Computer Architecture
Guidance

Keio University
AMANO, Hideharu
hunga@am. ics. keio. ac. jp

Contents

Techniques on Parallel Processing

- Parallel Architectures
- Parallel Programming → On real machines

- Advanced uni-processor architecture

  → Special Course of Microprocessors

  (by Prof. Yamasaki, Fall term)

Class

- Lecture using Powerpoint
  - The ppt file is uploaded on the web site http://www.am.ics.keio.ac.jp, and you can down load/print before the lecture.
  - Please check it on every Friday morning.
- Homework: mail to: hunga@am.ics.keio.ac.jp

Evaluation

- Exercise on Parallel Programming using GPU (50%)
  - Caution! If the program does not run, the unit cannot be given even if you finish all other exercises.
  - This year a new GPU P100 is now under preparation.
- Homework: after every lecture (50%)

glossary 1

- 英語の単語がさっぱりわからんとのことなので用語集を付けることにする。
- このglossaryは、コンピュータ分野に限り有効である。英語一般の使い方とかなり異なる場合がある。
- Parallel: 並列の　本当に同時に動かすことを意味する。並列に動いているように見えることを含める場合をconcurrent(並行）と呼び区別する。概念的にはconcurrent > parallelである。
- Exercise: ここでは授業の最後にやる簡単な演習を指す
- GPU: Graphic ProcessingUnit　Cell Broadband Engineを使って来たが、2012年からGPUを導入した。今年は新型でより高速のを使う予定

Computer Architecture 1

Introduction to Parallel Architectures

Keio University
AMANO, Hideharu
hunga@am. ics. keio. ac. jp

Parallel Architecture

A parallel architecture consists of multiple processing units which work simultaneously.
→ Thread level parallelism

- Purposes
- Classifications
- Terms
- Trends

Boundary between Parallel machines and Uniprocessors

Uniprocessors

- ILP(Instruction Level Parallelism)
  - A single Program Counter
  - Parallelism Inside/Between instructions
- TLP(Thread Level Parallelism)
  - Multiple Program Counters
  - Parallelism between processes and jobs

Parallel Machines

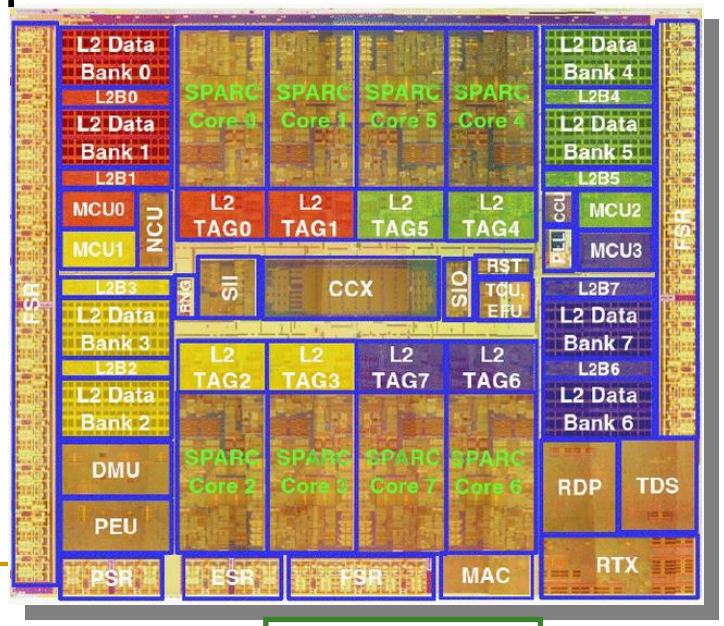
Definition
Hennessy & Petterson's
Computer Architecture: A quantitative approach

Multicore Revolution

1. The end of increasing clock frequency
   1. Consuming power becomes too much.
   2. A large wiring delay in recent processes.
   3. The gap between CPU performance and memory latency

2. The limitation of ILP
3. Since 2003, almost every computer became multi-core.

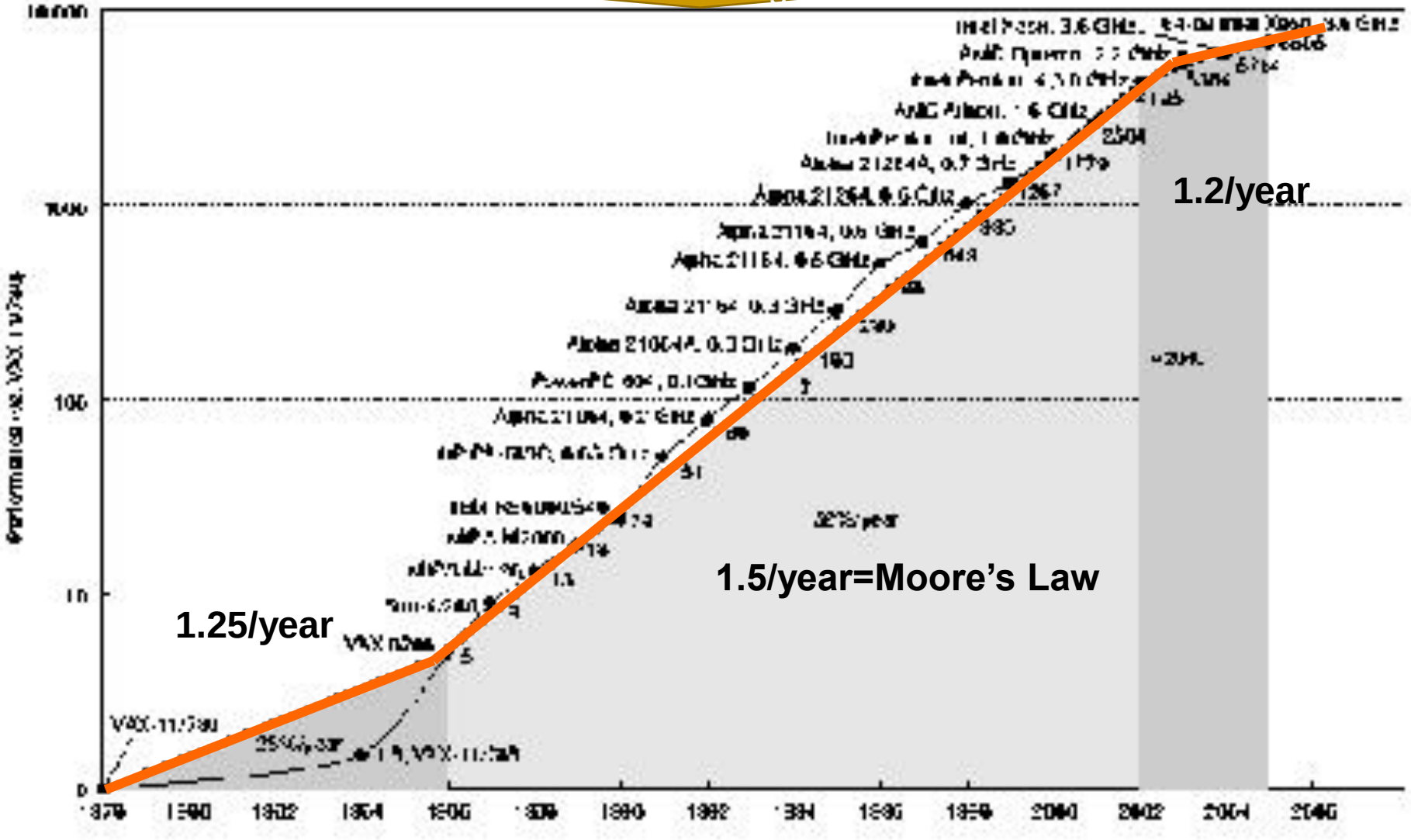
Even smartphones use 2-core/4-core CPU.



Niagara 2

End of Moore's Law in computer performance

No way to increase performance other than Increasing the number of cores



Purposes of providing multiple processors

- Performance
  - A job can be executed quickly with multiple processors
- Dependability
  - If a processing unit is damaged, total system can be available: Redundant systems
- Resource sharing
  - Multiple jobs share memory and/or I/O modules for cost effective processing: Distributed systems
- Low energy
  - High performance even with low frequency operation

Parallel Architecture: Performance Centric!

Low Power by using multiple processors

- *n* X performance with n processors,
  - but the power consumption is also n times.
  - If so, multiple processors does not contribute at all.
- $P_{dynamic} \propto Vdd^2 \times f$
- $f_{max} \propto Vdd$
- If *n* processor achieves *n* times performance, f_{max} can be *1/n*. → Vdd can be lowered. → $P_{dynamic}$ can be lowered.

Quiz

- Assume a processor which consumes 10W with 1.8V Vdd and 3GHz clock.
- You can improve performance by 10x with 10 processors, it means that the same performance can be achieved with 300MHz clock.
- In this case, Vdd can be 1.0V.
- How much power does the machine with 10 processors consume?

glossary 2

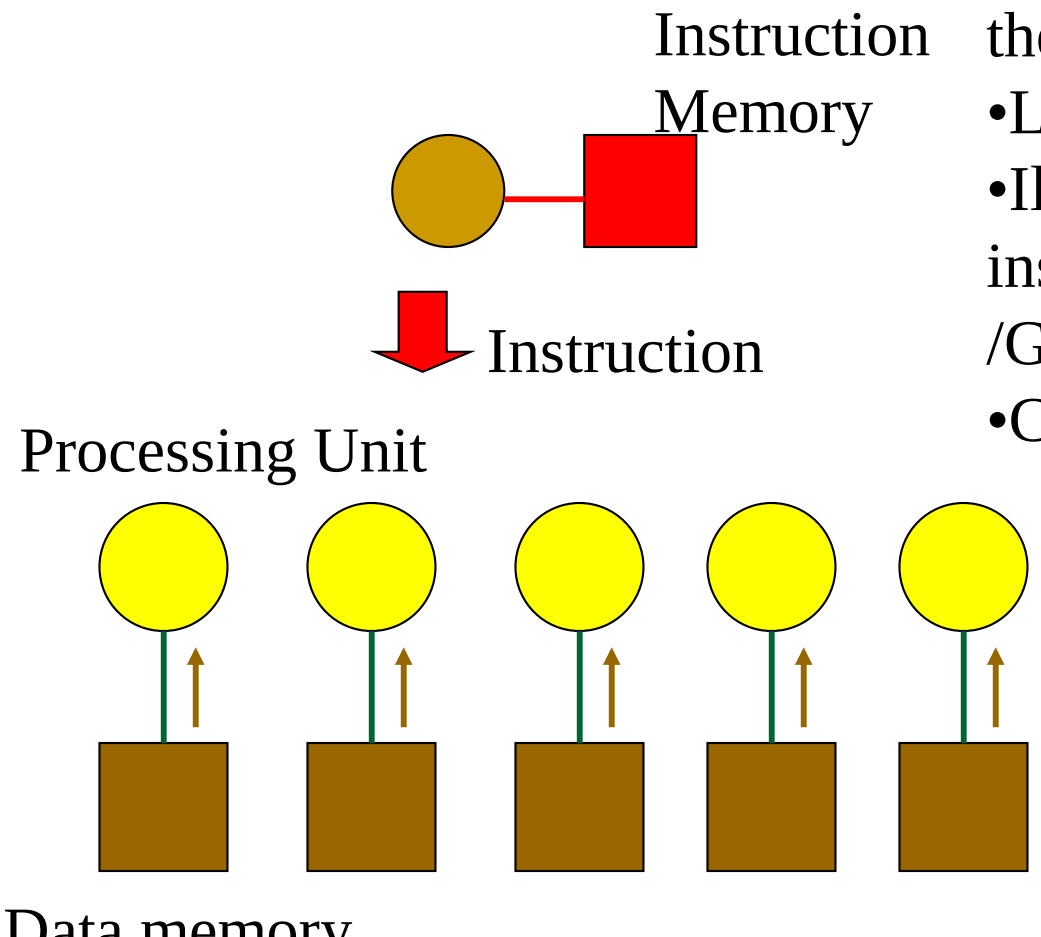
- Simultaneously: 同時に、という意味でin parallelとほとんど同じだが、ちょっとニュアンスが違う。in parallelだと同じようなことを同時にやる感じがするが、simultaneouslyだととにかく同時にやればよい感じがする。
- Thread: プログラムの一連の流れのこと。Thread level parallelism（TLP)は、Thread間の並列性のことで、ここではHennessy and Pattersonのテキストに従ってPCが独立している場合に使うが違った意味に使う人も居る。これに対してPCが単一で命令間にある並列性をILPと呼ぶ
- Dependability: 耐故障性、Reliability(信頼性）, Availability(可用性）双方を含み、要するに故障に強いこと。Redundant systemは冗長システムのことで、多めに資源を持つことで耐故障性を上げることができる。
- Distributed system:分散システム、分散して処理することにより効率的に処理をしたり耐故障性を上げたりする

