

Implementación de Clúster HPC como herramienta de cómputo distribuido en el Instituto Tecnológico de Acapulco

Gaddiel Fredy Flores Arteaga Ing.¹, Dr. Eduardo de la Cruz Gámez²,
M.T.I. Eloy Cadena Mendoza³ y M.C. Francisco Javier Gutierrez Mata⁴

Resumen—El presente artículo expone una estrategia para la implementación de un agrupamiento de computadoras (clúster) de alto rendimiento (HPC) por medio del cuál será posible realizar operaciones que exigen un alto poder de cómputo por su complejidad dentro del área científica, matemática o de ingeniería. Este artículo específicamente propone implementar esta herramienta para el desarrollo de criptoanálisis a sistemas de seguridad de sistemas operativos de categoría C1, aprovechando la ventajas del software de uso libre, así como equipo propio del Laboratorio de Cómputo del Instituto Tecnológico de Acapulco. Dentro del presente trabajo se mencionan los componentes de la infraestructura de comunicaciones, hardware y software más importantes del clúster HPC, así como la función que desempeñan dentro del mismo para poder ser implementados. Por último, se detalla su instalación, configuración, pruebas y resultados obtenidos.

Palabras clave—criptoanálisis, clúster, HPC.

Introducción

De acuerdo al portal Ranking web of world research centers (Cybermetrics Lab, 2018) el 95% de los centros de investigación en nuestro país forman parte de universidades o dependencias públicas. Día con día los investigadores realizan infinidad de operaciones y cálculos necesarios dentro de su campo de estudio, los cuales conforme avanzan se van tornando más complejos, ello obliga a ocupar mucho tiempo en procesos de cómputo para posteriormente poder obtener resultados satisfactorios. Como medida para optimizar tiempo en procesos de cómputo algunos centros de investigación de otros países tienen la capacidad económica para adquirir supercomputadoras, las cuales pueden estar en el rango de los millones de dólares.

Un clúster de computadoras de alto rendimiento utiliza potentes herramientas y realiza cómputo distribuido de aplicaciones de tal manera que puede analizar datos en investigaciones académicas casi con la misma velocidad y potencia de una supercomputadora, pero con un costo inferior y evitando el elevado costo por mantenimiento e inactividad.

Como una solución fiable y comparada a la obtención de una supercomputadora en este artículo se detallará la implementación un clúster de computadoras de alto rendimiento, el cual ha sido creado con hardware convencional que forma parte del Laboratorio de cómputo de la Maestría en Sistemas Computacionales del Instituto Tecnológico de Acapulco, así como herramientas de software de uso libre, que unidos por medio de una red de datos de alta velocidad ofrecerá a los investigadores ventajas sustanciales en comparación al desempeño ofrecido por computadoras personales, optimizando de esta manera el tiempo destinado a procesos de cómputo.

Estado del Arte

Existen actualmente varias universidades en México que han implementado estas herramientas de cómputo por las ventajas ofrecidas, a continuación, se detallan algunos ejemplos destacados.

Instituto Politécnico Nacional - CINVESTAV – Xihcoatl

El sistema está integrado actualmente por 253 servidores de tecnología híbrida. Es una supercomputadora con una capacidad de procesamiento teórica de 313 TFlop/s (313 billones de operaciones aritméticas por segundo). Cuenta con Procesamiento CPU x86 en 213 Nodos de cómputo. 67 Nodos AMD Interlagos 6274. 84 Nodos Intel X5675. 62 Nodos Intel E5-V4. Procesamiento GPU/Coprocesadores en 40 nodos de cómputo. 5 Nodos GPUs

¹ El Ing. Gaddiel Fredy Flores Arteaga es Estudiante de la Maestría en Sistemas Computacionales en el Instituto Tecnológico de Acapulco. gadflores@gmail.com

² El Dr. Eduardo de la Cruz Gámez es Jefe de la división de estudios de posgrado e investigación y Docente de la Maestría en Sistemas Computacionales en el Instituto Tecnológico de Acapulco. depi_acapulco@tecnm.mx

