiro	Brasil	http://portal.peq.coppe.ufrj.br/
71	SENAI CIMATEC	SENAI CIMATEC	Bahia	Brasil	http://www.senaicimatec.com.br/
72	SINMEC - Laboratório de Simulação Numérica em Mecânica dos Fluidos e Transferência de Calor,	Universidade Federal de Santa Catarina	Florianópolis - Santa Catarina	Brasil	http://www.sinmec.ufsc.br/site/index.php
73	Tanque de Provas Numérico - TPN	Universidade de São Paulo	Sao Paulo	Brasil	http://tpn.usp.br/
74	UERJ - Geologia	Universidade do Estado do Rio de Janeiro	Rio de Janeiro	Brasil	http://www.fgel.uerj.br/
75	UFF – Eng. Mecânica	Universidade Federal Fluminense	Rio de Janeiro	Brasil	www.uff.br

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

76	Universidad Federal de Alagoas	Universidad Federal de Alagoas	Alagoas	Brasil	http://www.ufal.edu.br/
77	Universidad Federal de Lavras	Universidad Federal de Lavras	Minas Gerais	Brasil	http://www.ufla.br/
78	Universidade Estadual do Norte Fluminense Darcy Ribeiro	Universidade Estadual do Norte Fluminense Darcy Ribeiro	Rio De Janeiro	Brasil	http://www.uenf.br/
79	Universidade Federal de Ouro Preto	Universidade Federal de Ouro Preto	Minas Gerais	Brasil	www.cpgg.ufba.br
80	CECAD - Centro de Computación de Alto Desempeño	Universidad Distrital "Francisco José de Caldas"	Bogota	Colombia	http://cecad.udistrital.edu.co/
81	Centro de Bioinformática y Biología Computacional (CBCB)	Centro de Bioinformática y Biología Computacional (CBCB)	Manizales	Colombia	http://bios.org.co
82	Centro de Computación Científica	Universidad EAFIT	Medellín	Colombia	http://www.eafit.edu.co/
83	Centro Internacional de Secuenciación Genómica y Bioingeniería de Colombia	Universidad del Cauca	Popayan	Colombia	http://cisgebi.org/
84	Computación de Alto Desempeño	Universidad de los Andes	Bogotá	Colombia	http://www.uniandes.edu.co/
85	Unidad de Supercomputación y Cálculo Científico	Universidad Industrial de Santander	Santander	Colombia	www.uis.edu.co/
86	Universidad Autónoma de Manizales	Universidad Autónoma de Manizales	Manizales	Colombia	http://www.autonoma.edu.co/
87	ZINE – Centro de Alto Rendimiento Computacional Javeriano	Pontificia Universidad Javeriana	Bogotá	Colombia	http://zine.javeriana.edu.co/
88	Centro Avanzado de Simulación de Procesos	Universidad del Bío-Bío	Región del Bío-Bío	Chile	http://industrial.ubiobio.cl/web/infraestructura/
89	Centro de Alto Rendimiento de la (CAR)	Universidad Católica del Norte	Antofagasta	Chile	http://www.ucn.cl/
90	Centro de Computación - CDC	Universidad de Antofagasta	Antofagasta	Chile	http://www.uantof.cl/
91	Centro de Innovación Tecnológica y de Computación de Alto Desempeño (CTI-HPC)	Universidad Técnica Federico Santa María	Valparaíso	Chile	http://www.usm.cl/investigacion/centros/innovacion-tecnologica-en-computacion-de-alto-desempeno/
92	Centro de Modelación y Computación Científica	Universidad de la Frontera - UFR	Temuco	Chile	http://www.ufro.cl/
93	Departamento de Astronomía y Astrofísica.	Pontificia Universidad Católica de Chile	Santiago de Chile	Chile	http://www.aiuc.puc.cl/index.php/computacion-de-alto-rendimiento-en-el-aiuc/
94	El Centro de Bioinformática y Simulación Molecular (CBSM)	Universidad de Talca	Talca	Chile	http://www.otalca.cl/link.cgi//CentrosTecnologicos/8044
95	Instituto Nacional de Hidráulica INH	Instituto Nacional de Hidráulica INH	Santiago De Chile	Chile	www.inh.cl
96	Laboratorio de Bioinformática y Matemática del Genoma (LBMG-Mathomics)	Universidad de Chile	Santiago de Chile	Chile	http://ingenieria.uchile.cl/investigacion/presentacion/laboratorios/centro-de-modelamiento-matematico/91399/laboratorio-de-bioinformatica-y-matematica-del-genoma-lbmg-mathomics
97	Laboratorio de Modelamiento en Imágenes Científicas y Visualización (MOTIV)	Universidad de Chile	Santiago de Chile	Chile	http://www.cmm.uchile.cl/?cmm_labs=motiv