Flynn's Classification

- The number of Instruction Stream: M(Multiple)/S(Single)
- The number of Data Stream: M/S
  - SISD
    - Uniprocessors (including Super scalar、VLIW)
  - MISD: Not existing (Analog Computer)
  - SIMD
  - MIMD

He gave a lecture at Keio in the last year

SIMD (Single Instruction Stream Multiple Data Streams



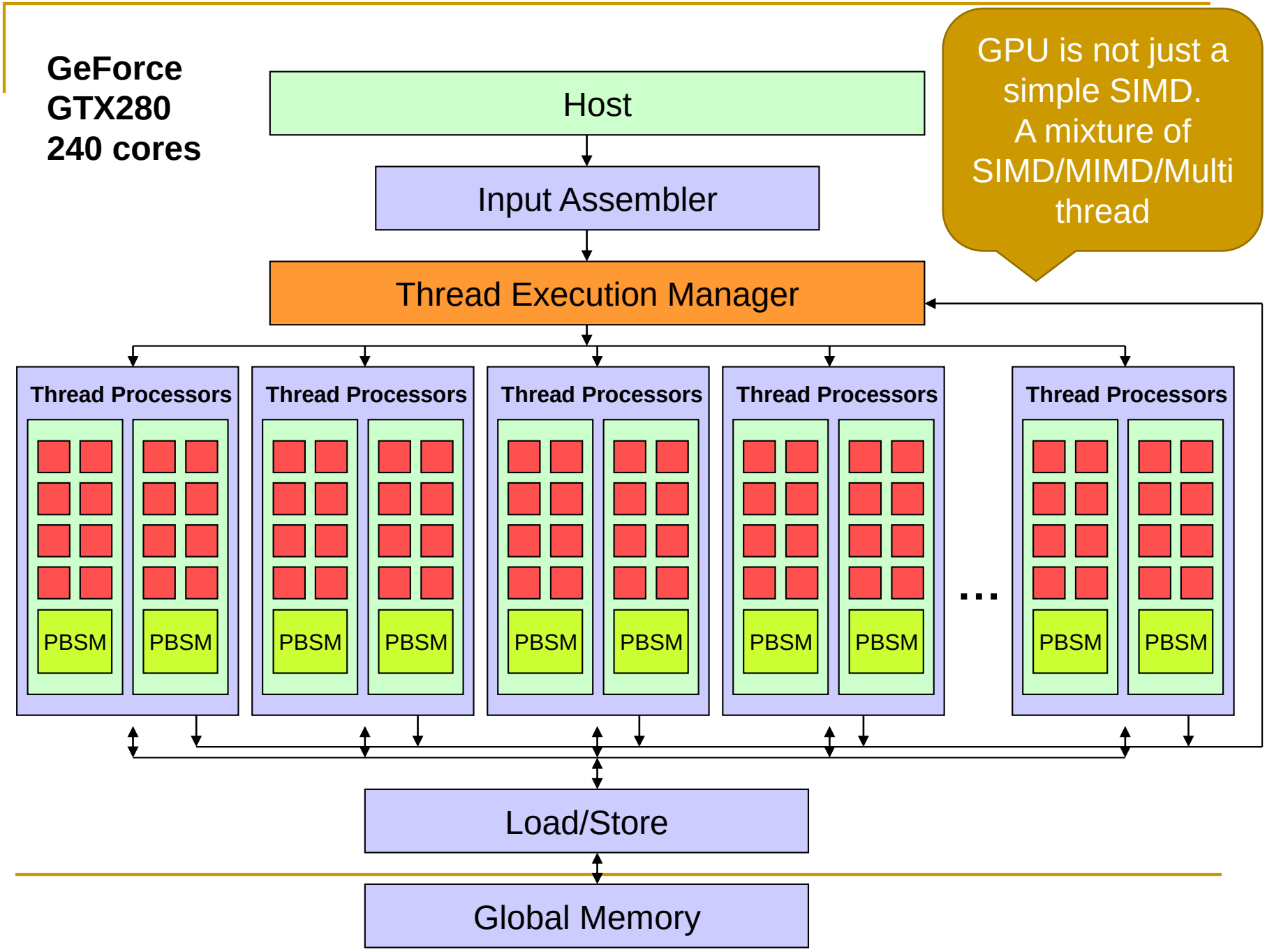
- All Processing Units executes the same instruction
- Low degree of flexibility
- Illiac-IV/MMX instructions/ClearSpeed/IMAP /GP-GPU（coarse grain）
- CM-2,（fine grain）

Two types of SIMD

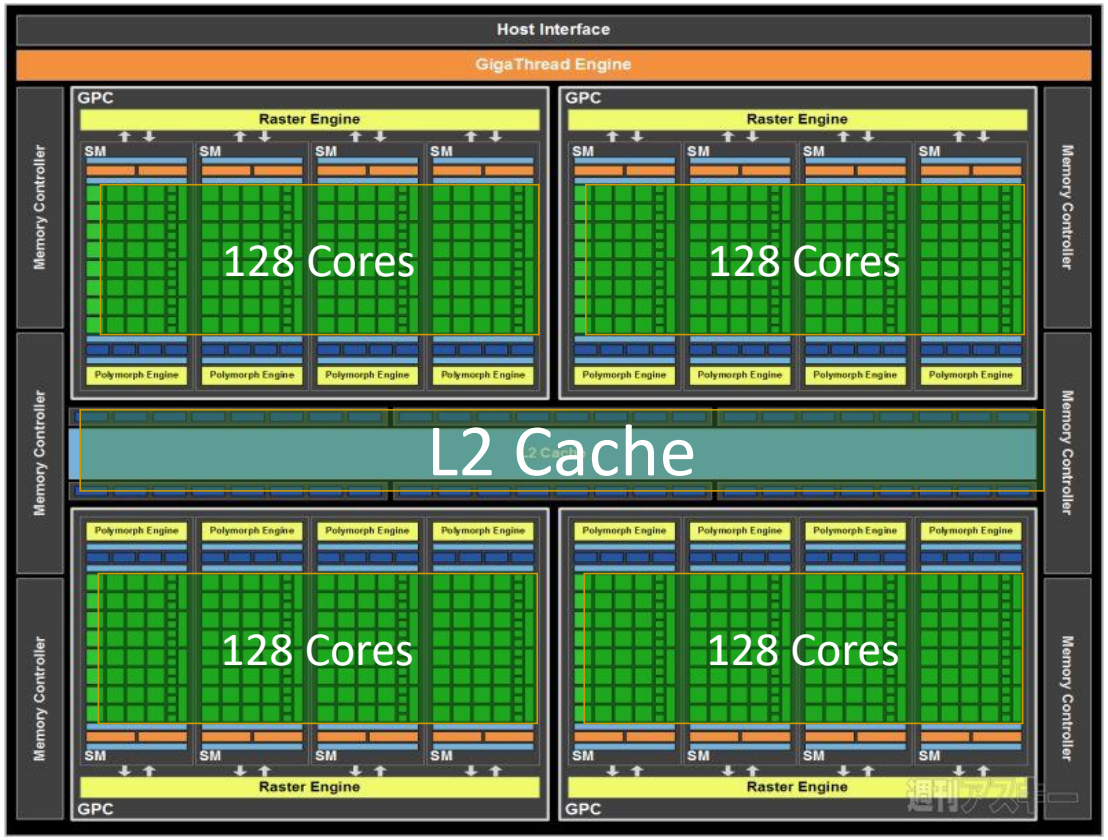
- Coarse grain：Each node performs floating point numerical operations
  - Old SuperComputers: ILLIAC-IV，BSP,GF-11
  - Multimedia instructions in recent high-end CPUs
  - Accelerator: GPU, ClearSpeed
  - Dedicated on-chip approach: NEC's IMAP
- Fine grain：Each node only performs a few bits operations
  - ICL DAP，CM-2，MP-2
  - Image/Signal Processing
  - Connection Machines （CM-2） extends the application to Artificial Intelligence (CmLisp)

GPGPU(General-Purpose computing on Graphic ProcessingUnit)

- Titan (NVIDIA K20X, 3rd place of Top500)
- TSUBAME2.5(NVIDIA K20X)
- A lot of supercomputers in Top500 use GPU.


