

Proyecto Básico: Clúster de ordenadores

Agenda

❖ Cluster de ordenadores

❖ Multicomputador

- ❖ Arquitectura y Componentes.

❖ Clusters de ordenadores

- ❖ Construir y desplegar un cluster
- ❖ Administración básica, planificación y balanceo de tareas.

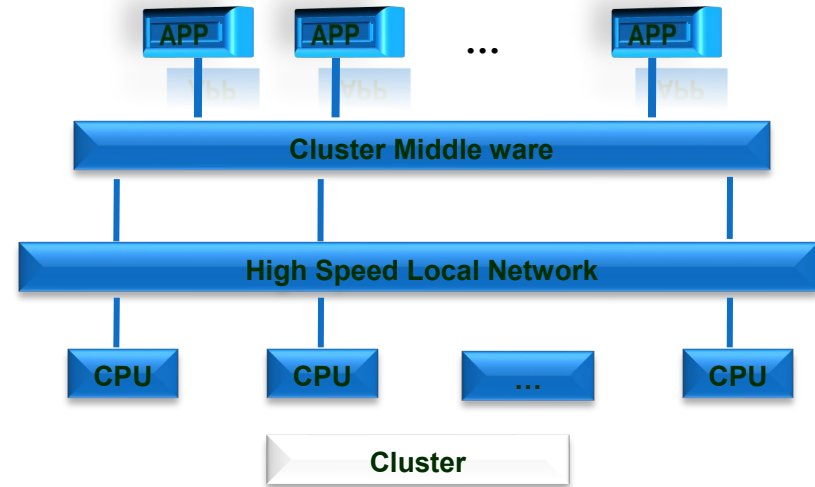
❖ Tipos de cluster y dominio de aplicación: HPC, Bigdata

Cluster de Ordenadores

- ❖ Un Cluster es la unión de múltiples ordenadores mediante una red de alta velocidad para formar un sistema de computación de alto rendimiento, seguro y fiable.

Cluster consta de :

- Nodes(master+computing)
- Network
- OS
- Cluster middleware



Características de un Cluster

- n Cada máquina en un cluster puede ser un sistema completo utilizable independientemente
- n Sistema fácilmente escalable
- n Sistema que se repara de manera muy simple, por sustitución de elementos

Arquitectura de un Cluster

- ❖ HW: Interconexión de un frontend con múltiples nodos de computo y recursos de almacenamiento.
- ❖ SW : Middleware que permita la ejecución distribuida y/o en paralelo de tareas.

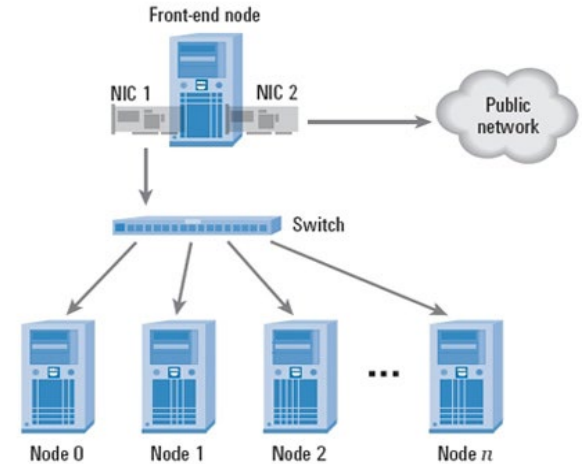
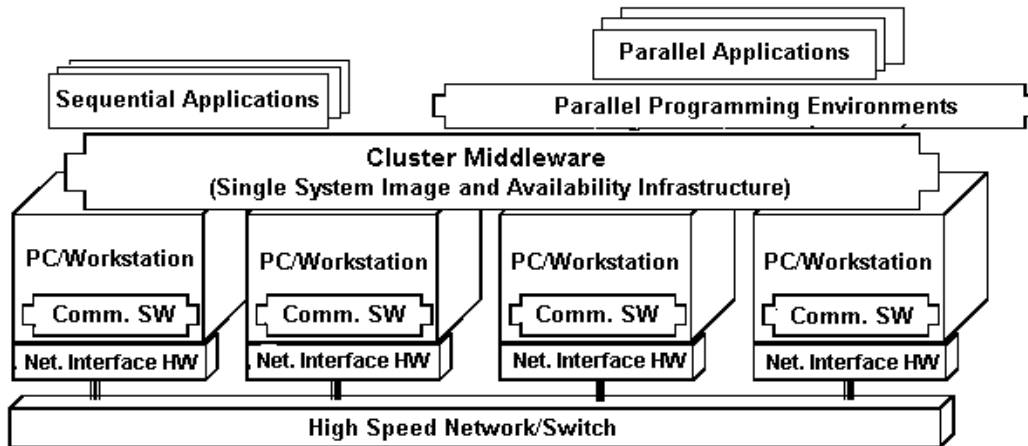
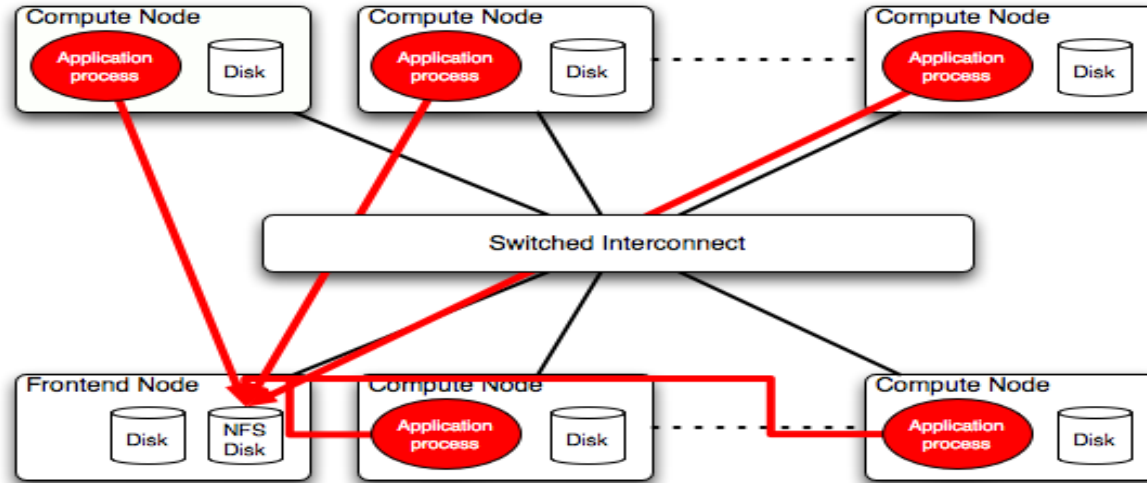


