Towards Energy-Centric Computing & Computer Architecture

Babak Falsafi

Parallel Systems Architecture Lab
EPFL
parsa.epfl.ch

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

PARALLEL SYSTEMS ARCHITECTURE LAB

Energy: Humanity’s Top Problem

Smalley’s top list of humanities’ problem for the next 50 years

1. Energy
2. Water
3. Food
4. Environment

....

Energy: Shaping IT's Future

- 40 years of energy scalability
  - Doubling transistors every two years
  - Quadratic reduction in energy from voltages

- But, while Moore's law continues
  - Voltages have started to level off
  - ITRS projections in 2000 for voltage levels in 2009 were 30% lower (www.itrs.net)

An exponential increase in energy usage every generation!!

Current Evolutionary Approaches

- Eliminating idle energy
- Dynamic voltage scaling (mobile)
- Novel technologies/same old architecture
- Make installations air-tight

None provides the orders of magnitude reduction in energy needed to scale!

Revolutionary Approach: Energy-Centric Computing

Minimize joules/work across computing

How? Careful vertical integration

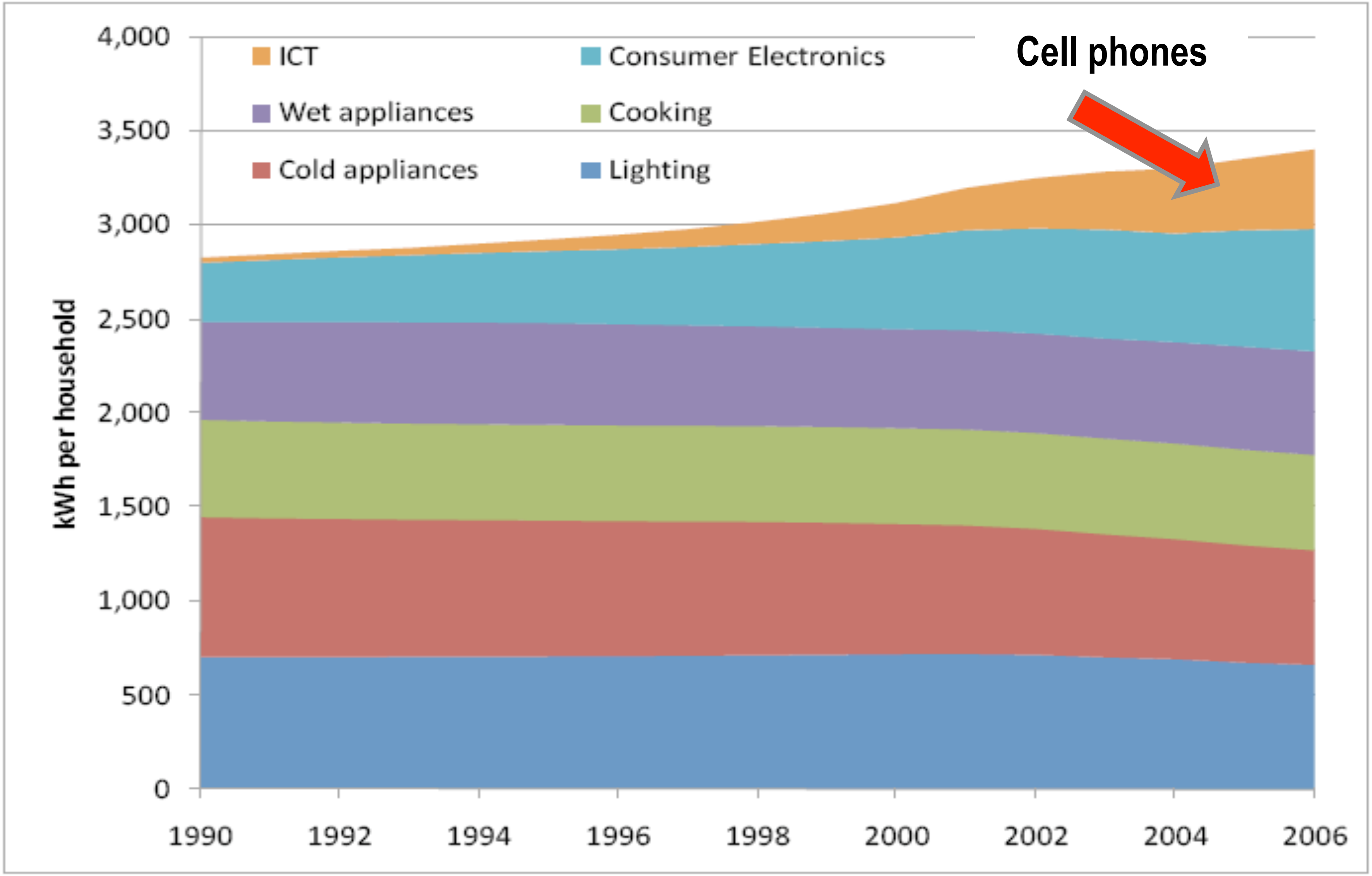
- Minimize energy at all levels
- Monitor and account energy
- Synergistic with infrastructure
  - Computing + delivery + cooling

Energy-centric computing provides orders of magnitude reduction needed to scale

Outline

- Overview
- → Where are we?
- Energy scalability for servers
- Where do we go from here?
- Activities at EPFL
- Summary

Household IT Energy Usage (from Sun)



Source: BERR (2008) *Energy consumption in the UK*

Enterprise IT Energy Usage

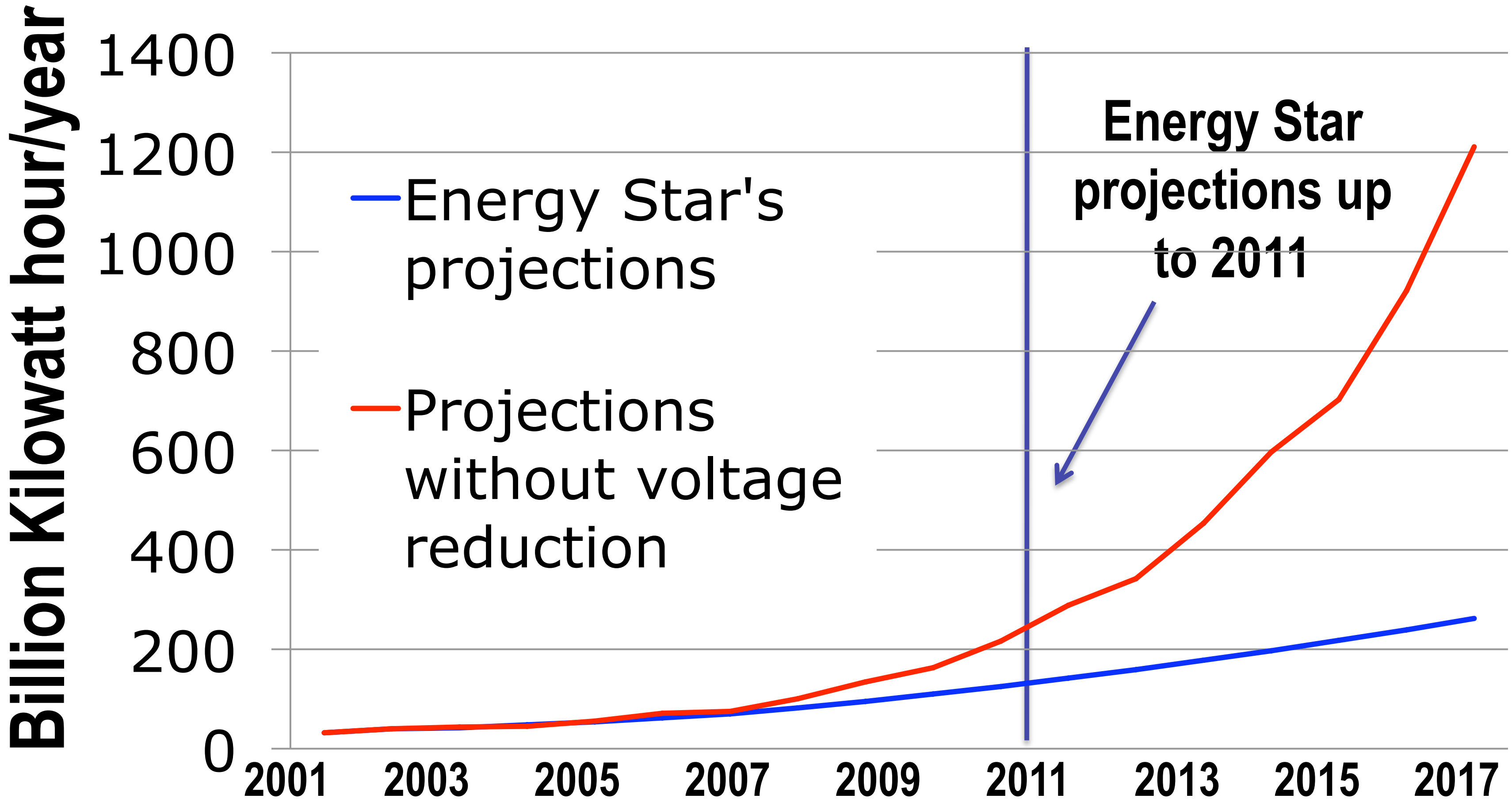
Kenneth Brill (Uptime Institute)