GeForce
GTX280
240 cores
Host
Input Assembler
Thread Execution Manager
GPU is not just a simple SIMD. A mixture of SIMD/MIMD/Multi thread
Thread Processors
PBSM
...
Load/Store
Global Memory

GPU(NVIDIA's GTX580)



512 GPU cores (128 X 4)

768 KB L2 cache

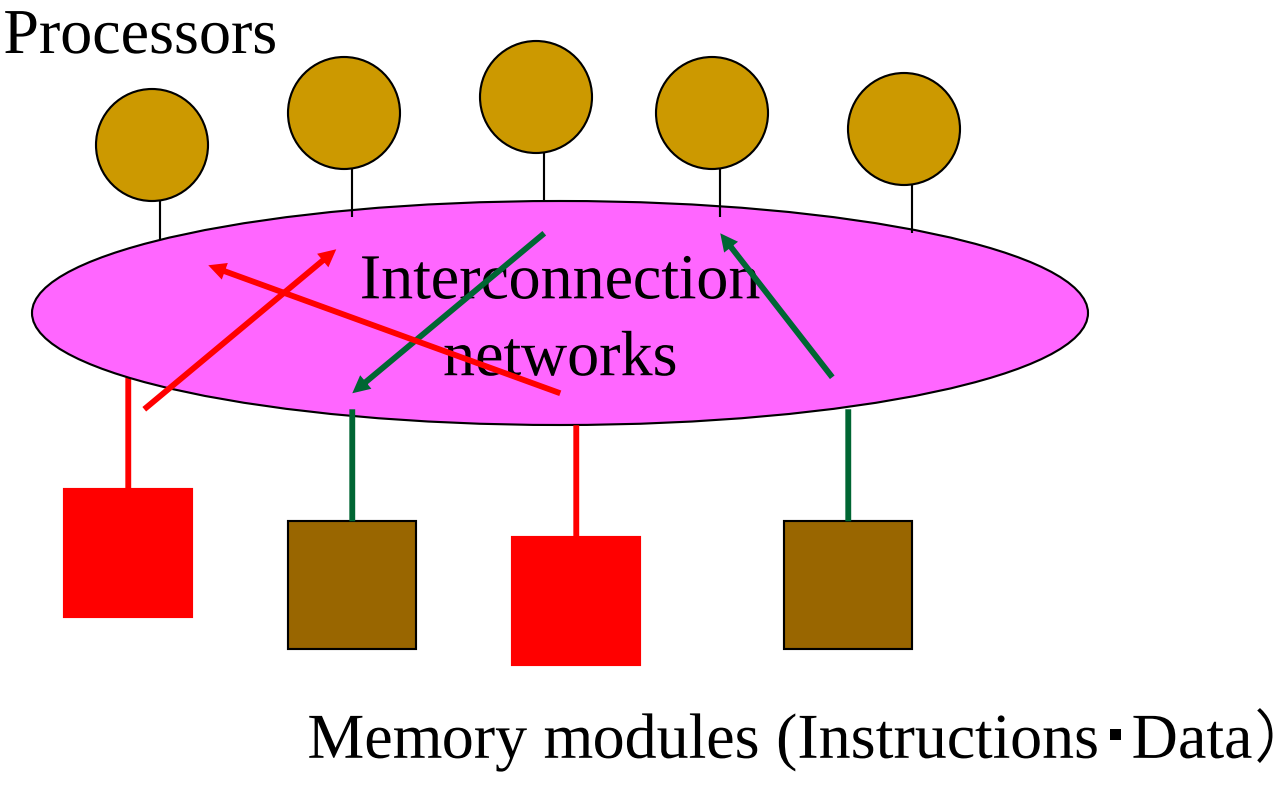
40nm CMOS 550 mm^2

The future of SIMD

- Coarse grain SIMD
  - GPGPU became a main stream of accelerators.
  - Other SIMD accelerators are hard to be survive.
  - Multi-media instructions will have been used in the future.
- Fine grain SIMD
  - Advantageous to specific applications like image processing
  - On-chip accelerator

MIMD

- Each processor executes individual instructions
- Synchronization is required
- High degree of flexibility
- Various structures are possible



Classification of MIMD machines

Structure of shared memory

- UMA(Uniform Memory Access Model)

  provides shared memory which can be accessed from all processors with the same manner.

- NUMA(Non-Uniform Memory Access Model)

  provides shared memory but not uniformly accessed.

- NORA/NORMA(No Remote Memory Access Model)

  provides no shared memory. Communication is done with message passing.

UMA

- The simplest structure of shared memory machine
- The extension of uniprocessors
- OS which is an extension for single processor can be used.
- Programming is easy.
- System size is limited.
  - Bus connected
  - Switch connected
- A total system can be implemented on a single chip

On-chip multiprocessor

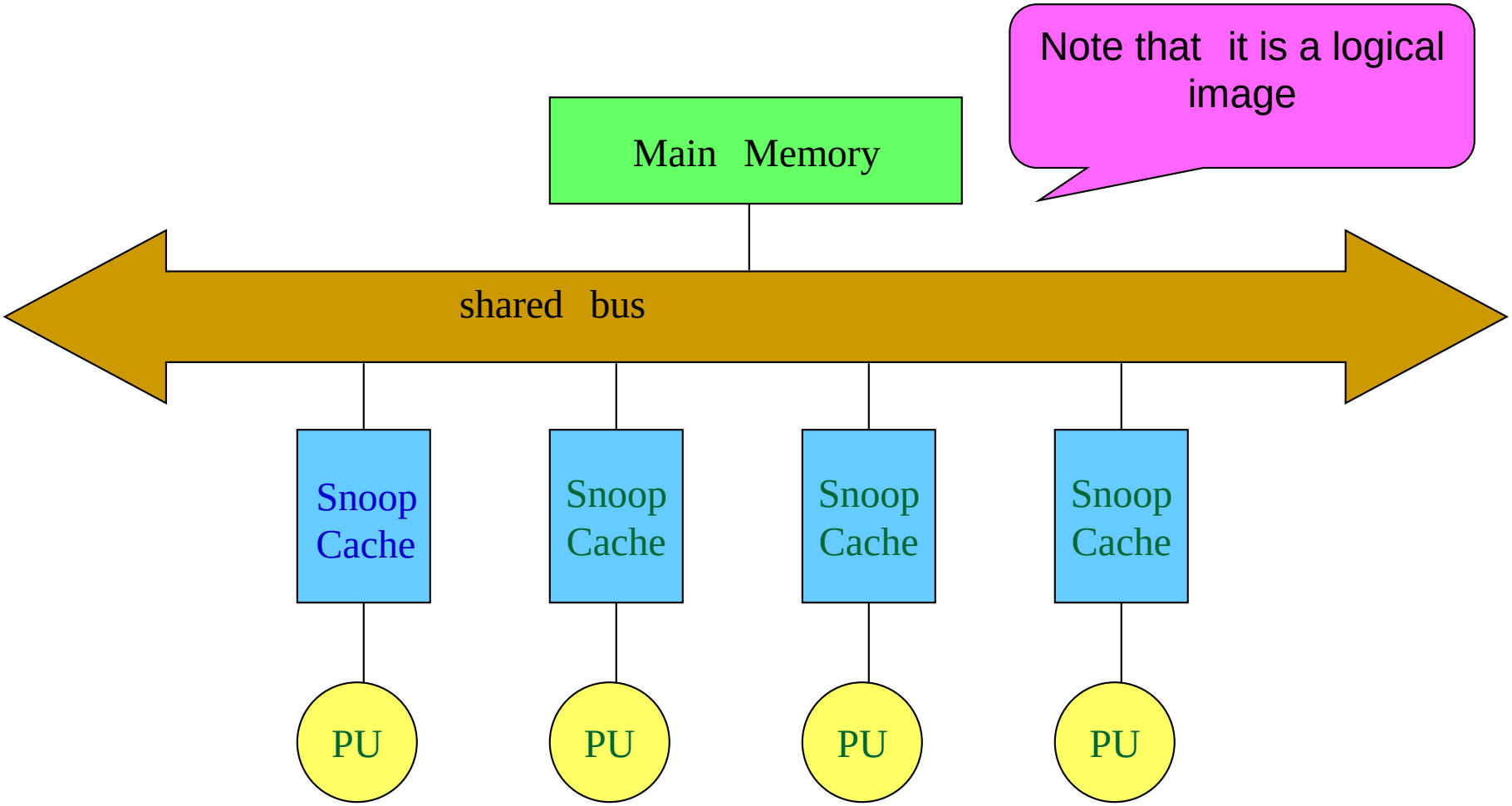
Chip multiprocessor

Single chip multiprocessor→ Multicore

IBM Power series

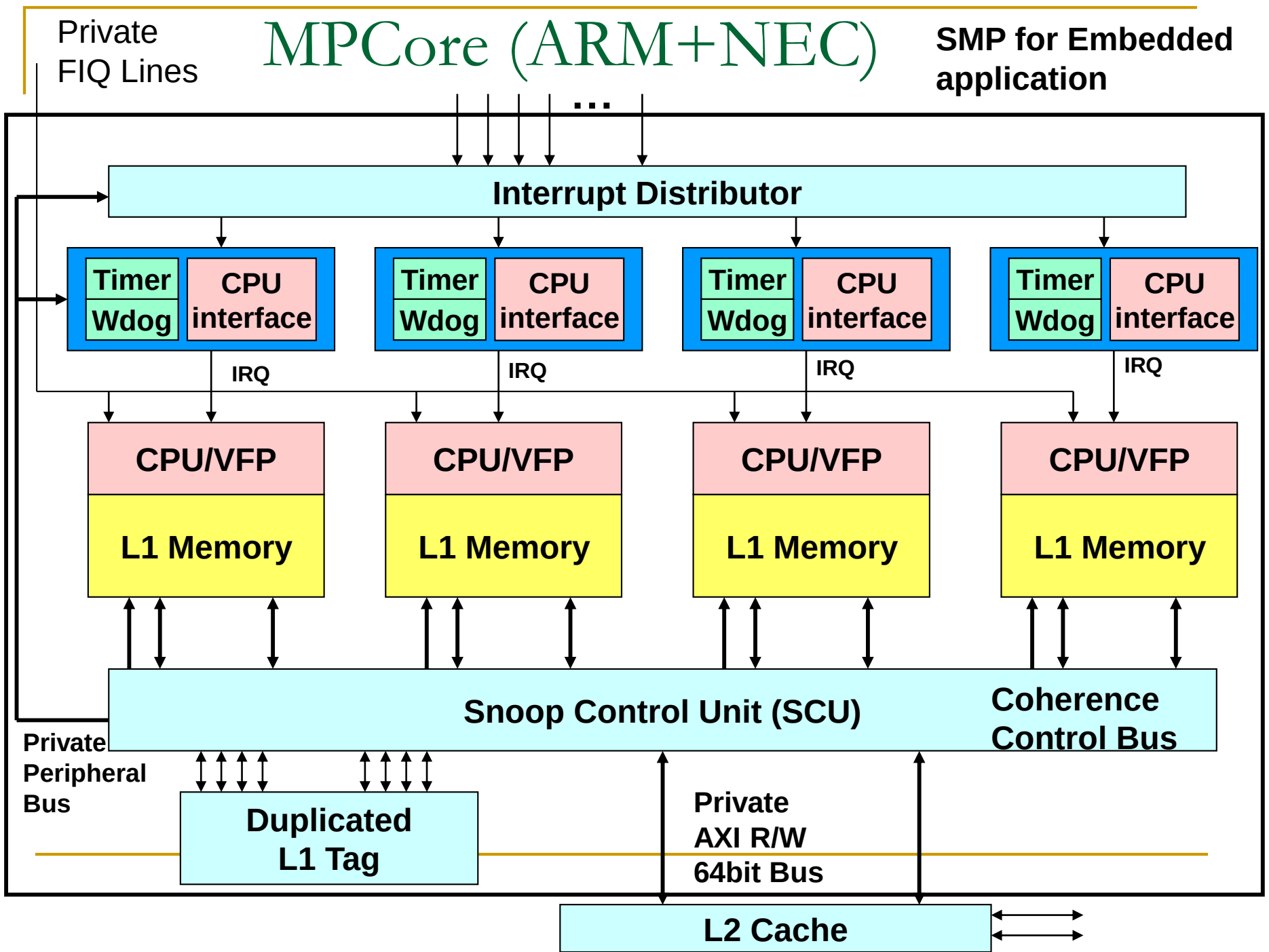
NEC/ARM chip multiprocessor for embedded systems

