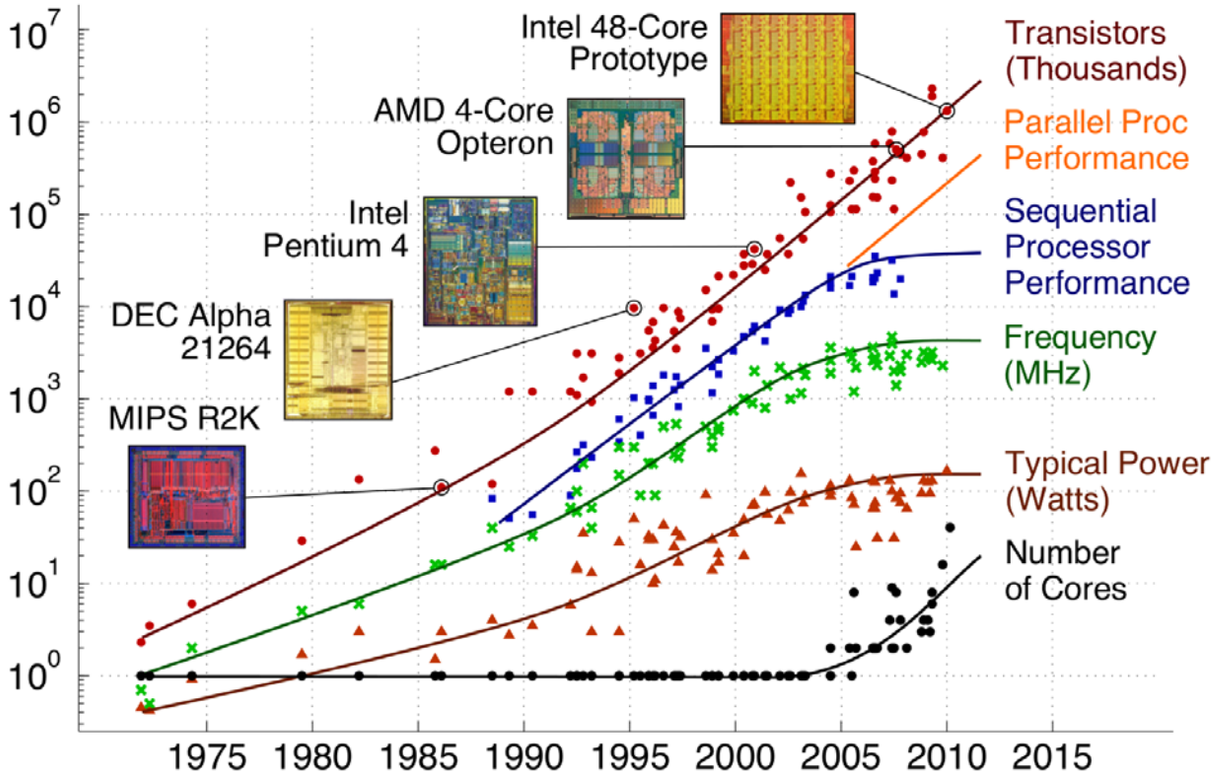
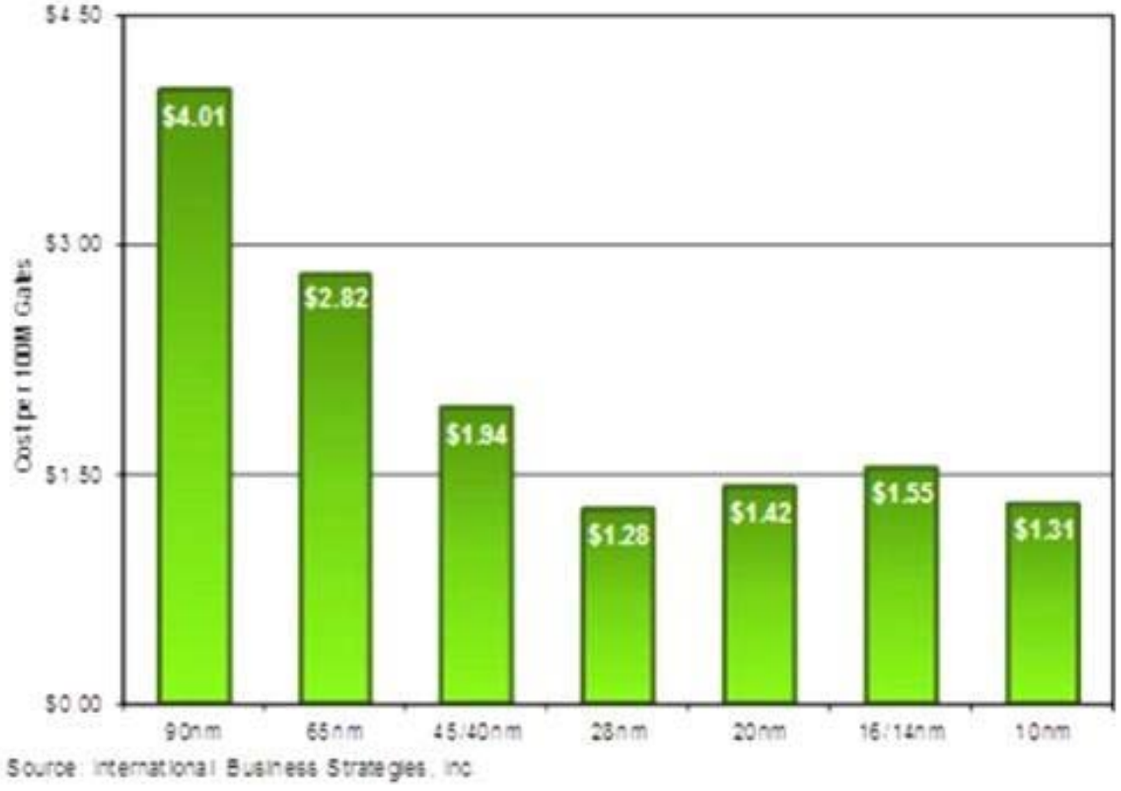