- “Economic Meltdown of Moore’s Law”
- In 2012: Energy/server lifetime 50% more than price/server
  - And 2% of all Carbon footprint in the US

Energy Star report to Congress:

- Datacenter energy 2x from 2000 to 2006
- Roughly 2% of all electricity & growing

Example Projections for Datacenters



- Projections for 2011 are already off
- Exponential increase in usage

Where does server energy go?

Many sources of power consumption:

- Server only [Fan, ISCA'07]
  - Processors chips (37%)
  - Memory (17%)
  - Peripherals (29%)
  - …
- Infrastructure (another 50%)
  - Cooling
  - Power distribution

How did we get here? CMOS & Leakage

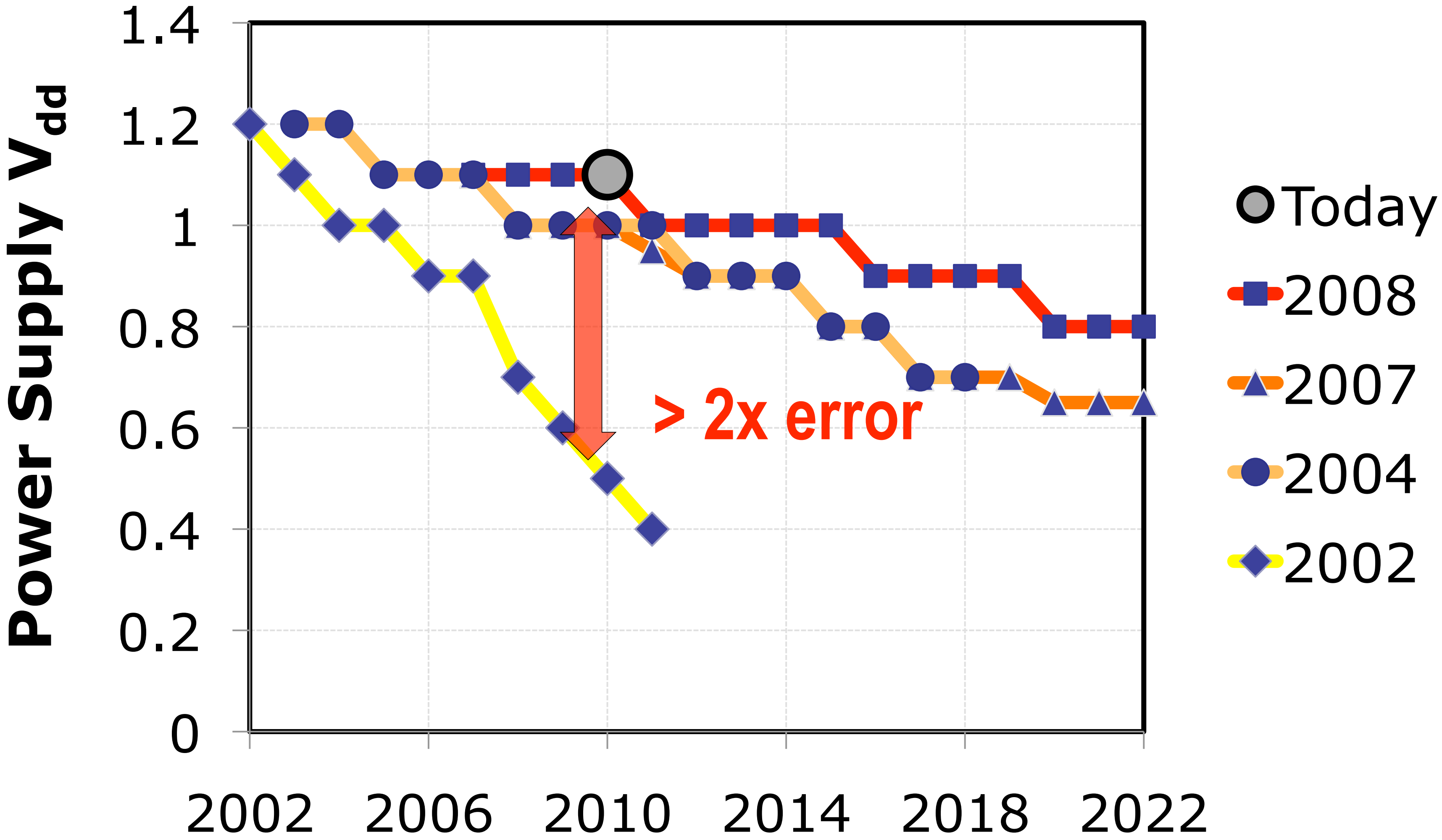
Energy scaling via voltages has slowed

Lower supply voltages?

- Need lower threshold to switch fast
- Lower threshold ➔ exponential “leakage”
- Leakage also exponential in temperature
  - Leakage has runaway effect

But, voltages are leveling off!

Voltages are leveling off



ITRS estimates for today were off by > 2x

What lies ahead?

For the next ten years:

- CMOS is still the cheapest technology

But,

- need ~100x reduction in energy just to keep up with Moore's Law

Current big research agenda?

Parallelism via "popular" programming

But, can we get there with parallelism (alone)?

Outline

- Overview
- Where are we?

