

Cluster Bicentenario, hacia una infraestructura de Supercomputador

A. Sarghel – I. Hansen – A. Fraga – P. Subia – A. Lamenza – A. Busso – P. Frias – S. A. Rodriguez

UTN FRT

Universidad Tecnológica Nacional, Facultad Regional Tucumán

Resumen: En este trabajo se discute el diseño y aplicación del Cluster Bicentenario, un cluster computacional de alta prestación y disponibilidad, con mas de 100 CPU, basado en Software Libre. Sus características (diseño, prestaciones, trabajo colaborativo e impacto) hacen de este grupo único. Para que este diseño se transforme, en realidad, debemos reciclar recursos en desuso, colaborar con varias ramas de la ciencia, mejorar las actuales prestaciones computacionales para lograr un gran impacto en la sociedad. El conjunto diverso de necesidades de los usuarios, mas la realidad socio económica en nuestro ámbito nos ha llevado a un diseño que difiere significativamente de la mayoría de los grupos que hemos visto en otros lugares. En este trabajo se puede observar a todos los integrantes, que colaboran Ad honorem con una metodología asociativa, compuesta de subgrupos interrelacionados por el intercambio de experiencias logradas en la vida real. En abril del 2009 se autó una visión de llegar a los 200 años de la Revolución Argentina (25 de mayo, 2010), con mas de 100 CPU, dejando al descubierto el nombre Cluster Bicentenario, hacia una infraestructura de Supercomputador.

Palabras claves: cluster, bicentenario, supercomputador, trabajo colaborativo, infraestructura, software libre, reciclado de hardware.

Abstract: In this paper we discuss the design and implementation of the Bicentennial Cluster, a high performance computational cluster and availability, a 100+ CPU, based on Open Source. Its characteristics (design, performance, collaborative work and impact) makes this group unique. For this design becomes a reality we must recycle unused resources, collaborate with several branches of science, improve current performance computing to achieve a great impact on society. The diverse set of user needs, but the socio-economic reality in our environment has led to a design that differs significantly from most of the groups we have seen elsewhere. In this paper we can see all the members, who work honorary, with an associated methodology, composed of subgroups interconnected

with the exchange of experiences gained in real life. In April 2009 a vision pattern reach 200 years of the Argentina Revolution (May 25, 2010), with more than 100 CPU, revealing the name Cluster Bicentennial, to a supercomputer infrastructure.

Keywords: cluster, bicentennial, supercomputer, collaborative work, infrastructure, open source, hardware recycling.

Introducción

Para la mayoría, en la última década la orientación principal en computación de alto rendimiento, ha sido en la dirección de grupos de productos básicos, comúnmente conocidos como clusters Beowulf [1]. Estos clusters agrupan hardware del mercado hogareño, para igualar o superar el rendimiento de los superordenadores tradicionales. En muchos casos cuesta como diez veces menos [2]. No todas las aplicaciones son adecuadas para los clusters, pero una importante fracción pueden adaptarse a ellos. Observando el ecosistema que nos rodea, analizamos las necesidades computacionales no explotadas, el escaso compromiso de la comunidad hacia el medio ambiente, la casi nula presencia de software libre en la región, la falta de consciencia para compartir documentación científica ilimitadamente, mediante el uso de licencias libres, combinando con la ausencia de recursos culturales y económicos. Esto nos conduce hacia una problemática de grave índole, en donde el impacto se refleja a toda la comunidad. En abril del 2009 nos encontramos en una encrucijada de ideas, donde los pensamientos no estaban plasmados de una forma clara. Comenzamos con la idea de armar un cluster compuesto por PCs de escritorio, obtenidas mediante donaciones. El fin se tornaba una investigación técnica, sin un objetivo completo. Al ver como avanzaba el desarrollo del estado del arte, nos encontramos empapados de nuevos conocimientos que no satisfacían nuestra vocación, considerábamos que el trabajo no retribuía de manera tangible a la comunidad. Esto nos llevó a reformular el objetivo del grupo, fijando nuevas pautas, donde nos comprometemos a reciclar hardware adquirido por donaciones, utilizar

sólo software libre, compartir la documentación con licencias libres, contar con un director para el lineamiento general del proyecto, áreas de software, hardware, documentación y calidad. Los miembros del proyecto participan del trabajo Ad honorem, convergiendo en un impacto de carácter ecológico, fraterno, científico y social, donde el límite es el tiempo.