³ El M.T.I. Eloy Cadena Mendoza es Jefe del Laboratorio Cómputo y Docente de la Maestría en Sistemas Computacionales en el Instituto Tecnológico de Acapulco. eloy_cadena@yahoo.com

⁴ El M.C. Francisco Javier Gutierrez Mata es Jefe del Centro de Cómputo y Docente de la Maestría en Sistemas Computacionales en el Instituto Tecnológico de Acapulco. fcomata84@hotmail.com

NVIDIA 2070/2075 e Intel X5675. 12 Nodos GPUs NVIDIA K40 e Intel E5-2650L v3. 4 Nodos Xeon-Phi 7120P. 19 Nodos GPUs NVIDIA K80 Intel E5-2660 v3. (Centro de Investigación y Estudios Avanzados, 2019) (Centro de Investigación y Estudios Avanzados, 2019)

Laboratorio Nacional de Supercómputo del Sureste (LNS). – BUAP - Cuetlaxcoapan

Cuenta con cerca de 6796 núcleos en CPU, una RAM de 2048 GB y unos 11520 núcleos CUDA gracias a la incorporación de tarjetas K40 Nvidia. Su almacenamiento es de 1.2 PB. Según pruebas, su potencia que alcanza es de 153.4 teraflops teóricos. (Laboratorio Nacional de Supercómputo del Sureste de México, 2018)

Universidad Nacional Autónoma de México – Miztli

El sistema HP Cluster Platform 3000SL "Miztli" es una supercomputadora con una capacidad de procesamiento de 118 TFlop/s (118 billones de operaciones aritméticas por segundo). Cuenta con 5,312 núcleos de procesamiento Intel E5-2670, 16 tarjetas NVIDIA m2090, una memoria RAM total de 15,000 Gbytes y un sistema de almacenamiento masivo de 750 Terabytes (Universidad Nacional Autónoma de México, 2016).

Universidad Nacional Autónoma de México – KanBalam

El sistema HP Cluster Platform 4000, "KanBalam" es una de las supercomputadoras paralelas más poderosas de América Latina con una capacidad de procesamiento de 7.113 Teraflops (7.113 billones de operaciones aritméticas por segundo). Cuenta con 1,368 procesadores (cores AMD Opteron de 2.6 GHz), una memoria RAM total de 3,000 Gbytes y un sistema de almacenamiento masivo de 160 Terabytes. (Universidad Nacional Autónoma de México, 2016)

Centro de Investigación en Matemáticas A.C. – El Insurgente

Cuenta con un nodo maestro AMD Quad Core Opteron 2350 HE con 8 núcleos, 16 Gb de memoria RAM, 2 TB de almacenamiento y 3 tarjetas de red Gigabit; 15 nodos de cómputo AMD Quad Core Opteron 2350 HE con 8 núcleos, 16 Gb de memoria RAM, 120 Gb de almacenamiento y 2 tarjetas de red Gigabit y 16 nodos más de cómputo INTEL Xeon CPU E5502 con 4 núcleos, 16 Gb de memoria RAM, 160 Gb de almacenamiento y 2 tarjetas de red Gigabit. El clúster "El Insurgente" se ha utilizado para la solución de problemas estadísticos, resistencia de materiales, algoritmos complejos y problemas que involucran un alto número de variables haciendo uso de la programación en paralelo (Centro de Investigación en Matemáticas, 2013).

Universidad Autónoma del Estado de Hidalgo

Cuenta con dos nodos maestro con sistema operativo Linux CentOS 5.8 con kernel 2.6.18-308.13.1.el5, además de las configuraciones de la conectividad de las tarjetas de red y de la comunicación segura con SSH, para la interconexión de los nodos se usa un Switch con capacidad de soportar Gigabit Ethernet (1000 Mbps). En nuestro primer prototipo de implementación se ha podido comprobar la disponibilidad de los datos mediante la configuración RAID10, debido a que, si un disco falla, existe el disco de reemplazo que almacena las copias de los datos. El sistema de archivos PVFS2 da soporte a un acceso a los datos de una manera rápida por la distribución de los mismos en todos los servidores de E/S y con el uso de Gigabit Ethernet permite alcanzar 1 Gbps en las transferencias (Hernández Palacios, Núñez Cárdenas, Hervert Hernández, & De la Cruz Bautista, 2012).