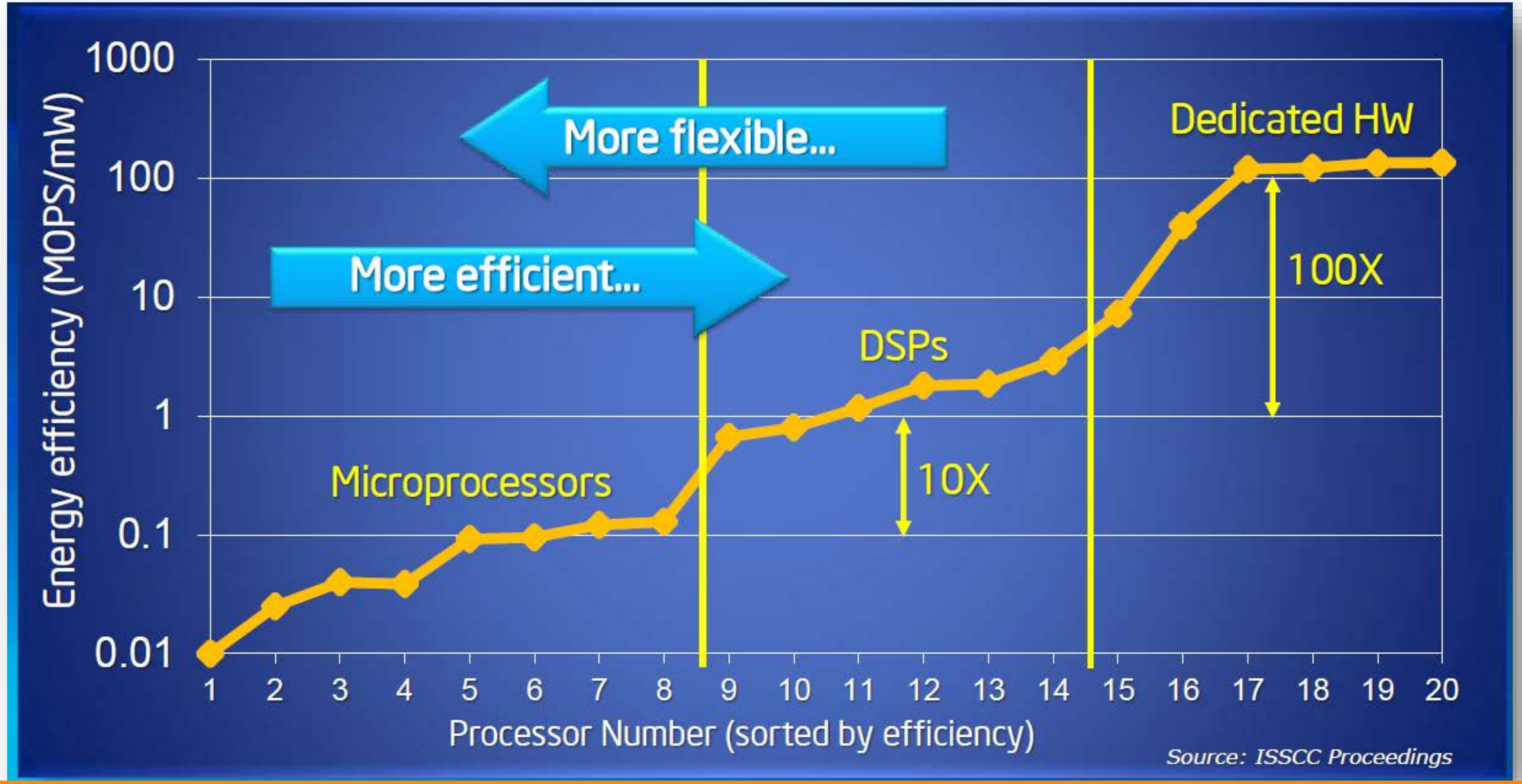

