

# Cómputo de Alta Disponibilidad

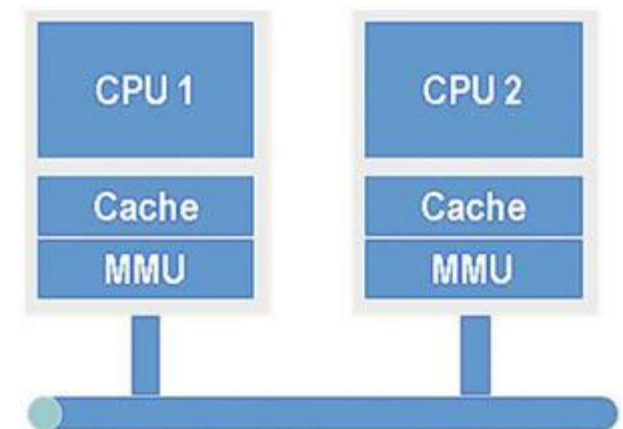
Abril 2015

# Introducción

- Actualmente existen diversas arquitecturas de hardware que mejoran su capacidad de escalabilidad, por ejemplo:
  - Sistemas multiprocesador que trabajan con más de un procesador compartiendo una memoria principal global y el mismo sistema de E/S.

# Multiprocesador vs. Multinúcleo

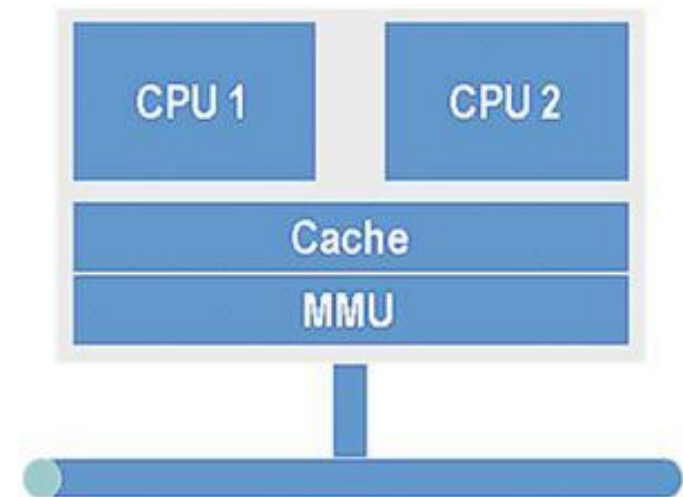
- Multiprocesador
  - Un sistema multiprocesador contiene más de un CPU que no se encuentran en el mismo chip
  - Los sistemas multiprocesador son menos complejos que los sistemas multinúcleo
  - La desventaja con los sistemas de multiprocesador es que son más caros porque requieren múltiples chips que una solución de chip único.



# Multiprocesador vs. Multinúcleo

- Dual-Core, Multinúcleo

- Los procesadores de doble núcleo son dos procesadores en un solo chip.
- Los procesadores de múltiples núcleos son una familia de procesadores que contienen cualquier número de CPUs (2, 4 y 8) en un solo chip.
- 
- El reto con procesadores multinúcleo está en el área de desarrollo de software.
- El rendimiento está directamente relacionado con la forma de hacer paralelo el código fuente de una aplicación (multi-threading).



# Hyperthreading

- Es una tecnología que se introdujo por Intel
- Tiene el propósito principal de mejorar el código multihilo
- Permite un uso más eficiente de los recursos del CPU mediante la ejecución de subprocesos “en paralelo” en un único procesador
- Un procesador hyperthreading pretende ser dos procesadores “lógicos” en un único procesador
- Procesadores Pentium 4 son un ejemplo de las CPU que implementan hyperthreading



# Motivación

- Intercambio de recursos
  - Compartir e imprimir archivos
  - Procesamiento en una base de datos distribuida
  - Uso de hardware especializado de forma remota
- Aceleración de cómputo – Balanceo de cargas
- Confiabilidad – Detección y recuperación de fallas
- Comunicación – Paso de mensajes