Marco Teórico

Un cúmulo, granja o clúster de computadoras, se puede definir como un sistema de procesamiento paralelo o distribuido. Consta de un conjunto de computadoras independientes, interconectadas entre sí, de tal manera que funcionan como un solo recurso computacional. A cada uno de los elementos del clúster se le conoce como nodo (Revista UNAM, 2003).

Los clústers se clasifican por la configuración del hardware en homogéneos si todos los nodos poseen la misma arquitectura física y ejecutan el mismo sistema operativo y, heterogéneos si son nodos con hardware diferente entre sí y pueden tener instalado diferentes sistemas operativos. Además, existen tres tipos de clúster basados en sus características y tareas a realizar:

Clúster de balanceo de cargas (Load Balancing)

Método de trabajo distribuido en el cual los procesos se encuentran en todos los nodos, permitiendo de esta manera que las cargas de trabajo estén balanceadas, de manera que ningún nodo sea el encargado de procesar toda la carga de trabajo, por el contrario, es distribuida de manera equitativa entre todos los nodos, como se muestra

en la figura 1. Con este balanceo de carga de trabajo se evita un posible fallo por sobrecarga en algún nodo y se minimizan los tiempos de espera de procesos.

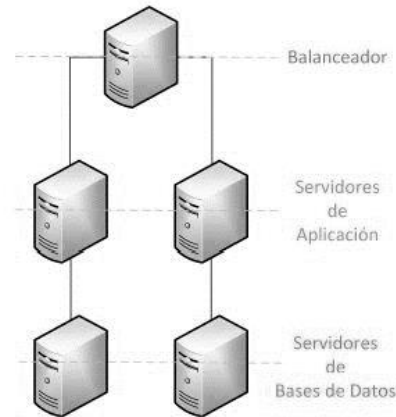


Figura 1. Clúster de balanceo de cargas (Jiménez & Medina, 2014)

Clúster de alta disponibilidad (High Availability)

Los clústers de alta disponibilidad están orientados a prestar el servicio de manera ininterrumpida. Por ejemplo, en los casos donde únicamente existe un servidor en producción, si existe alguna falla crítica, el servicio que ofrece el servidor queda interrumpido y los usuarios se ven afectados por negación de servicio, por el contrario, implementando este tipo de clúster si existe alguna falla crítica y un equipo se ve afectado, en el peor de los casos significaría, que el rendimiento del clúster se vería degradado, no así el servicio ofrecido por el nodo en cuestión, debido a que los demás nodos que están en operación suplen los servicios del nodo afectado de manera que el usuario no ve afectada su productividad a causa de la negación de servicio, como se muestra en la figura 2. De manera específica este tipo de clúster es utilizado para maximizar la disponibilidad del servicio con un rendimiento sostenido, concretamente para bases de datos que necesitan estar 99.99% disponibles para los usuarios, lo que significa 24/7 los 365 días del año.

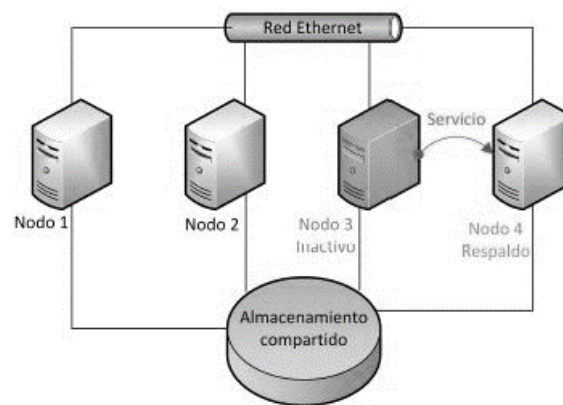


Figura 2. Clúster de alta disponibilidad (Jiménez & Medina, 2014)