An example of UMA : Bus connected

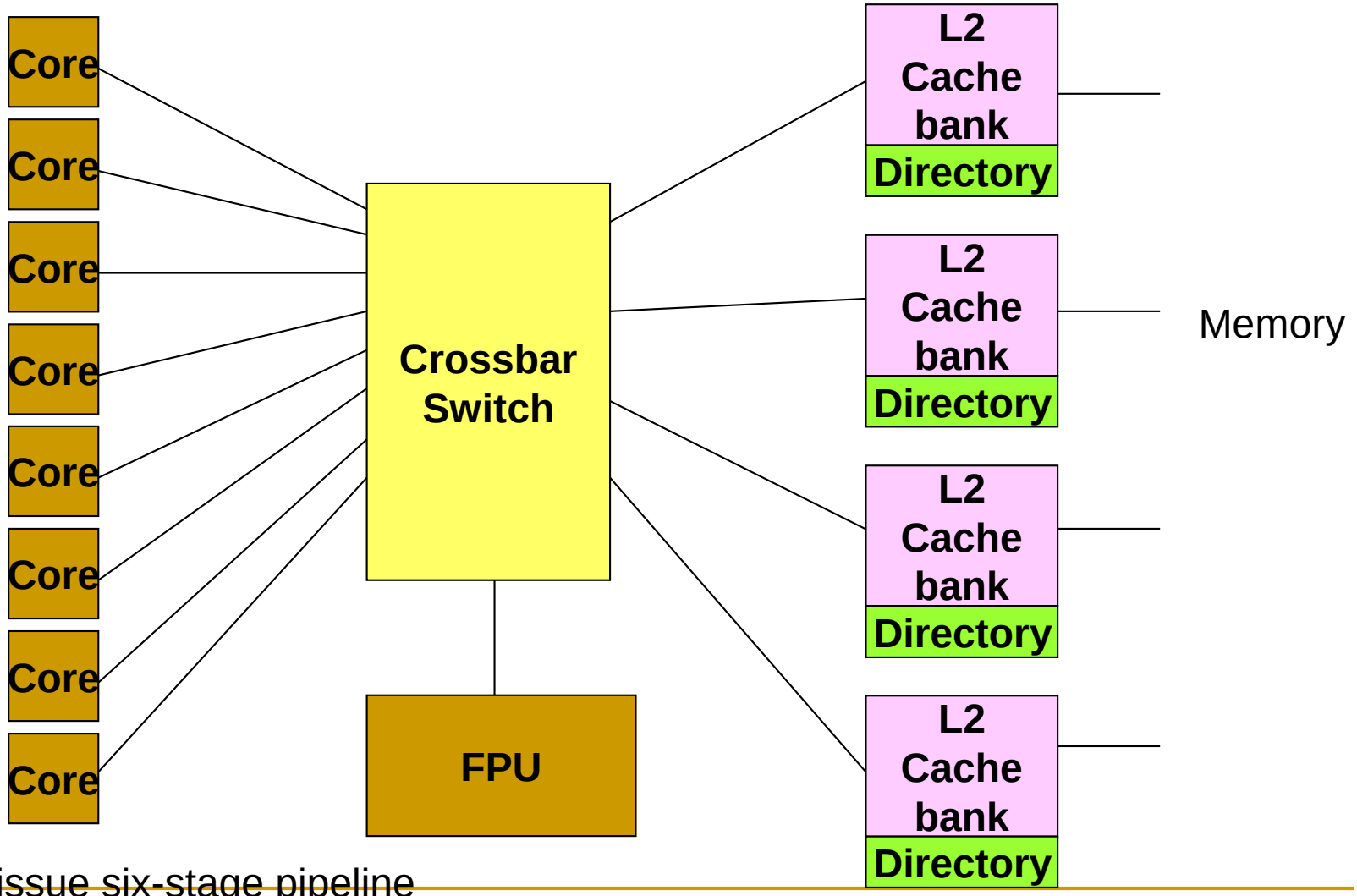


SMP (Symmetric MultiProcessor),

On chip multiprocessor or multicore


Private
FIQ Lines
MPCore (ARM+NEC)
SMP for Embedded application
...
Interrupt Distributor
Timer
Wdog
CPU interface
IRQ
CPU/VFP
L1 Memory
Snoop Control Unit (SCU)
Coherence Control Bus
Private Peripheral Bus
Duplicated L1 Tag
Private AXI R/W 64bit Bus
L2 Cache

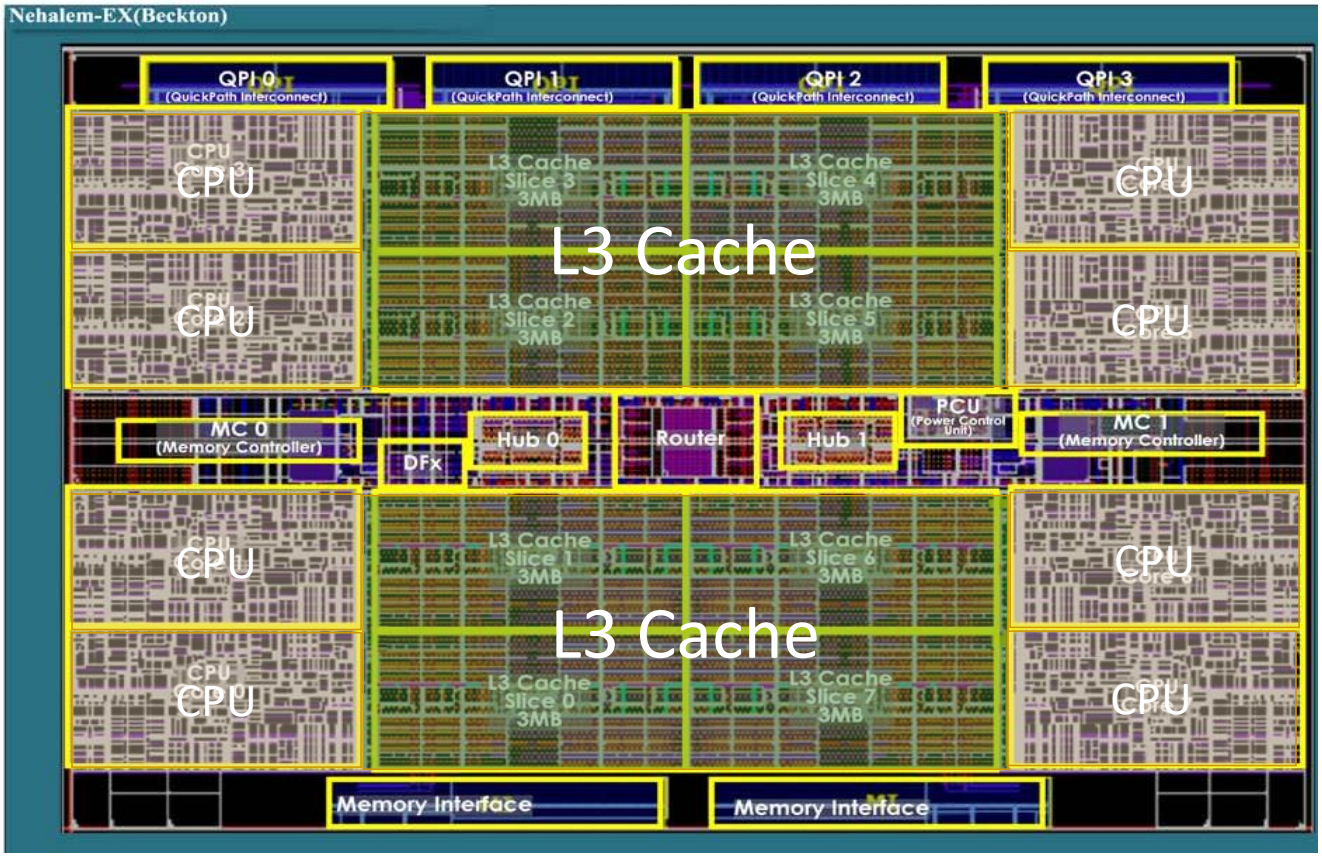
SUN T1



Single issue six-stage pipeline RISC with 16KB Instruction cache/ 8KB Data cache for L1

Total 3MB, 64byte Interleaved

Multi-Core (Intel's Nehalem-EX)



