

Arquitecturas Escalables

Bloque II. Tema 1

Gestión de Sistemas e Informática

Curso 2017 - 18

Juan José Aguado Gil

20 Julio 2017

1. Tipos de Procesadores

1.1. CISC: Complex Instruccion Set Computer

Ordenador con un Conjunto Completo de Instrucciones. Arquitectura tradicional de los PC donde la CPU utiliza microcódigo para ejecutar un conjunto muy amplio de instrucciones. Éstas pueden ser de longitud variable y usar todos los modos de direccionamiento, requiriendo para ello complejos circuitos decodificadores que ralentizan la ejecución de cada instrucción. Utilizan una Unidad de Control microprogramada.

1.2. RISC: Reduced Instruccion Set Computer

Ordenador con un Conjunto Reducido de Instrucciones. Utilizan instrucciones de longitud fija y los modos de direccionamiento son escasos y sencillos (no utilizan direccionamiento indirecto). Reducen el conjunto de instrucciones significativamente. Los procesadores RISC ejecutan instrucciones con mucha rapidez por la sencillez de las mismas. Utilizan una Unidad de Control cableada.

2. Taxonomía de Flynn

Clasificación de una arquitectura de ordenadores basada en el número de instrucciones concurrentes y los flujos de datos disponibles.



Figura 1: Mr. Michael Flynn.

		Instrucciones	
		SI	MI
Datos	SD	SISD	(MISD)
	MD	SIMD	MIMD

Figura 2: Taxonomía de Flynn (1966).

2.1. SISD: Single Instruction, Single Data

No se explota el paralelismo a nivel de instrucciones ni de datos, Ej: un PC de un solo núcleo. La unidad de control toma instrucciones de un flujo de instrucciones desde la memoria. La UC genera señales de control apropiadas para operar sobre flujos de datos sencillos, es decir, una operación al mismo tiempo.

2.2. SIMD: Single Instruction, Multiple Data

Se explota el paralelismo a nivel de flujos de datos. Se ejecutan secuencialmente instrucciones de un único flujo de instrucciones, que van operando simultáneamente sobre diversos flujos de datos. Ej: GPU (Graphics Processing Units), procesadores vectoriales y matriciales.

2.3. MISD: Multiple Instruction, Single Data

Múltiples flujos de instrucciones operan sobre un solo flujo de datos. Es la estructura menos común que normalmente se dirige a una arquitectura tolerante a fallos, como el control de vuelo

del Space Shuttle.

2.4. MIMD: Multiple Instruction, Multiple Data

Varias unidades funcionales realizan diferentes operaciones sobre diferentes datos.

3. Tipos de Arquitecturas Multi-Procesador

3.1. MPP: Massive Parallel Processing

Computador masivamente paralelo. Formado por muchos nodos individuales, cada uno de los cuales es esencialmente un ordenador independiente con al menos un procesador, su propia memoria y un enlace de red que lo une con el resto de nodos. El término masivo supone cientos o miles de nodos que se comunican por paso de mensajes, usando estándares como *MPI (Message Passing Interface)*.

3.2. SMP: Symmetric Multiprocessing

Arquitectura multiprocesador donde dos o más procesadores idénticos se conectan a una memoria principal compartida. Están controlados por una única instancia de sistema operativo. Ej: PCs de hoy en día, donde cada chip es de doble o cuádruple núcleo.

4. Acceso a Memoria en Sistemas Multi-Procesador

4.1. NUMA: Non Uniform Memory Access

NUMA es un diseño de memoria donde es más lento acceder a unas regiones de memoria que a otras. Es debido al hecho de que algunas regiones de memoria están conectadas físicamente a distintos buses que otras. Este diseño de memoria fue diseñado para resolver los problemas de escalabilidad que aparecen en arquitecturas SMP. En SMP todos los accesos a memoria se realizan a través del mismo bus de memoria compartida. Esto funciona bien para un número relativamente pequeño de procesadores, pero cuando se tienen docenas o incluso cientos de CPUs compitiendo por el acceso al mismo bus de memoria compartida aparecen grandes problemas de rendimiento. Bajo NUMA un procesador puede acceder a su propia memoria local de forma más rápida que a la memoria no local (memoria local de otro procesador o memoria compartida entre procesadores).