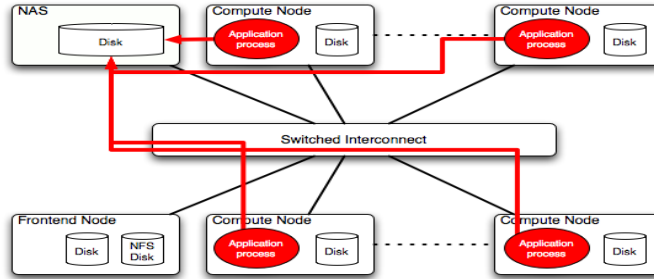
Figure 1. A typical Rocks cluster layout

Arquitectura del almacenamiento: Disco NFS Local



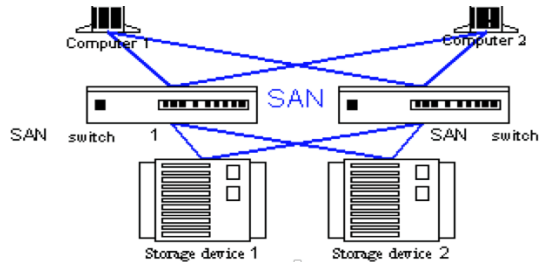
- ❖ El disco está exportado a los nodos de compute via NFS

NAS: Network Attached Storage



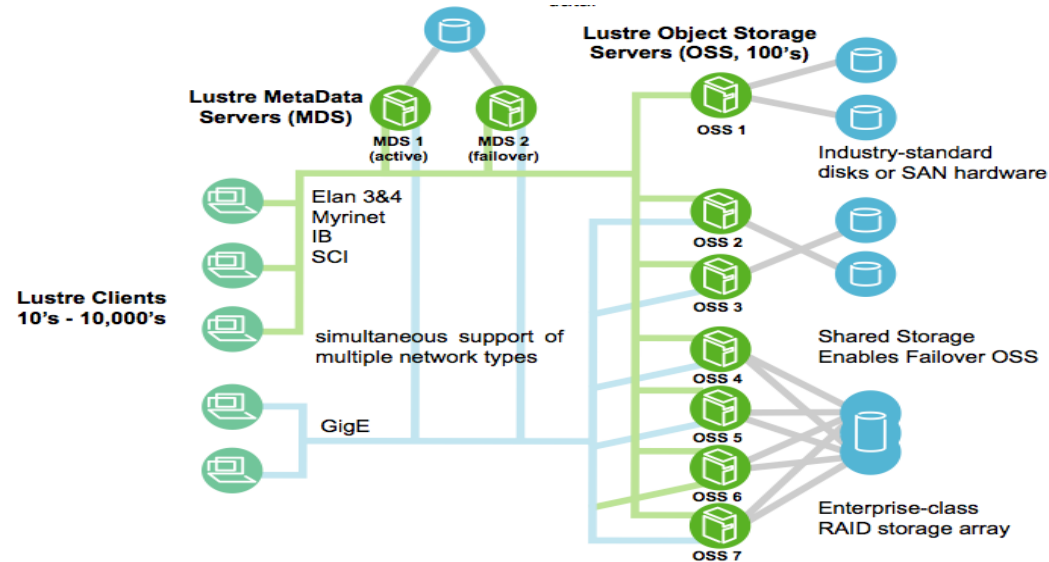
- ❖ Se puede incorporar como recurso de disco un NAS, que para el cluster es como una caja negra que funciona como un “appliance NFS”

SAN: Storage Area Network



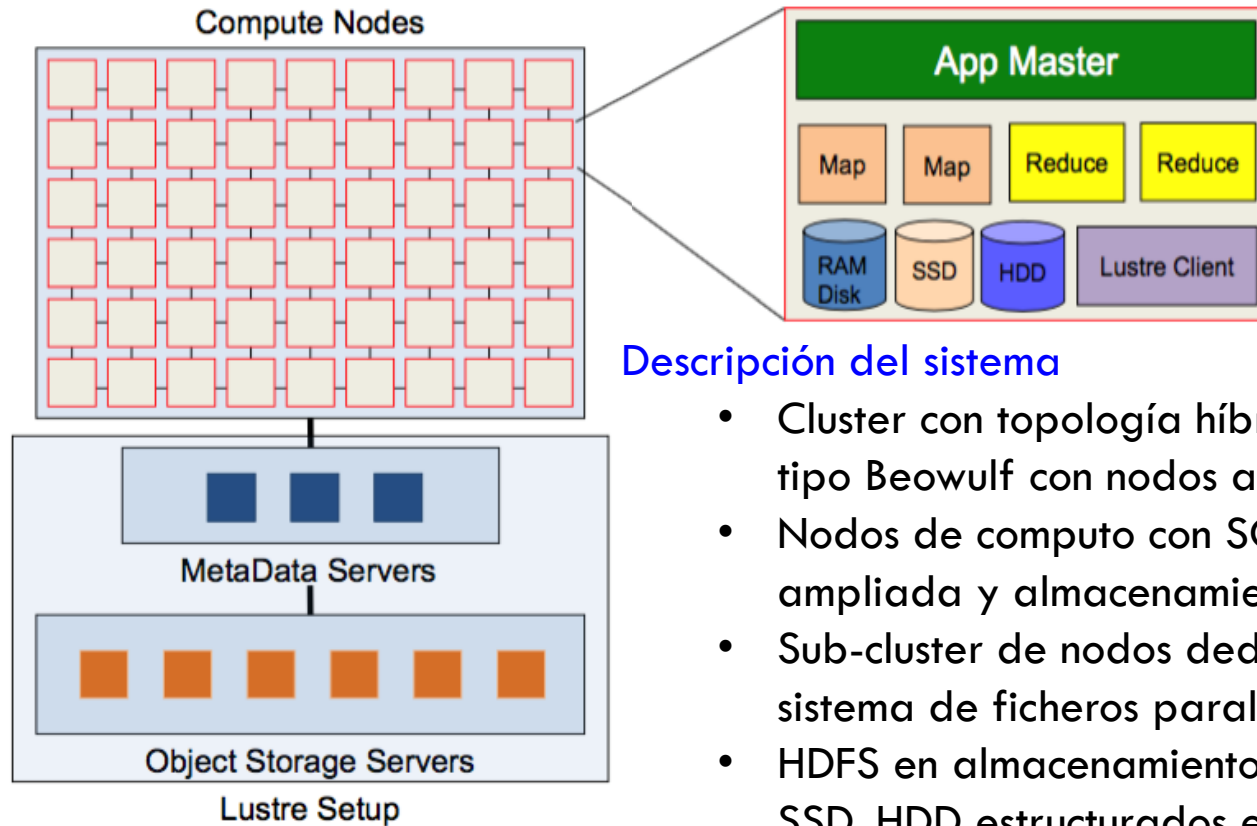
- SAN: El acceso a disco está sobre una conexión de red especializada (Fibre Channel / Ethernet)
- Los discos compartidos están fuera de de los servidores y son un elemento más de la red
- Se requiere un servicio central para coordinar las operaciones del sistema de ficheros y suele estar replicado para evitar punto de fallo único.

