A report from the Intel Developer Forum 2009 (San Francisco)

Andrzej Nowak

October 20th 2009

Overview

- **Keynote contents and comments**
- **Hardware part**
  - 32nm and 45nm process status
  - Data center efficiency
  - Future SSD innovations
- **Software part**
  - Scaling application performance (multi-core)
  - Tools and methods for scaling (multi-core)
  - Optimizing for Atom

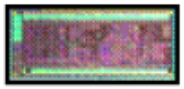
IDF 2009 keynotes

- **Keynotes brought a very small amount of exciting news**
- **Paul Otellini + Sean Maloney:**
  - New Atom cores on the way from TSMC ("not for capacity but for outreach")
  - +22% performance, 10x leakage per process step
  - Sandy Bridge silicon taped out 2 months ago
  - Silicon evolution between 2006 and 2009:
    - 90% reduction in power
    - 85% reduction in size
    - 65% reduction in cost
  - Real time systems running on IA thanks to WindRiver acquisition
  - Server growth projections 2008-2012
    - Users: 3x increase, data: 4.5x increase, devices: 7.7x increase
  - 22nm SRAM cells and test logic developed
- **Justin Rattner's (Intel Labs) keynote limited to the future of television**

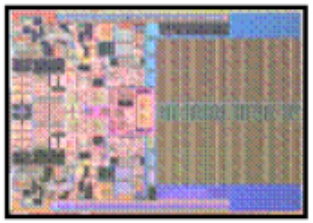
Part 1: HARDWARE

32nm process technology (Mark Bohr) (1)
Current 45 nm offerings

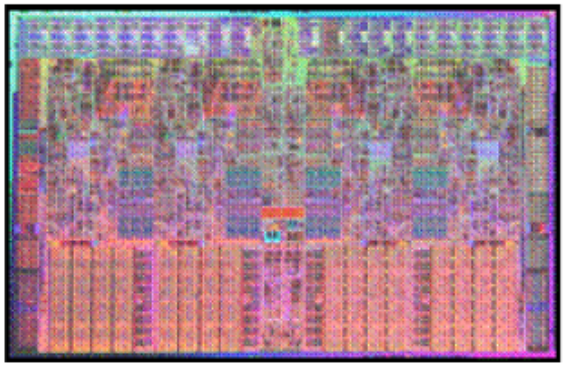
Single Core

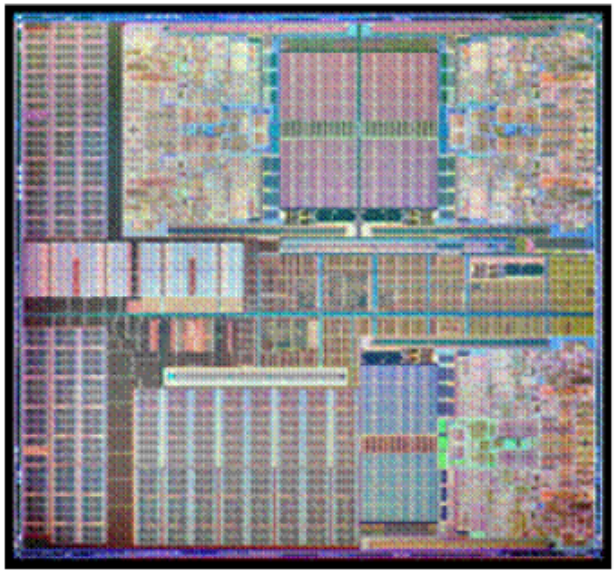


Dual Core

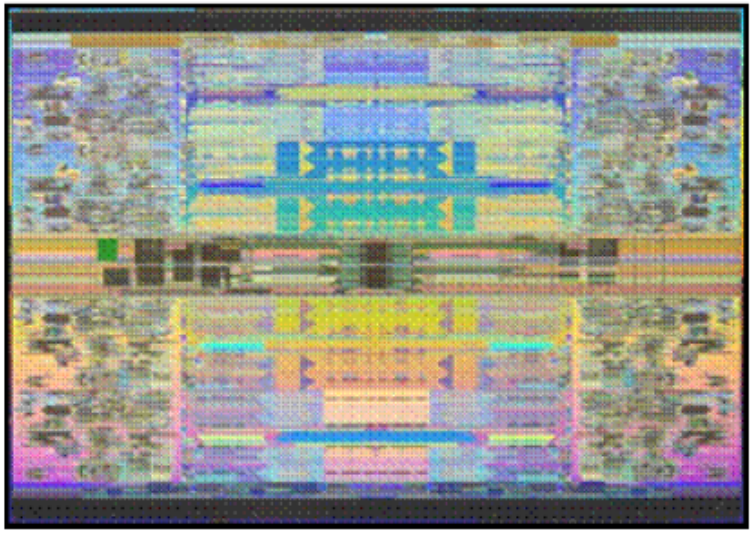


Quad Core





6 Core



8 Core

Image: IDF 2009 materials

32nm process technology (Mark Bohr) (2)
Defect density trend on 32nm

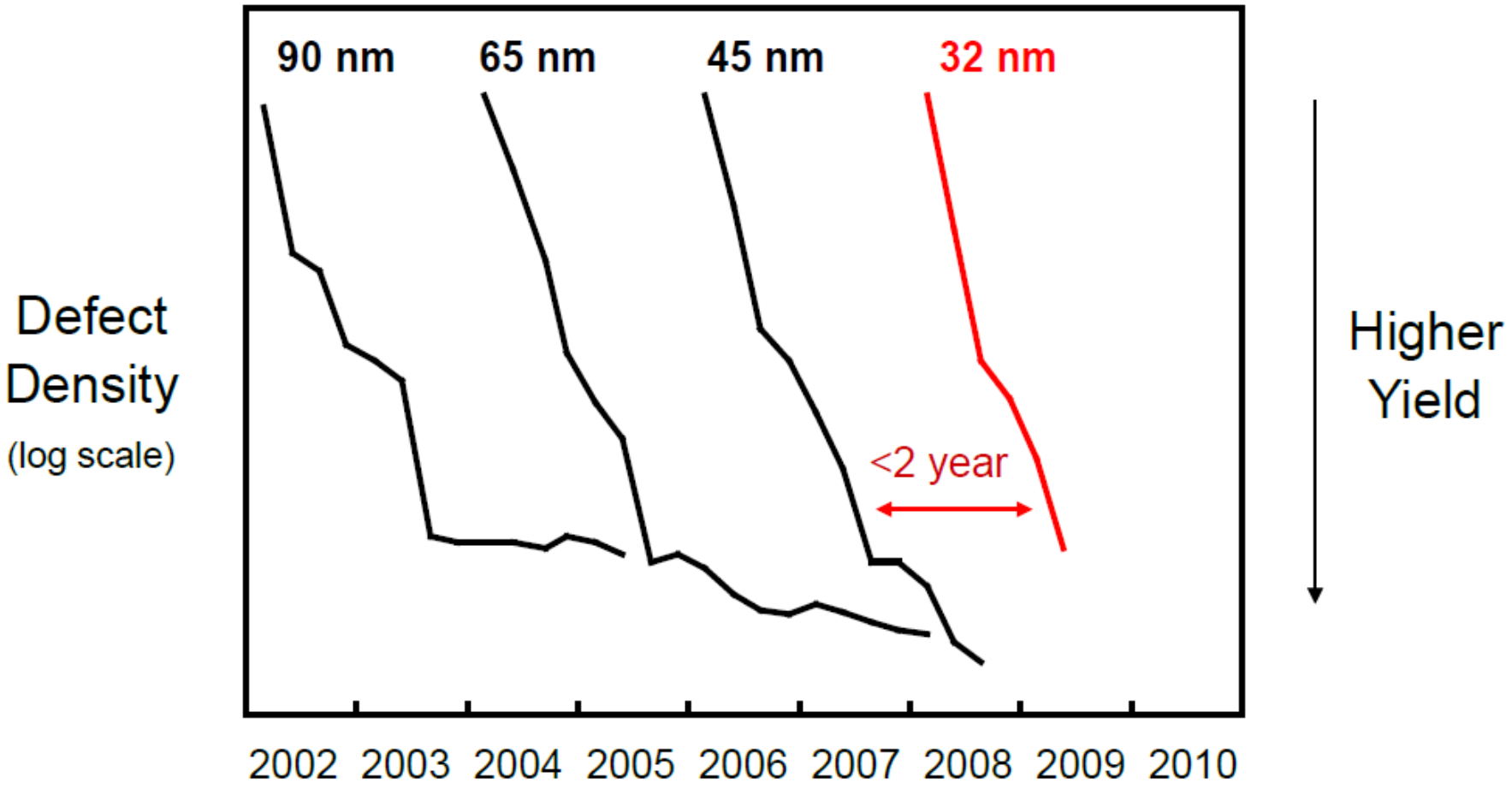


Image: IDF 2009 materials

32nm process technology (Mark Bohr) (3)

- **The ~2 year cycle continues**
- **32nm brings:**
  - 2nd generation high-K + metal transistors
  - 0.7x minimum pitch scaling
  - Pb-free and halogen-free packaging
- **32nm "revenue production" this quarter**
  - 4 fabs working on 32nm
- **Interesting tidbit: 32nm silicon from Intel offers a mix-and-match model with:**
  - 3 types of logic transistors
  - 3 types of i/o transistors
  - 2 metal layer choices
  - 3 types of advanced passives
  - 3 types of embedded memory