→ Energy scalability for servers

- Where do we go from here?
- Activities at EPFL
- Summary

A Study of Server Chip Scalability

Actual server workloads today

Actual physical char. of processors/memory

ITRS projections for technology nodes

Modeled power/performance across nodes

For server chips

- Bandwidth is near-term limiter

→**Energy is the ultimate limiter**

A few words about our model

Physical char. modeled after Niagara

Area: cores/caches (72% die)

- scaled across tech. nodes

Power:

- Active: projected V_{dd}/ITRS
  - Core=scaled, cache=f(miss), crossbar=f(hops)
- Leakage: projected V_{th}/ITRS, f(area), 62C

Performance:

- Parameters from real server workloads (DB2, Oracle, Apache, Zeus)
- Cache miss rate model (validated)
- CPI model based on miss rate

Caveat: Simple Parallelizable Workloads

Workloads are assumed parallel

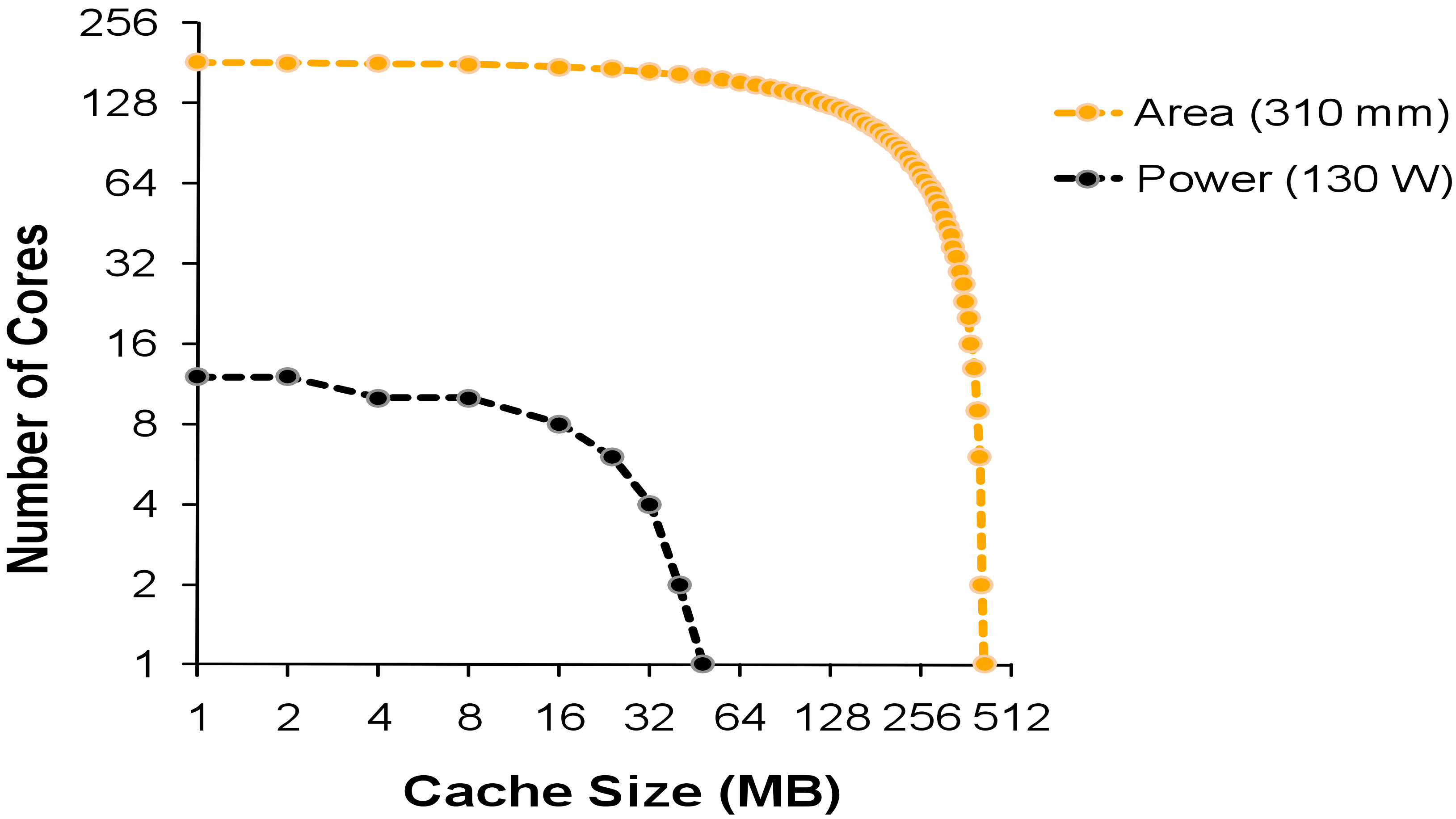
- Scaling server workloads is reasonable

CPI model:

- Works well for workloads with low MLP
- OLTP, Web & DSS are mostly memory-latency dependent

Future servers will run a mix of workloads

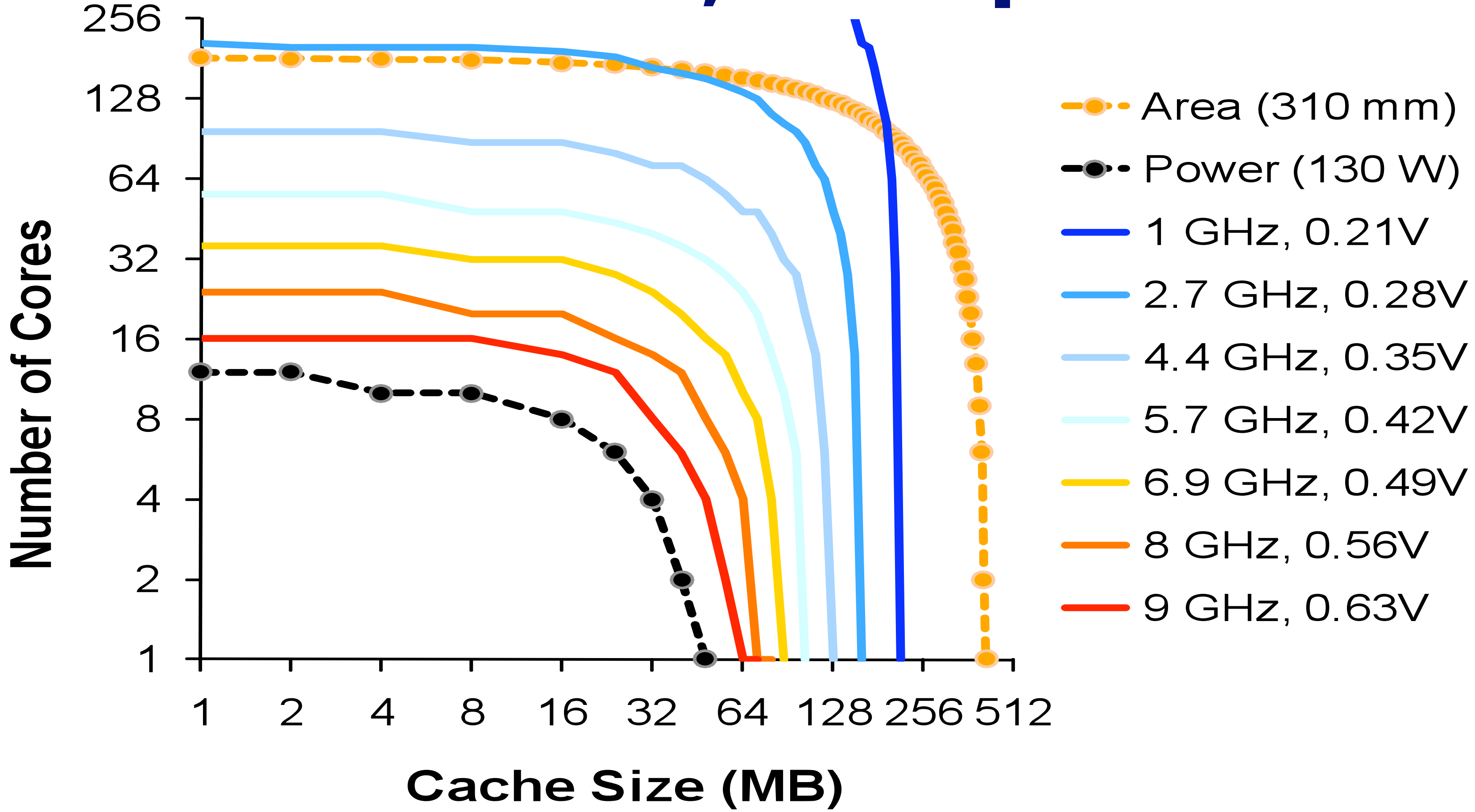
Area vs. Power Envelope (22nm)



