



UNIVERSIDAD DE LAS FUERZAS ARMADAS

COMPUTACIÓN PARALELA

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

---

## Instalación de un Clúster

---

*Estudiantes:*

Josué Merino, Angelo Sánchez, Justin Villarroel

*Docente:*

Ing. Carlos Andrés Pillajo Bolagay

## Objetivos

### Objetivos Generales

- Comprender el funcionamiento de un clúster usando ABCLinux.
- Instalar 4 máquinas virtuales con la distribución ABC\_GNULinux. Una máquina como master y tres máquinas parte del clúster.

### Objetivos Específicos

- Analizar las configuraciones de nodos y su integración en un clúster para optimizar el rendimiento computacional.
- Evaluar la efectividad del uso de herramientas de monitoreo y gestión de clústeres en la administración de recursos distribuidos.

## Resultados de Aprendizaje

1. **Análisis de ingeniería.** La capacidad de identificar, formular y resolver problemas de ingeniería en su especialidad; elegir y aplicar de forma adecuada métodos analíticos, de cálculo y experimentales ya establecidos; reconocer la importancia de las restricciones sociales, de salud y de seguridad, ambientales, económicas e industriales.
2. **Proyectos de ingeniería.** Capacidad para proyectar, diseñar y desarrollar productos complejos (piezas, componentes, productos acabados, etc.) procesos y sistemas de su especialidad, que cumplen con los requisitos establecidos, incluyendo tener conciencia de los aspectos sociales, de salud y seguridad, ambientales, económicos e industriales; así como seleccionar y aplicar métodos de proyectos apropiados.
3. **Aplicación práctica de la ingeniería.** Comprensión de las técnicas aplicables y métodos de análisis, proyecto e investigación y sus lineamientos en el ámbito de su especialidad.
4. **Comunicación y trabajo en equipo.** Capacidad para comunicar eficazmente información, ideas, problemas y soluciones en el ámbito de ingeniería y con la sociedad en general.

## Herramientas Utilizadas

- PC
- Máquina Virtual para simular Linux
- Ubuntu

## Marco Teórico

Un cúmulo, granja o cluster de computadoras, lo podemos 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 cluster se le conoce como nodo. Estos son aparatos o torres que pueden tener uno o varios procesadores, memoria RAM, interfaces de red, dispositivos de entrada y salida, y sistema operativo. Los nodos pueden estar contenidos e interconectados en un solo gabinete, o, como en muchos casos, acoplados a través de una red de área local (LAN (Local Area Network)). (1) Otro componente básico en un cluster es la interfaz de la red, la cual es responsable de transmitir y recibir los paquetes de datos, que viajan a través de la red entre los nodos. Finalmente el lograr que todos estos elementos funcionen como un solo sistema, es la meta a la que se quiere llegar para dar origen a un cluster. Comúnmente, en los clusters existe una máquina (con monitor, teclado, ratón, etcétera) que funciona como nodo-maestro y se encarga de administrar, controlar y monitorear todas las aplicaciones y recursos del sistema, en tanto que el resto de los nodos está dedicado al procesamiento de datos o a ejecutar operaciones aritméticas. Se les conoce como nodos-esclavos.

Las aplicaciones que instale en un clúster deben poder ejecutarse en cualquier servidor de aplicaciones que sea miembro de ese clúster. Dado que un grupo de nodos forma los límites de un clúster, todos los miembros de un clúster deben ser miembros del mismo grupo de nodos. Por consiguiente, para que la aplicación desplegada se ejecute correctamente, todos los miembros de un clúster deben encontrarse en nodos que cumplan los requisitos de dicha aplicación. (2)

En una célula que tiene muchas configuraciones de servidor distintas, puede ser difícil determinar qué nodos pueden albergar la aplicación. Un grupo de nodos puede utilizarse para definir grupos de nodos que tienen suficiente en común como para albergar miembros de un clúster determinado. Todos los miembros de un clúster deben estar en el mismo grupo de nodos. (2)

Todos los nodos son miembros de al menos un grupo de nodos. Al crear un clúster, el primer servidor de aplicaciones que se añade a dicho clúster define el grupo de nodos en el que deben residir los demás miembros de clúster. Todos los otros miembros de clúster

que añade al clúster sólo pueden estar en nodos que sean miembros de este mismo grupo de nodos. Cuando crea un nuevo miembro de clúster en la consola administrativa, sólo puede crear el servidor de aplicaciones en un nodo que sea miembro del grupo de nodos para ese clúster. (2)