# Motivación tecnológica

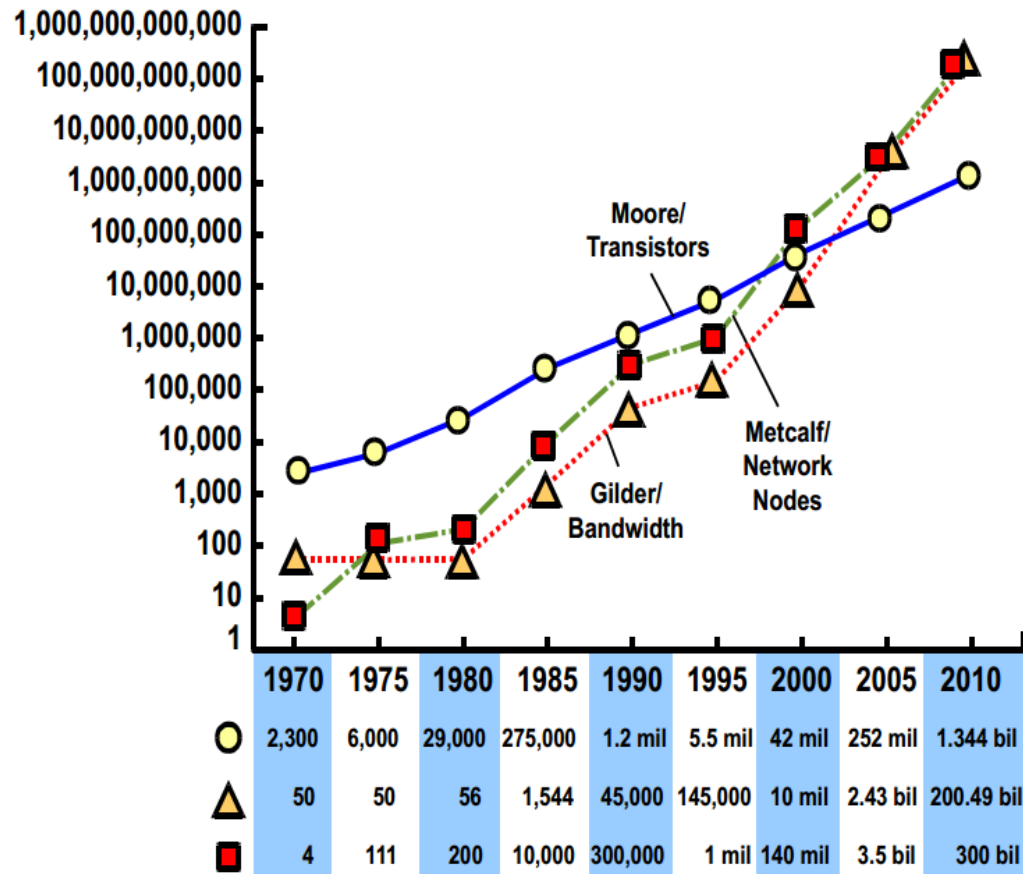
- Computadoras más baratas y veloces
  - En 1945 las computadoras ejecutaban 1 instrucción por segundo y costaban \$10,000,000.00 USD
  - Ahora superan el millón de instrucciones por segundo y cuestan \$1000.00 USD aprox.
- El desarrollo de redes de alta velocidad
  - Tecnologías LAN de 100, 1000, 10000 Mbps
  - Tecnologías WAN de 10 Gbps

# Leyes de cómputo

- Ley de Moore
  - El poder de procesamiento de un chip se duplica cada 18 meses mientras el costo se mantiene constante.” **Gordon Moore, Fundador de Intel**
- Ley de Metcalfe
  - El valor de una red para la comunidad es igual al cuadrado del número de usuarios. **Robert Metcalfe, Fundador de 3Com**
- Ley de Gilder
  - El ancho de banda se duplica cada año. **George Gilder, Economista Especializado en IT**
- Ley de Horn
  - El costo de almacenar un bit es inferior a un centavo y bajará a una velocidad mayor que aquella a la que crece el ancho de banda. **Paul Horn, Director de Investigación de IBM**