8 CPU cores

24MB L3 cache

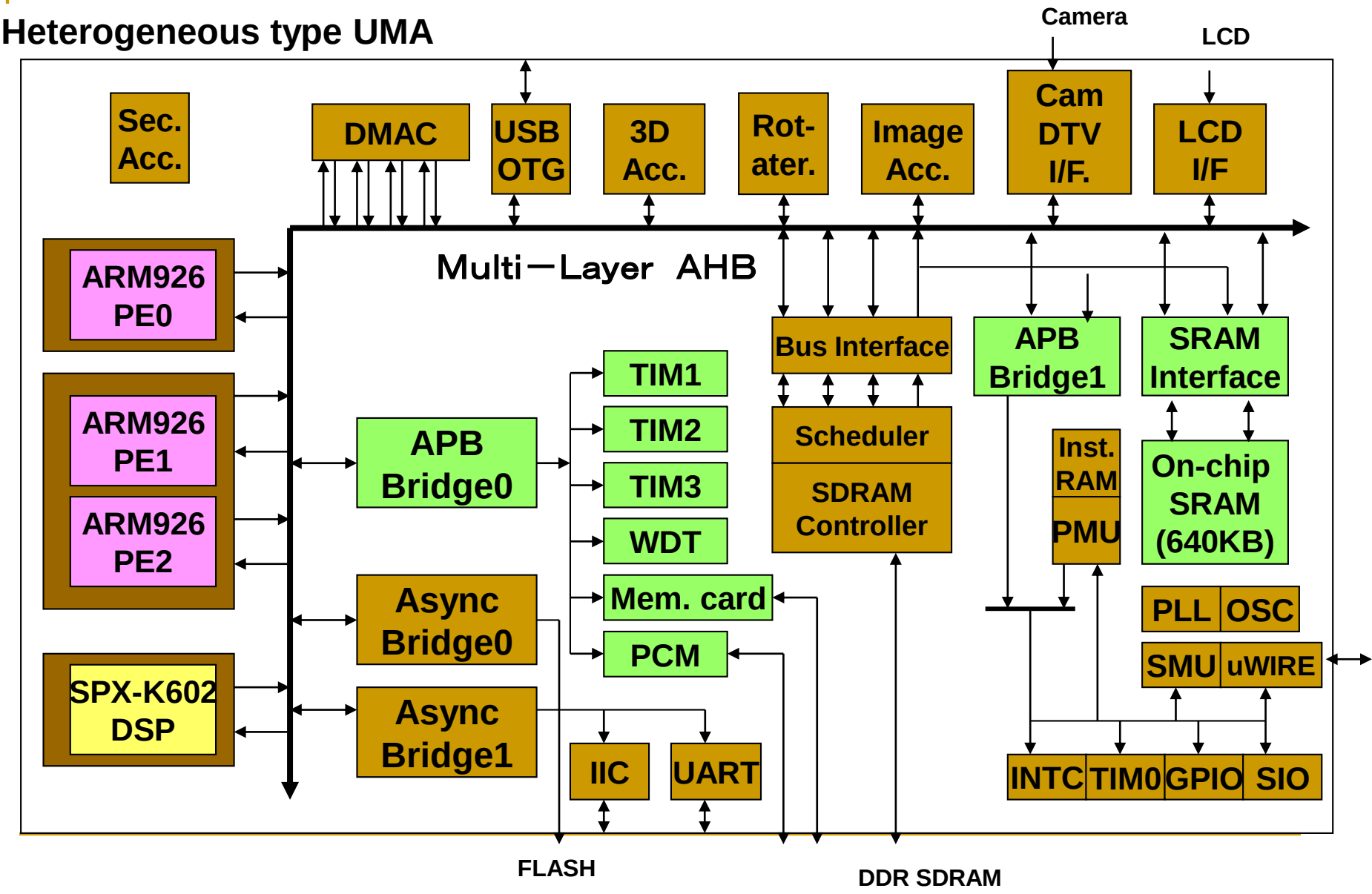
45nm CMOS 600 mm^2

Heterogeneous vs. Homogeneous

- Homogeneous: consisting of the same processing elements
  - A single task can be easily executed in parallel.
  - Unique programming environment
- Heterogeneous: consisting of various types of processing elements
  - Mainly for task-level parallel processing
  - High performance per cost
  - Most recent high-end processors for cellular phone use this structure
  - However, programming is difficult.

NEC MP211

Heterogeneous type UMA

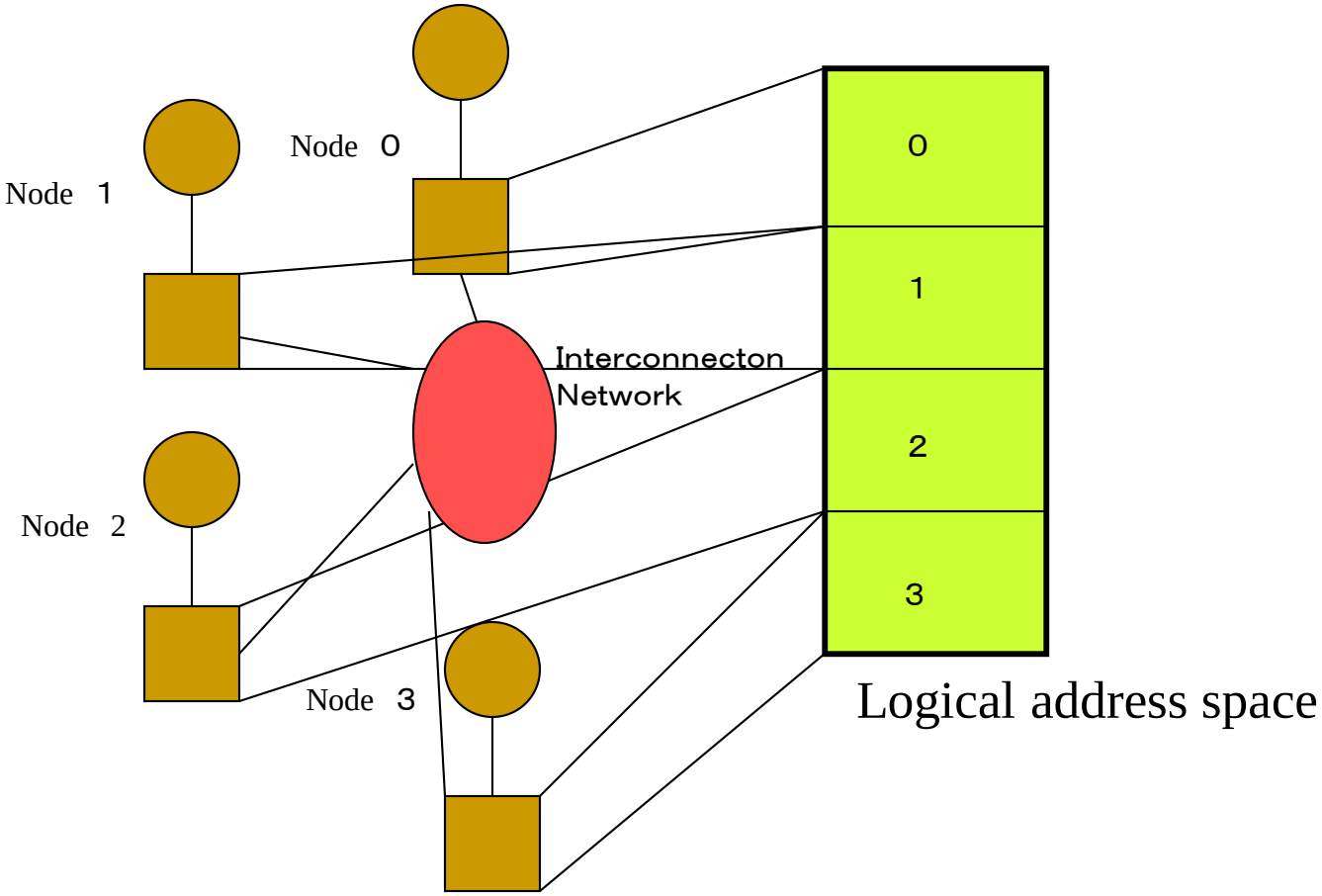


NUMA

- Each processor provides a local memory, and accesses other processors' memory through the network.
- Address translation and cache control often make the hardware structure complicated.
- Scalable:
  - Programs for UMA can run without modification.
  - The performance is improved as the system size.

Competitive to WS/PC clusters with Software DSM

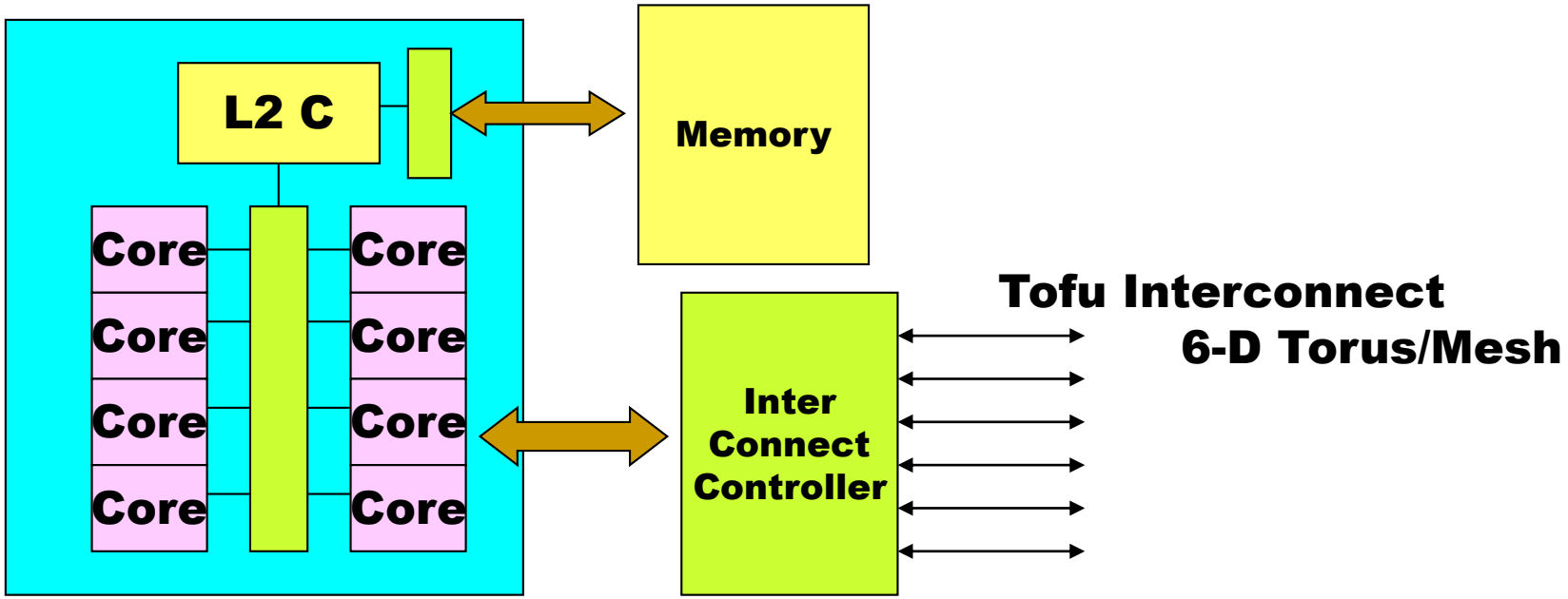
Typical structure of NUMA



Classification of NUMA

- Simple NUMA：
  - Remote memory is not cached.
  - Simple structure but access cost of remote memory is large.
- CC-NUMA：Cache Coherent
  - Cache consistency is maintained with hardware.
  - The structure tends to be complicated.
- COMA:Cache Only Memory Architecture
  - No home memory
  - Complicated control mechanism

Supercomputer「K」



SPARC64 VIIIfx Chip

4 nodes/board

24boards/Lack

96nodes/Lack

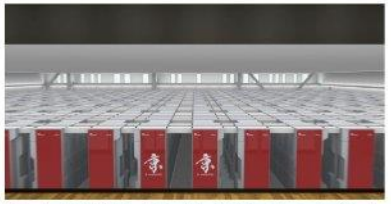
RDMA mechanism
NUMA or UMA+NORMA

ファイル(F) 編集(E) 表示(V) 移動(G) ヘルプ(H)