Arquitectura de archivos distribuido: Lustre



- ❖ Open Source “Object-based” storage
 - ❖ Ficheros son objetos, no bloques

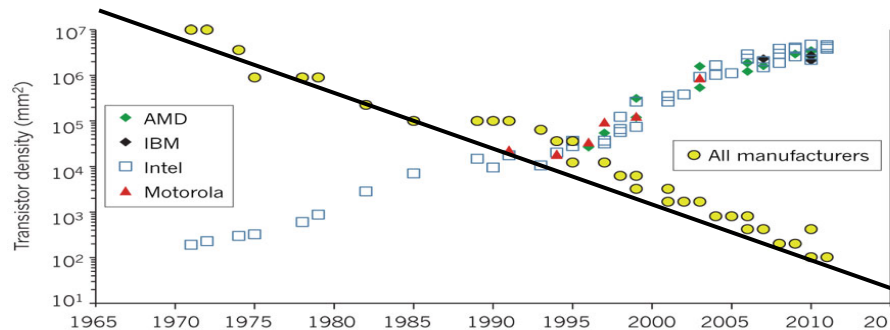
Cluster para HPC y Bigdata: Arquitectura de referencia



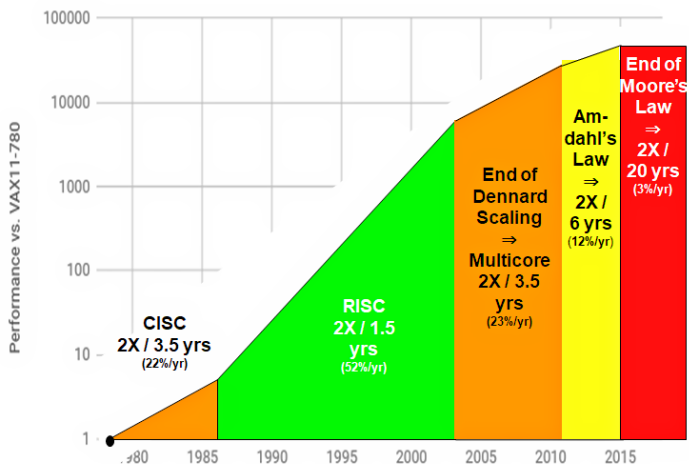
Descripción del sistema

- Cluster con topología híbrida de una arquitectura tipo Beowulf con nodos adicionales para I/O.
- Nodos de computo con SO versión ligera; memoria ampliada y almacenamiento local pequeño.
- Sub-cluster de nodos dedicados para I/O con un sistema de ficheros paralelo, (en la figura Lustre)
- HDFS en almacenamiento heterogéneo: RAMDisk, SSD, HDD estructurados en RAID, JBOD,...

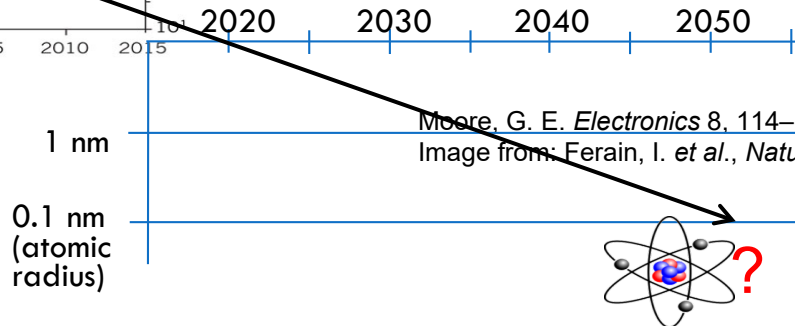
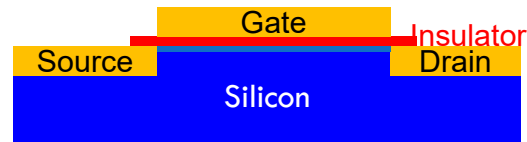
Rendimiento: Futuro de los procesadores



40 years of Processor Performance



... pronto entraremos en los límites físicos de los materiales



El rendimiento de los procesadores de propósito general se está estancando:

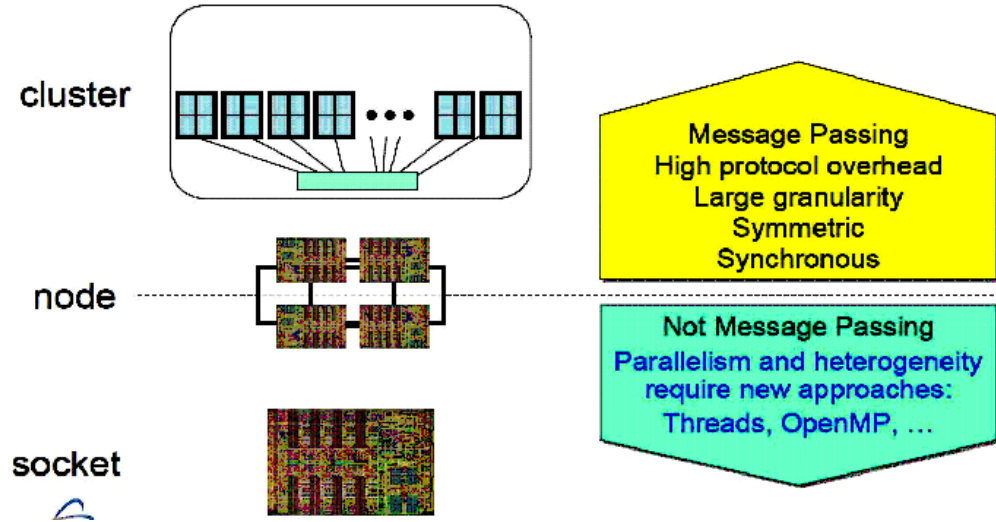
- Arquitecturas específicas
- Tecnologías más disruptivas: procesadores cuánticos.

