5.5 A Wire-Speed Power™ Processor: 2.3GHz 45nm SOI with 16 Cores and 64 Threads

Charles Johnson, David H. Allen, Jeff Brown,
Steve Vanderwiel, Russ Hoover, Heather Achilles,
Chen-Yong Cher, George A. May,
Hubertus Franke, Jimi Xenidis, Claude Basso

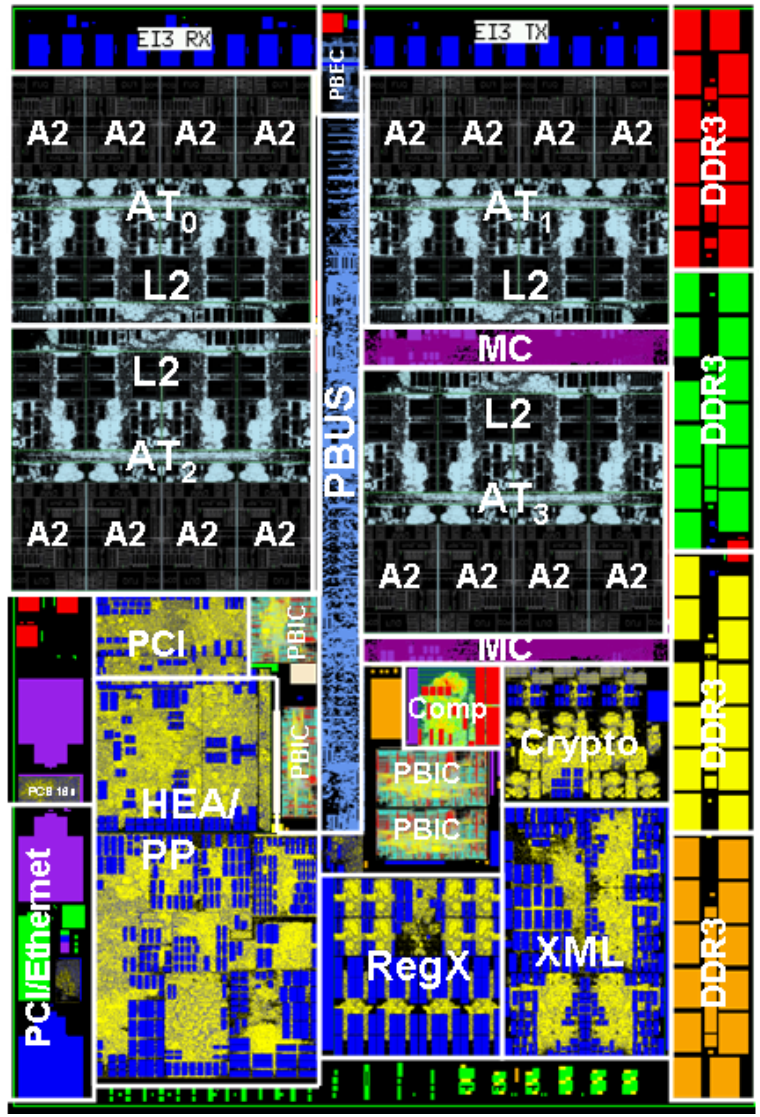
IBM Research & System Technology Group

Wire-Speed Processor: Goals

- Address new evolving market
- Address performance issues relating to “new” semiconductor scaling properties
- Reduced application power requirements
  - Achieve >50% thru Architecture
  - Achieve >50% thru Implementation

Wire-Speed Processor

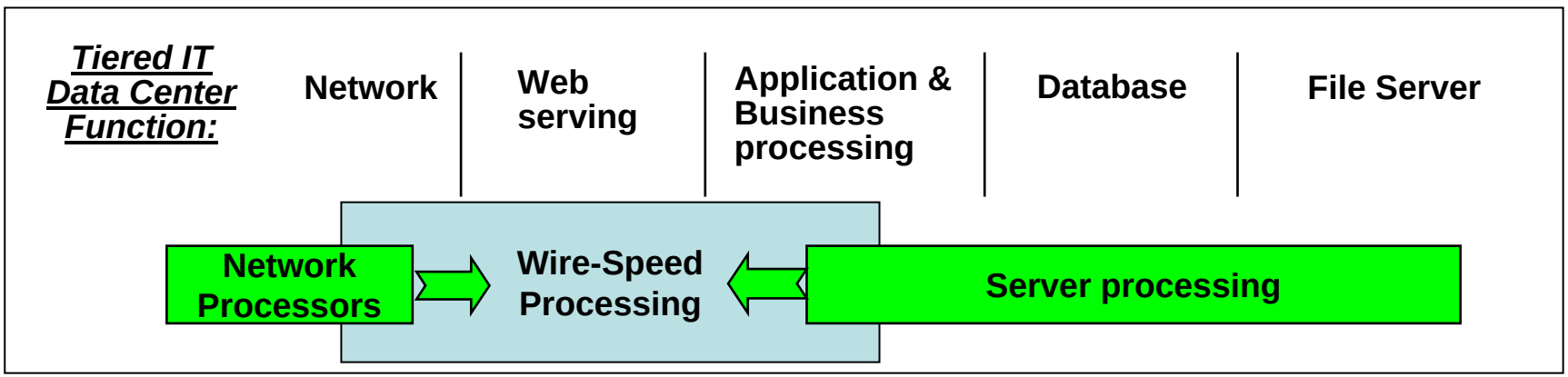
Technology	**IBM 45nm SOI**
Core Frequency	**2.3GHz @ 0.97V (Worst Case Process)**
Chip size	**428 mm2 (including kerf)**
Chip Power (4-AT node) **Chip Power (1-AT node)**	**65W @ 2.0GHz, 0.85V Max Single Chip** **20W @ 1.4GHz, 0.77V Min Single Chip**
Main Voltage (VDD)	**0.7V to 1.1V**
Metal Layers	**11 Cu (3-1x, 2-1.3x, 3-2x, 1-4x, 2-10x)**
Latch Count	**3.2M**
Transistor Count	**1.43B**
A2 Cores / Threads	**16 / 64**
L1 I & D Cache	**16 x (16KB + 16KB) SRAM**
L2 Cache	**4 x 2MB eDRAM**
Hardware Accelerators	**Crypto, Compression, RegX, XML**
Intelligent Network Interfaces	**Host Ethernet Adapter/Packet Processor** **2 Modes: Endpoint & Network**
Memory Bandwidth	**2x DDR3 controllers** **4 Channels @ 800-1600MHz**
System I/O Bandwidth	**4x 10G Ethernet, 2x PCI Gen2**
Chip-to-Chip Bandwidth	**3 Links, 20GB/s per link**
Chip Scaling	**4 Chip SMP**
Package	**50mm FCPBGA (4 or 6 layers)**



Wire-Speed Processor Attributes

A blurring of the Network and Server worlds

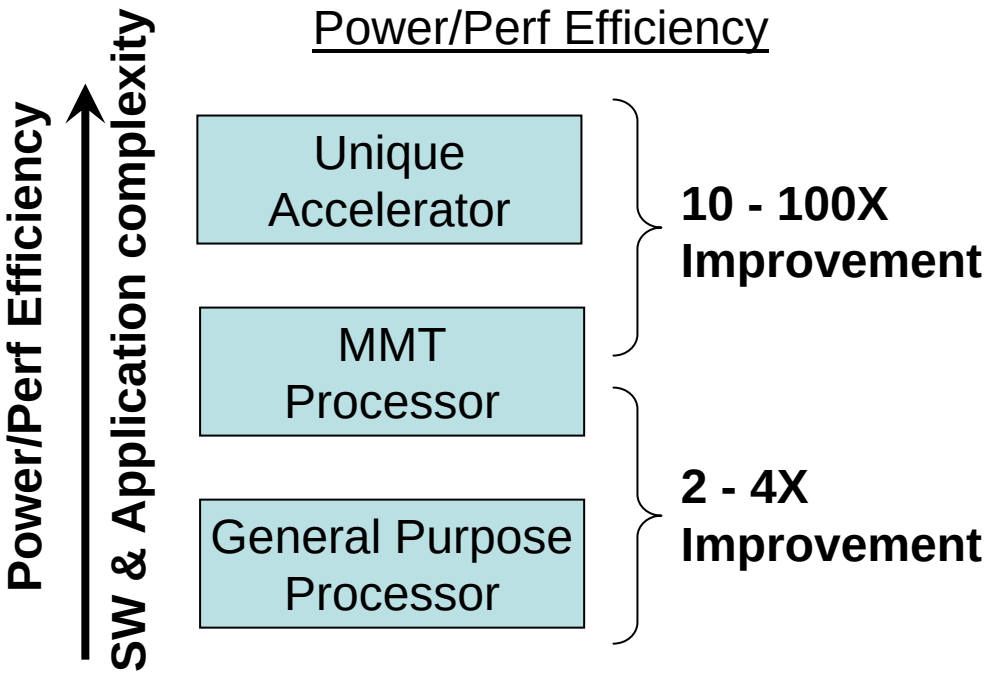
- Highly-multi-threaded low power cores with full ISA
- Standard programming models with OS's & hypervisors
- Smaller memory line sizes
- Accelerators: for both Networking & Application tiers
- Integrated Network system & Memory I/O
- Virtualization support
- Server RAS & infrastructure
- Low total power solution based on thruput optimization



