Beyond Multi-core: The Dawning of the Era of Tera

Intel™ 80-core Tera-scale Research Processor

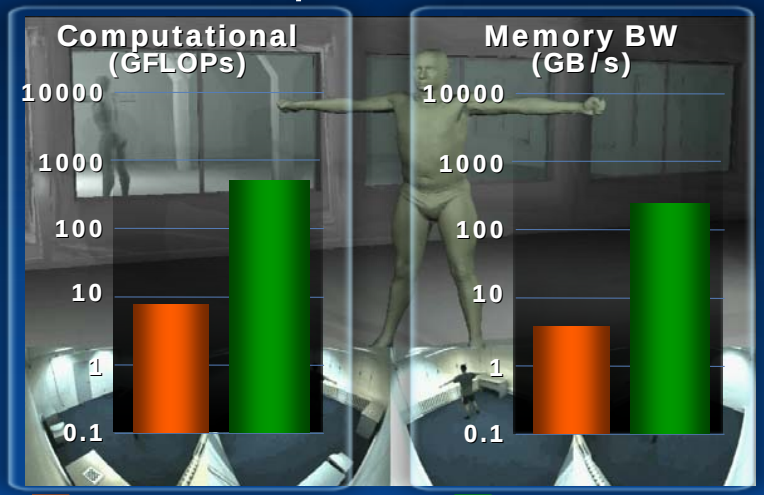
Jim Held
Intel Fellow & Director,
Tera-scale Computing Research
Intel Corporation

Agenda

- Tera-scale Computing
  - Motivation
  - Platform Vision
- Teraflops Research Processor
  - Key Ingredients
  - Power Management
  - Performance
  - Programming
  - Key Learnings
- Work in progress
- Summary