4.2. UMA: Uniform Memory Access

En este diseño, la memoria física es uniformemente compartida por todos los procesadores, a través del mismo bus y con idéntico tiempo de acceso para acceder a cualquier posición de memoria.

5. Clusters

Un grupo de ordenadores enlazados, trabajando juntos, colaborando estrechamente, formando uno solo en muchos aspectos. Conectados por redes de área local rápidas. Su propósito es mejorar la disponibilidad y el rendimiento.

- Clústers de Alta Disponibilidad. Ejemplo Linux-HA.
- Clústers de Balanceo de Carga.

5.1. Propiedades

- Escalabilidad absoluta e incremental.
- Alta disponibilidad.
- Mejor relación precio/prestaciones que los SMP.
- Dificultad de gestión y configuración.

5.2. Propiedad Importante de las Tecnologías de Clustering

La propiedad de un sistema que oculta la naturaleza heterogénea y distribuida de los recursos, y los presenta a los usuarios y a las aplicaciones como un recurso computacional unificado y sencillo se denomina **SSI (Single System Image)**.

6. Grid Computing

La computación grid (o en malla) es una tecnología que permite utilizar de forma coordinada recursos heterogéneos (cómputo, almacenamiento y aplicaciones específicas) que no están sujetos a un control centralizado. Es una forma de computación distribuida, en la cual los nodos participantes pueden ser de iguales o distintas arquitecturas y cubrir toda la gama de potencia de cómputo.

Un proyecto conocido de grid computing es SETI@home (Search Extraterrestrial Intelligence at Home), en el cual los PC de usuarios distribuidos a lo largo y ancho del mundo donan los ciclos de sus procesadores no utilizados, empleando Internet como red de conexión. El proyecto SETI funciona en la plataforma informática BOINC (Berkeley Open Infrastructure for Network Computing).

6.1. Diferencias con clustering

- Débilmente acoplado.
- Heterogéneo.

- Geográficamente disperso.
- Los recursos NO se administran de forma centralizada.

6.2. Estándares y Protocolos relacionados con la Tecnología Grid

- **OGSA (Open Grid Services Architecture)**: describe una arquitectura orientada a servicio en entornos de Grid Computing para uso empresarial y científico.
- **OGSI (Open Grid Services Infrastructure)**: proporciona una infraestructura de niveles para la OGSA.
- **WSRF (Web Services Resource Framework)**: familia de especificaciones para servicios web. Fue anunciada en el 2004 como el sucesor de OGSI.

6.3. Globus Toolkit

Globus Toolkit es un conjunto de herramientas open source para construir grids computacionales. Proporciona implementaciones, entre otros, para los estándares OGSA, OGSI y WSRF. Está considerado como un estándar de facto para grid computing.

6.4. Niveles de la Arquitectura Grid

- **Fabric**: proporciona los recursos computacionales y de almacenamiento que son compartidos por los protocolos grid.
- **Connectivity**: define el núcleo de comunicaciones y los protocolos de acceso requeridos para las transacciones de red.
- **Resource**: basado en la conectividad y en los protocolos de autenticación, control el acceso a los recursos.
- **Collective**: gestiona servicios de directorio, servicios de scheduling, servicios de replicación de datos y servicios de monitorización y diagnóstico.
- **User Applications**: comprende las aplicaciones de usuario que operan dentro de la organización virtual.