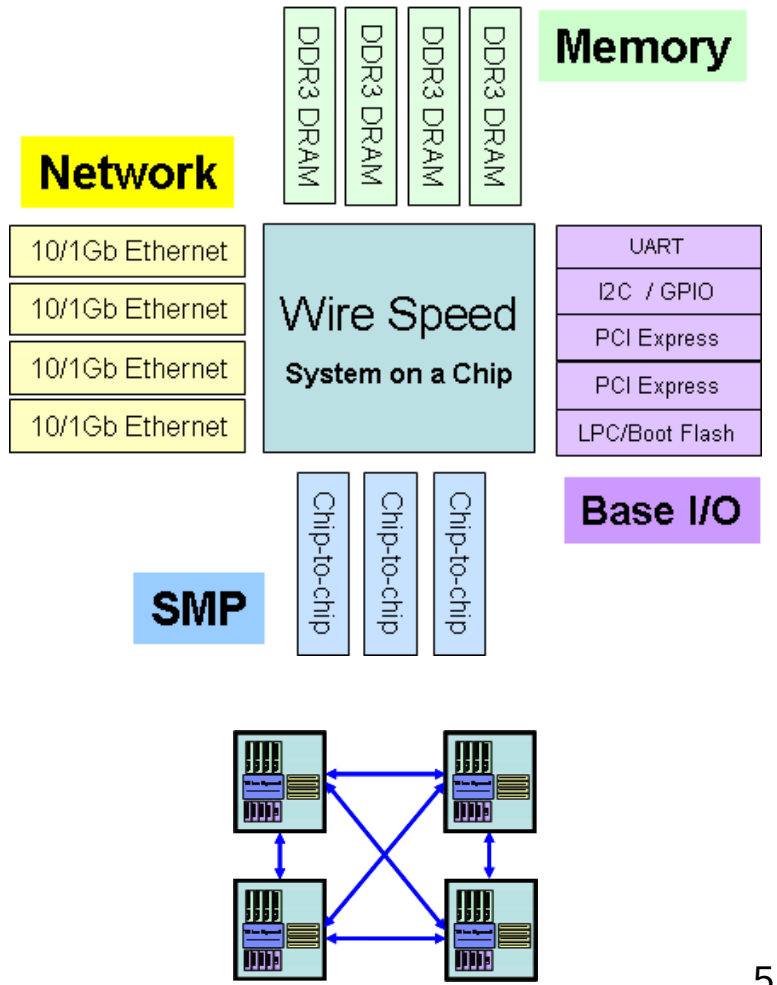
Power/Performance Scaling

Compute Scaling

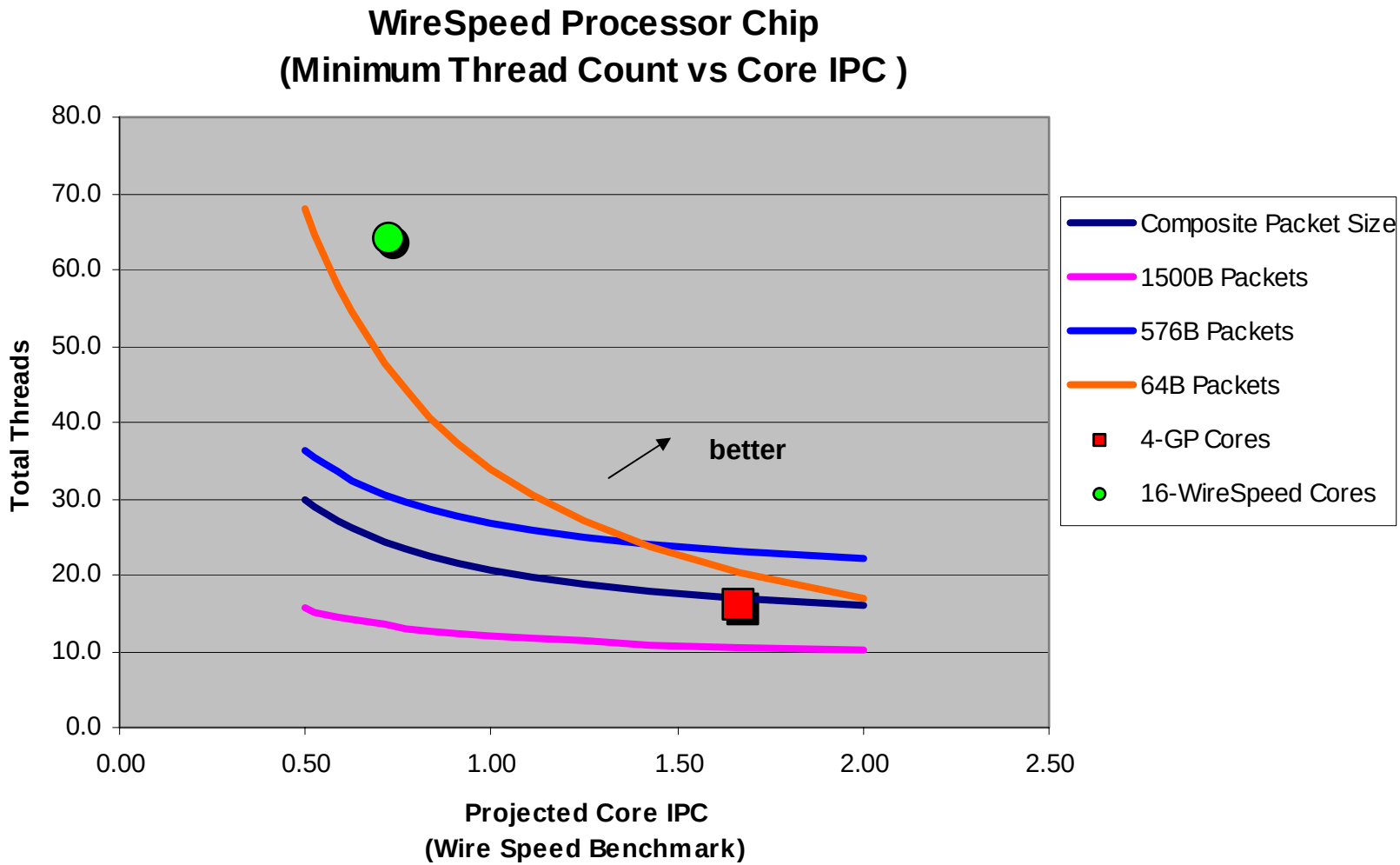
- Choose the right processor
- Choose the right mix of processors & accelerators
  - Accelerators have best power/perf efficiency, but only do one function



