COMPUTACIÓN DE ALTA PERFORMANCE

Curso 2016

Sergio Nesmachnow (sergion@fing.edu.uy)
Santiago Iturriaga (siturria@fing.edu.uy)
Nestor Rocchetti (nrocchetti@fing.edu.uy)

Centro de Cálculo

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

TEMA 2
ARQUITECTURAS PARALELAS

CONTENIDO

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Arquitecturas secuenciales y paralelas
  - Clasificación de Flynn
  - Modelo SIMD
  - Modelo SISD
  - Modelo SIMD
  - Arquitectura MIMD
    - MIMD con memoria compartida
    - MIMD con memoria distribuida
- Factores que determinan la eficiencia
- Máquina paralela virtual
- Clusters
- Arquitecturas multinúcleo

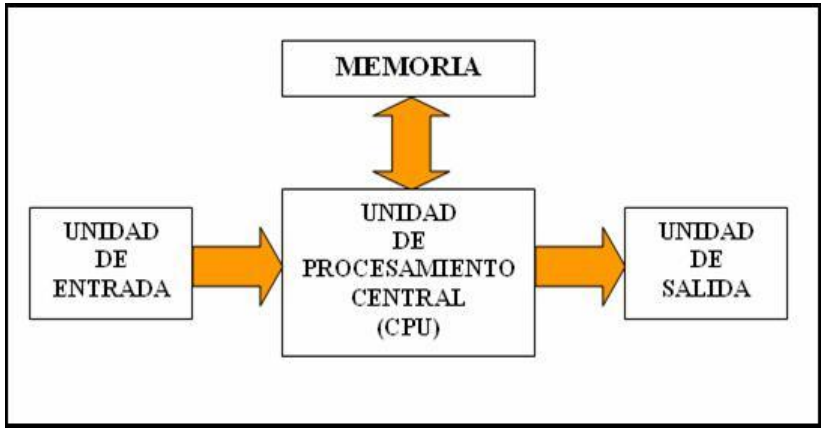
2.1: ARQUITECTURAS SECUENCIALES Y PARALELAS

ARQUITECTURAS PARALELAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Modelo estándar de computación:

  Arquitectura de Von Neumann
  - CPU única
    - Ejecuta un programa (único)
    - Accede a memoria
  - Memoria única
    - Operaciones read/write
  - Dispositivos



Arquitectura de Von Neumann

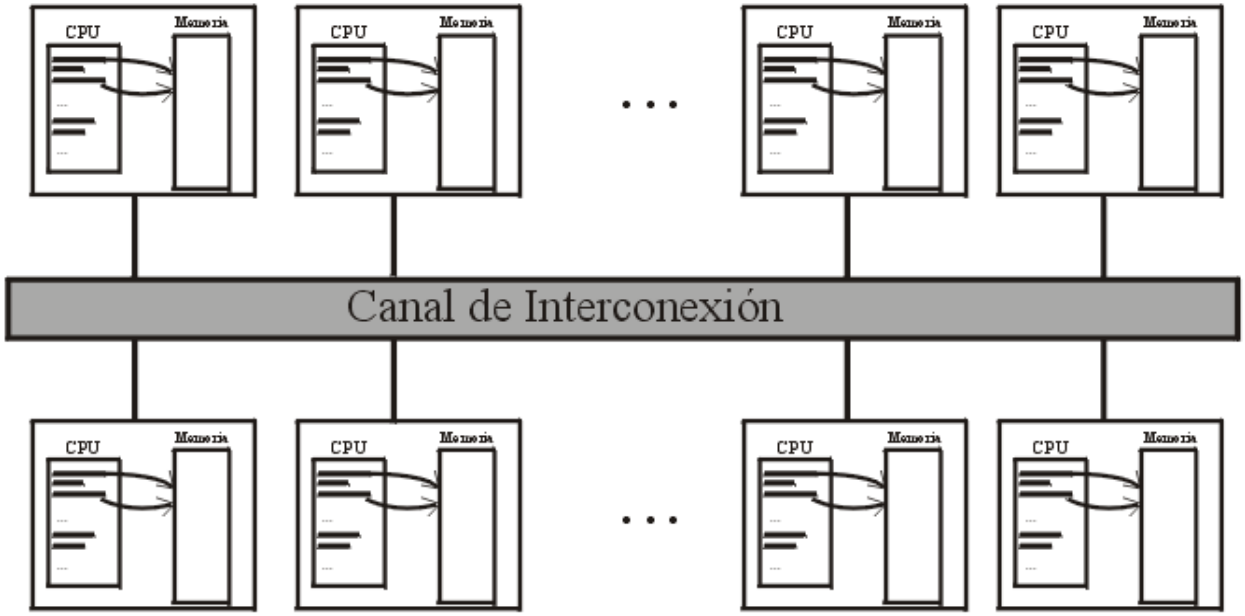
Neumann János

- Modelo robusto, independiza al programador de la arquitectura subyacente.
- Permitió el desarrollo de las técnicas de programación (estándar)

ARQUITECTURAS PARALELAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Extendiendo el modelo a la computación paralela, para lograr abstraer el hardware subyacente
- Existen varias alternativas, genéricamente contempladas en el modelo del multicomputador:
  - Varios nodos (CPUs de Von Neumann)
  - Un mecanismo de interconexión entre los nodos



Multicomputador (de memoria distribuida)

ARQUITECTURAS PARALELAS

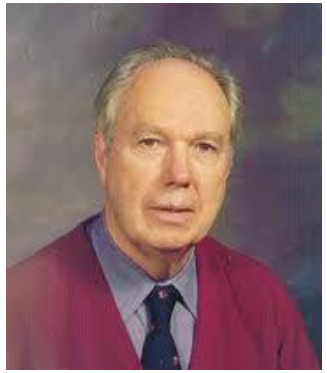
UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Extendiendo el modelo a la computación paralela ...
- Otras alternativas
  - Multiprocesador de memoria compartida
    - Nodos de Von Neumann
    - Memoria única
  - Computador masivamente paralelo
    - Muchísimos nodos (sencillas CPUs estilo Von Neumann)
    - Topología específica para interconexión entre los nodos
  - Cluster
    - Multiprocesador que utiliza una red LAN como mecanismo de interconexión entre sus nodos

CATEGORIZACIÓN DE FLYNN

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Clasificación de arquitecturas paralelas que considera la manera de aplicación de las instrucciones y el manejo de los datos



Michael Flynn

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

Taxonomía de Flynn (1966)

S=single, M=multi, I=Instrucción, D=Datos

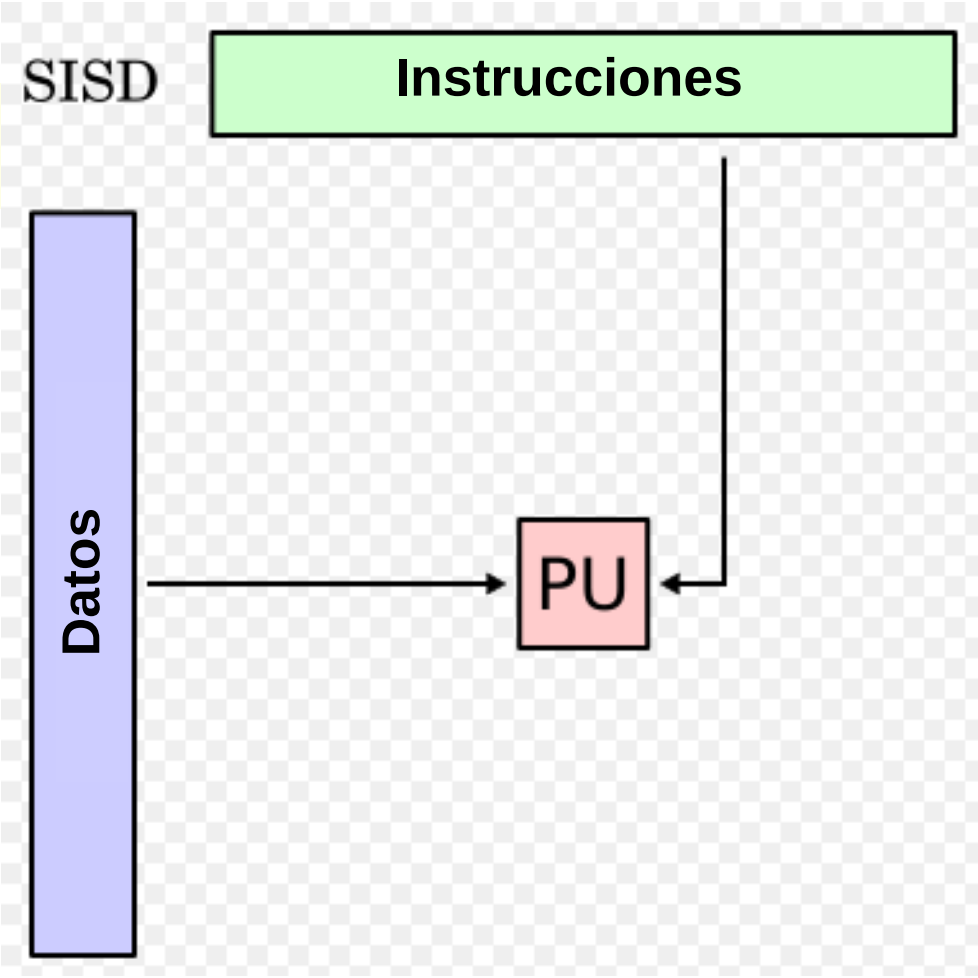
CATEGORIZACIÓN DE FLYNN

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

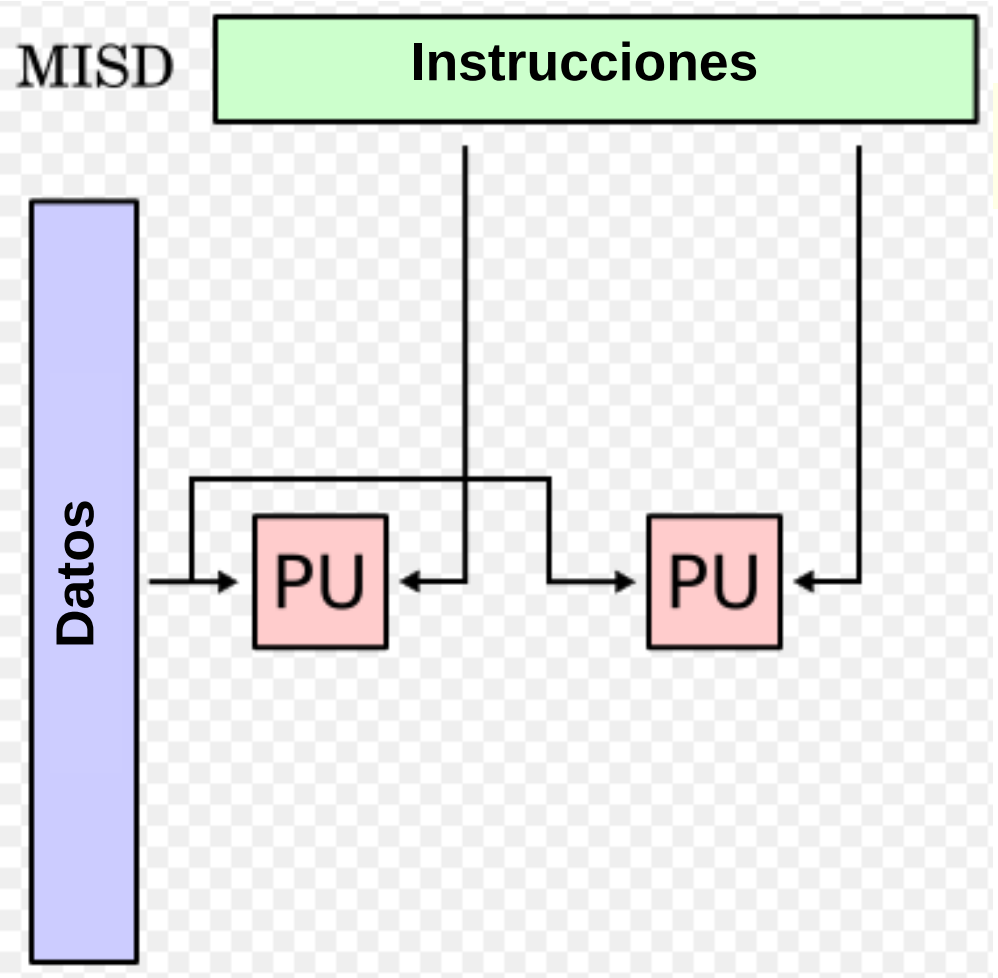
- SISD – Modelo convencional de Von Neumann
- SIMD – Paralelismo de datos, computación vectorial
- MISD – Pipelines, arrays sistólicos
- MIMD – Modelo general, varias implementaciones

- El curso se enfocará en el modelo MIMD, utilizando procesadores de propósito general o clusters de computadores
- El modelo SIMD se estudiará enfocado en el procesamiento de propósito general en procesadores gráficos

CATEGORIZACIÓN DE FLYNN

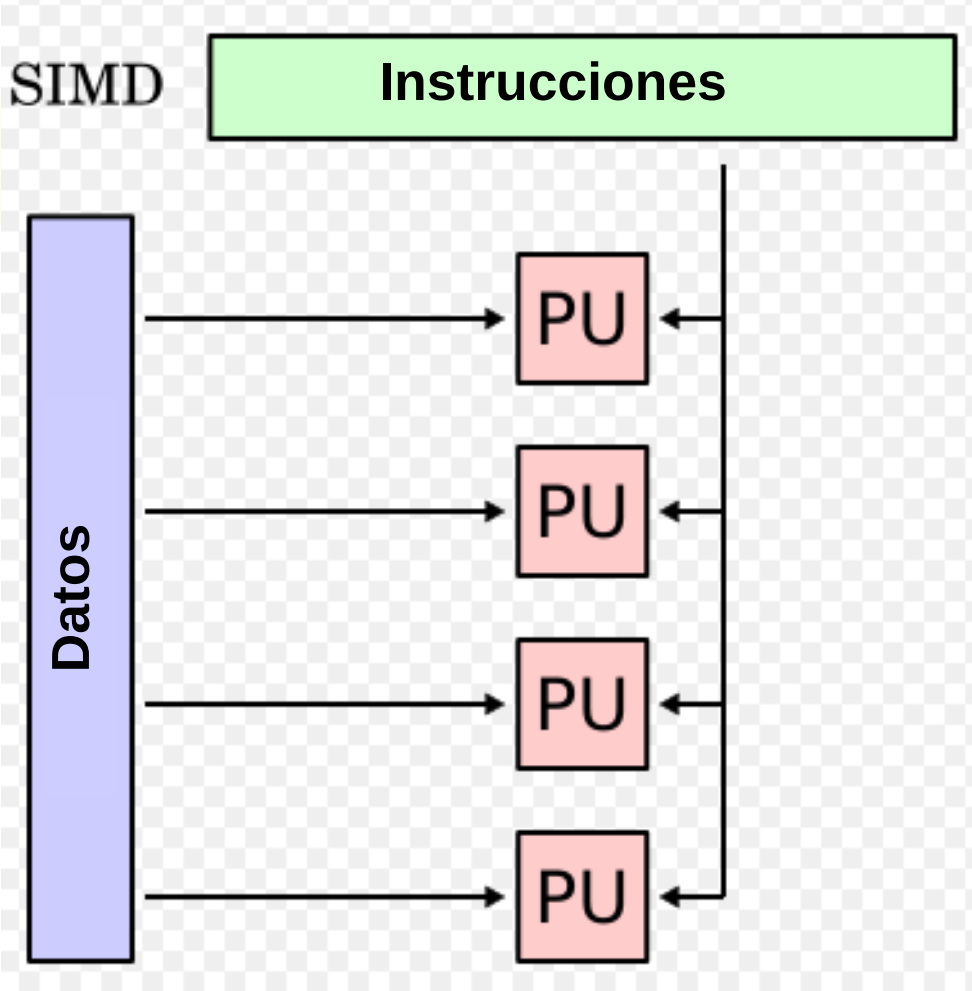


Single Instruction Single Data

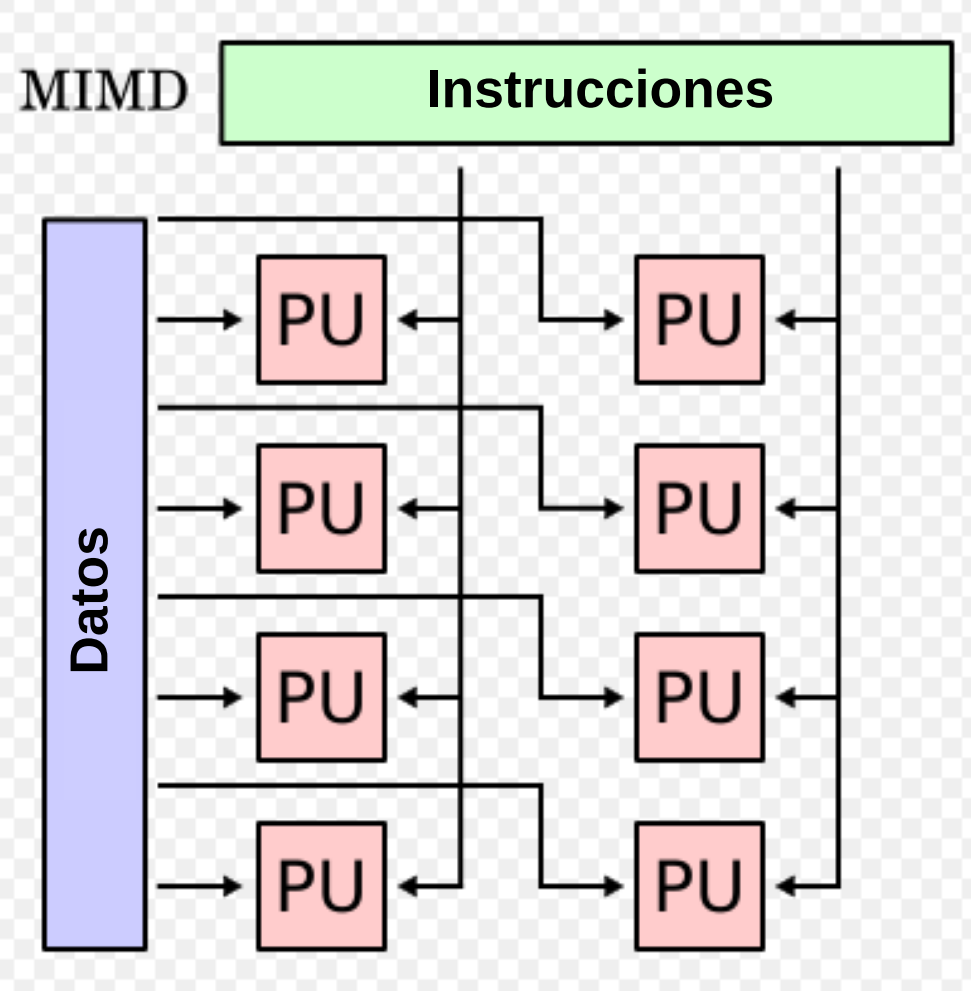


Multiple Instruction Single Data

CATEGORIZACIÓN DE FLYNN



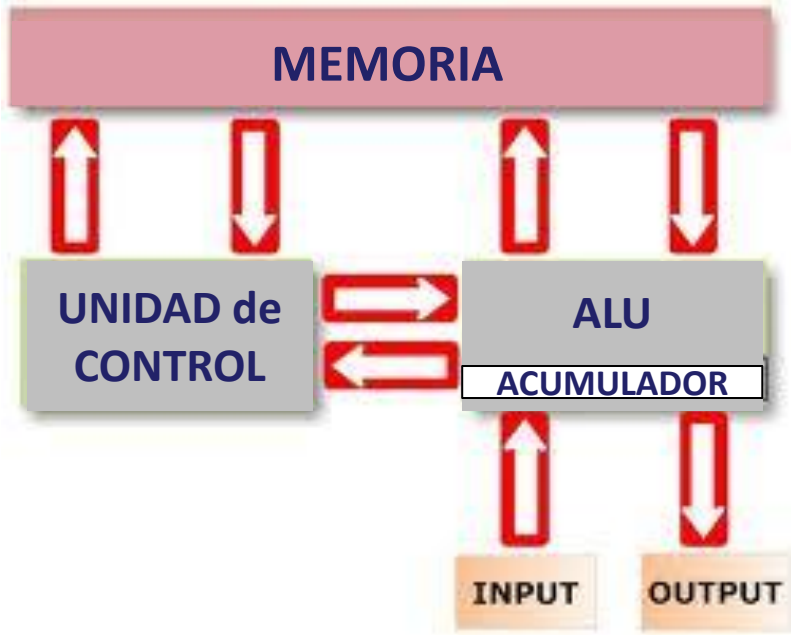
Single Instruction Multiple Data



Multiple Instruction Multiple Data

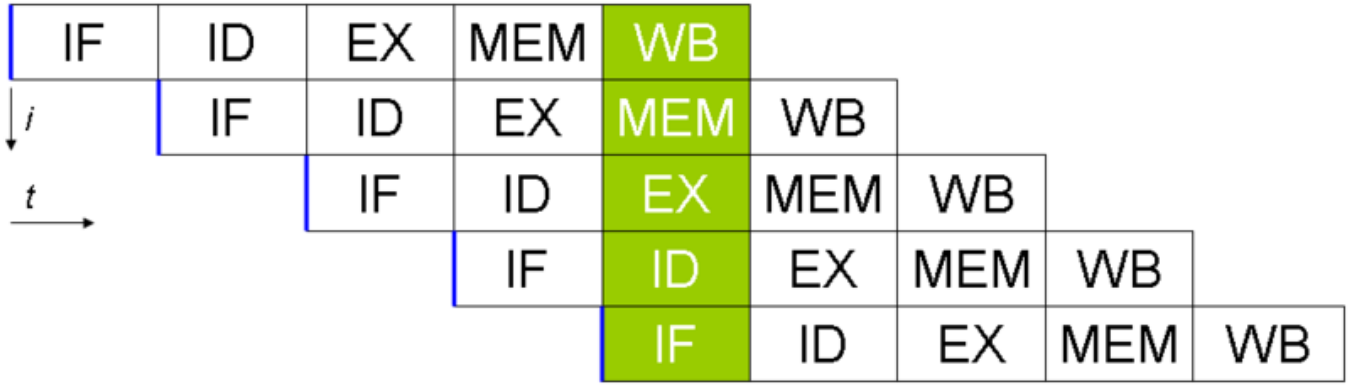
ARQUITECTURA SISD

- Máquina de Von Neumann
  - Un procesador capaz de realizar operaciones aritmético-lógicas secuencialmente, controlado por un programa que se encuentra almacenado en una memoria conectada al procesador
  - Este hardware está diseñado para dar soporte al procesamiento secuencial clásico, basado en el intercambio de datos entre memoria y registros del procesador, y la realización de operaciones aritmético-lógicas en ellos



