

Córdoba 15 de mayo del 2020

Centro de Computación de Alto Desempeño

Informe de gestión 2019

Director: Oscar Reula

El Centro de Computación de Alto Desempeño (CCAD) de la Universidad Nacional de Córdoba es un organismo creado en 2013 que tiene como finalidad la de administrar los recursos de computación de alto desempeño dentro de la Universidad Nacional de Córdoba. Consta de: un Directorio -integrado por miembros de las Facultades fundadoras, de la Prosecretaría de Informática y del CCT-Córdoba; un Director; Secretarios Técnico y Administrativo, y un Consejo Científico encargado de diseñar las políticas de uso y asignación de tiempo de cómputo. El CCAD posee una página web (<http://ccad.unc.edu.ar/>) con información detallada de sus equipos y actividades.

Durante el 2019 el Centro de Computación de Alto Desempeño realizó las siguientes tareas:

1. Equipamiento:

a. Incorporación de Nabucodonosor 2

Se adaptó un servidor de Mendieta para poder alojar 3 placas NVIDIA GeForce GTX 1080 Ti de 10 TFLOPS de potencia de cálculo pico de simple precisión. Dos placas fueron donadas por la Oficina Córdoba de Mercado Libre mientras que la tercera por el particular Hernán Amiune. La máquina tiene 2 TiB de almacenamiento de estado sólido donadas por el Córdoba Technology Cluster.

b. Armado de Nabucodonosor 3

Se inició el armado de Nabucodonosor 3 con tecnología más nueva. Se mantiene el uso de los servidores de cálculo de Mendieta, pero se actualizan las GPUs a GeForce RTX 2080 Ti y discos SSD de 2 TiB SATA3. Esto duplicará la potencia de cálculo y la capacidad de almacenamiento. La empresa Oixio dona una GPU y la empresa Bitlogic dona un SSD. Faltan 2 GPUs y 2 discos para armar un nuevo nodo.

c. Ampliación de Eulogia

Incorporación de un nuevo nodo de 4 KNLs 7250 en Eulogia con una potencia de 4.85 TFLOPS, sumando en conjunto 24.25 TFLOPS Rmax.

d. Ampliación Eulogia

Se inicia la ampliación de Eulogia con nodos a un precio muy bajo conseguidos en EEUU. Se realiza la compra de 3 chasis con 12 nodos KNL 7120 sin RAM y sin disco por un total de USD 6000 a la empresa Lambda Labs de San Francisco, California. Los nodos llegarán en 2020.

e. Adquisición de un nuevo NAS

El equipamiento son dos servidores de archivos idénticos con las siguientes características técnicas: SuperStorage 6049P-E1CR24L de Supermicro con 24 bahías 3.5" SAS3/SATA3, 2 Xeon Silver 4108 de 8 núcleos, 12 memorias 8 GiB DDR4 ECC 2400, totalizando 96 GiB de RAM, Disco estado sólido Intel DC P3100 256GiB PCIe 3.0 4x M.2.

f. Ampliación de almacenamiento de estado sólido

Se adquirieron 4 dispositivos de almacenamiento de estado sólido (SSD) de alto desempeño y durabilidad. Cada unidad es un Intel Optane SSD 900p de 280 GiB de capacidad. También se adquirió un SSD WD Blue 1 TiB SATA3 para sumar al segundo nodo de Nabucodonosor y completar los 3 TiB de almacenamiento sólido.

2. Formación de Recursos Humanos / Capacitación

a. Curso de Posgrado: R en el CCAD

Organizado por Maestría en Estadística Aplicada y el CCAD.
Disertantes: Lic. Juan Cruz Rodriguez (CIDIE-CONICET, FaMAF-UN) y Lic. Marcos Mazzini (CCAD, CONICET, IFEG, FaMAF)
Fecha: 20 y 21 de marzo y 24 y 25 de abril de 2019.

b. Minicurso de HPC en el Centro Atómico Bariloche

Del 5 al 7 de agosto de 2019, Carlos Bederián y Nicolás Wolovick, miembros del CCAD, fueron invitados por la Gerencia de Física del Centro Atómico Bariloche a dar un curso de HPC.