SMP Scaling



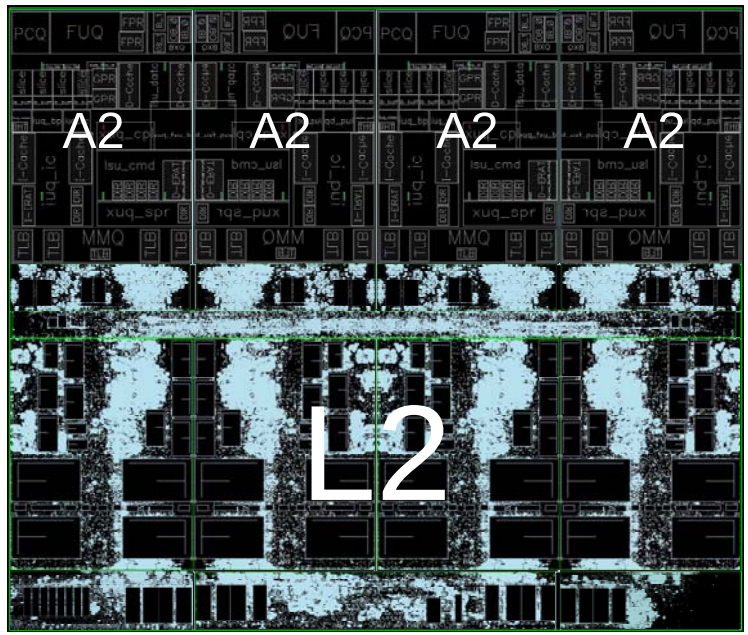
Wire-Speed Computing: Threads vs IPC



A2 Core/AT Node

- A2 Core
  - Full 64b Power ISA Embedded Architecture
    - Virtualization / Hypervisor support
    - DP Floating Point
  - Power efficient throughput optimized micro architecture
    - 4 way SMT, In-order 2 way concurrent issue
  - Enhanced instruction set for low latency accelerator communications

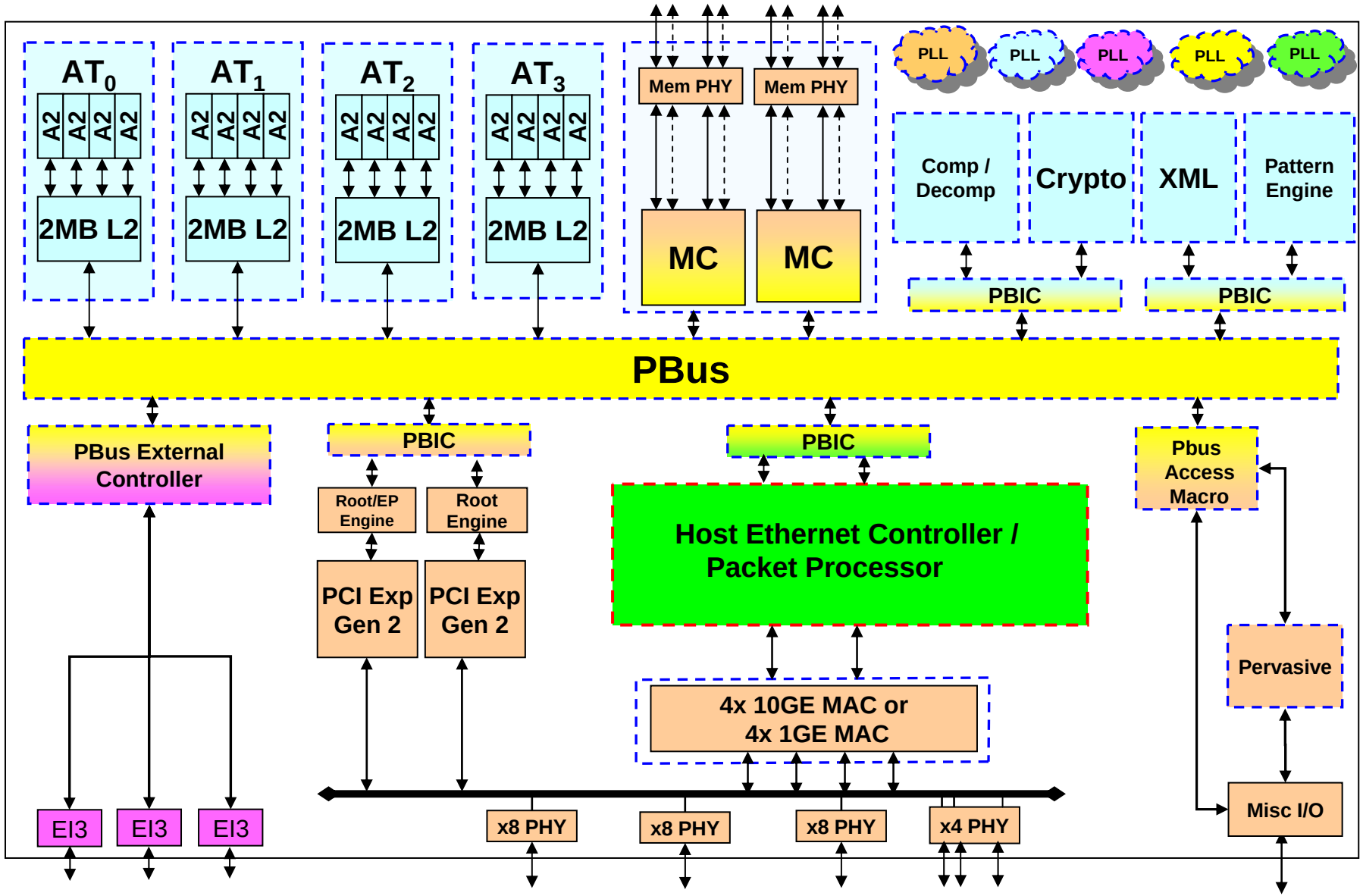
- AT Node
  - Fully coherent shared 2MB L2 cache
  - High speed interconnect for many cores
  - Power efficient EDRAM
  - Lots of cache controls
    - Line locking
    - Streaming
    - Slave memory
    - Injection



Wire-Speed Accelerator Bandwidths

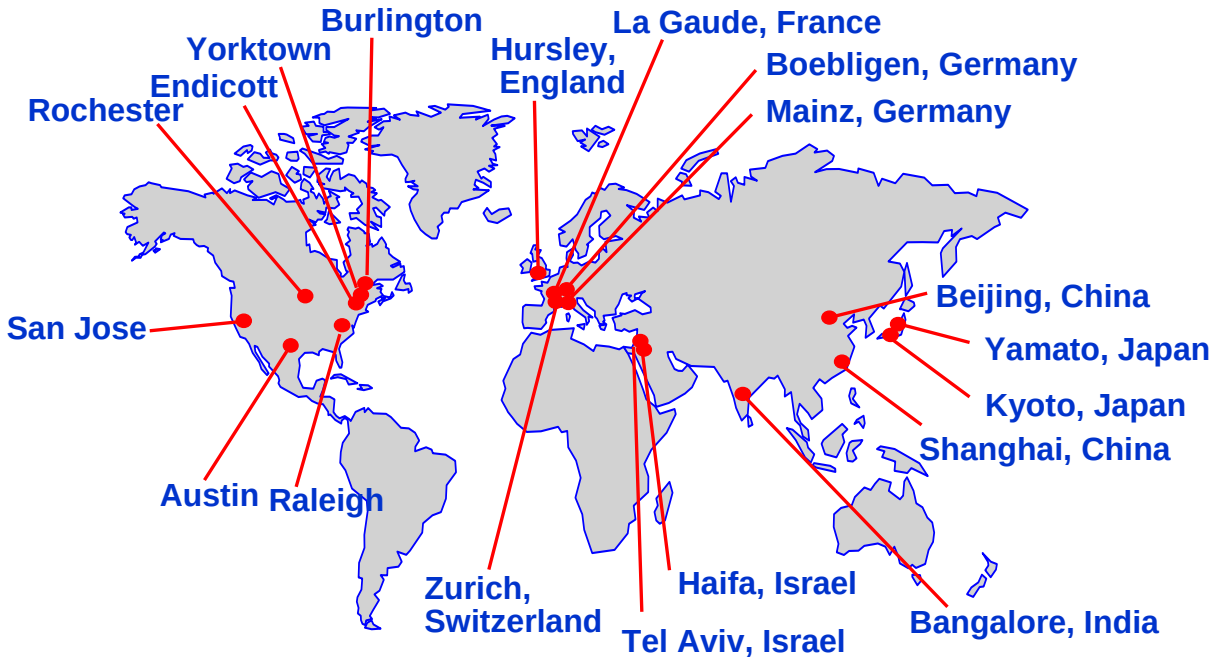
Accelerator Unit	Algorithm	# of Engines	Projected Bandwidth	
			Typical	Peak
HEA	network node mode	4	40 Gbps	40 Gbps
	endpoint mode	4	40 Gbps	40 Gbps
Compression	gzip (input bandwidth)	1	8 Gbps	16 Gbps
	gunzip (output bandwidth)	1	8 Gbps	16 Gbps
Encryption	AES	3	41 Gbps	60 Gbps
	TDES	8	19 Gbps	
	ARC4	1	5.1 Gbps	
	Kasumi	1	5.9 Gbps	
	SHA	6	23-37 Gbps	
	MD5	6	31 Gbps	
	AES/SHA	3	19-31 Gbps	
	RSA/ECC (RSA with 1024/2048 bit key)	3	45000/7260	
XML	Customer workload	4	10 Gbps	30 Gbps
	Benchmark workload	4	20 Gbps	
RegX	For typical pattern sets	8	20-40 Gbps	70 Gbps