ARQUITECTURA SISD

- En la década de 1980 se diseñaron computadores secuenciales que no correspondían estrictamente al modelo SISD
- A partir de la introducción de los procesadores RISC se comenzaron a utilizar varios conceptos de las arquitecturas paralelas, como el pipelining, la ejecución paralela de instrucciones no dependientes, el prefetching de los datos, etc., para lograr un incremento en la cantidad de operaciones realizadas por ciclo de reloj



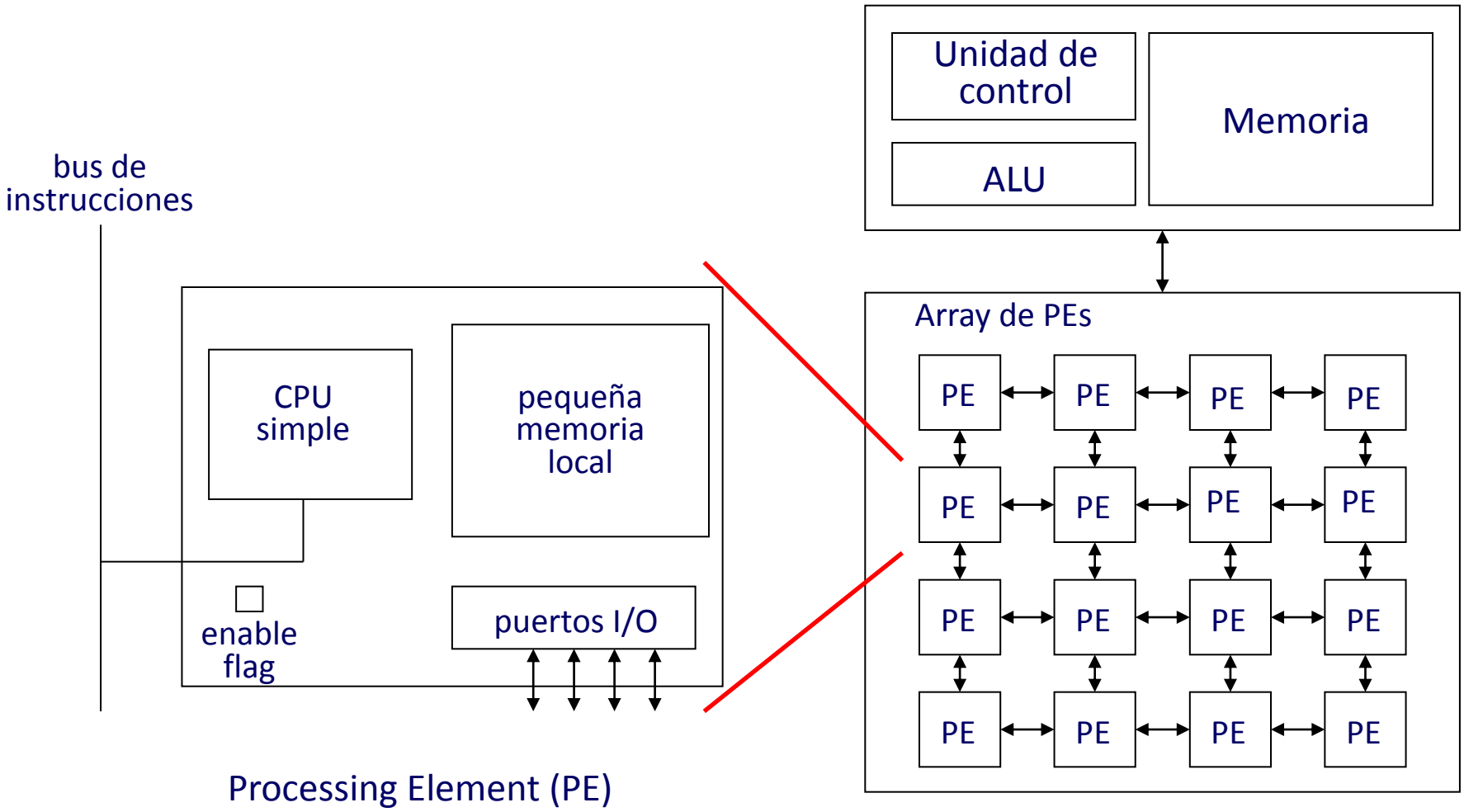
Pipeline

ARQUITECTURA SISD

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Aún mejorado, el modelo SISD no fue suficiente.
- Los problemas crecieron, o surgió la necesidad de resolver nuevos problemas de grandes dimensiones (manejando enormes volúmenes de datos, mejorando la precisión de las grillas, etc.).
- Si bien las maquinas SISD mejoraron su performance
  - Arquitecturas CISC y RISC.
  - Compiladores optimizadores de código.
  - Procesadores acelerando ciclos de relojes, etc.
- Aún no fue suficiente, y el ritmo de mejoramiento se desaceleró (principalmente debido a limitaciones físicas).
- En este contexto se desarrollaron los computadores paralelos.

ARQUITECTURA SIMD



ARQUITECTURA SIMD

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

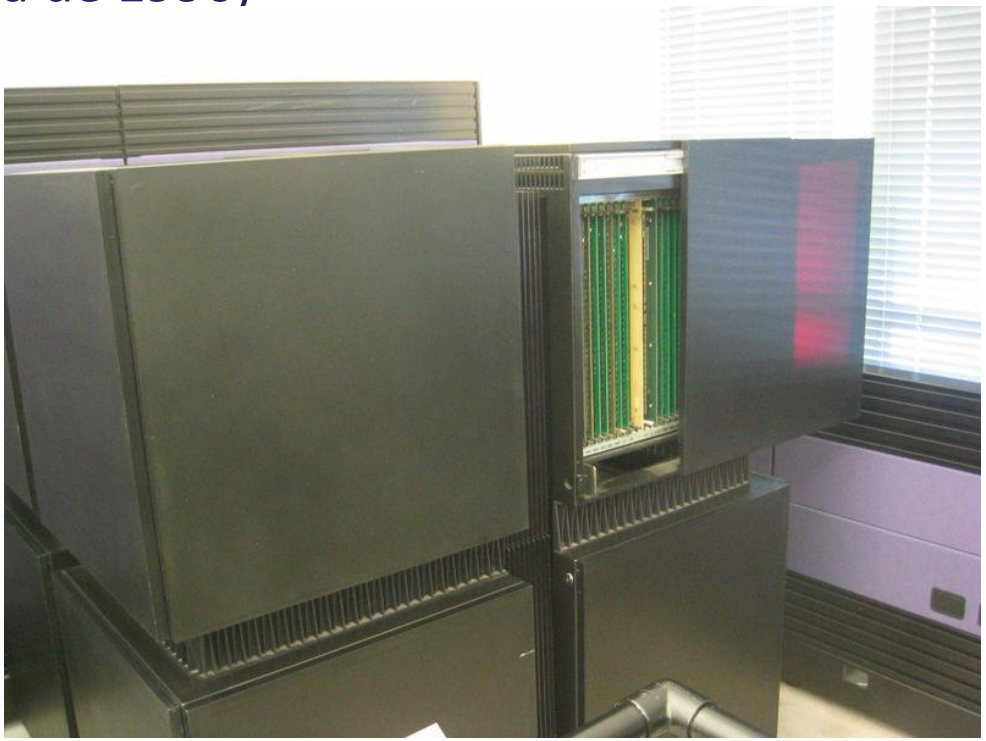
- Ún unico programa controla los procesadores.
- Útil en aplicaciones uniformes
  - Procesamiento de imágenes
  - Aplicaciones multimedia
  - Aplicaciones numéricas sobre grillas
- Su aplicabilidad está limitada por las comunicaciones entre procesadores.
  - La topología de interconexión es fija
- Los elementos de procesamiento tienen capacidad de cómputo limitada (1 bit a 8 bits), por lo que se pueden colocar una gran cantidad por chip (e.g. Connection Machine 2 con 64k PEs)
- Fueron los primeros multiprocesadores difundidos comercialmente (en la década de 1980)

ARQUITECTURA SIMD

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Ejemplos comerciales (históricos)
  - Cray X-MP (computador más potente entre 1983–1985)
  - Connection Machine (CM-2, CM-200, década de 1980)
  - MasPar MP2 (inicios de la década de 1990)

CRAY X-MP/24

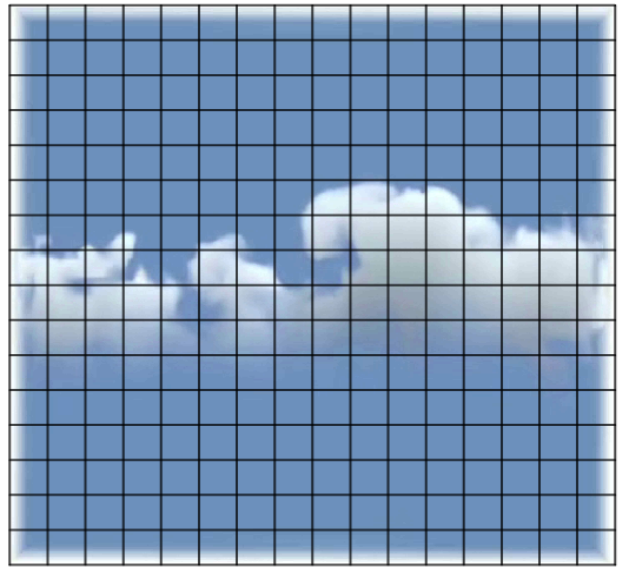


Connection Machine CM-2

GPU

- GPU = Graphics Processing Units
- Resurrección de la arquitectura SIMD
- Motivada por la industria de los juegos, televisión digital, etc.
- Tarjetas de video
  - Programables
    - Lenguaje CG, CUDA, OpenCL.
    - Interfases OpenGL, DirectX.
  - Paralelas
    - La misma “función” CUDA (CG, OpenCL) se aplica a muchos “pixels” (datos) al mismo tiempo.

GPU

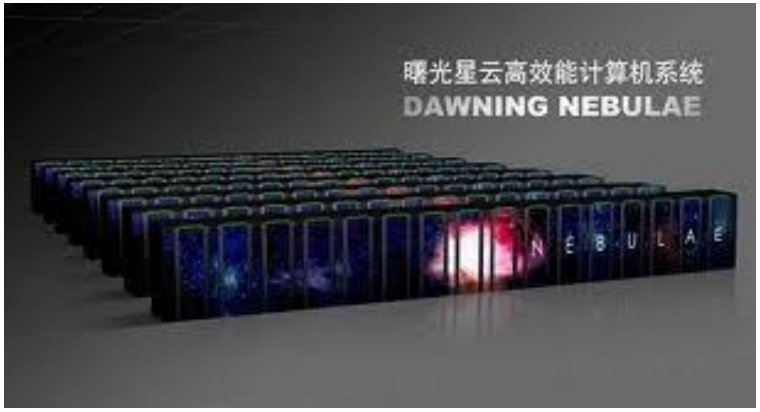


- La evolución de su performance no respeta la ley de Moore
  - Escalabilidad interna y externa simultáneamente
- La mayoría de las aplicaciones de imágenes son “trivialmente paralelas”
- Datos:
  - Una “imagen” es una “matriz”.
- Consecuencia:
  - Muy eficientes para calculo científico, en particular cuando involucran matrices

GPU

- La GPU puede verse como un multiprocesador de memoria compartida.
- Útil para computación de propósito general.

- Sitios de consulta:
  - General Purpose GPU (GPGPU)
    www.gpgpu.org
  - Parallel processing with CUDA
    http://www.nvidia.com/object/cuda_home.html



ARQUITECTURA MISD

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

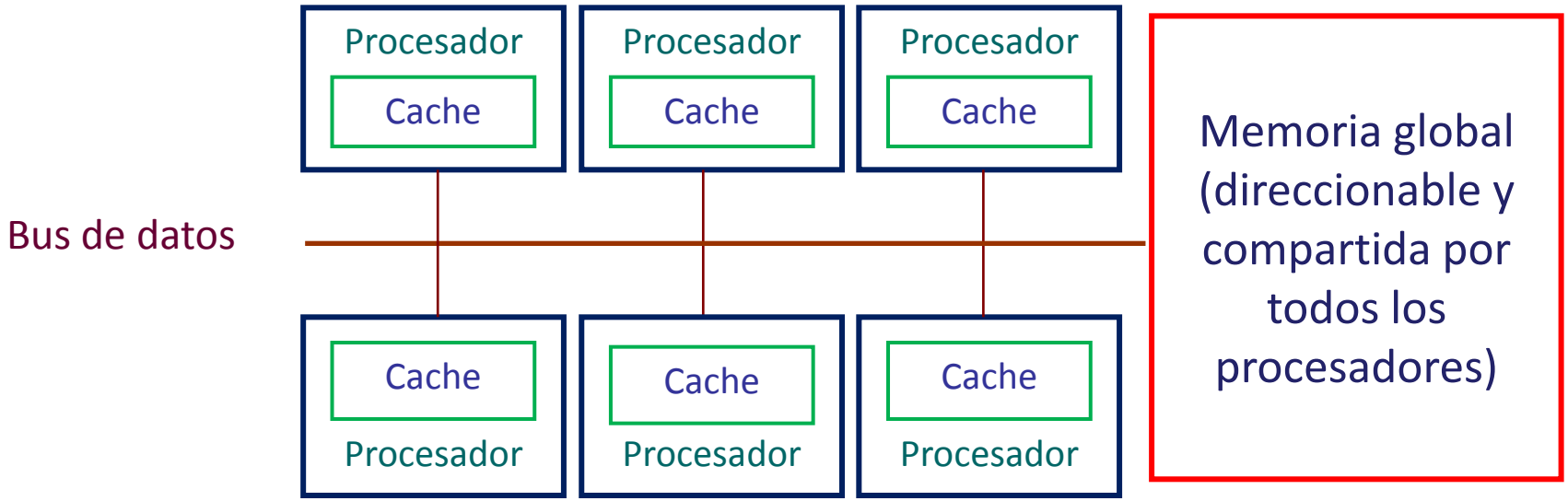
- Varias unidades funcionales ejecutan diferentes operaciones sobre el mismo conjunto de datos.
- Las arquitecturas de tipo pipeline pertenecen a esta clasificación
  - Aunque no puramente, ya que pueden modificar los datos sobre los que operan.
- Otros modelos: arrays sistólicos, FPGA celulares.
- También pertenecen al modelo MISD los computadores tolerantes a fallos que utilizan ejecución redundante para detectar y enmascarar errores.
- No existen otras implementaciones específicas.
- Los modelos MIMD y SIMD son más apropiados para la aplicación del paralelismo tanto a nivel de datos como de control.

ARQUITECTURA MIMD

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Consistieron en el “siguiente paso” en la evolución de las arquitecturas paralelas.
  - Fueron desplazando lentamente al modelo SIMD.
- A diferencia de los modelos SISD y MISD, las computadoras MIMD pueden trabajar asincrónicamente
  - Los procesadores tienen la capacidad de funcionamiento semi-autónomo.
- Existen dos tipos de computadores SIMD, de acuerdo al mecanismo utilizado para comunicación y sincronización:
  - MIMD de memoria compartida (fuertemente acopladas).
  - MIMD de memoria distribuída (poco acopladas).

ARQUITECTURA MIMD CON MEMORIA COMPARTIDA



Arquitectura MIMD con memoria compartida.

- Procesadores autónomos, trabajan asincrónicamente
- Comunicación entre procesadores a través del recurso compartido
  - Comunicación y sincronización se realiza en forma explícita
  - Emisor escribe y receptor lee de la memoria global

MIMD CON MEMORIA COMPARTIDA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Fáciles de construir
  - SO convencionales de los SISD son portables.
- Buena solución para procesamiento transaccional (sistemas multiusuario, bases de datos, etc.)
- Limitaciones: confiabilidad y escalabilidad
  - Un fallo de memoria de algún componente puede causar un fallo total del sistema.
- Incrementar el número de procesadores puede llevar a problemas en el acceso a memoria
  - Caso de supercomputadores Silicon Graphics
- El bus (cuello de botella) limita la escalabilidad a un máximo de pocas decenas de procesadores.
  - Se puede mejorar usando caches locales, pero se introduce el problema de "coherencia de cache"

MIMD CON MEMORIA COMPARTIDA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Mecanismos para compartir los datos
- UMA = Uniform Memory Access
  - Acceso uniforme (todos los procesadores tienen el mismo tiempo de acceso a la memoria)
  - Multiprocesadores simétricos (SMP)
  - Pocos procesadores (32, 64, 128, por problemas de ancho de banda del canal de acceso)
- NUMA = Non-Uniform Memory Access.
  - Colección de memorias separadas que forman un espacio de memoria direccionable
  - Algunos accesos a memoria son más rápidos que otros, como consecuencia de la disposición física de las memorias (distribuidas físicamente)
  - Multiprocesadores masivamente paralelos (MPP)

MIMD CON MEMORIA COMPARTIDA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Ejemplos:
  - Encore MULTIMAX
    - Inicios de la década de 1990, hasta 20 procesadores
  - Sequent Symmetry
    - Década de 1990, de 10 a 30 procesadores
  - Sun SPARCcenter 2000
    - Escalable hasta 20 procesadores
    - Sun-4d (d por *dragon*, nombre código del SPARCcenter 2000)
    - La Facultad de Ingeniería tuvo un supercomputador de este tipo desde 2000 a 2008

SUN SPARC CENTER 2000

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Arquitectura SPARC (Scalable Processor ARChitecture), versión 8
  - Arquitectura RISC
  - Microprocesadres SuperSPARC II 85 Mhz
  - Direccionamiento 32 bits
  - Arquitectura de bus multinivel
- Clasificación:
  - MIMD de memoria compartida
- Sistema altamente acoplado (tighthly coupled system)
- Diseño Modular
  - Unidad Procesador
  - Unidad Memoria
  - Unidad Entrada/Salida

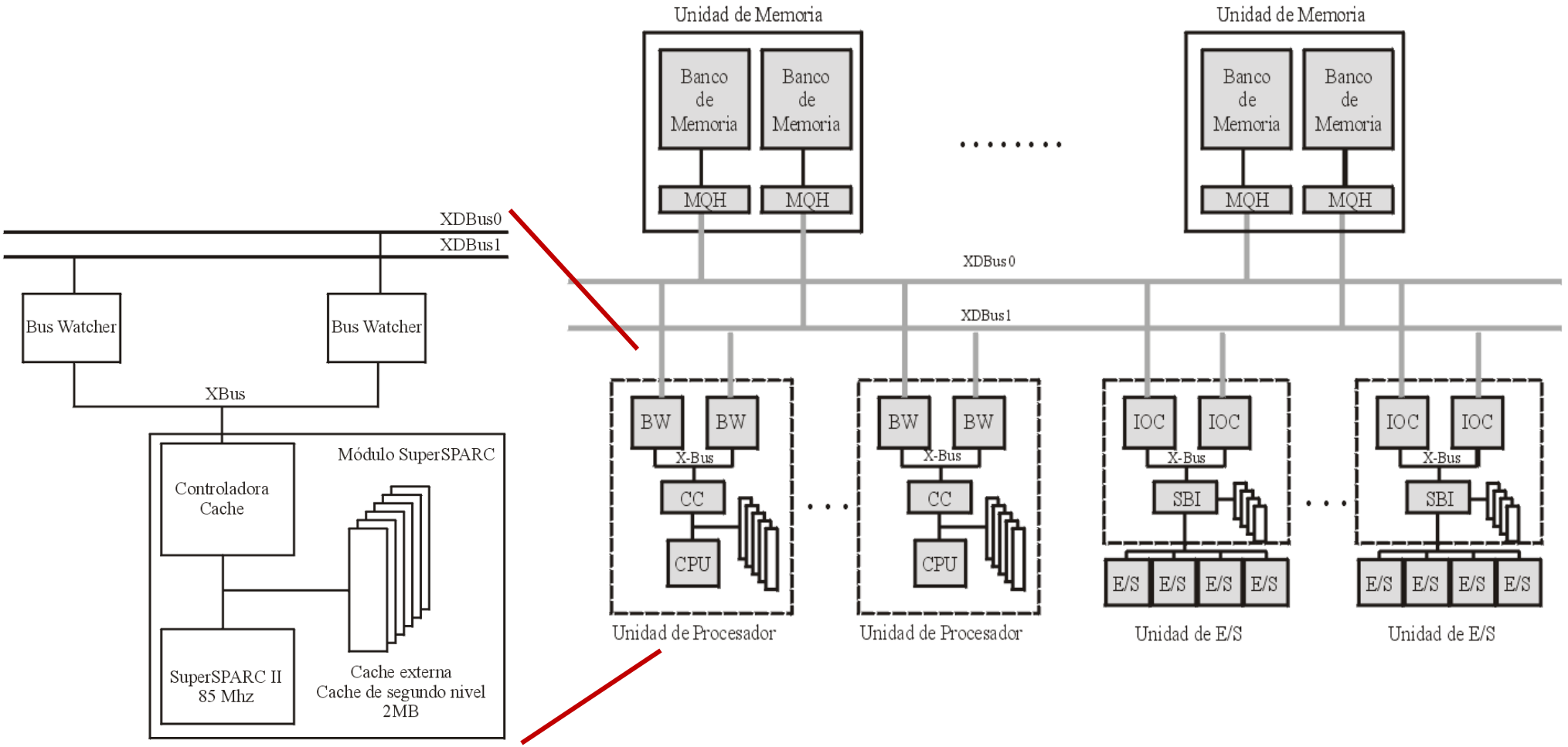
SUN SPARC CENTER 2000

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Componentes:
  - Placa de control
  - 10 placas de sistema
  - 20 procesadores SuperSPARC II (85 Mhz), 2 por placa de sistema
  - 5 GB de memoria RAM, 512 MB por placa de sistema
  - 40 dispositivos de entrada/salida, 4 por placa de sistema
  - 2 MB Level 2 cache por CPU
  - Bus Multinivel:
    - XDBUS, conecta unidades lógicas, canal de 72 bits (64 datos + 8 paridad), 400 Mb/s, packet-switched, sincrónico, coherencia de caché
    - XBUS, conecta BW y CC, similar al XDBUS
    - SBUS, conecta placa y dispositivos de entrada/salida