3. Otras Gestiones:

- a. Visita del Secretario de Articulación Científica de la Nación al Data Center
Secretario de Articulación Científica de la Nación, Agustín Campero, visitó el UNC Data Center para conocer el estado de avance de las obras. La visita la acompañaron Patricia Paredes (representante del CIN en el consejo asesor del SNCAD y miembro del CCAD-UNC), Carla Giacomelli (Secretaría de CyT-UNC), Miguel Montes (Prosecretario de Informática-UNC) y Marcos Mazzini (CPA, CCAD-UNC).

- b. Licitación Pública para la compra equipamiento de alto desempeño.
A través de la Resolución Rectoral [RR-2019-2617](#) el Rector de la UNC Hugo Juri convoca a la Licitación Pública para la adquisición de equipamiento de Computación de Alto Desempeño (CAD) por un monto de 386.000 dólares del proyecto PAGE dirigido por el Dr. Marcelo Mariscal.
- c. Participación en el NODO de Inteligencia Artificial CTC-FaMAF-CCAD.
- d. Participación del Consejo de la Diplomatura en Ciencia de Datos de la FaMAF.
- e. Traslado de lugar de trabajo y oficina del Director al Edificio Virtual de la Universidad.
- f. Re-empadronamiento de usuarios.
- g. Administración de uso exclusivo Invierno 2019 y Verano 2020.

4. CCAD en los medios

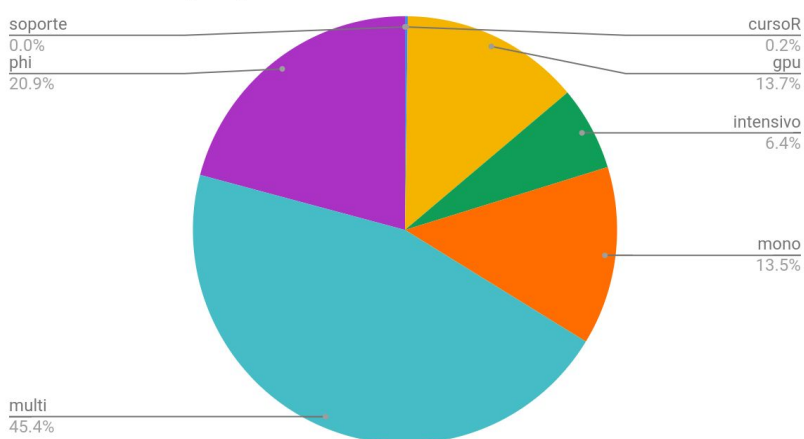
- a. Publicación en Nature (vol. 569, 2 de mayo de 2019) [LINK](#)
El Dr. Nicolás Wolovick miembro del CCAD-UNC cuenta cómo hacemos para sostener una infraestructura de HPC decente con un magro presupuesto.
- b. Artículo periodístico en Página 12 sobre HPC en Argentina [LINK](#)
El diario de tirada nacional Página 12 publicó en el suplemento Universidad del día 24 de octubre el artículo «Los cerebros de la Ciencia». En el, Ezequiel Bergonzi describe el panorama del HPC argentino con intervenciones de Esteban Mocskos (UBA, CONICET, CSC) y Nicolás Wolovick (UNC, FaMAF, CCAD).