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

98	Laboratorio Nacional de Computación de Alto Rendimiento (NLHPC)	Universidad de Chile	Santiago de Chile	Chile	http://ingenieria.uchile.cl/investigacion/presentacion/laboratorios/centro-de-modelamiento-matematico/91346/laboratorio-nacional-de-computacion-de-alto-rendimiento-nlhpc
99	Servicio de Impuestos Internos-SII	Servicio de Impuestos Internos- SII	Santiago de Chile	Chile	www.sii.cl
100	Universidad Andres Bello	Universidad Andres Bello	Santiago De Chile	Chile	http://www.unab.cl/
101	Clúster de Supercomputación ESPE	Universidad de las Fuerzas Armadas	Quito	Ecuador	http://ugi.espe.edu.ec/
102	Laboratorio De Modelización Computacional De Biomoléculas	Escuela Politécnica Nacional	Quito	Ecuador	http://www.epn.edu.ec/laboratorio-de-modelizacion-computacional-de-biomoleculas/
103	Laboratorio Nacional De Cálculo Científico	Centro de Modelización Matemática en Áreas Clave para el Desarrollo (MODEMAT)	Quito	Ecuador	http://www.modemat.epn.edu.ec
104	Red Nacional de Investigación y Educación de Ecuador	Red Nacional de Investigación y Educación de Ecuador	Cuenca	Ecuador	https://www.cedia.org.ec/cluster/
105	Universidad de Investigación de Tecnología Experimental Yachay	Universidad de Investigación de Tecnología Experimental Yachay	Yachay	Ecuador	http://www.yachaytech.edu.ec/home/
106	Centro de Alto Rendimiento Computacional de la Amazonía Peruana	Instituto de Investigaciones de la Amazonía Peruana	Iquitos	Perú	www.iiap.org.pe/manati/
107	Corporación Peruana de Aeropuertos y Aviación Comercial (CORPAC)	Corporación Peruana de Aeropuertos y Aviación Comercial (CORPAC)	Lima	Perú	Sin datos
108	Facultad	Universidad San Luis Gonzaga de Ica	Ica	Perú	Sin datos
109	Laboratorio de Dinámica de Fluidos Geofísicos Computacional	Instituto Geofísico del Perú (IGP)	Lima	Perú	www.igp.gob.pe
110	NASA Laser Tracking Station Arequipa	Universidad Nacional San Agustín de Arequipa - UNSA	Arequipa	Perú	www.unsa.edu.pe
111	Subdirección de Modelamiento Numérico de la Atmosfera de la Dirección de Redes de Observación y Datos	Servicio Nacional de Meteorología e Hidrología de Perú (SENAMHI)	Lima	Perú	www.senamhi.gob.pe
112	Centro Nacional de Cálculo Científico - CeCaICULA	Universidad de Los Andes	Mérida	Venezuela	www.ula.ve/
113	Petroleos de Venezuela SA	Petroleos de Venezuela SA	Caracas	Venezuela	http://www.pdvsa.com/
114	Cluster FING	Universidad de la República	Montevideo	Uruguay	https://www.fing.edu.uy

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

ANEXO 02: Listado de las 100 primeras supercomputadoras del Top500 Noviembre del 2016

Ran k	Name	Computer	Site	Manufactur er	Rmax	Rpeak	Operating System
1	Sunway TaihuLight	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	National Supercomputing Center in Wuxi	NRCPC	93014593.9	125435904	Sunway RaiseOS 2.0.5
2	Tianhe-2 (MilkyWay-2)	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	National Super Computer Center in Guangzhou	NUDT	33862700	54902400	Kylin Linux
3	Titan	Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	DOE/SC/Oak Ridge National Laboratory	Cray Inc.	17590000	27112550	Cray Linux Environment
4	Sequoia	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	DOE/NNSA/LLNL	IBM	17173224	20132659.2	Linux
5	Cori	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	DOE/SC/LBNL/NE RSC	Cray Inc.	14014700	27880653	Cray Linux Environment
6	Oakforest-PACS	PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	Joint Center for Advanced High Performance Computing	Fujitsu	13554600	24913459	Linux
7		K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	RIKEN Advanced Institute for Computational Science (AICS)	Fujitsu	10510000	11280384	Linux
8	Piz Daint	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100	Swiss National Supercomputing Centre (CSCS)	Cray Inc.	9779000	15987968	Cray Linux Environment
9	Mira	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	DOE/SC/Argonne National Laboratory	IBM	8586612	10066330	Linux
10	Trinity	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	DOE/NNSA/LANL/ SNL	Cray Inc.	8100900	11078861	Cray Linux Environment
11		Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	United Kingdom Meteorological Office	Cray Inc.	6765250	8128512	Cray Linux Environment
12	Marconi Intel Xeon Phi	CINECA Cluster, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	CINECA	Lenovo	6223040	10832998	Linux
13	Pleiades	SGI ICE X, Intel Xeon E5-2670/E5-2680v2/E5-2680v3/E5-2680v4 2.6/2.8/2.5/2.4 GHz, Infiniband FDR	NASA/Ames Research Center/NAS	HPE/SGI	5951550	7107148.8	SUSE Linux Enterprise Server 11
14	Hazel Hen	Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect	HLRS - Höchstleistungsrechenzentrum Stuttgart	Cray Inc.	5640170	7403520	Cray Linux Environment
15	Shaheen II	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	King Abdullah University of Science and Technology	Cray Inc.	5536990	7235174	Cray Linux Environment