「京」のシステム構成

システム全体

計算ラック×864



1京回/秒≒10.6ペタフロップス
1PB以上

計算ラック群

計算ラック×8
ディスクラック×2

98.4兆回/秒
12TB

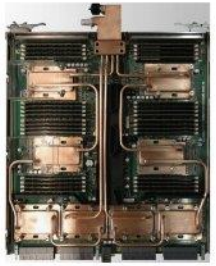
計算ラック

システムボード×24
IOシステムボード×6

12.3兆回/秒
1.5TB

システムボード

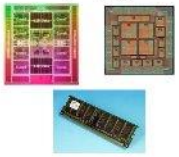
ノード×4



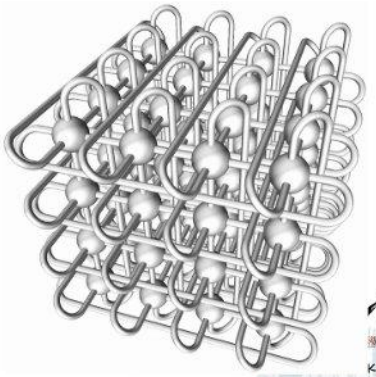
5120億回/秒
64GB

ノード

CPU×1
ICC×1
メモリ



演算性能:1280億回/秒
メモリ容量:16GB

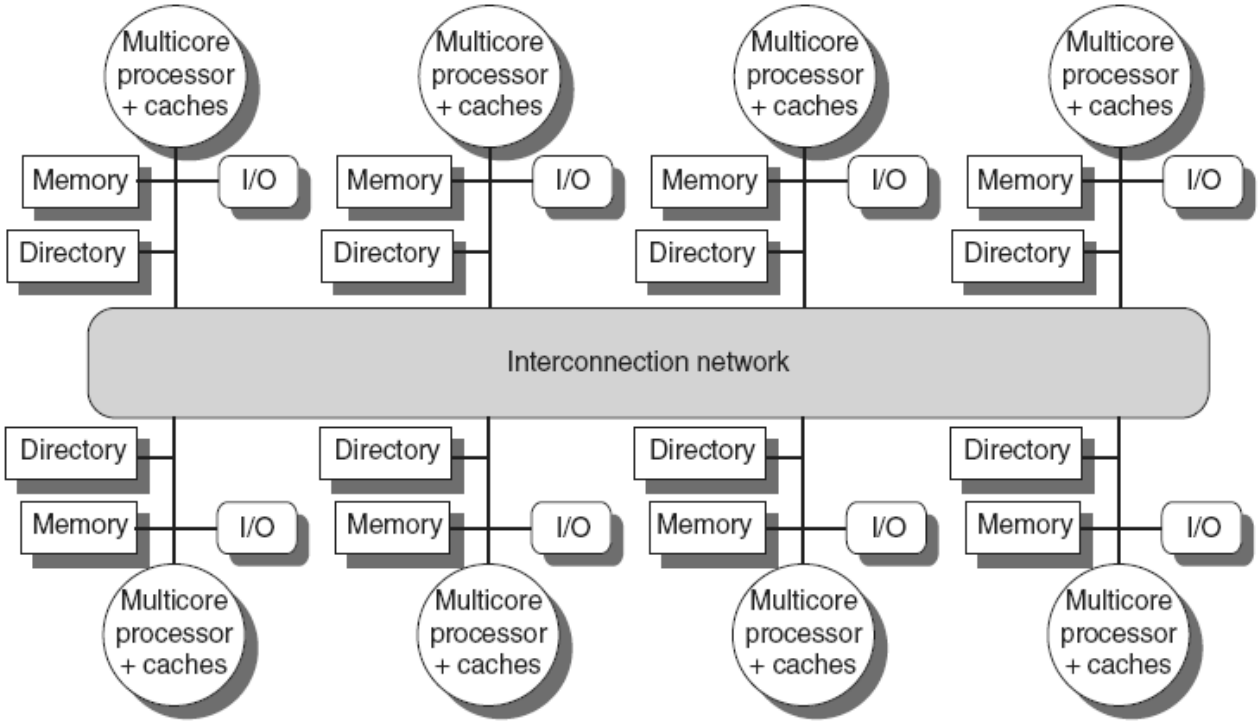


SACSIS2012 Invited speech

Multicore Based systems

- Implementing in shared L3 cache
  - Keep bit vector of size = # cores for each block in L3
  - Not scalable beyond shared L3

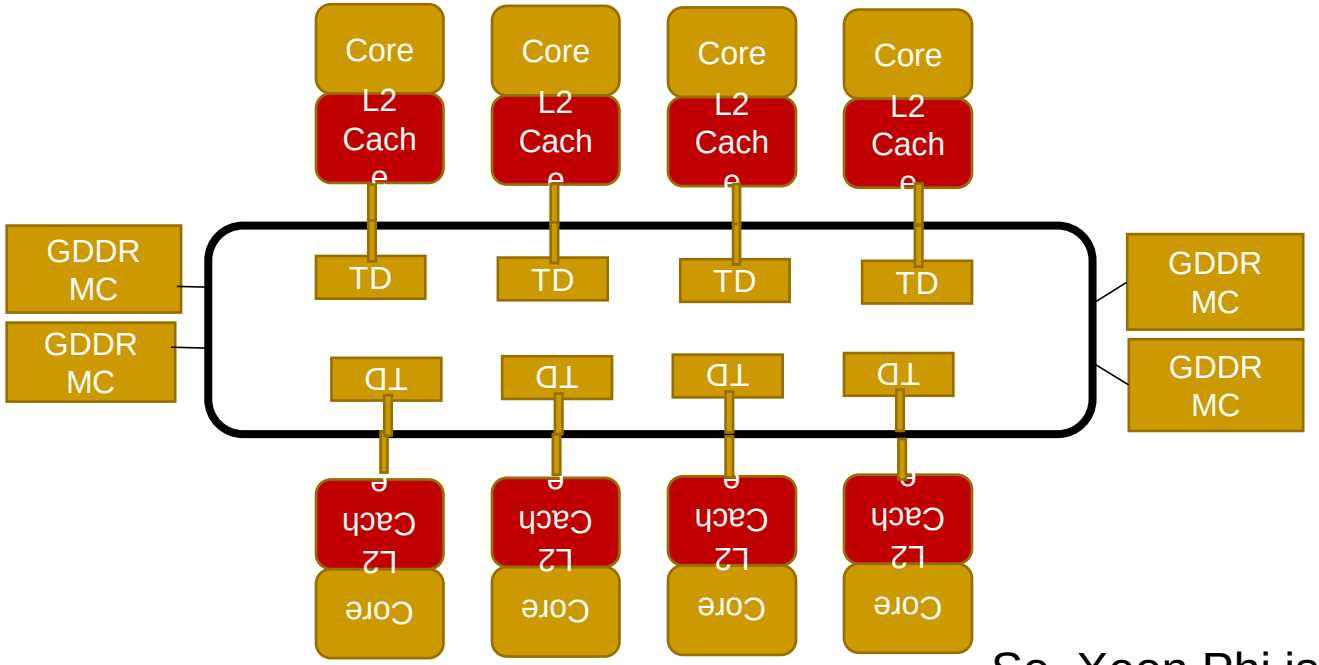
IBM Power 7
AMD Opteron 8430



Copyright © 2012, Elsevier Inc. All rights reserved.

Xeon Phi Microarchitecture

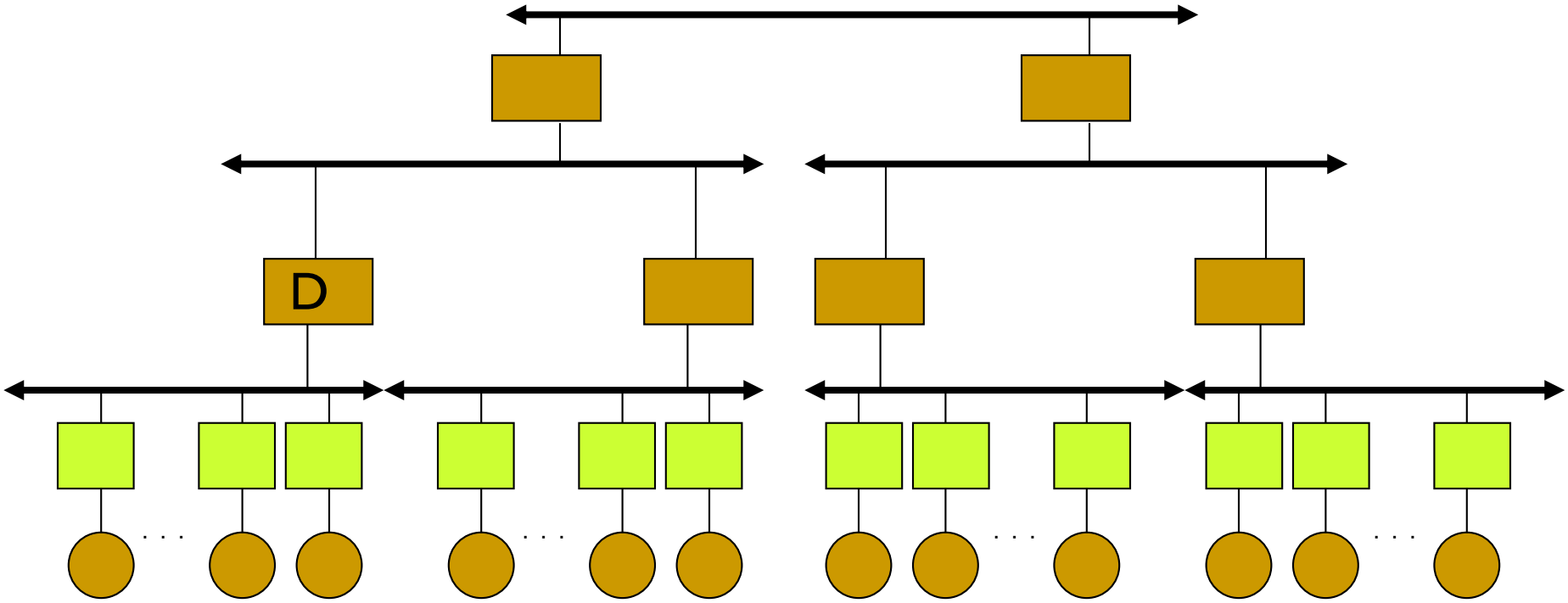
All cores are connected through the ring interconnect.
All L2 caches are coherent with directory based management.



So, Xeon Phi is classified into CC (Cache Coherent) NUMA.

Of course, all cores are multithreaded, and provide 512 SIMD instructions.

DDM(Data Diffusion Machine)



NORA/NORMA

- No shared memory
- Communication is done with message passing
- Simple structure but high peak performance

Cost effective solution.