Cluster: Arquitecturas y sus modelos de programación

Reto:

Entender como aprovechar las posibilidades de ejecución paralela que incorporan todos los procesadores y sistemas actuales y aprender a sacar partido a sus modelos de programación.

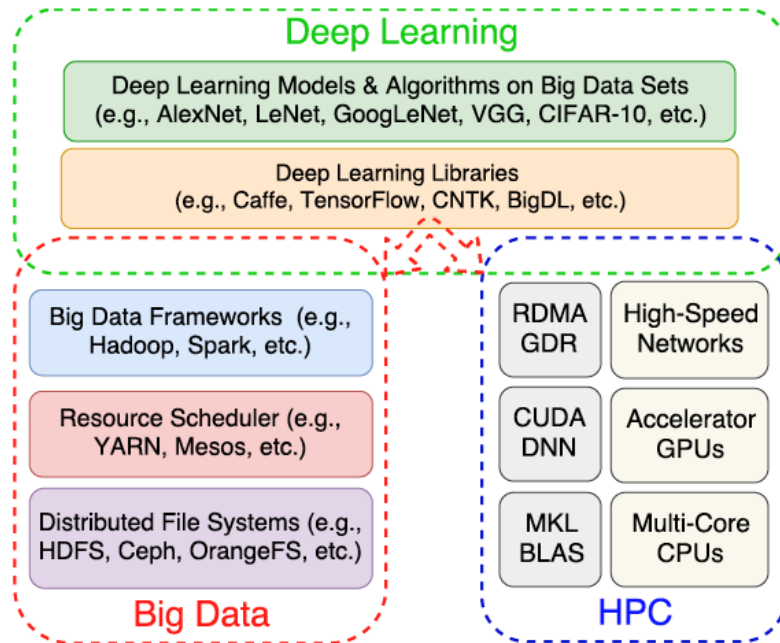
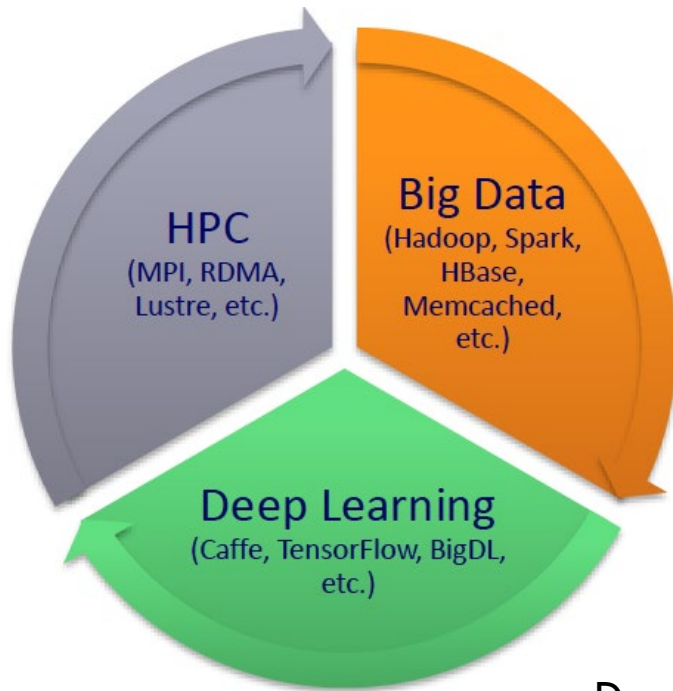
Arquitecturas paralelas y sus modelos de programación, diferenciando:



- **Hasta decenas de cores: Multicore de memoria compartida con OpenMP**
- **Cientos de cores y aceleradores: Manycore con programación GPGPU**
- **Con muchas máquinas: Multicomputador y programación con MPI**
- **Tecnologías para BigData: Hadoop y Spark**

Sistemas para BigData, HPC y DL

Convergencia de High Performance Computing(HPC), Big Data, y Deep Learning (DL)



Deep Learning es un subconjunto de Machine Learning,
que está revolucionando los entornos de Big Data

Arquitecturas para BigData: Aceleradores/Coprocesadores GPU

Distributed Deep Learning

100s of servers with GPUs

scale of the computational infrastructure enabled by IBM's communication library for Distributed Deep Learning training

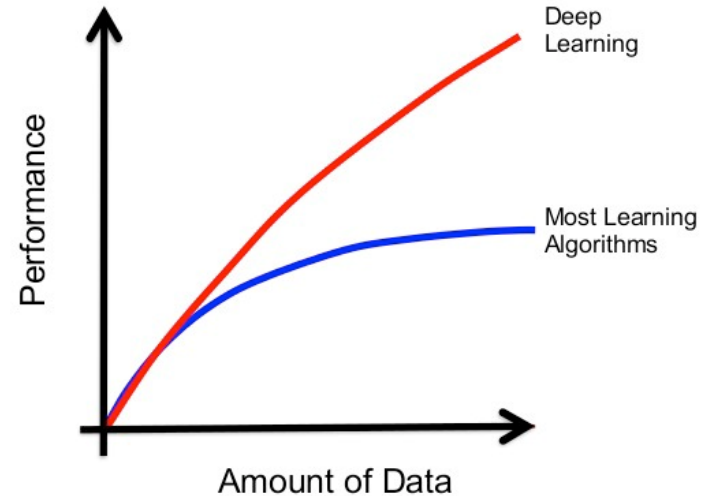
95%

scaling efficiency achieved by IBM @ 256 P100 GPUs

+4%

increase in image recognition accuracy over previous best result

BIG DATA & DEEP LEARNING



Infraestructura de los sistemas para BigData:

Cluster de ordenadores:



Solución *low-cost*

Opciones de integración de arquitecturas Big Data

Arquitecturas físicas

Plataformas Cloud

Almacenamiento datos en la nube

Ahorro en el hardware

Mantenimiento por parte del
proveedor

Menor control del sistema
desarrollado

Servidor especializado:



Ejemplo:

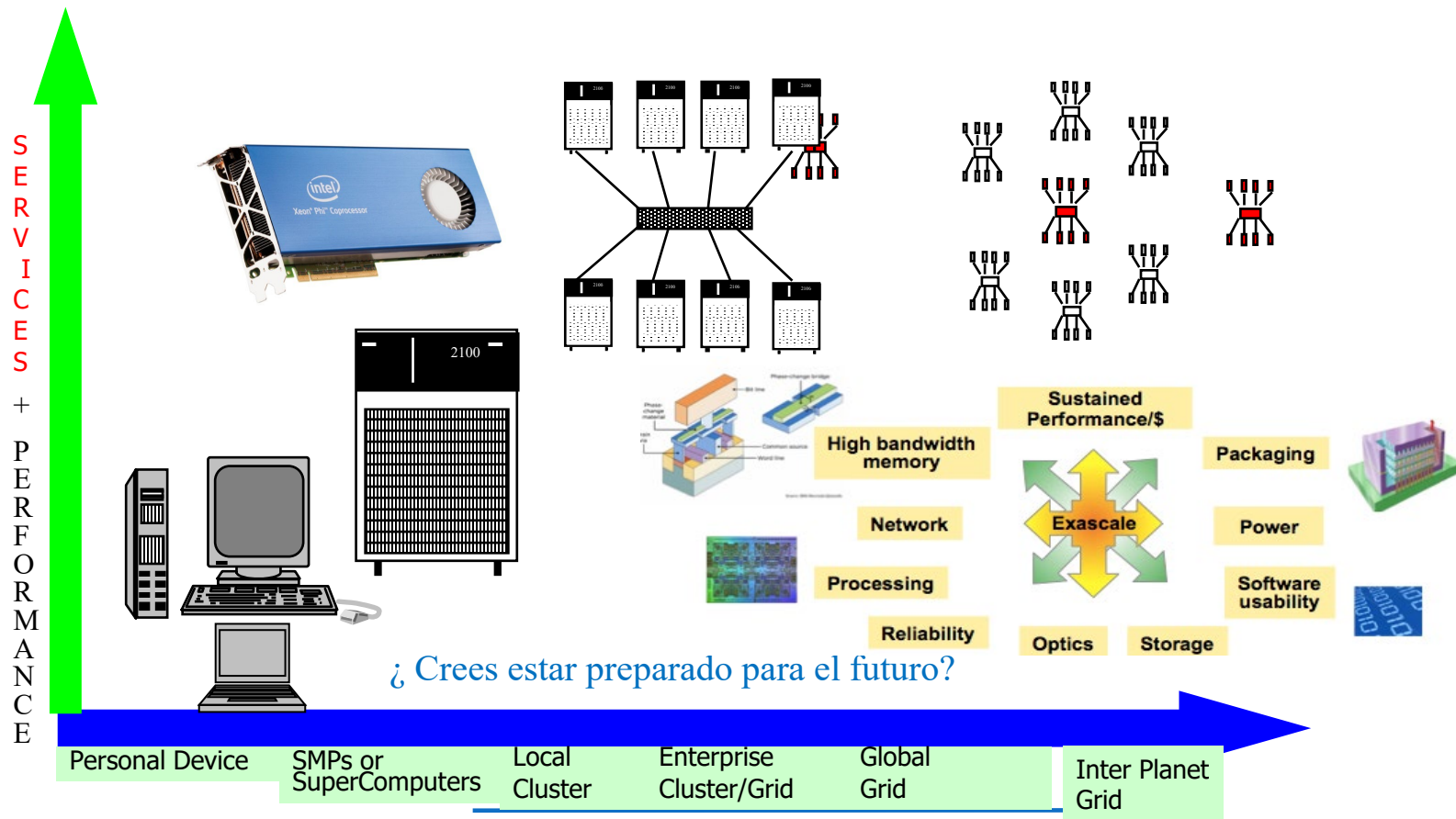


Coste fijo + coste variable

Soluciones adaptables

Escalabilidad automática

Cómo seguimos evolucionando... www.top500.org



Tipos de cluster por su finalidad

Objetivo 1: resolver grandes desafíos o aplicaciones muy exigentes en recursos

- ❖ **Cluster de alto rendimiento para HPC**

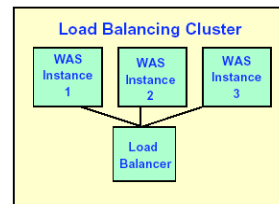
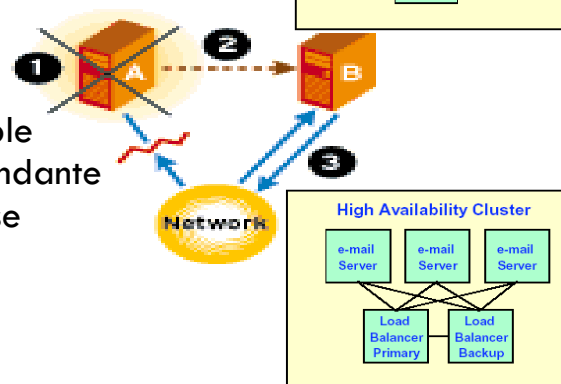
- ❖ Enlazar muchos nodos de computación (ordenadores) para conseguir que funcionen en equipo y obtengan la solución de un problema más rápidamente trabajando todos juntos de manera independiente.



Objetivo 2: que funcionen aplicaciones críticas.

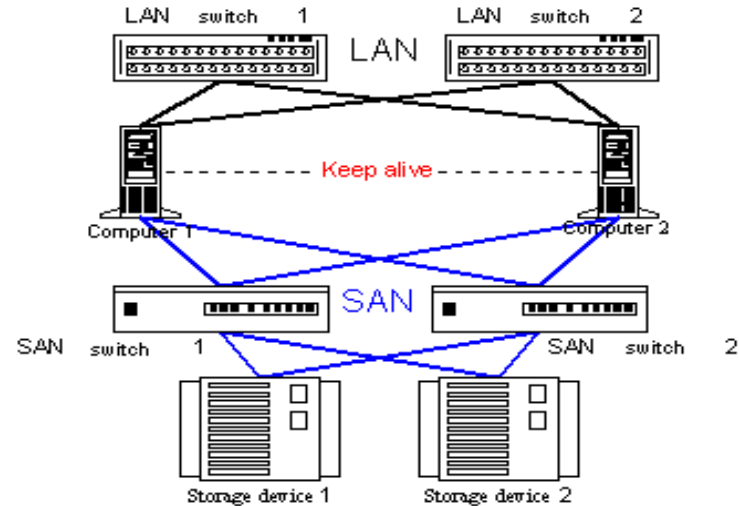
- ❖ **Cluster de alta disponibilidad**

- ❖ Conseguir un sistema de ordenadores mas fiable compartiendo trabajos y funcionamiento redundante de tal manera que si un ordenador falla otro se encarga de realizar su trabajo.