Clúster de alto rendimiento (High Performance Computing)

Estos tipos de clústers están orientados a ejecutar de manera simultánea grandes tareas, las cuales por su naturaleza necesitan gran poder de procesamiento. A modo de desafío la tendencia de los clústers de alto rendimiento es obtener cada vez mayor poder de cómputo, es por esto que está muy enfocado al desarrollo de supercomputadoras, así como el desarrollo de algoritmos de cómputo paralelo. Para lograr lo anterior como se muestra en la figura 3, se utiliza agrupamiento de equipos interconectados entre sí a través de redes de datos de alta velocidad logrando de esta manera un rendimiento muy superior al alcanzado por cualquier computadora personal y con un costo accesible por

instancias de investigación donde requieren cómputo distribuido. Específicamente este tipo de clúster es utilizado para realizar tareas demandantes como por ejemplo la interpretación (renderizado) de imágenes, cálculos matemáticos, simulación y predicción de tiempo.

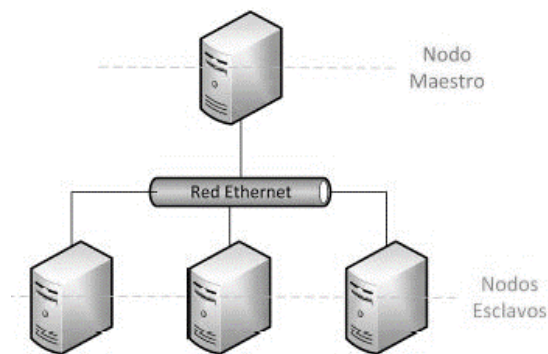


Figura 3. Clúster de alto rendimiento (Jiménez & Medina, 2014)

Metodología

Diseño del clúster de alto rendimiento

Una vez concluida la etapa en la cual se realizó un estudio exhaustivo de los diferentes métodos de clustering y analizando la tarea principal para la cual será implementado el clúster, que es la de realizar criptoanálisis, lo cual implica gran demanda de cómputo, se determinó la implementación de un clúster de alto rendimiento homogéneo debido a que los equipos disponibles tienen las mismas características físicas.

Como paso siguiente fue la gestión y acceso al equipo que se tendrá a disposición para la realización del proyecto, así como el software necesario dentro del cual se puede citar: sistema operativo, librerías, programas y utilerías que son ejecutadas para la correcta implementación del clúster. No se debe olvidar que un clúster es cómputo distribuido por medio de una red interna, para la cual se utilizará un switch y se pondrá en operación una red con topología de estrella como la que se muestra en la figura 4.

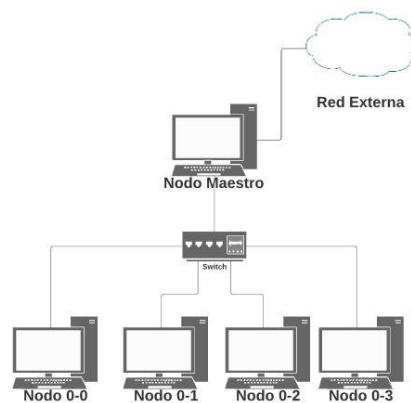


Figura 4. Topología de red utilizada para la comunicación entre los nodos (Fuente propia)

Descripción del equipo

- 1 Nodo Maestro. Computadora Acer AXC-605. Procesador Intel I3 de 4 núcleos, disco duro 1TB, memoria RAM 6GB, 2 tarjetas de red 1 Gbps.
- 16 Nodos esclavos. Computadora Acer AXC-605. Procesador Intel I3 de 4 núcleos, disco duro 1TB, memoria RAM 6GB, tarjeta de red 1 Gbps.
- 1 switch Cisco SG2600-26 Gigabit