Wire-Speed Processor



Wire-Speed Processor:

SoC Design Enabled Distributed Development Team

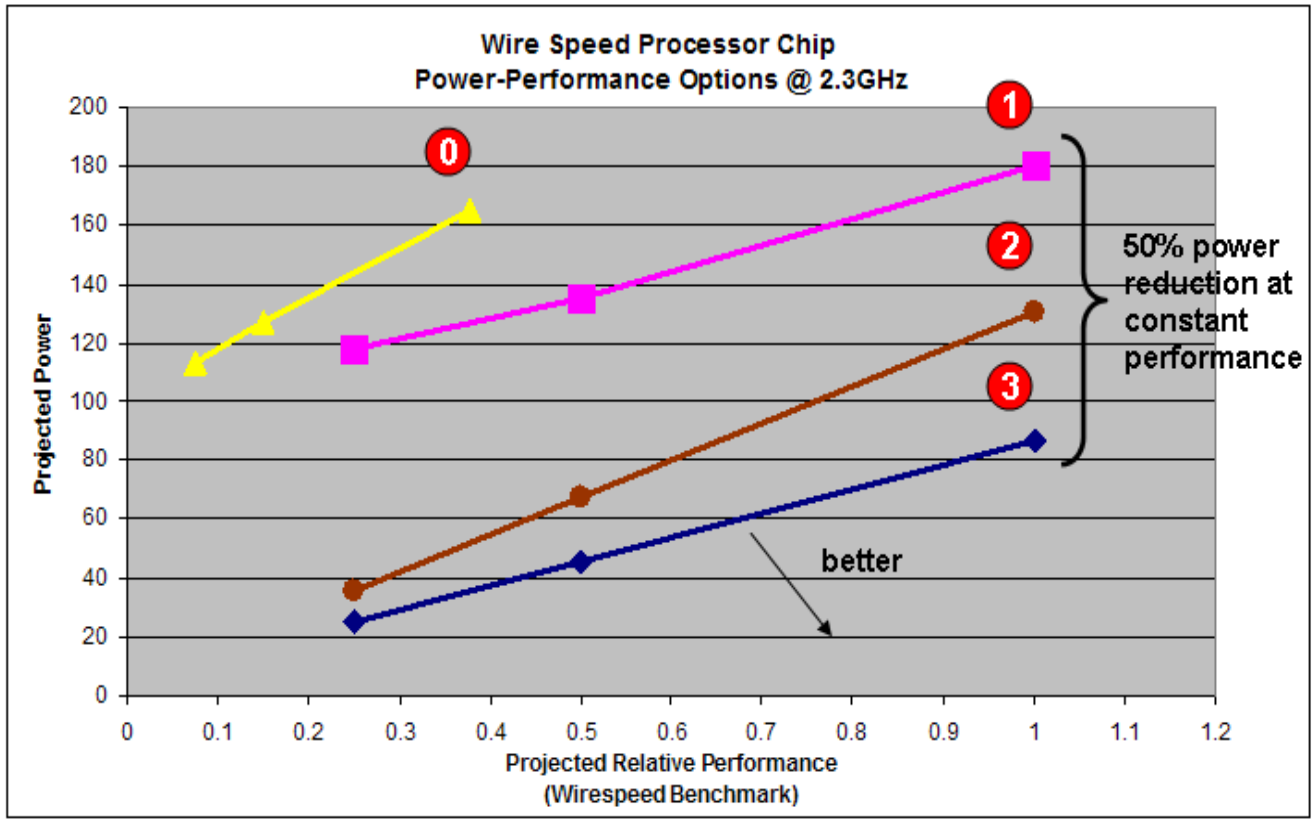


- **Verification is key to complex SoC**
  - **1.43 billion transistors**
  - **1.5 teracycles Chip / Unit Sim**
  - **22 teracycles System Sim**
- **19 World Wide Locations Participating**

Evolution of a Wire-Speed Processor

Power Savings from Architecture / Integration

- **Low Power Integrated Design**
  - 1 -- Accelerators
  - 2 -- I/O (Ethernet/HEA)
  - 3 -- Cores/Threads
  - -- In 25-75 W Sockets
- **Eliminates I/O Bottlenecks**
- **Improves Latency & State Replication**
- **Scalable Solution**
  - -- 10 => 40 => 100 GbE



[General Purpose Processor chip]

⇩

[WireSpeed chip]