✓ Good news: can fit hundreds of cores

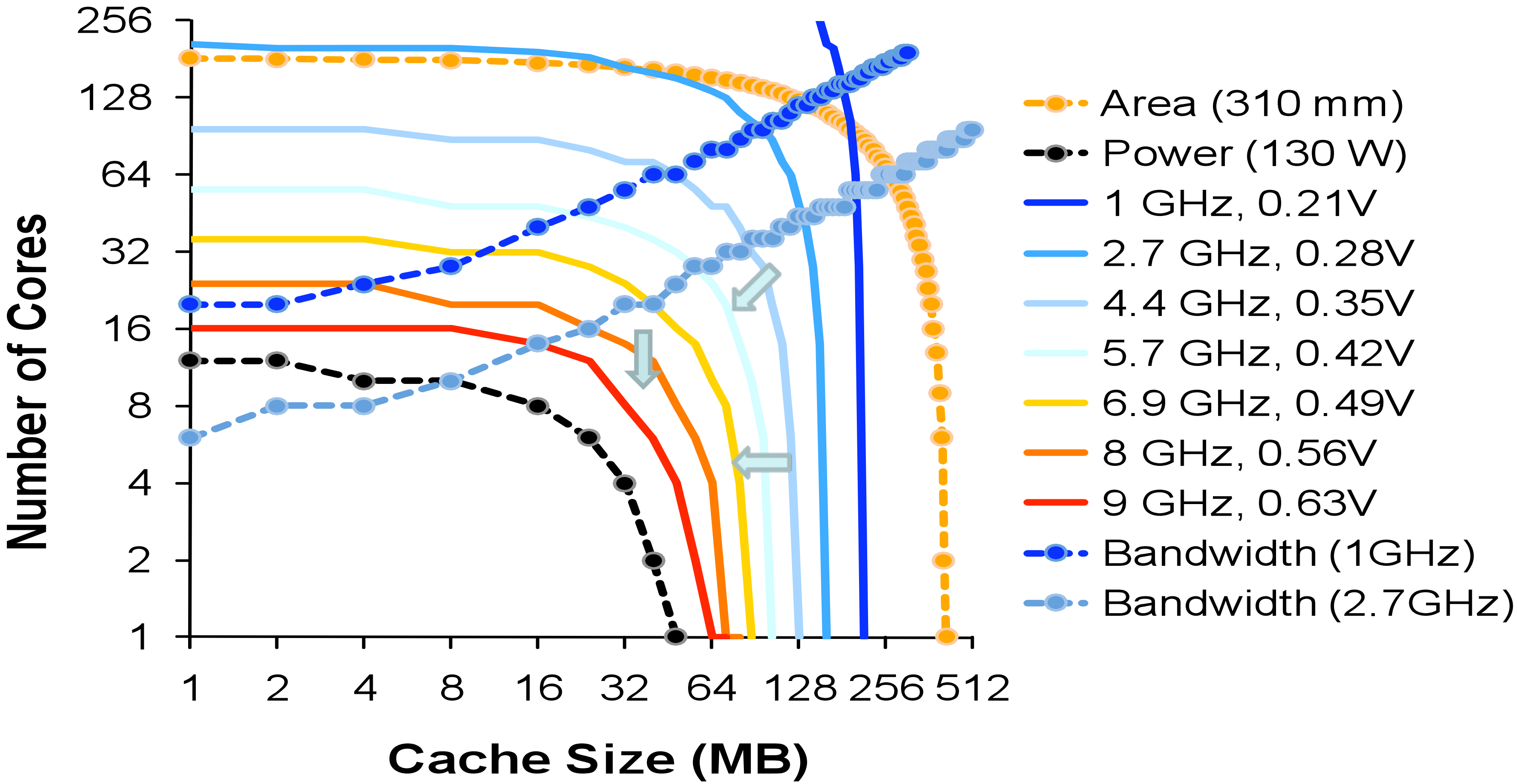
× Can not use them all at highest speed

Of course one could pack more slower cores, cheaper cache



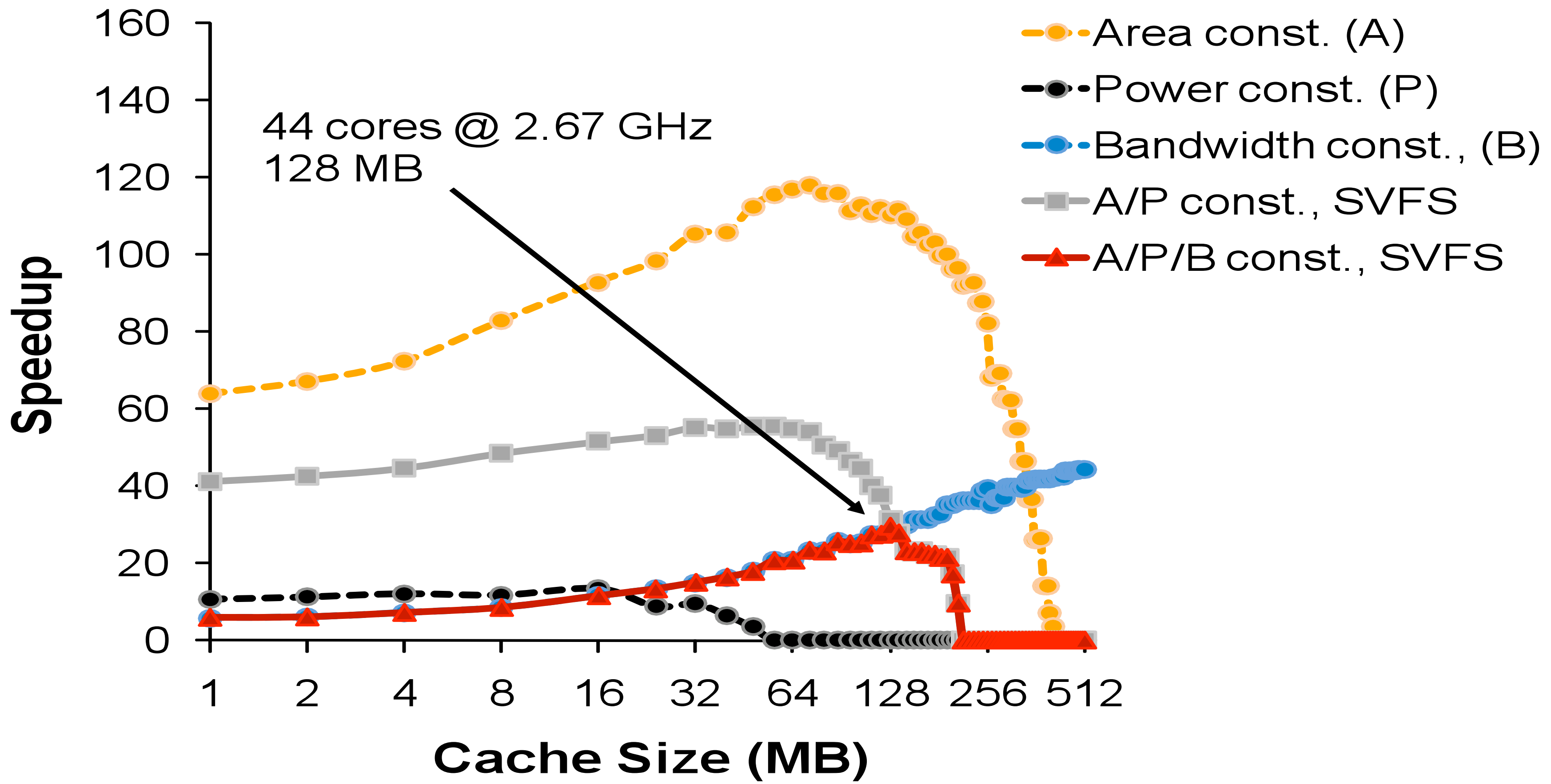
- Result: a performance/power trade-off
- Assuming bandwidth is unlimited