IMAGES VIDEOS MAPS NEWS SEARCH HISTORY MORE MSN OUTLOOK.COM
Make Bing my homepage
2
250
Sign in
bing
Accelerating Large-Scale Datacenter Services
Andrew Putnam – Microsoft
FPL 2015 – RC4Masses
© Image Credits
© 2014 Microsoft | Privacy and Cookies | Legal | Advertise | About our ads | Help | Feedback

Moore's Law is Ending





Capabilities, Costs $\propto \frac{Performance/Watt}{\$}$



Increase Efficiency with Hardware Specialization

Datacenter Environment

- Machines last 3 years, purchased on a rolling basis
- Machines usually repurposed multiple times
- Little/no HW maintenance, no accessibility
- Homogeneity is highly desirable

The paradox: Specialization *and* homogeneity

What is “The Datacenter”?

- Massive scale – 10k, 100k, 1M+
- Many diverse, fast changing workloads

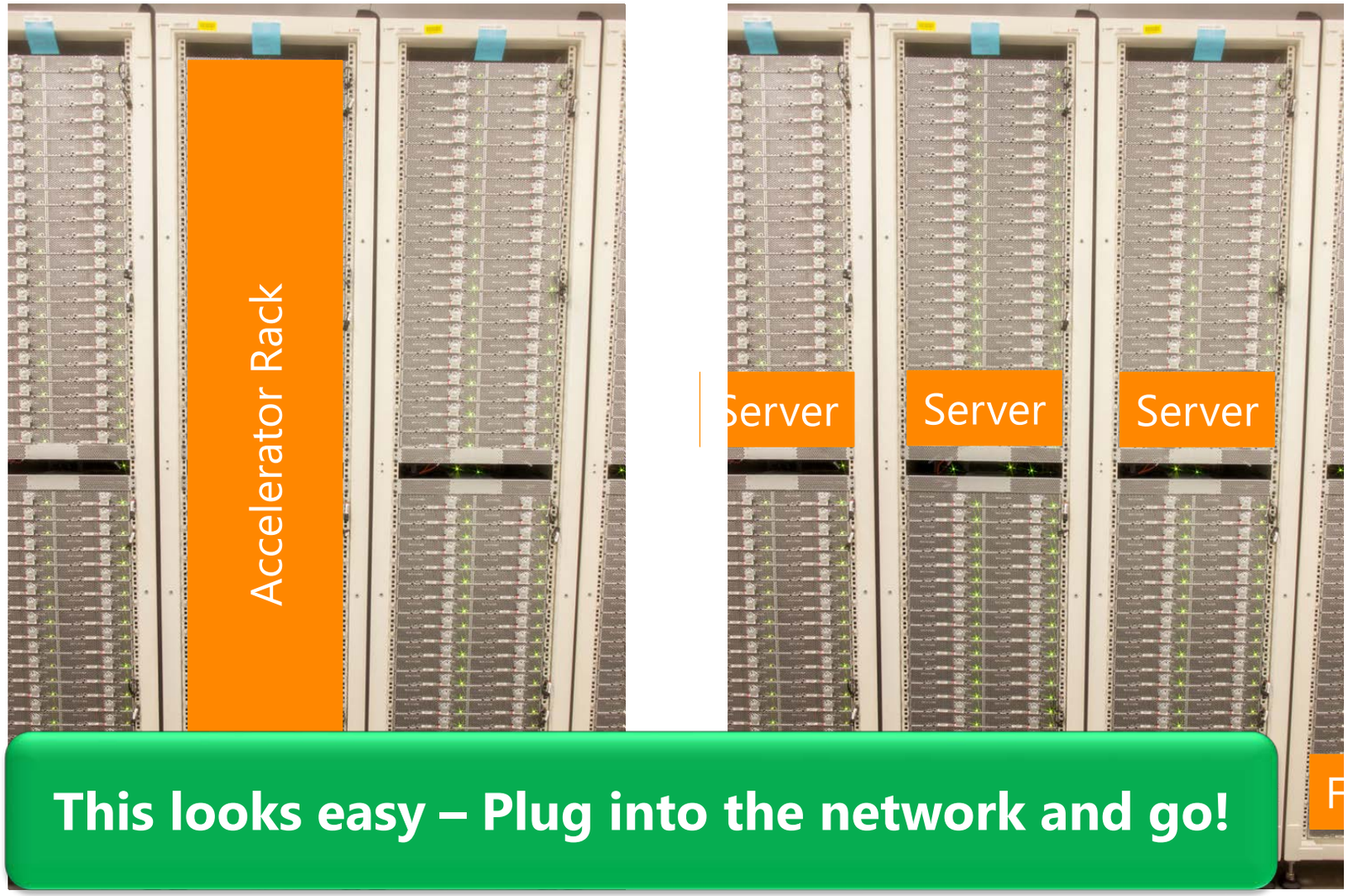
Outlook.com Windows Azure msn bing Microsoft SQL Server XBOX LIVE Microsoft Office 365 skype COSMOS OneDrive

What is “The Datacenter”?

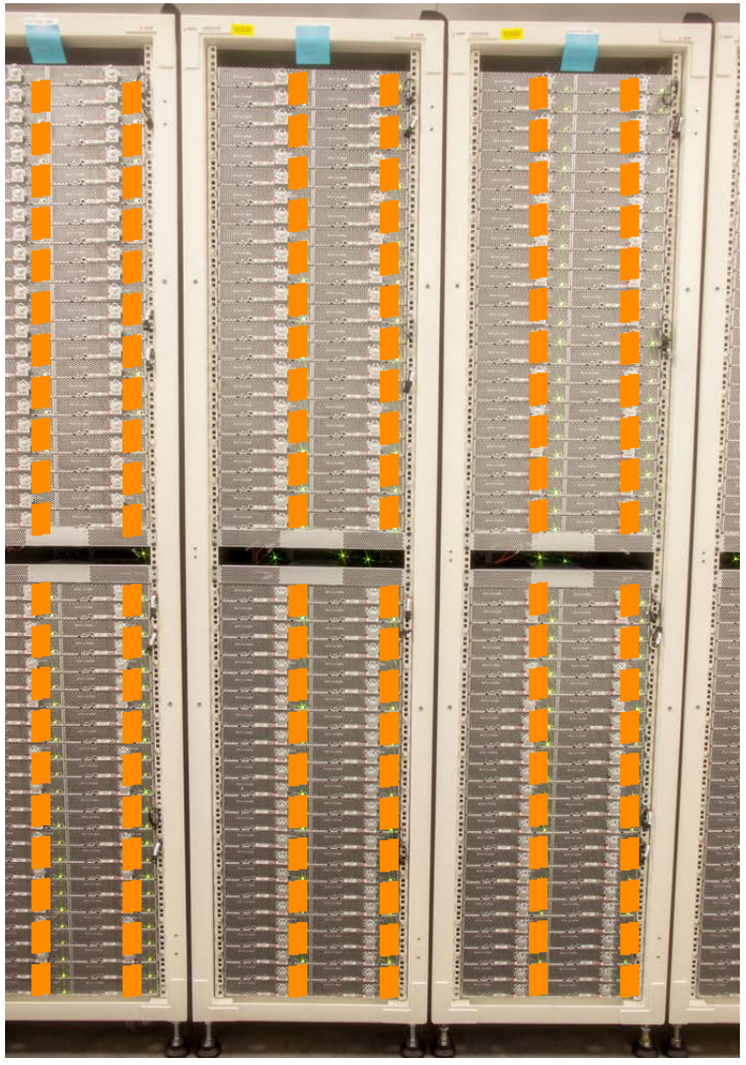
- Massive scale – 10k, 100k, 1M+
- Many diverse, fast changing workloads

Outlook.com Windows Azure msn bing Microsoft SQL Server XBOX LIVE Microsoft Office 365 skype COSMOS OneDrive