0 : 8- GP cores

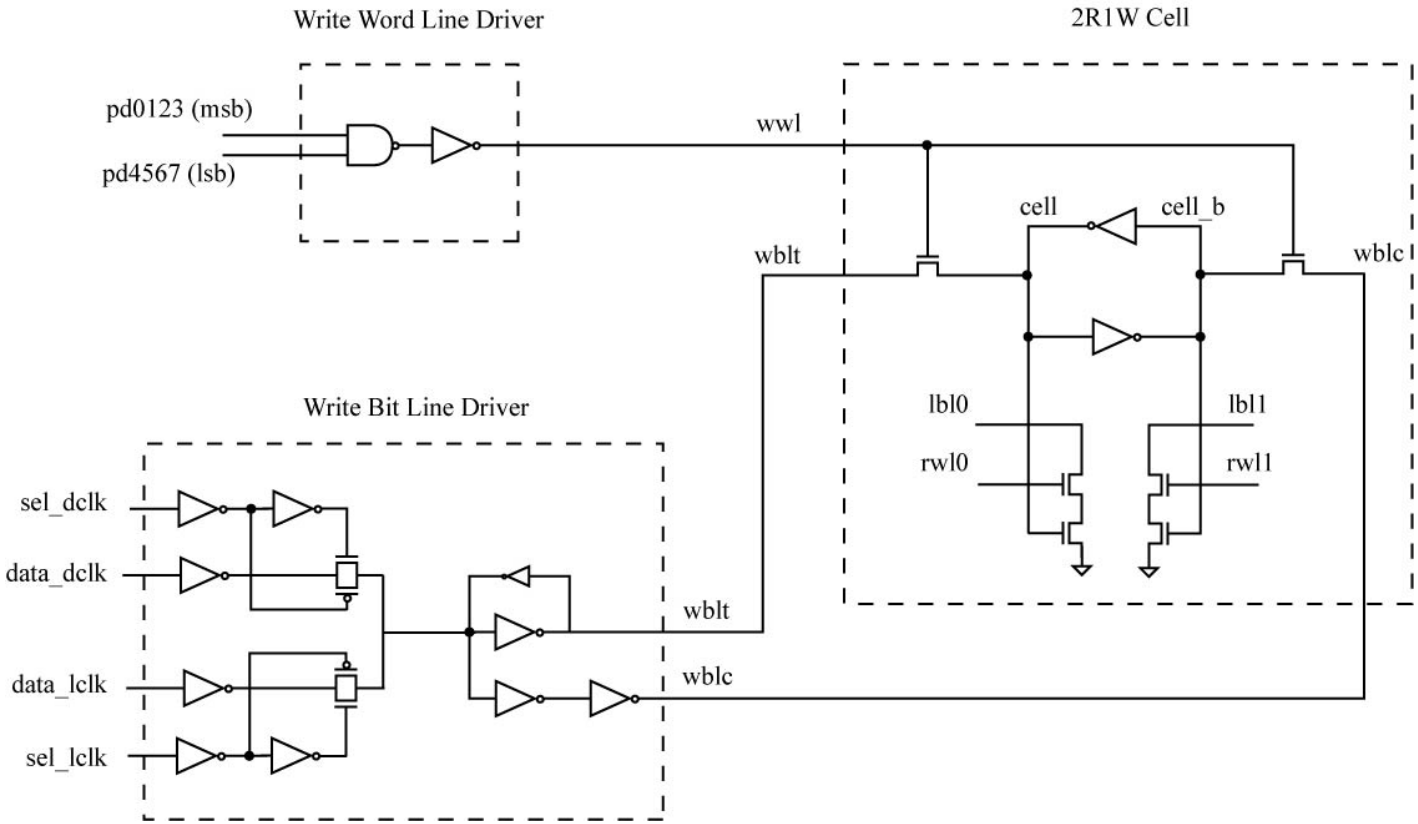
1 : 8- GP cores + 4 Accelerators

2 : 8- GP cores + 4 Accelerators + Integrated I/O (DDR3, 10Ge, HEA)

3 : 16- A2 cores + 4 Accelerators + Integrated I/O (DDR3, 10Ge, HEA)

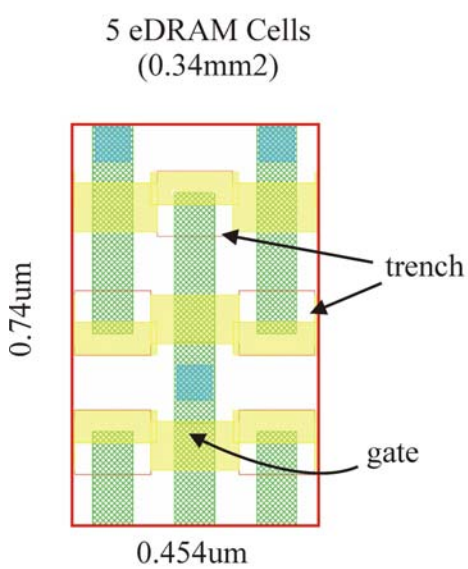
Double Pumped Register File

- 64 total register file instances required in the 16-A2 cores.
- Double pumping a 2R1W cell provided 2R2W functionality
  - at 55% of the area and 85% of the power.
- Additional power savings from smaller cell and reduced performance requirements on critical read access logic

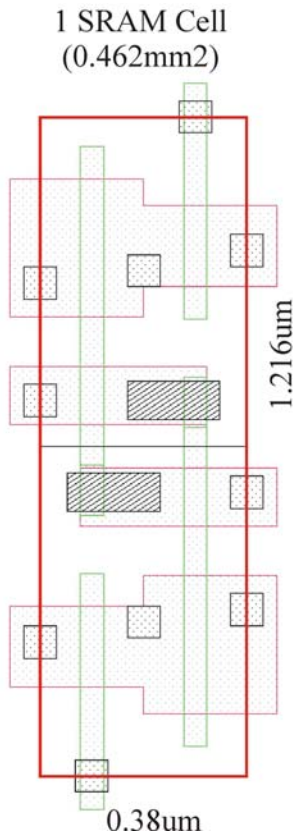


45nm SOI eDRAM Cells Compared to SRAM Cell

- Density Ratio
  - 6.8x at cell level
  - 3.1x at array level
- Power Ratio
  - 0.2x /mm2 at array level
- Performance Ratio
  - Cycle: 3.2x (2.0x application)
  - Latency: 2.5x (1.3x application)
- Application net result
  - Similar performance in <0.5x area and <0.2x power



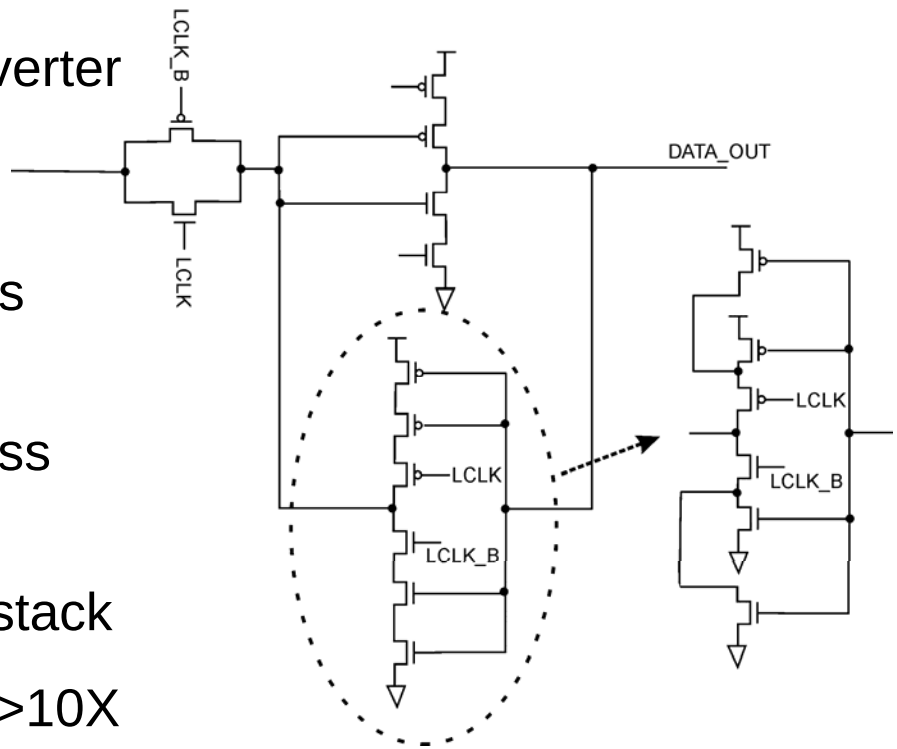
Cycle = 1.7ns
Latency = 1.35ns



Cycle = 0.53ns
Latency = 0.53ns

Asynchronous Domain Crossing Latch Optimization

- Soft error rate (SER): asynchronous latches used triple-stack feedback inverter to reduce diffusion cross-section.
- Latch metastability: large numbers of asynchronous clock domain crossings increased the failure rate
- SER rate was orders of magnitude less than the metastability failure rate
- Latches redesigned to remove triple stack
- Latch Metastability rate improved by >10X
- Modification required no size, pin location, or blockage changes on the latches



Wire-Speed: Transistor Characteristics

- Device Scaling (0.7V)
  - 1.50x Perf
  - 5.65x Leakage Power
- Device Scaling (0.9V)
  - 1.37x Perf
  - 4.61x Leakage Power

	TT	TT		TT	TT
	85C	85C		85C	125C
	0.7V	0.9V		0.9V	0.9V
Transistor	Performance (GHz/stage)			Leakage (nA/stages)	
RVT	48	73		1300	2300
HVT	37	61		550	950
SVT	32	53		230	500
	Performance (Normalized)			Leakage (Normalized)	
RVT	1.50	2.28		5.65	10.00
HVT	1.16	1.91		2.39	4.13
SVT	1.00	1.66		1.00	2.17