Los nodos pueden ser miembros de varios grupos de nodos. Si el primer miembro que añade a un clúster tiene definidos varios grupos de nodos, el sistema selecciona automáticamente el grupo de nodos que enlaza con el clúster. Puede cambiar el grupo de nodos modificando los valores del clúster. Utilice la página Valores de clúster de servidores para cambiar el grupo de nodos. (2)

## Desarrollo

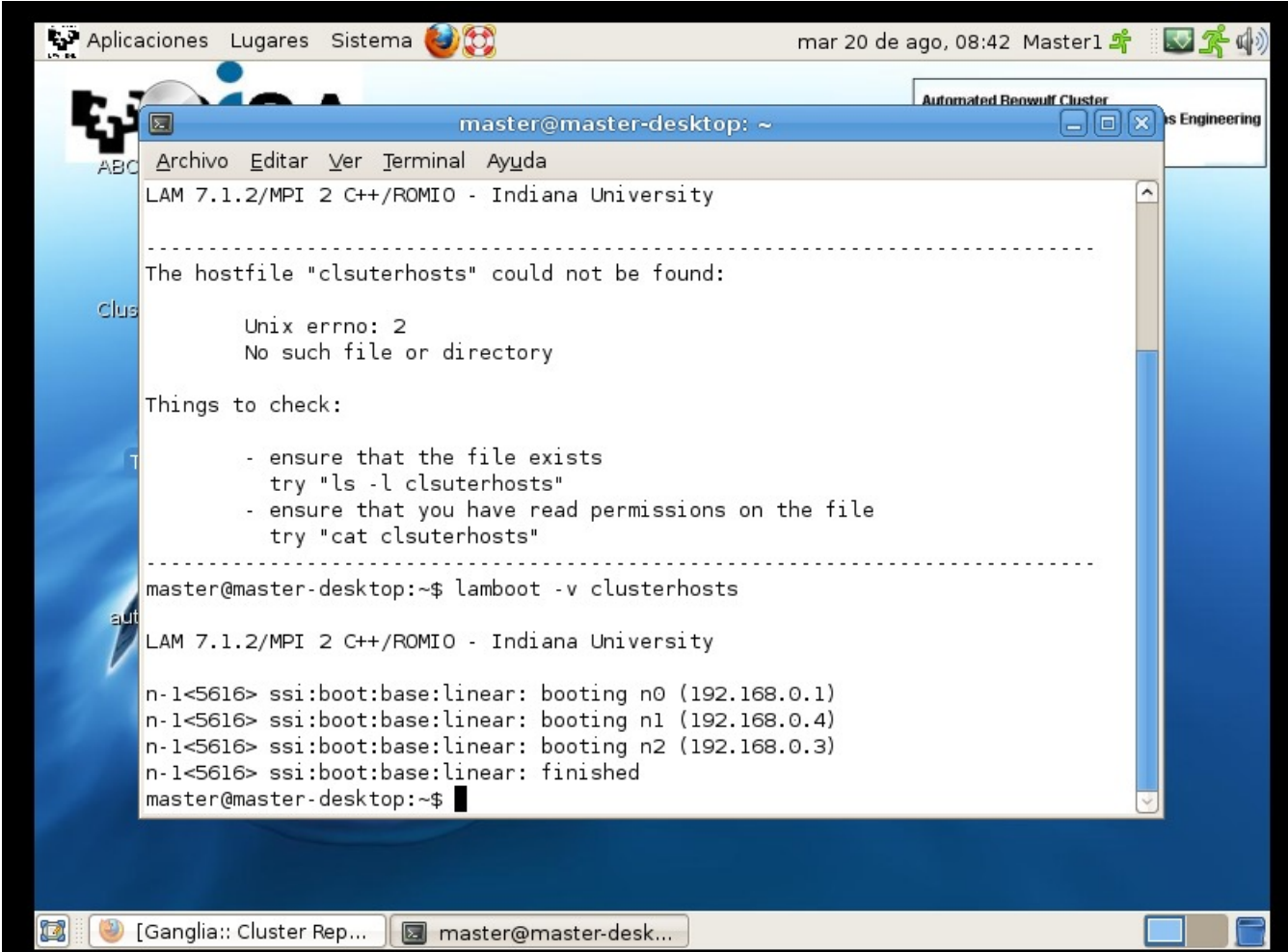
Se configuran las virtual networks de VMware, en donde se tiene una máquina máster y otros dos equipos que funcionarán como nodos.

```
* Already running. [ OK ]
* Starting Common Unix Printing System: cupsd [ OK ]
Starting Ganglia Monitor Daemon: gmond.
Starting Ganglia Monitor Meta-Daemon: gmetad.
* Starting NFS common utilities [ OK ]
* Exporting directories for NFS kernel daemon... [ OK ]
* Starting NFS kernel daemon [ OK ]
* Starting powernowd...
* CPU frequency scaling not supported... [ OK ]
Not starting slurm-llnl
slurm.conf was not found in /etc/slurm-llnl
Please follow the instructions in /usr/share/doc/slurm-llnl/README.Debian
* Starting Hardware abstraction layer hald [ OK ]
* Starting System Tools Backends system-tools-backends [ OK ]
landscape-client is not configured, please run landscape-config.
* Starting anac(h)ronistic cron anacron [ OK ]
* Starting deferred execution scheduler atd [ OK ]
* Starting periodic command scheduler crond [ OK ]

Automated Beowulf Cluster node tty1
node login: master
Password:
Last login: Tue Aug 20 13:39:38 CEST 2024 on tty1
Linux node 2.6.27-17-generic #1 SMP Fri Mar 12 03:09:00 UTC 2010 i686
master@node:~$ _
```