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

16	Pangea	SGI ICE X, Xeon Xeon E5-2670/ E5-2680v3 12C 2.5GHz, Infiniband FDR	Total Exploration Production	HPE/SGI	5283110	6712320	SUSE Linux Enterprise Server 11
17	Stampede	PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P	Texas Advanced Computing Center/Univ. of Texas	Dell	5168110	8520111.6	Linux
18	Theta	Cray XC40, Intel Xeon Phi 7230 64C 1.3GHz, Aries interconnect	DOE/SC/Argonne National Laboratory	Cray Inc.	5095840	8626176	Cray Linux Environment
19	JUQUEEN	BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	Forschungszentrum Juelich (FZJ)	IBM	5008857	5872025.6	Linux
20	Cheyenne	SGI ICE XA, Xeon E5-2697v4 18C 2.3GHz, Infiniband EDR	National Center for Atmospheric Research (NCAR)	HPE/SGI	4788190	5332320	SUSE Linux Enterprise Server 12 SP1
21	Vulcan	BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	DOE/NNSA/LLNL	IBM	4293306	5033165	Linux
22	Abel	Cray XC30, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	Petroleum Geo-Services	Cray Inc.	4042460	5369856	Cray Linux Environment
23		Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	ECMWF	Cray Inc.	3944680	4249325	Cray Linux Environment
24		Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	ECMWF	Cray Inc.	3944680	4249325	Cray Linux Environment
25		Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40	Government	Cray Inc.	3577000	6131840	Linux
26		Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40	Government	Cray Inc.	3577000	6131840	Linux
27	Topaz	SGI ICE X, Xeon E5-2699v3 18C 2.3GHz, Infiniband FDR	ERDC DSRC	HPE/SGI	3318950	4570560	SUSE Linux Enterprise Server 11
28	DGX SATURNV	NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	NVIDIA Corporation	Nvidia	3307000	4896512	Ubuntu 14.04
29	HPC2	iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, NVIDIA K20x	Exploration & Production - Eni S.p.A.	IBM	3188000	4605000	Linux
30	SORA-MA	Fujitsu PRIMEHPC FX100, SPARC64 Xlfx 32C 1.98GHz, Tofu interconnect 2	Japan Aerospace eXploration Agency	Fujitsu	3157000	3481056	Linux
31		Cray XC30, Intel Xeon E5-2697v2 12C 2.7GHz, Aries interconnect	Government	Cray Inc.	3143520	4881254.4	Cray Linux Environment
32	Thunder	SGI ICE X, Xeon E5-2699v3/E5-2697 v3, Infiniband FDR, NVIDIA	Air Force Research Laboratory	HPE/SGI	3126240	5610481	SUSE Linux Enterprise Server 11

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

33	Camphor 2	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	Academic Center for Computing and Media Studies (ACCMS), Kyoto University	Cray Inc.	3057350	5483520	Cray Linux Environment
34	Mistral	bullx DLC 720, Xeon E5-2680v3 12C 2.5GHz/E5-2695V4 18C 2.1GHz, Infiniband FDR	DKRZ - Deutsches Klimarechenzentrum	Bull, Atos Group	3010680	3962880	Bullx Linux
35		Fujitsu PRIMEHPC FX100, SPARC64 Xlfx 32C 2.2GHz, Tofu interconnect 2	Information Technology Center, Nagoya University	Fujitsu	2910000	3244032	Linux
36	SuperMUC	iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR	Leibniz Rechenzentrum	IBM/Lenovo	2897000	3185050	Linux
37	SuperMUC Phase 2	NeXtScale nx360M5, Xeon E5-2697v3 14C 2.6GHz, Infiniband FDR14	Leibniz Rechenzentrum	Lenovo/IBM	2813620	3578266	Linux
38		Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	United Kingdom Meteorological Office	Cray Inc.	2801782	3019162	Cray Linux Environment
39		Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	United Kingdom Meteorological Office	Cray Inc.	2801782	3019162	Cray Linux Environment
40	TSUBAME 2.5	Cluster Platform SL390s G7, Xeon X5670 6C 2.93GHz, Infiniband QDR, NVIDIA K20x	GSIC Center, Tokyo Institute of Technology	NEC/HPE	2785000	5617483	SUSE Linux Enterprise Server 11
41	LLNL CTS-1 Quartz	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Lawrence Livermore National Laboratory	Penguin Computing	2632510	3207859.2	TOSS
42	LLNL/NNSA CTS-1 Jade	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Lawrence Livermore National Laboratory	Penguin Computing	2632510	3207859.2	TOSS
43	Tianhe-1A	NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050	National Supercomputing Center in Tianjin	NUDT	2566000	4701000	Linux
44	cascade	Atipa Visione IF442 Blade Server, Xeon E5-2670 8C 2.600GHz, Infiniband FDR, Intel Xeon Phi 5110P	DOE/SC/Pacific Northwest National Laboratory	Atipa Technology	2539130	3388032	Linux
45	Excalibur	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	Army Research Laboratory DoD Supercomputing Resource Center (ARL DSRC)	Cray Inc.	2485000	3682355.2	Cray Linux Environment
46	Nuri	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	Korea Meteorological Administration	Cray Inc.	2395680	2895360	Cray Linux Environment
47	Miri	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	Korea Meteorological Administration	Cray Inc.	2395680	2895360	Cray Linux Environment
48	Plasma Simulator	Fujitsu PRIMEHPC FX100, SPARC64 Xlfx 32C 1.98GHz, Tofu interconnect 2	National Institute for Fusion Science (NIFS)	Fujitsu	2376000	2621030	Linux