But, limited pin b/w favors fewer cores + more cache



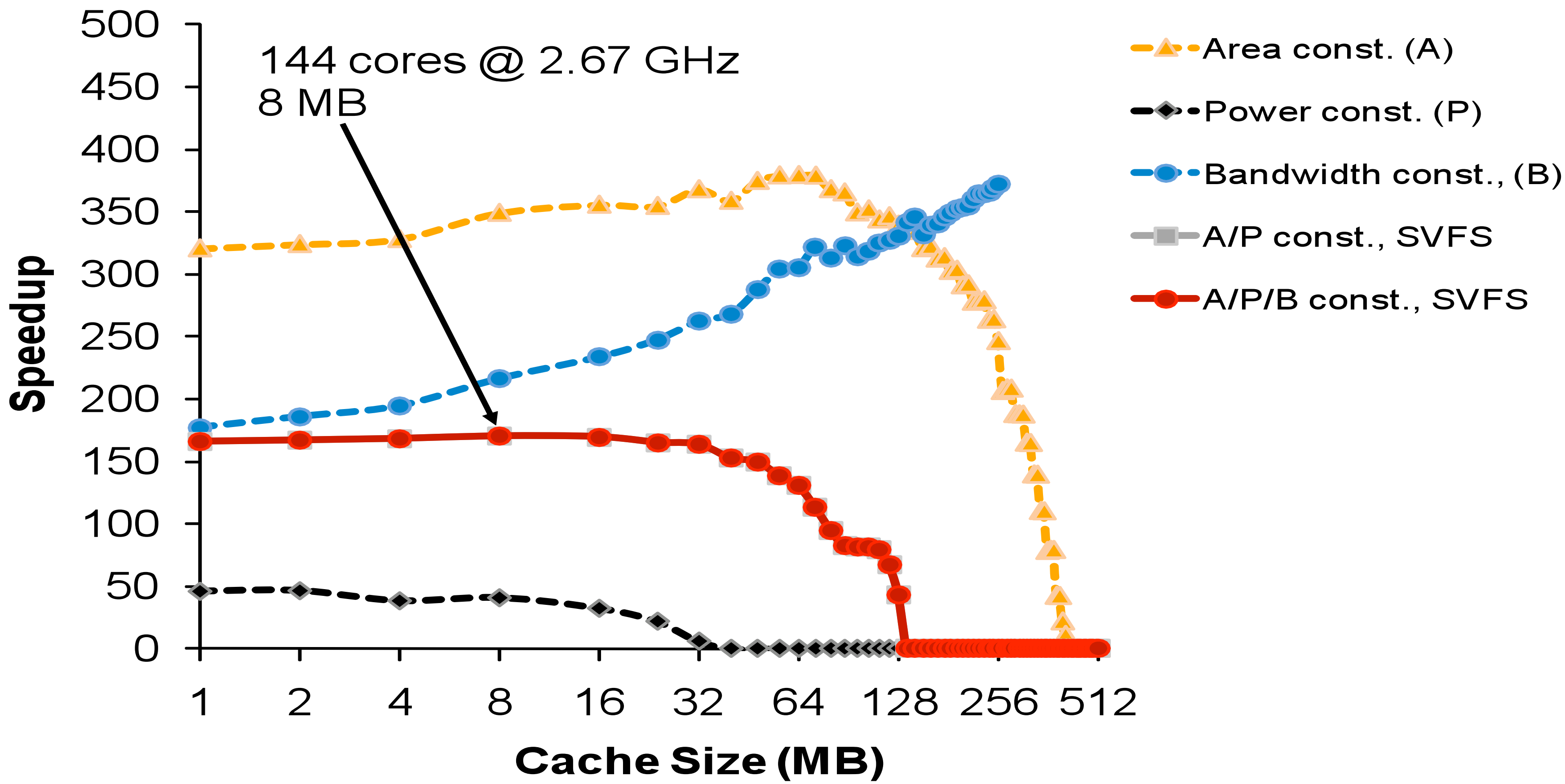
- For clarity, only showing two bandwidth lines
- Where would the best performance be?

Peak Performing with Conventional Memory



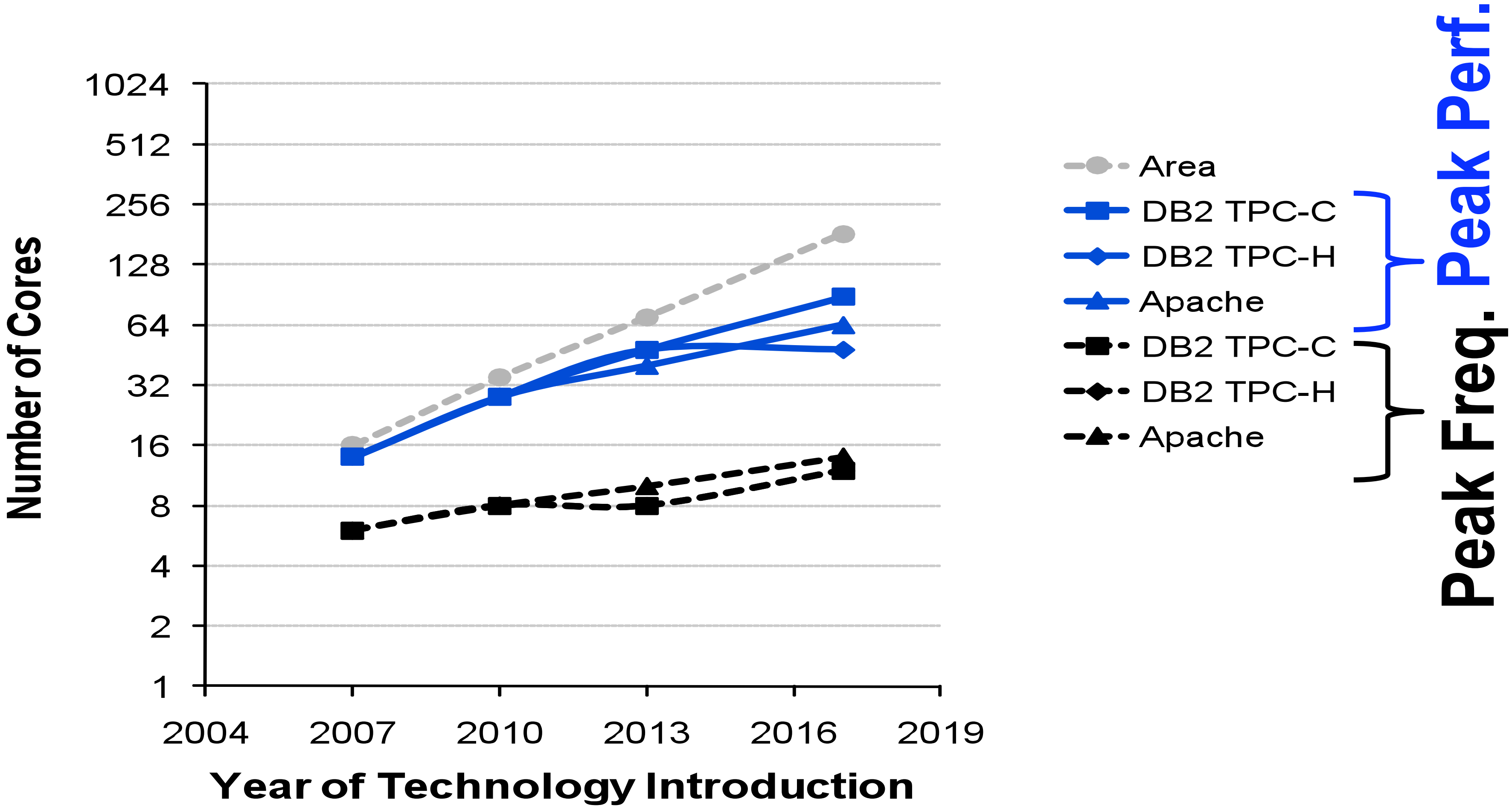
- B/W constrained, then power constrained
- Fewer slower cores, lots of cache