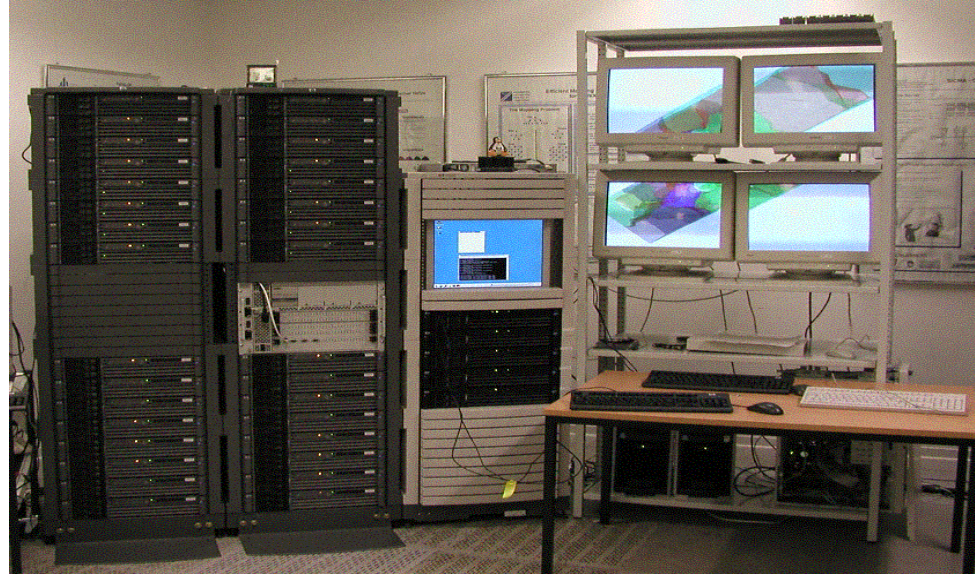


Arquitectura de un cluster de alta disponibilidad

- ❖ Replicación:
 - ❖ Se diseña con componentes redundantes y múltiples caminos de comunicación.
 - ❖ Evitar puntos de fallo único



Otras arquitecturas: Cluster de Visualización



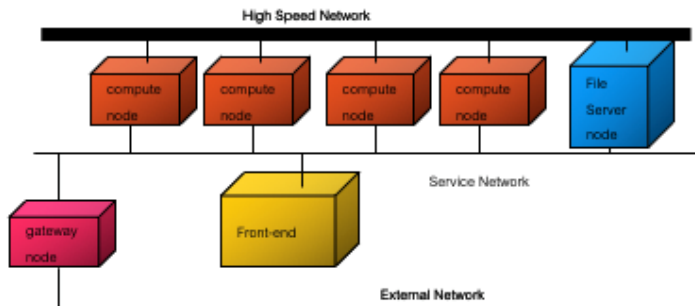
- ❖ Cada nodo en el cluster gestiona un monitor

Tipos de cluster por su configuración

Cluster cerrado:

Se oculta el cluster detrás de un nodo gateway

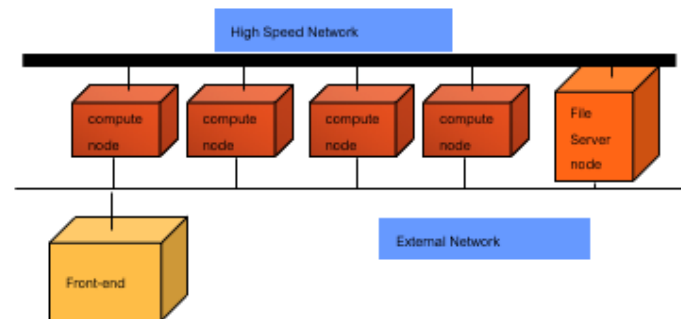
- ❖ Se necesitan menos direcciones IP.
- ❖ Mas seguridad
- ❖ Tareas computacionales grandes.



Cluster abierto:

Todos los nodos son visibles desde el exterior

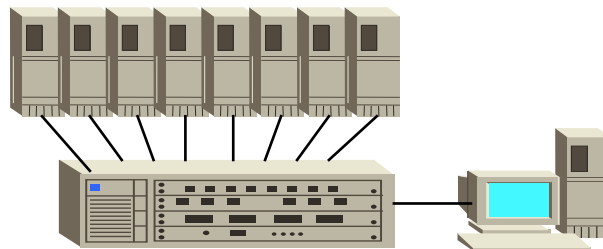
- ❖ Se necesitan muchas direcciones IPs,
- ❖ Mas difícil de controlar aspectos de seguridad
- ❖ Más flexible
- ❖ Tareas como servidor de internet/web



Cluster Beowulf: filosofía en requisitos y objetivos

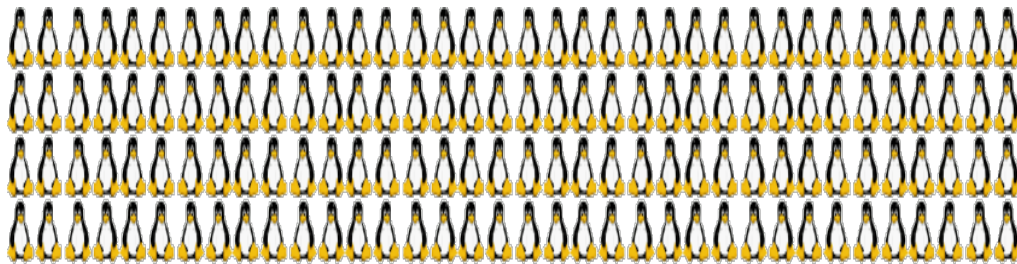
Requisitos de los componentes de cluster Beowulf

1. Ordenadores: PCs, Estaciones de trabajo, plataformas multiprocesador...
 - n Independencia del fabricante y bajo coste
2. Sistema Operativo que admita funcionamiento en cluster
 - n Libre distribución y Open source
3. Red de interconexión
 - n Rapidez
 - n Bajo coste



Beneficios potenciales:

- ❖ Rendimiento
- ❖ Disponibilidad
- ❖ Balanceo de carga
- ❖ Escalabilidad
- ❖ Fiabilidad y Seguridad



El proyecto Beowulf

- ❖ Proyecto en el Center of Excellence and Information Systems Sciences (CESDIS) de NASA Goddard Space Center,



· Dr. Thomas Sterling, Donald Becker que lo definen como:

"Beowulf is a project to produce the software for off-the-shelf clustered workstations based on commodity PC-class hardware, bandwidth internal network, and the Linux operating system." <http://www.beowulf.org/>



Clasificación de los cluster

High availability clusters (HA) (Linux)

**Mission critical
applications**

High-availability clusters (also known as Failover Clusters) are implemented for the purpose of improving the availability of services which the cluster provides.

provide redundancy

**eliminate single points
of failure.**

Network Load balancing clusters

**operate by distributing
a workload evenly
over multiple back
end nodes.**

**Typically the cluster
will be configured
with multiple
redundant load-
balancing front ends.**

**all available servers
process requests.**

**Web servers, mail
servers,..**

Science Clusters

**High-
performance
(HPC) clusters**

**Beowulf
Special
purpose**

*to provide
greater
computational
power than a
single computer
can provide.*

A decorative graphic on the left side of the slide. It features a vertical orange bar on the far left. To its right, a series of orange and blue squares are arranged in a stepped pattern, resembling a staircase. A solid blue horizontal bar extends from the right edge of the squares across the top of the slide.

