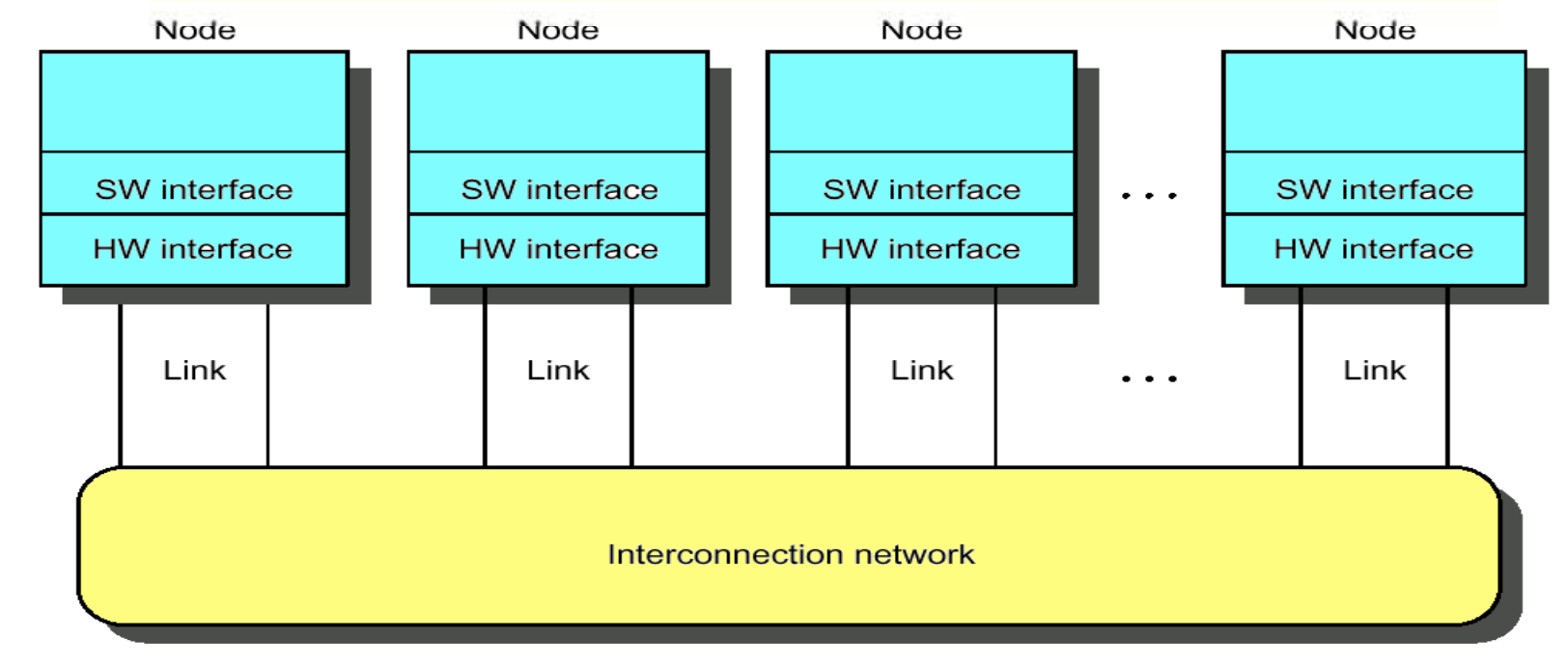
CMSC 611: Advanced Computer Architecture

Parallel Systems

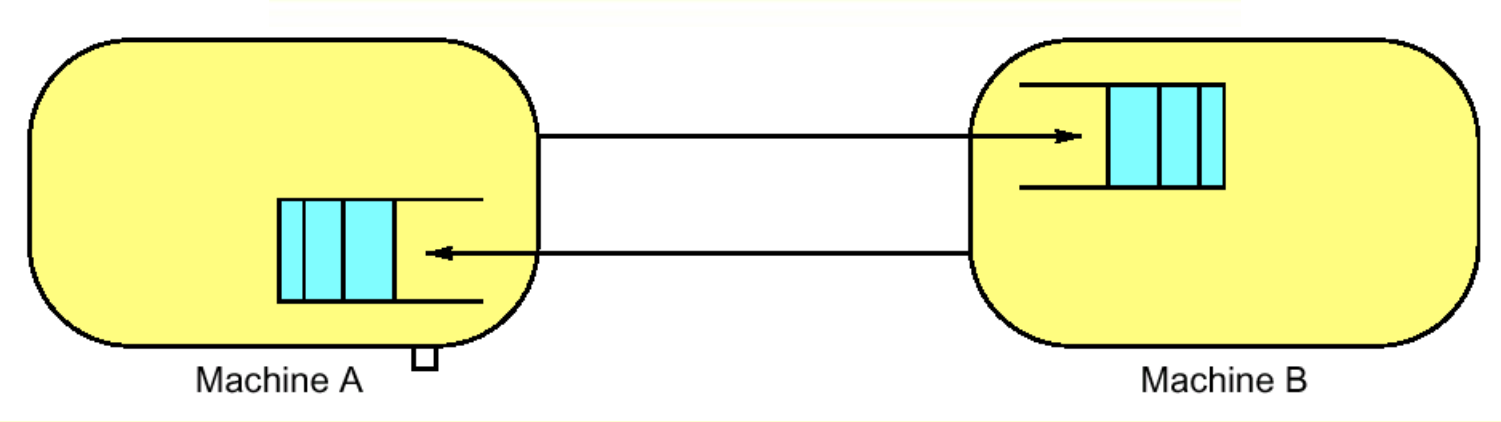
Most slides adapted from David Patterson. Some from Mohomed Younis

Interconnection Networks



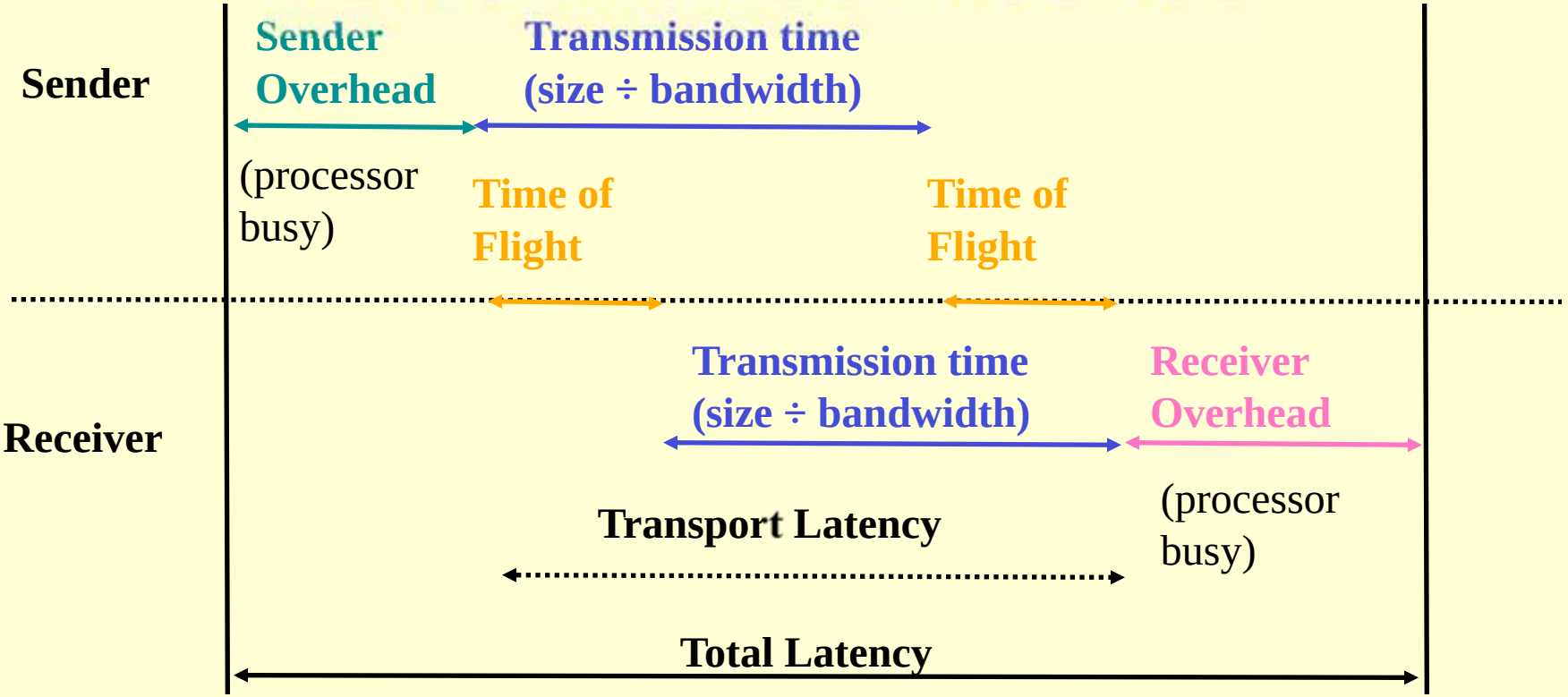
- Massively processor networks (MPP)
  - Thousands of nodes
  - Short distance (<~25m)
  - Traffic among nodes
- Local area network (LAN)
  - Hundreds of computers
  - A few kilometers
  - Many-to-one (clients-server)
- Wide area network (WAN)
  - Thousands of computers
  - Thousands of kilometers