32nm process technology (Mark Bohr) (4)
22nm test chip – first shot

SRAM, Logic, Mixed-Signal
Test Circuits

SRAM, Logic, Mixed-Signal
Test Circuits

Discrete
Test Structures

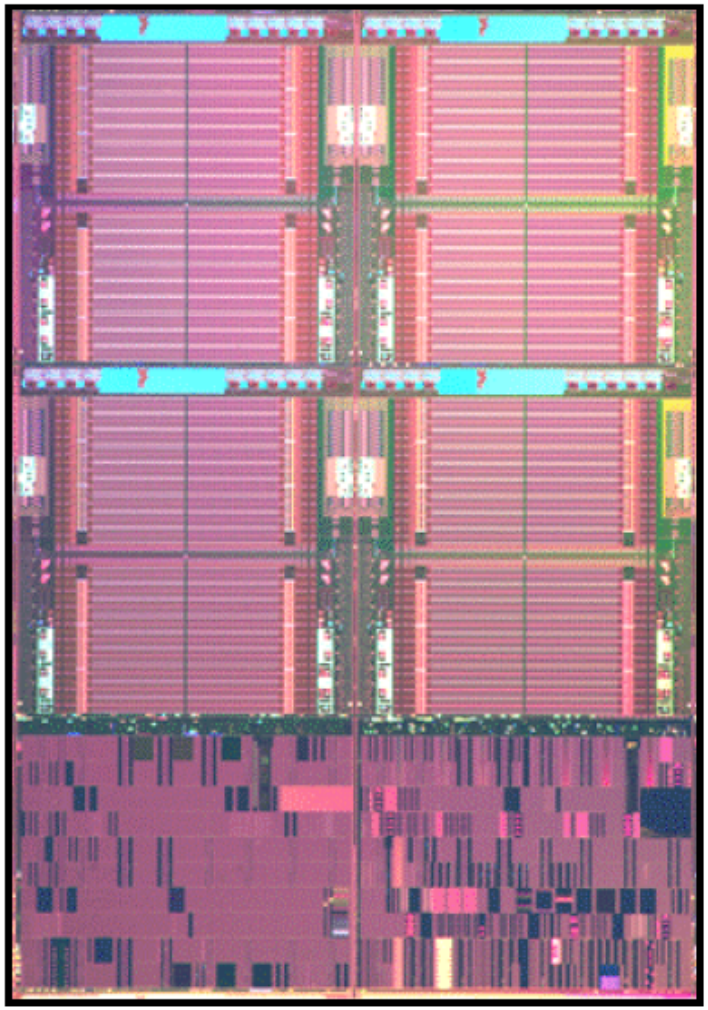
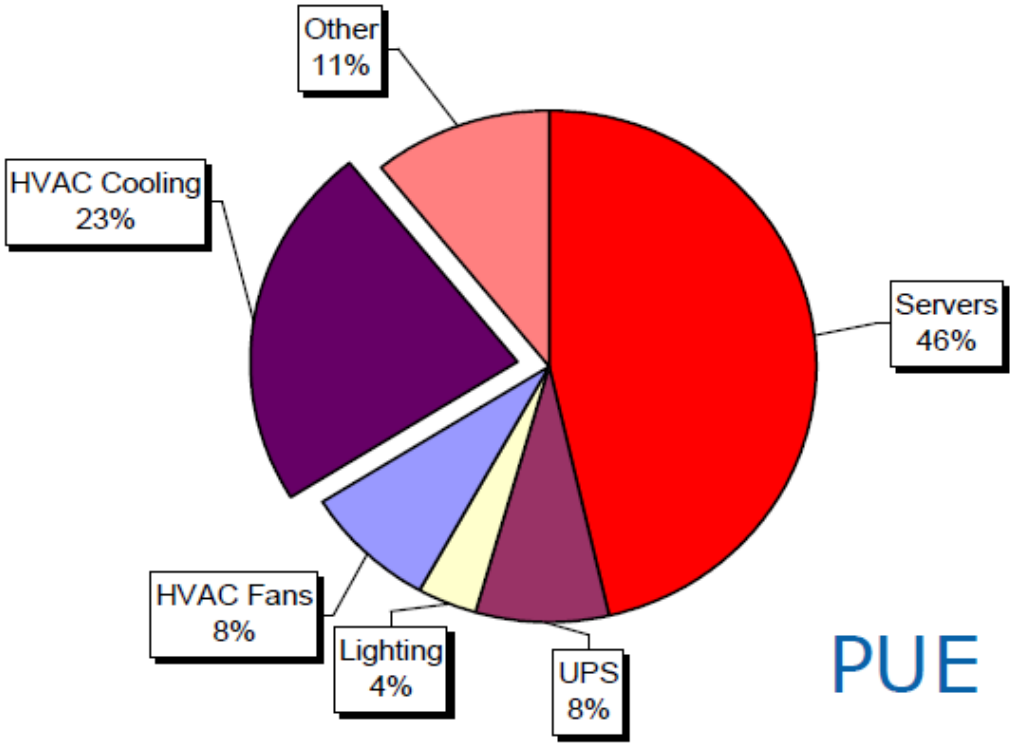


Image: IDF 2009 materials

Data center efficiency (Claude Fiori, Allyson Klein, Markus Leberecht) (1)

Average Data Center Power Allocation



PUE = 2.17

7

Source: http://tc99.ashraetcs.org/

IDF2009
INTEL DEVELOPER FORUM

Data center efficiency (Claude Fiori, Allyson Klein, Markus Leberecht) (2)

- **T-Systems infrastructure test lab:**
  - Analysis of different cooling concepts
  - Evaluation of different powering systems
  - Developing and testing ideas for infrastructure improvements
  - Stretch current kW/m^2 limits
  - Evaluate new computing hardware
- **Interesting features:**
  - 180 servers, 8 racks
  - Cold water supply, liquid cooling
  - Humidity control
  - Adjustable ceiling and walls
  - Coated walls
  - Smoke generators

Data center efficiency (Claude Fiori, Allyson Klein, Markus Leberecht) (3)

> Degrees of freedom:

- Enclosure: none / hot aisles / cold aisles / both
- Power supply: AC/DC
- Room height
- Floor leakage
- Rack leakage
- Room humidity
- Room temperature
- Cpu load
- Fan speed, water temperature, water flow

Future SSD innovations (Knut Grimsrud, Chris Saleski)