Uso de los recursos computacionales

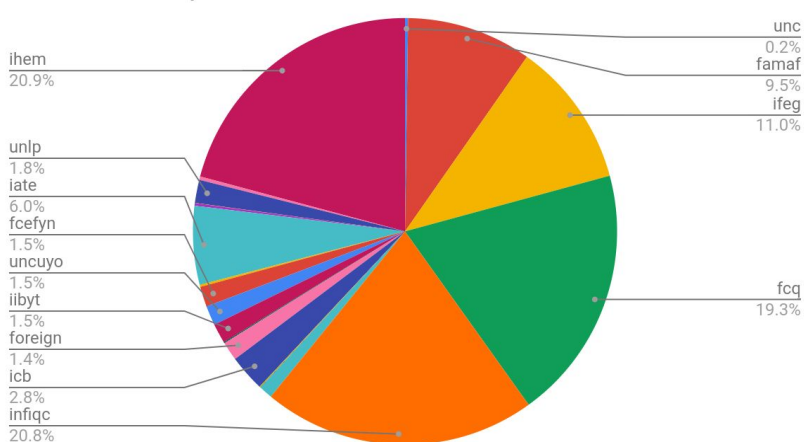
El **CCAD** administra varios recursos computacionales. Entre ellos, el de mayor importancia es el Cluster Mendieta del cual brindamos información sobre su uso ya que por su tamaño es el más relevante en cuanto a su impacto en el medio. La puesta en funcionamiento del mismo permitió el acceso a recursos informáticos de primer nivel a muchos investigadores que anteriormente no contaban con facilidades similares, suministrando 1,38 millones de horas de cómputo en el transcurso del año.

En la siguiente gráfica se muestra el uso del cluster durante el año 2019 de acuerdo a la pertenencia de los investigadores en distintas unidades. El gráfico reporta mayormente unidades ejecutivas del CCT-Córdoba ya que es una descripción más detallada que si la hiciéramos por Facultad.