ABCs of Networks



- Rules for communication are called the “protocol”, message header and data called a "packet"
  - What if more than 2 computers want to communicate?
    - Need computer “address field” (destination) in packet
  - What if packet is garbled in transit?
    - Add “error detection field” in packet (e.g., CRC)
  - What if packet is lost?
    - Time-out, retransmit; ACK & NACK
  - What if multiple processes/machine?
    - Queue per process to provide protection

Performance Metrics

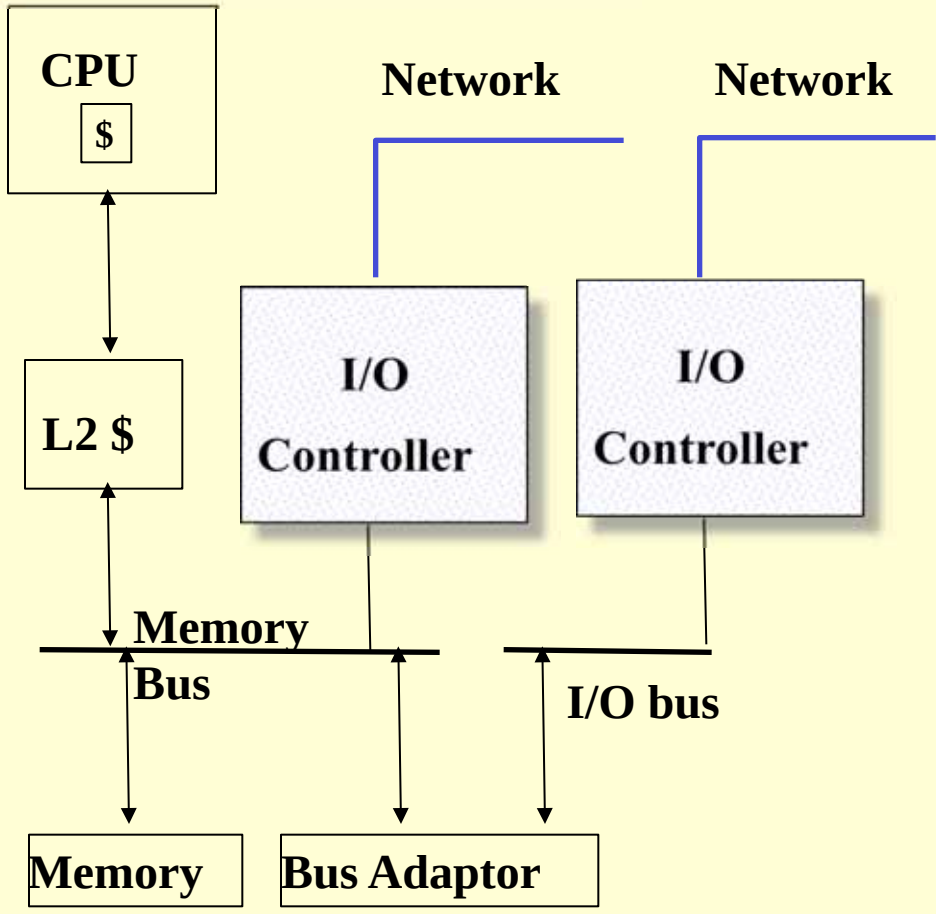


$$\text{Total latency} = \text{Sender Overhead} + \text{Time of flight} + \frac{\text{Message size}}{\text{Bandwidth}} + \text{Receiver overhead}$$

- Bandwidth: maximum rate of propagating information
- Time of flight: time for 1st bit to reach destination
- Overhead: software & hardware time for encoding/decoding, interrupt handling, etc.

Network Interface Issues

- Where to connect network to computer?
  - Cache consistency to avoid flushes
    - memory bus
  - Low latency and high bandwidth
    - memory bus
  - Standard interface card?
    - I/O bus
  - Typically, MPP uses memory bus; while LAN, WAN connect through I/O bus



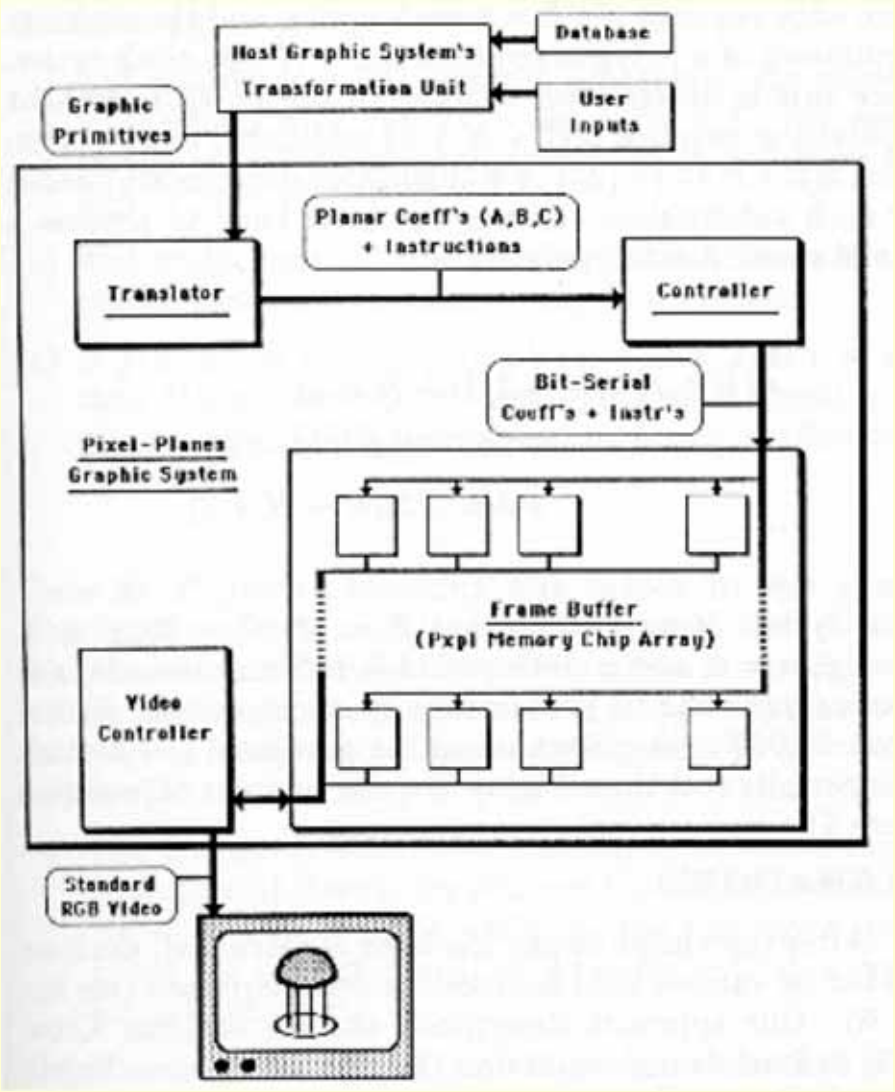
Ideal: high bandwidth, low latency, standard interface

Some Graphics Examples

- Pixel-Planes 4
- Pixel-Planes 5
- Pixel-Flow
- NVIDIA GeForce 6 series
- NVIDIA GeForce 8 series
- Intel Larrabee

Pixel-Planes 4

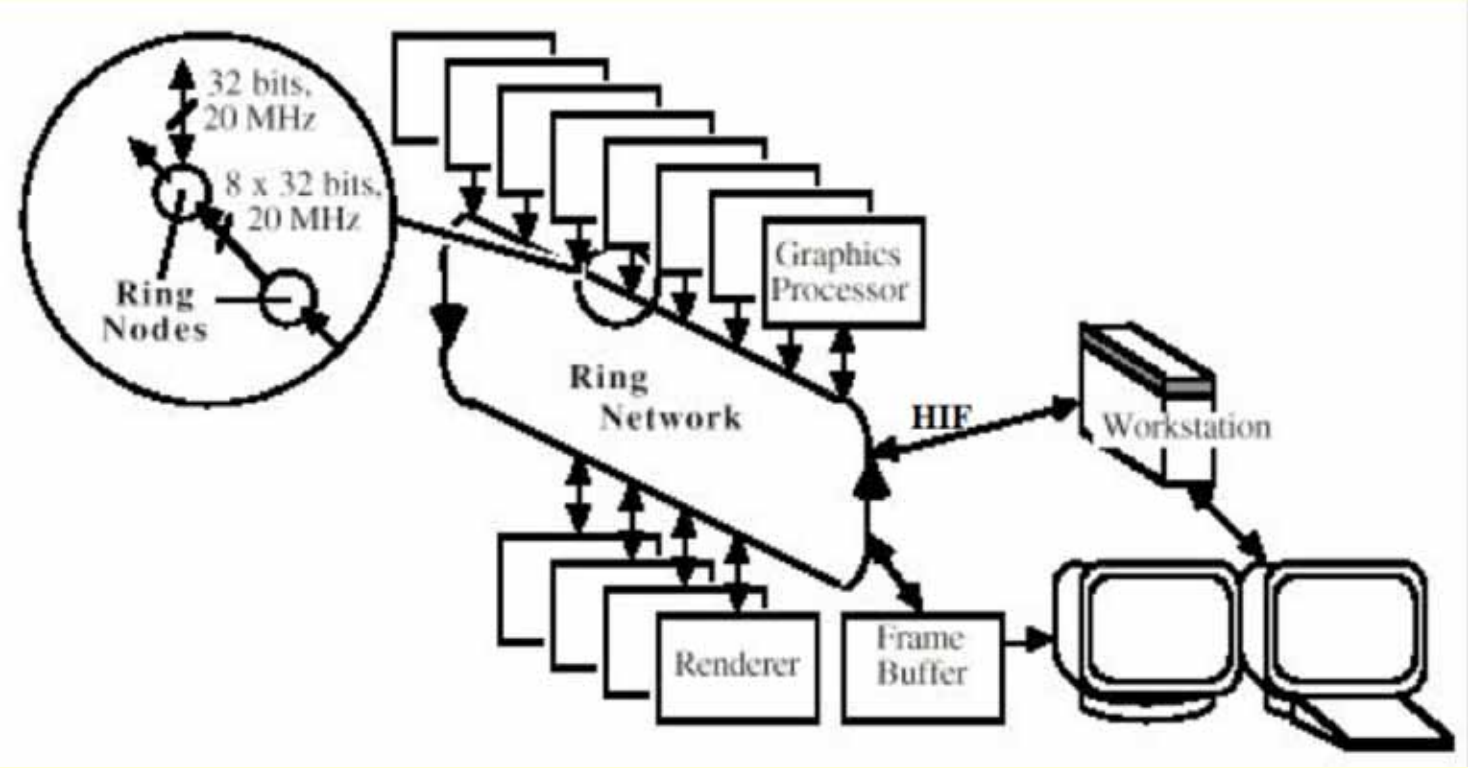
- 512x512 SIMD array (full screen)