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

49	Makman-2	Dell PowerEdge R630, Xeon E5-2680v3 12C 2.5GHz, Infiniband QDR	Saudi Aramco	Dell	2249680	3041280	Redhat Enterprise Linux 6.5
50	Prolix2	bullx DLC 720, Xeon E5-2698v4 20C 2.2GHz, Infiniband FDR	Meteo France	Bull, Atos Group	2167990	2534400	bullx SCS
51	Beaufix2	bullx DLC 720, Xeon E5-2698v4 20C 2.2GHz, Infiniband FDR	Meteo France	Bull, Atos Group	2157410	2585088	bullx SCS
52	Lomonosov 2	T-Platform A-Class Cluster, Xeon E5-2697v3 14C 2.6GHz, Infiniband FDR, Nvidia K40m	Moscow State University - Research Computing Center	T-Platforms	2102000	2962253	Linux
53	Tianhe-2 LvLiang Solution	Tianhe-2 LvLiang, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Intel Xeon Phi 31S1P	LvLiang Cloud Computing Center	NUDT	2071390	3074534.4	Kylin Linux
54		SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR	Japan Atomic Energy Agency (JAEA)	HPE/SGI	1929370	2409600	SUSE Linux Enterprise Server 11
55	Tera-1000-1	bullx DLC 720, Xeon E5-2698v3 16C 2.3GHz, Infiniband FDR	Commissariat a l'Energie Atomique (CEA)	Bull, Atos Group	1871000	2586010	Bullx Linux
56	Fermi	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	CINECA	IBM	1788878	2097152	Linux
57	Marconi Intel Xeon	Lenovo NeXtScale nx360M5, Xeon E5-2697v4 18C 2.3GHz, Omni-Path	CINECA	Lenovo	1723890	2003098	Linux
58	SwiftLucy	Cluster Platform 3000 BL460c Gen9, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR	Government	HPE	1703280	2304000	Linux
59	Prometheus	HP Apollo 8000, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, NVIDIA Tesla K40	Cyfronet	HPE	1670090	2348640	Linux
60	Edison	Cray XC30, Intel Xeon E5-2695v2 12C 2.4GHz, Aries interconnect	DOE/SC/LBNL/NE RSC	Cray Inc.	1654702	2569420.8	Cray Linux Environment
61	ARCHER	Cray XC30, Intel Xeon E5 v2 12C 2.700GHz, Aries interconnect	EPSRC/University of Edinburgh	Cray Inc.	1642536	2550528	Cray Linux Environment
62	Luna	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	NOAA	Cray Inc.	1635020	2036736	Cray Linux Environment
63	Surge	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	NOAA	Cray Inc.	1635020	2036736	Cray Linux Environment
64	Occigen	bullx DLC, Xeon E5-2690v3 12C 2.6GHz, Infiniband FDR	Grand Equipement National de Calcul Intensif - Centre Informatique National de	Bull, Atos Group	1628770	2102630.4	Redhat Enterprise Linux 6.4

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

66	LANL CTS-1 Grizzly	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Los Alamos National Laboratory	Penguin Computing	1524720	1792627	TOSS
67	Salomon	SGI ICE X, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR, Intel Xeon Phi 7120P	IT4Innovations National Supercomputing Center, VSB-Technical University of Ostrava	HPE/SGI	1457730	2011641	CentOS
68	Blue Joule	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	Science and Technology Facilities Council - Daresbury Laboratory	IBM	1431102	1677722	Linux
69	JURECA	T-Platforms V-Class, Xeon E5-2680v3 12C 2.5GHz, Infiniband EDR/ParTec ParaStation ClusterSuite, NVIDIA Tesla K80/K40	Forschungszentrum Juelich (FZJ)	T-Platforms	1424720	1693440	CentOS
70	Spirit	SGI ICE X, Xeon E5-2670 8C 2.600GHz, Infiniband FDR	Air Force Research Laboratory	HPE/SGI	1415470	1530547.2	Redhat Enterprise Linux 6
71	Piz Dora	Cray XC40, Xeon E5-2695v4 18C 2.1GHz, Aries interconnect	Swiss National Supercomputing Centre (CSCS)	Cray Inc.	1410730	1509581	Cray Linux Environment
72	Beskow	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect	KTH - Royal Institute of Technology	Cray Inc.	1397000	1973657.6	Cray Linux Environment
73	Sid	bullx DLC 720, Xeon E5-2695v4 18C 2.1GHz, Infiniband FDR	Atos	Bull, Atos Group	1363480	1676506	bullx SCS
74	Curie thin nodes	Bullx B510, Xeon E5-2680 8C 2.700GHz, Infiniband QDR	CEA/TGCC-GENCI	Bull, Atos Group	1359000	1667174	bullx SUpperCOMputer Suite A.E.2.1
75	Cobalt	bullx DLC 720, Xeon E5-2680v4 14C 2.4GHz, Infiniband EDR	Commissariat a l'Energie Atomique (CEA)/CCRT	Bull, Atos Group	1299470	1479475	bullx SCS
76	Conrad	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Intel Xeon Phi 5120D	Navy DSRC	Cray Inc.	1285600	2049602	Cray Linux Environment
77	Gordon	Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect , Intel Xeon Phi 5120D	Navy DSRC	Cray Inc.	1285600	2049602	Cray Linux Environment
78		iDataPlex DX360M4, Intel Xeon E5-2680v2 10C 2.800GHz, Infiniband FDR	Max-Planck-Gesellschaft MPI/IPP	Lenovo/IBM	1283311.9	1463168	Linux
79	Nebulae	Dawning TC3600 Blade System, Xeon X5650 6C 2.66GHz, Infiniband QDR, NVIDIA 2050	National Supercomputing Centre in Shenzhen (NSCS)	Sugon	1271000	2984300	Linux
80	Yellowstone	iDataPlex DX360M4, Xeon E5-2670 8C	National Center for Atmospheric Research (NCAR)	IBM	1257615	1503590	Linux
81	Sisu	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	CSC (Center for Scientific Computing)	Cray Inc.	1250000	1689292.8	Cray Linux Environment