SUN SPARC CENTER 2000

- Esquema de arquitectura



MIMD con MEMORIA DISTRIBUIDA

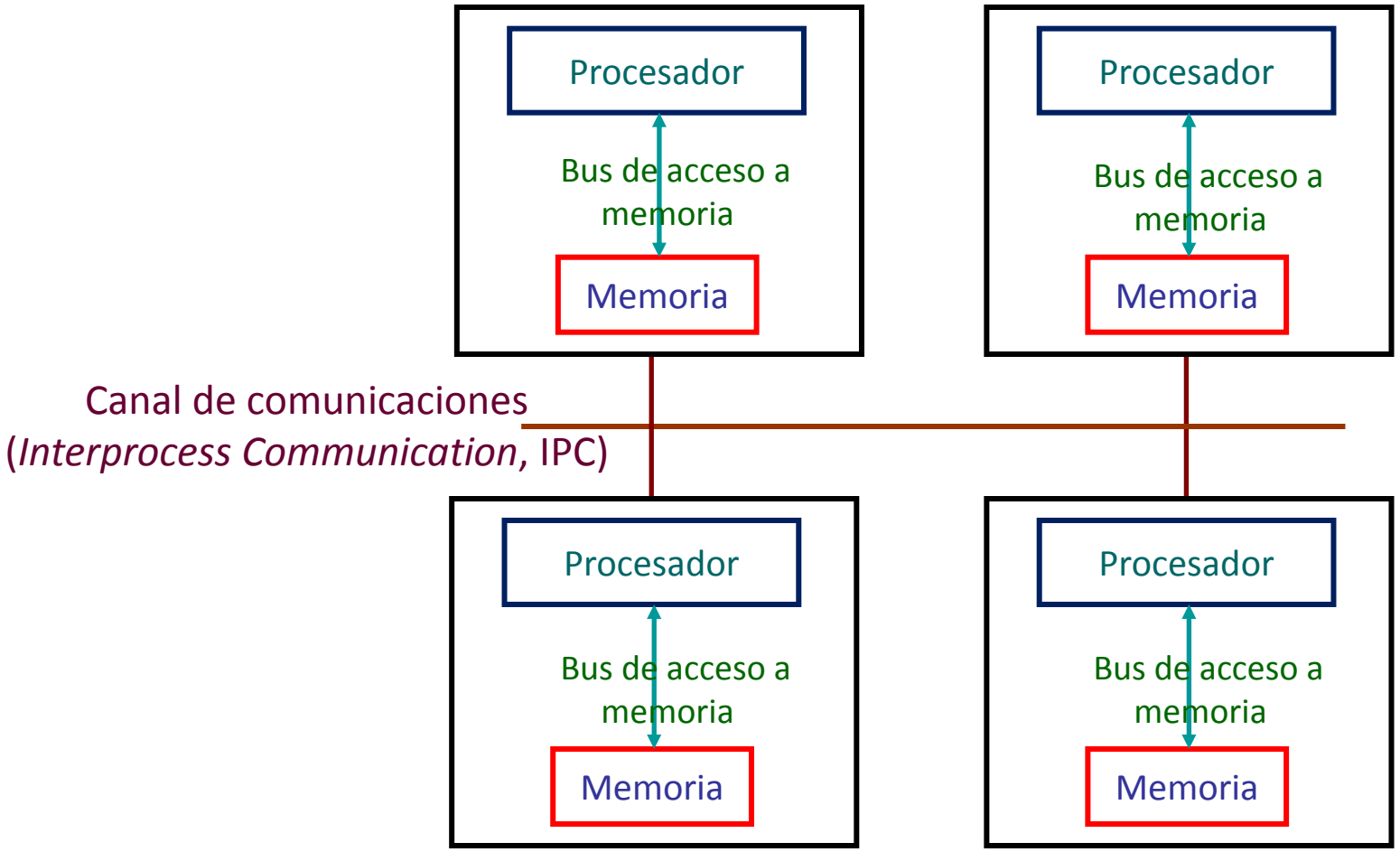
UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- No existe el concepto de memoria global
- Comunicación y sincronización:
  - Mecanismos explícitos de IPC (pasaje de mensajes) sobre una red (en escenario óptimo, red de alta velocidad)
  - Tienen mayor costo que en MIMD de memoria compartida
- Las comunicaciones pueden ser cuellos de botella.

- Arquitectura escalable para aplicaciones apropiadas:
  - Decenas de miles de procesadores
  - Popularizadas en clusters

- Ventajas respecto a MIMD de memoria compartida
  - Fácilmente escalable
  - Alta disponibilidad: el fallo de una CPU individual no afecta a todo el sistema

ARQUITECTURA MIMD CON MEMORIA DISTRIBUIDA



Arquitectura MIMD con memoria distribuida

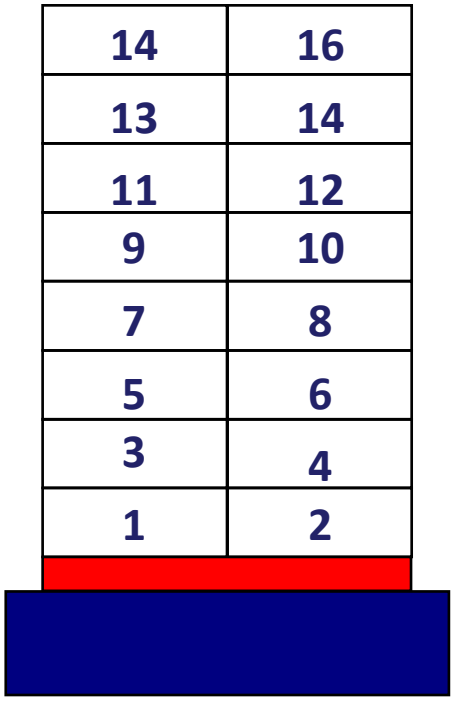
MIMD CON MEMORIA DISTRIBUIDA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Ejemplos comerciales
  - Connection Machine CM-5 (1991, 16k procesadores).
  - Intel Paragon (1992: 2048 procesadores, luego 4000).
  - Cray.
    - Luego de fusión con SGI: Scalable Node SN1, SN2.
    - T3E, 1997, hasta 2048 procesadores.
  - IBM SP (IBM Scalable POWER parallel)
    - Incluía tecnología High Performance Switch (HPS) para comunicación entre nodos.
    - Década de 1990 e inicios de los 2000
    - En 1999 incorporan procesadores POWER3, en 2001 POWER4 y en 2004 POWER5.

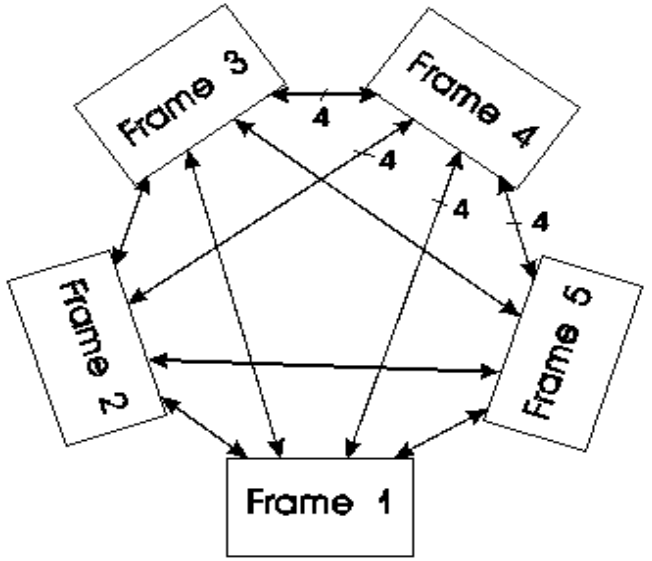
IBM SP-2

- Uno o más SP (Scalable POWER) frames.
  - SP: Basado en RS/6000
  - 2 a 16 máquinas por frame
  - High Performance Switch



HIGH PERFORMANCE SWITCH

- Conexión bi-direccional todos con todos.
  - Packet-switched network (versus circuit-switched).
- Ancho de banda: 150 MB.
- Tiempo de latencia: 500 ns.
- Instalaciones en Uruguay:
  - BPS, DGI, UTE, IBM.



SMP y MPP

- SMP: Multiprocesamiento simétrico
  - Paralelismo sobre memoria compartida
  - Modelo SIMD y modelo MIMD UMA
  - Primera implementación en 1961
  - Dominante hasta mediados de la década de 1990
  - En 2006, aparecen los PCs dual-core

- MPP: procesamiento paralelo masivo
  - Sistema con muchas (decenas, cientos) unidades de procesamiento (ALUs, procesadores)
  - Unidades integradas en una única arquitectura
  - Dominaron el espectro de HPC hasta los inicios del 2000

2.2: FACTORES QUE DETERMINAN LA EFICIENCIA

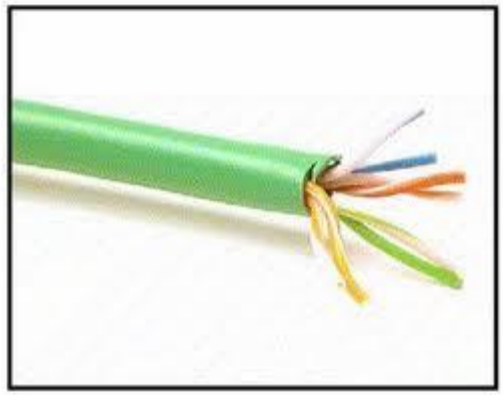
CONECTIVIDAD ENTRE PROCESADORES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

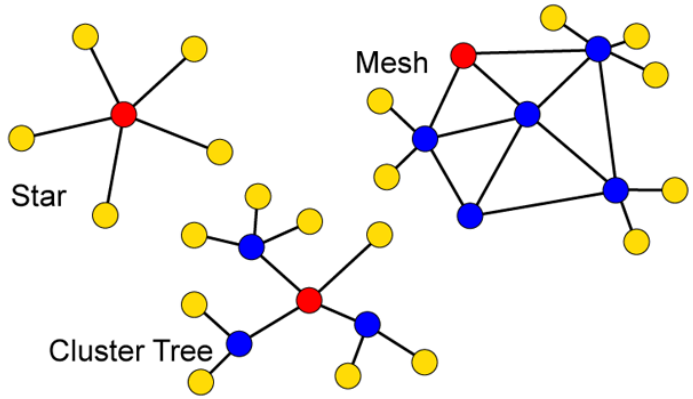
- ESTÁTICA
  - Caminos prefijados entre procesadores.
  - Usualmente canales de comunicación P2P entre procesadores.

- DINÁMICA
  - Los caminos se determinan dinámicamente.
  - Se implementa a través de switches.
    - Simplifica la programación al evitar problemas de comunicación.
    - Garantiza igualdad de latencia para comunicaciones entre distintos procesadores a una distancia fija.
    - Permite comunicaciones “all to all”.
    - Generan topologías fácilmente escalables.

FACTORES QUE DETERMINAN LA EFICIENCIA



- Ancho de banda
  - Número de bits capaces de transmitirse por unidad de tiempo
- Latencia de la red
  - Tiempo que toma a un mensaje transmitirse a través de la red
- Latencia de las comunicaciones
  - Incluye tiempos de trabajo del software y retardo de la interfaz
- Latencia del mensaje
  - Tiempo que toma enviar un mensaje de longitud cero
- Valencia de un nodo
  - Número de canales convergentes a un nodo
- Diámetro de la red
  - Número mínimo de saltos entre los nodos más alejados
  - Permite calcular el peor caso de retardo de un mensaje
- Largo máximo de un tramo de comunicación.



FACTORES QUE DETERMINAN LA EFICIENCIA

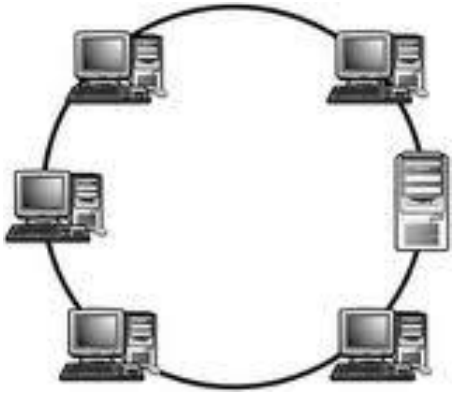
- Ancho de bisección
  - Número mínimo de enlaces que en caso de no existir la red se separaría en dos componentes conexas
- Costo
  - Cantidad de enlaces de comunicación

- CONFIGURACIÓN ÓPTIMA
  - Ancho de banda grande
  - Latencias (de red, comunicación y mensaje) bajas
  - Diámetro de la red reducido
  - Ancho de bisección grande
  - Valencia constante e independiente del tamaño de la red
  - Largo máximo de tramo reducido, constante e independiente del tamaño de la red
  - Costo mínimo

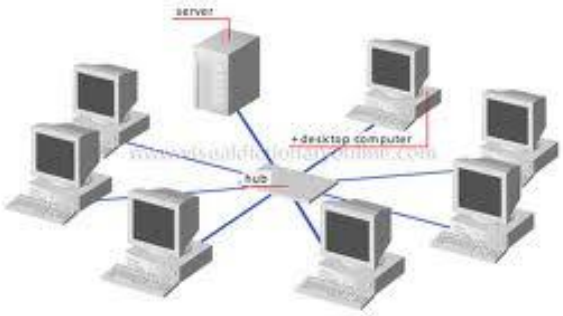
CONECTIVIDAD ENTRE PROCESADORES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

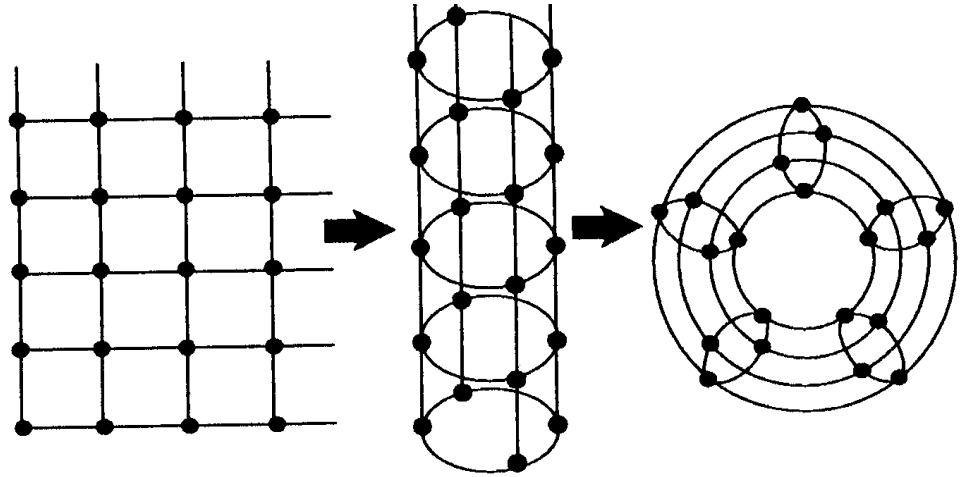
Modelos



Anillo



Estrella

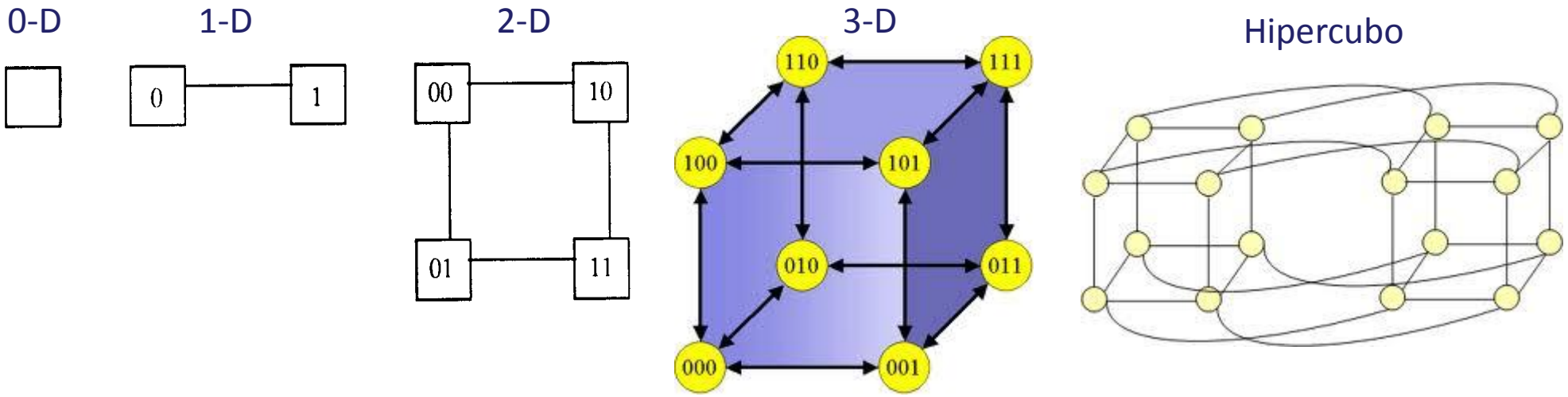


Grilla (2D) Cilindro Toro

Topologías geométricas (mallas)

CONECTIVIDAD ENTRE PROCESADORES

- Modelos de conectividad

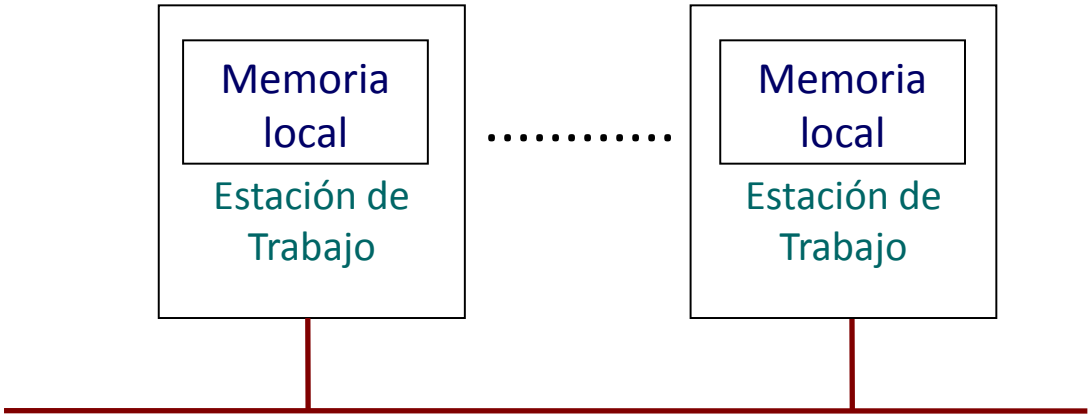


Topologías “dimensionales” (distancia constante)

MIMD CON MEMORIA DISTRIBUIDA

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

CASO PARTICULAR (1)



MÁQUINA PARALELA VIRTUAL

Red de datos
(LAN, WAN, Ethernet, FastEthernet, Gigabit Ethernet, FDDI, etc).

VENTAJAS

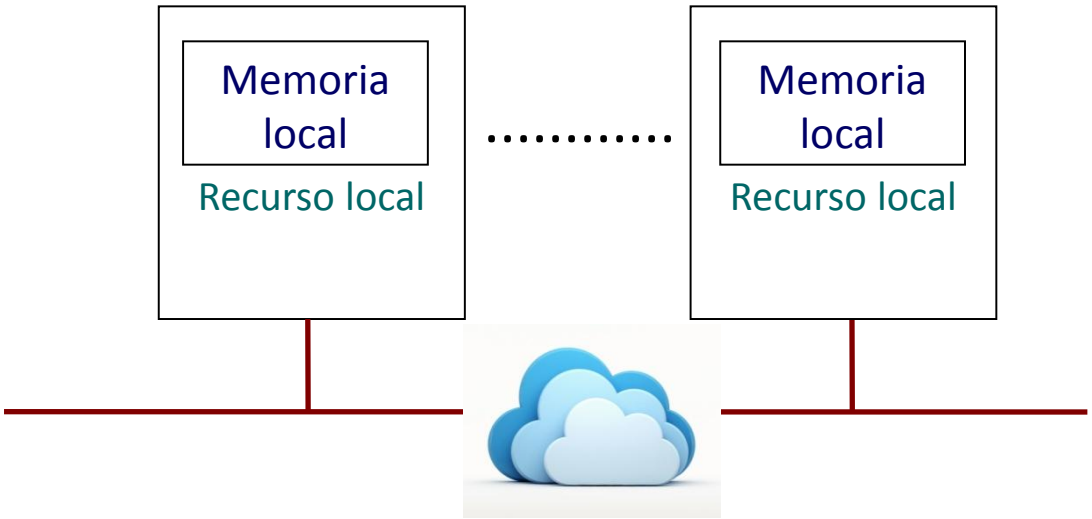
- Usa infraestructura existente
- Sistema escalable
- Fácilmente programable

DESVENTAJAS

- Grandes latencias en las comunicaciones
- Disponibilidad, seguridad

MIMD CON MEMORIA DISTRIBUIDA

CASO PARTICULAR (2)



GRIDS y CLOUDS computacionales

VENTAJAS

- Permite agregación e integración
- Sistema totalmente escalable
- Existen mecanismos de programación

DESVENTAJAS

- Enormes latencias en las comunicaciones

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

2.3: CLUSTERS

Basado en el artículo “Cluster computing at a glance”.