Implementación en entorno de Laboratorio

Instalación del sistema operativo

Se evaluaron varias distribuciones libres del sistema operativo Linux, y se seleccionó la distribución de Linux llamada Rocks en su versión 7, la cual se puede descargar desde la página oficial. Primero se instaló el nodo maestro seleccionando para ellos los diferentes *rolls* (utilerías) necesarios para la ejecución de herramientas de monitoreo y ejecución del clúster. Posteriormente, una vez estando en ejecución el nodo maestro, se procedió a instalar cada uno de los 16 nodos esclavo por medio de la red, tomando los archivos de instalación procedentes del nodo maestro.

Configuración de la Red Interna

Para que el clúster pueda tener la comunicación óptima necesaria se tuvo que configurar la red interna, por medio de la cual se comunicarán los nodos esclavos desde su propia instalación con el nodo maestro, posteriormente para realizar el cómputo distribuido la red cobra vital importancia para la realización del procesamiento en paralelo de cada una de las tareas a realizar. Se utilizaron tarjetas de red con tazas de transferencia de 1 Gb en cada uno de los nodos, un switch Gigabit y cable UTP categoría 6 que soporta esas tazas de transferencia. Se utilizaron direcciones IP privadas. (10.1.1.1)

Herramienta de administración del clúster (Ganglia)

Ganglia es un proyecto de código abierto con licencia BSD (permite el uso del código fuente en software no libre) que surgió del Proyecto Berkeley Millennium de la Universidad de California. Es un sistema de monitoreo distribuido escalable para sistemas de cómputo de alto rendimiento. Aprovecha las tecnologías más utilizadas, como XML para la representación de datos, XDR para el transporte de datos compacto y portátil, y RRDtool para el almacenamiento y la visualización de datos. Utiliza estructuras de datos y algoritmos cuidadosamente diseñados para lograr gastos generales muy bajos por nodo y alta concurrencia (GANGLIA, 2016).

Pruebas de funcionamiento

Se realizaron pruebas en una pc y en el clúster para determinar la capacidad de procesamiento obtenido con esta arquitectura, por medio de la ejecución de diferentes aplicaciones que por su naturaleza requieren alta capacidad computacional. Para dichas pruebas realizaron dos proyectos y a continuación se muestran sus tablas comparativas:

1. Criptoanálisis de archivo SAM de Windows 8

Para realizar proyecto se utilizó la herramienta John the ripper, por medio del cual se criptoanalizó este importante archivo de Windows dentro del cual el sistema guarda de manera encriptada las contraseñas de los usuarios. Se criptoanalizaron contraseñas con una longitud de 7 caracteres alfanuméricos; esto quiere decir, 26 letras mayúsculas, 26 letras minúsculas, 10 dígitos y 20 caracteres especiales lo que da un total de 24.9 billones de posibles combinaciones.

Nodos	Núcleos	Tiempo de Procesamiento (Horas)
1 (PC)	4	107:53:16
17 (Clúster)	68	12:27:27

Tabla 1. Resultados de Criptoanálisis de archivos SAM (Fuente propia)

2. Obtención de números primos.

Por medio de un programa creado en lenguaje C y a través de la implementación de las librerías MPI y open MPI propias de la arquitectura de clúster de computadoras se ejecutó este programa en los mismo equipos descritos anteriormente para obtener todos los números primos encontrados dentro de un intervalo dado. En este caso el intervalo dado fue de 10 000 000.