Peak Performing w/ 3D-stacked Memory



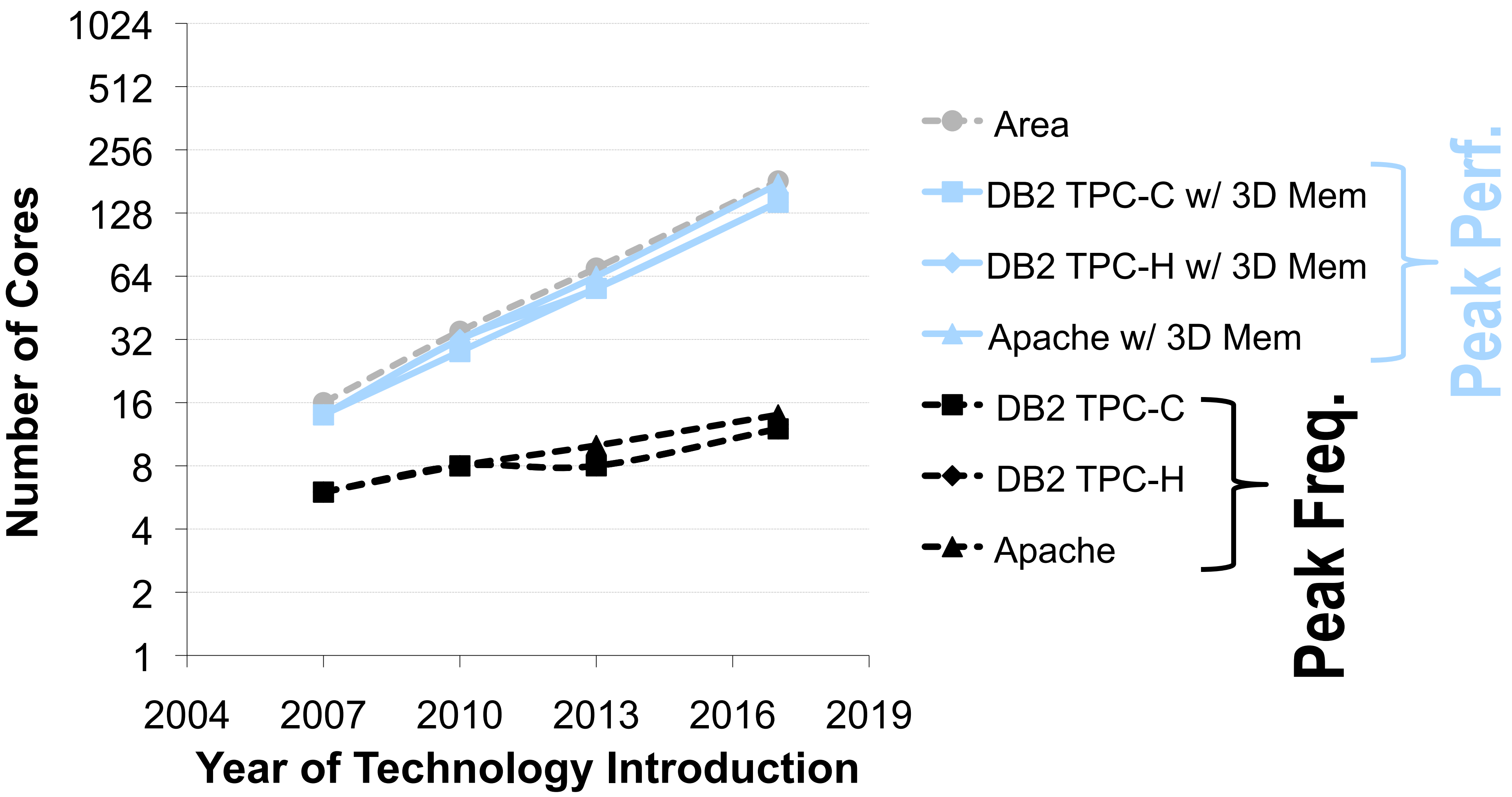
- Only power-constrained
- Virtually eliminates on-chip cache

Core Scaling across Technologies



- Assumes a 130-Watt chip envelope
- Pin b/w keeps Niagara from scaling

Niagara + 3D-stacked Memory



- Power limits Niagara to 75% area!

But, even Niagara is an overkill!