Emerging Applications will demand Tera-scale performance

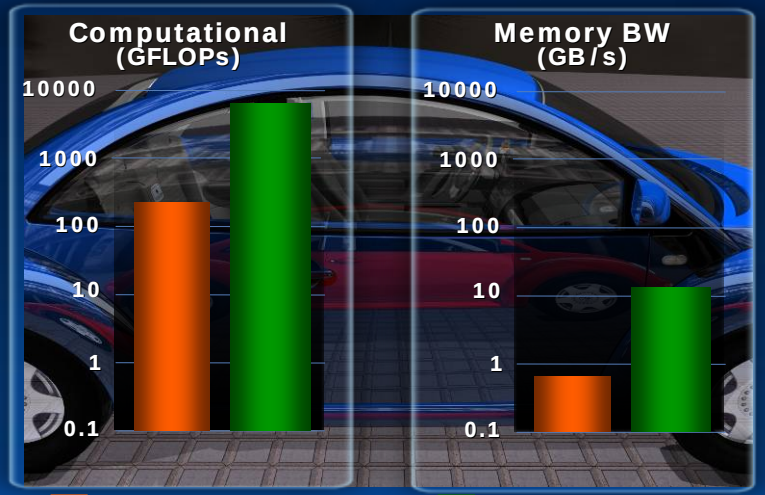
Computer Vision



2 Webcam Video Surveillance

4 Camera Body Tracking

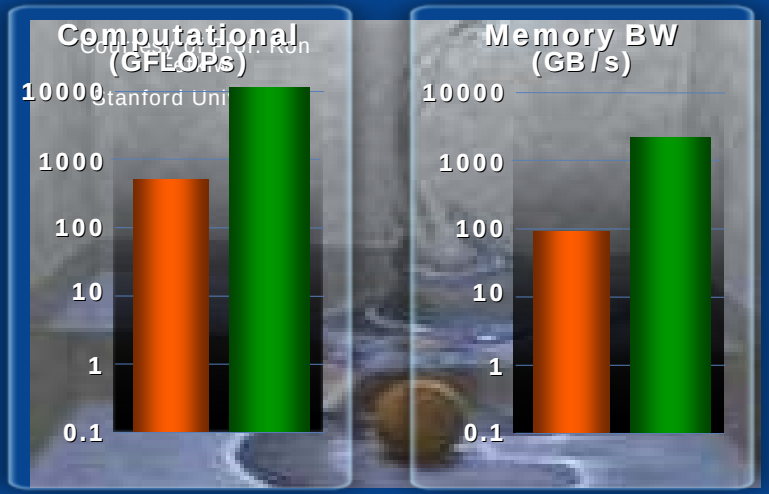
Ray-Tracing



Bar Scene, 1M pixel

Beetle Scene 1M pixel

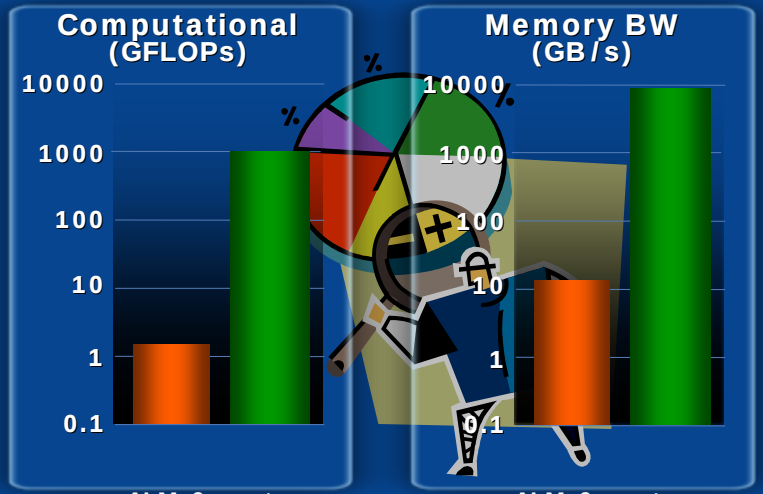
Physical Simulation



CFD, 75x50x50, 10 fps

CFD, 150x100x100, 30 fps

Financial Analytics

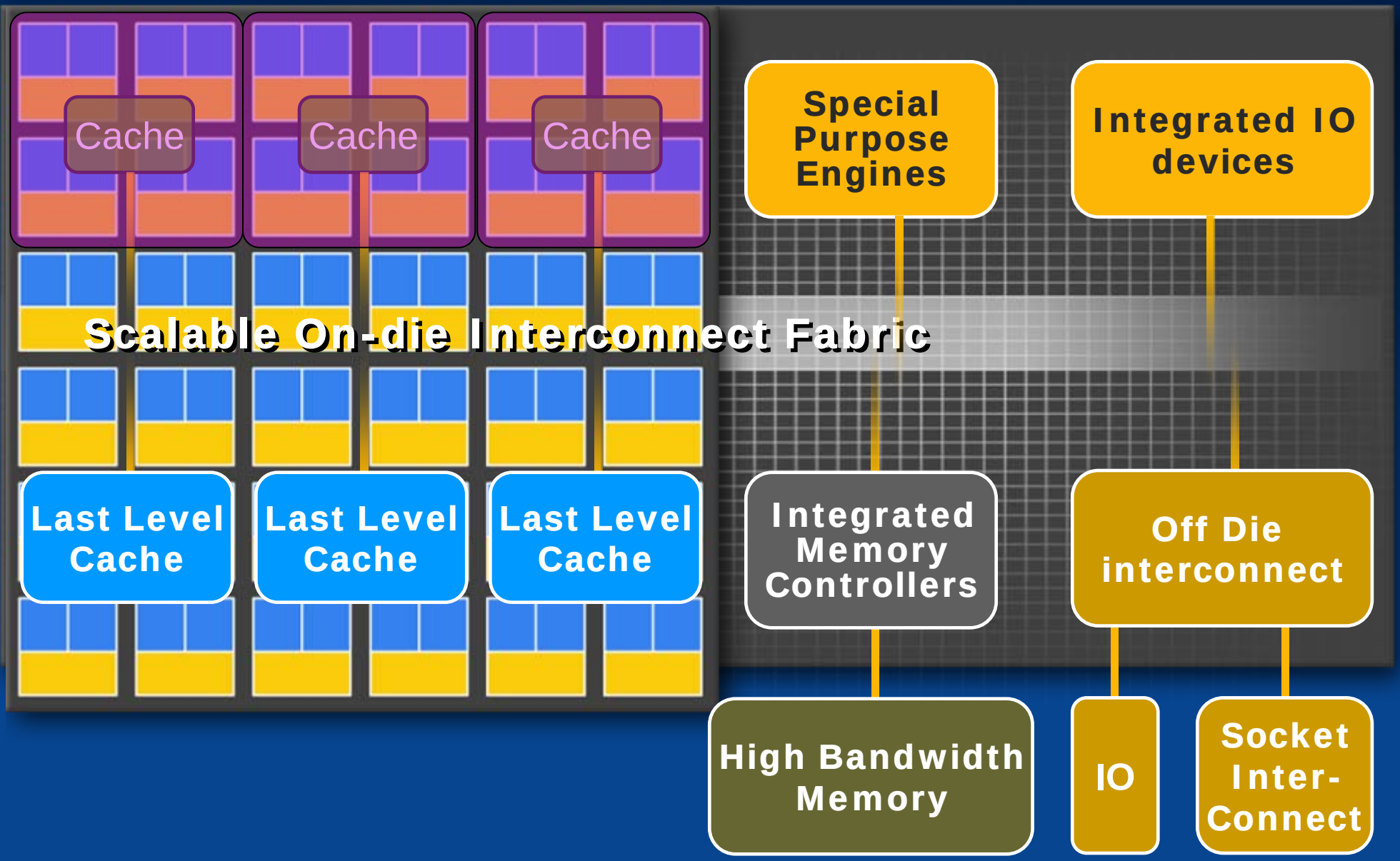