Figura 1: Instalación Nodo

Se evidencia que la versión del software de administración del clúster se encuentra instalado, escribiendo en la terminal: `lambboot -v`



```
master@master-desktop: ~  
Archivo Editar Ver Terminal Ayuda  
LAM 7.1.2/MPI 2 C++/ROMIO - Indiana University  
-----  
The hostfile "clsuterhosts" could not be found:  
  
    Unix errno: 2  
    No such file or directory  
  
Things to check:  
  
    - ensure that the file exists  
      try "ls -l clsuterhosts"  
    - ensure that you have read permissions on the file  
      try "cat clsuterhosts"  
-----  
master@master-desktop:~$ lamboot -v clusterhosts  
LAM 7.1.2/MPI 2 C++/ROMIO - Indiana University  
  
n-1<5616> ssi:boot:base:linear: booting n0 (192.168.0.1)  
n-1<5616> ssi:boot:base:linear: booting n1 (192.168.0.4)  
n-1<5616> ssi:boot:base:linear: booting n2 (192.168.0.3)  
n-1<5616> ssi:boot:base:linear: finished  
master@master-desktop:~$
```

Figura 2: lamboot -v

A través del programa Cluster Monitor, se obtienen las siguientes gráficas y resultados:

## Instalación de un Clúster

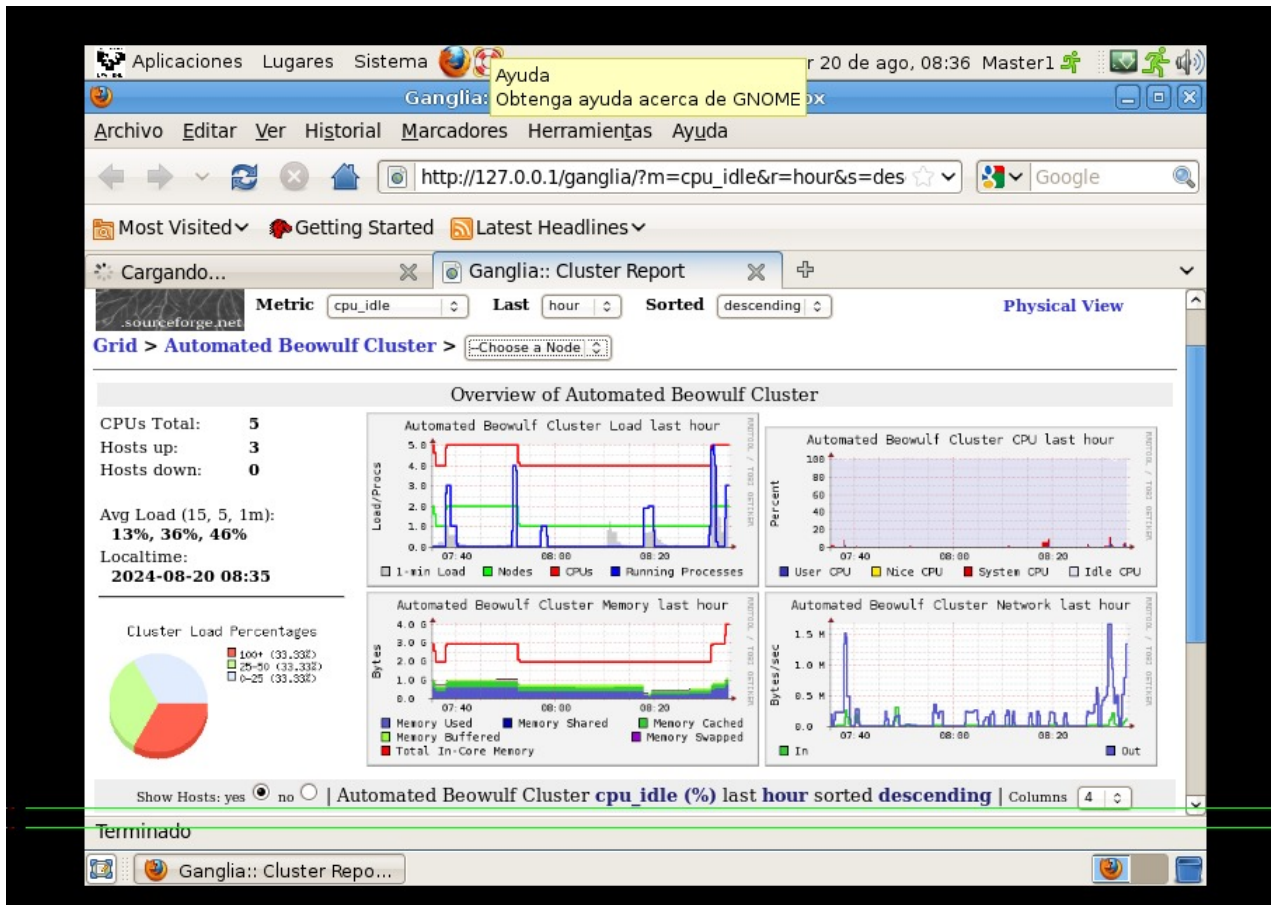


Figura 3: Cluster Master

A través del Host Report se muestran los siguientes resultados:

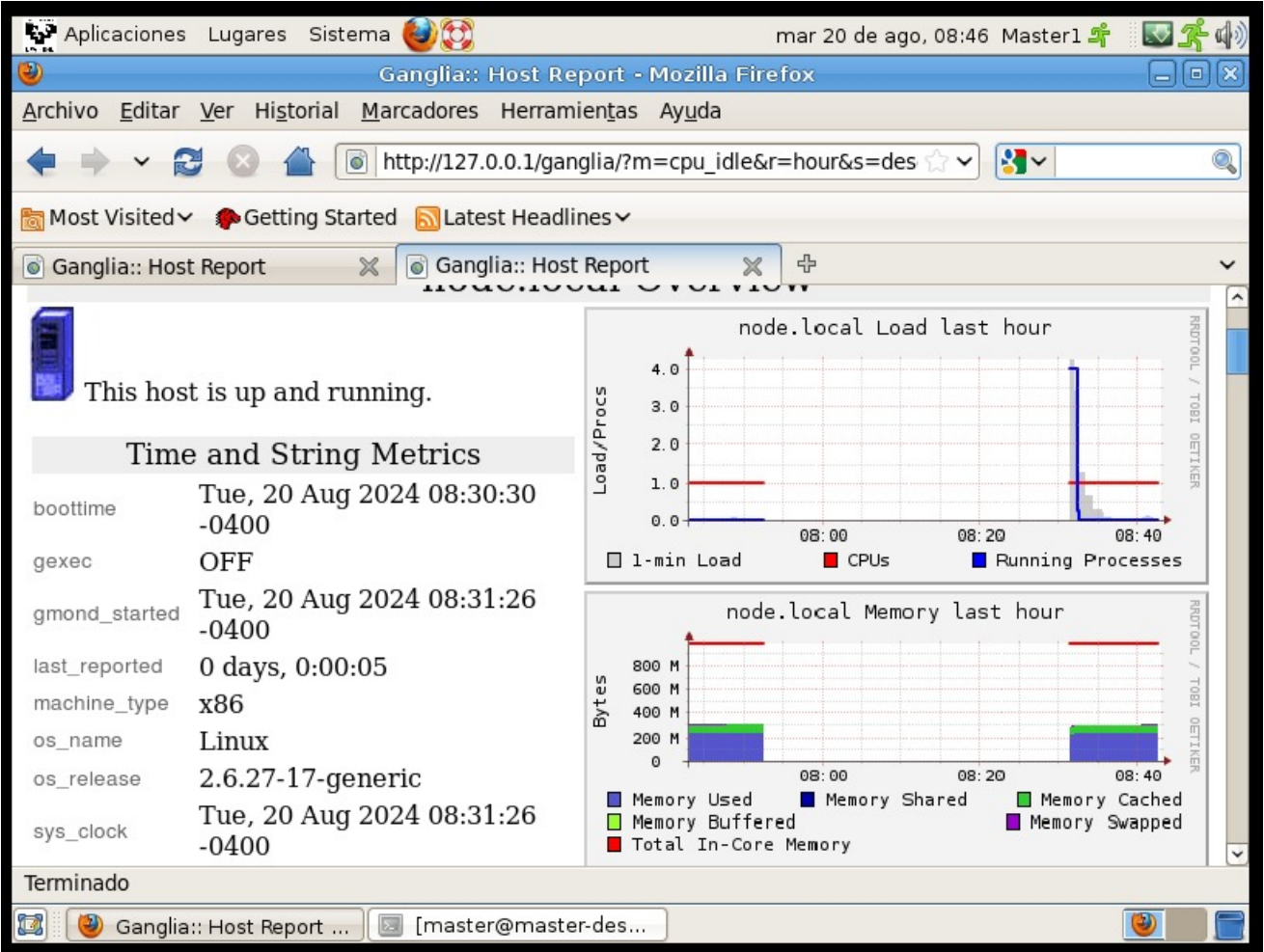


Figura 4: Métricas Nodo

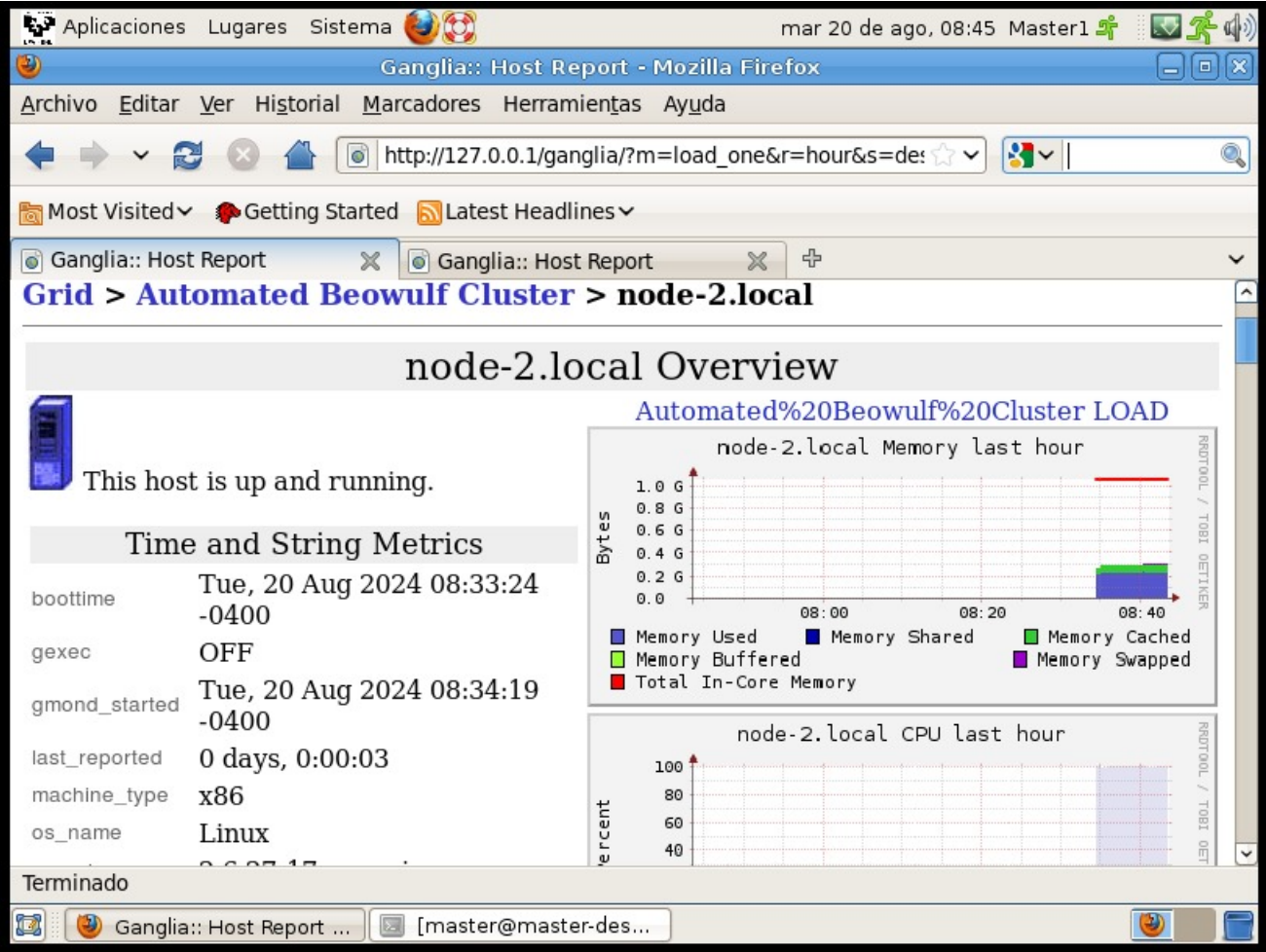


Figura 5: Nodo 2



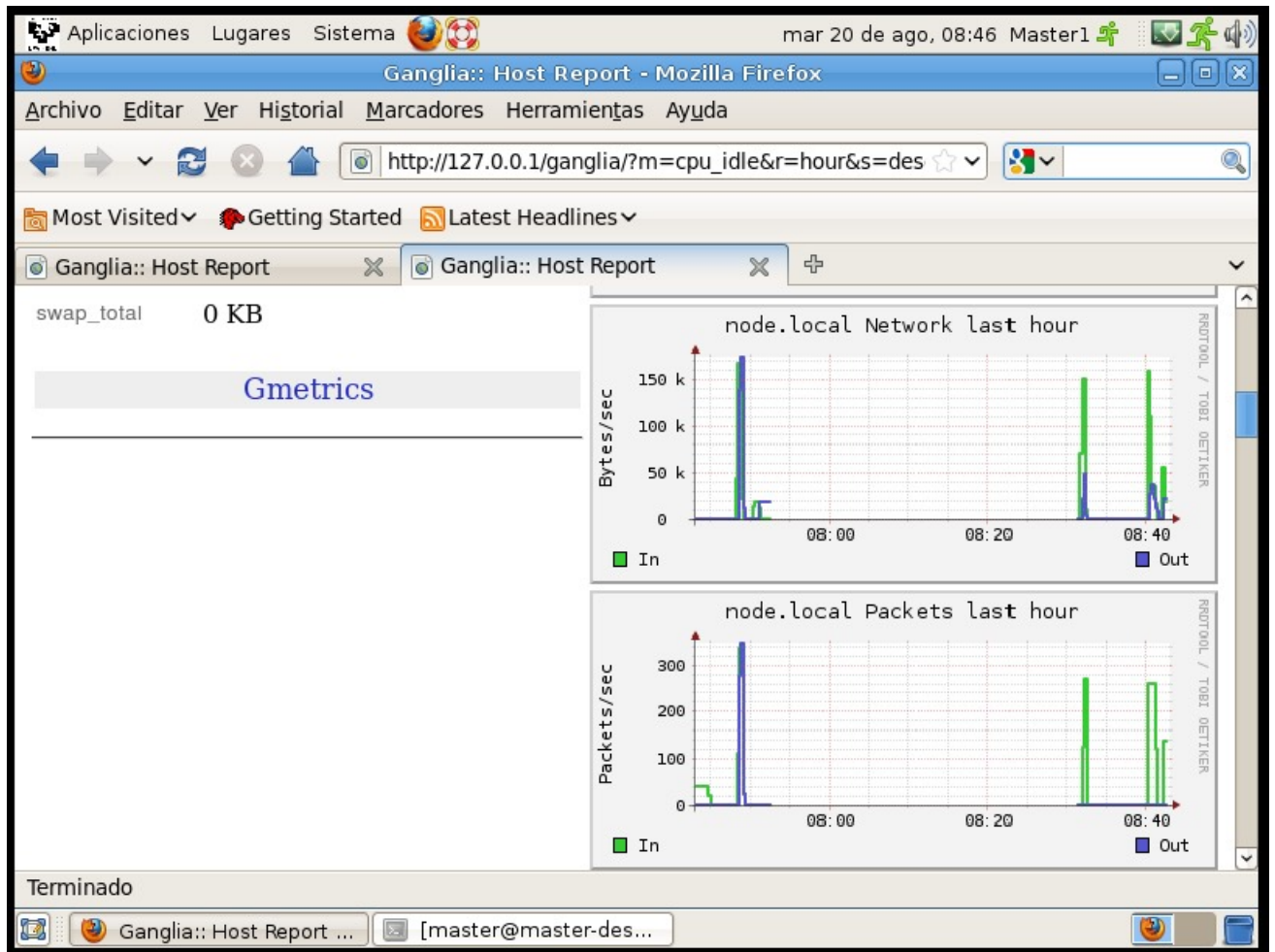


Figura 6: Gmetrics Nodo local

## Conclusiones

- Capacidades de Resolución de Problemas en Ingeniería: La instalación y configuración del clúster no solo permitieron comprender la arquitectura y el funcionamiento de un clúster, sino que también se aplicaron métodos analíticos, de cálculo, y experimentales para solucionar problemas específicos durante el proceso. Este enfoque demostró la capacidad para abordar y resolver problemas complejos en la ingeniería computacional, aplicando métodos establecidos de manera efectiva.