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

82	Helios	Bullx B510, Xeon E5-2680 8C 2.700GHz, Infiniband QDR	International Fusion Energy Research Centre (IFERC), EU (F4E) - Japan Broader Approach collaboration	Bull, Atos Group	1237000	1524096	bullx SuperCOmputer Suite A.E.2.1
83	Diego	bullx DLC 720, Xeon E5-2698v4 20C 2.2GHz, Infiniband FDR	Atos	Bull, Atos Group	1225280	1647360	bullx SCS
84	SNL/NNSA CTS-1 Serrano	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Sandia National Laboratories	Penguin Computing	1223700	1357171	TOSS
85	SNL/NNSA CTS-1 Cayenne	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Sandia National Laboratories	Penguin Computing	1223700	1357171	TOSS
86	Sekirei	SGI ICE XA, Xeon E5-2680v3 12C 2.5GHz, Infiniband FDR	University of Tokyo/Institute for Solid State Physics	HPE/SGI	1178300	1520640	SUSE Linux Enterprise Server 11
87	Garnet	Cray XE6, Opteron 16C 2.500GHz, Cray Gemini interconnect	ERDC DSRC	Cray Inc.	1167000	1505280	Linux
88		Lenovo RD450, Xeon E5-2670v3 12C 2.3GHz, 10G Ethernet	Service Provider T	Lenovo	1149570	2870400	Linux
89	LANL/NNSA CTS-1 Ice	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Los Alamos National Laboratory	Penguin Computing	1144840	1332979	TOSS
90	LANL/NNSA CTS-1 Fire	Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path	Los Alamos National Laboratory	Penguin Computing	1144840	1332979	TOSS
91	Cielo	Cray XE6, Opteron 6136 8C 2.40GHz, Custom	DOE/NNSA/LANL/SNL	Cray Inc.	1110000	1365811.2	Linux
92		Dell C4130, Xeon E5-2630v3 8C 2.4GHz, Infiniband FDR, NVIDIA Tesla K80	Energy Company (DD)	Dell	1106000	1748390	CentOS
93		Lenovo ThinkServer RD650, Xeon E5-2680v3 12C 2.5GHz, 10G Ethernet	Internet Company A	Lenovo	1101600	3360000	Linux
94		Lenovo ThinkServer RD650, Xeon E5-2680v3 12C 2.5GHz, 10G Ethernet	Internet Company A	Lenovo	1101600	3360000	Linux
95	Magnus	Cray XC40, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect	Pawsey Supercomputing Centre, Kensington, Western Australia	Cray Inc.	1097558	1485619.2	Cray Linux Environment
96	Electra	SGI ICE X, Xeon E5-2680v4 14C 2.4GHz, Infiniband FDR	NASA/Ames Research Center/NAS	HPE/SGI	1096020	1238630	SUSE Linux Enterprise Server 11
97	Cartesius 2	Bullx DLC B710/B720 Blades, Xeon E5-2695 v2 2.4GHz E5-2690v3 2.6GHz, Infiniband FDR	SURFsara	Bull, Atos Group	1088510	1327104	bullx SCS
98		HP Apollo XL7x0f, Xeon E5-2697v3 14C 2.6GHz, Infiniband FDR	IT Service Provider	HPE	1085460	1509581	Linux
99	Turing	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	CNRS/IDRIS-GENCI	IBM	1073327	1258291	Linux
100	DiRAC	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	University of Edinburgh	IBM	1073327	1258291	Linux

ANEXO 03: Acceso de usuarios al MANATI IIAP.

Nombres y apellidos	Institución	Temática de estudio
Julio Rubén Barrios Llacuachaqui	Universidad Nacional de Ingeniería	Cálculos de Química Computacional
David Alfonso Laván Quiroz	INICTEL-UNI	Bioinformatica
Cristian José López Del Alamo	Universidad La Salle	Triangulaciones de Delaunay.
Jose Antonio Fiestas Iquira	Universidad de Ingeniería	Simulaciones de N-cuerpos
Jorge Ibáñez Gijón	Aix-Marseille Universite	Modelos biomecanicos de control del movimiento humano
Manuel Castillo-Cara	Universidad Nacional de Ingeniería	Para la localización en indoor tenemos 65000 posibles combinaciones de potencias y posiciones, estas combinaciones tardan aproximadamente 2 días.
Rel Guzman Apaza	Universidad Nacional San Agustín	Training & hyperparameter tuning de algoritmos de Machine Learning
Cesar Marchelo Bragagnini Mendizabal	Universidad Católica San Pablo	Deep Learning, Natural Language Processing, Image Caption, Sentiment Analysis, Computer Vision
Christian Solis Calero	Universidad Nacional Mayor de San Marcos	Química Computacional
Rayner Harold Montes Condori	Universidad de São Paulo	Procesamiento de imágenes.
Julio Cesar Mendoza Bobadilla	University of Campinas	Deep Learning
Josimar Edinson Chire Saire	Universidad Católica San Pablo	Experimentos de algoritmo evolutivo de inspiración cuántica con funciones benchmark y alta dimensionalidad
Elisban Flores Quenaya	Universidad Nacional del Altiplano/ Universidad Católica San Pablo	Evaluación feno-genotípica de alpacas
Mario Meza Hidalgo	Universidad Nacional Autónoma de Alto Amazonas	Prácticas en HPC
Junior Usca Huacasi	Universidad Nacional de San Agustín	Análisis de datos geoespaciales distribuidos de tweets
Gerardo Cristian Jacome Vergaray	SENAMHI	Predicción de eventos climáticos.
Disraeli Fausto Ari Mamani	Universidad Nacional de San Agustín Arequipa	Trabajos en procesamiento de lenguaje natural, específicamente el procesamiento en "Spanish Billion Words Corpus and Embeddings"
María Ximena Tagle Casapía	IIAP	Modelos de distribución de especies
Disraeli Fausto Ari Mamani	Universidad Nacional de San Agustín Arequipa	Deep Learning
Ricardo Rildo Coronado Pérez	Universidad Nacional San Agustín	Deep Learning en imágenes medicas
Alexander Ocsa	Universidad Nacional San Agustín	Deep Learning y Base de Datos en GPU
César Martín Cruz Salazar	Facultad de Ciencias - Universidad Nacional de Ingeniería	Ejecución de algoritmos matemáticos para cálculo de primos y valores propios de matrices de 400 x 400