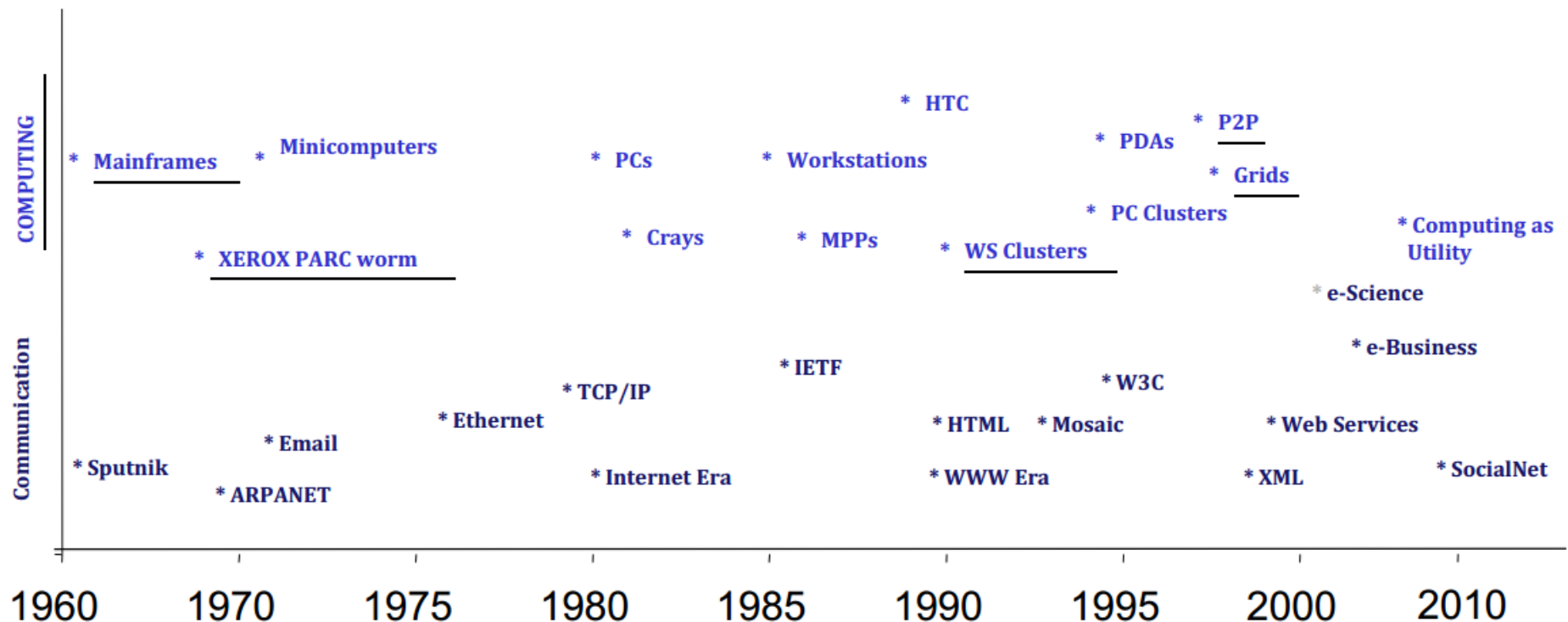
# Leyes de cómputo



Source: Cambridge Energy Resource Associates

# Cómputo y Comunicaciones

## Evolución tecnológica: 1960-2010



# Aplicaciones que demandan recursos

- Ciencias de la salud
- Aplicaciones militares
- Internet
- Comercio
- Aeroespacial

