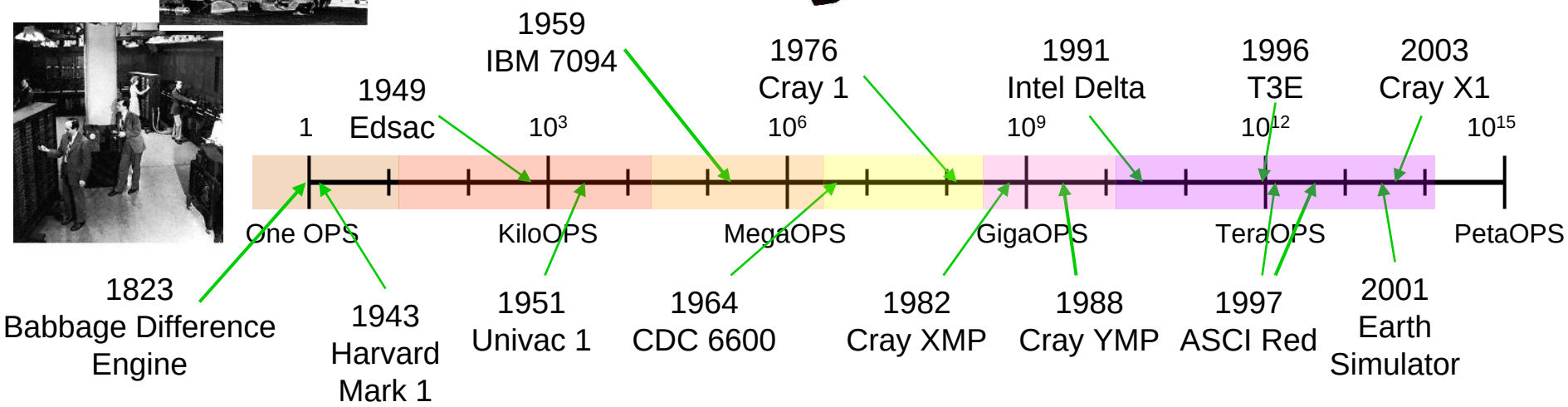
Exotic Technologies:
The *Vision2020* Exaflops Architecture

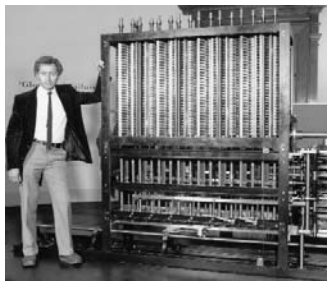
Thomas Sterling
Department of Computer Science
November 16, 2006

Topics

- Introduction
- Linpack projections
- Technology projections
- Challenges
- Paradigm shift in execution model
- Vision2020 architecture proposal
- Conclusions

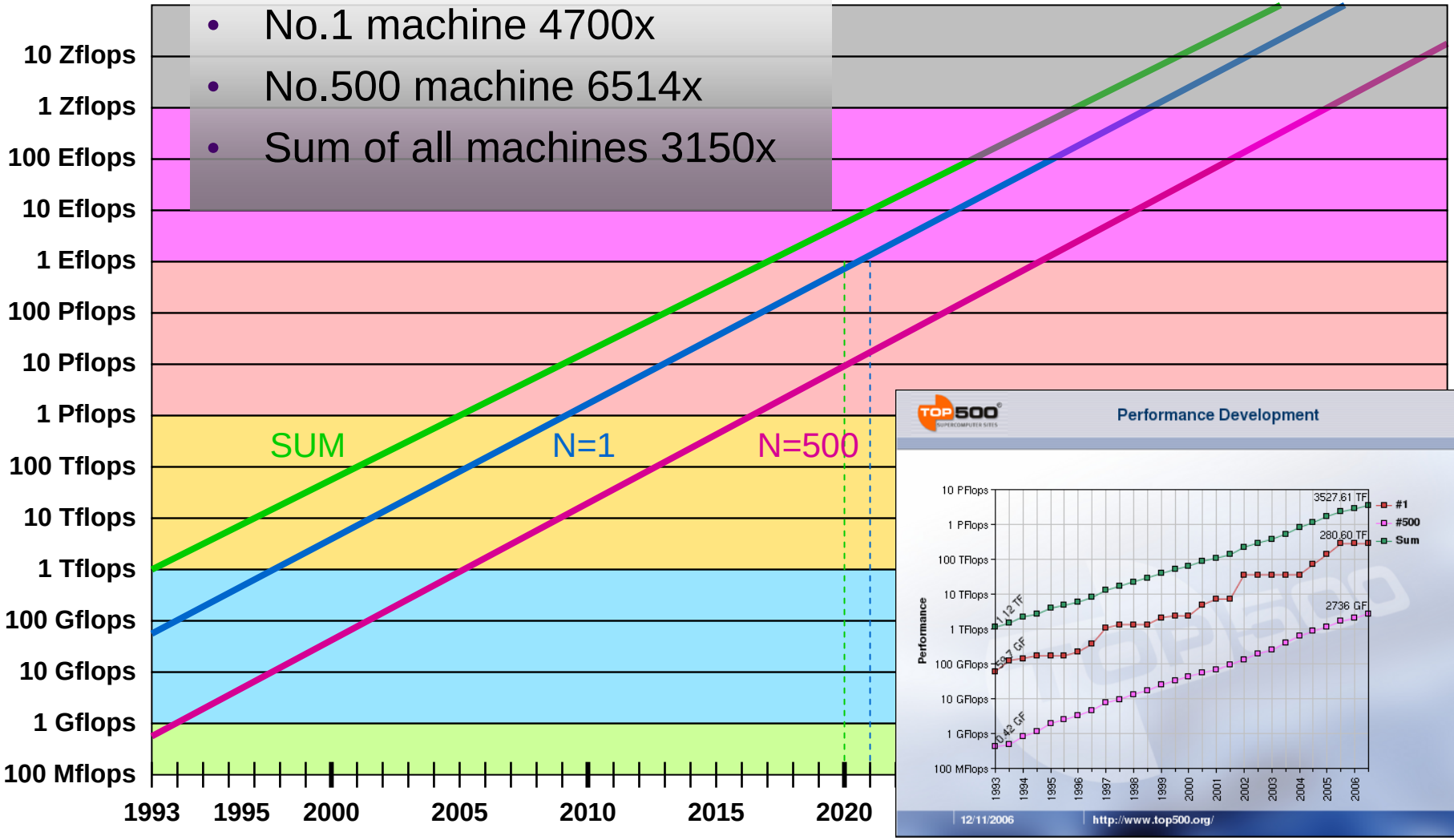
A Growth-Factor of a Billion in Performance in a Single Lifetime