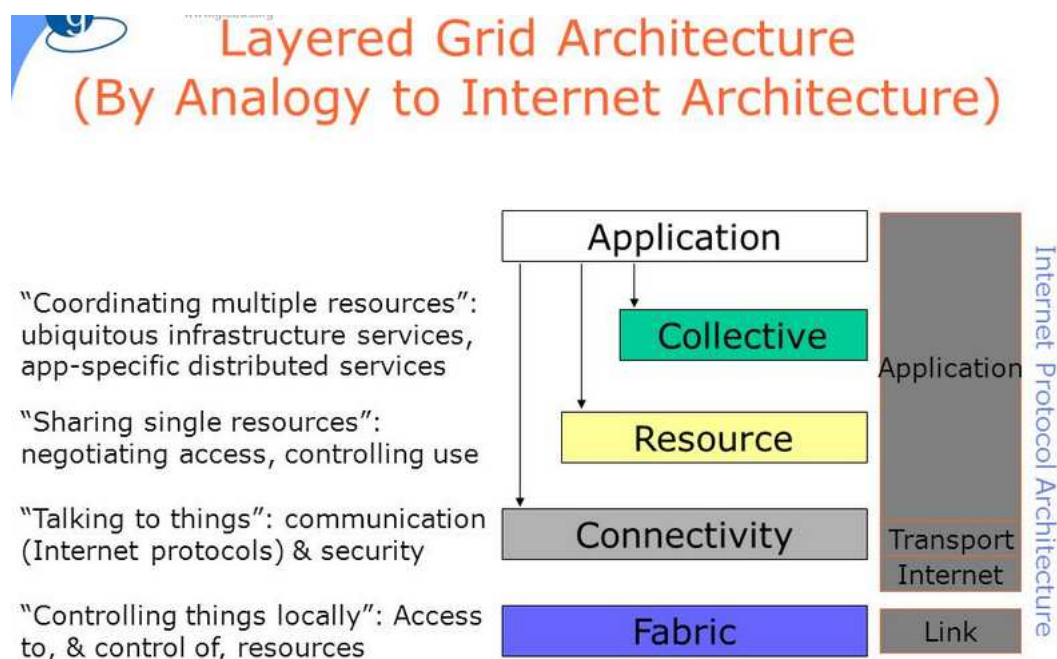


Figura 3: Niveles en la Arquitectura Grid vs Arquitectura TCP/IP.

6.5. Grid Semántica

Es un tipo de grid computing donde la información, recursos computacionales y servicios son descritos usando el modelo semántico de datos. En este modelo los datos y metadatos son expresados como hechos (pequeñas sentencias), directamente entendibles por los humanos. La especificación **S-OGSA** define el modelo, las capacidades y los mecanismos de una grid semántica.

7. Leyes Relacionadas con Arquitectura de Ordenadores

7.1. Ley de Amdahl

Se usa para averiguar la mejora máxima de un sistema cuando sólo una parte de éste es mejorado. Establece que: *“La mejora obtenida en el rendimiento de un sistema debido a la alteración de uno de sus componentes está limitada por la fracción de tiempo que se utiliza dicho componente”*.

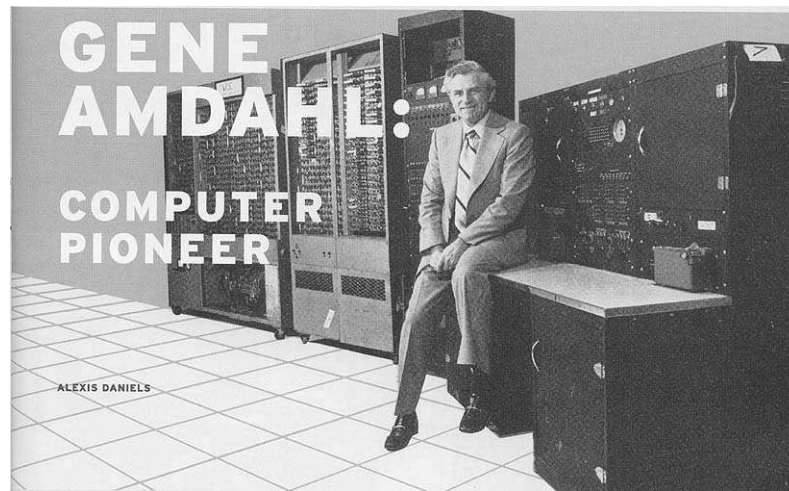


Figura 4: Mr. Gene Amdahl.