Integrating Accelerators in the Datacenter



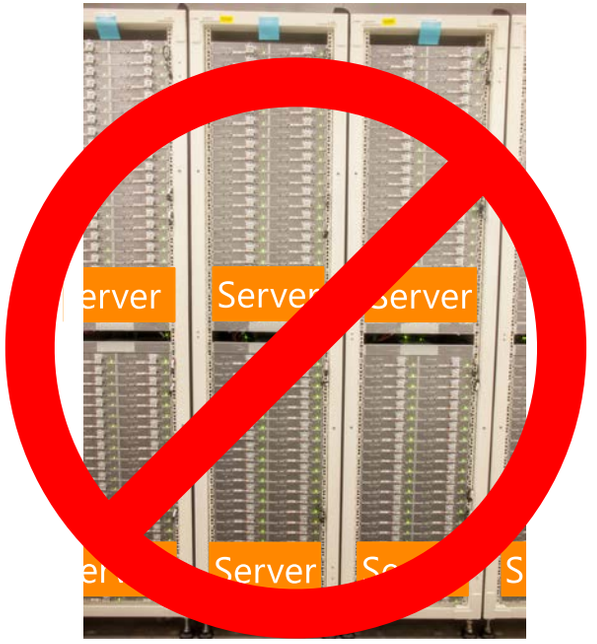
This looks easy – Plug into the network and go!



Centralized **Distributed**

Centralized Model Unsuitable for Datacenter

- Single point of failure
- Complicates rack design, thermals, maintainability
- Network communication for any use of accelerator
  - Definition of the Network In-cast problem
  - Requires larger chunks of code for offload
  - Precludes many latency-sensitive workloads
- Limited elasticity



Datacenter Servers

- Microsoft Open Compute Server
- 1U, ½ wide servers
- Enough space & power for ½ height, ½ length PCIe card
- Squeeze in a single FPGA
- Won't fit (or power) GPU

http://www.globalfoundationservices.com/posts/2014/january/27/microsoft-contributes-cloud-server-specification-to-open-compute-project.aspx

Microsoft Open Compute Server

- Two 8-core Xeon 2.1 GHz CPUs
- 64 GB DRAM
- 4 HDDs @ 2 TB, 2 SSDs @ 512 GB
- 10 Gb Ethernet
- No cable attachments to server