- Configuración de Recursos Distribuidos: El proceso de configuración de las máquinas virtuales y su integración en el clúster permitió observar cómo un conjunto de nodos puede actuar como un único recurso computacional, lo que es crucial

para la ejecución eficiente de aplicaciones distribuidas. El uso de herramientas como VMware y Cluster Monitor facilitó la verificación y monitoreo de la correcta integración de los nodos en el clúster.

- **Importancia de la Estandarización en la Configuración de Clústeres:** Se evidenció la importancia de mantener estándares claros y precisos en la configuración de clústeres, especialmente en la selección de nodos y en la asignación de roles dentro del clúster. Esto asegura la compatibilidad y la correcta ejecución de las aplicaciones distribuidas.

## Recomendaciones

- **Mejora Continua en la Documentación de Proyectos:** Es fundamental documentar cada paso del proceso de configuración y resolución de problemas durante la instalación del clúster. Esta documentación no solo servirá como una guía para futuras implementaciones, sino que también ayudará a identificar y corregir posibles errores que puedan surgir en el futuro.
- **Implementación de Políticas de Seguridad en el Clúster:** Considerar la implementación de políticas de seguridad robustas para proteger la integridad de los datos y la operación del clúster. Esto incluye la gestión adecuada de accesos, la configuración de firewalls, y la monitorización continua de la red.
- **Optimización del Uso de Recursos en Máquinas Virtuales:** Ajustar los recursos asignados a las máquinas virtuales, como la memoria y el procesamiento, basados en el tipo de cargas de trabajo que se espera que maneje el clúster. Esta optimización no solo mejorará el rendimiento del sistema, sino que también puede reducir costos y consumo energético.
- **Capacitación Adicional en Gestión de Clústeres:** Realizar capacitaciones adicionales en la gestión y administración de clústeres es esencial para asegurar que los operadores estén bien preparados para manejar situaciones inesperadas y para optimizar el rendimiento del clúster. Esto incluye el conocimiento de herramientas avanzadas de monitoreo y la capacidad de realizar ajustes dinámicos en la configuración del clúster.

## Referencias

- [1] ¿Qué es un cluster? (s. f.). <https://www.revista.unam.mx/vol.4/num2/art3/cluster.htm>
- [2] WebSphere Application Server for z/OS traditional 9.0.5.x. (s. f.). <https://www.ibm.com/docs/es/was-zos/9.0.5?topic=servers-introduction-clusters>