INTRODUCCIÓN A LA SUPERCOMPUTACIÓN EN EL PERÚ

Christian Alex Mayhua Tijera	Universidad Católica San Pablo	Deep Learning
Aaron Eyzaguirre Valles	Facultad de Ingeniería de Sistemas e Informática (UNAP)	desarrollo de graficos en CUDA para problemas de física y matematicas
Billy Jack Alvarado Vela	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Deep Learning aplicada al estudio del clima
Jose Martin Ruiz Babilonia	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Procesamiento de imagen para uso de reconocimiento de aguajales
Martin Alonso Ocmin	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Pre - Procesamiento de Imagen satelitales
Kevin Roy Villacorta Valdez	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Procesamiento de Imagenes Satelitales
Adrian Vasquez Ocmin	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Crear un software que facilite el conteo de huevos de pescado.
Karen Isuiza Del Aguila	Facultad de Ingeniería de Sistemas e Informática (UNAP)	GESTIÓN DE LA SUPERCOMPUTADORA
Julio Martin Rojas Tenazoa	UNAP	Modelo matemático del PageRank para generar la importancia de cada pagina, usando Matlab
Luis Antonio Lomas Valles	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Deep Learning
Bryan Franco Acurio Rengifo	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Procesamiento de imagen
Fernando Andres Perea Quispe	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Adaptacion De Analizador De Secuencias A HPC
Kiper Pereyra Trigoso	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Renderizado de material multimedia
Marco Antonio Paredes Riveros	SENAMHI	Modelos del Tiempo
Jorge Walter Zvietcovich Diaz	SENAMHI	Implementación, adaptación y parametrización para el modelamiento climático.
Luis Pablo Marcelo Perea	Facultad de Ingeniería de Sistemas e Informática (UNAP)	Procesamiento de datos
Jemmy Dennis Vargas Maytahuari	Instituto de Investigaciones de la Amazonía Peruana	Análisis de datos

ANEXO 04: Cálculo de flops por computadora.

Fórmula para calcular los flops en un CPU [7]

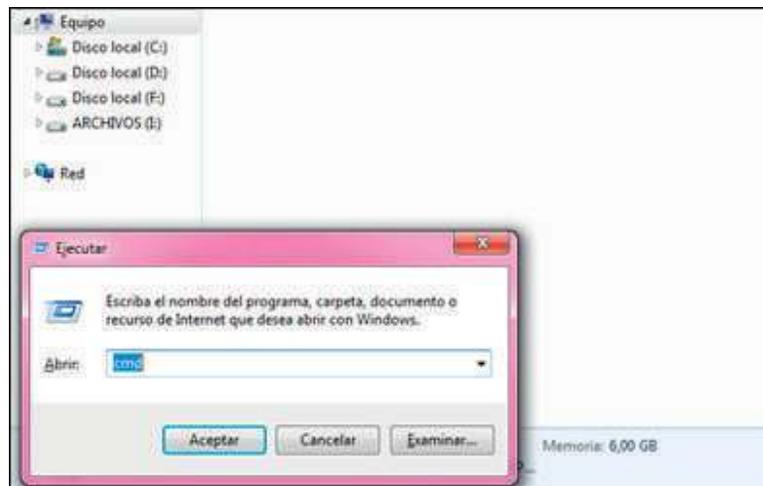
$$\text{GFLOP} = (\text{Velocidad de la CPU en GHz}) \times (\text{n}^\circ \text{ núcleos}) \times (\text{instrucciones por ciclo})$$

1. Para conocer los datos requeridos

1.1 Para Windows

a. Abrir símbolo del sistema

Tecla Windows + R



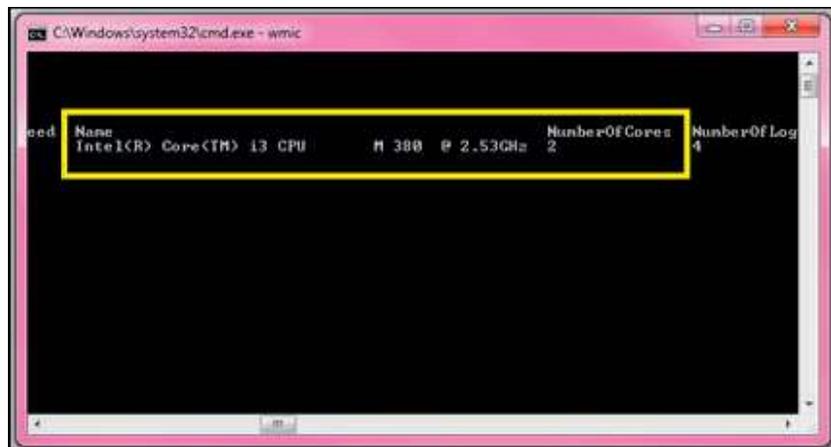
b. Escribir "wmi"



WMIC significa Windows Management Instrumentation Command-line [85], es una herramienta de administración para Windows que permite no solo obtener información sino realizar acciones.