Diseño

Seleccionamos un esquema de AD (alta disponibilidad) por la gran probabilidad de fallo en algún nodo, debido a que contamos con hardware reciclado en su mayoría. Al mismo tiempo nos vemos limitados al no poseer todo el hardware necesario para este tipo de diseño, por este motivo le damos prioridad a la AD de los CF (cortafuegos) y los NM (nodos maestros) Fig. 1. Colocamos 2 CF y dos swtichs, para lograr la AD en el paso de la información hacia los NM. A medida que logremos obtener mas donaciones, buscaremos mejorar la AD de los nodos, aumentando el número de swtichs Fig. 2.

Prestaciones

Para lograr una AP (alta prestación) con nodos de escasa capacidad de cálculo, es necesario aumentar la cantidad de nodos, utilizar la menor cantidad de recursos en cada uno y optimizar la red de datos. A medida que aumenta la capacidad de proceso de datos, incrementamos los nodos sin alterar el diseño, una vez alcanzados los 100 nodos, se agrega un nodo maestro. Para llegar a una buena prestación mediante un modelo flexible contamos con, 2 CF , 2 NM y 5 Nodos. Se utiliza como SO (sistema operativo) FreeBSD 7.2

RELEASE [3], excepto en los MF donde se implementa OpenBSD 4.5 [4]. Como administrador de recursos contamos con Sun Grid Engine [5], la interconexión de equipos se realiza con 3 swtichs a través una red Ethernet 10/100Mb. Se comparten los recursos utilizando NFS [6] (345,8 GB SCSI) Fig.4. La virtualización se realiza con Sun Virtualbox [7]. Los backups se realizan con un DLT de 20/40 GB y un DAT de 4/8GB. La salida hacia internet se realiza por una Red de Altas Prestaciones.

Detalle de los núcleos utilizados en Fig. 2.

Tabla 1: tabla de detalles

Tipos de CPU	Equipos	CPUs
Pentium II 266 MHz	1	1
Celeron 600 MHz	1	1
Pentium III 600 MHz	3	3
Xeon 300 MHz	1	1
Xeon 400 Mhz	1	4
Xeon 500 Mhz	1	2
Xeon 700 Mhz	1	2
Total	9	14