ALM, 6 assets, 7 branches, 1 min

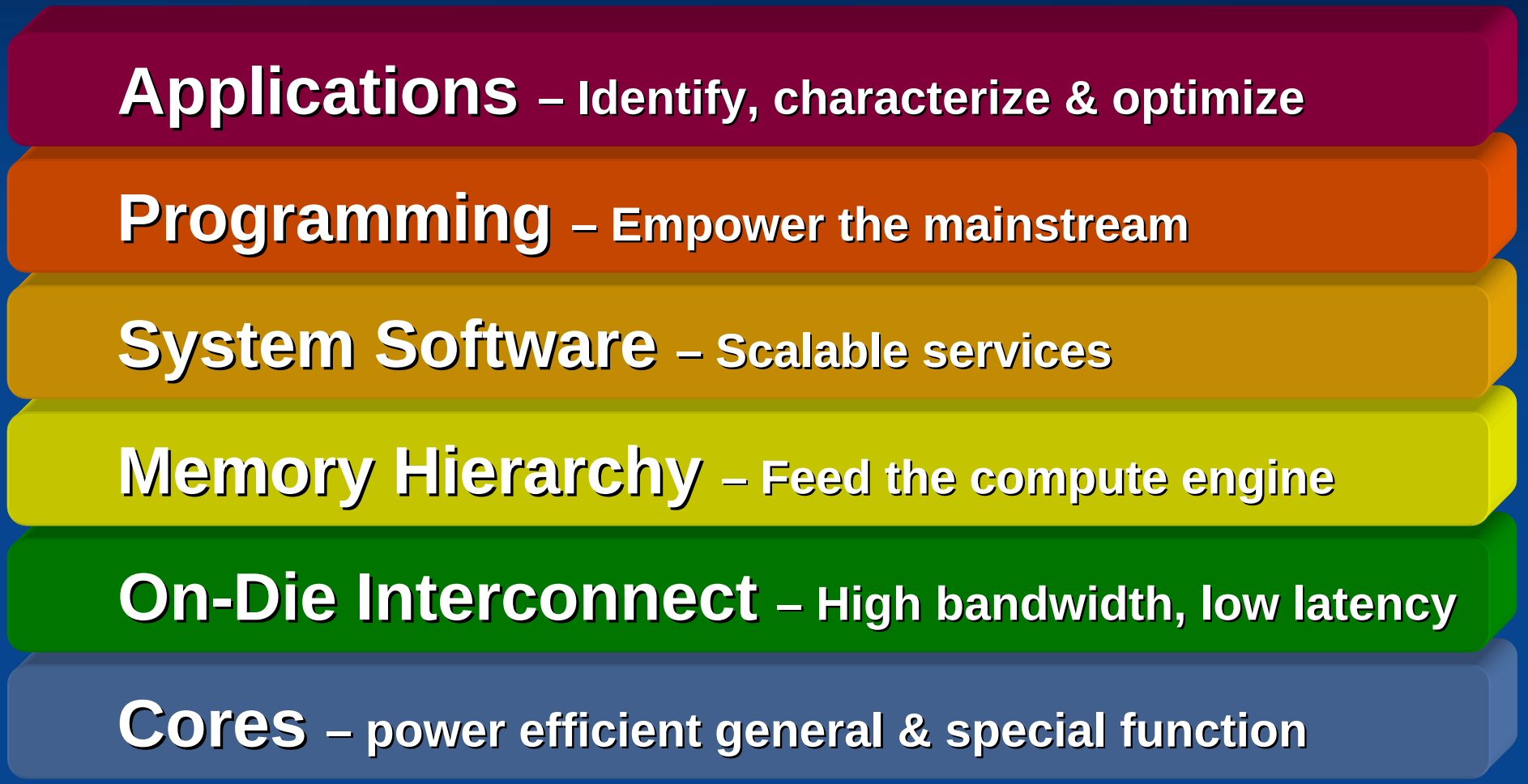
ALM, 6 assets, 10 branches, 1 sec

intel

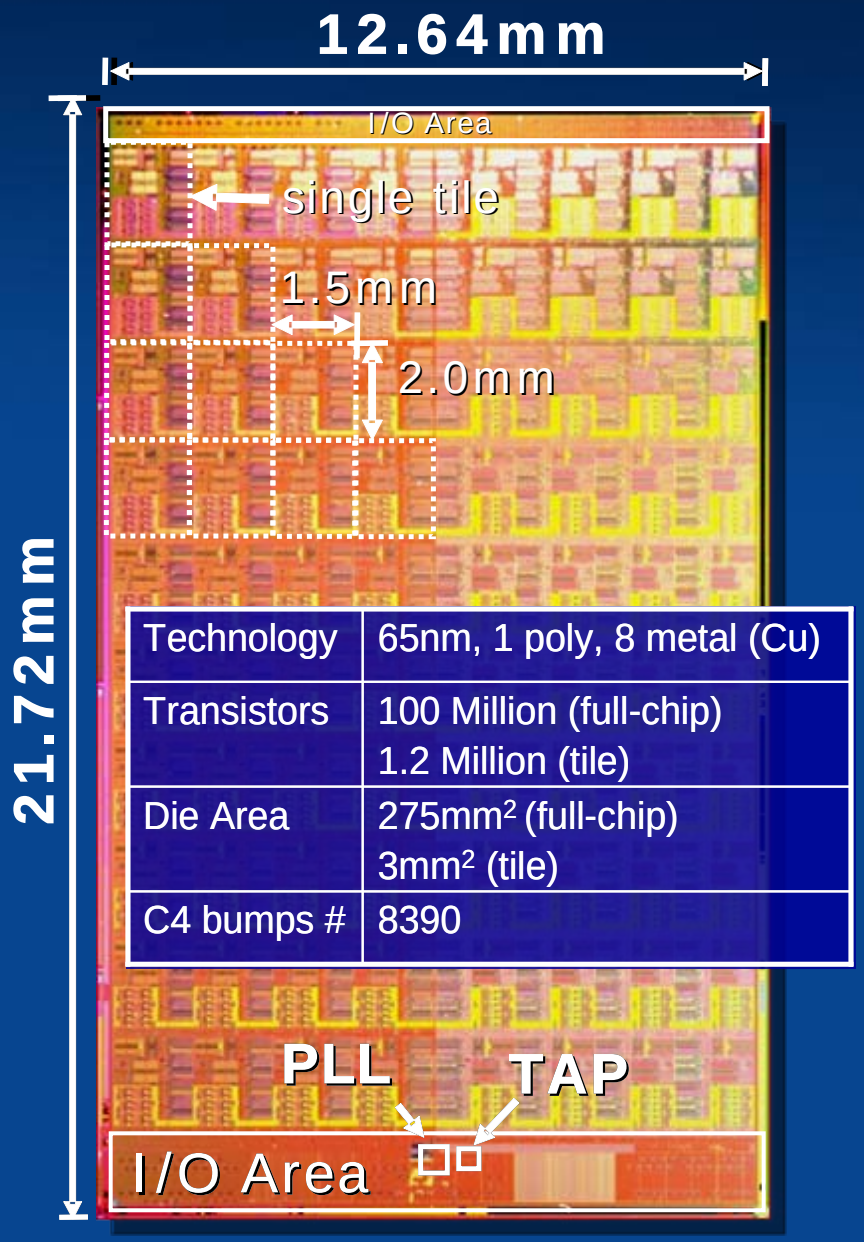
A Tera-scale Platform Vision



Tera-scale Computing Research



Teraflops Research Processor

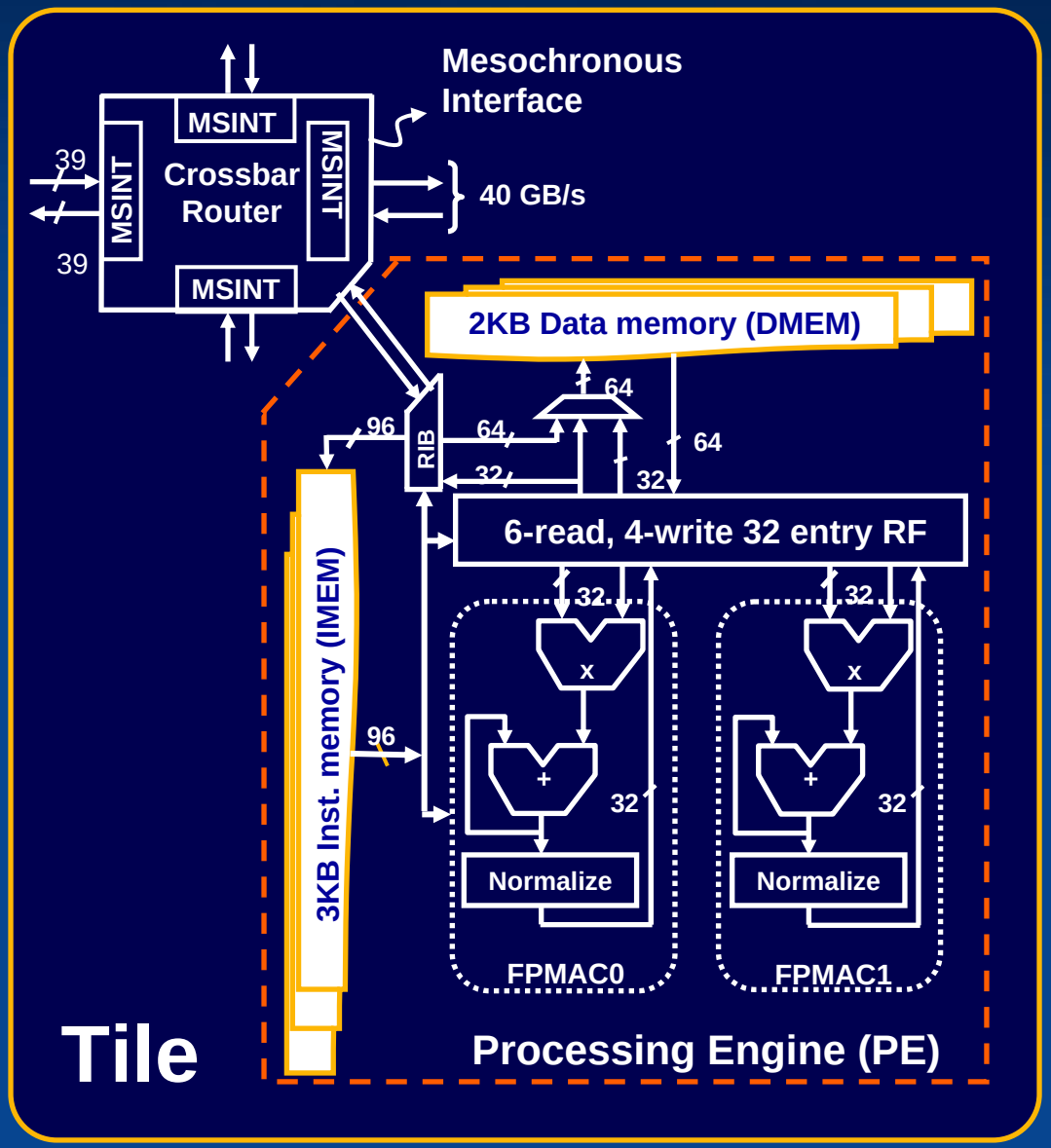


Technology	65nm, 1 poly, 8 metal (Cu)
Transistors	100 Million (full-chip) 1.2 Million (tile)
Die Area	275mm^2 (full-chip) 3mm^2 (tile)
C4 bumps #	8390