Fuchs, et al., "Fast Spheres, Shadows, Textures, Transparencies, and Image Enhancements in Pixel-Planes", SIGGRAPH 1985

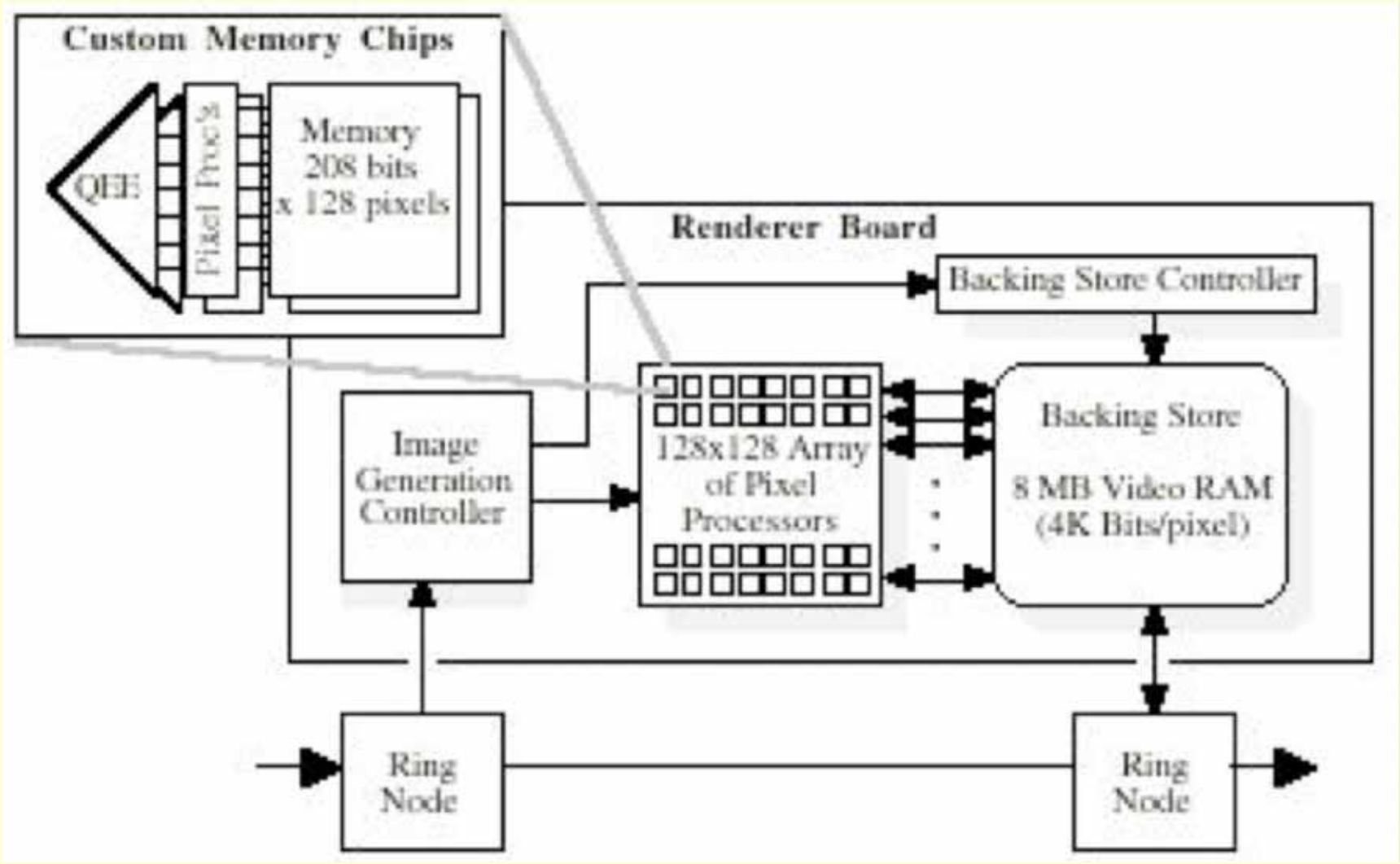
Pixel-Planes 5

- Message-passing
- ~40 i860 CPUs
- ~20 128x128 SIMD arrays (~80 tiles/screen)



Fuchs, et al., "Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor Enhanced Memories", SIGGRAPH 89

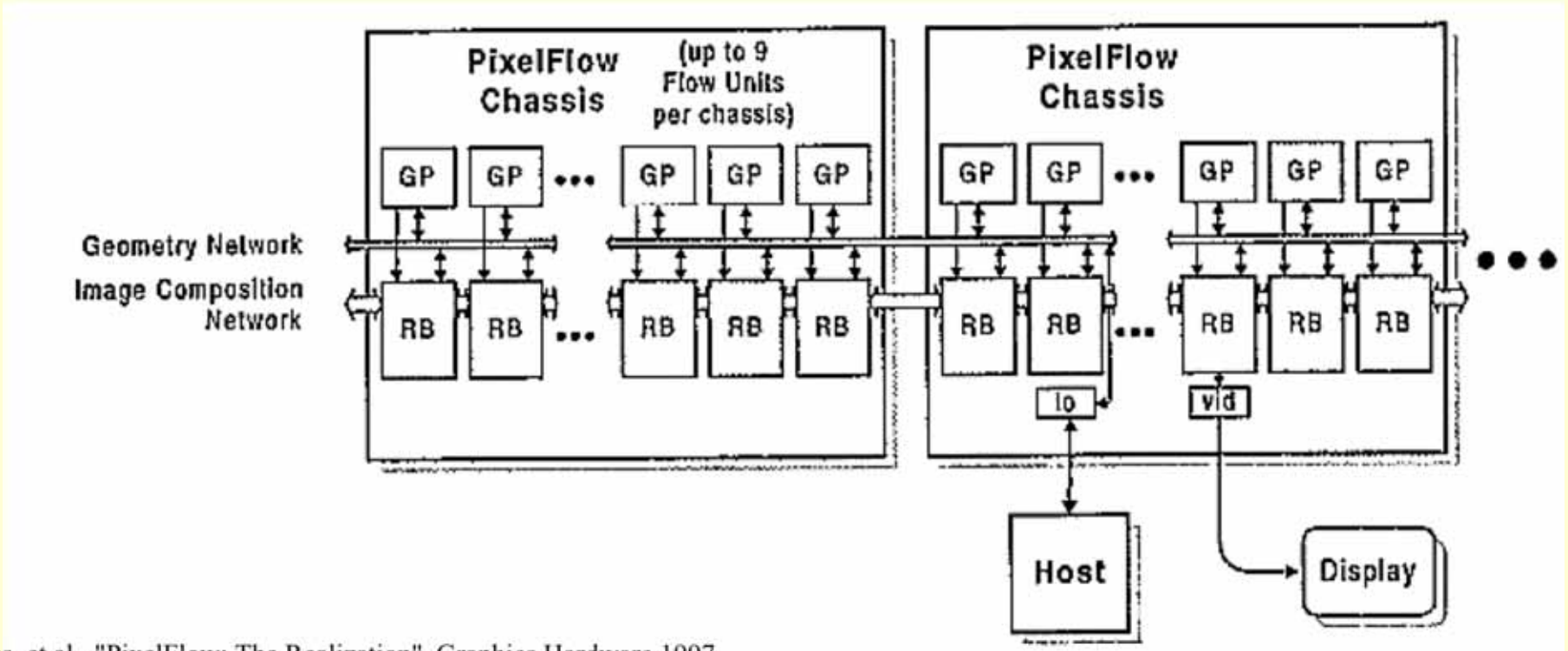
Pixel-Planes 5



Fuchs, et al., "Pixel-Planes 5: A Heterogeneous Multiprocessor Graphics System Using Processor Enhanced Memories", SIGGRAPH 89