M. Baker, R. Buyya

¿CÓMO MEJORAR EL DESEMPEÑO?

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Básicamente, hay 3 maneras:
  1. Trabajar más intensamente
  2. Trabajar más inteligentemente
  3. Solicitar ayuda
- Analogía para la computación
  1. Utilizar hardware especial, por ejemplo reducir el tiempo por instrucción con un procesador de mayor ciclo de reloj
  2. Optimizar algoritmos y técnicas de programación
  3. Utilizar múltiples recursos de cómputo para resolver el problema, ejecutando más instrucciones en el mismo tiempo

SECUENCIAL vs PARALELO

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Limitaciones de las arquitecturas secuenciales
  - Se alcanzaron las limitaciones físicas (velocidad de la luz, termodinámicas, cuánticas)
  - Mejoras de hardware como el pipelining, el procesador superscalar, etc., no son escalables y requieren una complicada tecnología de fabricación e instrumentación
  - El procesamiento vectorial sólo funciona adecuadamente para cierta clase de problemas
- Procesamiento paralelo en la década de 2000
  - La tecnología del procesamiento paralelo maduró y fue explotada comercialmente
  - Existía un amplio trabajo de investigación y desarrollo en herramientas y entornos de programación
  - El desarrollo significativo de la tecnología de redes permitió el avance de la computación heterogénea

CLUSTERS: MOTIVACIÓN

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Analogía biológica:
  - Procesamiento paralelo en estructuras cerebrales
  - La “velocidad global” con la cual las millones de neuronas del cerebro humano resuelven problemas muy complejos es asombrosa, aún cuando individualmente, el tiempo de respuesta de una neurona es lento (del orden de ms)

  - Este argumento sugiere la potencial utilidad de utilizar múltiples recursos de cómputo funcionando de modo coordinado para resolver un problema global

- Motivación para utilizar clusters:
  - Permiten reutilizar equipamiento disponible (no dedicado)
  - El ancho de banda para comunicaciones entre workstations ha crecido al desarrollarse nuevas tecnologías y protocolos e implementarse en LANs and WANs
  - Los clusters de workstations son más sencillos de integrar en los entornos de desarrollo y producción que las supercomputadoras

RESOLUCIÓN DE PROBLEMAS COMPLEJOS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- ¿Se necesitan más recursos de cómputo para resolver problemas complejos?
- Considerando:
  - La gran cantidad de equipos subutilizados.
  - La enorme cantidad de ciclos de procesador libres, a los cuales puede darse un uso práctico.
  - El costo desmedido de un supercomputador.
  - Los recursos de cómputo distribuídos encajan con el modelo de clusters.
- ¿Se necesitan más recursos de cómputo para resolver problemas complejos? ……

SI

- Pero no siempre es necesario HARDWARE ESPECIAL
- Es posible implementar soluciones con el equipamiento disponible (de oficina, empresa, centro educativo).

MOTIVACIÓN PARA USAR CLUSTERS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- La brecha entre el poder de cómputo de clusters y supercomputadoras se redujo considerablemente.
  - La performance de workstations y PCs mejoró rápidamente.
- Estudios muestran que el uso de ciclos de CPU promedio de las estaciones de trabajo es típicamente menor al 10% de su capacidad.
  - Como la performance mejora, el uso (porcentual) de CPU decrece aún más.
- Se hace cada vez más difícil justificar la inversión importante para adquirir un supercomputador y sus herramientas de desarrollo.

- Las herramientas de desarrollo para PCS y workstations están exhaustivamente analizadas y probadas (inclusive algunas estandarizadas).
- Esta situación contrasta con las soluciones propietarias de las supercomputadoras, muchas de ellas no estandarizadas.
- Los clusters de workstations son escalables. Por relativamente poco costo adicional es posible agregar nuevos recursos de cómputo.

USOS DE CPU

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

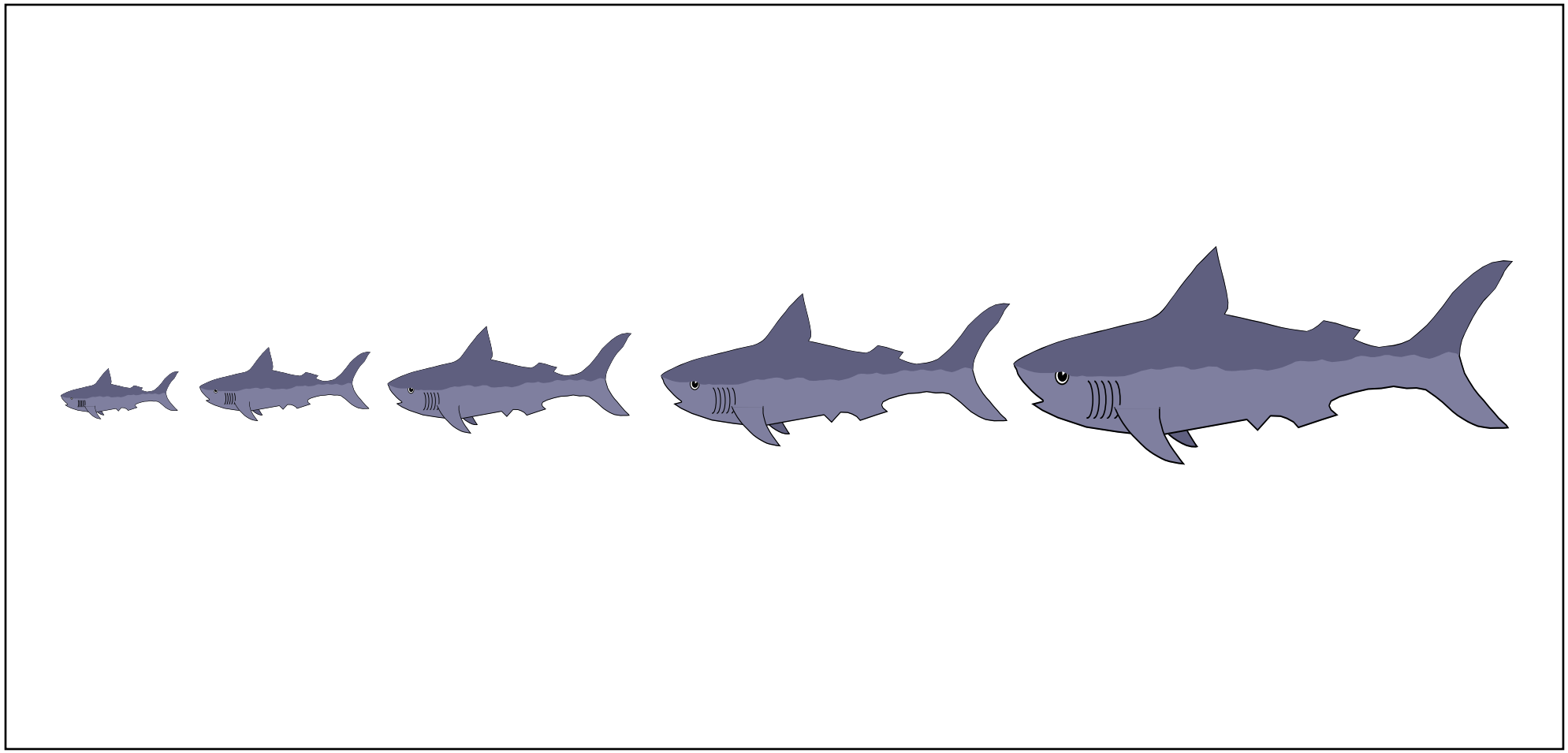
- Usualmente una workstation será propiedad de un indivuduo, grupo, departamento u organización, en el sentido de que estarán DEDICADAS al uso de su(s) propietario(s).
- ¿Cómo implementar un “cluster de workstations” para ejecución de aplicaciones paralelas y distribuídas?
- Típicamente, hay varios tipos de “propietarios”, de acuerdo al uso que dan a “su” CPU.
  1. Estudiantes o trabajadores ocasionales.
  2. Aplicaciones “de oficina” (email, documentos).
  3. Desarrollo de software (edición, compilación, test).
  4. Ejecución de aplicaciones que requieren cómputo intensivo.

EL “ROBO DE CICLOS” DE CPU

- Las técnicas de computación en clusters no dedicados tratan de “robar” ciclos de CPU a los propietarios que realizan tareas “livianas” [(1), (2) y (3)] de manera de proveer a los que ejecutan tareas “pesadas” [(4)].

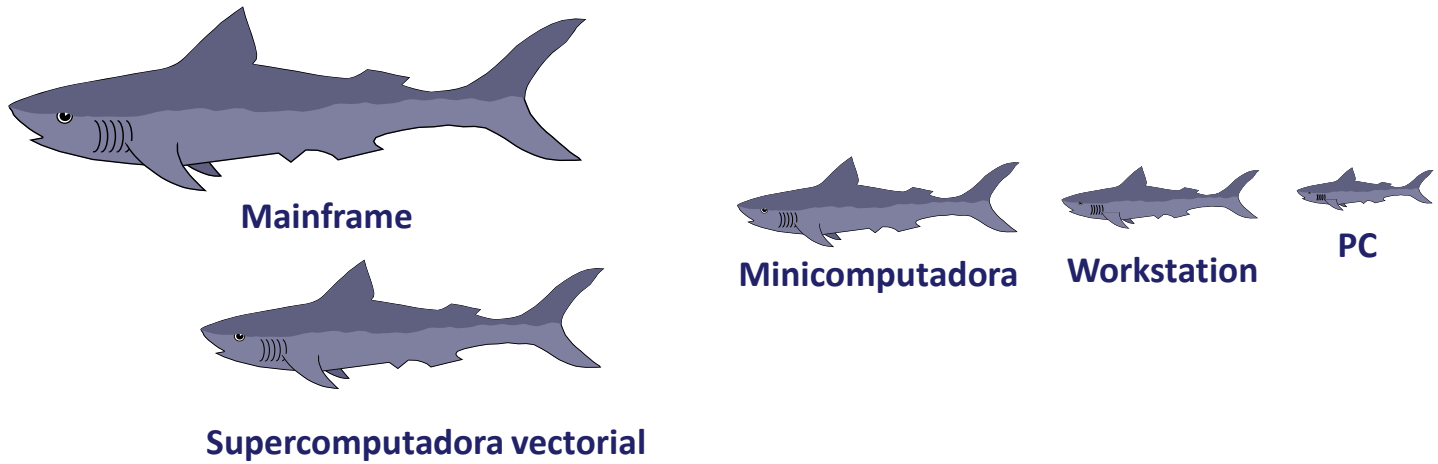
- Sin embargo, esta práctica requiere superar los egoísmos personales (los usuarios son muy protectores de “sus” equipos).
- Usualmente se requiere de autorización institucional para utilizar las computadoras de esta manera.
- El robo de ciclos de CPU realizado fuera de las horas estándar de trabajo es viable.
- El robo de ciclos de CPU realizado en horario laboral, sin impactar el uso interactivo de CPU y memoria es más difícil de implementar.

¿DÓNDE UBICAR A LOS CLUSTERS?



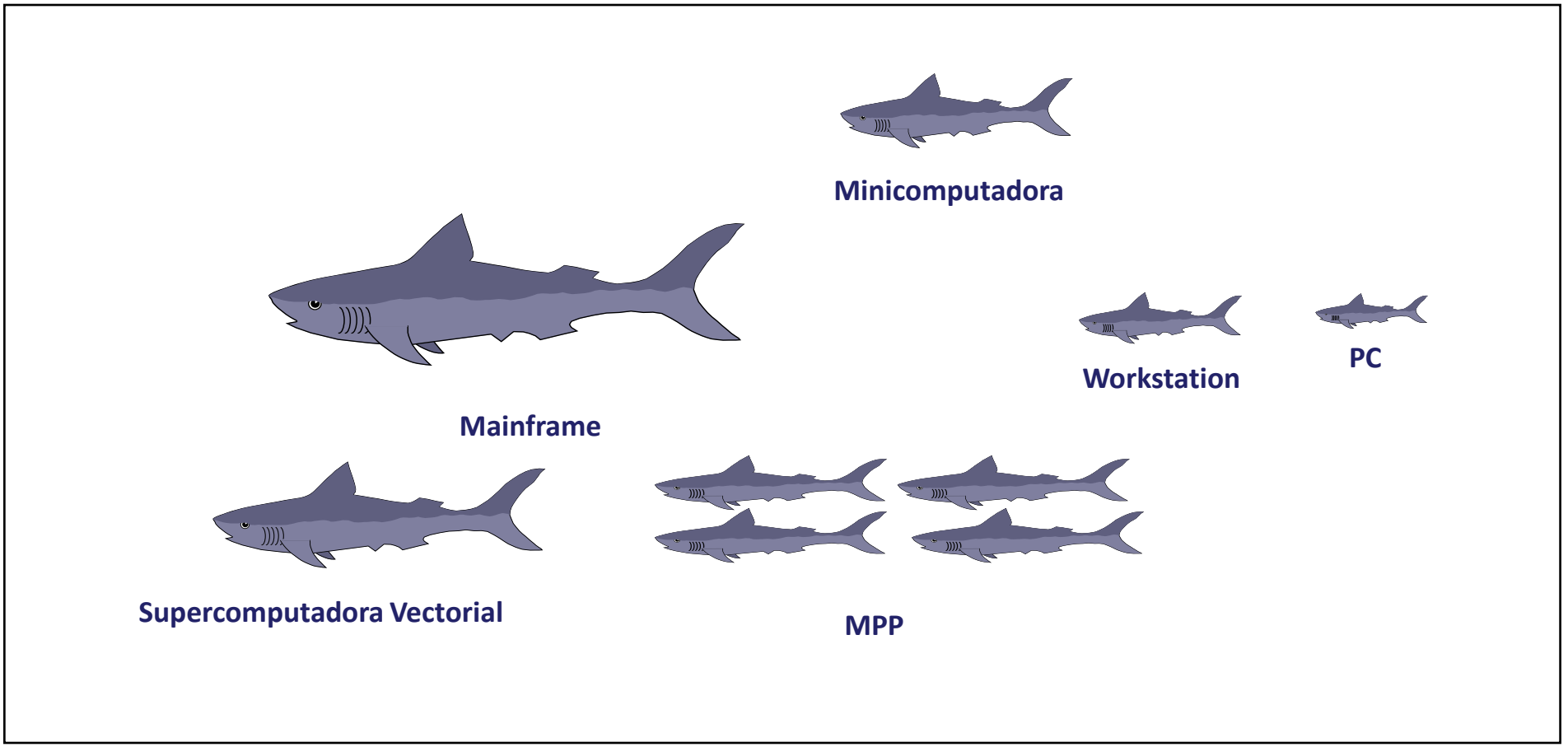
Cadena alimenticia real

¿DÓNDE UBICAR A LOS CLUSTERS?



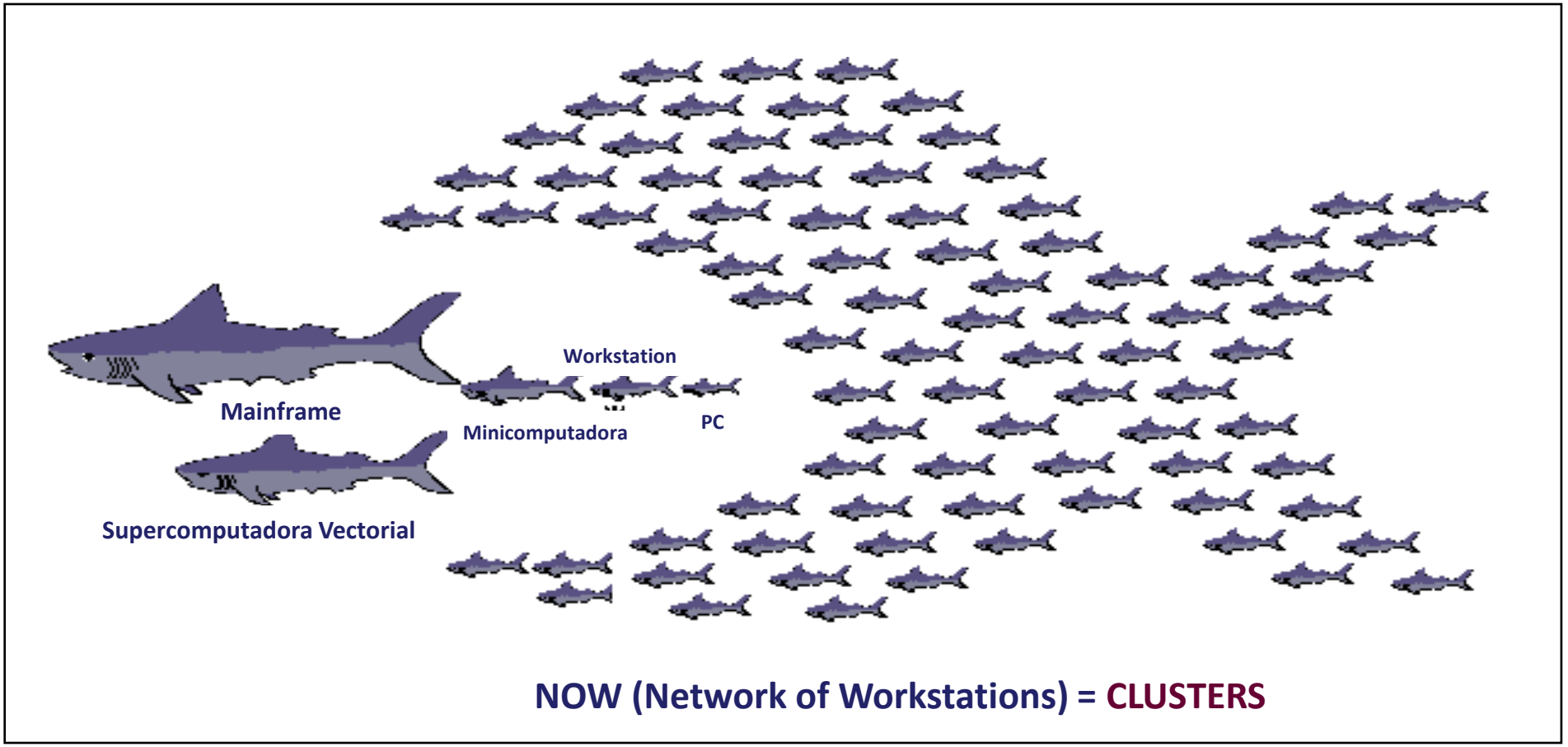
Cadena de las computadoras (década de 1990)

¿DÓNDE UBICAR A LOS CLUSTERS?



Cadena de las computadoras (década de 2000)

¿DÓNDE UBICAR A LOS CLUSTERS?



Cadena de las computadoras (presente y futuro)

FORMALMENTE: ¿QUÉ ES UN CLUSTER?

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Un cluster consiste en un tipo de sistema de procesamiento paralelo o distribuido, compuesto por un conjunto de computadoras que son capaces de trabajar cooperativamente como un único e integrado recurso de cómputo.