Goals:

- Deliver Tera-scale performance
  - Single precision TFLOP at desktop power
  - Frequency target 5GHz
  - Bi-section B/W order of Terabits/s
  - Link bandwidth in hundreds of GB/s
- Prototype two key technologies
  - On-die interconnect fabric
  - 3D stacked memory
- Develop a scalable design methodology
  - Tiled design approach
  - Mesochronous clocking
  - Power-aware capability

Key Ingredients



- Special Purpose Cores
  - High performance Dual FPMACs
- 2D Mesh Interconnect
  - High bandwidth low latency router
  - Phase-tolerant tile to tile communication
- Mesochronous Clocking
  - Modular & scalable
  - Lower power
- Workload-aware Power Management
  - Sleep instructions and Packets
  - Chip voltage & freq. control

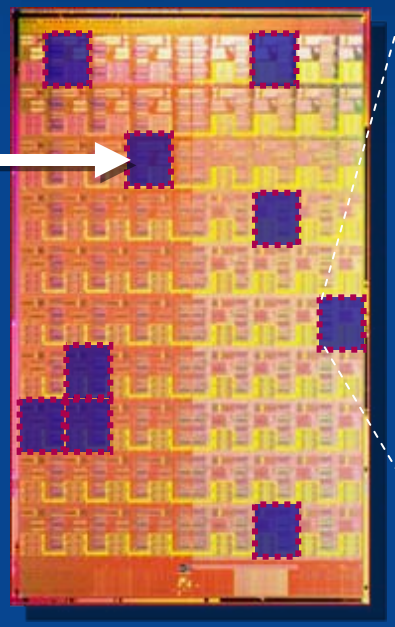
Fine Grain Power Management

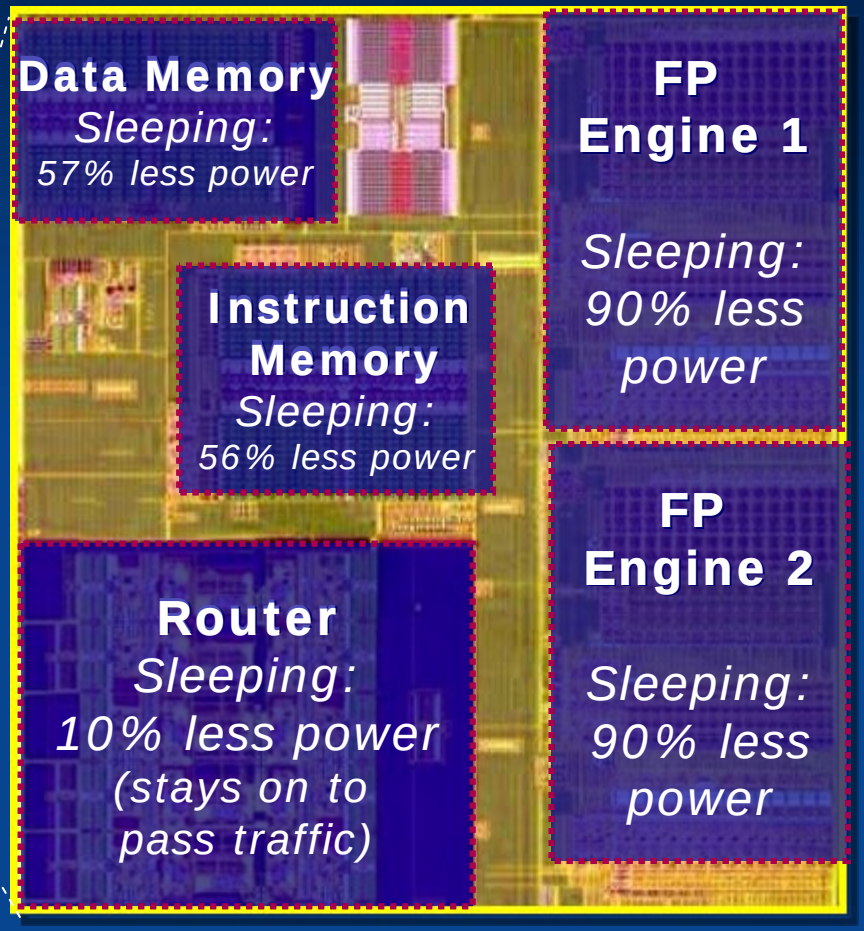
21 sleep regions per tile *(not all shown)*