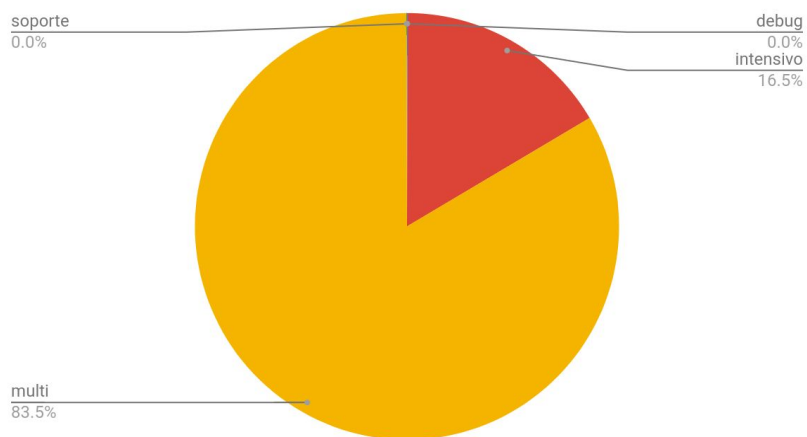
Uso Mendieta por partición



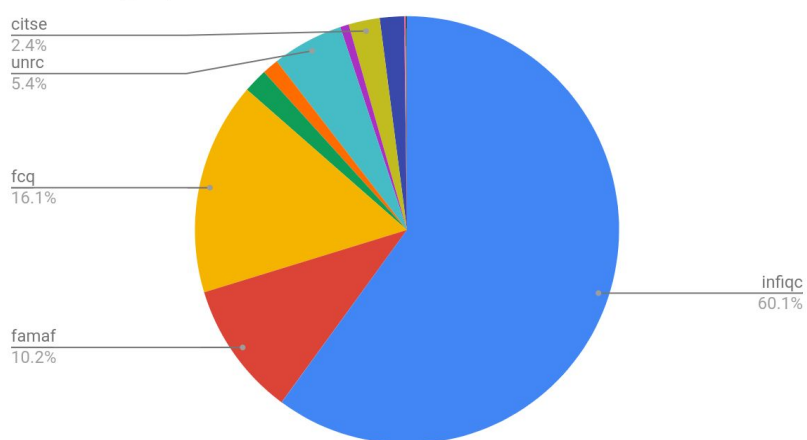
Uso Mendieta por Unidad Académica



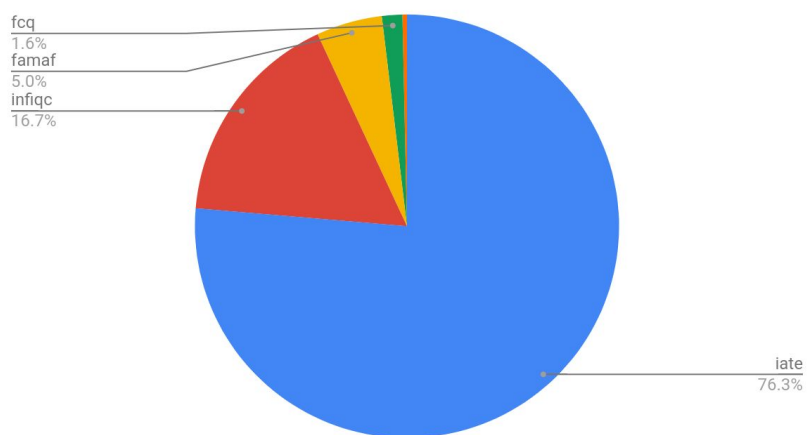
Uso Eulogia por partición



Uso Eulogia por Unidad Académica



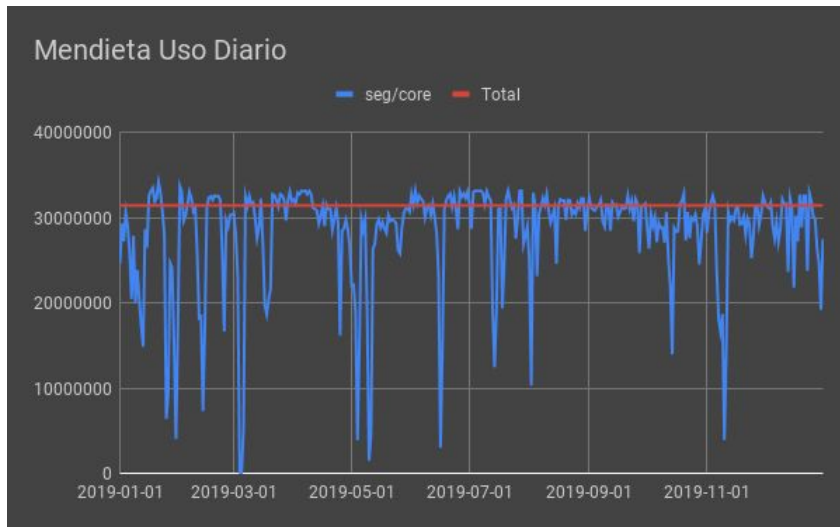
Uso Mulatona por Unidad Académica



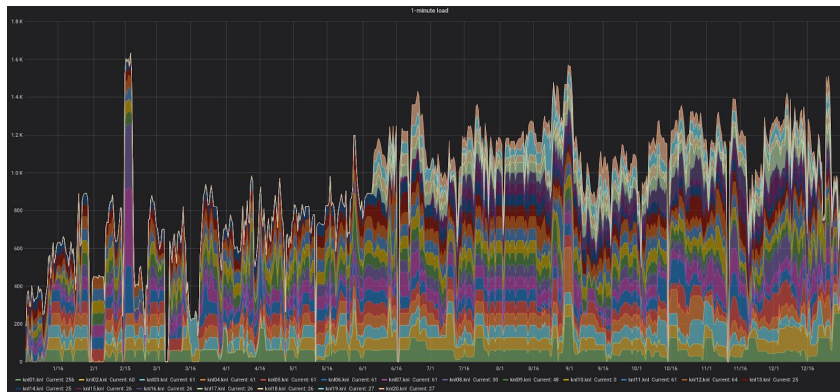
Cabe destacar que el uso de estos recursos no se circunscribe a la región, teniendo usuarios regulares de otras provincias, como los de la Facultad de Ciencias Exactas y Naturales (previamente Instituto de Ciencias Básicas) de la Universidad de Cuyo, del Centro de Investigación de Métodos Computacionales de Santa Fe, de otros países de la región

como los de la Universidad de la República de Uruguay, y de empresas privadas a las que el centro les presta servicios a través de acuerdos de vinculación tecnológica.