- Características de un típico cluster :
  - Red: rápida, mejor que una típica LAN
  - Protocolos de comunicación de latencia baja
  - Menor conexión que un SMP

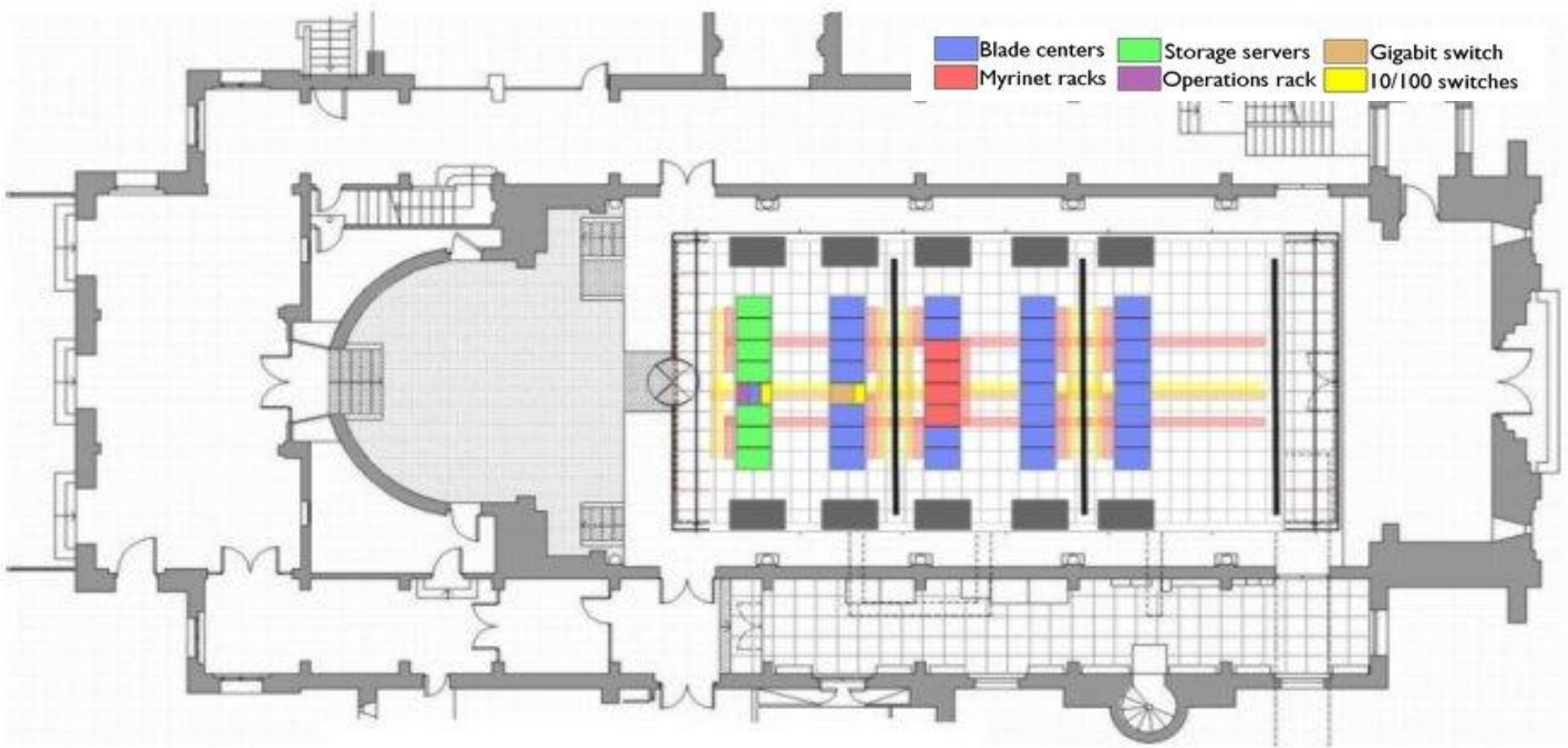
CLUSTERS: MOTIVOS DE SU DESARROLLO

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- El desempeño de los PCs se incrementó.
  - Ley de Moore: capacidad de procesamiento se duplica cada 18 meses aproximadamente.
- Las redes se hicieron cada vez más veloces.
  - Aumenta ancho de banda, interfaces simples.
- RAID (Almacenamiento redundante de bajo costo).
  - Alta disponibilidad y escalabilidad.
- Los clusters tienen escalabilidad incremental
  - Desempeño de nodos individuales puede mejorarse con recursos adicionales (memoria, disco).
  - Pueden agregarse nuevos nodos y reemplazar otros.
  - Clusters de clusters (metacomputadoras).
- Herramientas de software completas.
  - Threads, PVM, MPI, C, C++, Java, .NET, Compiladores, Debuggers, SOs, etc.
- Amplia gama de aplicaciones.

EJEMPLOS DE CLUSTERS: MARENOSTRUM

- MareNostrum es un cluster de procesadores PowerPC (arquitectura BladeCenter), SO Linux, y con red de interconexión Myrinet



#5 en Top500 (Junio de 2005): 27910 GFlops

EJEMPLOS DE CLUSTERS: MARENOSTRUM

- 42.144 Teraflops de rendimiento de pico teórico (42.144×10^{12} = 42 billones de operaciones por segundo).
- 4.800 procesadores PowerPC 970FX en 2400 Nodos duales de 2.2 GHz.
- 9.6 TB de memoria.
- 236 TB de almacenamiento en disco.
- 3 redes de interconexión:
  - Myrinet.
  - Gigabit Ethernet.
  - Ethernet 10/100.

EJEMPLOS DE CLUSTERS: THUNDER

#7 Top500 (Junio 2005): 19940 GFlops

- 1024 Nodos
- 4 CPU/Nodo
- Itanium 2 1.4 Ghz CPU
- 8GB RAM/Nodo
- Discos: 75GB/Nodo
- SO: CHAOS

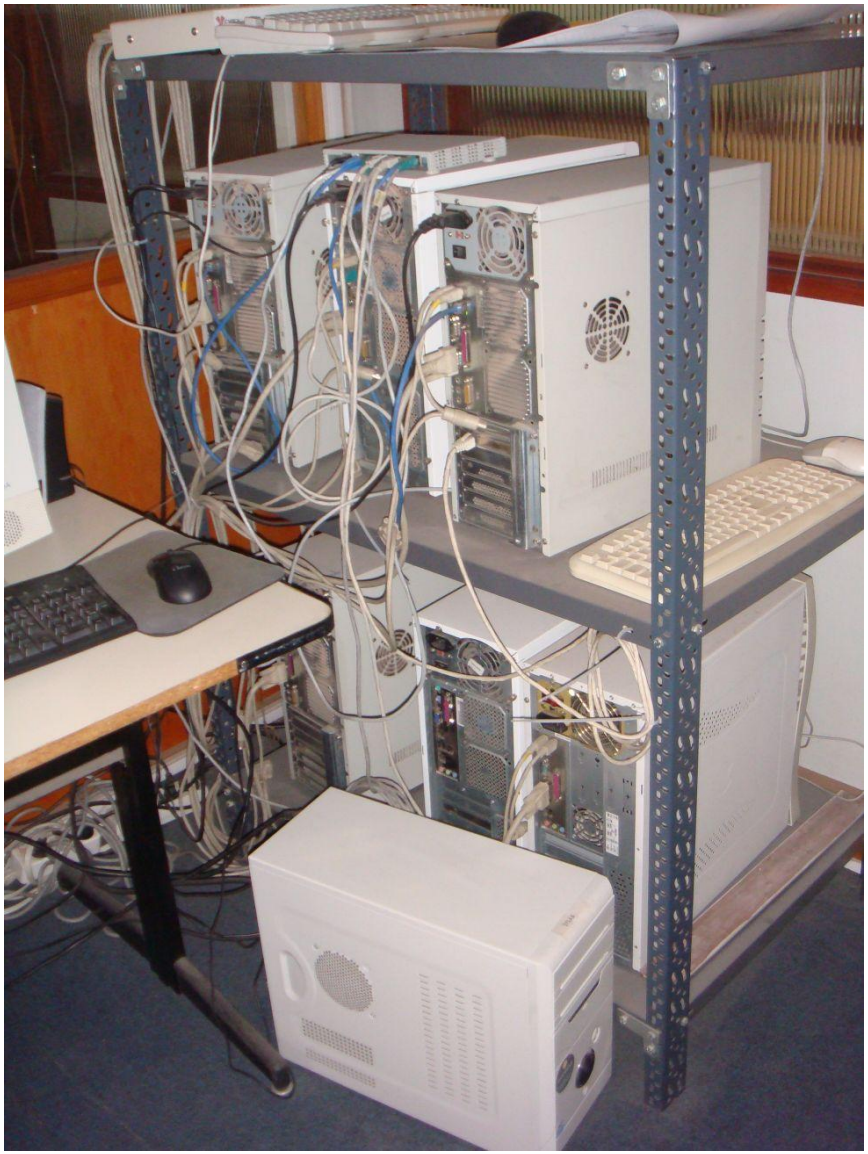
MILLENIUM PC CLUMPS

- Clusters baratos, fáciles de administrar
- Replicados en varios departamentos
- Prototipo para grandes clusters de PC

BEOWULF CLUSTERS

- Sterling y Becker (NASA), 1994
- Computadores comerciales, idénticos, que ejecutan software libre, sistema operativo tipo Unix (BSD, Linux, Solaris)
- Conectados por TCP/IP sobre LAN
- Tienen instalado software que permiten el procesamiento cooperativo
  - PVM, MPI, OSCAR

EL CLUSTER ROCK



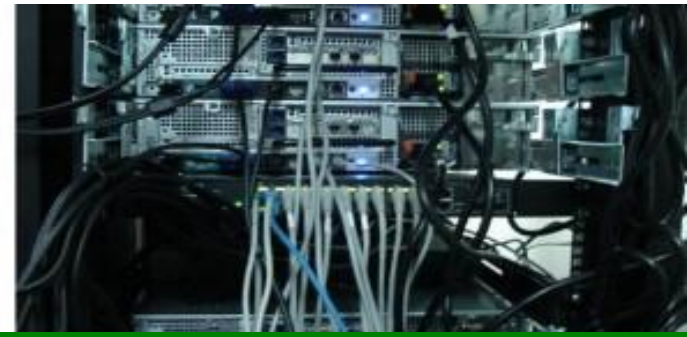
- Centro de Cálculo, Facultad de Ingeniería
- Primer cluster operativo de la Facultad (desde inicios de los 2000)
- Hasta 8 PCs conectados

EL CLUSTER FING

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

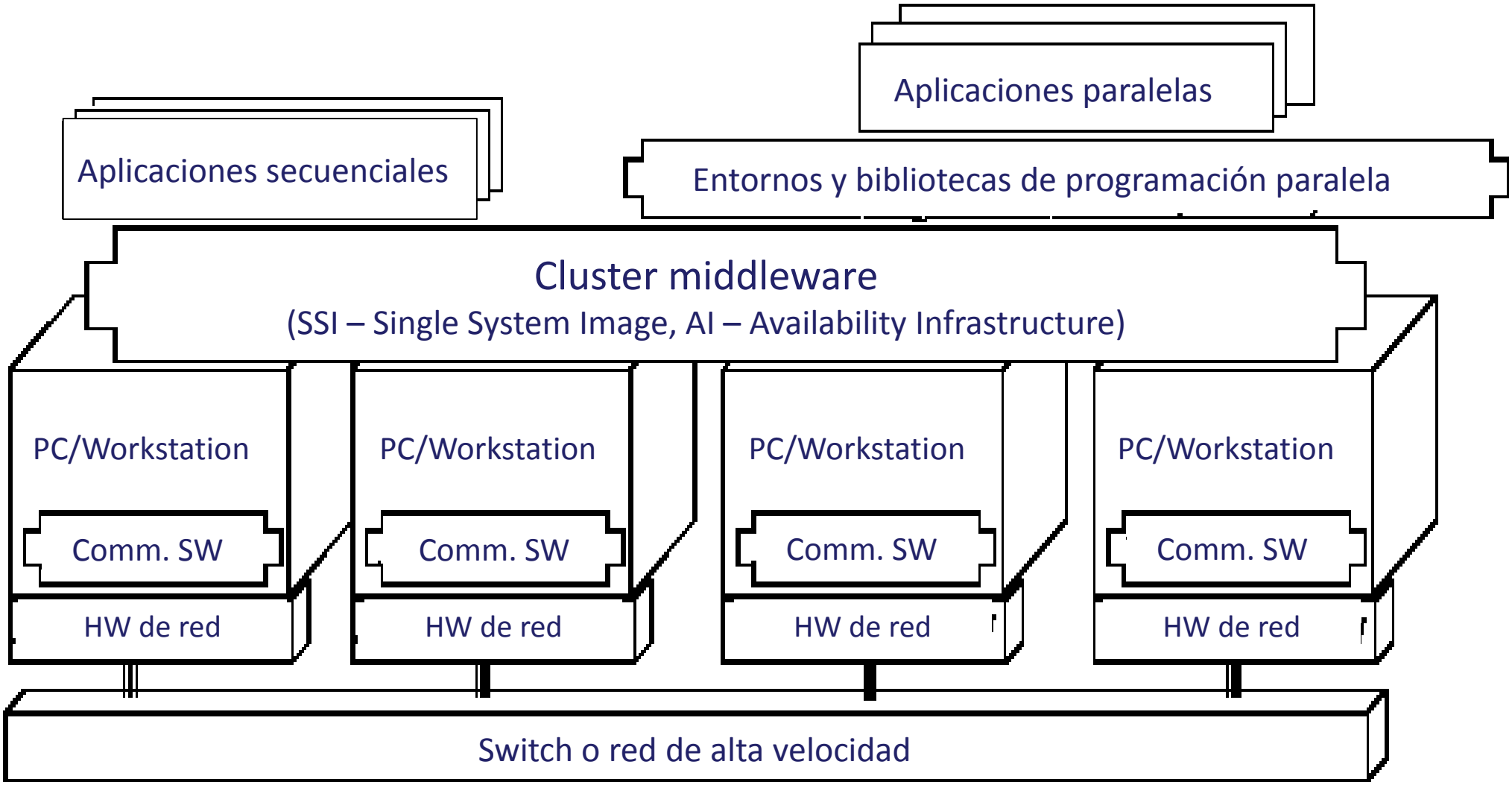
- Infraestructura para computación científica de alto desempeño, UdelaR.
- Operacional desde marzo de 2009
  - Autofinanciada
  - Autogestionada
- 1740 cores (540 de CPU, 960 de GPU, 240 Xeon Phi)
  - >1 TB de memoria RAM, >250 TB RAID storage, 34 kVA batería
  - Pico de performance: 6000 GFLOPS (6×10^{12} operaciones de punto flotante por segundo), *el mayor poder de cómputo disponible en el país*





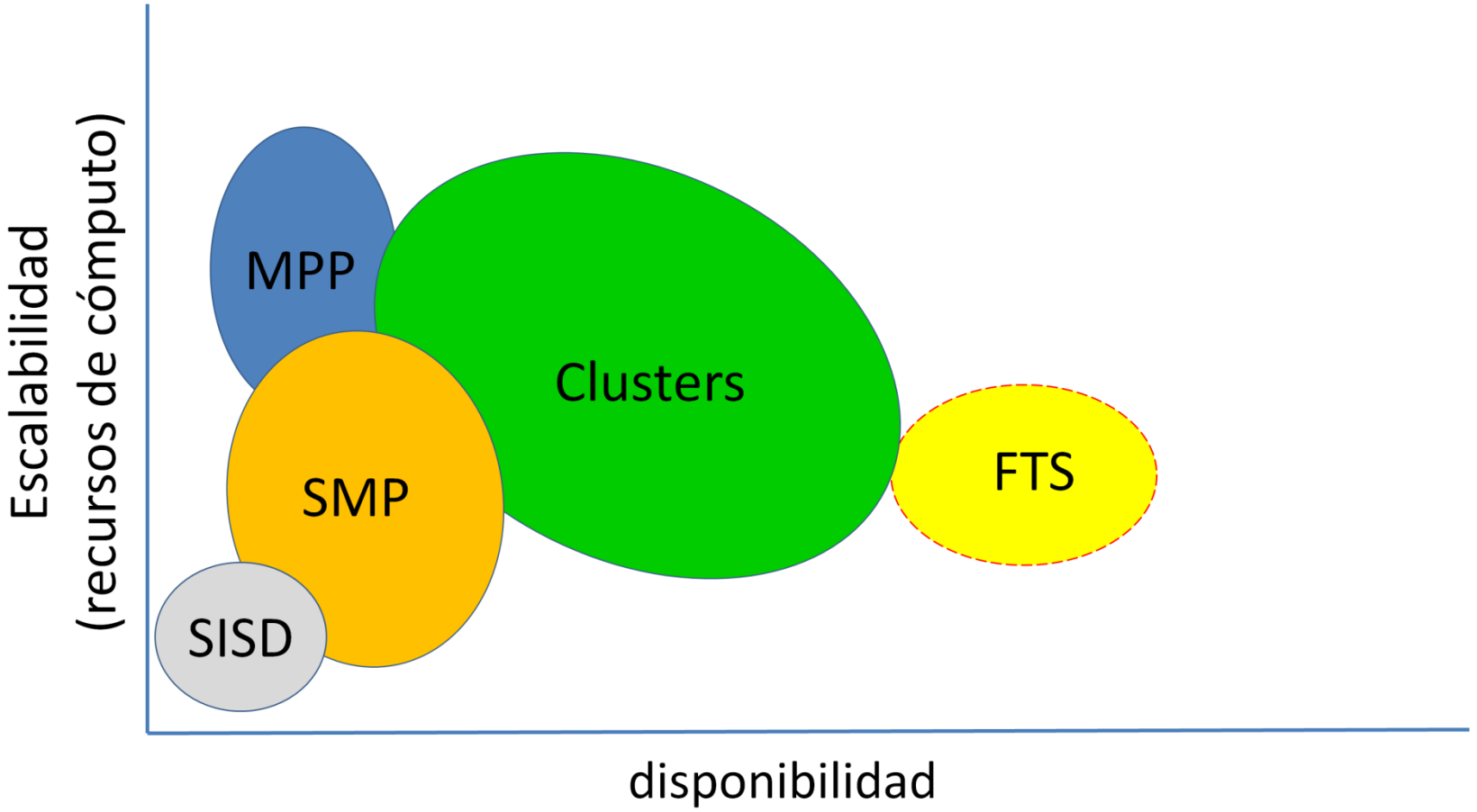
Cluster FING: >7.500.000 hs de cómputo efectivo (2016)
http://www.fing.edu.uy/cluster

ARQUITECTURA DE UN CLUSTER



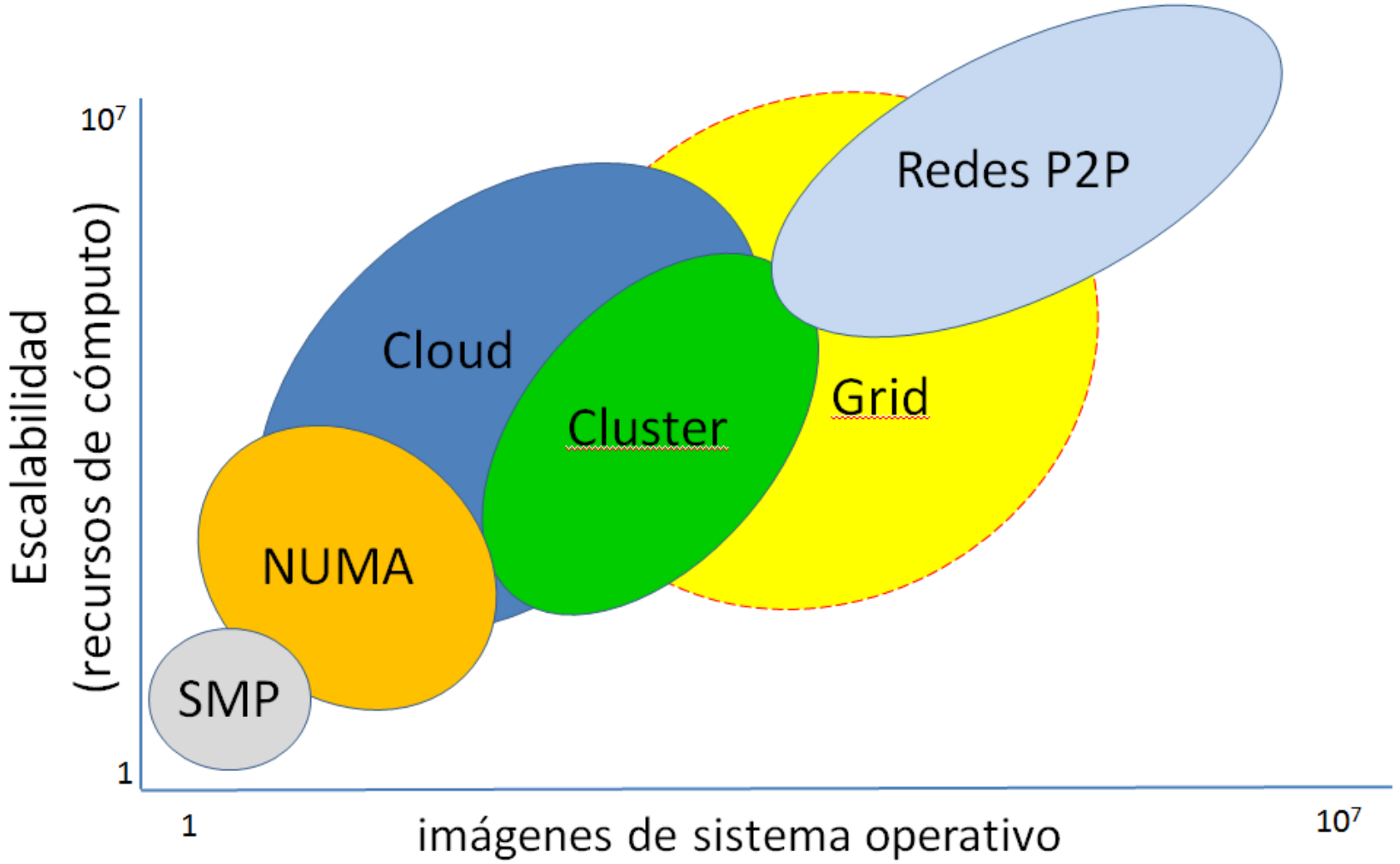
CLUSTERS: EFICIENCIA Y ROBUSTEZ

- Escalabilidad versus disponibilidad



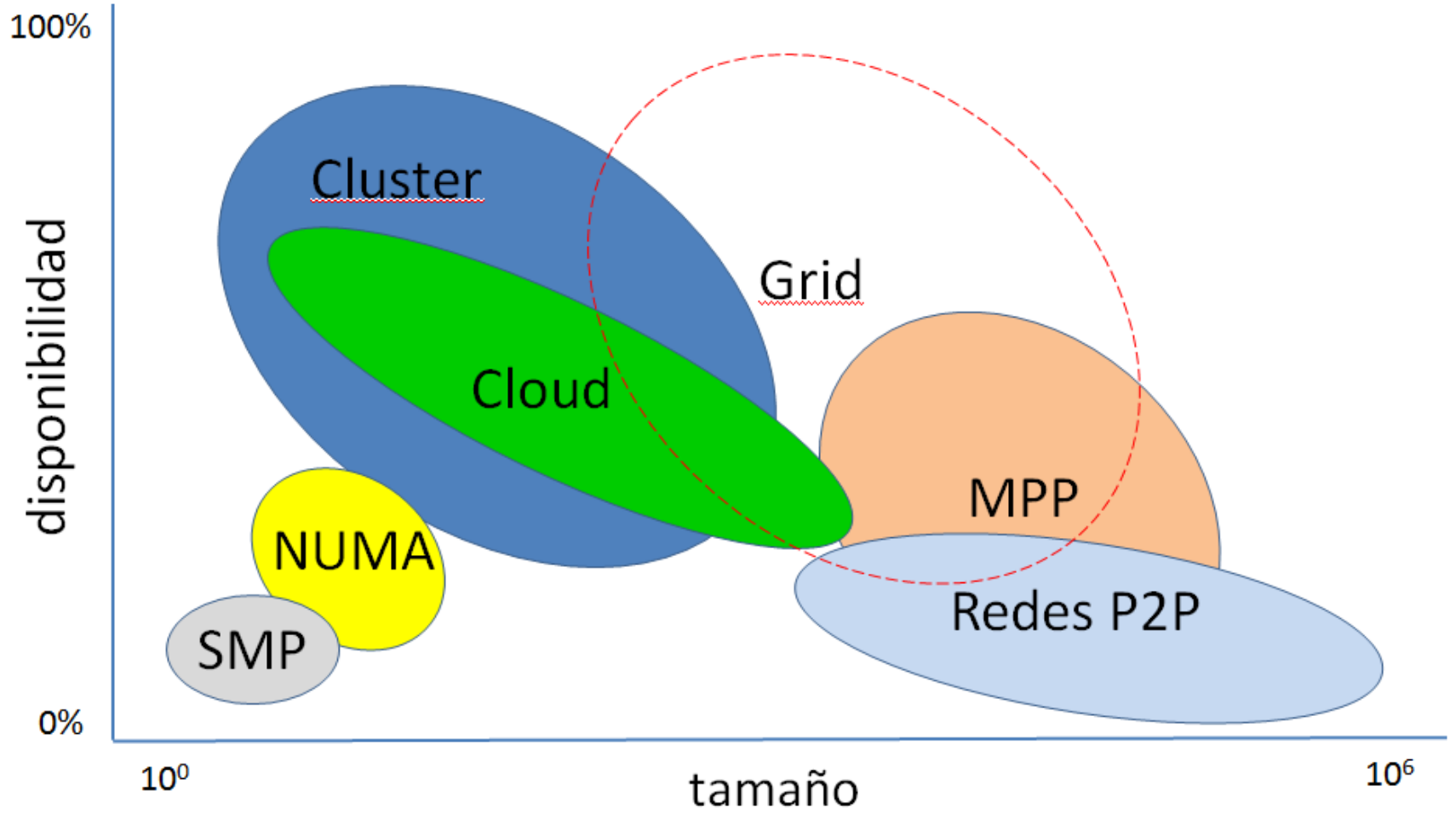
CLUSTERS: ESCALABILIDAD

- Escalabilidad versus imágenes del sistema operativo



CLUSTERS: DISPONIBILIDAD

- Disponibilidad versus tamaño



COMPONENTES DE UN CLUSTER

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Los componentes de un cluster incluyen:
  1. Nodos
     - Elementos de cómputo y almacenamiento de datos
  2. Software de base
     - Sistemas operativos
  3. Comunicaciones
     - Redes de alta velocidad
     - Interfaces y software para comunicaciones
  4. Middleware
  5. Entornos de programación
     - Bibliotecas y herramientas de desarrollo