Dynamic sleep

STANDBY:
- Memory retains data
- **50%** less power/tile

FULL SLEEP:
- Memories fully off
- **80%** less power/tile

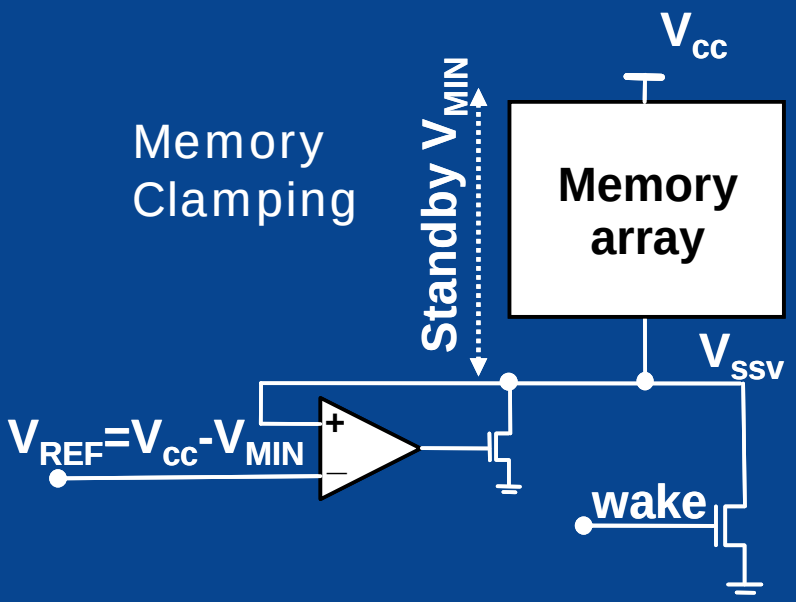


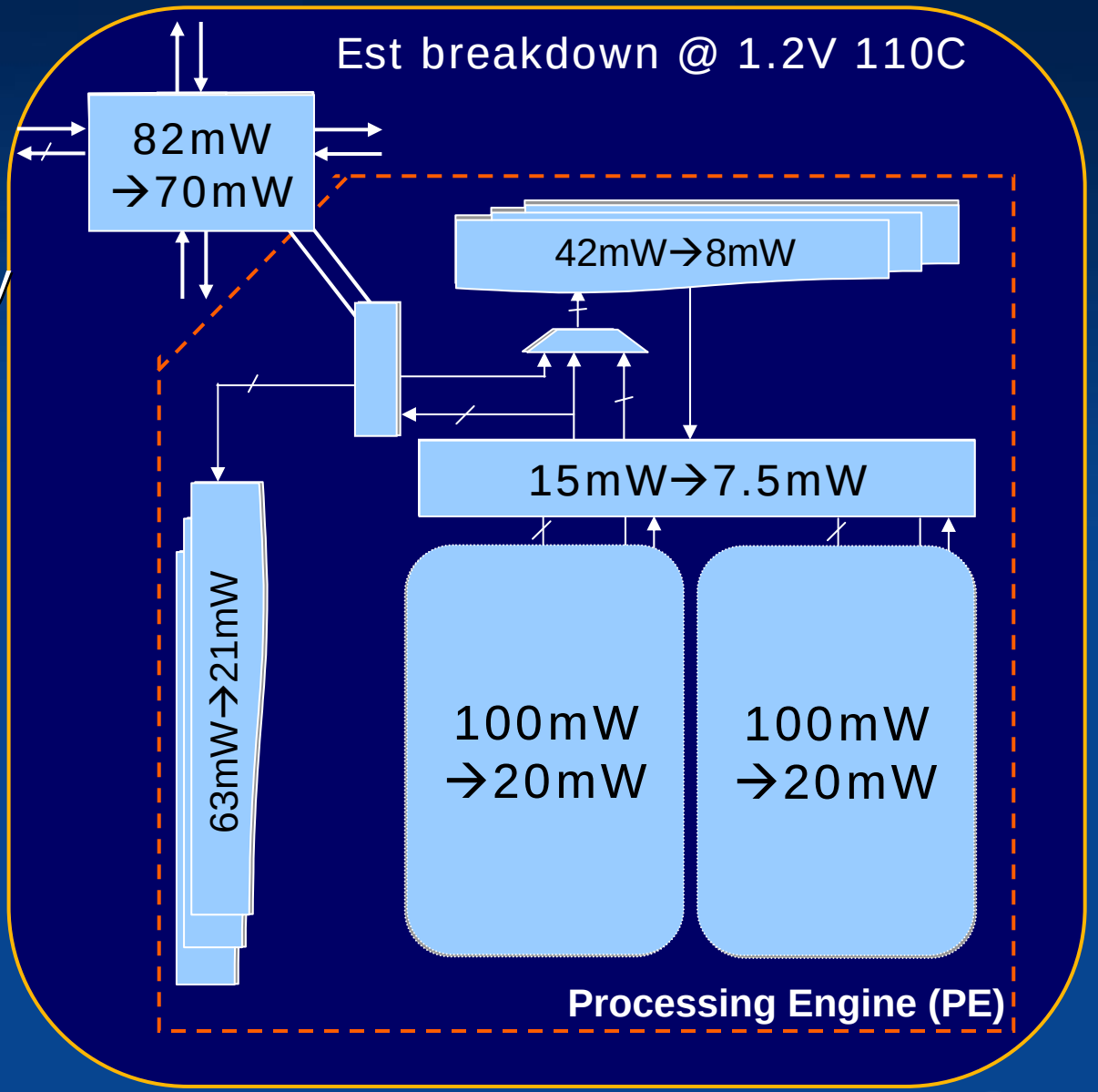


Scalable power to match workload demands

Leakage Savings

- Dynamic sleep
  - Total measured idle power 13W → ~7W sleeping
- Regulated sleep for memory arrays
  - State retention