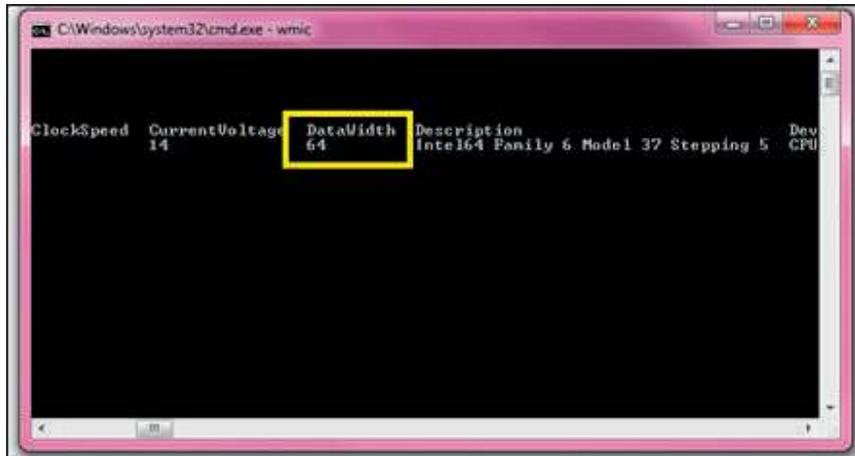


c. Escribir "cpu", a continuación se muestran todos los datos de la CPU, buscar: Name, NumberOfCore.



d. Ahora tenemos los datos:

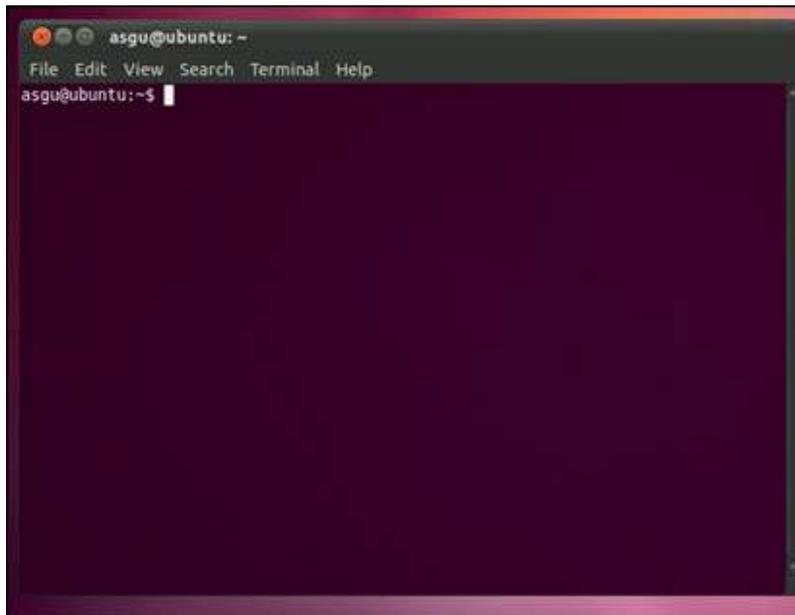
- Velocidad de la CPU: 2.53 GHz
- Número de Núcleos (NumberOfCores): 2
- Familia: Intel Core
- Arquitectura: 64 bits (Precisión Doble)