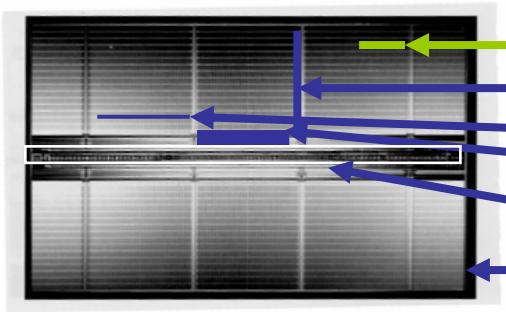


Linpack 1 Exaflops in 2020

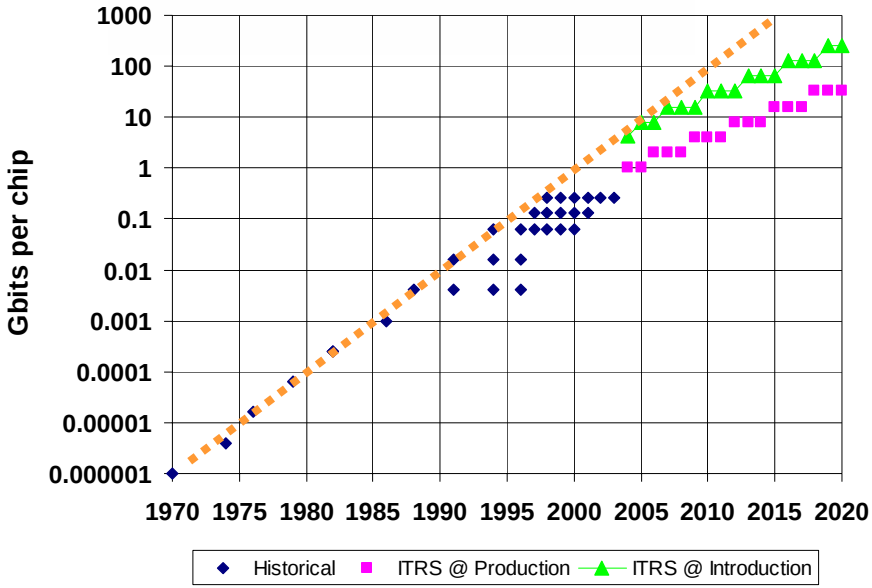
- Spans 13.5 years
- No.1 machine 4700x
- No.500 machine 6514x
- Sum of all machines 3150x

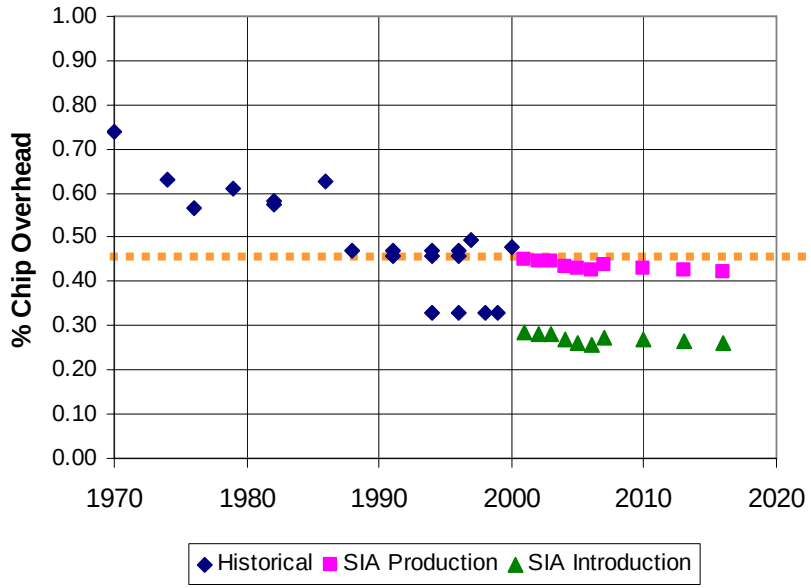


Classical DRAM



- Memory mats: ~ 1 Mbit each
- Row Decoders
- Primary Sense Amps
- Secondary sense amps & “page” multiplexing
- Timing, BIST, Interface
- Kerf