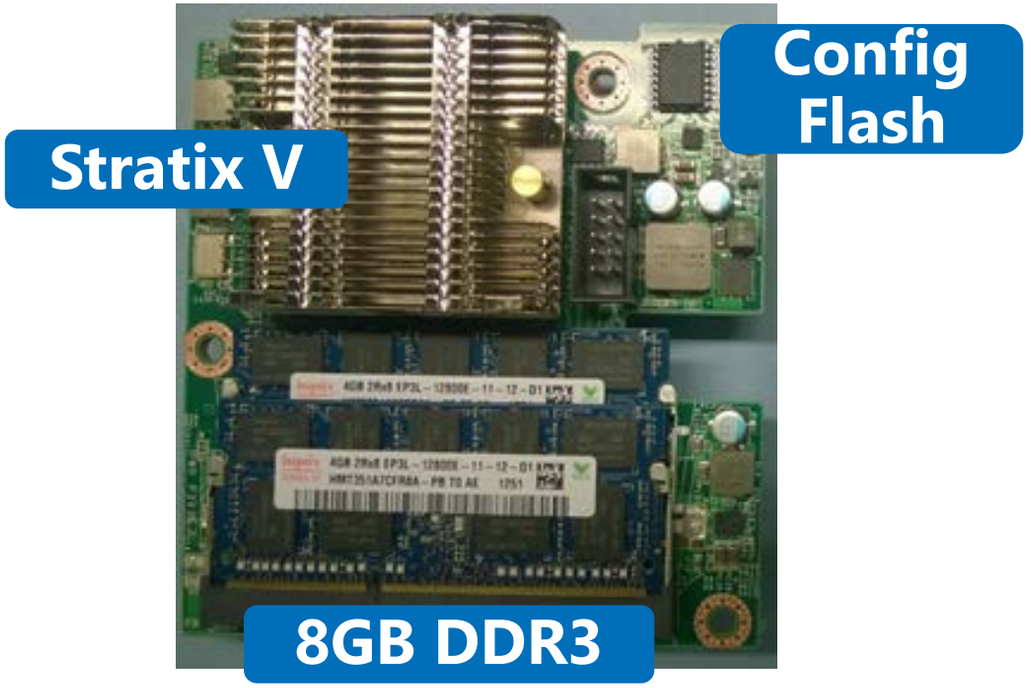
Air flow

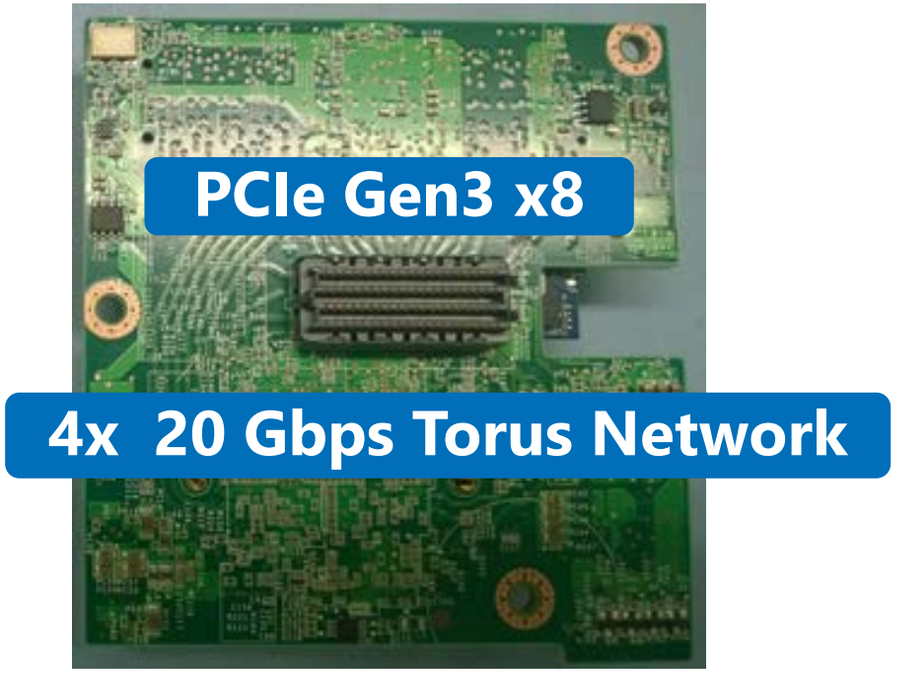
200 LFM

70 ^{0}C Inlet

Catapult FPGA Accelerator Card

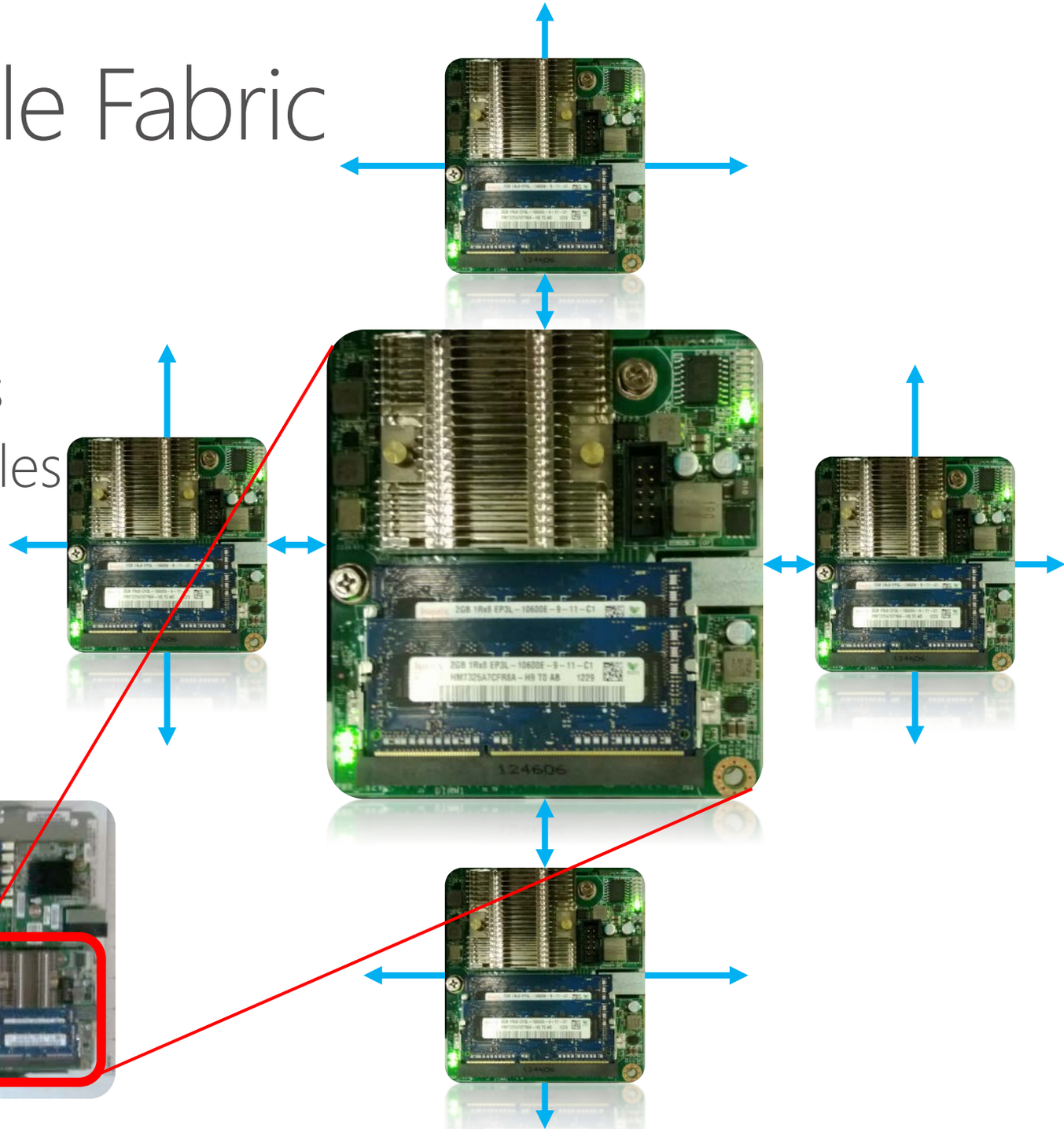
- Altera Stratix V GS D5
  - 172k ALMs, 2,014 M20Ks, 1,590 DSPs
- 8GB DDR3-1333
- 32 MB Configuration Flash
- PCIe Gen 3 x8
- 8 lanes to Mini-SAS SFF-8088 connectors
- Powered by PCIe slot