Servers mostly access memory

Benefit little from core complexity

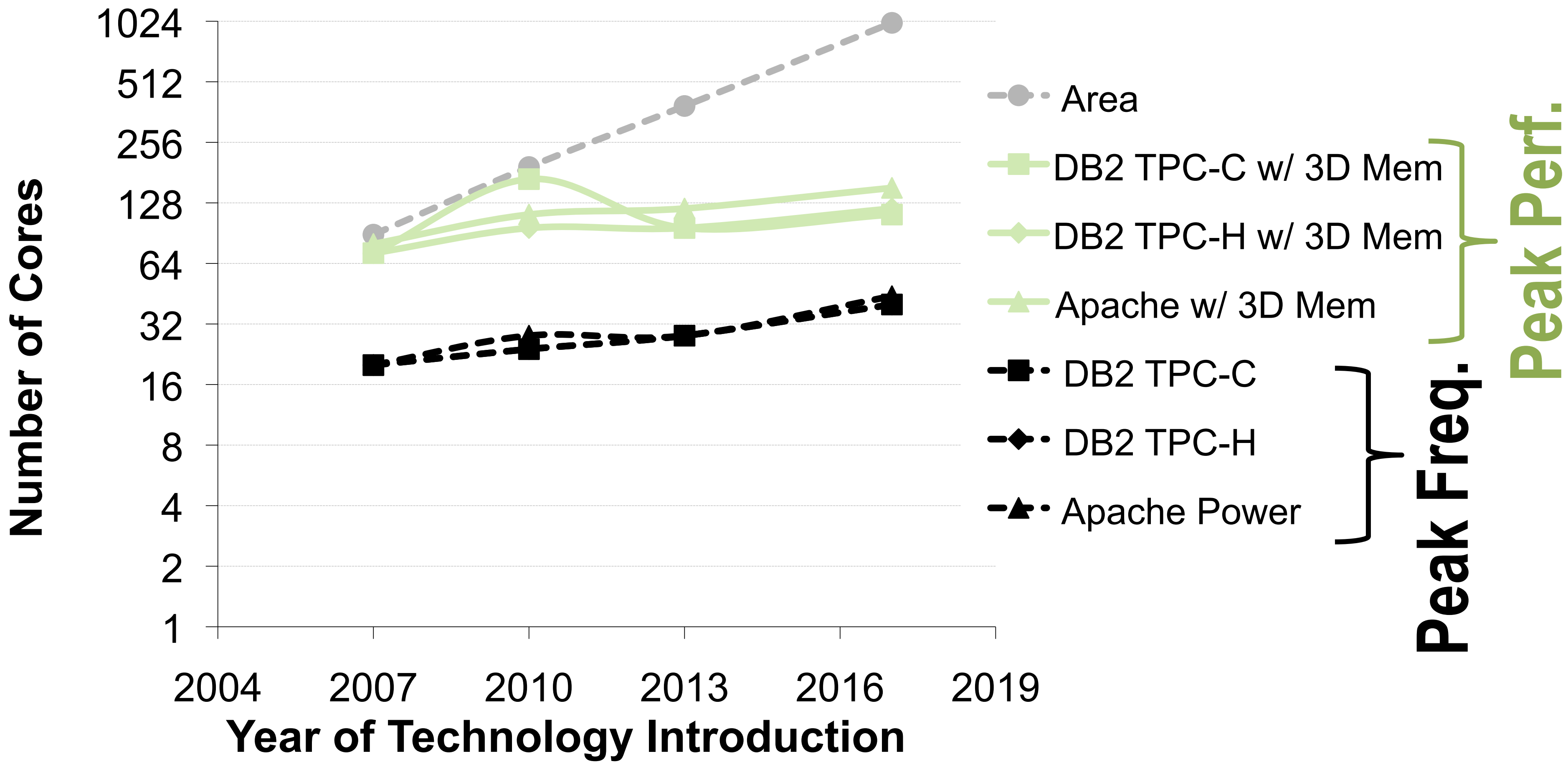
Niagara cores are too big!

E.g., Kgil et al., ASPLOS06:

- Servers on embedded cores + 3D

Can we run servers with embedded cores?

ARM11 + 3D-stacked Memory



- Can not scale with a 130-Watt envelope!!!
- On-chip hierarchy + interconnect not scalable

Outline

- Overview
- Where are we?
- Energy scalability for servers
- Where do we go from here?

➔Activities at EPFL

- Summary

EcoCloud Center @ EPFL
(soon to come @ ecocloud.ch)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

½ dozen labs & industrial affiliates

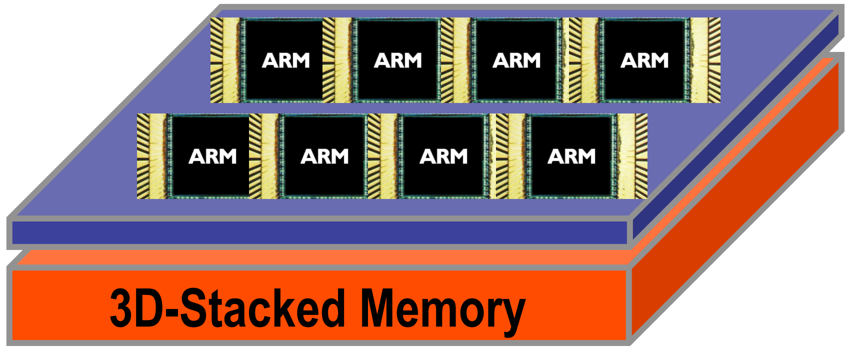
Research:
- Energy-proportional data processing, communication & storage
- Scalable cloud applications & services
- Vertically-integrated computing & cooling

Making tomorrow's clouds green & sustainable

Where do we go from here? Short-term: EuroCloud (Falsafi)

Datacenters with mobile processors

- ARM cores
  - Will likely have to be multithreaded!
- 3D-stacked memory
- Nokia's Ovi Cloud applications



Your 1-Watt Future Datacenter Chip

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE ARM NOKIA Connecting People imec UCyprus

Where do we go from here?
Mid-term: VISA (Ailamaki, Falsafi, Piguet)

Can not power up entire chip!

Vertically-integrated server architecture (VISA)

- Identify services which are energy hogs
- Integrate SW/HW to minimize energy/service
- E.g., Intel's TCP/IP processor @ 1W

Good places to start:

- OS services
- Database services
- Search

Where do we go from here? Long-term: Energy-Scalable IT (Atienza, Kostic, Thome)

Reach beyond IT

- Servers are only part of the equation

Holistic energy scalability

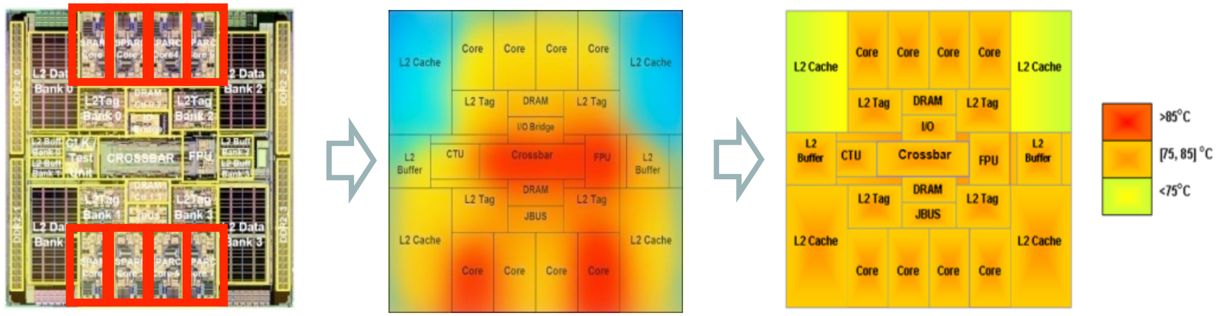
- Computing + generation, delivery, cooling
- E.g., Swiss grocery chain's plan: cook, refrigerate next to datacenter!!!

Much to learn from embedded computing

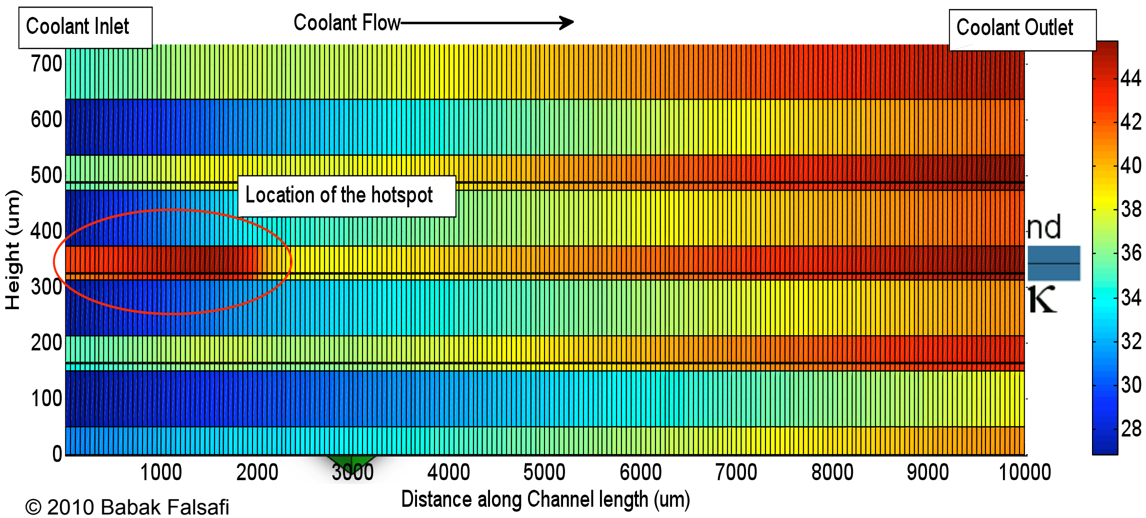
The next decade will be the era of
Energy-Centric IT