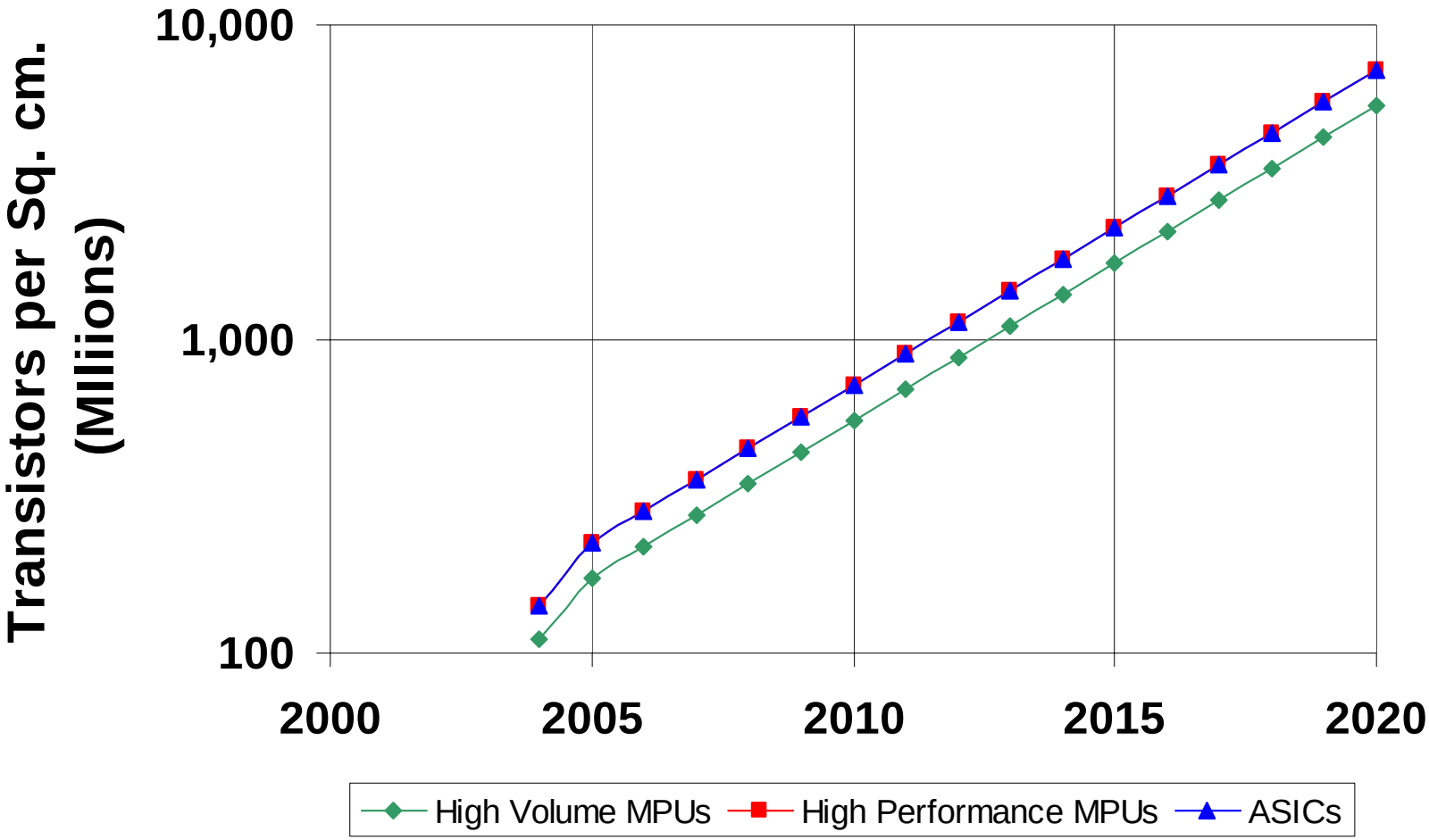




Density/Chip has dropped below 4X/3yrs

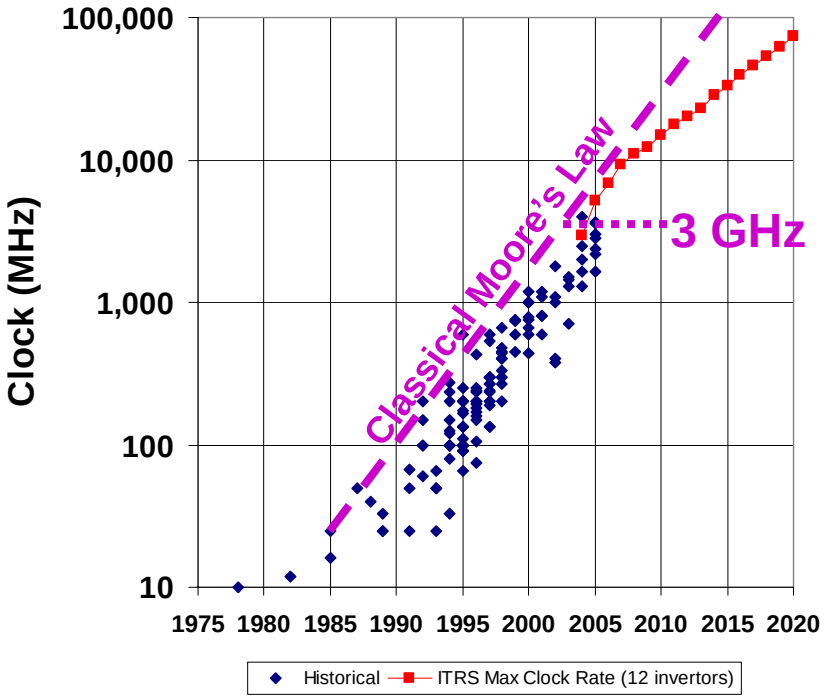
And 45% of Die is *Non-Memory*

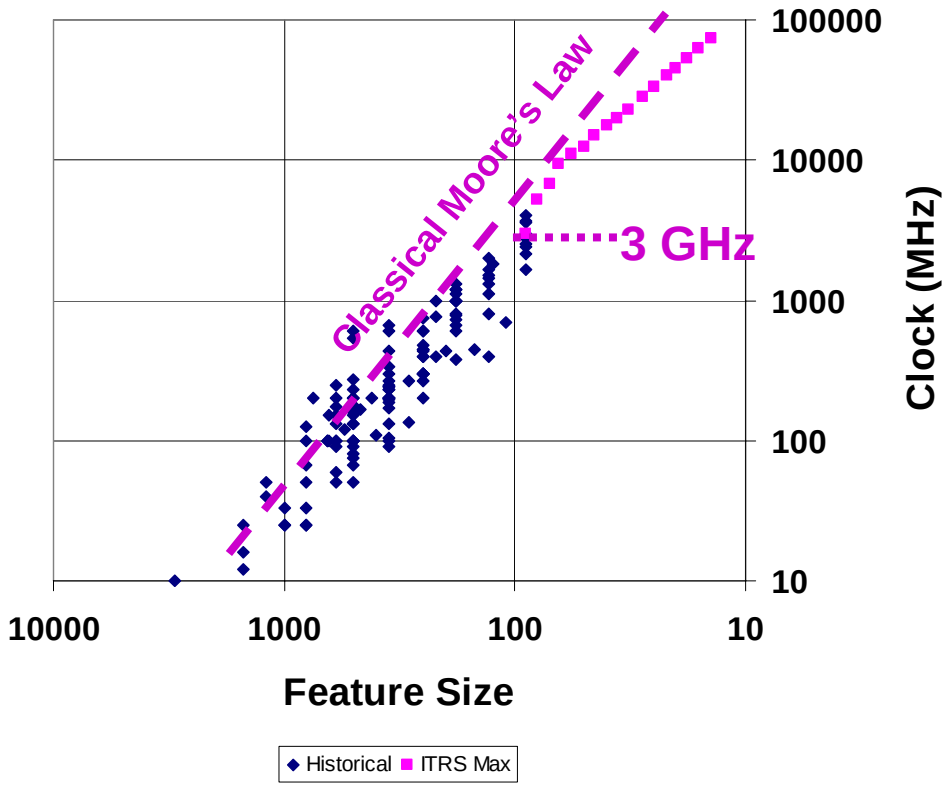
Logic Chip Density Scaling



Logic functions per unit area: ~2x every 3 years

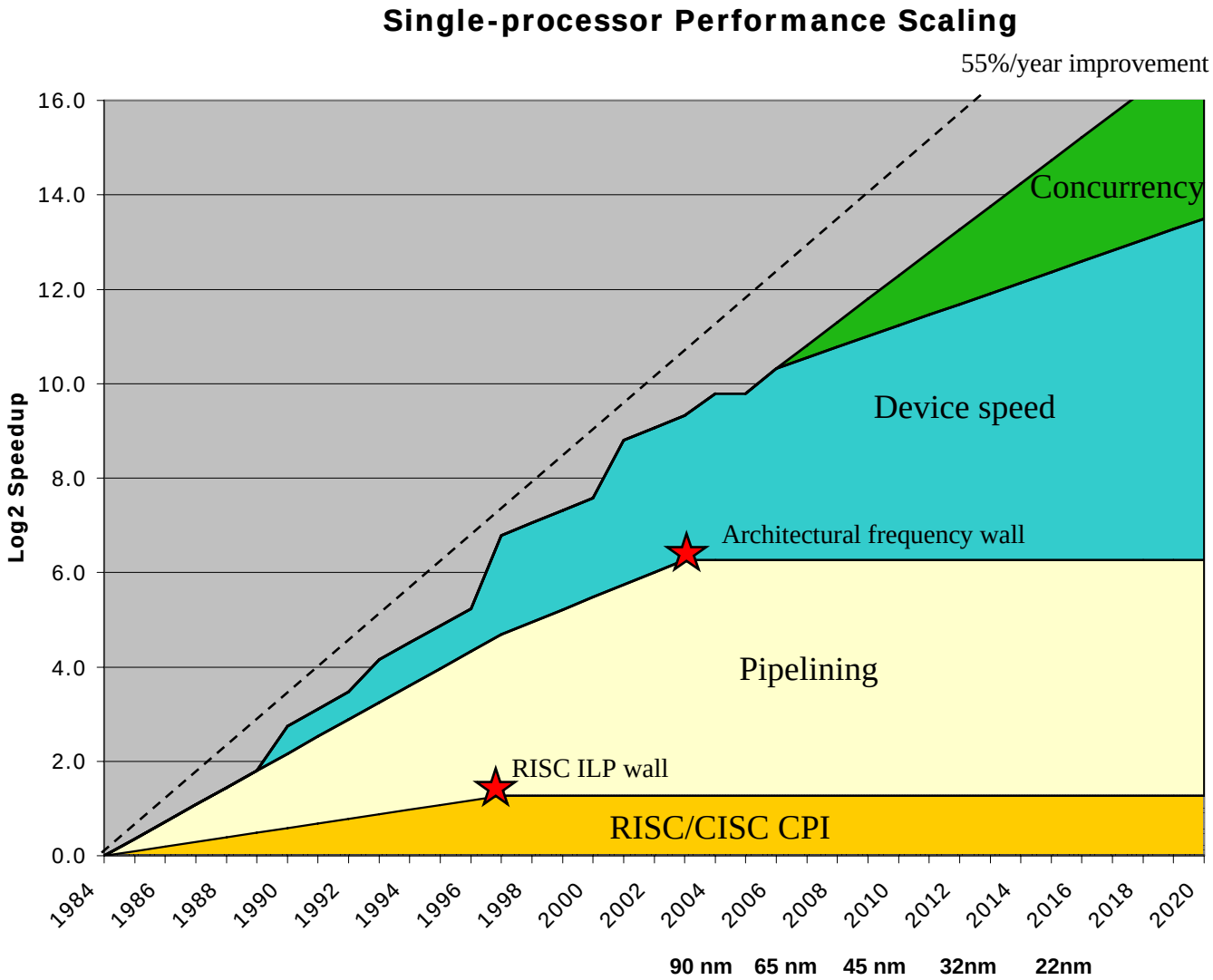
Peak Logic Clock Rates