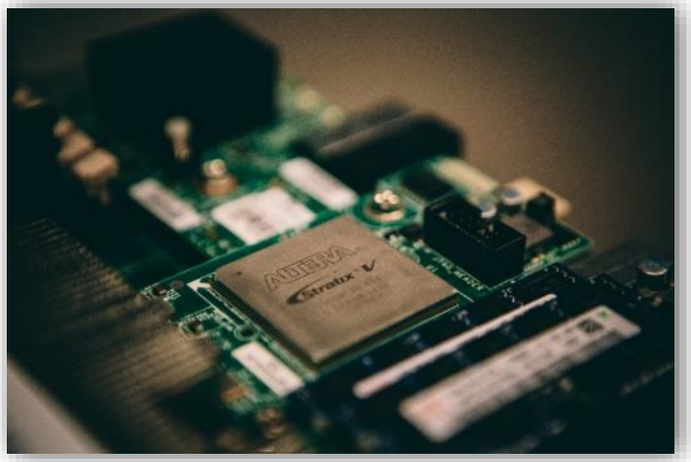


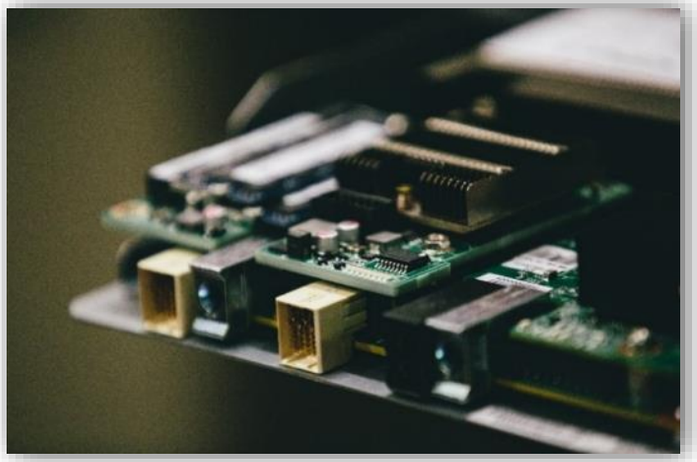
Scalable Reconfigurable Fabric

- 1 FPGA board per Server
- 48 Servers per ½ Rack
- 6x8 Torus Network among FPGAs
  - 20 Gb over SAS SFF-8088 cables



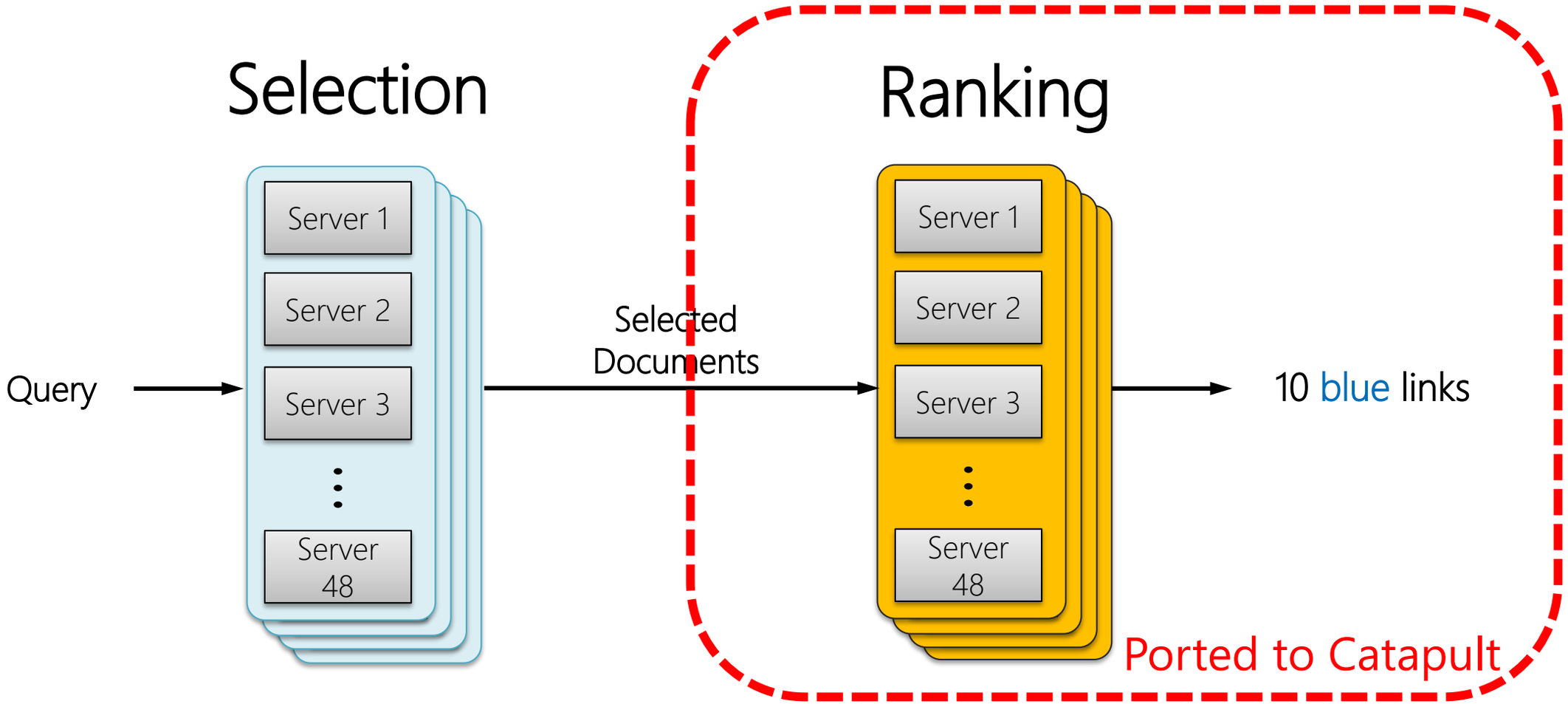
Data Center Server (1U, ½ width)