Rocks

Rocks : Objetivo principal

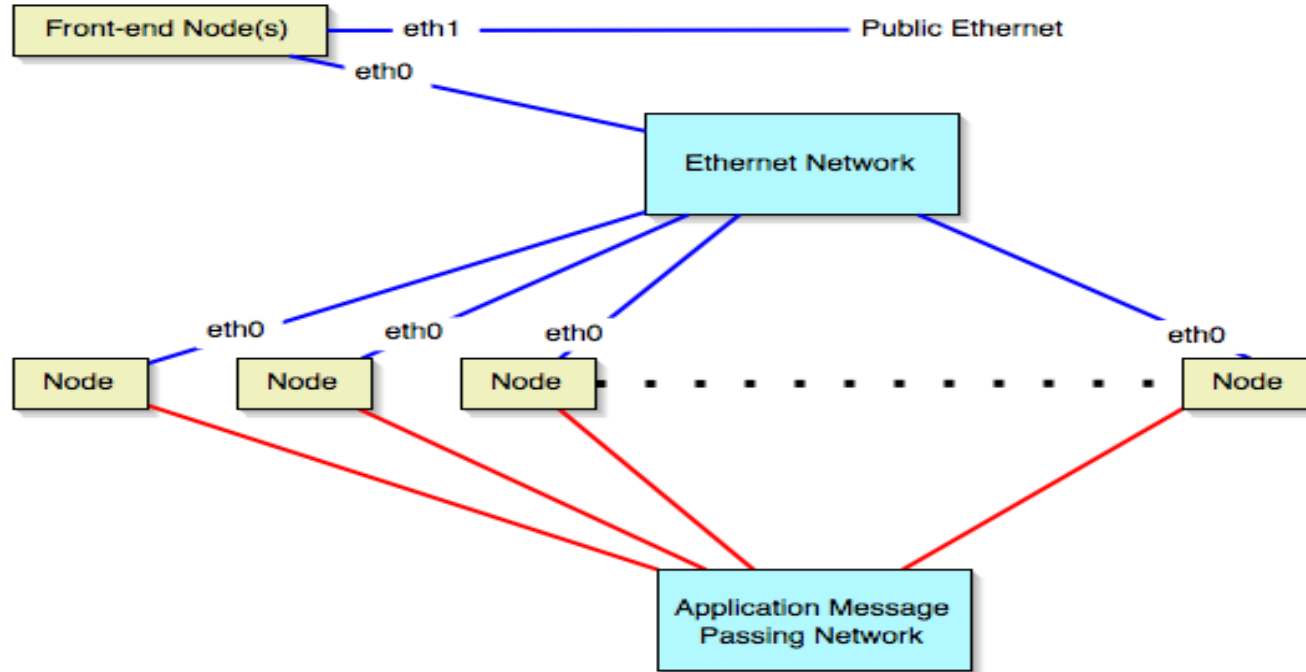
- ❖ Facilitar la instalación de un cluster.

Make clusters easy !



- ❖ Audiencia: Científicos que desean tener un alta computación en su propio laboratorio.

Arquitectura de un cluster Rocks para HPC (High Performance Cluster)



Proceso de Instalación

❖ Instalar a frontend

1. Insert Rocks Base CD
2. Insert Roll CDs (optional components)
3. Answer 7 screens of configuration data
4. Drink coffee (takes about 30 minutes to install)

❖ Instalar los nodos de computo:

1. Login to frontend
2. Execute insert-ethers
3. Boot compute node with Rocks Base CD (or PXE)
4. Insert-ethers discovers nodes
5. Goto step 3

❖ Añadir usuarios

❖ Empezar a utilizarlo

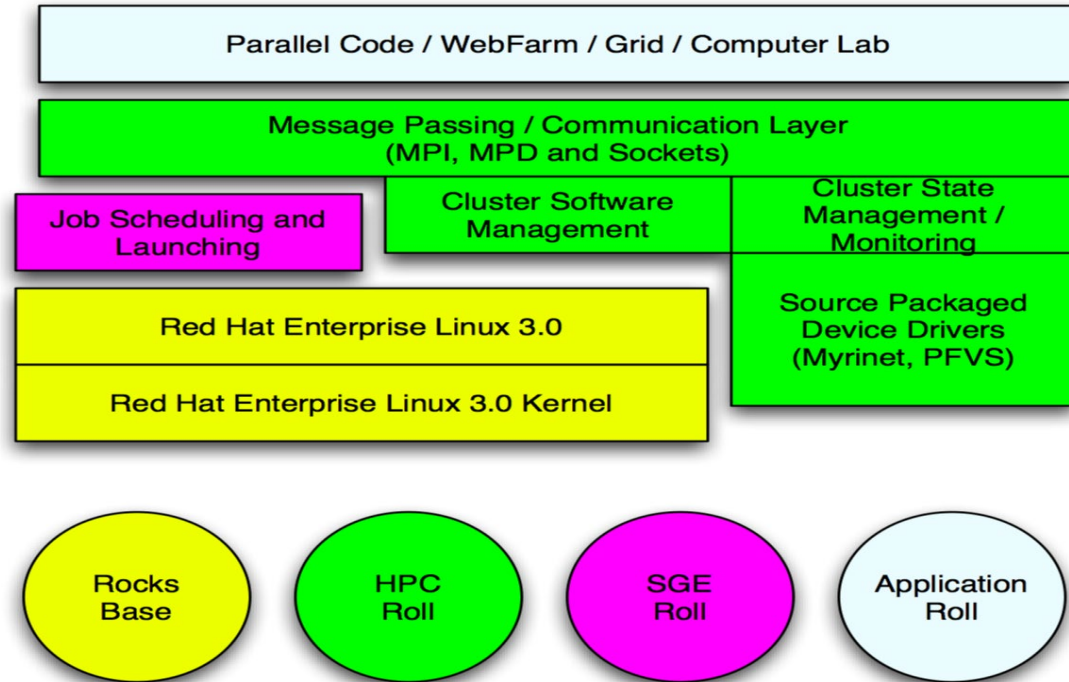


Opcional (Rolls)

