

## **Desarrollo e Implementación de Modelos de Supercomputación**

Erika López

Directora del Centro de Tecnología de la Información y las Comunicaciones

Universidad Don Bosco, El Salvador

Erika.lopez@udb.edu.sv

**Resumen.** Este documento es la presentación del proyecto de “Desarrollo e Implementación de Modelos de Supercomputación” realizado durante el periodo de octubre de 2008 a septiembre de 2009, siendo el coordinador y ejecutor del proyecto, el Ing. Carlos Bran.

Este documento recopila las actividades realizadas durante su ejecución, presentando el problema a resolver y como se logro el cumplimiento de sus objetivos. Se presenta además, la metodología que siguió cada una de las actividades.

**Palabras Clave:** VoIP, Voz sobre IP, RAICES, Redes Avanzadas, Internet2, Telefonía.

### **Descripción del proyecto:**

El problema a resolver era el cómo lograr que nodos no homogéneos (Distintas arquitecturas, capacidades y marcas), puedan operar de forma distribuida para resolver problemas y sostener servicios, manteniendo de forma óptima las características de: Economía, Velocidad, Disponibilidad, Escalabilidad, y Transparencia.

El proyecto consistió en investigar modelos de agrupamiento de computadoras para desarrollar infraestructuras de procesamiento paralelo, que posibiliten la supercomputación de alto rendimiento y alta disponibilidad, basándose en tecnologías de código abierto existentes y adaptándolas para poder hacerlas fáciles de desplegar, administrar y escalar; así como la evaluación de aplicaciones detonantes que puedan ejecutarse sobre este tipo de infraestructura.

### **Objetivo general:**

Diseñar un modelo escalable y eficiente de reciclado de computadoras, para ser usadas en infraestructuras de procesamiento paralelo de alto rendimiento o de alta disponibilidad.

### **Objetivos específicos:**

1. Construir un clúster experimental de supercomputación que aproveche el poder de procesamiento del hardware de cualquier computadora de escritorio discontinuada, sobre plataformas de software de código abierto; así como, la capacidad de interconexión de redes de alta capacidad y que sirva de núcleo de pruebas para el establecimiento de otros nodos a nivel nacional.
2. Desarrollar procedimientos y secuencias de instalación amigables que faciliten el despliegue mantenimiento y crecimiento de clúster de supercomputación, en ambientes académicos y empresariales.

3. Establecer una batería de aplicación y/o servicios de uso académico y empresarial que puede implementarse sobre una plataforma de computación paralela; las que servirán de detonante para la solución de otros problemas científicos y tecnológicos que puedan aprovechar las ventajas de la computación paralela.

Tales objetivos implicaron una fuerte investigación de soluciones de núcleo, así como el trabajo con sistemas operativos nativos a Unix, ya que este tipo de soluciones implican modificaciones profundas de la forma como las estaciones distribuyen y migran sus recursos no solo a sus procesadores internos, sino a otros externos conectados a ellos, usando redes de datos de alta velocidad.

## **Metodología:**

### **1. Investigación de evaluación de tecnología**

Durante el primer bimestre de trabajo, el componente fuerte se enfocó en investigar de forma documental y experimental las soluciones de código abierto para el despliegue de infraestructuras distribuidas, con el propósito de entender mejor su funcionamiento y evaluar la facilidad de despliegue de las mismas, para elegir la que reúna las mejores características y tomarla como base de la solución que se pretende desarrollar, para luego ser adaptada y desplegada en la infraestructura experimental.

En esta fase por lo tanto se evaluaron 4 tecnologías de núcleo para el montaje de clúster de alto rendimiento, 2 tecnología de clúster de alta disponibilidad, y 2 tecnologías de grillas (GRID) de alto rendimiento, esta tecnología no estaba considerada en los objetivos originales del proyecto, pero dada su importancia actual y sus posibilidades futuras, se consideró importante incluirla en esta investigación.

### **2. Instalación física y lógica**

El principal enfoque para este período fué el de crear los instaladores para los diferentes clúster, lo que implicaba usar la documentación de evaluación de los núcleos desarrollada en la investigación y evaluación de tecnología y utilizar herramientas para empaquetarlo, sobre un instalador eficiente y amigable, por otro lado implicaba el montaje de la infraestructura física del clúster adquirida a partir del esquema desarrollado en el período previo.

Además, se realizó la evaluación de aplicaciones con el propósito de efectuar pruebas básicas de las arquitecturas de supercomputación.

### **3. Depuración, Pruebas y Desarrollo Final**

Durante este último período se desarrollaron las actividades finales para la culminación de los objetivos propuestos en el proyecto, lo cual cubrió principalmente actividades de depuración y desarrollo final de los instaladores de los diferentes modelos de supercomputación, así como el trabajo con las aplicaciones prototipo que se lanzarán sobre el ambiente distribuido.