Nodos	Núcleos	Tiempo de Procesamiento (Segs.)
1 (PC)	4	5.078
17 (Clúster)	68	0.0013

Tabla 2. Resultados de ejecución de programa para la obtención de números primos (Fuente propia)

Conclusión

Al iniciar este trabajo se investigó se planteó la problemática de la necesidad de equipo de cómputo con gran capacidad de procesamiento dentro de ambientes de investigación como lo son las universidades. A medida que fue avanzando este trabajo, se demostró que la arquitectura de clúster tipo Beowulf es óptima dentro de este tipo de ambientes debido a que los recursos económicos son limitados, y es ahí donde se aprovecha al máximo los equipos con los actualmente cuenta la institución, así como herramientas de software libre, con lo cual se obtiene un ahorro generalizado en los costos de instalación, administración, mantenimiento y monitoreo del clúster en comparación a equipos de supercomputación con capacidades de procesamiento similares. Se realizaron pruebas con diferente cantidad de nodos y para poder obtener diferencias notables en capacidad de procesamiento se pudo lograr trabajar con 17 nodos con los cual hubo un rendimiento de hasta un 89.2 % del clúster con respecto a una computadora personal con las mismas características que un nodo, ejecutando un programa para criptoanalizar contraseñas. Además de lo anterior un clúster de computadoras puede ser utilizado para procesar grandes volúmenes de datos, lo cual puede ser muy útil en campos como la minería de datos, big data, programación paralela, reconocimiento de patrones, entre muchos otros, además de una amplia gama de aplicaciones dentro del área científica, matemática y de ingeniería.

Referencias

- Laboratorio Nacional de Supercómputo del Sureste de México. (01 de 01 de 2018). *LNS*. Recuperado el 02 de 2019, de LNS: <http://lms.org.mx/?q=infraestructura%20>
- Centro de Investigación en Matemáticas. (01 de 02 de 2013). *CIMAT*. Recuperado el 04 de 09 de 2018, de CIMAT: <http://personal.cimat.mx:8181/~chuche/hpc/>
- Centro de Investigación y Estudios Avanzados. (2019). *CINVESTAV*. Recuperado el 02 de 2019, de CINVESTAV: <https://clusterhibrido.cinvestav.mx/>
- Cybermetrics Lab. (17 de 02 de 2018). *Ranking Web of World Research Centers*. Recuperado el 15 de 08 de 2018, de Ranking Web of World Research Centers: <http://research.webometrics.info>
- Acosta Díaz, R., García Ruíz, M. Á., Banda Montes, C., Barajas Alcalá, O., Ramírez Alcaraz, J. M., Damián Reyes, P., & Bustos Mendoza, C. (06 de 2009). *ResearchGate*. Obtenido de ResearchGate: <https://www.researchgate.net/publication/255622841>
- GANGLIA. (2016). *Ganglia Monitoring System*. Obtenido de Ganglia Monitoring System: <http://ganglia.sourceforge.net/>
- Hernández Palacios, R., Núñez Cárdenas, F., Hervert Hernández, J., & De la Cruz Bautista, M. (2012). *Universidad Autónoma del Estado de Hidalgo*. Recuperado el 07 de 09 de 2018, de Universidad Autónoma del Estado de Hidalgo: <https://www.uaeh.edu.mx/scige/boletin/huejutla/n1/a2.html>
- Jiménez, D., & Medina, A. (2014). Cluster de alto Rendimiento. *Journal Innovación y Tecnología*(14), 16-27. Recuperado el 10 de 09 de 2018, de http://www.revistasbolivianas.org.bo/scielo.php?pid=S1234-12342014000100004&script=sci_arttext
- Revista UNAM. (2003). Conceptos básicos del clustering. *Revista UNAM*.
- Universidad Nacional Autónoma de México. (2015). *Instituto de Geofísica*. Recuperado el 04 de 09 de 2018, de Instituto de Geofísica: <http://www.geofisica.unam.mx/reclat/cluster.php>
- Universidad Nacional Autónoma de México. (01 de 01 de 2016). *UNAM*. Recuperado el 02 de 09 de 2018, de UNAM: <http://www.super.unam.mx/index.php/content-layouts>
- Universidad Nacional Autónoma de México. (01 de 01 de 2016). *UNAM*. Recuperado el 01 de 2019, de UNAM: <http://www.super.unam.mx/index.php/component/content/article/35-supercomputadoras/120-kan-balam>