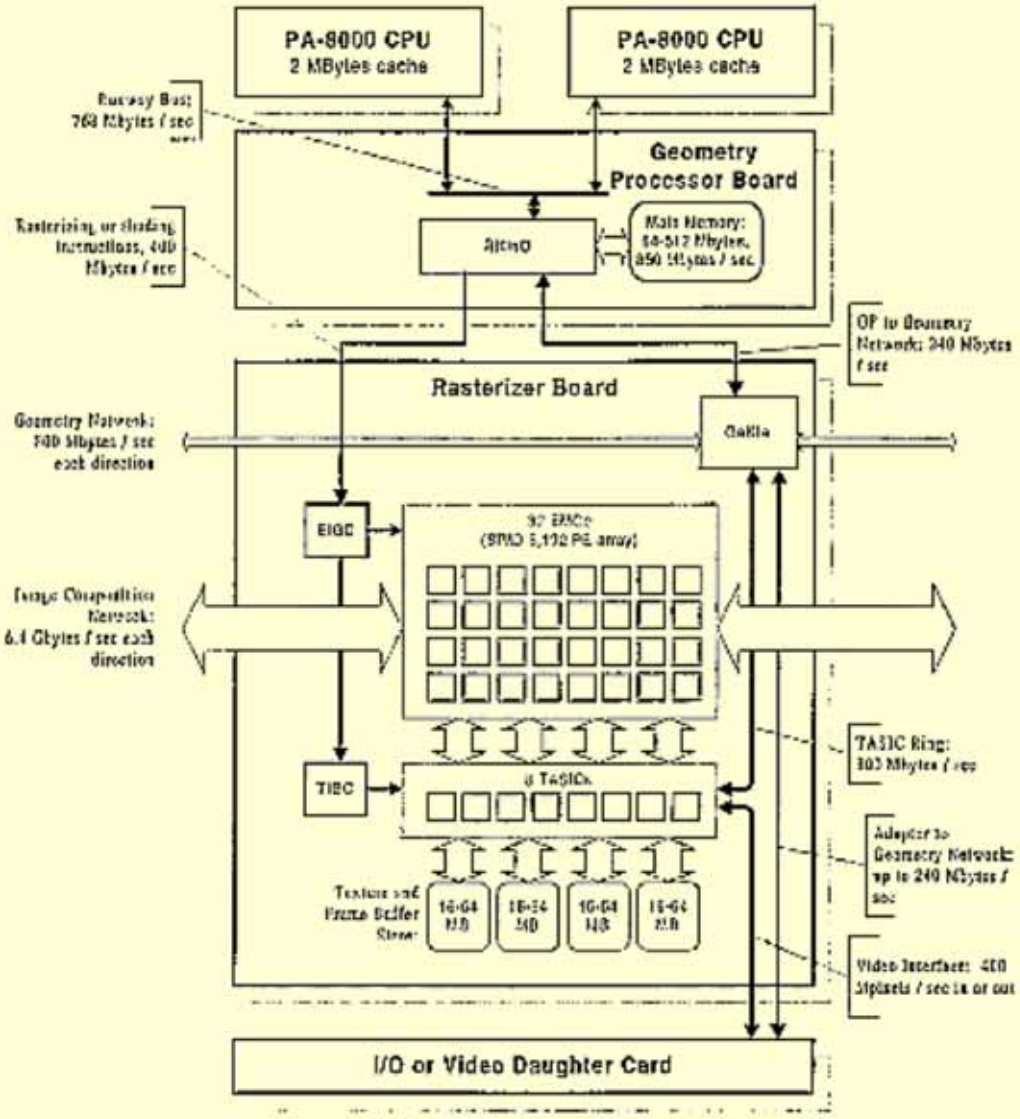
Pixel-Flow

- Message-passing
- ~35 nodes, each with
  - 2 HP-PA 8000 CPUs
  - 128x64 SIMD array (~160 tiles/screen)



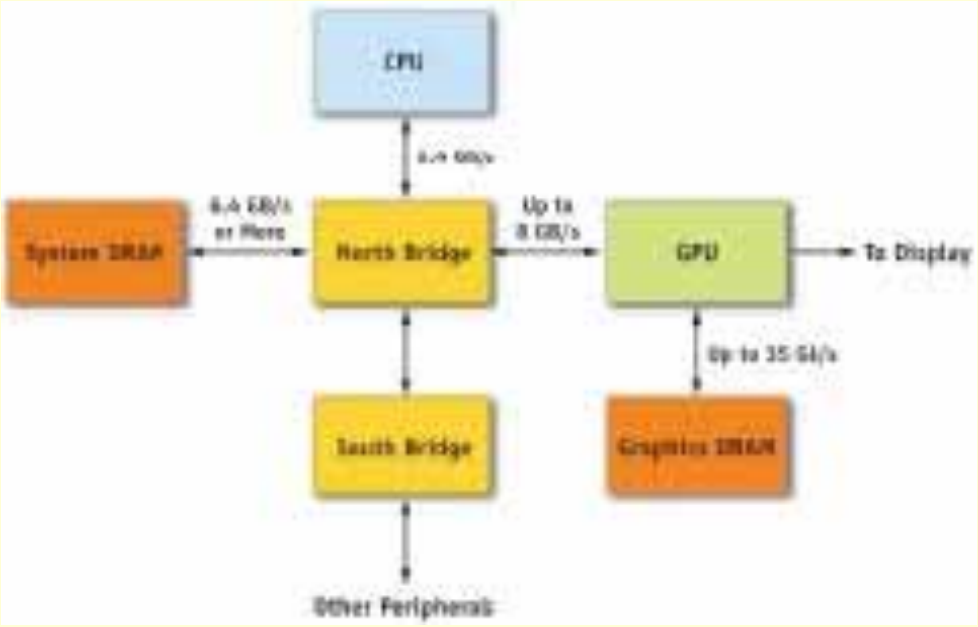
Eyles, et al., "PixelFlow: The Realization", Graphics Hardware 1997

Pixel-Flow



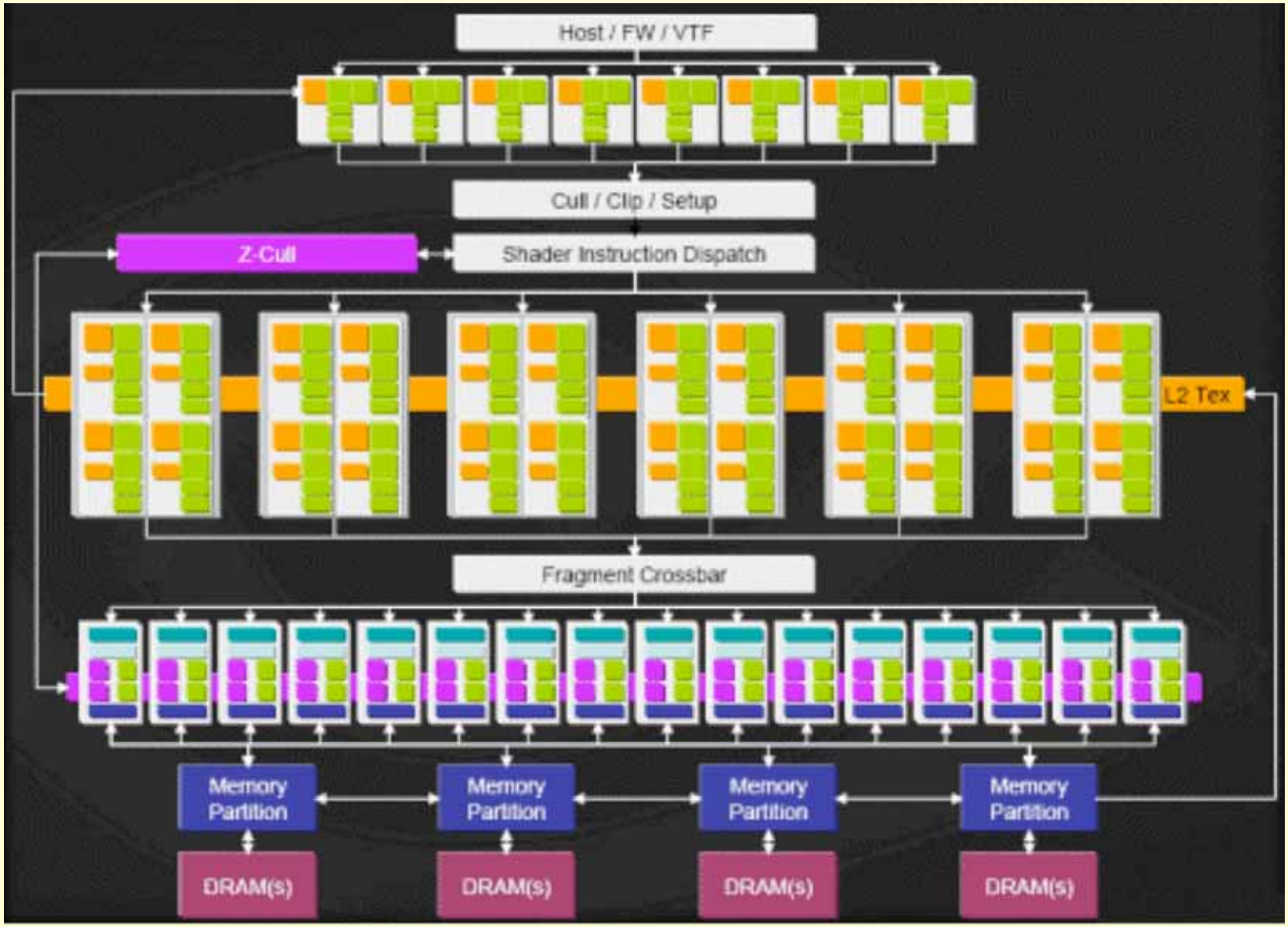
Eyles, et al., "PixelFlow: The Realization", Graphics Hardware 1997