1,632 Server Pilot Deployed in a Production Datacenter

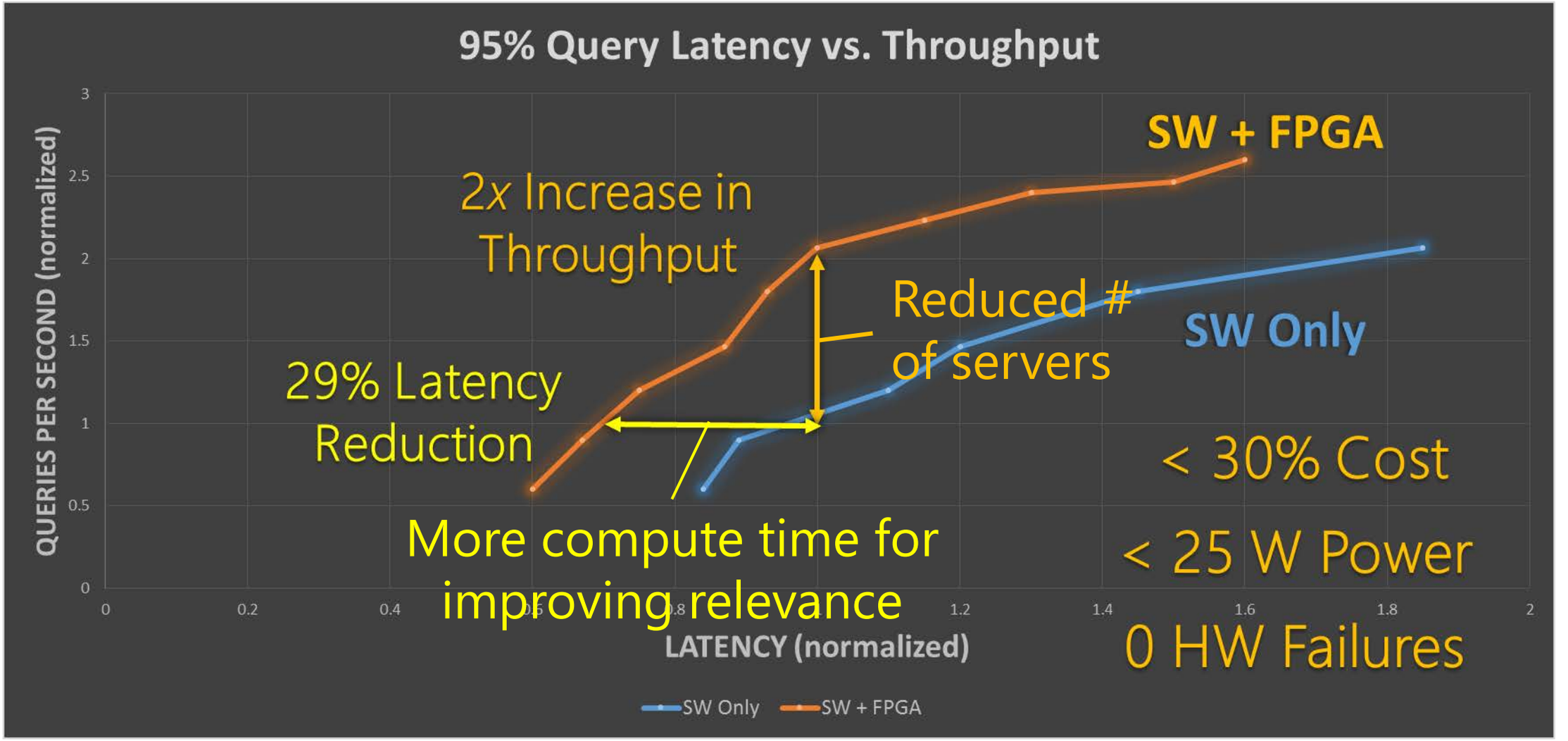
Bing Web Search



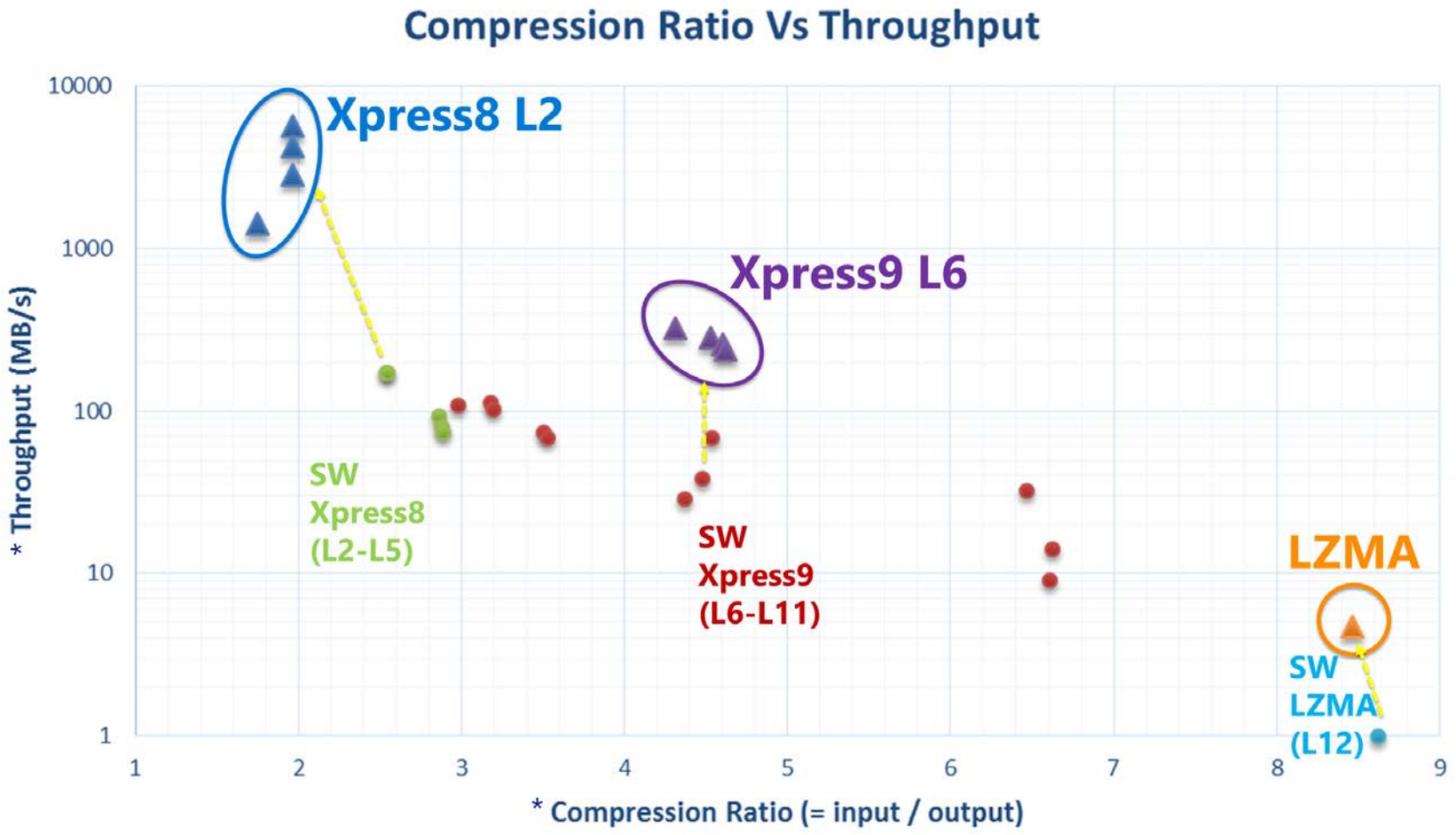
~30,000 lines of C++

Bing Page Ranking

1,632 Servers with FPGAs Running Bing Page Ranking Service (~30,000 lines of C++)