CLUSTERS: NODOS Y PROCESADORES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Nodos: múltiples componentes de alta performance
  - PCs y workstations, servers, GPUs, coprocesadores
  - Sistemas de HPC distribuidos, que conducen a la “metacomputación”
- Los componentes pueden ser de diferentes arquitecturas, y ejecutar con diferentes sistemas operativos (clusters heterogéneos)
- Procesadores de varios tipos: CISC/RISC/VLIW/Vectoriales
  - Intel: Xeon (Clovertown, Harpertown, Nehalem, Westmere)
  - AMD Opteron SixCore, QuadCore, DualCore
  - IBM Power6, PowerXCell
  - Otros: Sun SPARC y ULTRASPARC, SGI MPIS, Digital Alpha, etc.
- También se han utilizados procesadores que integran memoria, procesador y red en un único chip
  - IRAM (CPU y memoria, http://iram.cs.berkeley.edu)
  - Alpha 21366 (CPU, Controlador de memoria, NI)

CLUSTERS: SISTEMAS OPERATIVOS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Estado del arte según OS:
  - Linux (desde los clusters Beowulf. Hoy, el 93% del Top500)
- El resto se lo reparten AIX, CNK/SLES, SLES10+SGI ProPack 5, CNL, CentOS, WindowsHPC, otros
- Otros SO que han sido utilizados desde 1995:
  - Microsoft NT (Illinois HPVM), SUN Solaris (Berkeley NOW), IBM AIX (IBM SP2), HP UX (Illinois - PANDA), Mach (SO basado en microkernel), Cluster Operating Systems (Solaris MC, SCO Unixware, MOSIX y proyectos académicos), OS gluing layers: (Berkeley Glunix)

CLUSTERS: COMUNICACIONES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Redes de alta velocidad
  - Ethernet (10Mb/s), Fast Ethernet (100Mb/s)
  - Gigabit Ethernet (1Gb/s), 10 Gigabit Ethernet
  - SCI (Dolphin-MPI latencia: 12 microsegundos)
  - ATM
  - Myrinet (1.2 Gb/s), Myrinet 2000 (2 Gb/s) y Myri-10G (10 Gb/s), altamente escalable
  - Infiniband (10 Gb/s, enlaces añadidos permiten alcanzar 100 Gb/s)
  - Digital Memory Channel, FDDI, etc.
- Tarjetas de red
  - Myrinet tiene NIC (Network Interface Card)
  - Soporte de acceso a nivel de usuario
  - Existen procesadores que integran controlador de memoria e interfaz de red en un único chip

CLUSTERS: SOFTWARE DE COMUNICACIONES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Facilidades estándar de IPC de los sistemas operativos
  - Sockets (TCP/IP), pipes, etc.
- Protocolos a nivel de usuario
  - Active Messages (Berkeley)
  - Fast Messages (Illinois)
  - U-net (Cornell)
  - XTP (Virginia)
- Sistemas que pueden ser construidos sobre los protocolos de base

CLUSTERS: MIDDLEWARE

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Reside entre el SO y las aplicaciones, ofrece infraestructura para soportar:
  - Single System Image (SSI)
  - System Availability (SA)

- SSI permite que un cluster pueda ser visto como un único equipo ("globaliza" los recursos de un sistema)
  - Accceso a través de ssh cluster.myinstitute.edu
  - Monitoreo centralizado, sistema de archivos global

- SA provee disponibilidad
  - Check pointing y migración de procesos

CLUSTERS: ENTORNOS DE PROGRAMACIÓN

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Threads (PCs, SMPs, …)
  - POSIX Threads
  - Java Threads
- Bibliotecas de desarrollo de progrmas paralelos y distribuidos
  - PVM (Parallel Virtual Machine)
  - MPI (Message Passing Interface)
- Software DSMs (Shmem)
- Colas de mensajes
- Otras tecnologías: Java RMI, .NET, etc.

CLUSTERS: OTRAS HERRAMIENTAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Compiladores
  - C/C++
  - Java
  - FORTRAN
- RAD (rapid application development tools)
  - Herramientas basadas en GUI para modelos de programación paralela.
- Debuggers
- Herramientas de análisis de performance
- Herramientas de visualización

CLUSTERS: CONCLUSIONES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Clusters
  - Insfraestructura promisoria para contemplar necesidades importantes de cómputo en un entorno de recursos limitados

- Principales ventajas:
  - Relación costo/performance
  - Escalabilidad incremental
  - Sistema “multipropósito” (no dedicado)

SERÁ LA PRINCIPAL PLATAFORMA DE TRABAJO A UTILIZAR EN EL CURSO

2.4: ARQUITECTURAS MULTINÚCLEO

TECNOLOGÍAS DE PROCESADORES

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

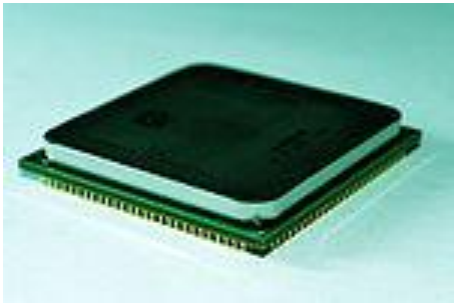
- Utilizando los avances tecnológicos, hasta la década de 2000 los fabricantes de procesadores prácticamente duplicaban la cantidad de transistores en un chip de 18 a 24 meses.
- Este ritmo se mantuvo hasta llegar a los límites físicos permitidos.
- El desarrollo de transistores se movió de un tamaño de 90nm a 65nm, lo que permite tener mas transistores en un chip.
- Sin embargo, existen reportes que pronostican que una vez que se alcance los 16nm en tamaño, el proceso no podrá controlar el flujo de los electrones a medida que el flujo se mueva a través de los transistores.
- De esa forma, llegará un momento que los chips no podrán ser más pequeños.
- Asimismo, al reducir su tamaño e incrementar su densidad, los chips generan mayor calor, causando errores en el procesamiento.

MULTINÚCLEOS (MULTICORE)

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- La tecnología de procesadores multinúcleo constituye una alternativa para mejorar la performance a pesar de las limitaciones físicas.
- Sin duda, los sistemas multinúcleo proponen mayores desafíos en cuanto al desarrollo de sistemas ya que se debe tener en cuenta que en el micro-tiempo se ejecuta más de una instrucción en el mismo equipamiento.
- Sin embargo, un buen uso de la tecnología puede implicar un beneficio importante en el poder de procesamiento.

Intel Core 2 Duo E6750



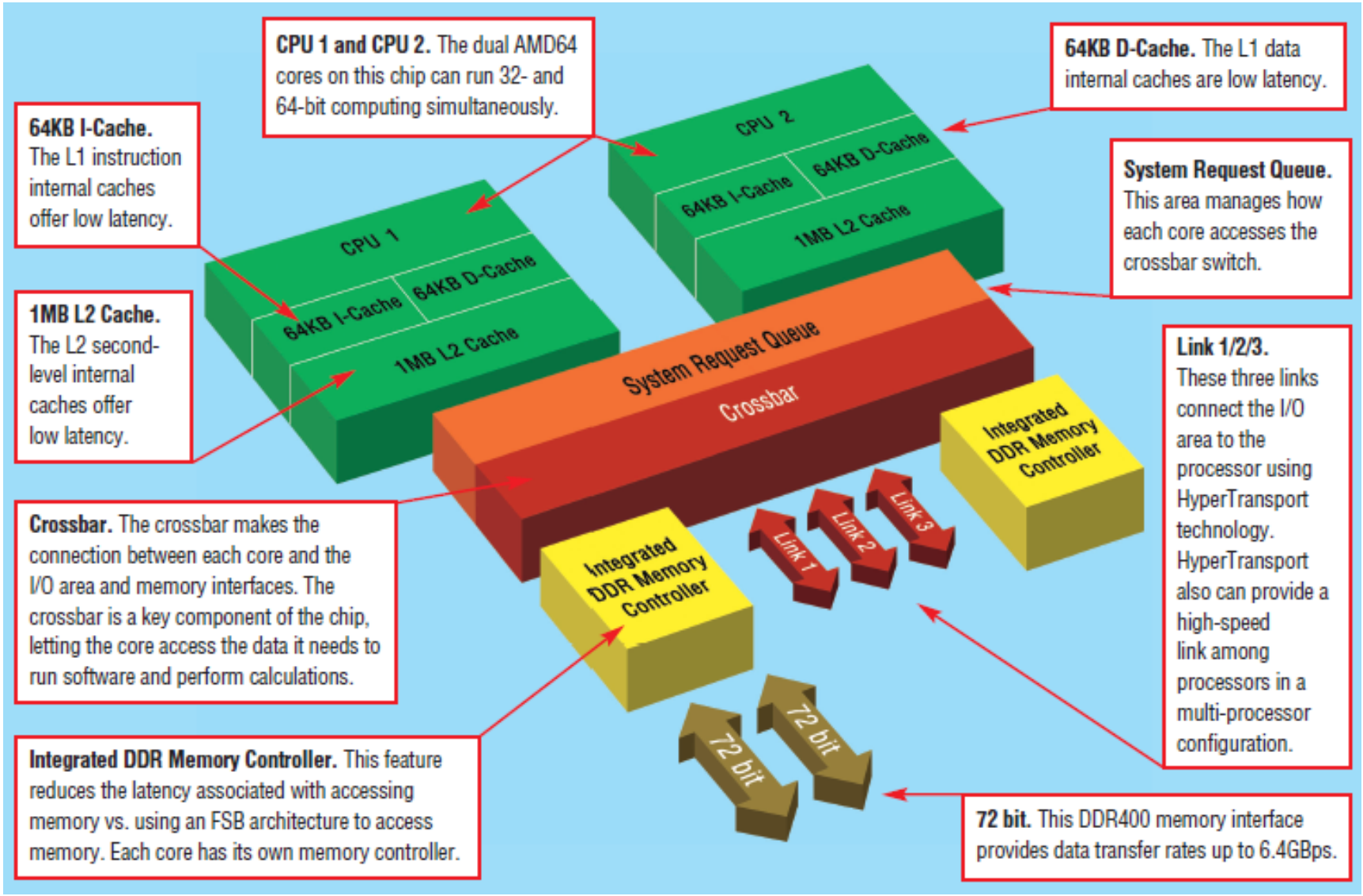
AMD Athlon X2 6400+

PRIMERAS TECNOLOGÍAS AMD MULTINÚCLEO

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- En 2005, AMD lanzó la línea de procesadores Opteron con una tecnología que denominó Direct Connect Arquitecture, la cual integraba el controlador de memoria en el mismo chip del procesador.
- AMD también dispuso de una memoria cache de segundo nivel (L2 cache) independiente para cada procesador.
- Para interconectar los dos procesadores se presentó un crossbar switch de alto desempeño que permitía el acceso cruzado a la memoria.
- La interconexión con los dispositivos de entrada/salida se realizaba a través de la tecnología HyperTransport.

AMD DUAL-CORE

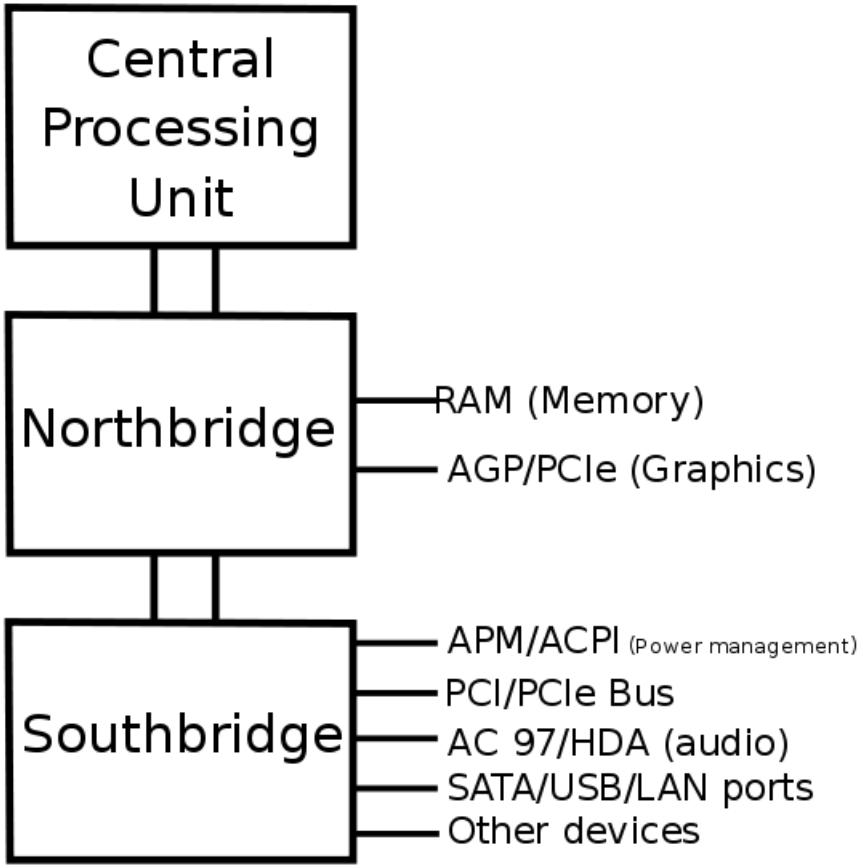


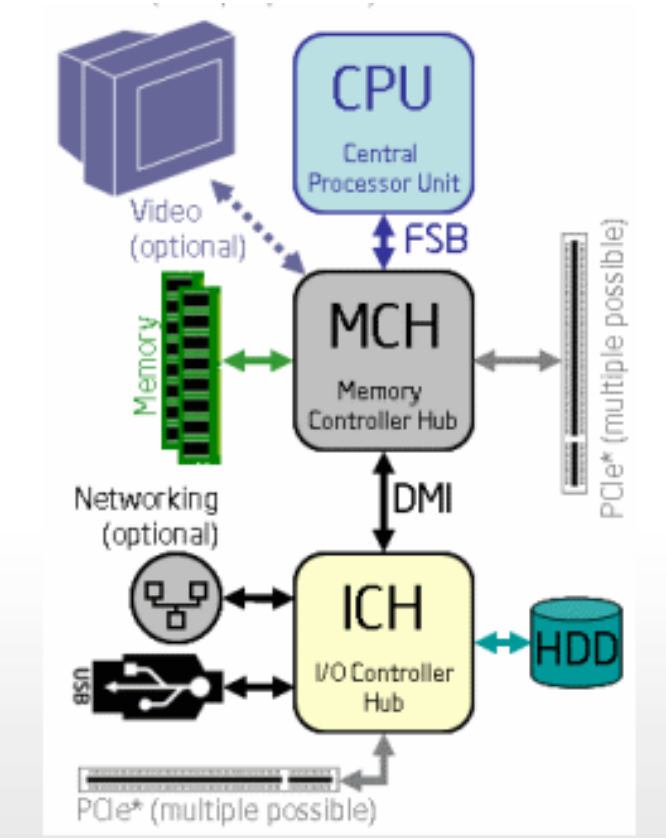
PRIMERAS TECNOLOGÍAS INTEL MULTINÚCLEO

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Los procesadores Intel mantuvieron la interconexión a través de un chipset compuesto por el NorthBridge y el SouthBridge.
- El NorthBridge contiene el controlador de acceso a la memoria RAM.
- El SouthBridge permite la interconexión con dispositivos de entrada/salida.
- El bus de interconexión entre los procesadores entre el procesador y el controlador de memoria es denominado FSB (Front Side Bus).
- De esta forma, los procesadores Intel mantuvieron el sistema UMA.
- En este caso, a diferencia de lo propuesto por AMD, los núcleos compartían la memoria cache de segundo nivel (L2 cache).

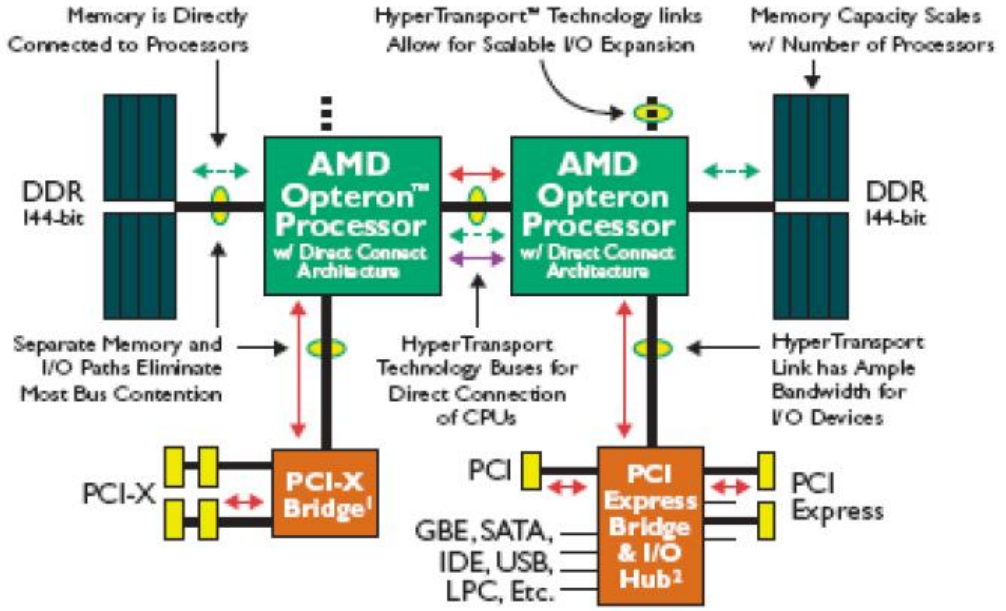
INTEL NORTH-SOUTH BRIDGE

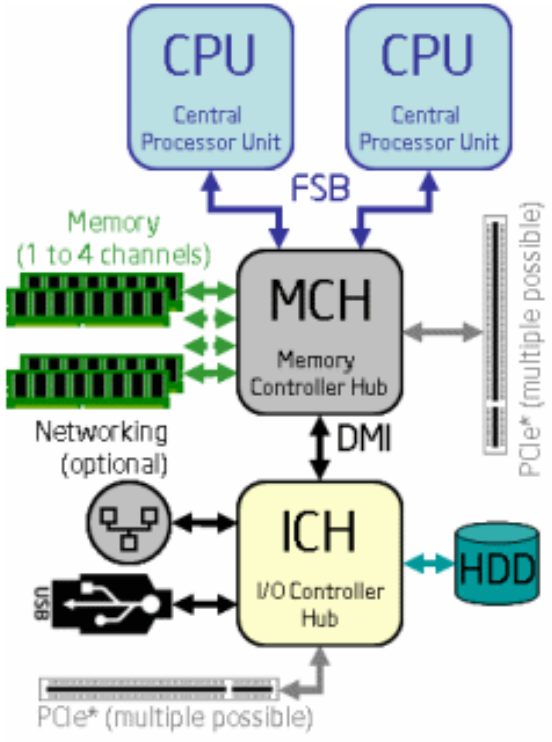




COMBINANDO PROCESADORES MULTINÚCLEO

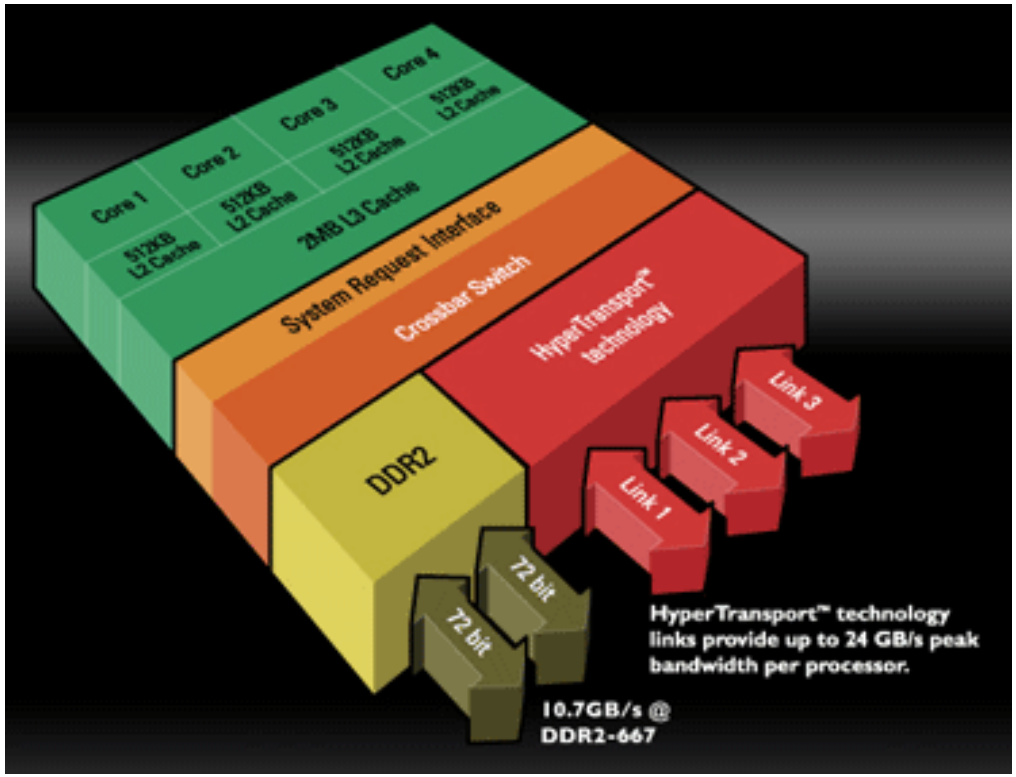
- A medida que la tecnología avanzó se incorporaron más procesadores por equipo, logrando combinar multiprocesadores con multinúcleo.
- La tecnología pasó a ser de un sistema NUMA.
- Los procesadores tienen acceso a toda la memoria (global), pero acceden a la memoria con diferentes velocidades.