PC Graphics Cards

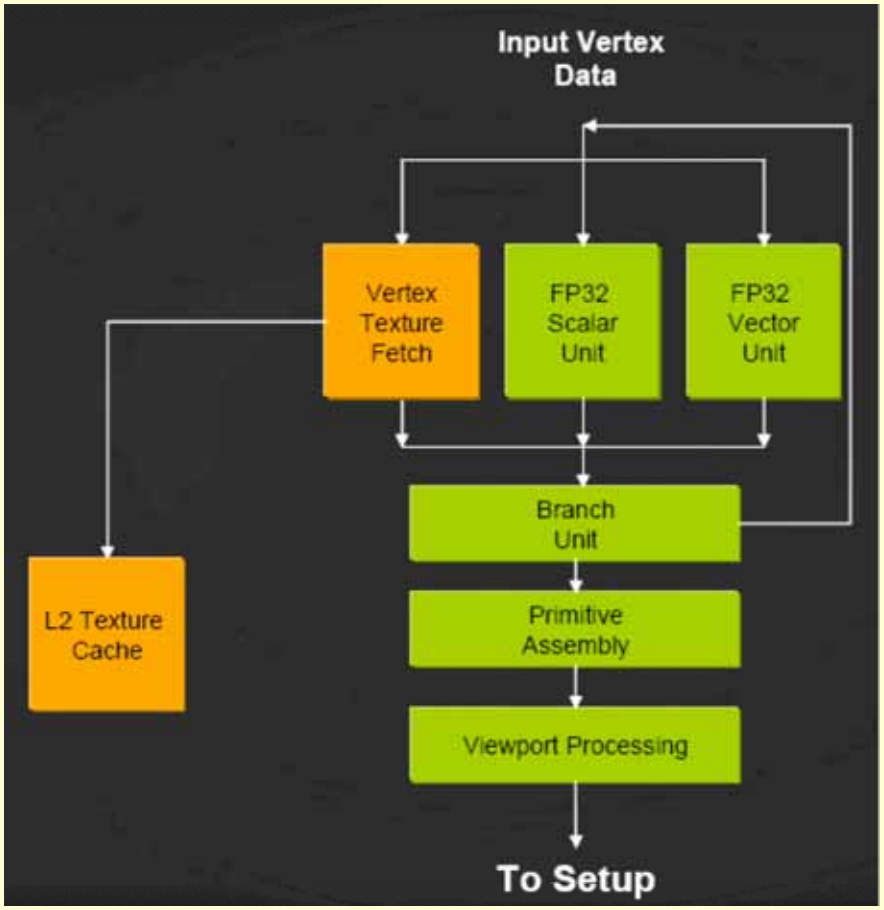


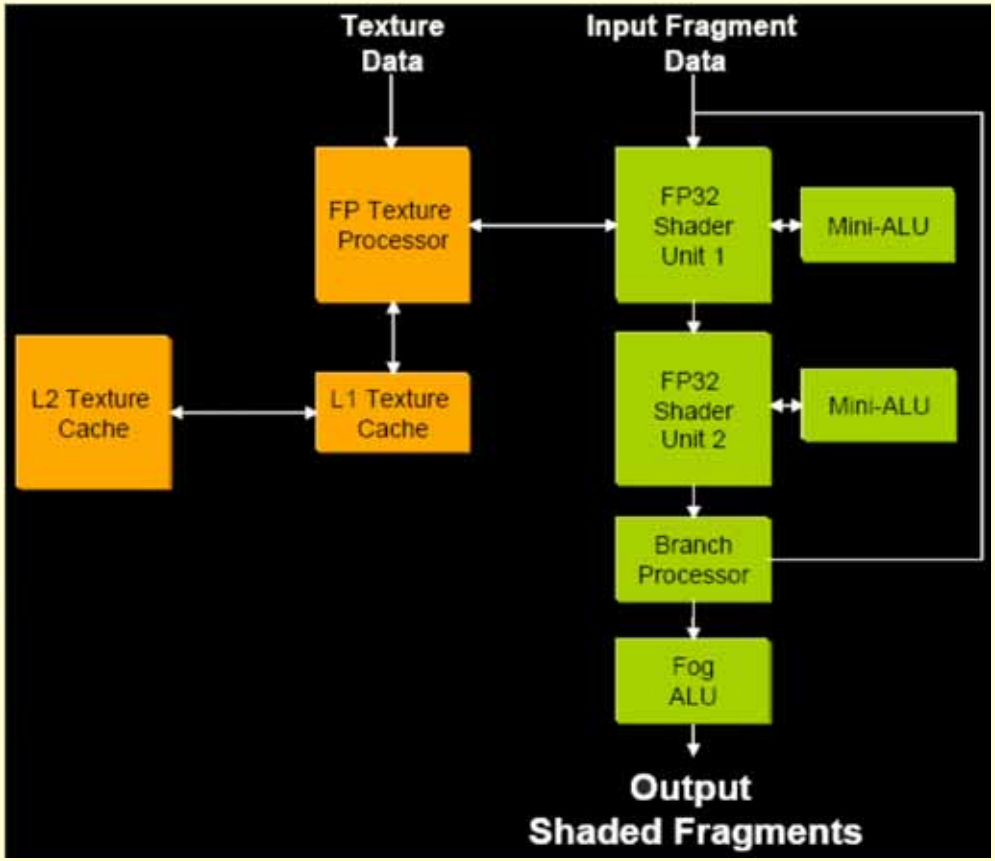
Kilgariff and Fernando, “The GeForce 6 Series Architecture”, GPU Gems 2