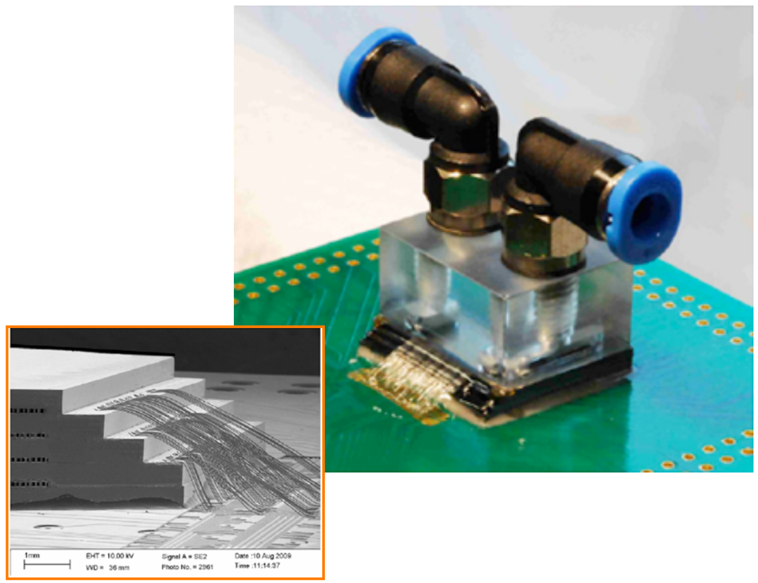
Integrated Thermal Modeling & Cooling (Atienza, Thome)

Heat flow prediction in 3D Niagara (w/Sun)

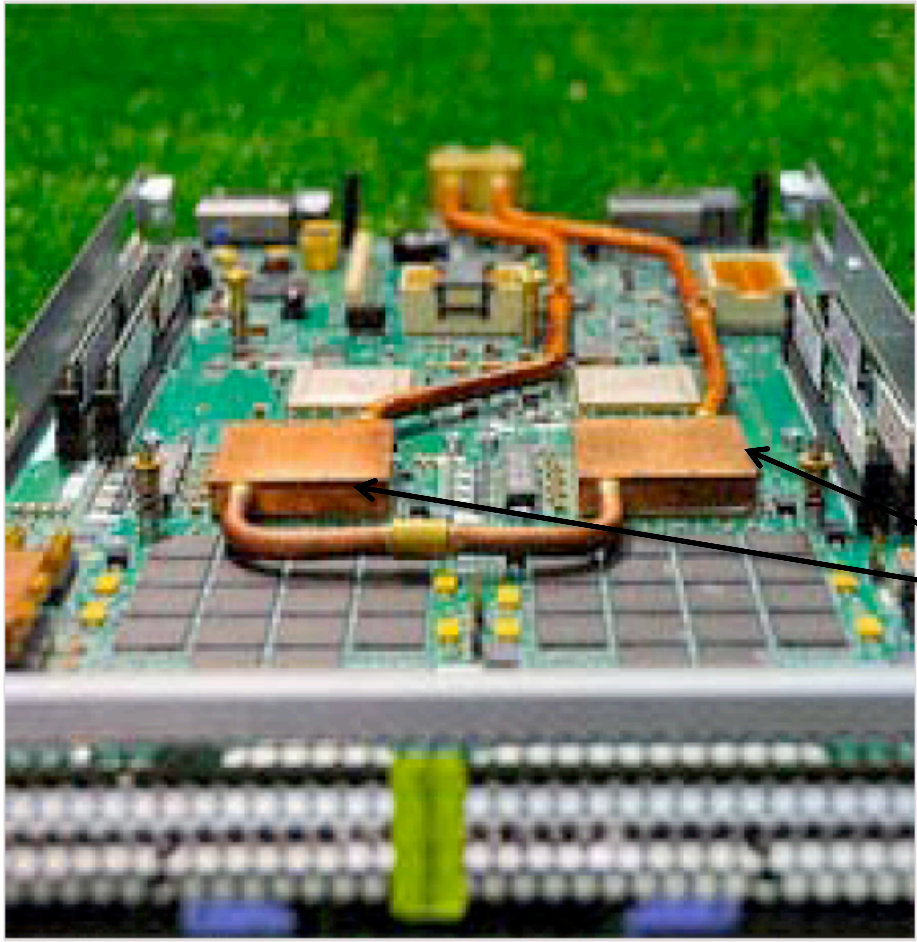


Thermal modeling with active cooling (w/ IBM)

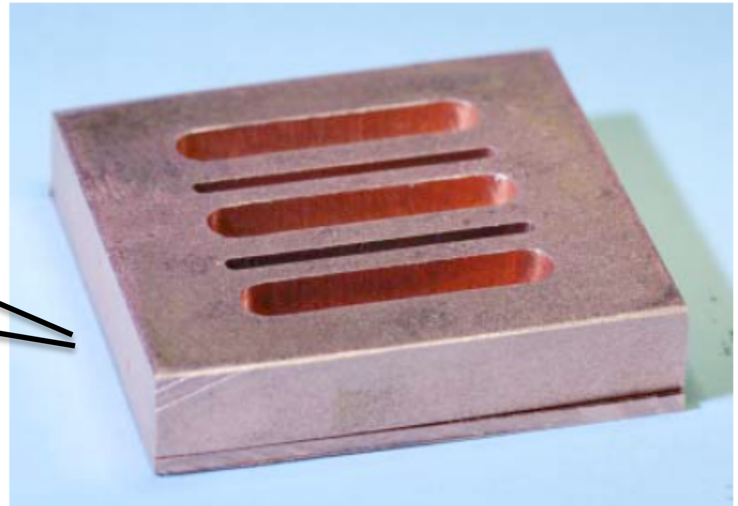




AQUASAR: Water-Cooled Blade Server (Thome)



Project Partners: EPFL, ETH, IBM
Left: Cooling loop with two CPU's
Below: Cooling element without its cover plate and piping ports



Above: Thome (EPFL) dual-purpose cooling element designed for water-cooled and two-phase cooled blades (*two elements in photo at left*).

Research Day @ EPFL
June 17, 2010

Theme: **ECO² Computing**

Location: Rolex Learning Center

Speakers:

- Dan Reed, Corporate Vice President, Microsoft
- Rob Rutenbar, Head of CS, University of Illinois
- Chandrakant Patel, Director of Sustainable IT, HP

For more info, ic.epfl.ch/researchday

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Summary

- Moore's law continues (for another decade)
- CMOS is still cheap
- But, energy scaling has slowed down

Recommendation: Energy-Centric Computing

- Can't get there with parallelism alone
- Holistic approach to energy

Time to put the "embedded" into all of computing!