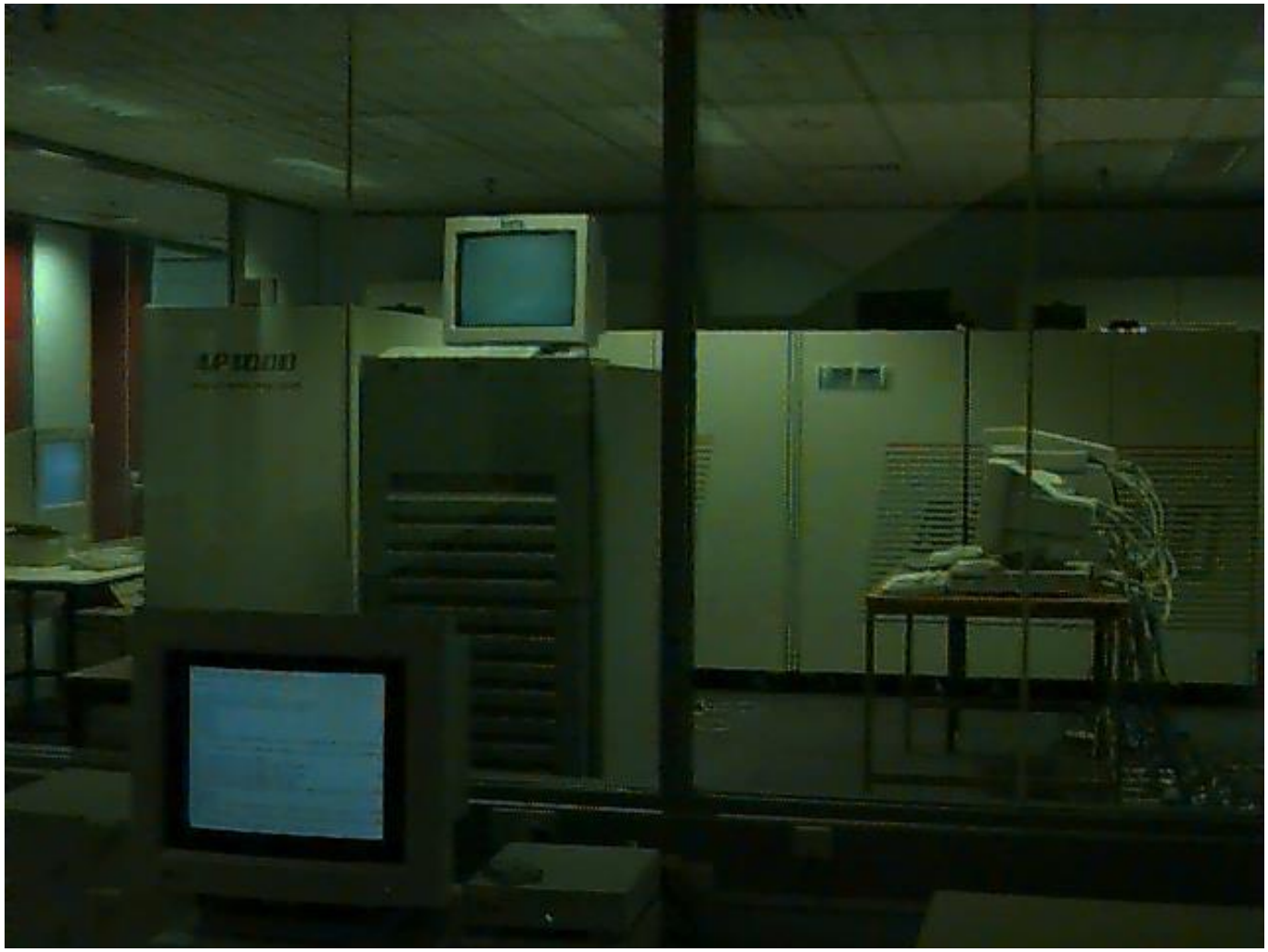
Hard for programming

Inter-PU communications

Cluster computing

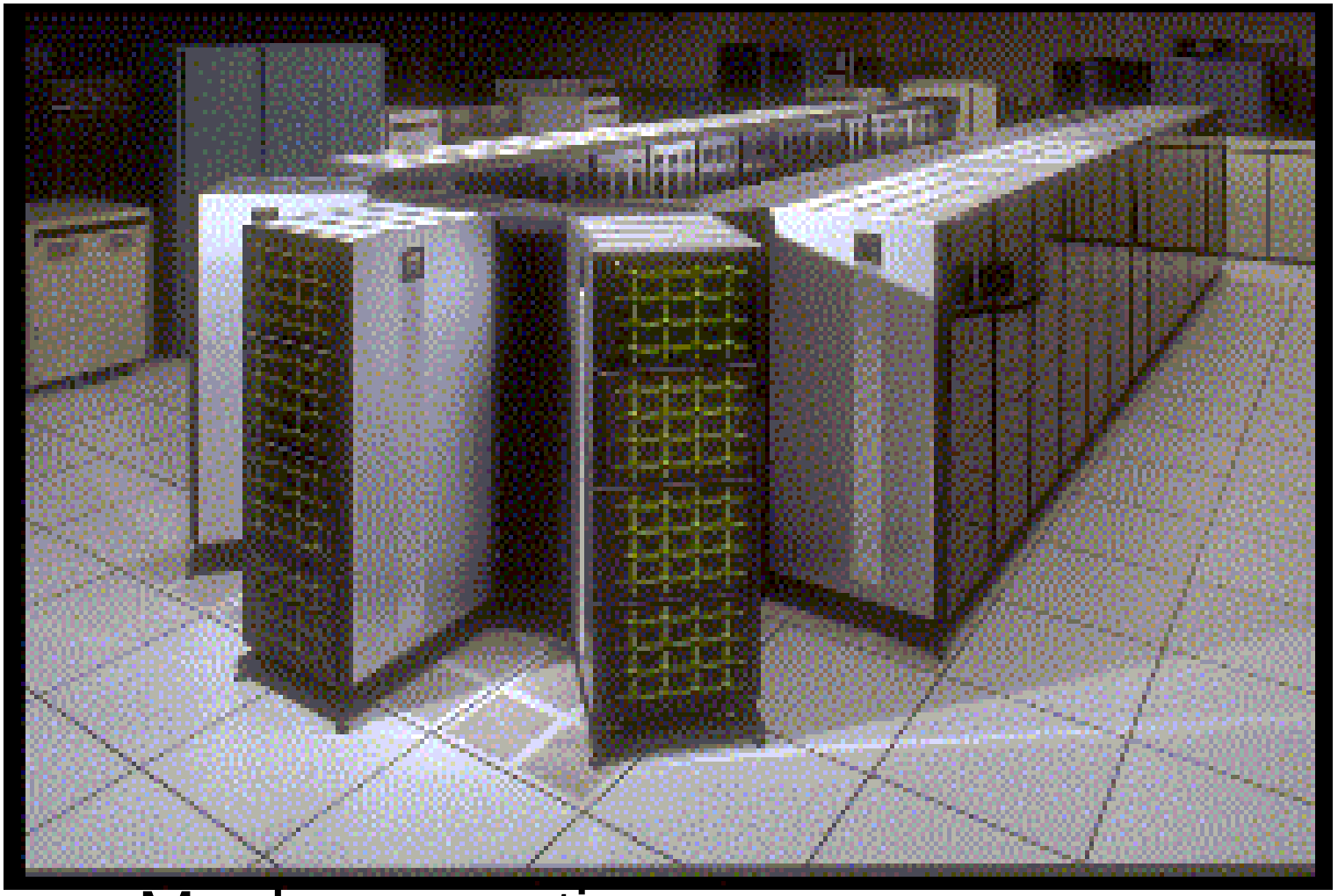
Early Hypercube machine nCUBE2

Fujitsu's NORA AP1000(1990)



- Mesh connection
- SPARC

Intel's Paragon XP/S(1991)

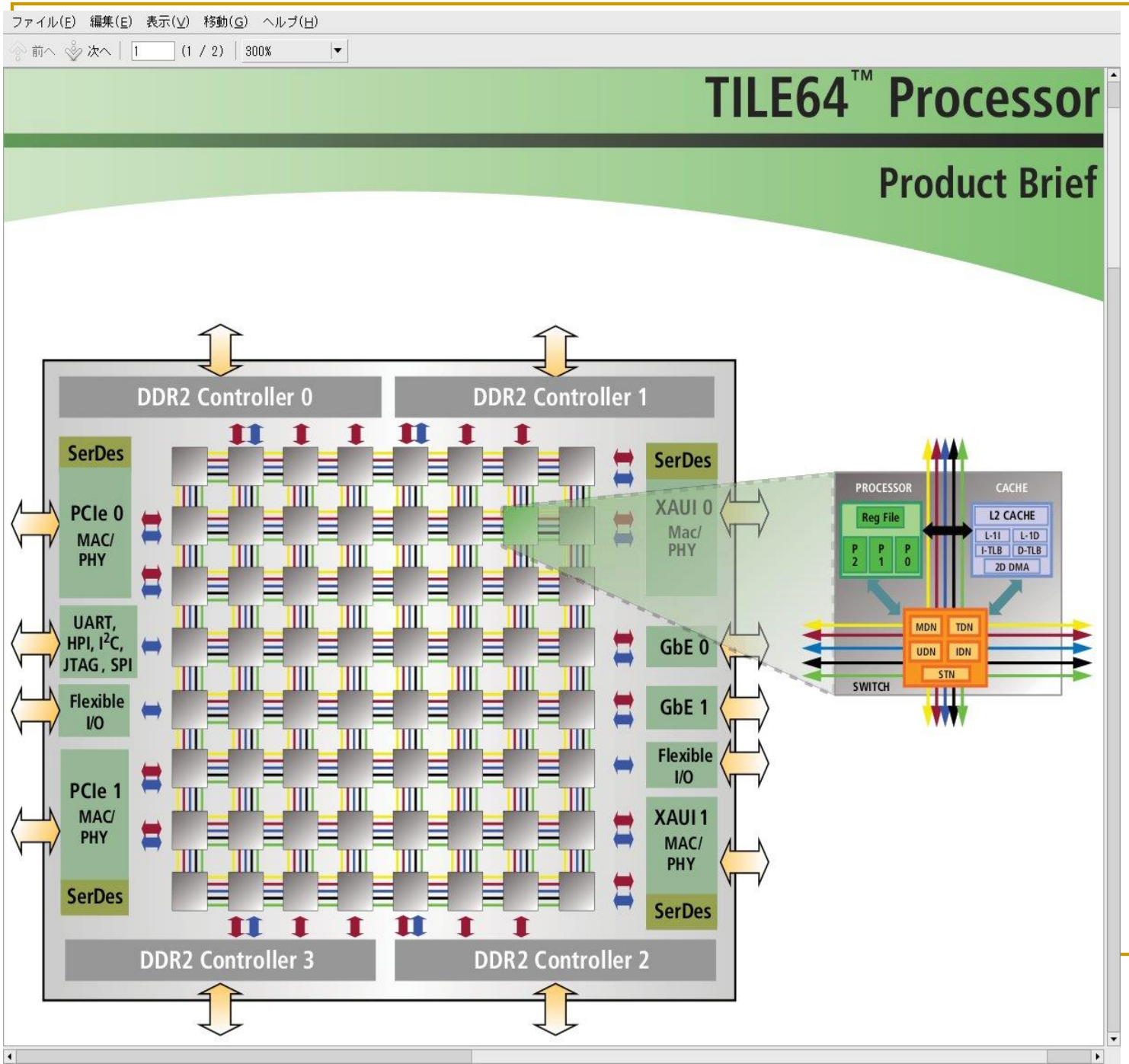


- Mesh connection
- i860

PC Cluster

- Beowulf Cluster (NASA's Beowulf Projects 1994, by Sterling)
  - Commodity components
  - TCP/IP
  - Free software
- Others
  - Commodity components
  - High performance networks like Myrinet / Infiniband
  - Dedicated software