- **Driver and stack optimization is different from HDD – it's a big efficiency problem alone**
  - For example prefetching: optimized only for HDDs; less data shipped = smaller power usage
- **Controller efficiency also a factor**
  - A ports are not the same as B ports! Up to 6x performance difference
- **Multi- and many-core: are the interrupts delivered to the right cores?**
- **As the drive fills up, performance decreases**
- **Peak vs. sustained performance: the minimum might be the most important figure**
- **There were cases where just by changing the driver you could get 3.5x performance improvement**
- **MSRP of the 50 nm generation is going down by 60%**
  - 34nm is the current generation

Part 2: SOFTWARE

Accelerating Application Performance and Scaling it Forward (Nash Palaniswamy, John Gustafson) (1)

- **Moore's law will be achieved by:**
  - More cores
  - More cache
  - More bandwidth
  - More instruction set enhancements
  - Scalable platform features (storage, I/O)
  - A combined focus on software
- **Nehalem-EX coming "early 2010"**
- **John Gustafson is now working at Intel Labs**
  - System balance is not about bytes per flops or mass storage to DRAM or any such ratio
  - It is also not about configuring so that no single component holds up the computation
  - it means selecting features where the % improvement in value is greater than the % increase of the TCO (which also includes the programming effort)

Accelerating Application Performance and Scaling it Forward (Nash Palaniswamy, John Gustafson) (2)

> Data manipulation – power consumption

64-bit mul-add	200pJ
Read 64 bits per cache	800pJ
Move 64 bits across a chip	2000pJ
Execute an instruction	7500pJ
Read 64 bits from DRAM	12000pJ

> 12'000 pJ @ 3GHz = 36 Watts

- Solution? Lower memory speed

Accelerating Application Performance and Scaling it Forward (Nash Palaniswamy, John Gustafson) (3)

- **Future HPC cluster:**
  - We want to have a teraflop at 20W (Chip/module)
  - Petaflop at 20kW (cabinet)
  - Exaflop at 20MW (data center)
  - No free lunch
  - Future clusters may be about 90% communication, 10% computation (like a brain)
- **Systems now over-provision floating point hardware to the point where only linpack sees benefit. Real scientific workloads have very little floating point math that is not overlapped with data motion**
- **Data motion and not floating point now limits performance on all but very few workloads**
- **Future directions:**
  - Westmere (32nm): More cores
  - Sandy Bridge (32nm): Higher integration

Developer tools for scaling performance forward (Paresh Rattani, Sanjay Goil)

- **Task level parallelism**
  - Task queues - TBB, Cilk++
  - Directives - OpenMP
- **Data level parallelism**
  - Language expression - Ct
  - Vectorizatoin - SSE4, AVX
  - Directives - OpenMP
- **Cluster level parallelism**
  - Message passing - MPI
- **Cilk++, Ct, Intel AVX is future functionality**
- **openlab is already collaborating on Ct and AVX with Intel**

Atom performance optimization (Uli Dumschat)

> Performance optimization principles:

- Use pragmas, not only compiler switches
- Use the PMU, Atom has a decent one
- Use the multimedia libraries which are highly optimized for Atom (IPP is Atom optimized)
- Use ICC 11.1 series:
  - with a 2nd generation in-order scheduler: 10%-40% improvements quoted
  - The FP unit is “more streamlined”
  - Convenient division is used
  - LEA is used a lot
  - Numerous heuristics applied
- Compiling the kernel with ICC yields little benefit (up to 3% quoted)

> Software available: Moblin developer suite

- Intel C++ Compiler, Intel Performance Primitives, JTAG, debuggers, VTune; beta tools on demand (CNDA)

Private meetings – Intel/openlab

- Ronak Singhal
- Michael Haedrich
- Nash Palaniswamy
- Anwar Ghuloum

Q & A

CERN openlab

Backup

> Session reference (http://intel.com/idf):

- DPTS004
- ECOS002
- MEMS002, MEMS004
- MOBS004
- RESS002
- SPCS009, SPCS011
- TCIS001, TCIS002