CLUSTER MENDIETA

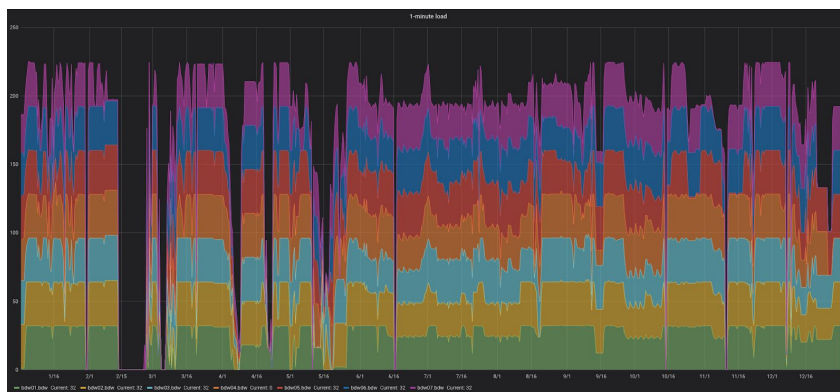


CLUSTER EULOGIA



NOTA: cabe destacar que, dado las características de las placas INTEL XEON PHI con 1312 cores total que pueden trabajar como hasta 5248 cores, la ocupación del cluster Eulogia no está correctamente reflejada en la gráfica anterior.

CLUSTER MULATONA



Conclusiones

El CCAD mantiene sus recursos computacionales con un porcentaje de uso muy alto, lo cual es bueno, ya que indica que estos, que caducan en un plazo bastante corto, son aprovechados de forma máxima. El impacto de estas facilidades ha sido muy grande, dando lugar a la culminación de numerosos trabajos científicos ([LINK](#)). A través del monitoreo del uso de los clusters vemos que el sistema está saturado de trabajos, siendo imperioso contar con mayores facilidades de cómputo en la Universidad, para evitar demoras en el proceso de investigación o incluso el uso de recursos computacionales de otras universidades.

Tareas para el corriente año (2020)

Durante este año los clusters han estado funcionando casi sin interrupción, en el presente año dedicaremos nuestros recursos en:

1. Mudanza del equipamiento al Data Center de la UNC
2. Puesta en funcionamiento de nuevo equipamiento
3. Reconversión de servidores para Inteligencia Artificial
4. Mejorar el sistema de almacenamiento de datos, dotándolo de una capacidad mayor de flujo de datos de entrada y salida y mayor capacidad de almacenamiento.
5. Gestionar diversas fuentes de financiamiento a fines de adquirir un nuevo Cluster que incremente los recursos de procesamiento y contribuya a desahogar el uso de los actuales.
6. Realización de Workshops y dictado de cursos de capacitación en programación paralela.

Anexo: Informe de la Encuesta a Usuarios

Durante el mes de Julio de 2019 se realizó una encuesta online a los usuarios sobre distintos puntos arrojando los siguientes resultados:

Metodología: Formulario on-line enviado por e-mail

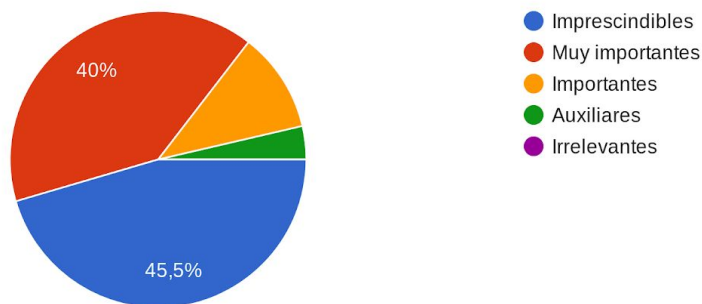
Cobertura: Se envió a **382** usuarios activos y no activos. Se recibieron **55 respuestas** de **17 instituciones** distintas.