2005 projection was for 5.2 GHz – and we didn't make it in production. Further, we're still stuck at 3+GHz in production.

Performance Scaling



New programming models needed?

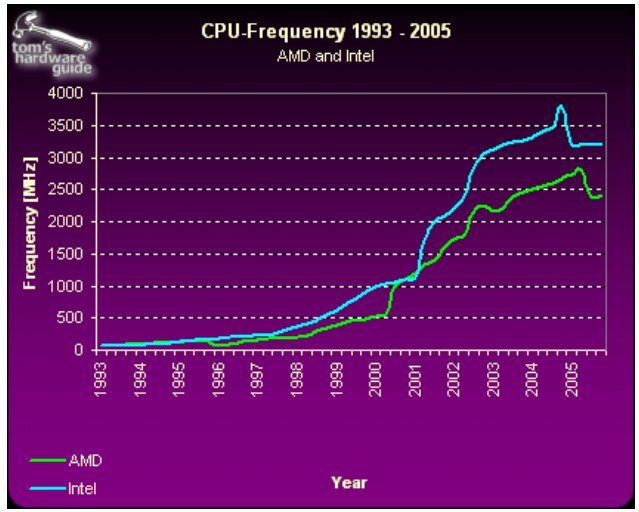
Assume successful 17%/year scaling

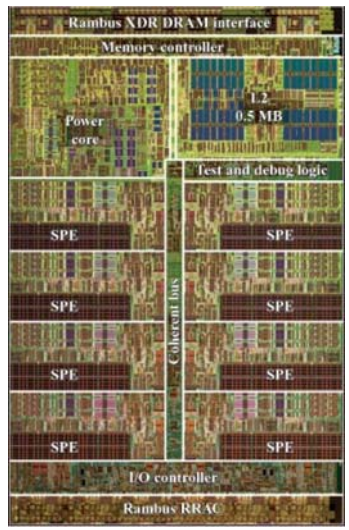
Conventional architectures cannot improve performance