TECNOLOGÍAS QUAD-CORE

- El siguiente paso fue abordar los procesadores quad-core.
- AMD propuso el chip denominado Barcelona que incluyó una memoria cache de segundo nivel (más reducida) particular de cada núcleo y una gran memoria cache de tercer nivel compartida por los cuatro núcleos.



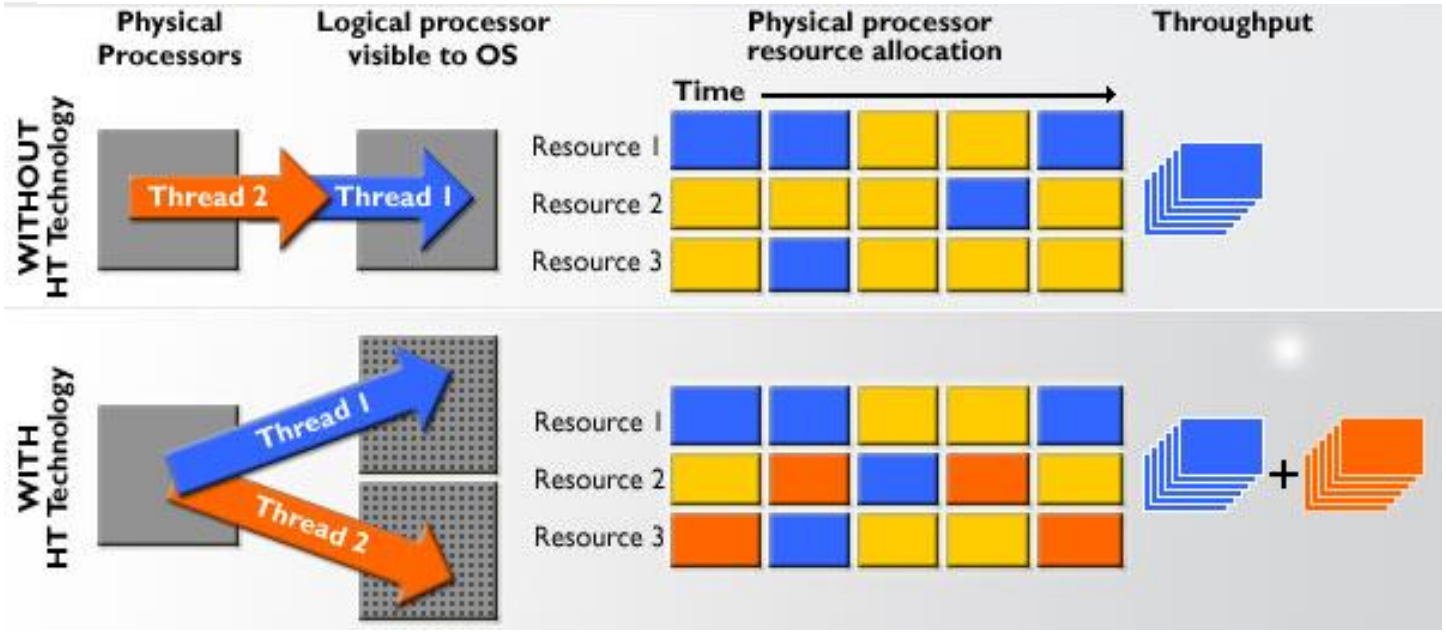
TECNOLOGÍAS QUAD-CORE

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Otro gran avance fue la inclusión de nuevas instrucciones (SSE128) en la arquitectura que permitieron operaciones de 128 bits (AMD, 2007/2008)
  - Estas operaciones permiten 4 operaciones de punto flotante de doble precisión por ciclo de reloj
  - Dos operaciones hechas a través de la multiplicación y dos a través de la suma
- Por otro lado, Intel siguió con su arquitectura de FSB pero con una cache de segundo nivel compartida por dos procesadores y de mayor tamaño
- A su vez, propuso el movimiento a procesadores de tamaño de 45nm
- Recién a partir de la familia de procesadores Nehalem (2010), Intel desecho el FSB para pasar la controladora de memoria en el procesador
  - Se incorporó una tecnología similar a la de AMD, pasando a un sistema NUMA
- Intel adoptó la tecnología de interconexión Quick-Path Interconnect (QPI) para la comunicación entre los procesadores y también para el acceso al chipset de entrada/salida

HYPERTHREADING

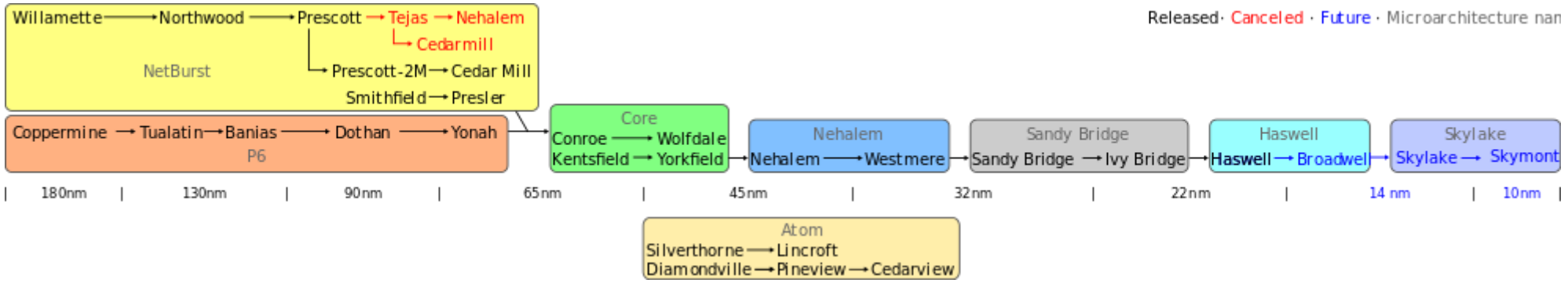
- Implementación propietaria de Intel para multithreading; introducida en 2002 con los procesadores Xeon (servidor) y Pentium 4 (desktop)
- Consiste en hacer que cada *núcleo físico* sea considerado por el sistema operativo como 2 *núcleos virtuales* independientes
- El sistema operativo puede despachar más de un proceso o hilo simultáneamente y los recursos del núcleo físico son compartidos entre los núcleos virtuales



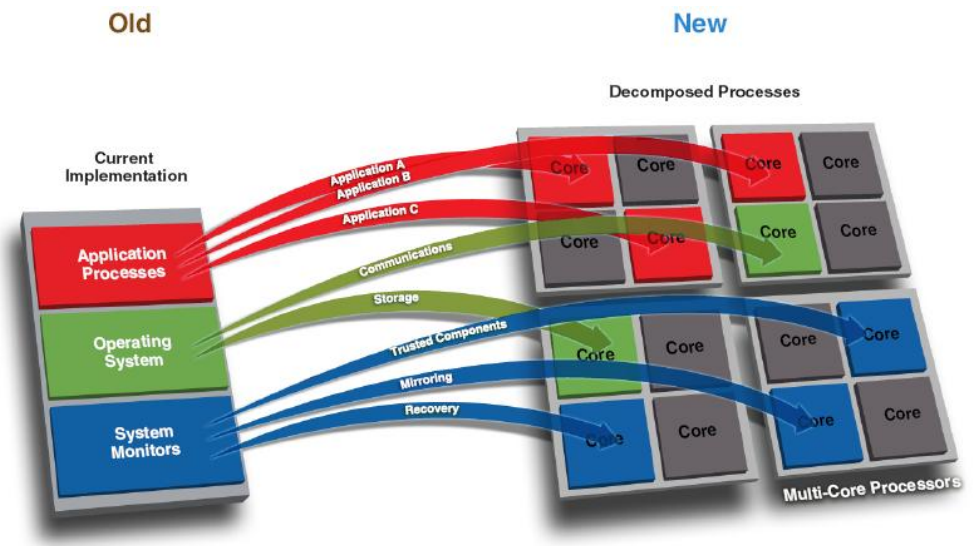
TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Actualmente Intel dispone de la familia *Haswell* (2014) y su sucesora Broadwell (2015-2016), con procesadores Xeon de 22 a 14 nm y hasta 18 núcleos físicos



- Los primeros chips Broadwell de 14 nm fueron lanzados en abril de 2016 (Broadwell-EP Xeon E5 V4) y junio de 2016 (familia Broadwell-E Core i7 69xx/68xx)
- En la familia Skylake (2015-2016), Intel ofrece el Core i7 6700K, con 4 cores a 4.0 GHz (hasta 4.2 GHz en single core)

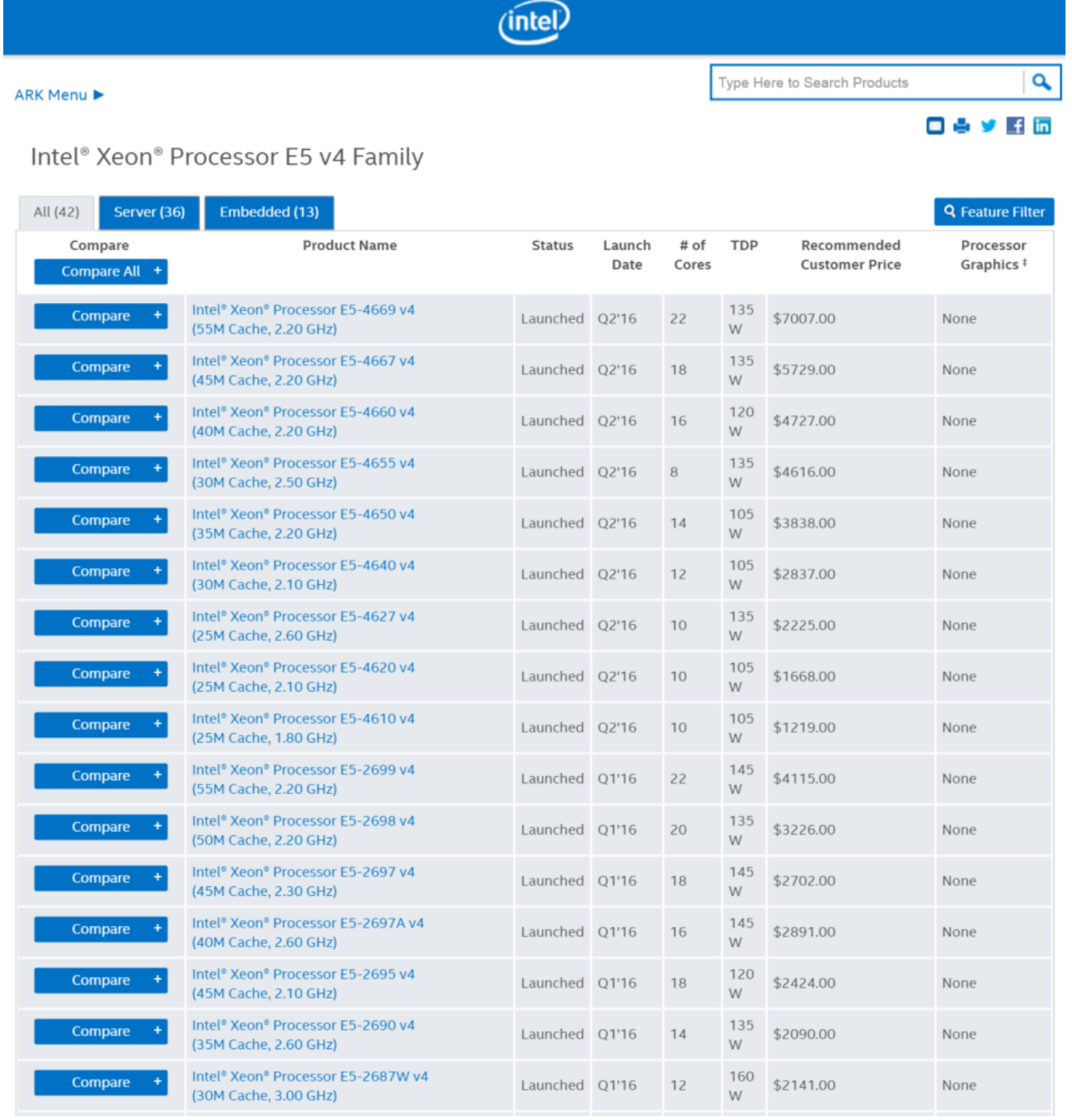


TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Desde a introducción del Xeon 5500 (Nehalem, marzo de 2009), Intel incrementó el numero de cores en los Xeon casi 50% cada 18 meses
  - Xeon 5500: 4
  - Xeon 5600: 6 cores (junio de 2010)
  - Xeon E5-2600 (Sandy Bridge): 8 cores [más rápidos] (marzo de 2012)
  - Xeon E5-2600 v2 (Ivy Bridge EP): 12 cores (setiembre de 2013)
  - Xeon E5-2600 v3 (Haswell-EP, setiembre de 2014): 18 cores (setiembre 2014)
- Por razones energéticas y de espacio, no es posible mantener el ritmo de incremento en el número de cores

- Xeon E5 v4 (Broadwell-EP, abril de 2016): servidor multi-socket con más cores, mayor ancho de banda a memoria y más caché
- Da el salto de 22nm a 14 nm e integra 24 cores (“solo” 22 activados)
- Reloj (turbo) de 3.6 GHz, relojes base más lentos
- El salto de desempeño es modesto



ARK Menu ▶

Type Here to Search Products

Intel® Xeon® Processor E5 v4 Family

All (42) | Server (36) | Embedded (13) | Feature Filter

Compare	Product Name	Status	Launch Date	# of Cores	TDP	Recommended Customer Price	Processor Graphics ‡
Compare All +							
Compare +	Intel® Xeon® Processor E5-4669 v4 (55M Cache, 2.20 GHz)	Launched	Q2'16	22	135 W	$7007.00	None
Compare +	Intel® Xeon® Processor E5-4667 v4 (45M Cache, 2.20 GHz)	Launched	Q2'16	18	135 W	$5729.00	None
Compare +	Intel® Xeon® Processor E5-4660 v4 (40M Cache, 2.20 GHz)	Launched	Q2'16	16	120 W	$4727.00	None
Compare +	Intel® Xeon® Processor E5-4655 v4 (30M Cache, 2.50 GHz)	Launched	Q2'16	8	135 W	$4616.00	None
Compare +	Intel® Xeon® Processor E5-4650 v4 (35M Cache, 2.20 GHz)	Launched	Q2'16	14	105 W	$3838.00	None
Compare +	Intel® Xeon® Processor E5-4640 v4 (30M Cache, 2.10 GHz)	Launched	Q2'16	12	105 W	$2837.00	None
Compare +	Intel® Xeon® Processor E5-4627 v4 (25M Cache, 2.60 GHz)	Launched	Q2'16	10	135 W	$2225.00	None
Compare +	Intel® Xeon® Processor E5-4620 v4 (25M Cache, 2.10 GHz)	Launched	Q2'16	10	105 W	$1668.00	None
Compare +	Intel® Xeon® Processor E5-4610 v4 (25M Cache, 1.80 GHz)	Launched	Q2'16	10	105 W	$1219.00	None
Compare +	Intel® Xeon® Processor E5-2699 v4 (55M Cache, 2.20 GHz)	Launched	Q1'16	22	145 W	$4115.00	None
Compare +	Intel® Xeon® Processor E5-2698 v4 (50M Cache, 2.20 GHz)	Launched	Q1'16	20	135 W	$3226.00	None
Compare +	Intel® Xeon® Processor E5-2697 v4 (45M Cache, 2.30 GHz)	Launched	Q1'16	18	145 W	$2702.00	None
Compare +	Intel® Xeon® Processor E5-2697A v4 (40M Cache, 2.60 GHz)	Launched	Q1'16	16	145 W	$2891.00	None
Compare +	Intel® Xeon® Processor E5-2695 v4 (45M Cache, 2.10 GHz)	Launched	Q1'16	18	120 W	$2424.00	None
Compare +	Intel® Xeon® Processor E5-2690 v4 (35M Cache, 2.60 GHz)	Launched	Q1'16	14	135 W	$2090.00	None
Compare +	Intel® Xeon® Processor E5-2687W v4 (30M Cache, 3.00 GHz)	Launched	Q1'16	12	160 W	$2141.00	None

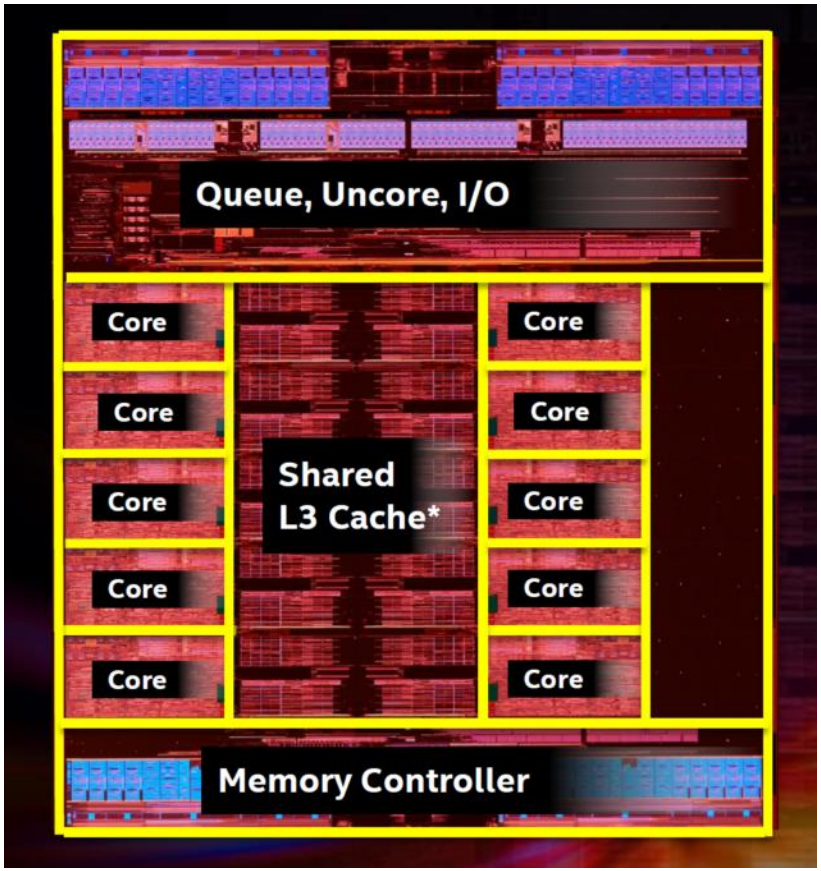
$7007.00
$5729.00
$4727.00
$4616.00
$3838.00
$2837.00
$2225.00
$1668.00
$1219.00
$4115.00
$3226.00
$2702.00

TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- En equipos de escritorio, Intel ofrece la familia Broadwell-E Core i7 69xx/68xx
- Core i7-6950X: 10 cores