En los últimos 3 años (01-01-2017) corrieron en mendieta **167** personas

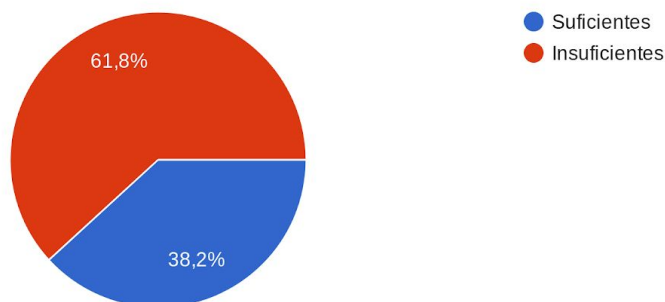
Usuarios activos 2019: **101**

Resumen de las respuestas:

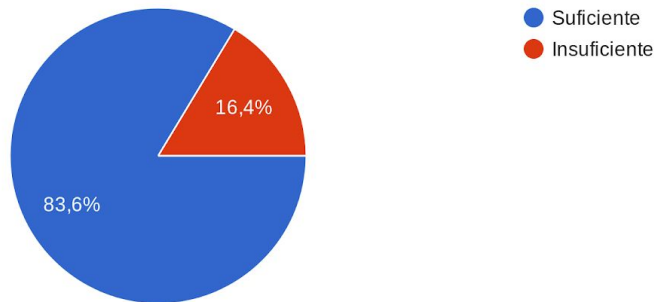
Los recursos del CCAD son para mi investigación
55 respuestas



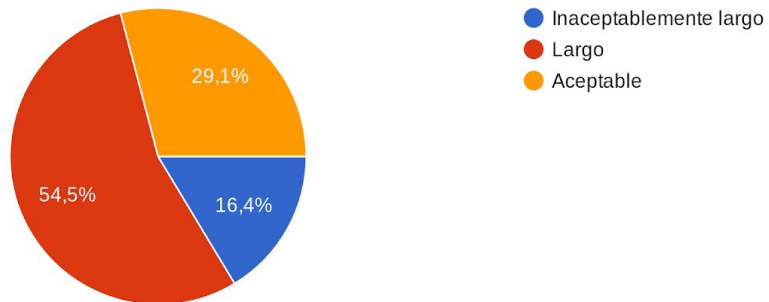
Los recursos de cómputo son
55 respuestas



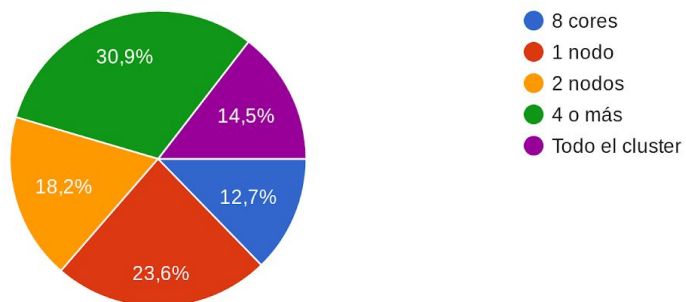
El almacenamiento es
55 respuestas



El tiempo de espera en cola es
55 respuestas



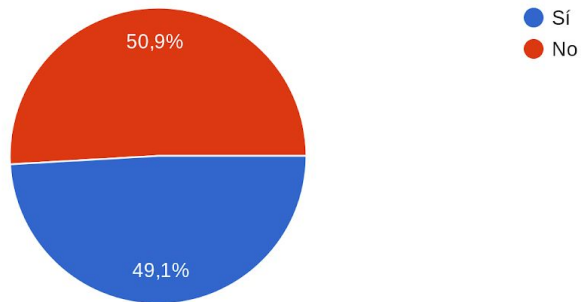
¿Cuántos recursos podría aprovechar si contara con uso exclusivo por 1 mes?
55 respuestas



La mitad de los usuarios programa:

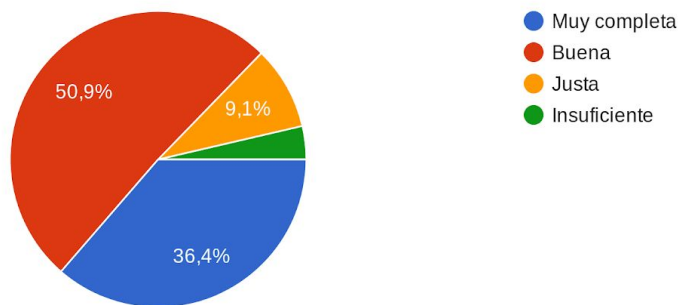
Desarrolla y mantiene códigos propios?