Industry shifts to frequency dominated strategy

SIA ITRS Projections

- Chip memory capacity
  - Projects 32 Gigabits/chip by 2020
  - 45% of chip is non-memory
  - Growth factor < 4X every 3 years
- Logic density 2X every 3 years
  - Factor of 25X
- Clock rate is uncertain
  - Projects 70+ GHz by 2020
  - Current projections not met
  - ~10X or 32 GHz
- Conclusions
  - Technology alone insufficient
  - Power consumption not considered
  - Massive memory/logic imbalance
  - Architecture must make up the difference



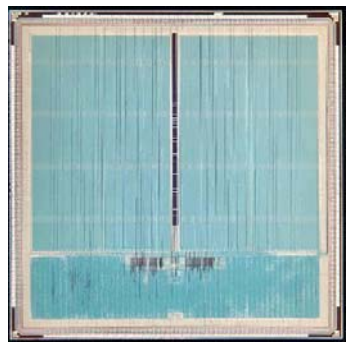


Challenges for *Vision2020* Machine

- Performance degradation
  - Latency (idle time due to round trip delays)
  - Overhead (critical path support mechanisms)
  - Contention (inadequate bandwidth)
  - Starvation (sufficient parallelism and load balancing)
- Power consumption
  - Just too much!
  - Dominating practical growth in mission critical domains
- Reliability
  - Single point failure modes cannot be tolerated
  - Reduced feature size and increased component count
- Changing application workload characteristics
  - Data (meta-data) intensive for sparse numerics and symbolics
- Programmability & ease of use
  - System complexity, scale and dynamics defy optimization by hand

The *Vision2020* System Strategy

- Cost imperatives:
  - High availability ALUs
  - High utilization of memory bandwidth
  - Percolation: cheap threads manage global parallel flow control
- New Model of Parallel computation (*ParalleX*)
  - Intrinsic latency hiding
  - Message-driven split-phase transaction processing
  - Near fine-grain in-memory synchronization local control objects
- Heterogeneous architecture for disparate temporal-locality modalities
  - High temporal locality: high ALU-density dataflow controlled
  - Low temporal locality: in memory threads
- Global name space
  - Address translation in meta-data
  - Copy semantics for targeted value-sets