# Películas procesadas usando clúster



# ¿Qué es un clúster?

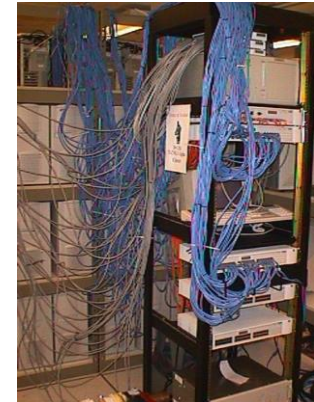
- Es un sistema de procesamiento paralelo o distribuido, que consiste en una colección de computadoras interconectadas que cooperan entre sí como un único e integrado recurso de cómputo.
- Las computadoras son interconectadas con dispositivos de alta velocidad.

# Clúster

- La construcción es sencilla y económica debido a su flexibilidad:
  - Pueden tener la misma configuración de hardware y sistema operativo **cluster homogéneo**
  - Diferente rendimiento pero con arquitecturas y sistemas operativos similares **cluster semihomogéneo**
  - Diferente hardware y sistema operativo **cluster heterogéneo**

07/04/2015

# ¿Qué es un clúster?



# Tipos de clúster

- ***Clúster computacional.*** Se utilizan para cualquier tarea que requieran enormes cantidades de cómputo. Entre sus prioridades se encuentra la escalabilidad.
- ***Clúster de alta disponibilidad.*** Son capaces de cubrir fallos que se generan en él. Cuando ocurre un error en un nodo, la aplicación ejecutada es migrada a otro nodo. La prioridad no es la escalabilidad sino la infraestructura y los servicios para ofrecer alta disponibilidad.



# Tipos de clúster

- ***Clúster de balanceo de carga.***
- Con el crecimiento de Internet en los últimos años el tráfico en la red ha aumentado junto con la carga de trabajo que ha de ser soportada por los servidores (p.e. Web).
- Para cumplir con estos requerimientos hay dos soluciones:
  - El servidor se basa en una máquina de alto poder de cómputo que a largo plazo probablemente quede obsoleta por el crecimiento de la carga
  - Un clúster de servidores. Cuando la carga de trabajo crezca, se añadirán más servidores al cluster.
- Una solución es utilizar un balanceo de cargas para distribuir el procesamiento entre los servidores del clúster.
- El proyecto **Linux Virtual Server** (LVS) permite construir un cluster de servidores que implementa alta disponibilidad y balanceo de carga sobre el sistema operativo GNU/Linux.

# Aplicaciones que usan clústers

- Científicas y de ingeniería
  - Predicciones climáticas
  - Simulaciones automotrices
- De negocios
  - E-commerce (amazon, eBay)
  - Bases de Datos (Oracle ERP)
- Internet
  - Aplicaciones proveedoras de servicios
  - Buscadores
  - E-Business
- Misión crítica
  - Sistemas de control
  - Bancos
  - Muelles

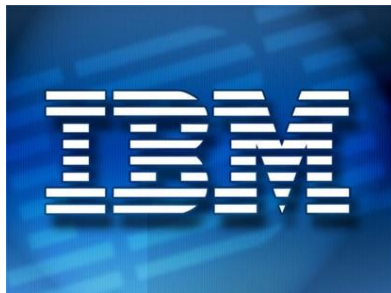
# Compañías detrás del clúster



**Microsoft®**

GRIDWARE  
STATE SPACE GRIDS

**COMPAQ**



**Sun**  
microsystems

**Google™**

**ORACLE®**

# Desventajas

- La mayoría de Clusters requieren arquitecturas y/o sistemas operativos homogéneos
- Si se agregan nuevos nodos deben cumplir con una regla homogeneidad

# Grid

- Un tipo de sistema paralelo y distribuido que habilita el intercambio, selección, agregación y compartición de recursos autónomos **distribuidos geográficamente**
  - Computadoras
  - Software
  - Bases de datos
  - Instrumentos o dispositivos especiales