7.2. Ley de Gustafson

Es una ley en ciencia de la computación que establece que cualquier problema suficientemente grande puede ser eficientemente paralelizado.

8. Preguntas de Exámenes

1. (2016.21). **Grid Computing es un tipo de arquitectura distribuida y paralela, de ámbito:**
 - a) local, en la que se premia la paralelización, y a continuación la distribución.
 - b) local, en la que se premia la distribución, y a continuación la paralelización.
 - c) extenso geográficamente, en la que se premia la paralelización, y a continuación la distribución..
 - d) extenso geográficamente, en la que se premia la distribución, y a continuación la paralelización.
2. (2016.24). **La taxonomía según la cual las arquitecturas de los computadores se clasifican en SISD, MISD, SIMD y MIMD se debe a:**
 - a) Tanenbaum.
 - b) Flynn.
 - c) Turing.
 - d) Dijkstra.
3. (2016.25). **En las tecnologías de clustering, la propiedad de un sistema que oculta la naturaleza heterogénea y distribuida de los recursos, y los presenta a los usuarios y a las aplicaciones como un recurso computacional unificado y sencillo se denomina:**
 - a) VCS (Virtual Cluster Software).
 - b) SSI (Single System Image).
 - c) SA (System Availability).
 - d) AVS (Automatic Virtual System).
4. (2015.25). **La clasificación de Flynn divide a las computadoras en cuatro grupos principales. Indique el ERRÓNEO:**
 - a) SISD.
 - b) SIMD.
 - c) MISD.
 - d) SSMD.
5. (2015.26). **El uso de estándares y protocolos abiertos ha favorecido el desarrollo de la tecnología Grid, entre ellos NO se encuentra:**
 - a) OGSI.

- b)* OGSA.
 - c)* GNoMo.
 - d)* WSRF.
- 6. (2014.22). **La arquitectura GRID está compuesta por capas. La capa más baja y donde se encuentran los recursos computacionales y de almacenamiento que serán compartidos por los protocolos grid se denomina:**
 - a)* Fabric.
 - b)* Connectivity.
 - c)* Resource.
 - d)* Collective.
- 7. (2013.21). **En los sistemas multiprocesador, los sistemas MIMD se caracterizan por:**
 - a)* Varias unidades funcionales realizan diferentes operaciones sobre diferentes datos.
 - b)* Una única unidad funcional realiza varias operaciones sobre diferentes datos.
 - c)* Varias unidades funcionales realizan operaciones similares sobre datos diferentes.
 - d)* Una única unidad funcional realiza una única operación sobre diferentes datos.
- 8. (2013.22). **¿Qué ley establece que "la mejora obtenida en el rendimiento de un sistema al utilizar algún modo de ejecución más rápido está limitada por la fracción de tiempo que se pueda utilizar ese modo más rápido?:**
 - a)* Ley de Gustafson.
 - b)* Ley de Metcalfe.
 - c)* Ley de Amdahl.
 - d)* Ley de Machrone.
- 9. (2011.27). **Un sistema multiprocesador en el que todas las CPU comparten la memoria corresponde a un modelo del tipo:**
 - a)* SMP.
 - b)* MPP o SMP.
 - c)* Ninguno, puesto que en un sistema multiprocesador las CPU nunca comparten memoria.
 - d)* MPP.
- 10. (2011.29). **Las siglas SSI (Single System Image) corresponden a:**
 - a)* Una interfaz única de acceso al sistema para un clúster de servidores.