RHiNET-2 cluster



Tilera's Tile64

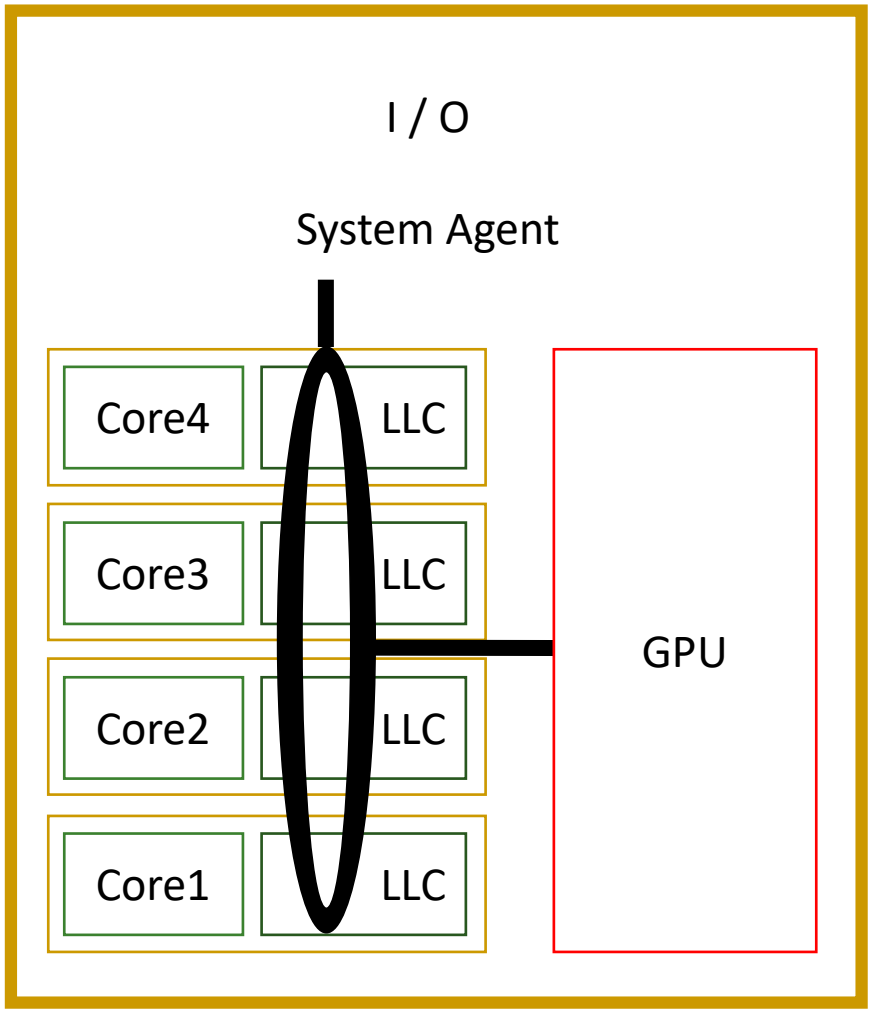
Tile Pro, Tile Gx

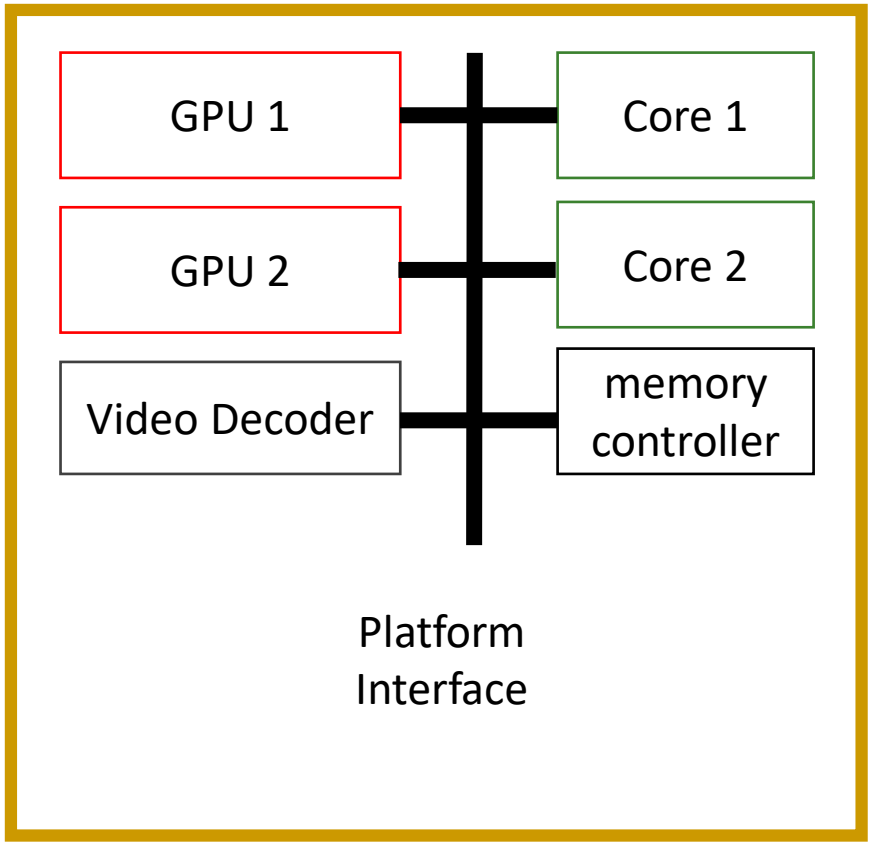
Linux runs in each core.

All techniques are combined

- Nodes with CPU (Multi-core) are connected with NORA/NORMA
  - Clusters in data-centers.
- Nodes with CPUs(Multi-core)+GPUs(SIMD/many-core) are connected with NORA/NORMA
  - Tsubame (TIT) and other supercomputers
- Nodes with Multi-core are connected with NUMA
  - K-supercomputer

Multi-core + Accelerator





Intel's Sandy Bridge

AMD's Fusion APU

glossary 3

- Flynn's Classification: Flynn(Stanford大の教授）が論文中に用いた分類、内容は本文を参照のこと
- Coarse grain：粗粒度、この場合はプロセッシングエレメントが浮動小数演算が可能な程度大きいこと。反対がFine grain（細粒度）で、数ビットの演算しかできないもの
- Illiac-IV, BSP, GF-11, Connection Machine　CM-2, MP-2などはマシン名。SIMDの往年の名機
- Synchronization:同期、Shared Memory:共有メモリ、この辺は後の授業で詳細を解説する
- Message passing:メッセージ交換。共有メモリを使わずにデータを直接交換する方法
- Embedded System:組み込みシステム
- Homogeneous:等質な　Heterogeneous：性質の異なったものから成る
- Coherent Cache:内容の一貫性が保障されたキャッシュ、Cache Consistencyは内容の一貫性、これも後の授業で解説する
- Commodity　Component:　標準部品、価格が安く入手が容易
- Power 5, Origin2000, Cray XD-1,AP1000,NCUBE などもマシン名。The earth simulatorは地球シミュレータ,IBM BlueGene/Lは現在のところ最速

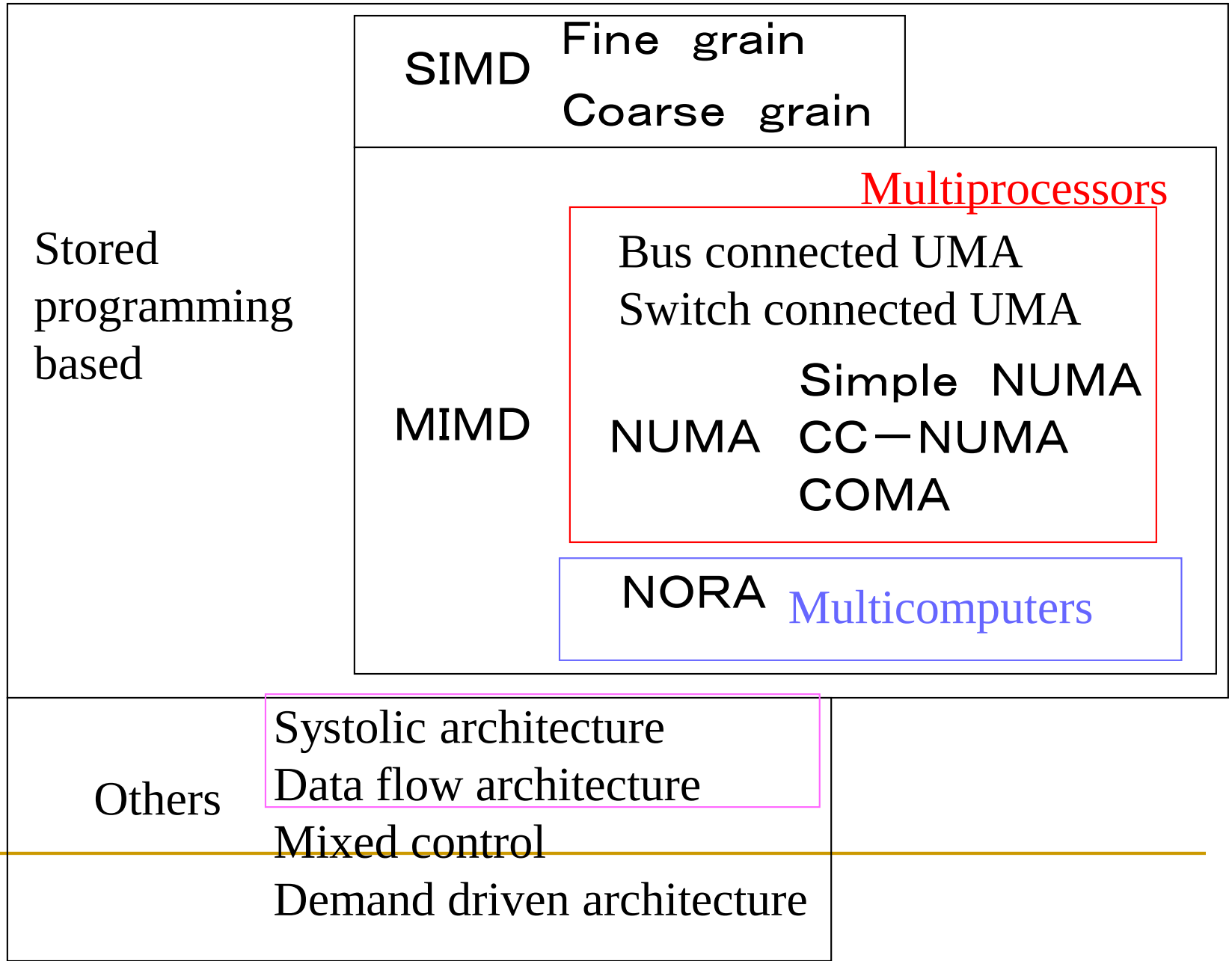
Terms(1)

- Multiprocessors：
  - MIMD machines with shared memory
  - （Strict definition：by Enslow Jr.）
    - Shared memory
    - Shared I/O
    - Distributed OS
    - Homogeneous
  - Extended definition: All parallel machines（Wrong usage）
- Multicomputer
  - MIMD machines without shared memory, that is NORA／NORMA

Term(2)

- Multicore
  - On-chip multiprocessor.
  - Mostly UMA.
  - Symmetric Multi-Processor SMP
    - Historically, SMP is used for multi-chip multiprocessor
- Manycore
  - On-chip multiprocessor with a lot of cores
  - GPUs are also referred as “Manycore”.

Classification



Exercise 1

- Pezy's supercomputer Gyoukou got the 4th place at TOP500 in 2017 November.
- But, the president of Pezy was arrested for crime of fraud later.
- It uses Pezy SC2 as an accelerator.
- Which type is Pezy SC2 classified into ?
- If you take this class, send the answer with your name and student number to hunga@am.ics.keio.ac.jp
- You can use either Japanese or English.