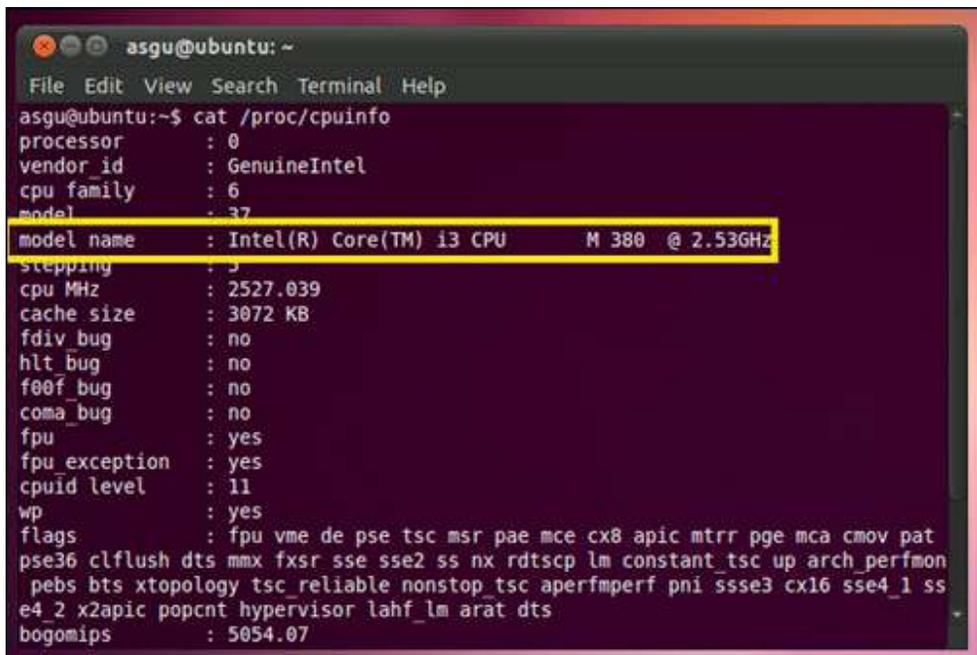
1.2. Para Linux

a. Acceder al terminal del sistema

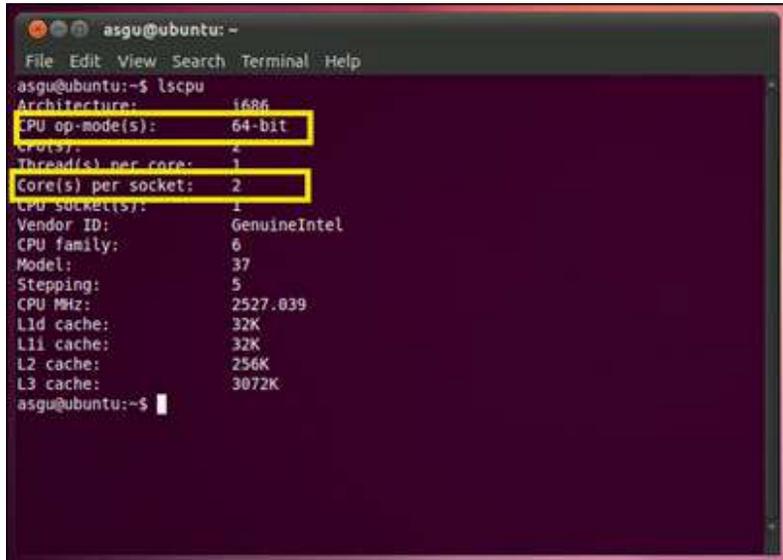




b. Teclear “cat proc/cpuinfo”



c. Teclear “lscpu”



```
asgu@ubuntu: ~  
File Edit View Search Terminal Help  
asgu@ubuntu:~$ lscpu  
Architecture:          i686  
CPU op-mode(s):      64-bit  
Cpus(s):              2  
Thread(s) per core:  1  
Core(s) per socket:  2  
CPU socket(s):       1  
Vendor ID:            GenuineIntel  
CPU family:           6  
Model:                37  
Stepping:             5  
CPU MHz:              2527.039  
L1d cache:            32K  
L1i cache:            32K  
L2 cache:             256K  
L3 cache:             3072K  
asgu@ubuntu:~$
```

d. Ahora tenemos los datos:

- Velocidad de la CPU: 2.53 GHz
- Número de Núcleos (Core(s) per socket): 2
- Familia: Intel Core
- Arquitectura: 64 bits (Precisión Doble)

2. Instrucciones por ciclo por familia de procesadores

Tabla 34: Instrucciones por ciclo (IPC) de algunas familias de procesadores.

Familia CPU	Precisión Doble	Precisión Simple
Intel Core and Intel Nehalem	4 DP IPC	8 SP IPC
Intel Sandy Bridge and Intel Ivy Bridge	8 DP IPC	16 SP IPC
Intel Haswell, Intel Broadwell and Intel Skylake	16 DP IPC	32 SP IPC
AMD K10	4 DP IPC	8 SP IPC
AMD Bulldozer, AMD Piledriver and AMD Steamroller, per module (two cores)	8 DP IPC	16 SP IPC
Intel Atom (Bonnell, Saltwell, Silvermont and Goldmont)	2 DP IPC	4 SP IPC
AMD Bobcat	2 DP IPC	4 SP IPC
AMD Jaguar	4 DP IPC	8 SP IPC
ARM Cortex-A7	1 DP IPC	8 SP IPC
ARM Cortex-A9	1 DP IPC	8 SP IPC
ARM Cortex-A15	1 DP IPC	8 SP IPC
ARM Cortex-A32	2 DP IPC	8 SP IPC
ARM Cortex-A35	2 DP IPC	8 SP IPC
ARM Cortex-A53	2 DP IPC	8 SP IPC
ARM Cortex-A57	2 DP IPC	8 SP IPC
ARM Cortex-A72	2 DP IPC	8 SP IPC
Qualcomm Krait	1 DP IPC	8 SP IPC
Qualcomm Kryo	2 DP IPC	8 SP IPC
IBM PowerPC A2 (Blue Gene/Q), per core	8 DP IPC	8 DP IPC
IBM PowerPC A2 (Blue Gene/Q), per thread	4 DP IPC	4 DP IPC
Intel Xeon Phi (Knights Corner), per core	16 DP IPC	32 SP IPC
Intel Xeon Phi (Knights Corner), per thread (two per core)	8 DP IPC	16 SP IPC

Fuente: <http://stakoverflow.com/questions/15655835/flops-per-cycle-for-sandy-bridge-and-haswell-sse2-avx-avx2>

3. Con los datos obtenidos, calculamos los gigaflops.

$$\text{GFLOPS} = 2.53 \times 2 \times 4$$

$$\text{GFLOPS} = 20.24$$

Es así que podemos concluir que el rendimiento de nuestra computadora es de 20.24 gigaflops.

PROGRAMA BIOINFO - INSTITUTO DE INVESTIGACIONES DE LA AMAZONIA PERUANA



PERÚ

Ministerio
del Ambiente



CONCYTEC
CONSEJO NACIONAL DE CIENCIA,
TECNOLOGÍA Y INNOVACIÓN TECNOLÓGICA



CIENCIAACTIVA

Scholarships and Grants by Concytec

ISBN: 978-9972-667-99-2



9 789972 667992