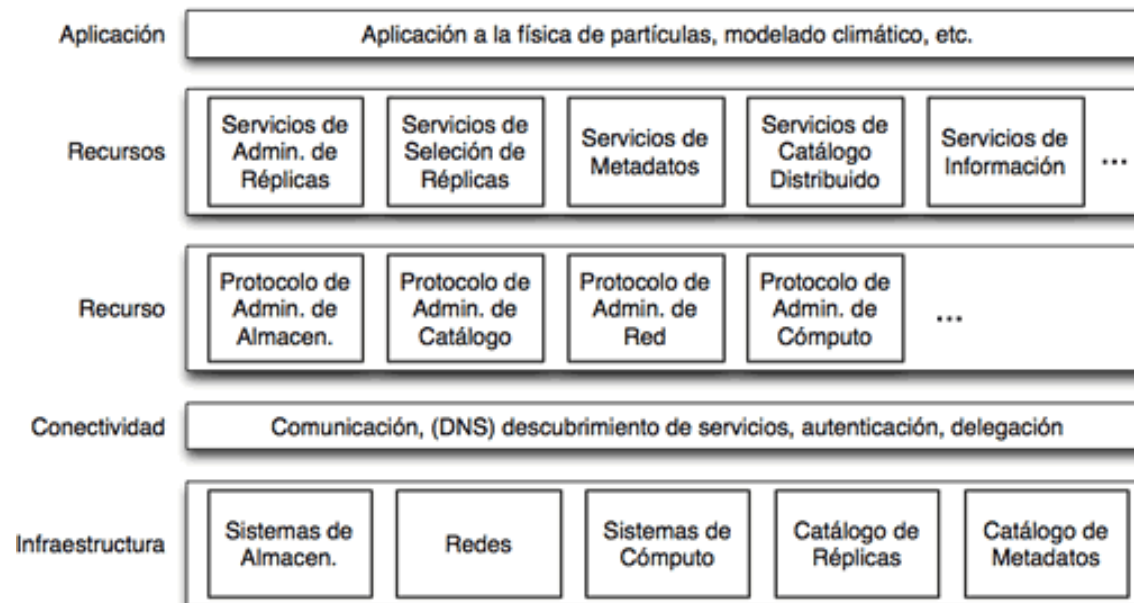
# Grid (malla)

- Un Grid consiste en la compartición de potencia computacional, aprovechando los ciclos de procesamiento no utilizados de cada nodo que pertenecen al Grid.

# Arquitectura

- **Nivel de infraestructura.** Recursos computacionales, computadoras, clúster, etc.
- **Nivel de conectividad.** Protocolos de comunicación (TCP/IP).
- **Nivel de recurso.** Protocolos para controlar el recurso local (proceso, tarea).
- **Nivel de recursos.** Servicios que permiten gestionar el conjunto de recursos. Por ejemplo, descubrir y ubicar recursos, un reloj distribuido, servicios de monitoreo.
- **Nivel de aplicaciones.** Se encuentran definido los protocolos que permiten el acceso al grid. Las aplicaciones se construyen en término de servicios.

# Arquitectura





# Características

- **Balanceo de sistemas**, no se calcula la capacidad de los sistemas en función al rendimiento, la capacidad se puede reasignar a donde sea necesario
- **Alta disponibilidad**, si un servidor falla, se reasignan los servicios en los servidores restantes.
- Reducción de costos, los servicios son gestionados por granjas de recursos.
- NO es necesario disponer de servidores con grandes recursos, se puede hacer uso de componentes de bajo costo.

# Características

- A diferencia de un clúster, en un grid los nodos pueden estar distribuidos a nivel mundial.
- Las computadoras pueden tener diferentes sistemas operativos.



# Desventajas

- Debido a que los nodos pueden ser distribuidos geográficamente la seguridad es un aspecto importante a considerar
- No todas las aplicaciones pueden ser paralelas o distribuidas
- La configuración de una Grid puede afectar la calidad y la fiabilidad de la infraestructura
- Comunicación lenta y no uniforme