NVIDIA 7800 / G70

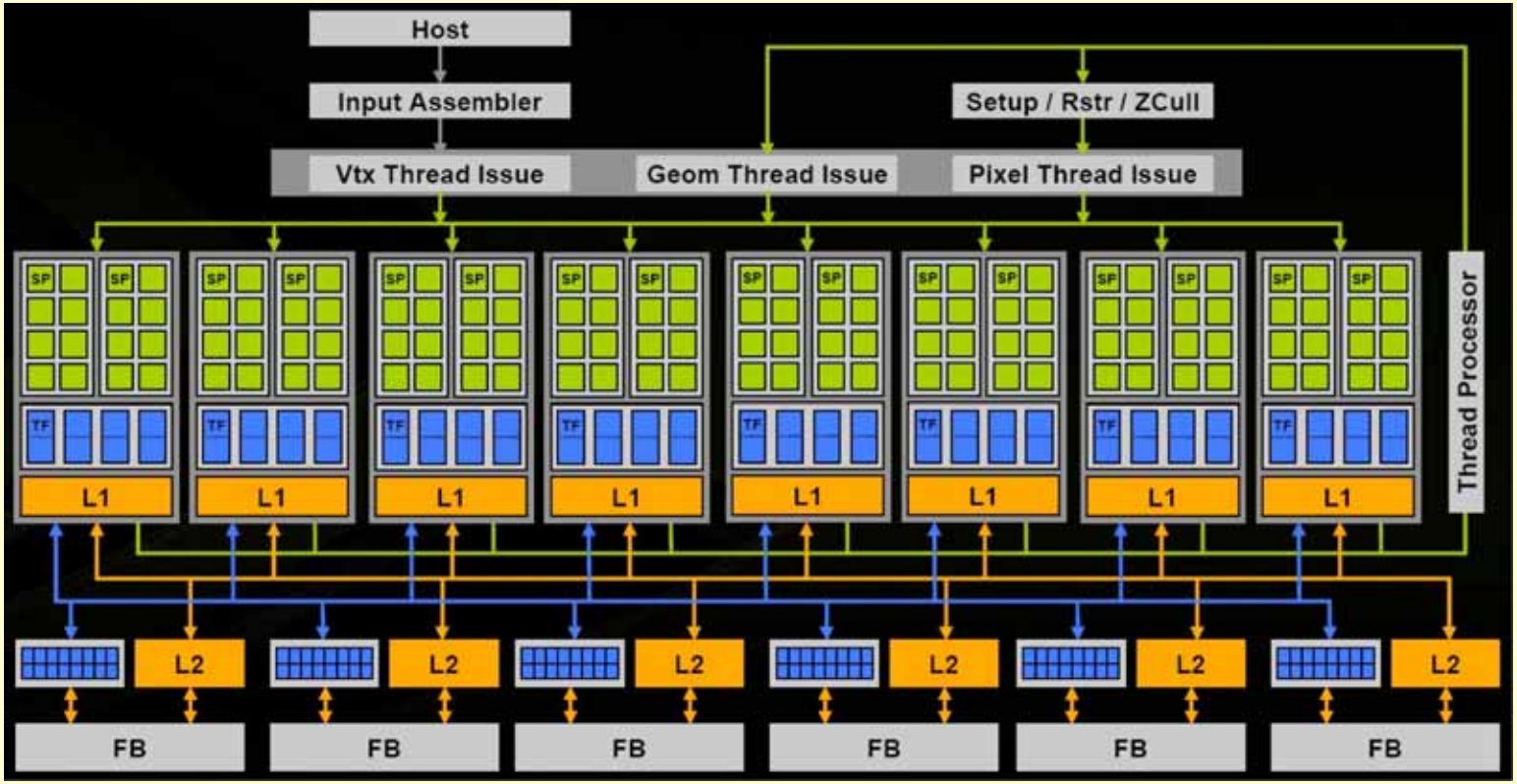


NVIDIA 7800 / G70



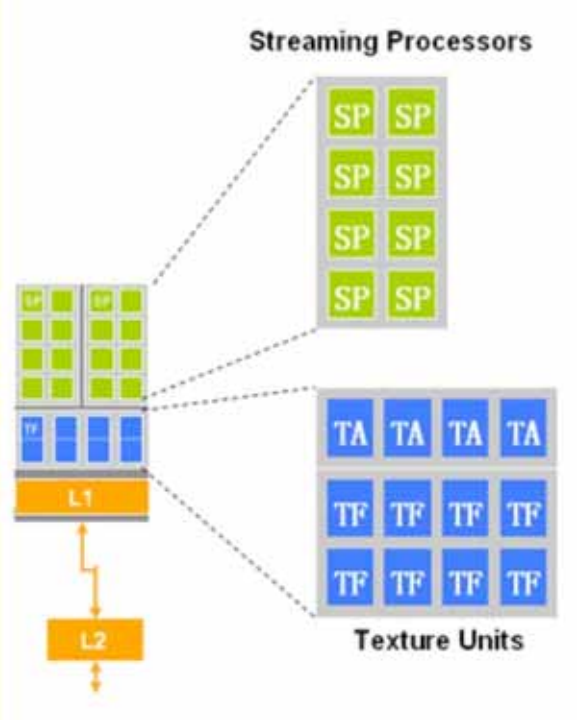


NVIDIA G80

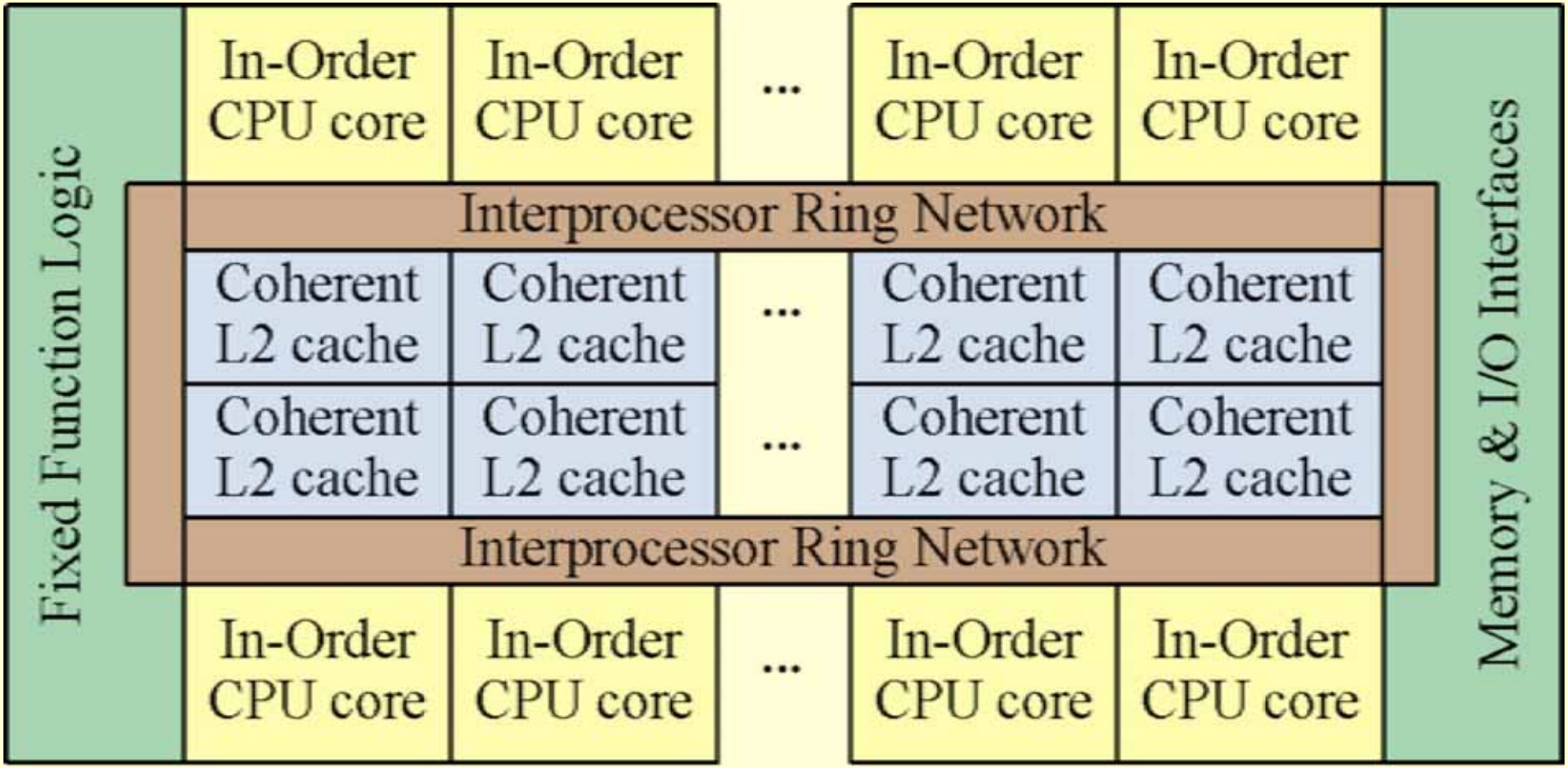


NVIDIA, “NVIDIA GeForce 8800 GPU Architecture Overview”, TB-02787-001_v01, November 2006

Streaming Processors

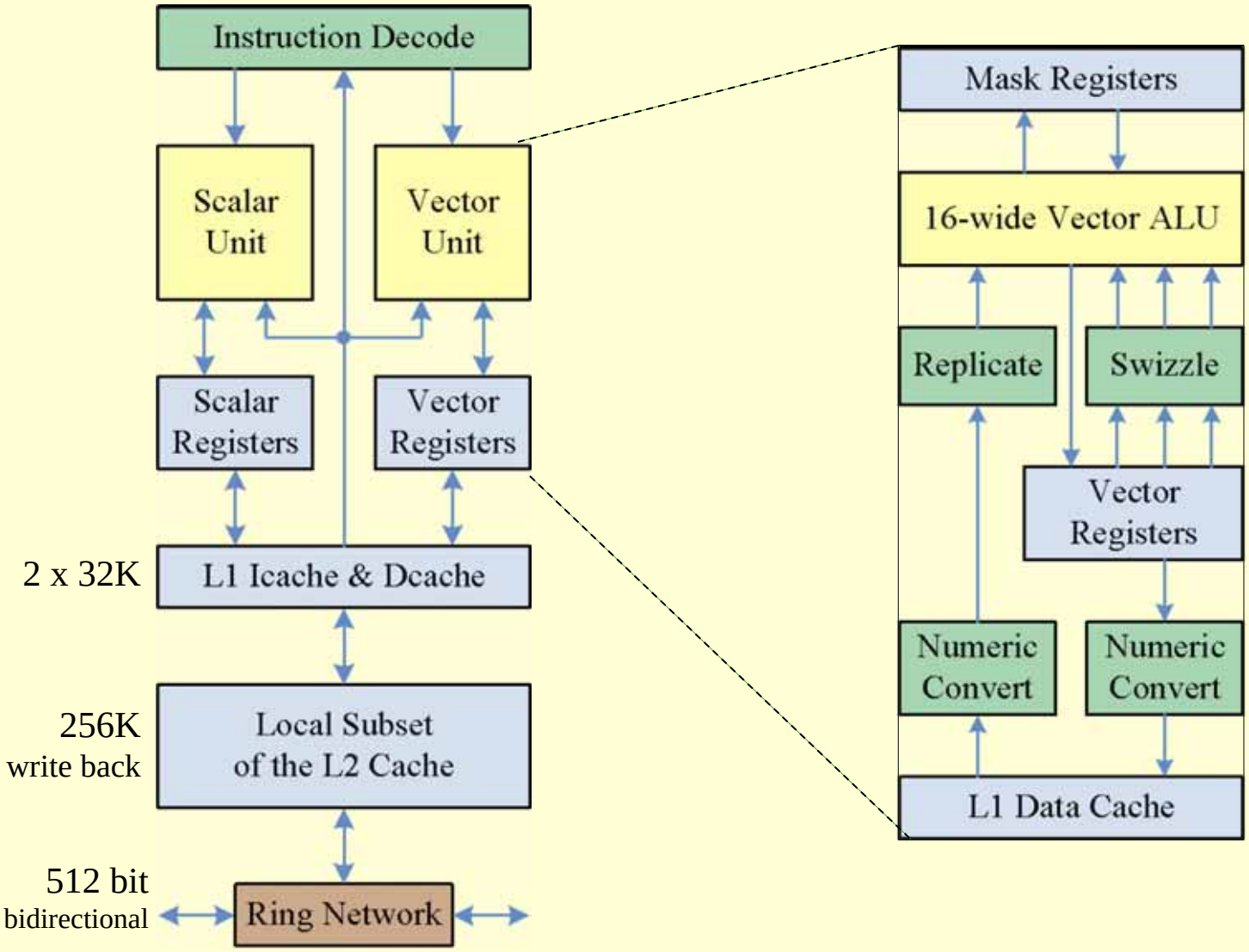


Intel Larrabee



Seiler et al., “Larrabee: A Many-Core x86 Architecture for Visual Computing”, ACM SIGGRAPH 2008