- ☐ **Condor**
- ☐ **Grid (based on NMI R4)**
- ☐ **Intel (compilers)**
- ☐ **Java**
- ☐ **SCE (developed in Thailand)**
- ☐ **Sun Grid Engine**
- ☐ **PBS (developed in Norway)**
- ☐ **Area51 (security monitoring tools)**

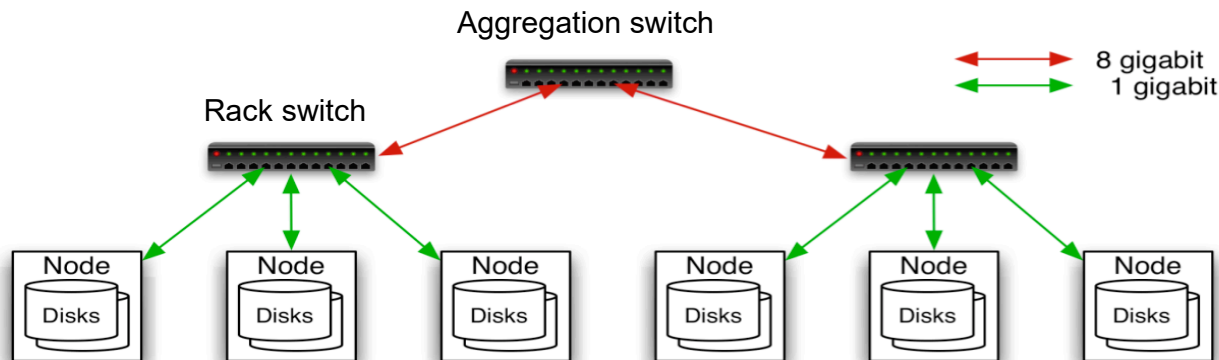
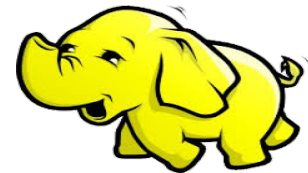
Rocks & ‘Rolls’

- ❖ Rolls son contenedores de paquetes software y los scripts de configuración asociados.

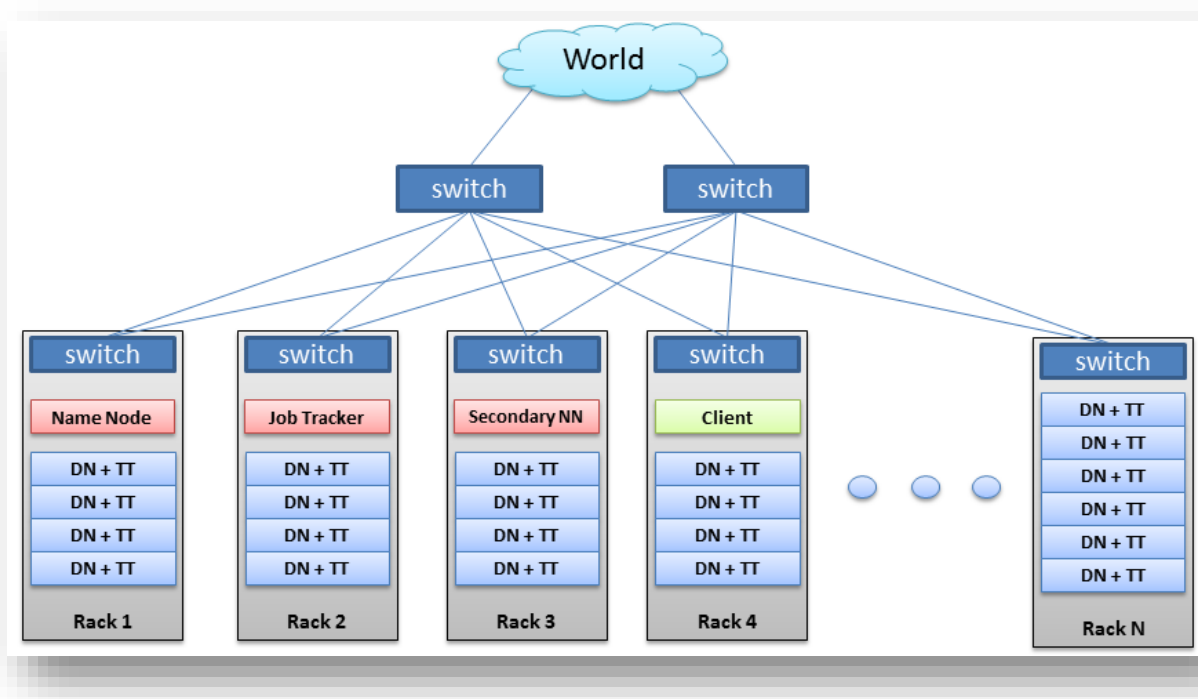


Cluster Hadoop

- ❖ Cluster creado con commodity Hardware;
 - ❖ Nodos inicialmente eran PCs
 - ❖ 30-40 nodos/rack
 - ❖ Red a 1 gigabit/s en rack



Arquitectura de almacenamiento HDFS



HDFS : Hadoop Distributed File System

- ❖ Sistema de Ficheros distribuido muy grande
 - ❖ 10K nodos, 100 millones de ficheros 10PB
- ❖ Realizado con “*Commodity Hardware*”
 - ❖ Ficheros replicados para tolerancia a fallos
 - ❖ Detecta fallos y recupera los datos.
- ❖ Optimizado para proceso por lotes (“*Batch Processing*”).
 - ❖ Expone la localización de los datos y así permite que la computación se pueda llevar cerca de los datos.
 - ❖ El ancho de banda agregado es muy alto.



Cluster de ordenadores

Ejercicio 1: Instalación de un cluster Rocks para HPC



Cluster ROCKS

- ❖ Instalación guiada utilizando Máquinas virtuales
- ❖ Ejemplos de administración básica

Bibliografía y enlaces de interés

❖ Bibliografía

- ❖ High Performance Cluster Computing: Architectures and Systems by Rajkumar Buyya Ed Prentice Hall PTR; ISBN: 0130137847; 1st edition (1999)
- ❖ Linux Cluster Architecture by Alex VreniosSams; ISBN: 0672323680; (2002)
- ❖ Linux Clustering: Building and Maintaining Linux Clusters by Charles Bookman Ed:New Riders Publishing; ISBN: 1578702747; 1st edition (2002)

❖ Enlaces

- ❖ <http://www.hispacluster.org/>
- ❖ <http://www.openclustergroup.org/>
- ❖ <http://www.beowulf.org/> y <http://www.beowulf-underground.org/>
- ❖ <http://www.mosix.com/>