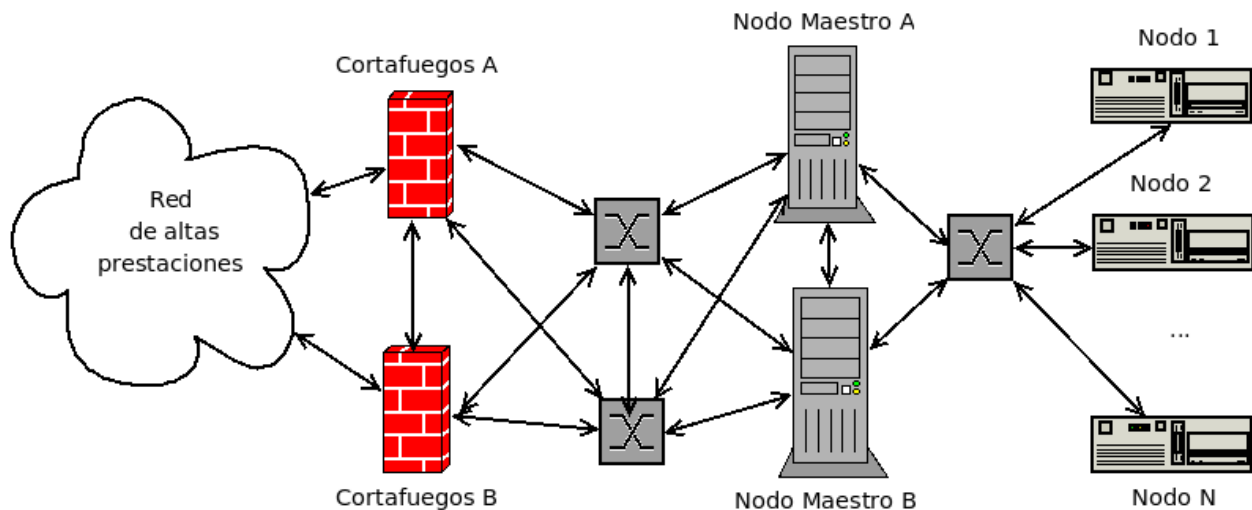


Fig.1 Diagrama de Estructura

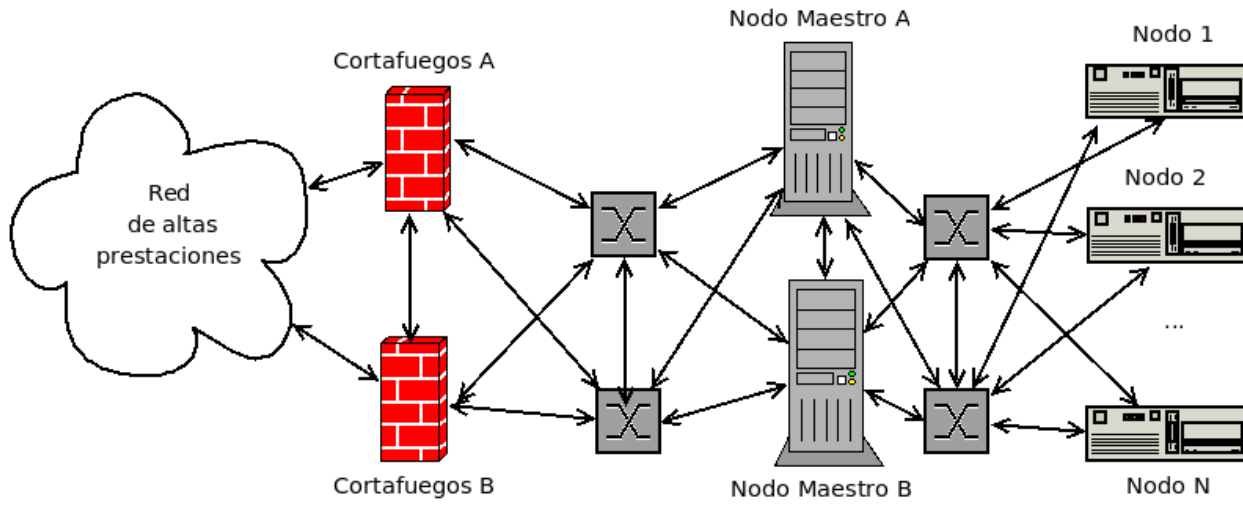


Fig.2 Diagrama a futuro

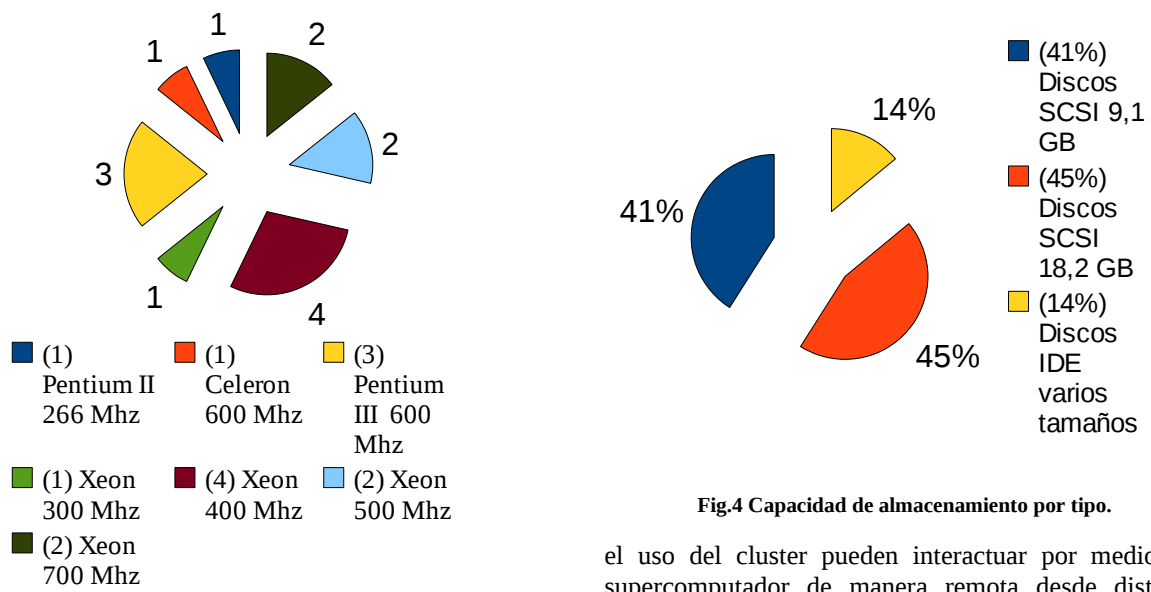


Fig.3 Cantidad de procesadores por tipo

Fig.4 Capacidad de almacenamiento por tipo.

Trabajo Colaborativo

El grupo del trabajo cuenta con un director general para el lineamiento del proyecto, supervisores de las áreas de software, hardware, documentación, gestión y calidad. En el desarrollo proponemos un trabajo colaborativo de macro niveles divididos en cuatro grandes grupos, usuarios finales, conglomerado de computadores reciclados, miembros del proyecto y la comunidad externa. Los usuarios finales que son beneficiados por

el uso del cluster pueden interactuar por medio del supercomputador de manera remota desde distintos puntos de acceso pudiendo trabajar en equipo y distintas tareas. Desde el punto de vista técnico el conglomerado de computadores trabajan en forma conjunta para alcanzar un alto grado de rendimiento cooperándose unos a otros para cubrir las necesidades de los usuarios. El equipo de desarrollo actúa en forma sinérgica donde en la suma de las partes obtenemos el todo para cualquier actividad a llevar a cabo tanto de investigación como de desarrollo. Completa el macro nivel la comunidad investigadora científica como beneficiada para simplificar a futuro trabajos similares obteniendo de nuestro trabajo un camino marcado por la experiencia y éxitos logrados. Formamos un ambiente donde cada nivel es completamente colaborativo en sus partes y a su vez cooperan entre ellos formando un conjunto de estrategias tendientes a

maximizar los resultados y minimizar la pérdida de tiempo en formación a beneficios comunes.



Fig.5 Estructura colaborativa de macro-nivel.

Trabajos relacionados

La iniciativa de computo avanzado del IFUNAM (Instituto de Física de la Universidad de México) tiene su origen en un proyecto de apoyo a la investigación en el año 1999 y que fue motivado por la tradición de vanguardia para explorar e implementar nuevas tecnologías de investigación científica, que son generadas en el mundo entero. [8]