Larrabee Core



Larrabee: In Order Core

#CPU Cores	2 out-of-order	10 in-order
Instruction issue	4 per clock	2 per clock
VPU per core	4-wide SSE	16-wide vector
Single stream	4 per clock	2 per clock
Vector	8 per clock	160 per clock

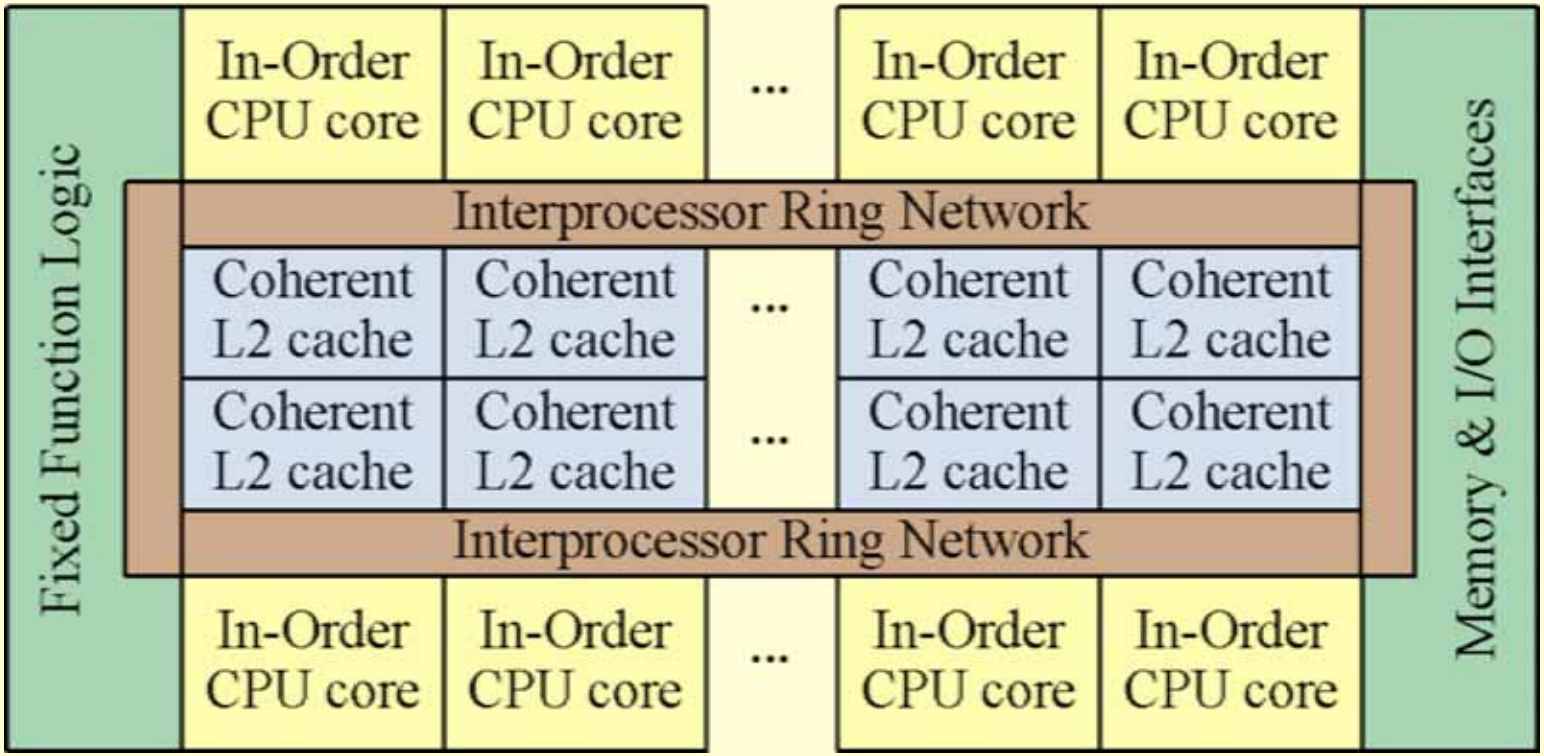
- Small, so fit more on chip

Larrabee ISA

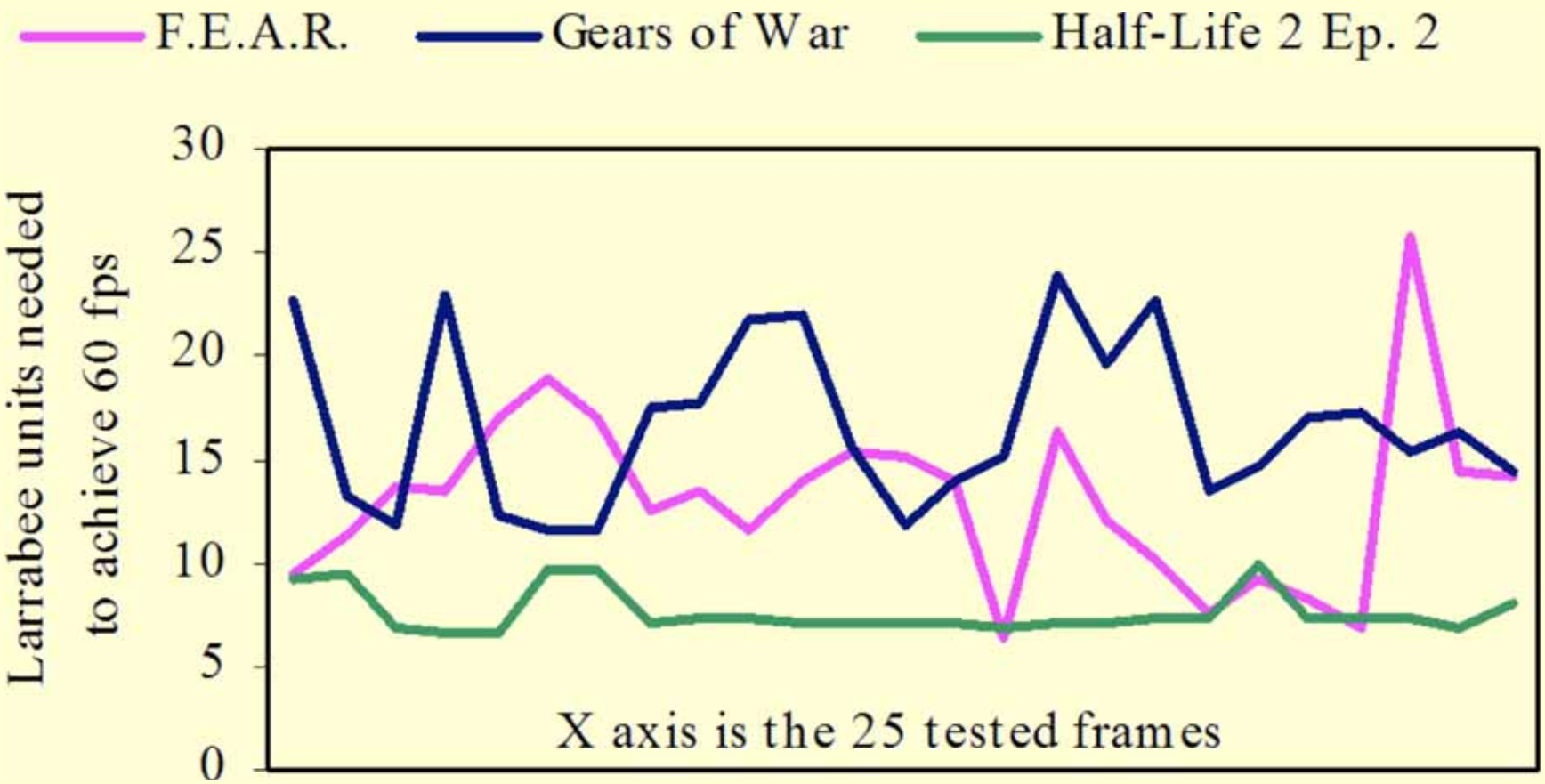
- x86 base
- Cache (instructions & modes)
  - prefetch
  - early eviction
  - Direct from L1 as fast as registers
- Exposed dual issue
  - 2nd restricted set for second instruction
- 4 threads w/ independent registers
- Vector instructions