- b)* Una imagen del sistema operativo de un ordenador que se utiliza para la creación de clones.
 - c)* Una copia completa de un ordenador personal Windows que sirve para restaurar el sistema.
 - d)* Una representación de la memoria del sistema que se utiliza en .NET.
- 11. (2010.24). **¿Cuál de las siguientes afirmaciones en referencia a la arquitectura de computadores CISC (Complex Instruction Set Computer) es FALSA?:**
 - a)* Se caracteriza por tener un juego de instrucciones máquina muy amplio.
 - b)* Tiene muchos modos de direccionamiento.
 - c)* Tiene varios formatos de instrucción (de longitud variable).
 - d)* Su unidad de control es cableada.
- 12. (2010.25). **En relación con la tecnología grid computing, señale la respuesta FALSA:**
 - a)* Los nodos no tienen que estar dedicados.
 - b)* Son sistema homogéneos (todos los nodos deben ser idénticos).
 - c)* Se consigue reducción de costes, no es necesario disponer de grandes y costosos servidores.
 - d)* La escalabilidad parametrizable es una característica muy importante de esta arquitectura.
- 13. (2008.26). **Atendiendo a la clasificación de arquitecturas MIMD (Multiple Instruction stream, Multiple Data stream) [Johnson88], señale cuál de las siguientes se correspondería con una implementación del tipo DMMP (Distributed Memory / Message Passing):**
 - a)* Arquitecturas de memoria compartida o SMP.
 - b)* Arquitecturas de memoria virtual compartida.
 - c)* No existen implementaciones de esta tipología.
 - d)* Clusters.
- 14. (2008.28). **Señale la afirmación correcta. Se entiende por cluster al sistema que sirve para:**
 - a)* Aumentar la confidencialidad duplicando la clave de entrada.
 - b)* Aumentar la disponibilidad duplicando los servidores.
 - c)* Disminuir el riesgo de fallo eléctrico y/o de inundación del CPD.
 - d)* Aumentar el riesgo por concatenación de eventos.

15. (2008.31). Los “cluster” son hoy en día una alternativa a los multiprocesadores simétricos (SMP). Señale cuál de las siguientes NO es una característica generalmente asociada a los mismos:
- a) Escalabilidad absoluta e incremental.
 - b) Alta disponibilidad.
 - c) Mejor relación precio/prestaciones que los SMP.
 - d) Mayor facilidad de gestión y configuración.
16. (2016.24.LI). En relación con la computación GRID, señale la opción FALSA:
- a) S-OGSA extiende la especificación OGSA para profundizar en aspectos de la seguridad en la computación GRID.
 - b) Globus Toolkit es el estándar de facto para grid computing basado en software de código abierto.
 - c) Una implementación de OGSA (Open Grid Service Architecture) es el Globus Toolkit.
 - d) seti@home es un proyecto de computación distribuida que funciona en la plataforma BOINC (Barkeley Open Infraestructure for Netwotk Computing).
17. (2014.22.LI). Globus Toolkit es:
- a) Un framework de IBM que proporciona características para desarrollar, hacer debug y testear aplicaciones AJAX.
 - b) Un conjunto de herramientas open source para construir grids computacionales.
 - c) Un framework que contiene APIs y widgets para facilitar el desarrollo de aplicaciones Web que utilicen tecnología AJAX y programación paralela.
 - d) Una librería Python, modular y extensible, que ofrece una API para desarrollar aplicaciones distribuidas.
18. (2011.23.LI). La arquitectura de proceso paralelo en la cual cada procesador tiene su propia memoria local y además puede acceder a las memorias de otros procesadores se conoce como:
- a) SMP.
 - b) Grid.
 - c) NUMA.
 - d) Cluster.
19. (2011.25.LI). Si se quisiera un middleware para una grid, se elegiría:
- a) Globus Toolkit.

- b)* Cluster Veritas.
- c)* Arduino.
- d)* Data protector.

9. Soluciones

1. D
2. B
3. B
4. D
5. C
6. A
7. A
8. C
9. A
10. A
11. D
12. B
13. D
14. B
15. D
16. A
17. B
18. C
19. A