Desde año 2002, la Dirección General de Servicios de Cómputo Académico de la UNAM adquirió un cluster Beowulf por la necesidad de cubrir a sus usuarios investigadores grandes cálculos numéricos para brindar una infraestructura completa y compleja [9].

Diferencia entre Cluster Bicentenario y los clusters de la actualidad

Existen grandes desigualdades entre los clusters actuales y el proyecto que aspiramos. La principal diferencia es emplear como materia prima hardware reciclado, por lo tanto encontraremos en el conglomerado de computadores una gran diversidad referida a marcas, potencias, velocidades, generaciones, etc. formando un cluster muy heterogéneo y por sobre todo, económico y ecológico.

Impacto

Gracias a los variados campos de aplicación los resultados de este proyecto impactan tanto en sectores

sociales como niveles informáticos puedan imaginarse, haciendo casi imposible medir cual de todos ellos es mas provechoso.

La construcción de un supercomputador reciclado constituye una valiosa herramienta que inmersa en el marco de la actividad científica de investigación, provee velocidad, multiplicidad y certeza en procesamiento, sumando el trabajo en simultaneo con una cantidad ilimitada de personas desde cualquier punto de acceso. Utilizada para actividades que requieran millones de cálculos y operaciones matemáticas, procesos simultáneos de grandes volúmenes de datos, y otras áreas como medicina, química, ciencias nucleares, geografía, ciencias de la atmósfera, ciencias aéreo espaciales, etc. Esta herramienta logra cubrir las necesidades que los investigadores a nivel cómputo y rendimiento aspiran en cualquier tipo de proyecto. Por otro lado, la humanidad inconscientemente destruye el planeta gracias a la falta de cultura ecológica, los resultados esperados con el reciclado de hardware y acondicionamiento de equipos deteriorados demostrarán que estos aún conservan un alto valor agregado trabajando en conjunto.