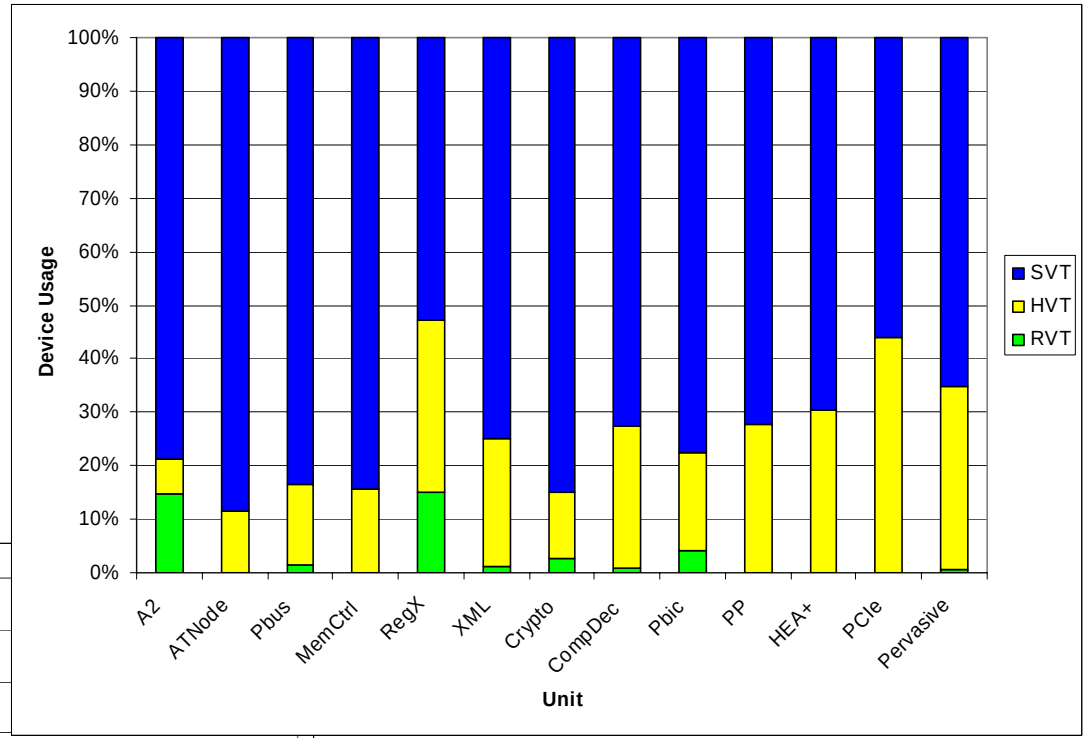
Power Design Parameters (Pass 1)

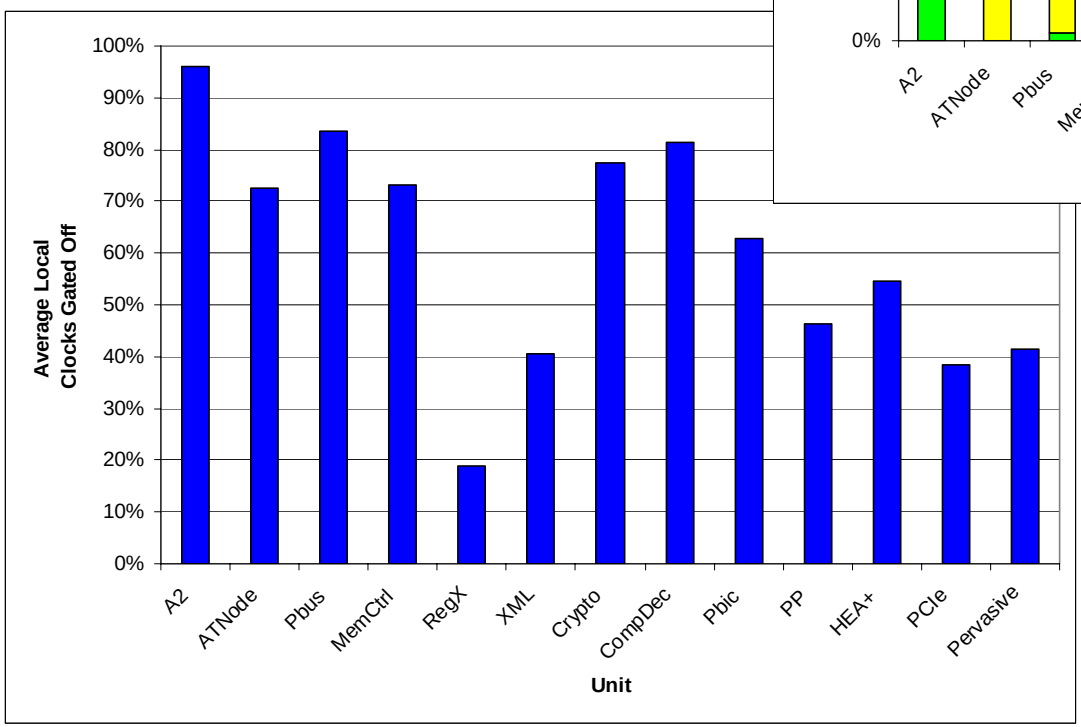
VT Usage Targets:

75% SVT

20% HVT

5% RVT





Clock Gating Target:

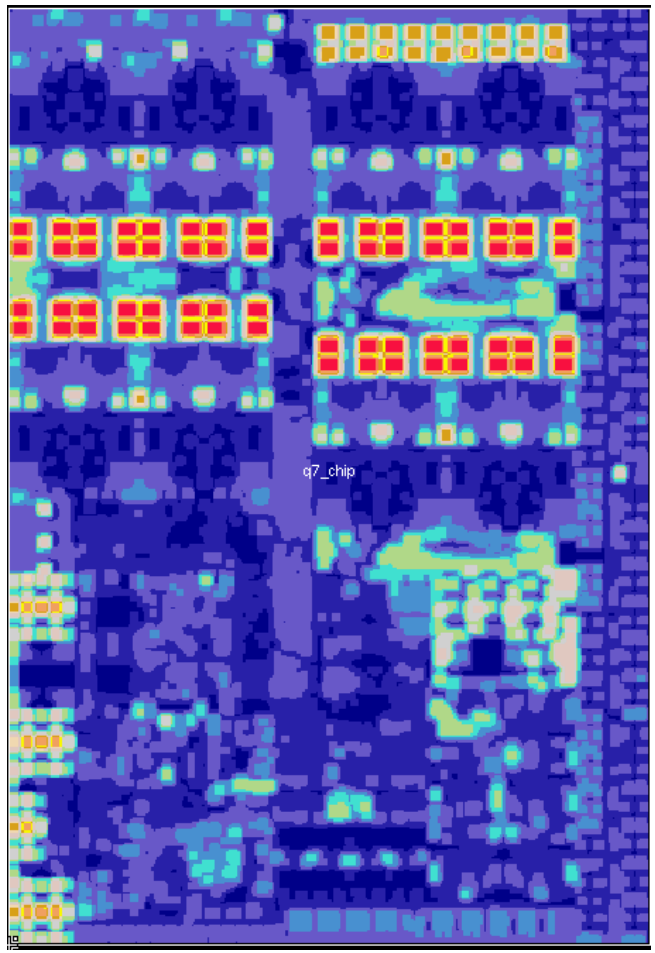
90%

Voltage Domains

Supply		Nom Voltage @BGA (V)	@C4 Vmin (V)	@Cct Vmin (V)	@Cct Vmax (V)
VDD	*Logic*	**1.20**			1.22
		1.01	0.96	0.90	n/a
		0.89	0.85	0.79	n/a
		0.70	0.67	0.63	n/a
	Combo Phy	**1.01**	0.93	0.86	1.03
		0.89	0.82	0.76	n/a
VCS	*Array*	**1.19**	1.08	1.05	1.18
		1.11	1.02	0.99	1.11
VMEM	*Memory*	**1.50**	1.37	1.29	1.56
VIO	*EI3 IO*	**1.20**	1.11	1.04	1.25
VTT	*Combo PHY termination*	**1.20**	1.09	1.01	1.23
VTTA	*Combo PHY drivers*	**1.20**	1.09	1.03	1.21
DVDD	*Low Speed Device I/O*	**3.30**	3.05	2.87	3.44
DVDD2	*Low Speed Device I/O Bias*	**1.80**	1.67	1.57	1.87
AVDD	*Various PLL analog*	**1.80**	1.72	1.62	1.87
GND	*Logic & array ground*	**0.00**			
AGND	*Various Analog PLL ground*	**0.00**			

DT Decoupling Capacitors

- Deep trench (DT) decoupling capacitors occupy 10.2% of chip area
- DT shapes ~4% of chip area
- ~6 uF decoupling cap on VDD
- 30mv noise reduction
  - 5W power savings