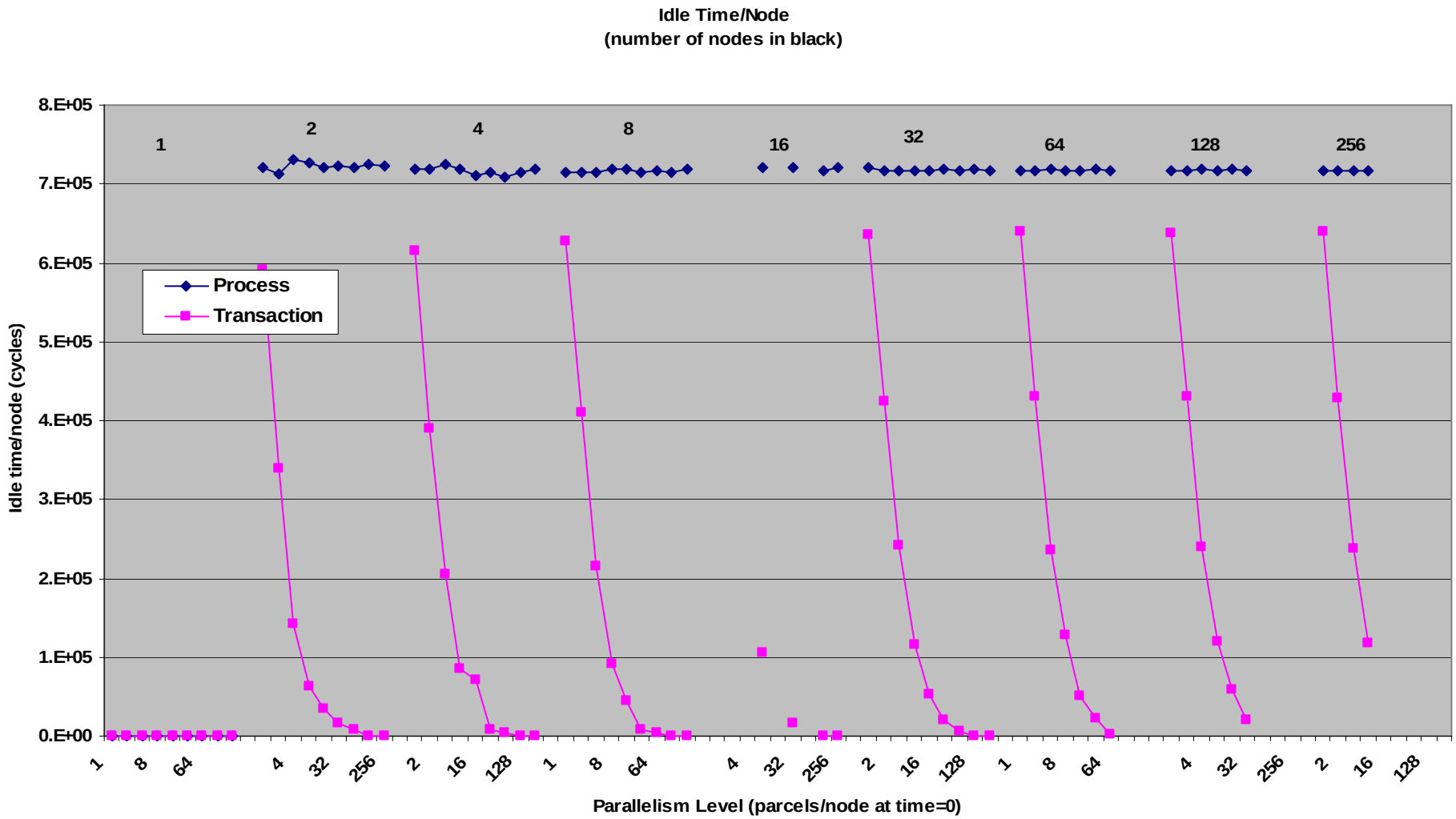


ParalleX Semantics

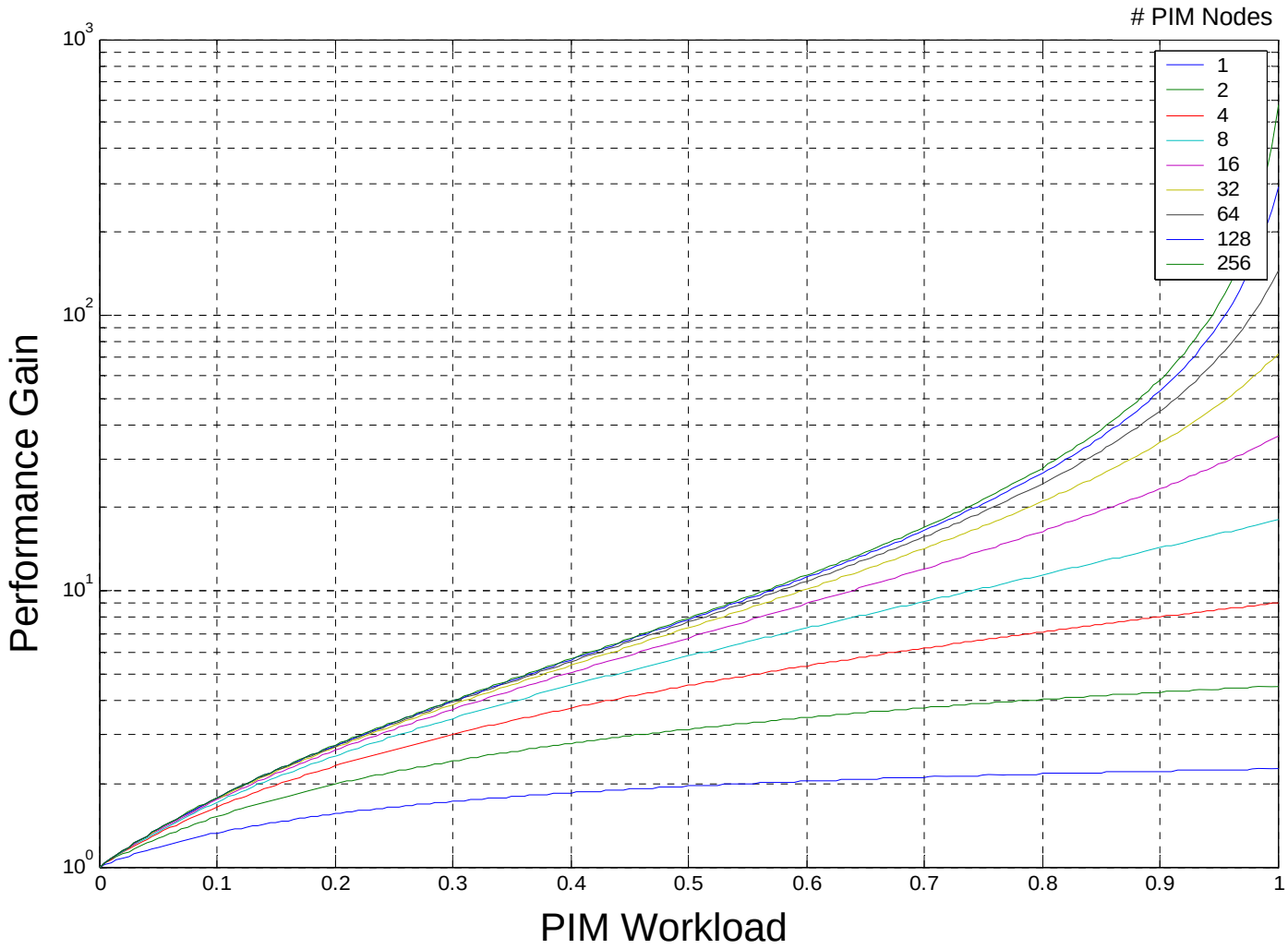
- Locality domains
  - Intra-locality: Controlled synchronous
  - Inter-locality: Asynchronous between localities
- Split-phase transactions
  - Work queue model
    - Only do work on local state
    - No blocking or idle time for remote access
- Message-driven computation
  - Parcels carry data, target address, action, continuation
- Multi-threaded
  - First class objects
  - Dataflow on transient values
- Local control objects
  - Futures
  - Dataflow
  - Data-directed parallel control
- Meta-data embedded address translation
- Failure-oriented with micro-checkpointing