Esperamos con el funcionamiento del cluster y sus beneficios que en un futuro no muy lejano la clasificación de los fabricantes de hardware sea de acuerdo al respeto hacia el medio ambiente.

Trabajo futuro

Como siguiente paso, isolaremos el cluster para mayor provecho de su capacidad, obteniendo la posibilidad de brindar prioridades a los servicios. Automatizar los recursos virtualizados, mejorar la conectividad, aumentar la capacidad de almacenamiento. Sustentar requerimientos con mayor cantidad de nodos y mejoras en el diseño. Divulgar el trabajo colaborativo que proponemos, aclarando que no solamente está referido al trabajo de un grupo de nodos computacionales, sino también a sumar personas que aporten conocimientos y dedicación, para formar parte de una comunidad investigadora particular. Buscar nuevas formas de construir un supercomputador llegando al top 500 de cluster mundial.

Conclusión

Brindar una infraestructura informática de alta capacidad de proceso y acceso, mediante la unión de conocimientos, objetivos y filosofías de trabajo, permitiendo a la humanidad, cuidar el medio ambiente y difundir conocimientos.

Referencias

- [1] Beowulf.org is a collection of resources for the expanding universe of users and designers of Beowulf class cluster computers. These enterprise systems are built on commodity hardware deploying Linux OS and open source software. <http://www.beowulf.org/>
- [2] Brooks Davis, Michael AuYeung, Gary Green, Craig Lee. Building a High-performance Computing Cluster Using FreeBSD.
17 julio 2006
<http://people.freebsd.org/~brooks/papers/>
20 Marzo 2008
<http://www.nycbug.org/index.php?NAV=Home:SUBM=10151>
<http://2009.asiabsdcon.org/papers/index.html>
- [3] FreeBSD es un avanzado sistema operativo derivado de BSD, la versión de UNIX desarrollada en la Universidad de California, Berkeley.
<http://www.freebsd.org/es/>
- [4] OpenBsd. El proyecto OpenBSD produce un sistema operativo libre tipo Unix, multiplataforma, basado en 4.4BSD. Su empeño se orientan principalmente a la portabilidad, cumplimiento de estadares, corrección, seguridad y criptografía integrada. <http://www.openbsd.org/es/>
- [5] Sun Grid Engine
The Grid Engine project provides enabling distributed resource management software for wide ranging requirements from compute farms to grid computing.
<http://gridengine.sunsource.net/>
- [6] NFS
The Sun Network Filesystem protocol provides transparent remote access to share files across the networks.
<http://www.ietf.org/rfc/rfc1094.txt>
- [7] Sun Virtualbox es un software de virtualización.
<http://www.virtualbox.org>
- [8] Instituto de Física, Universidad Nacional de México, Proyecto de Supercómputo Avanzado. Abril de 1999.
<http://cluster.fisica.unam.mx/>
- [9] Revista UNAM, Revista digital Universitaria. Maestros y Esclavos, una aplicación a los cúmulos de computadoras Volumen 4, Numero 2 - Junio de 2003 Hernández Cervantes , Santillán González y Caballero Cruz Reyna.
<http://www.revista.unam.mx/vol.4/num2/art3/art3.htm>

Nombre Completo del Autor

Alexis Sarghel – Iván Hansen – Alvaro Fraga – Pedro Subia – Agustín Lamenza – Agustín Busso – Pablo

Frias – Sebastian A. Rodriguez

Dirección del autor

E-mail: cluster-bicentenario@googlegroups.com

<http://groups.google.com.ar/group/cluster-bicentenario>

Alexis Sarghel: alexis@sarsys.com.ar

Iván Hansen: ivanjhansen@gmail.com

Alvaro Fraga: gioalvaro@gmail.com

Pedro Subia: subiapedro@gmail.com

Agustín Lamenza: agustinlamenza@gmail.com

Agustín Busso: bussoagustin@gmail.com

Pablo Frias: frias.pablo@gmail.com

Sebastian A. Rodriguez:

sebastianrodriguez@sebastianrodriguez.com.ar