### **4. Revisión de mecanismos de seguridad**

Este punto no fue considerado originalmente en el proyecto, pero dada la envergadura de la capacidad de procesamiento disponible, fué necesario desarrollar análisis de vulnerabilidades, con el propósito de crear una malla perimetral fortalecida de los modelos, para ello se aprovechó, la disponibilidad de los firewalls de los laboratorios de redes para efectuar pruebas y fortalecer la seguridad de los servidores de núcleo y nodos. Para ésto se usó un dispositivo de seguridad adaptativa ASA5510 el cual se instaló posterior a la finalización del proyecto ya que

se propuso al FIES una redistribución de los fondos para la adquisición del equipo, el cual fué aprobado hasta el mes de septiembre y debido a que estos equipos se tardaron en entregar se hizo uso de un mecanismo de protección virtual para garantizar la seguridad de los equipos y aplicaciones mientras se instalaba el servidor definitivo.

### **Instrumentos utilizados:**

1. Bitácora diaria de actividades.
2. Cuadro de indicadores de medición de las diferentes actividades realizadas según periodo, indicando los avances de los resultados logrados.
3. Análisis basados en los objetivos del proyecto indicando las Barreras que limitaron la ejecución del proyecto y claves de éxito que lo potenciaron.
4. Reportes de control en base a los objetivos indicando el porcentaje de avance de las actividades.
5. Realización de reportes gráficos.

### **Resultados:**

Durante el primer bimestre de trabajo, el componente fuerte se enfocó en investigar de forma documental y experimental las soluciones de código abierto para el despliegue de infraestructuras distribuidas, con el propósito de entender mejor su funcionamiento y evaluar la facilidad de despliegue de las mismas, para elegir la que reúna las mejores características y tomarla como base de la solución que se pretende desarrollar, para luego ser adaptada y desplegada en la infraestructura experimental.

Los resultados obtenidos mediante la ejecución de cada una de las actividades son:

1) Contar con una plataforma experimental de supercomputación operativa y lista para la solución de múltiples problemas de carácter académico o empresarial.



2) Instaladores en disco compacto con guía digital para facilidad de los usuarios de: Nodos monitores y celdas de un clúster de supercomputación; Herramientas de administración y monitoreo del clúster de supercomputación; Batería de aplicaciones y/o servicios que pueden ejecutarse en ambientes distribuidos.

3) Manuales de proceso de: Pre instalación e instalación del clúster; Mantenimiento del clúster; Operaciones del clúster; Proceso de actualización y escalación del clúster.

4) Aceptación de El Salvador en el proyecto EELA-2 con lo que se podrá expandir la plataforma de nodos a la Grilla a más 40,000 nodos.

### **Conclusión:**

- Obtención de experiencia en el trabajo con clúster.
- Contar con laboratorio de estaciones disponibles de la universidad, permitió hacer pruebas de forma previa, con ésto lograr hacer las evaluaciones experimentales para sustentar los conocimientos obtenidos.
- Conocer la gran diversidad de herramientas para procesamiento y cálculo numérico aplicado en ámbitos científicos y de ingeniería disponibles.
- Experiencia en re compilación de aplicaciones de código abierto.

## Referencias

1. Todas las referencias usadas fueron mediante sitios web para realizar la investigación documental y experimental.
2. ORGANISMO BEOWULF. [en línea] <<http://www.beowulf.org/>>
3. ORGANISMO THE GLOBUS ALLIANCE. [en línea] <<http://www.globus.org/>>
4. OPEN SOURCE CLUSTER APPLICATION RESOURCES. [en línea] <<http://svn.oscar.openclustergroup.org/trac/oscar>>
5. ORGANISMO CLUSTER AND MULTI-CLUSTER MANAGEMENT. [en línea] <<http://www.mosix.org/>>
6. GLITE - LIGHTWEIGHT MIDDLEWARE FOR GRID COMPUTING. [en línea] <<http://glite.cern.ch/>>
7. SISTEMA DE LENGUAJE DE INTERPRETACION DE ALTO NIVEL: OCTAVE. [en línea] <<http://www.gnu.org/software/octave/>>
8. SISTEMA LIBRE DE COMPUTACION NUMERICO: SCILAB. [en línea] <<http://www.scilab.org/>>
9. SISTEMA DE DISEÑO DE IMÁGENES MARCA AUTODESK: MAYA 3D. [en línea] <<http://usa.autodesk.com/adsk/servlet/pc/index?id=13577897&siteID=123112>>
10. SISTEMA DE CREACION DE IMÁGENES DE CONTENIDO EN 3D: BLENDER. [en línea] <<http://www.blender.org/>>
11. SIMULADOR DE RED GRAFICO: GNS3. [en línea] <<http://www.gns3.net/>>
12. E-SCIENCE GRID FACILITY FOR EUROPE AND LATIN AMERICA. [en línea] <[http://www.eu-eela.eu/index.php?option=com\\_content&task=view&id=155&Itemid=162](http://www.eu-eela.eu/index.php?option=com_content&task=view&id=155&Itemid=162)>

### Más información sobre el autor:

El Ing. Carlos Guillermo Bran, quien fue el Coordinador e Investigador del proyecto, trabajo durante 12 meses en la elaboración del mismo junto a 7 investigadores auxiliares. Actualmente el Ing. Bran se encuentra realizando sus estudios de Maestría en una universidad de España, con lo que posteriormente, se busca darle seguimiento a este proyecto.

Ing. Carlos Guillermo Bran  
cbran@udb.edu.sv

Depto. I+D+i. Centro de Tecnología de la Información y las Comunicaciones – CTIC