Latency Hiding with Parcels

Idle Time with respect to Degree of Parallelism



Speedup with Memory Accelerators

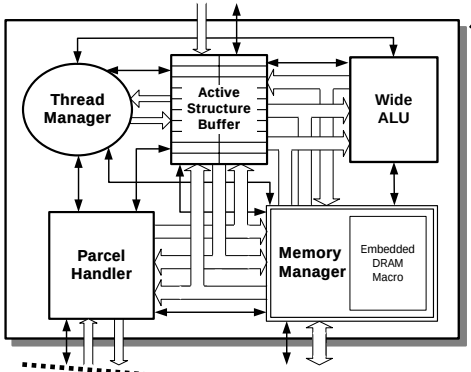


Vision2020 System Elements

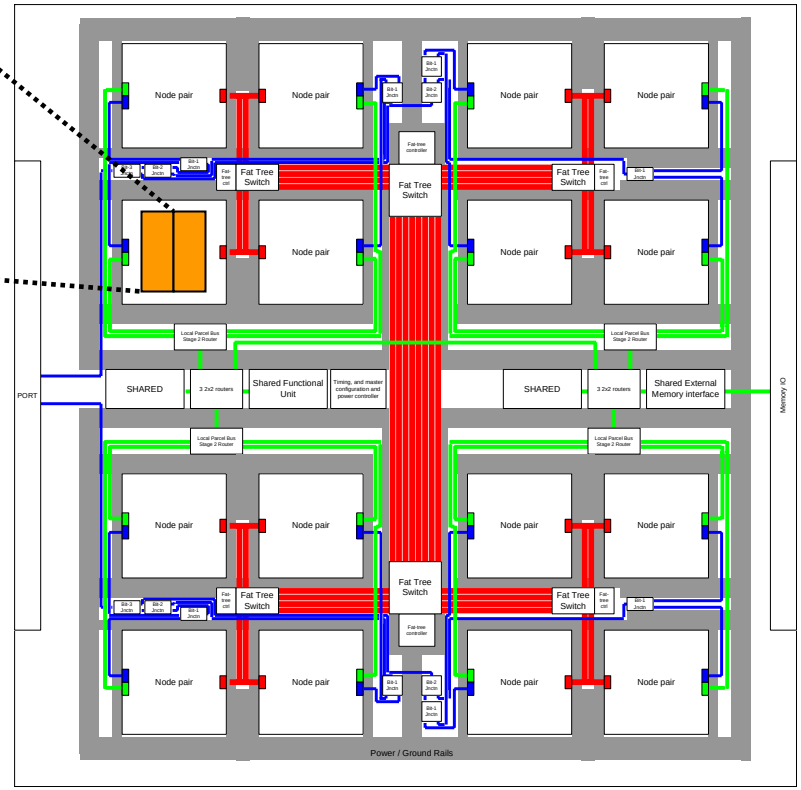
- Executable Memory
  - Supports low-temporal (e.g. touch once) locality global data operations
  - Threads in memory with wide ALUs
- Dataflow Accelerator
  - Supports high-temporal locality operations
  - Very high throughput low latency processing
  - Low power per operation
- Data Vortex optical network
  - Innovative topology
  - Low latency, low logic
  - Graceful degradation of injection rate with traffic density
  - High degree switches
- Penultimate store
  - Fast backing store for core computing
  - Exploits highest density semiconductor memory
  - Reconfigurable for fault tolerance