Larrabee Fixed Function

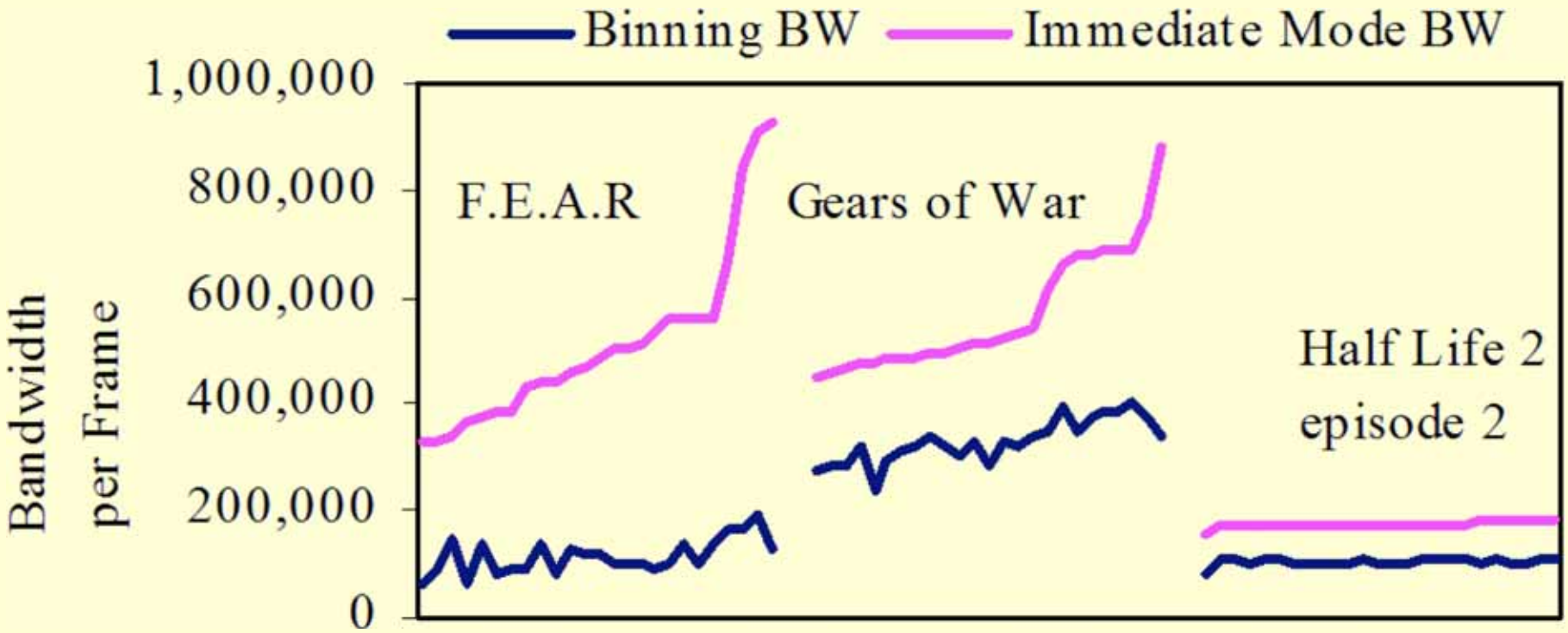
- Extra application-specific units
- Texture filtering
  - 12-40x faster than software



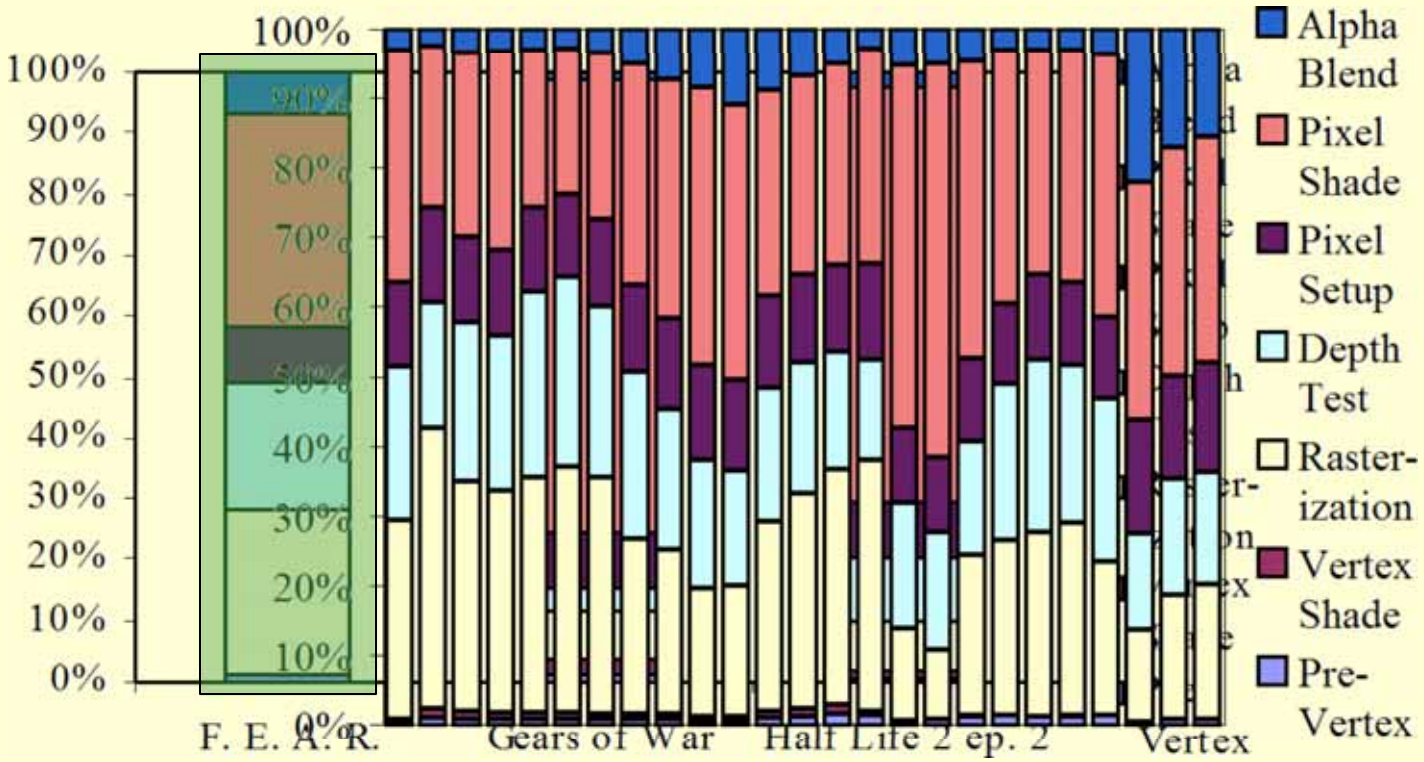
Larrabee Size



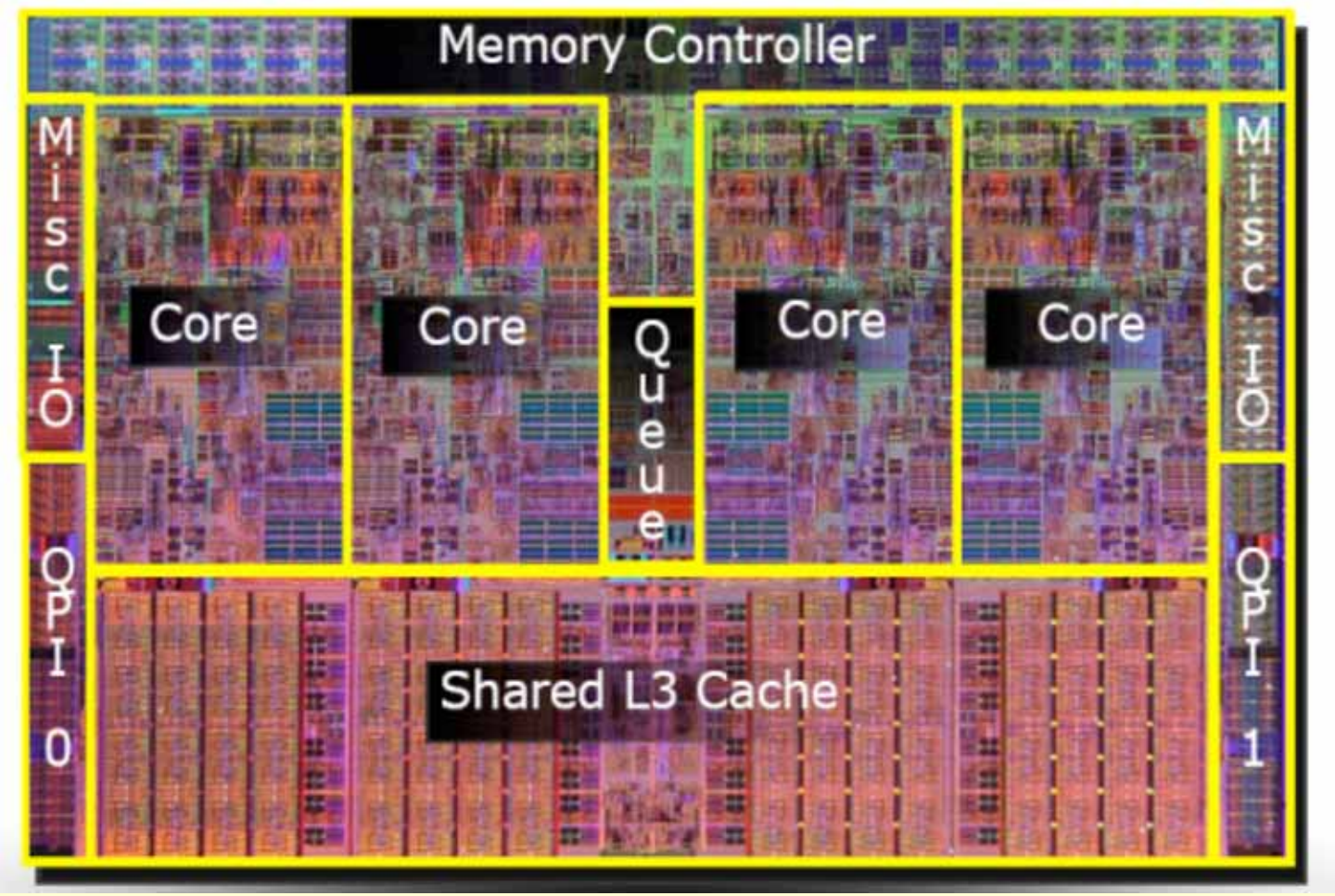
Larrabee Bandwidth



Larrabee Processing



Nehalem


Memory Controller
Misc IO
Core
Core
Queue
Core
Core
Misc IO
QPI 0
Shared L3 Cache
QPI 1