DT density

AT Node: Vdd-Freq-Power Scaling

- 11% Vdd Reduction
- 22% Frequency Reduction
- 41% Power Reduction
  - -- 42% AC Power Reduction
  - -- 39% DC Power Reduction

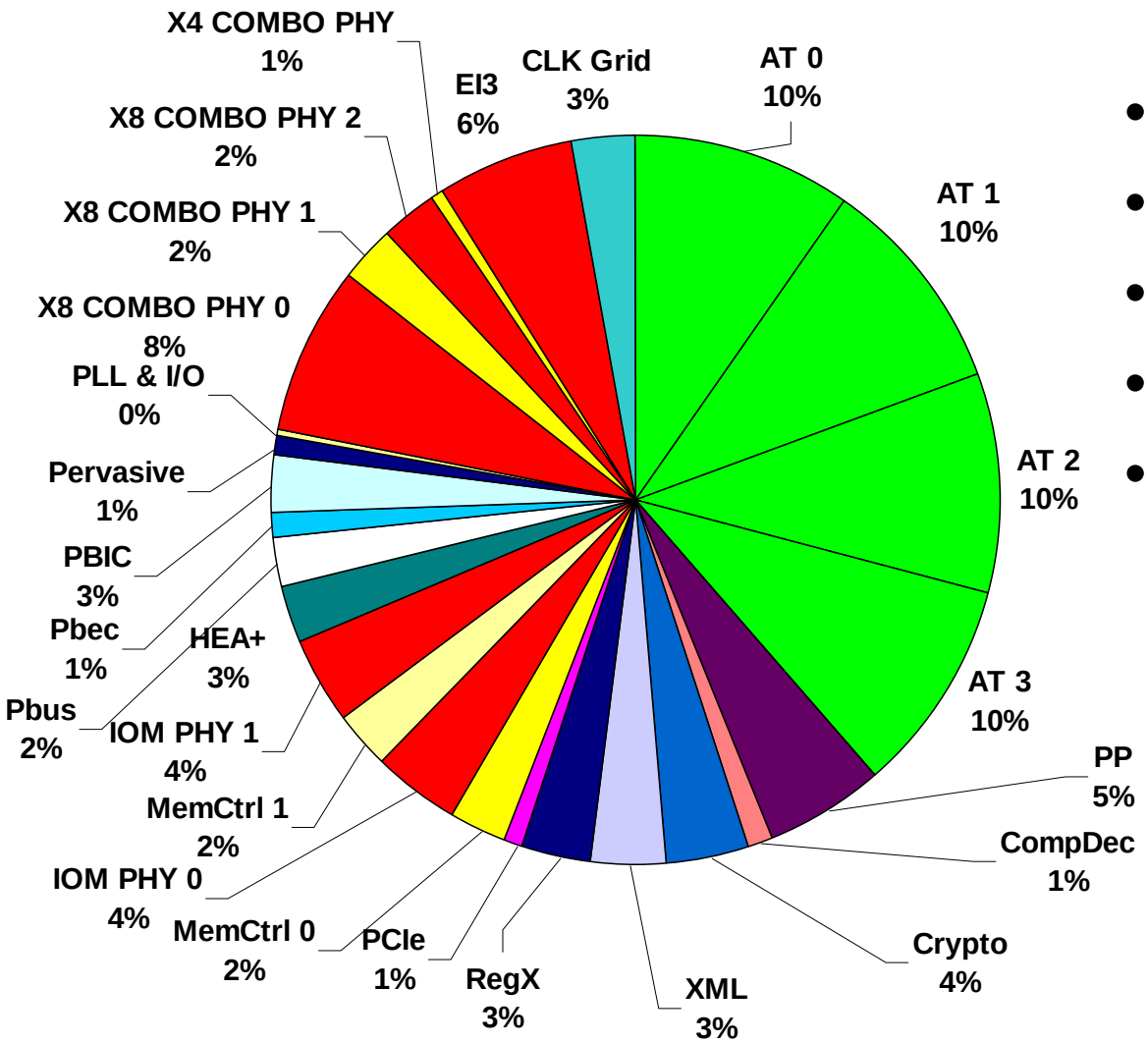
	2.3 Ghz		2.0 Ghz		1.8 Ghz	
	A2	L2 nest	A2	L2 nest	A2	L2 nest
Temperature (ºC)	85		85		85	
Frequency (Ghz)	2.30		2.00		1.81	
VOLTAGE @ C4	0.97		0.89		0.86	
VOLTAGE @ Circuit	0.89		0.82		0.79	
Instances	4	1	4	1	4	1
Logic AC (W)	3.612	2.294	2.517	1.598	2.058	1.307
SRAM AC (W)	0.959	0.196	0.702	0.143	0.587	0.120
DRAM AC (W)		0.259		0.194		0.164
Logic DC	1.181	0.727	0.835	0.520	0.715	0.448
SRAM DC	0.129	0.121	0.090	0.085	0.077	0.073
DRAM DC		0.079		0.055		0.047
CLK Grid (A2s + AT)		0.772		0.551		0.455
AT Power		**10.329**		**7.290**		**6.052**

Wire-Speed Processor:

Projected Chip Power at Various Frequencies / Functionality

Power	AT Nodes	Frequency	Comments
85W	4	2.3 GHz	Full function/frequency
65 W 56 W	4 2	2.0 GHz 2.0 GHz	Power limiter in operation, disabled EI3, PBEC, XML
50 W 41 W 36 W	4 2 1	1.8 GHz 1.8 GHz 1.8 GHz	disabled EI3, PBEC, XML PPI, PCI-E @ 50% frequency DDR3 @ 1333MHz
20-22 W	1	1.4 GHz	1 mem controller & IOM PHY @ 800MHz No accelerators enabled. PBus @ 700MHz. HEA/PP grid @ 312.5GHz, (2) 1 Gb ethernet ports. PCI-E @ 250MHz, I/O is GEN 1 @ 2.5Gbps, 4 lanes. PBEC & EI3 links disabled.

Wire-Speed Processor:
Projected power breakdown by function @ 2.0GHz



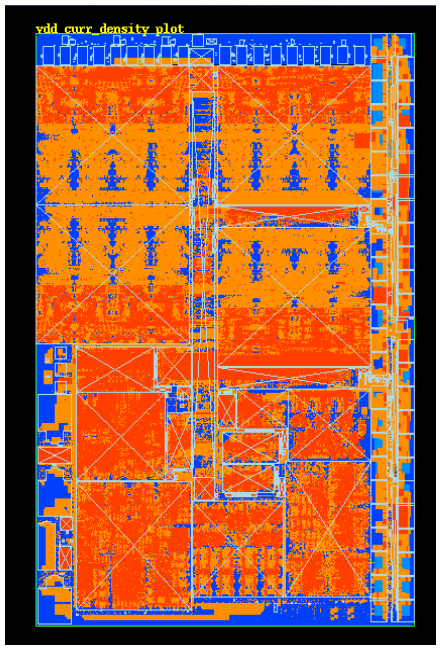
- Cores/Caches: 40%
- Accelerators: 11%
- I/O & Bus Logic: 19%
- PHY: 27%
- Clock Grid: 3%

Power Reduction “Big Hitters”

Technique	**Impact**	**Power Reduction (W)**
Static voltage scaling	Lower voltage & 50% reduced DC leakage on fast parts	10-12W
Clock skew reduction	0.4W/ps	4-6W
VT optimization	Improved timing on critical paths, lower VDD @ frequency	8-12W
Deep trench decoupling capacitors	30mV lower AC noise	5W
Clock gating	50% lower AC power	18-22W
Net		**45-57W**