55 respuestas

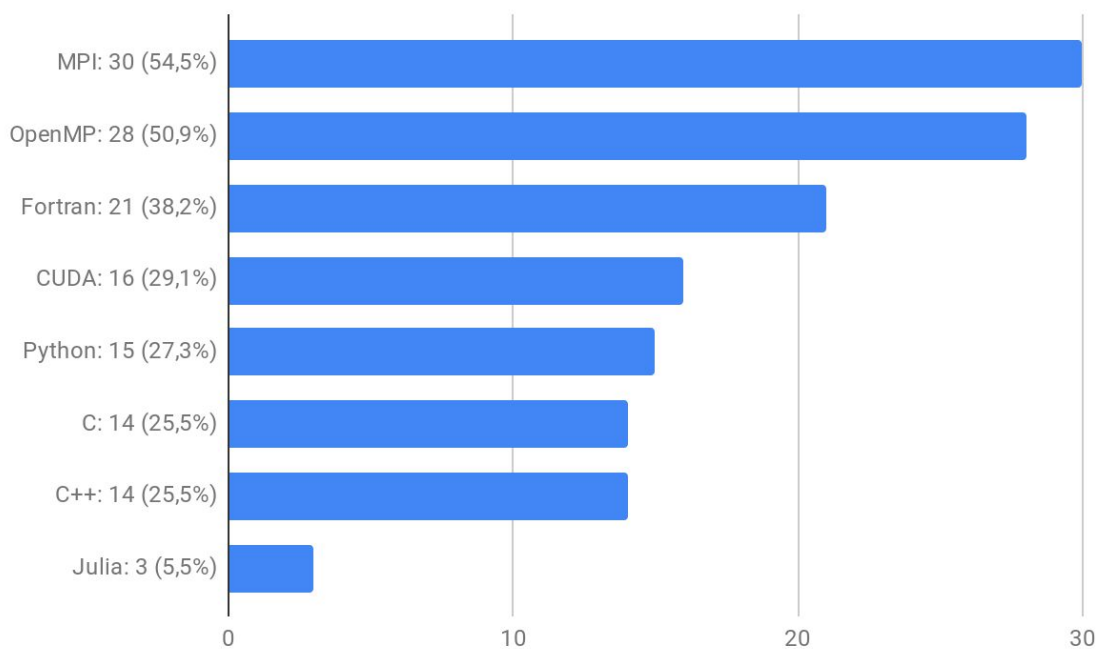


La variedad de herramientas (aplicaciones, compiladores, librerías) es

55 respuestas

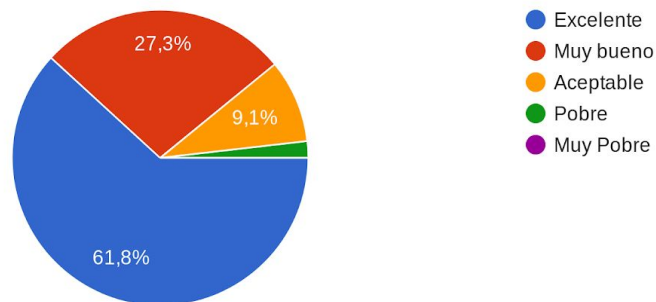


Lenguajes y Herramientas de paralelización (admite múltiples respuestas)



¿Como evalúa el desempeño del Soporte Técnico?

55 respuestas



Resumen de las sugerencias:

- Más cursos sobre uso básico de un cluster.
- Que el grupo de google funcione como foro para consultas (potenciar el intercambio entre usuarios)
- GPUs nuevas
- Fat node con 1tb de RAM
- Cola debug en eulogia (2 pedidos)
- Mejorar el throughput para jobs trivialmente paralelos