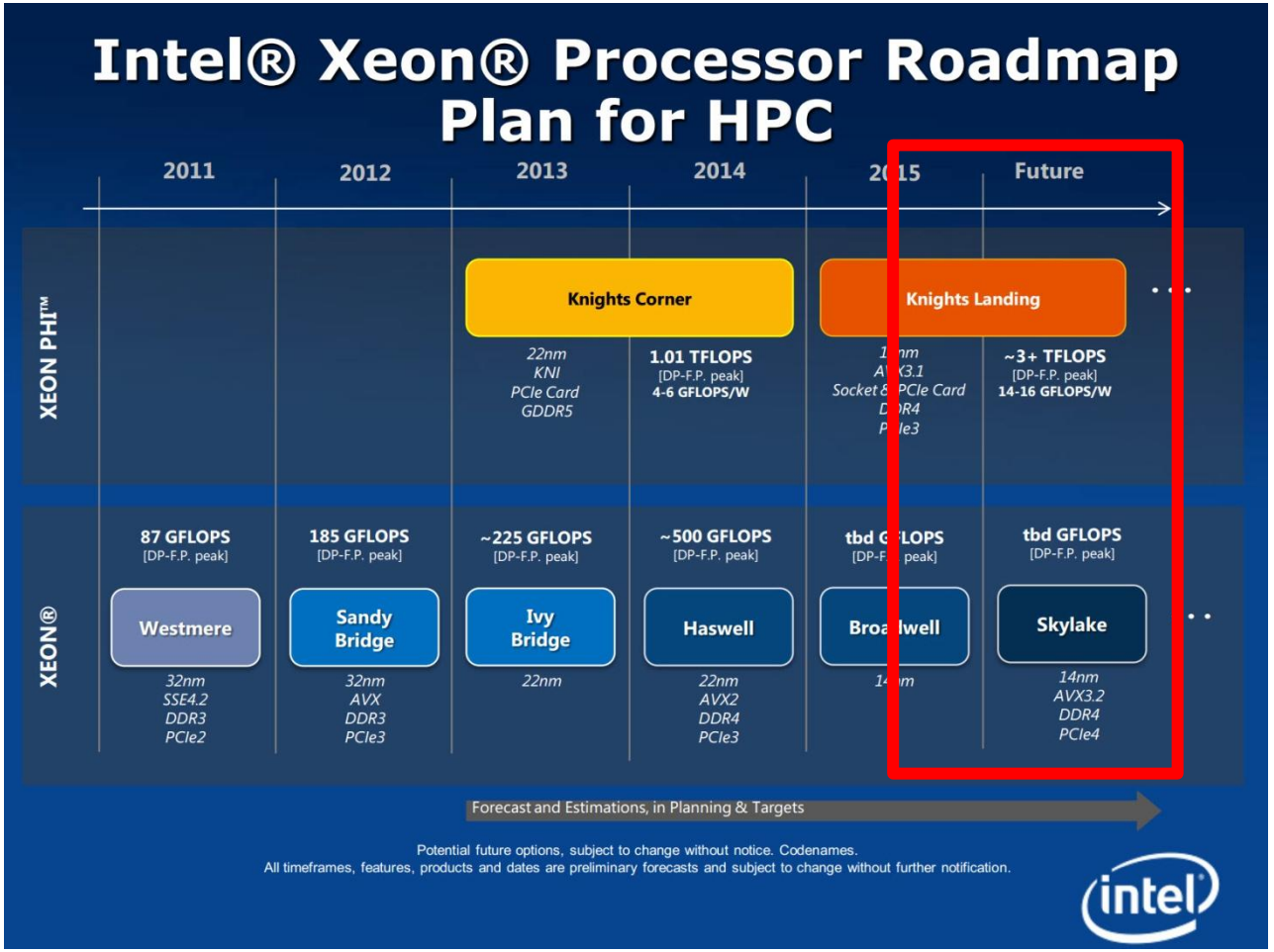
Core i7	6950X	6900K	6850K	6800K	5960X	4960X	3960X
Cores	10	8	6	6	8	6	6
Threads	20	16	12	12	16	12	12
Base CPU Freq	3.0 GHz	3.2 GHz	3.6 GHz	3.4 GHz	3.0 GHz	3.6 GHz	3.3 GHz
Turbo CPU Freq	3.5 GHz	3.7 GHz	3.8 GHz	3.6 GHz	3.5 GHz	4.0 GHz	3.9 GHz
TDP	140W					130W	
Memory Freq.	DDR4-2400				DDR4 2133	DDR3 1866	DDR3 1600
L3 Cache	25MB	20MB	15MB	15MB	20MB	15MB	15MB
PCIe Lanes	40	40	40	28	40	40	40
Arch.	Broadwell-E				HSW-E	IVB-E	SNB-E
Price	$1723	$1089	$617	$434	$999	$990	$990



TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- En 2015, Intel lanzó al mercado la familia *Skylake*
- Se mantiene la tecnología multinúcleo (8 núcleos) y el tamaño del procesador (14nm)
- Implementan la extensión AVX (Advanced Vector eXtensions) 3.2 para registros SSE de 256 bits y utiliza memoria RAM DDR4



TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- AMD brinda, a través de la familia *Piledriver* (APU, AMD FX y Opteron) procesadores Warsaw de 32nm y 16 núcleos físicos

Modelo	Cores	Frecuencia			Cache		HT	TDP	Lanzado
		Base	Full Load turbo	Half Load turbo	L2	L3			
Warsaw – sixteen core									
Opteron 6370P	16	2.0 GHz	2.2 GHz	2.5 GHz	8 × 2 MB	2 × 8 MB	3.2 GHz	99 W	enero 2014
Opteron 6376	16	2.3 GHz	2.6 GHz	3.2 GHz	8 × 2 MB	2 × 8 MB	3.2 GHz	115 W	noviembre 2012
Opteron 6378	16	2.4 GHz	2.7 GHz	3.3 GHz	8 × 2 MB	2 × 8 MB	3.2 GHz	115 W	noviembre 2012
Opteron 6380	16	2.5 GHz	2.8 GHz	3.4 GHz	8 × 2 MB	2 × 8 MB	3.2 GHz	115 W	noviembre 2012
Opteron 6386 SE	16	2.8 GHz	3.2 GHz	3.5 GHz	8 × 2 MB	2 × 8 MB	3.2 GHz	140 W	noviembre 2012

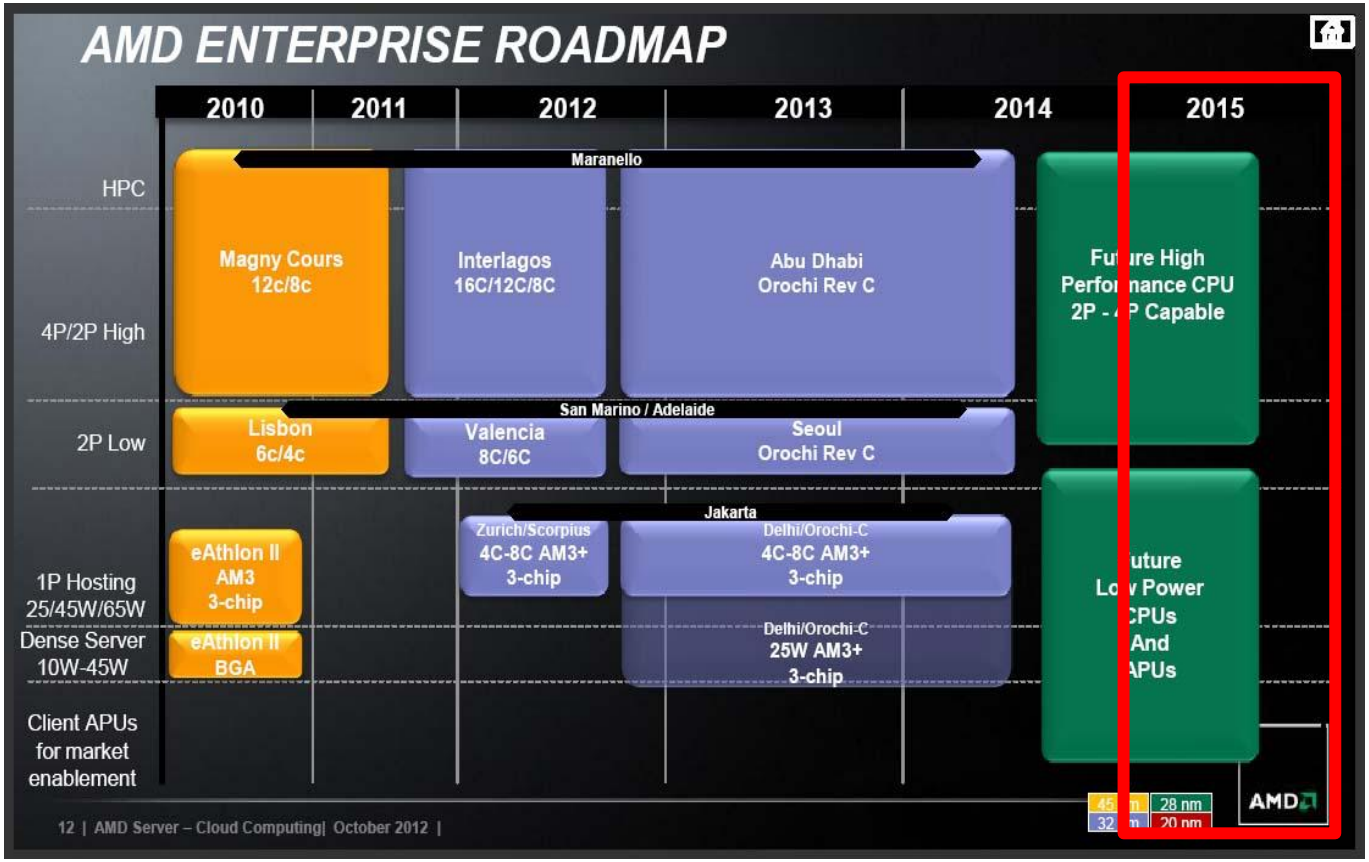
- La familia con microarquitectura Jaguar ofrece procesadores de 28 nm
  - Kyoto (X1150): 4 núcleos, 1,2–2 GHz, 9–17 W (mayo de 2013)
  - Kyoto (X2150): 4 núcleos APU a 1,1–1,9 GHz y 128 núcleos GPU a 266– 600 MHz, 11–22W (mayo de 2013)

TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- AMD introdujo procesadores *Steamroller* a comienzos de 2014 y *Excavator* en 2015 (basados en tecnología *Piledriver*, sucesora de *Bulldozer*)

- La tecnología es de 28 nm y hasta 16 núcleos
- Se mantiene la misma capacidad de cómputo usando menos energía
- Posee FPU con mejor desempeño que los procesadores anteriores

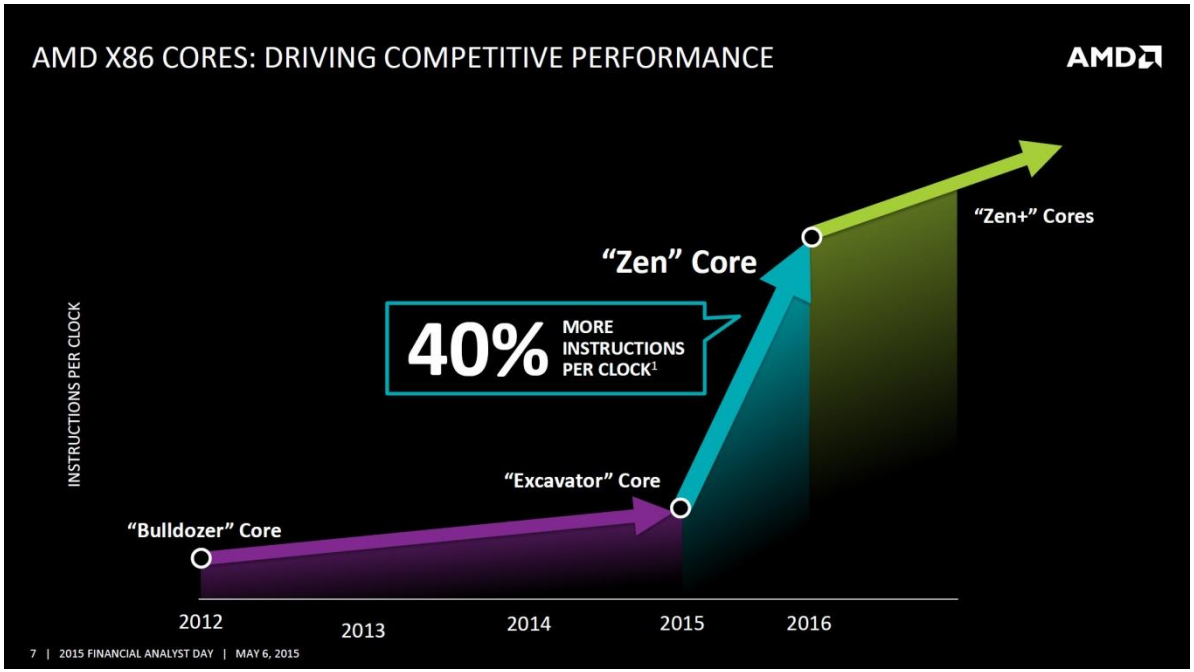


TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- AMD prepara el lanzamiento de la microarquitectura Zen (lanzamiento esperado para octubre de 2016)
- Nuevo diseño: procesador de 14 nm y nuevo socket AM4, full system on chip
- Incremento en performance por core utilizando *multithreading simultáneo* (threads, procesos separados, etc.)

- Cada core tiene cuatro unidades enteras, dos unidades de generación de direcciones y cuatro unidades de punto flotante, permite implementar *clusters multithreading*



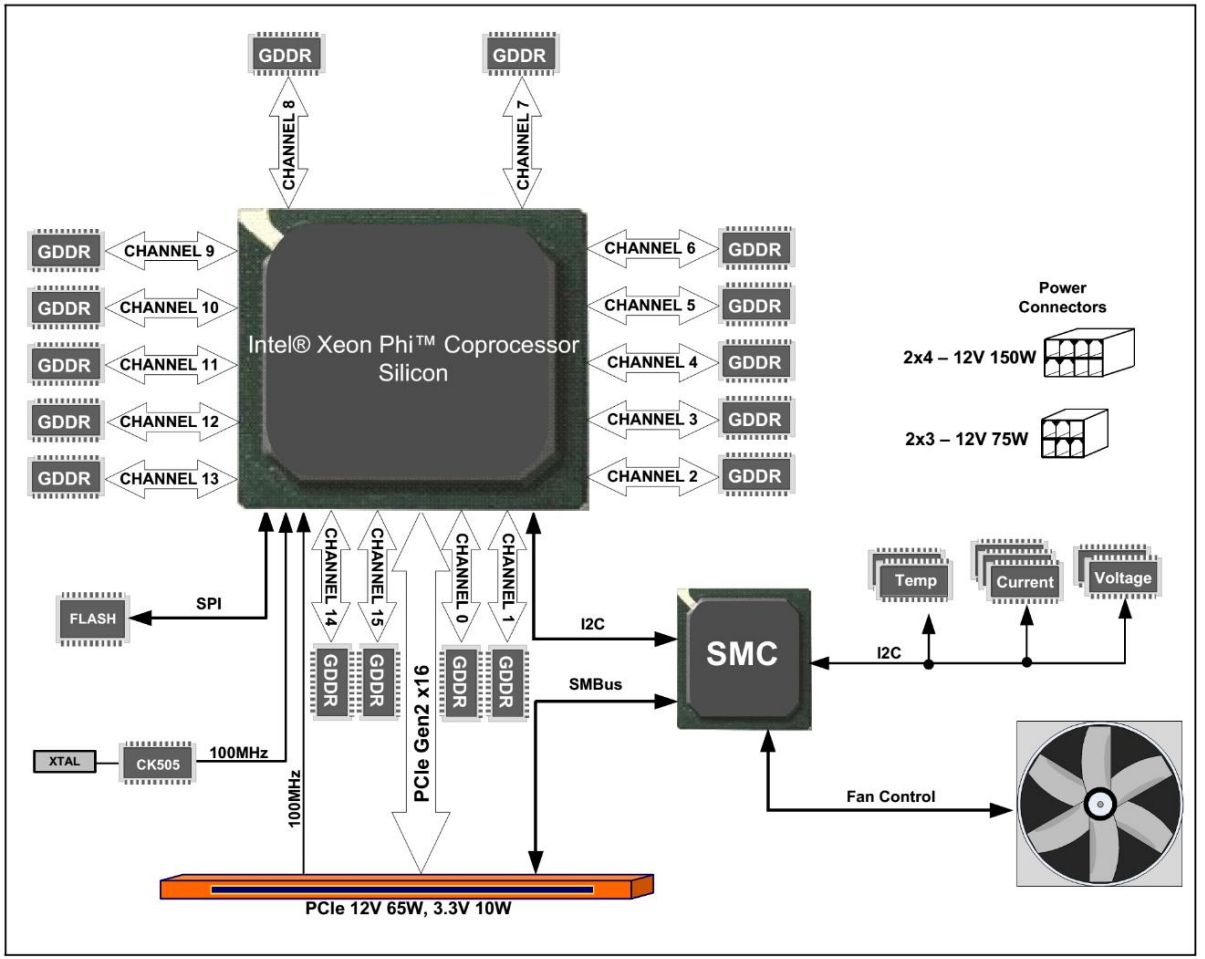
TECNOLOGÍAS ACTUALES Y FUTURAS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- SPARC M7 Server (Oracle), julio de 2015
- Procesador
  - 32 cores, 4.13 GHz, 256 threads (8 por core)
  - 32 unidades de punto flotante por procesador (una por core)
  - 32 aceleradores de encriptado con algoritmos criptográficos
  - 8 aceleradores por procesador, cada uno soporta cuatro consultas concurrentes con descompresión
- Un generador de números aleatorios
- Caché:
  - L1: 16 KB de instrucciones y 16 KB de datos
  - L2: 256 KB de instrucciones y 16 KB de datos cada cuatro cores
  - L3: 64 MB
- Dos a ocho procesadores por sistema
- 16 módulos DIMM de hasta 64 GB por procesador: memoria máxima 8 TB

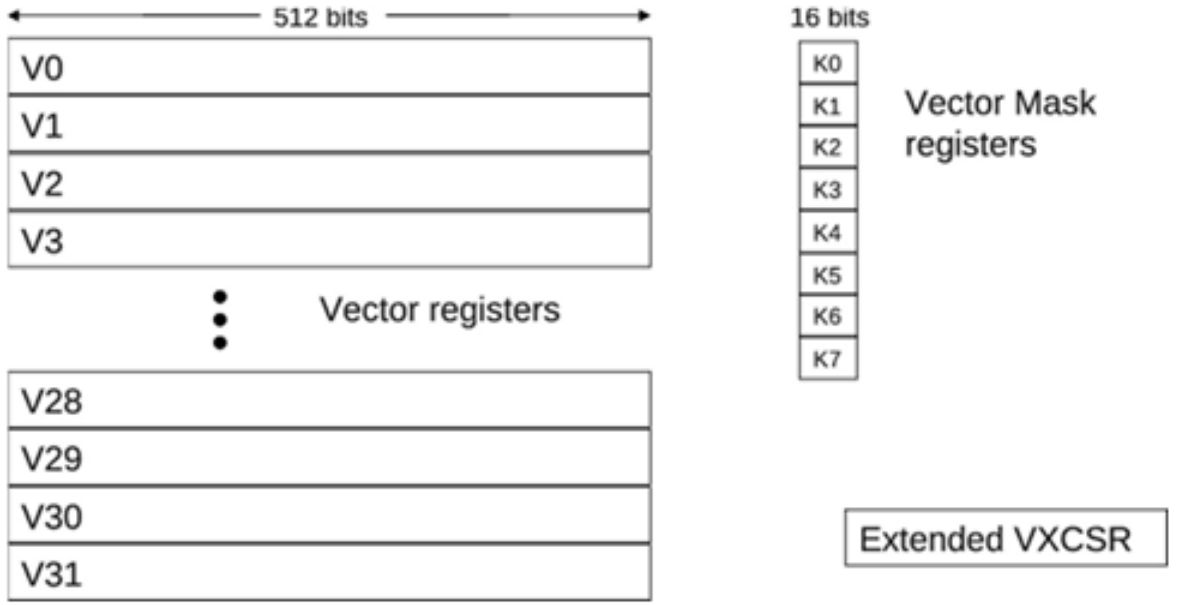
XEON PHI

- Introducida por Intel en noviembre de 2012; ofrece hasta 61 núcleos y está desarrollada para trabajar en conjunto con los procesadores Xeon
- Cada núcleo soporta 32 vectores de 512 bits y posee una *Vector Processing Unit* (VPU) y una *Extended Math Unit* (EMU) propias
- Un coprocesador tiene hasta 8 controladores de memoria RAM y cada controlador soporta 2 bancos de memoria GDDR5



XEON PHI

- Hay disponibles 2 caches de nivel 1 (L1) de 32KB por cache (para instrucciones y para datos); cada núcleo posee un cache de nivel 2 (L2) de 512KB para uso local
- Soporta *memory prefetching* de datos (por software o hardware) a las caches L1 y L2
- En cada ciclo de reloj se realizan 16 (8) operaciones de simple (doble) precisión (hay 32 vectores de 512 bits)
- El estado de los vectores se controla de forma condicional (*Vector Mask Register* y *Extended VXCSR*)
- Conjunto de instrucciones para operaciones vectoriales



XEON PHI

- Otras características destacadas:
  - Comunicación directa entre dispositivos PCI Express sin depender del anfitrión
  - Hasta 8 canales DMA operando simultáneamente
  - Control de consumo energético según el estado de cada núcleo
  - Sensores para control del estado del hardware (temperatura, velocidad de los ventiladores, potencia de entrada del slot PCI Express y los conectores de energía, etc)

Estudiaremos programación paralela en Xeon Phi en el curso

MULTI-CORE y MULTI-THREADS

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

- Las arquitecturas multinúcleo son útiles y eficientes para implementar programas multi-threads
- Los threads (o hilos) son las unidades de procesamiento
  - Múltiples threads de un proceso son capaces de compartir estado e información (memoria y otros recursos)
  - Los threads comparten el espacio de direccionamiento (variables)
  - Los threads son capaces de comunicarse sin utilizar mecanismos explícitos de IPC
  - El cambio de contexto entre threads es más veloz que entre procesos

La programación multithreading será estudiada en el curso