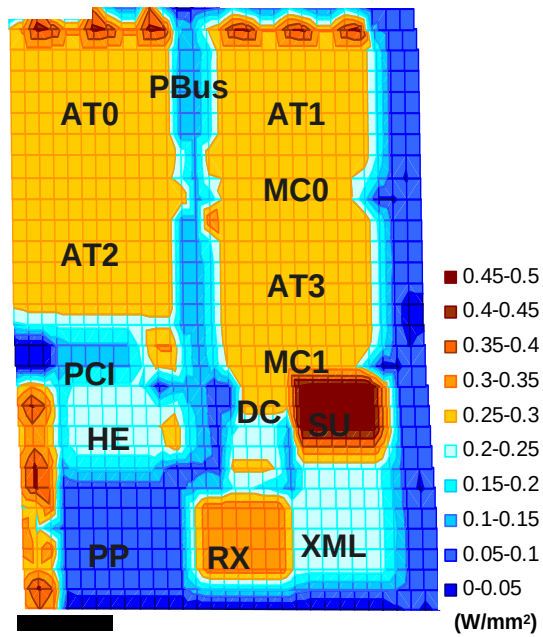
Wire-Speed: Power/Current Maps

Vdd Current Density



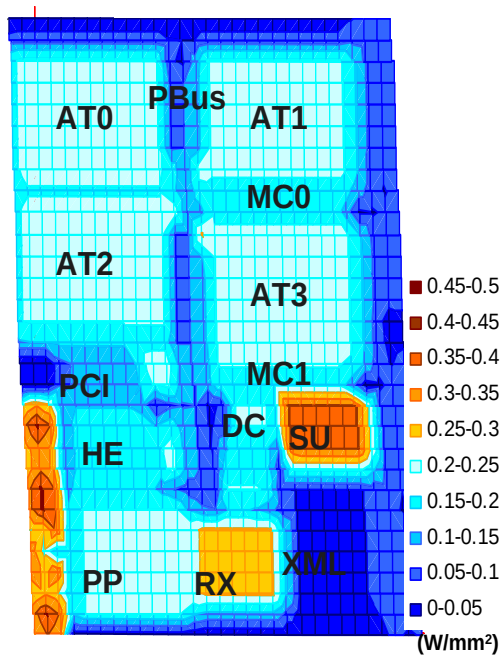
0.925V, 2.3GHZ

Multi-Chip Endpoint mode _2.3GHz



Total power: 87.3W

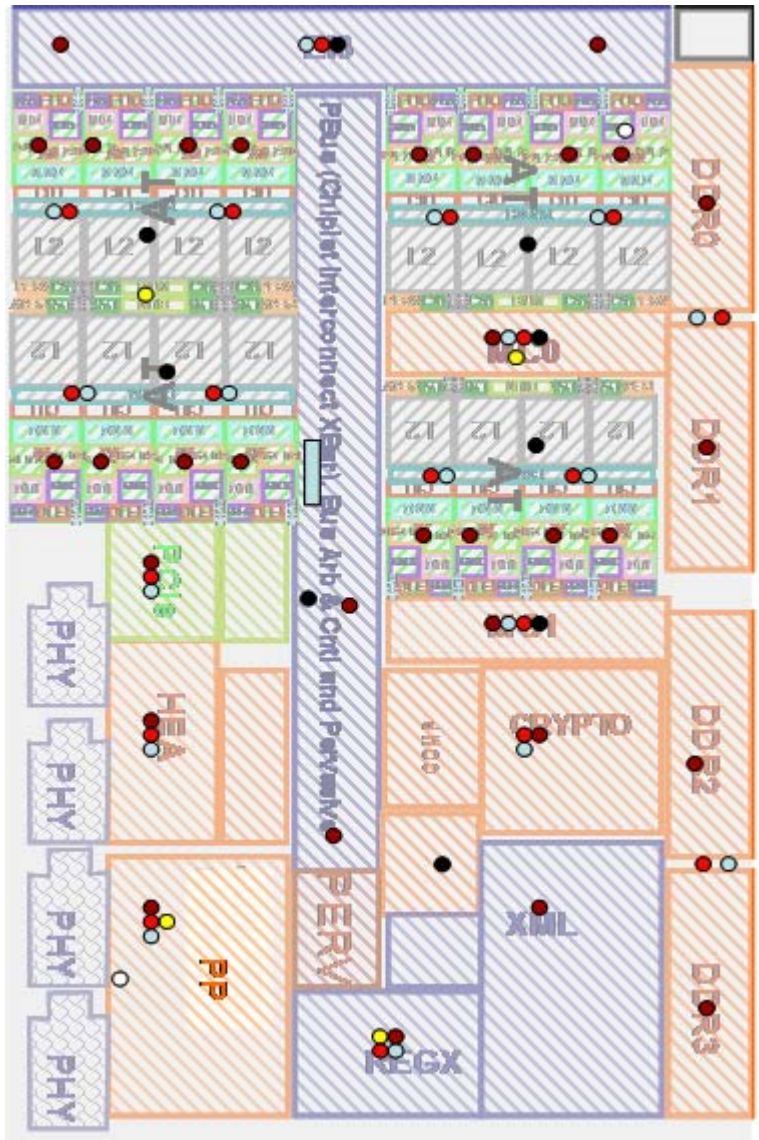
Single-Chip Packet mode 2.0GHZ



Total power: 65.4W

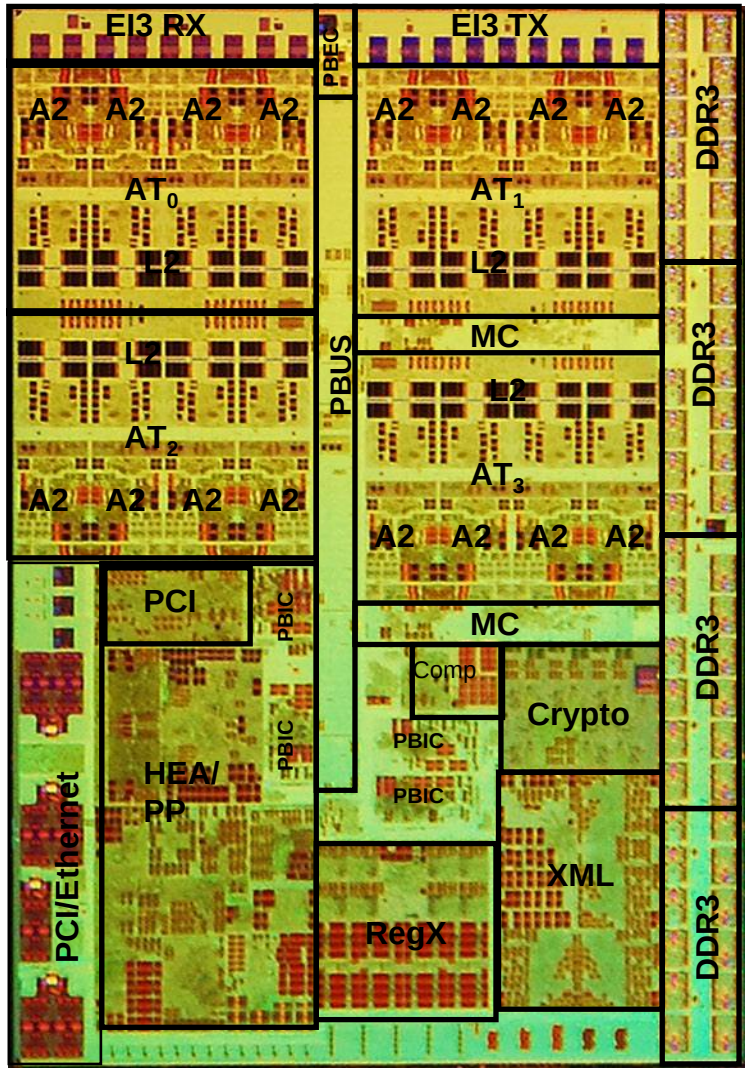
Sensor Summary

- PSRO (32)
  Performance Screen Ring Oscillator
- MHCRO (1)
  Model Hardware Correlation Ring Oscillator
- Skitter (10)
- DTS2 (18)
  Digital Thermal Sensor
- Thermal Diodes (4)
- CPM (18)
  Critical Path Monitor
- Kelvin Probe (2)



Wire-Speed Power™ Processor: Summary

- New architecture addresses converging market between the network and servers
- Low power, highly scalable design
  - 25-75W chip at full frequency (2.3GHz)
- Significant power reduction thru architecture & implementation
  - Architecture: simple highly threaded cores, accelerators & integrated I/O
  - Implementation: static voltage scaling, clock skew reduction, VT optimization, deep trench decoupling capacitors, clock gating



Special thanks to
STG Development Team