MIND elements

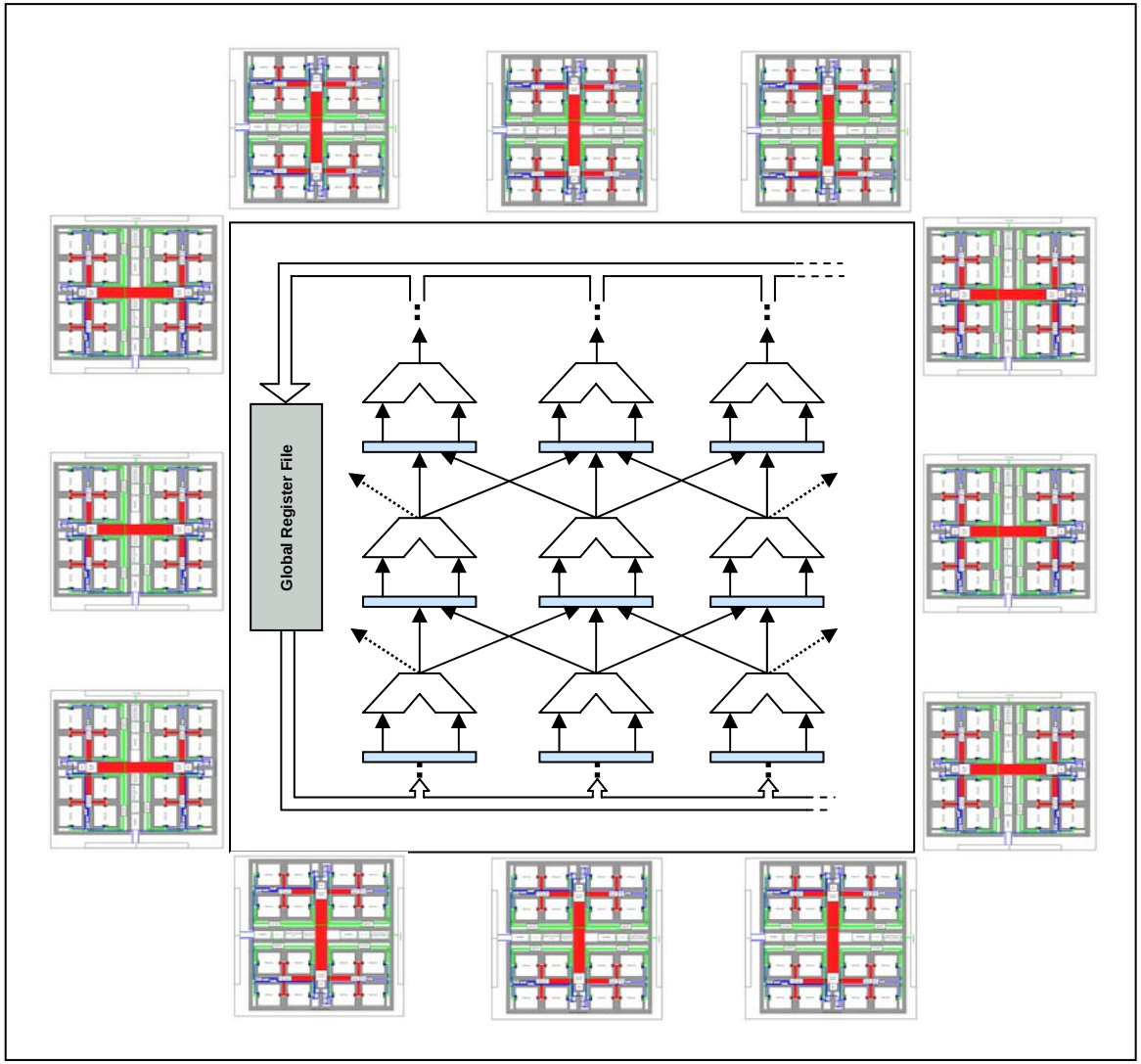
MIND memory accelerator



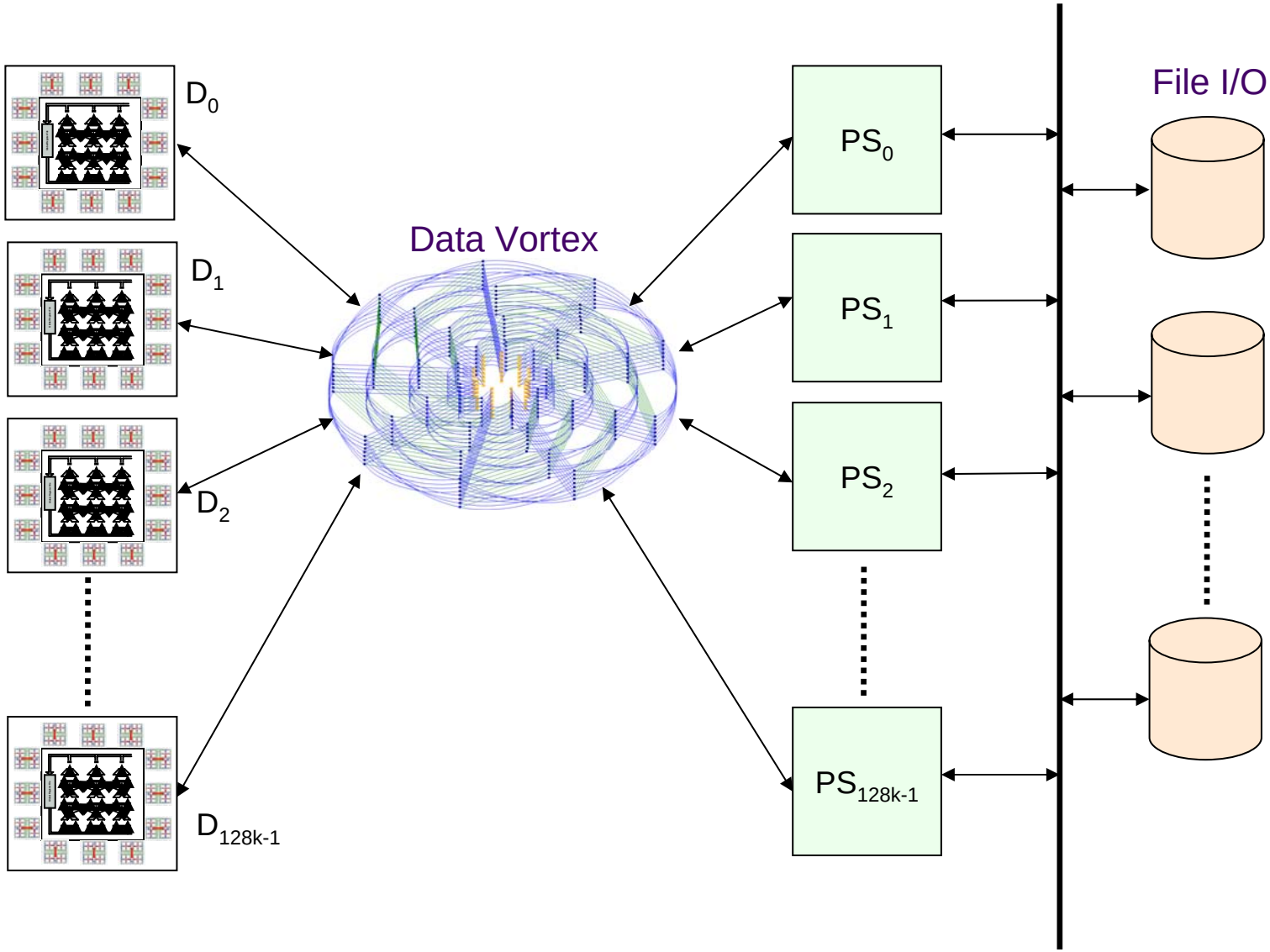
Gilgamesh



Dodecatron

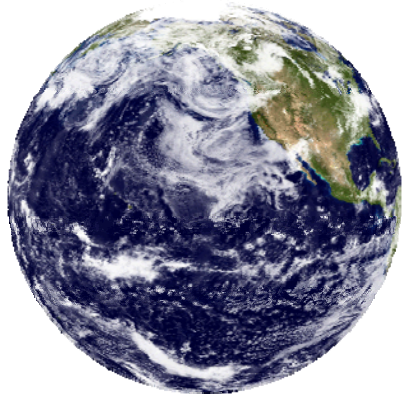
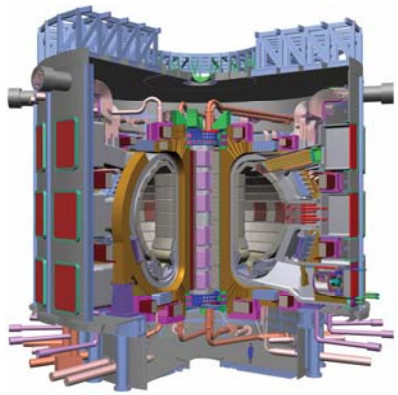
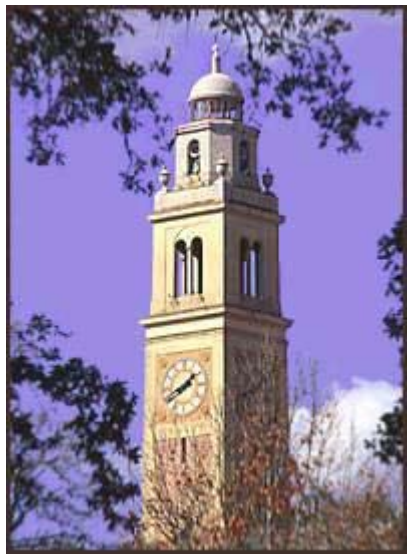


Vision2020 System



Summary : Vision2020 Characteristics

Parameter	Value
Gilgamesh component	
Clock frequency	16 GHz
MIND accelerators	32
FP operations per cycle (1 ALU)	8
Peak performance	4 TFLOPS
Memory capacity	512 MB
Dataflow component	
Clock frequency	32 GHz
FP operations per cycle (1 ALU)	1
Number of ALUs	256
Peak performance	8 TFLOPS
Dodecatron (single chip)	
Gilgamesh components	12
Dataflow components	1
Peak performance	56 TFLOPS
Memory capacity	6 GB
System	
Number of chips	128 K
Peak performance	> 7 EXAFLOPS
Dodecatron memory capacity	768 TB
Penultimate storage capacity	128 PB

Summary and Conclusions

- By the year 2020 Moore's Law in CMOS will be near flat lined around 10 nanometers
- The logic density & clock rate product projected to about 256X
- Temporal locality to drive computer architecture
- New execution model, mechanisms, and structures required to enhance efficiency, scalability, and work/energy
- Multi-billion-way parallelism
- Power TBD
- 1 Exaflops Linpack performance