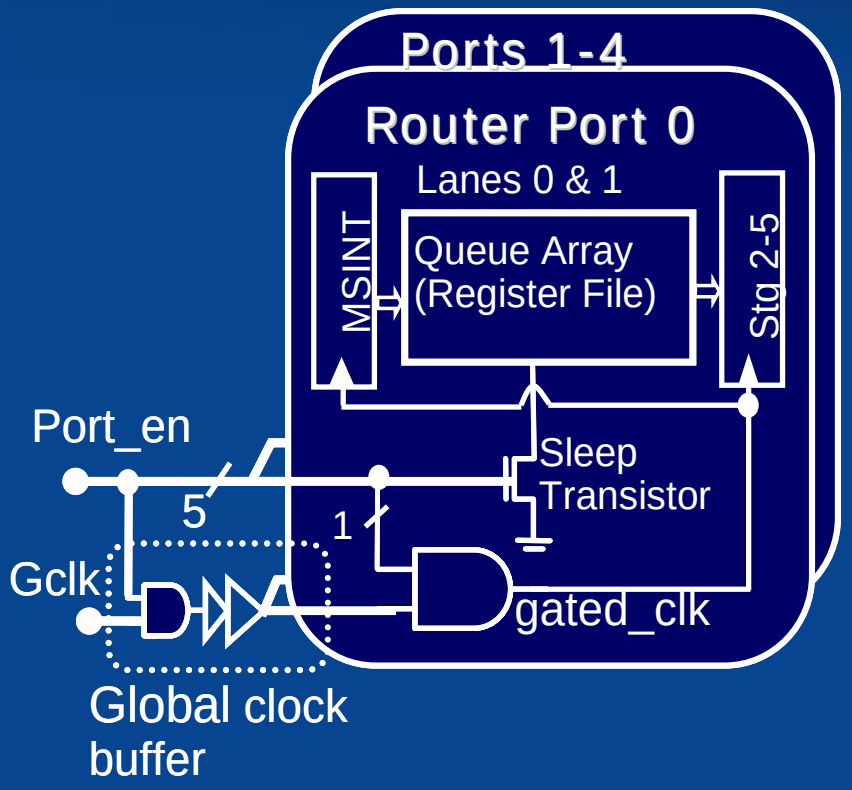


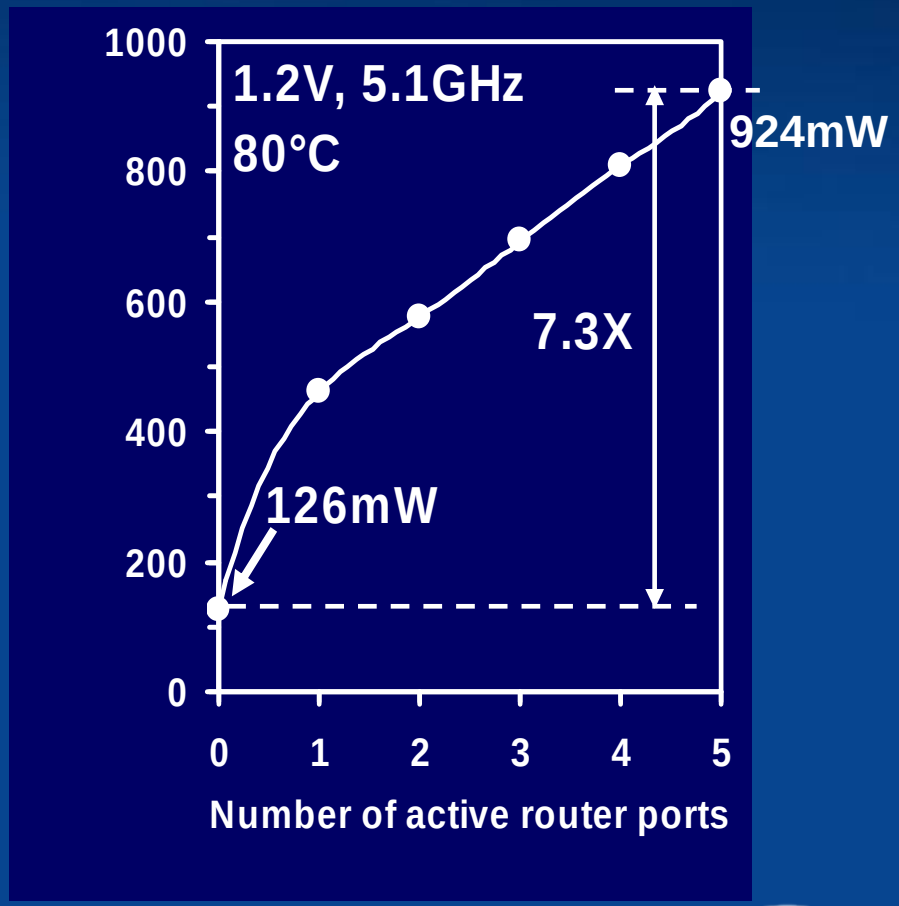


2X-5X leakage power reduction

Router Power Management

- Activity based power management
- Individual port enables
  - Queues on sleep and clock gated when port idle

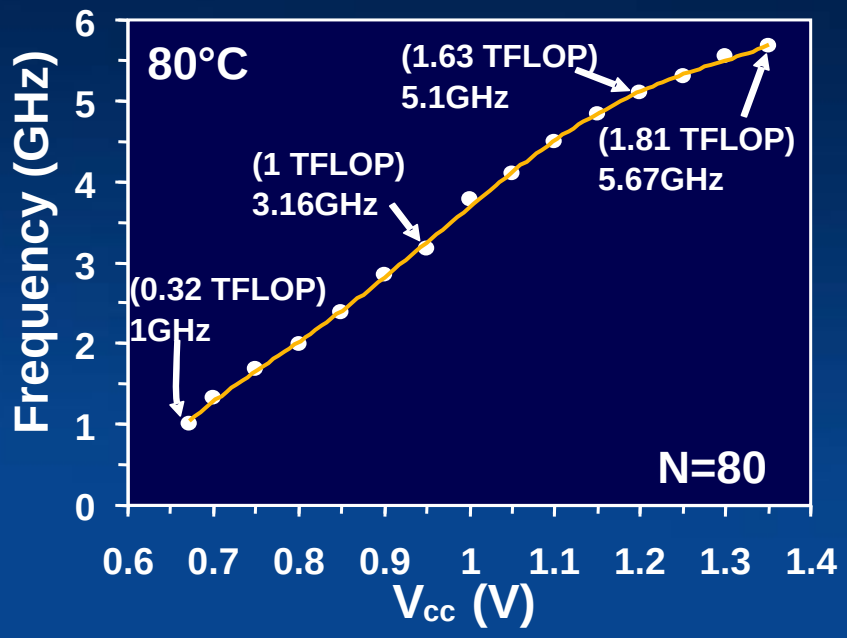




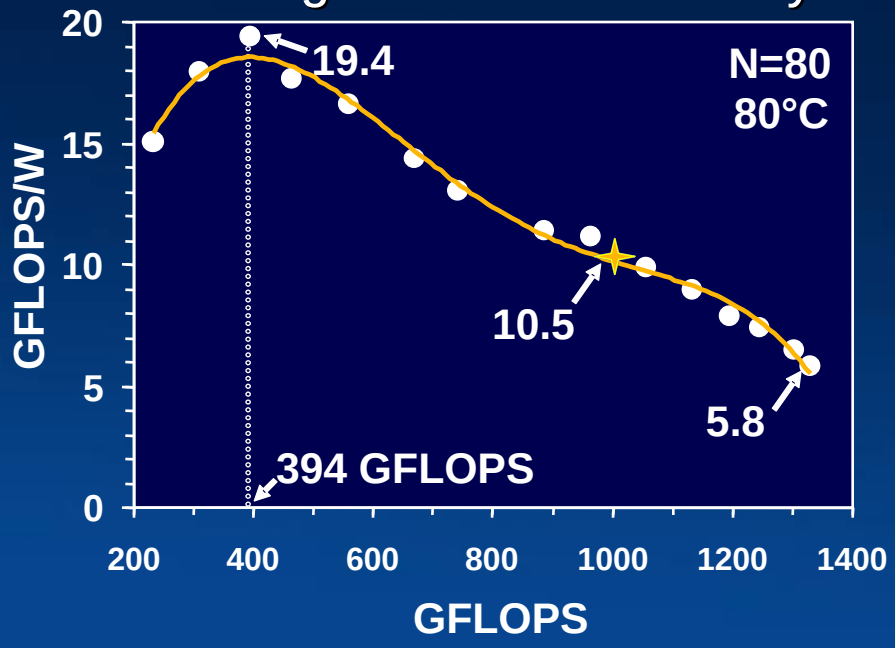
7X power reduction for idle routers

Power Performance Results

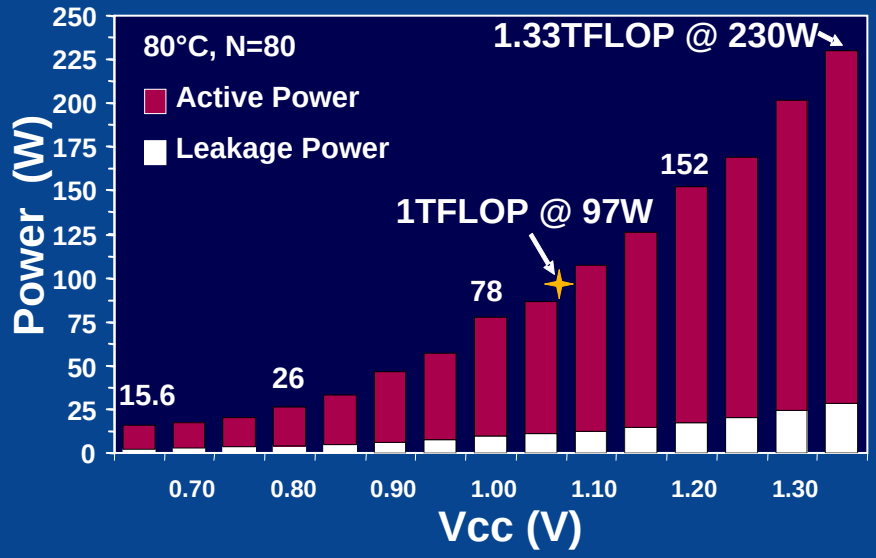
Peak Performance



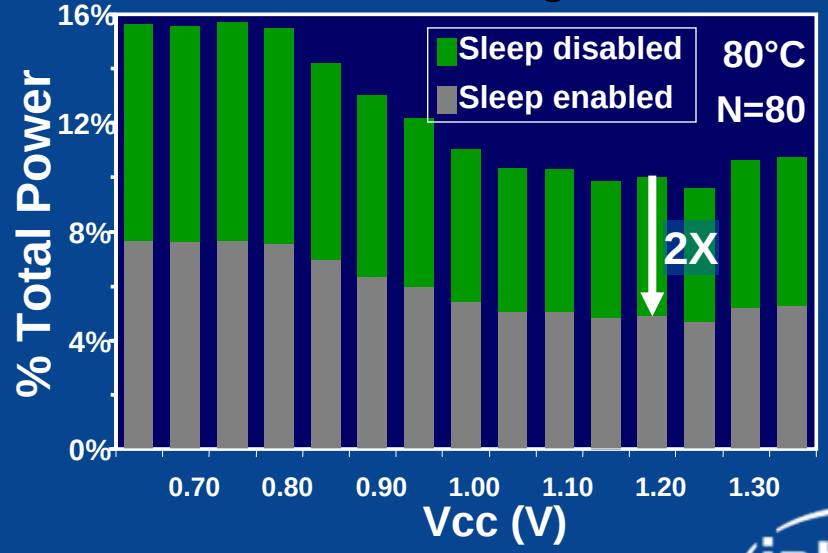
Average Power Efficiency



Measured Power



Leakage

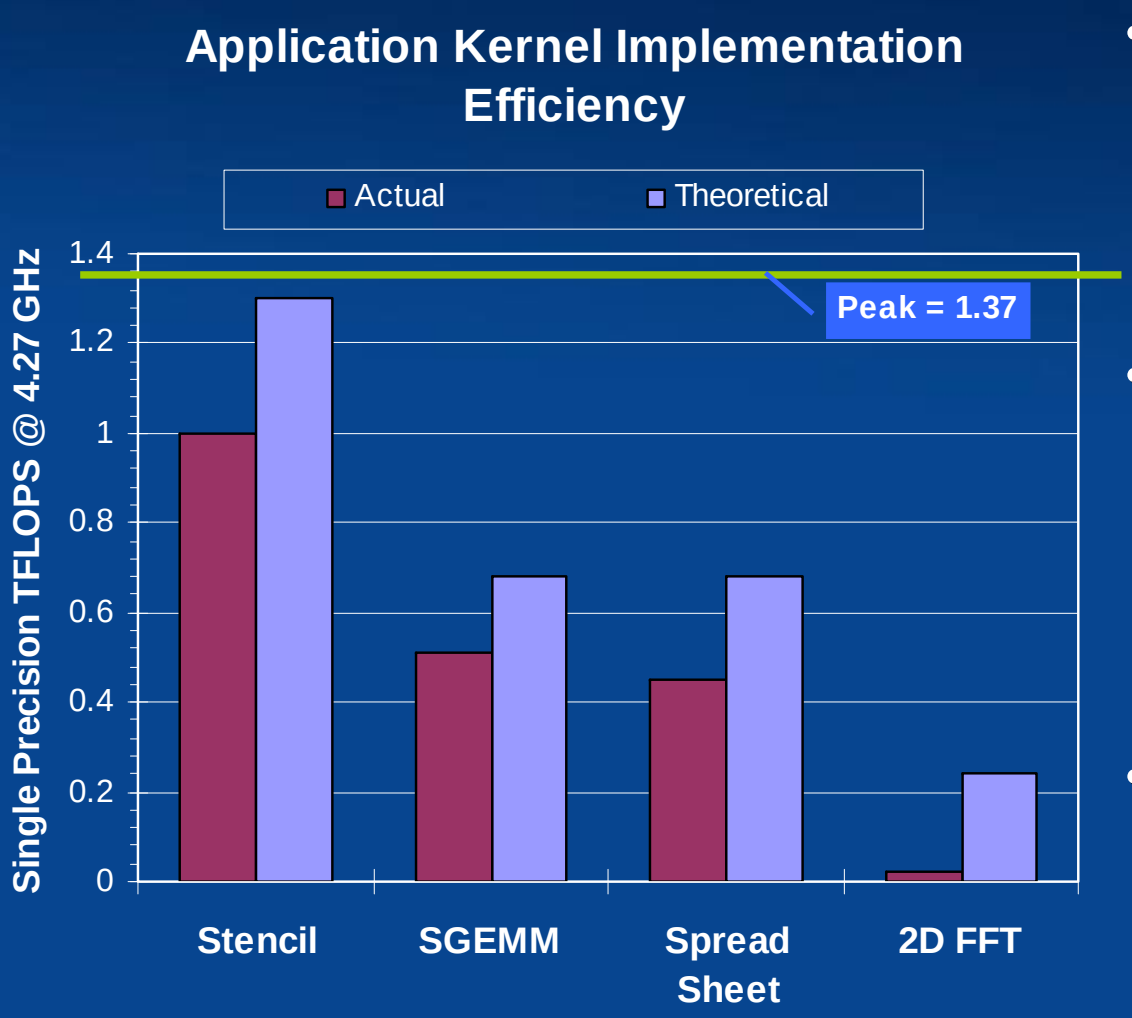


Stencil: 1TFLOP @ 97W, 1.07V;

All tiles awake/asleep

intel

Programming Results



- Not designed as a general Software Development Vehicle
  - Small memory
  - ISA limitations
  - Limited data ports
- Four kernels hand-coded to explore delivered performance:
  - Stencil 2D heat diffusion equation
  - SGEMM for 100x100 matrices
  - Spreadsheet doing weighted sums
  - 64 point 2D FFT (w 64 tiles)
- Demonstrated utility and high scalability of message passing programming models on many core

Key Learnings

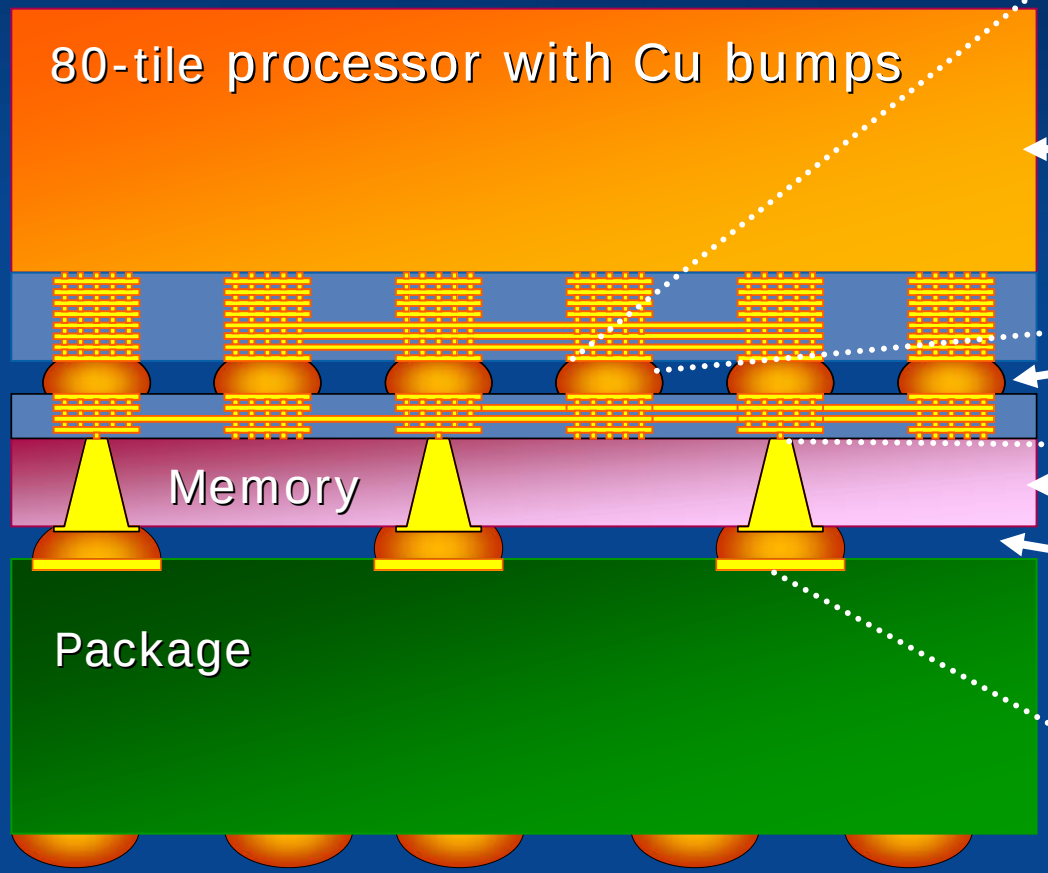
- Teraflop performance is possible within a mainstream power envelope
  - Peak of 1.01 Teraflops at 62 watts
  - Measured peak power efficiency of 19.4 GFLOPS/Watt
- Tile-based methodology fulfilled its promise
  - Design possible with ½ the team in ½ the time
  - Pre & Post-Si debug reduced – fully functional on A0
- Fine-grained power management pays off
  - Hierarchical clock gating and sleep transistor techniques
  - Up to 3X measured reduction in standby leakage power
  - Scalable low-power mesochronous clocking
- Excellent SW performance possible in this message-based architecture
  - Further improvements possible with additional instructions, larger memory, wider data ports

Work in Progress: Stacked Memory Prototype

256 KB SRAM per core

4X C4 bump density

3200 thru-silicon vias

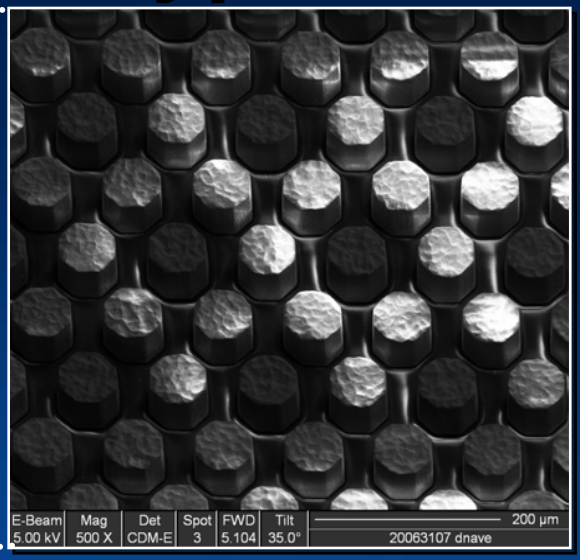


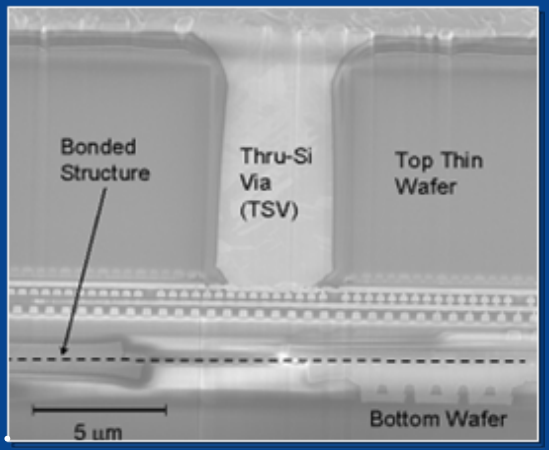
"Polaris"

Denser than C4 pitch

"Freya"

C4 pitch

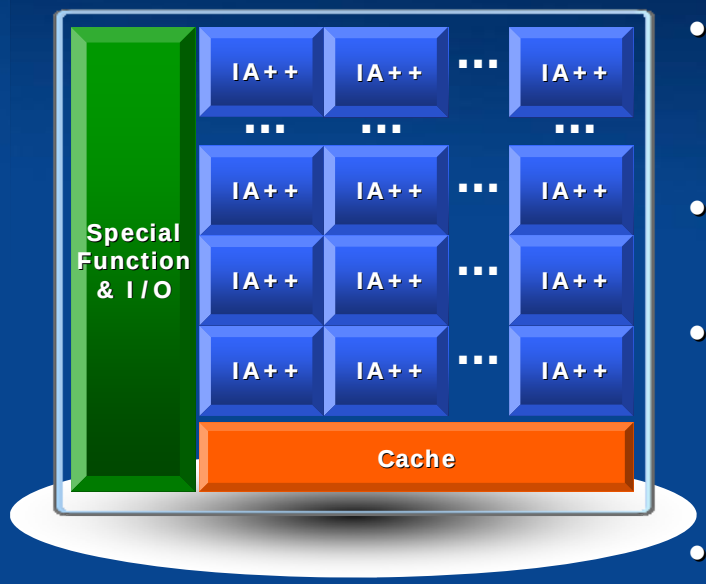




Memory access to match the compute power

Teraflops on IA

Bringing IA Programmability and Parallelism to High Throughput Computing



- Highly parallel, IA programmable architecture in development
- Ease of scaling for software ecosystem
- Array of enhanced IA cores
  - Scientific Computing, RMS, Visualization, Financial Analytics & Health applications
- Teraflops of performance

Pat Gelsinger – Intel Developer Forum 2007

Summary

- Emerging applications will demand teraflop performance
- Teraflop performance is possible within a mainstream power envelope
- Intel is developing technologies to enable Tera-scale computing

Questions

Acknowledgments

Sriram Vangal, Jason Howard, Gregory Ruhl, Saurabh Dighe, Howard Wilson, James Tschanz, David Finan, Priya Iyer, Arvind Singh, Tiju Jacob, Shailendra Jain, Sriram Venkataraman, Yatin Hoskote, Nitin Borkar, Rob van der Wijngaart, Michael Frumkin and Tim Mattson