Data Compression



*- Measured on Canterbury dataset

Xpress8 L2 (5.6GB/s)

- 30x throughput
- 20% compression loss
- In-line compression

Xpress9 L6 (300MB/s)

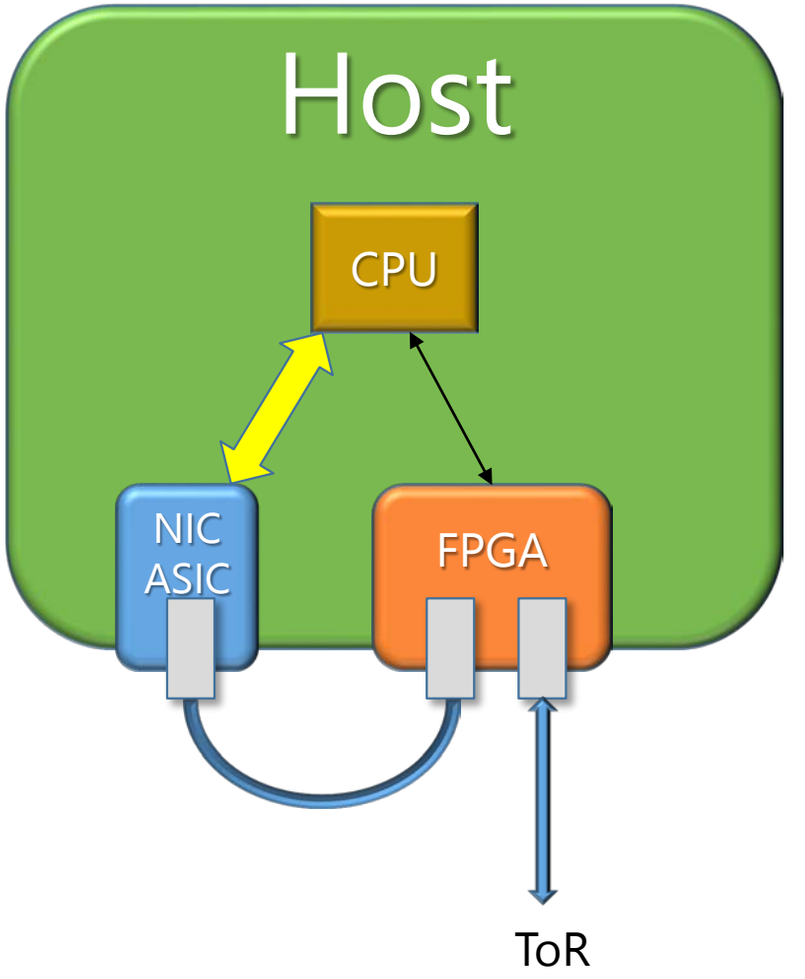
- 4x throughput
- No compression loss
- Short/mid-term data

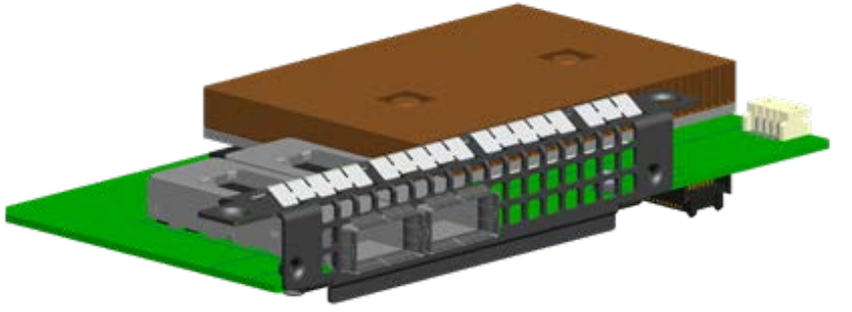
LZMA (5MB/s)

- 5x throughput
- 5% compression loss
- Long-term storage

Azure SmartNIC

- Use Catapult FPGAs for reconfigurable functions
  - Already used in Bing
  - Roll out Hardware as we do software
- Programmed using Generic Flow Tables (GFT)
  - Language for defining Software Defined Networks (SDNs)
- SmartNIC can do Crypto, QoS, storage acceleration, and more…





Deep Learning -- Image Classification via CNN

2x 8-core 2.10 GHz Xeon (95W TDP)

One Stratix V D5 FPGA (25 W)

Projected Improvements with Tuning

Platform	Library/OS	ImageNet 1K Inference Throughput	Peak TFLOPs	Effective TFLOPs	Estimated Peak Power for **CNN Computation**	Estimated GOPs/J (assuming peak power)
CPU **16-core, 2-socket Xeon E5-2450, 2.1GHz**	Caffe + Intel MKL Ubuntu 14.04.1*	53 images/s	0.27T	0.074T (27%)	~225W	~0.3
FPGA **Arria 10 GX1150**	Windows Server 2012	369 images/s **~880 images/s**	1.366T	0.51 T (38%) ~1.2T (89%)	~37W ~40W	~12.8 ~30.6
GPU **NervanaSys-32 on NVIDIA Titan X**	NervanaSys-32 on Ubuntu 14.0.4	4129 images/s[2]	6.1T	5.75T (94%)	~250W	~23.0

**-Projected Results*

[1]Dense layer time estimated

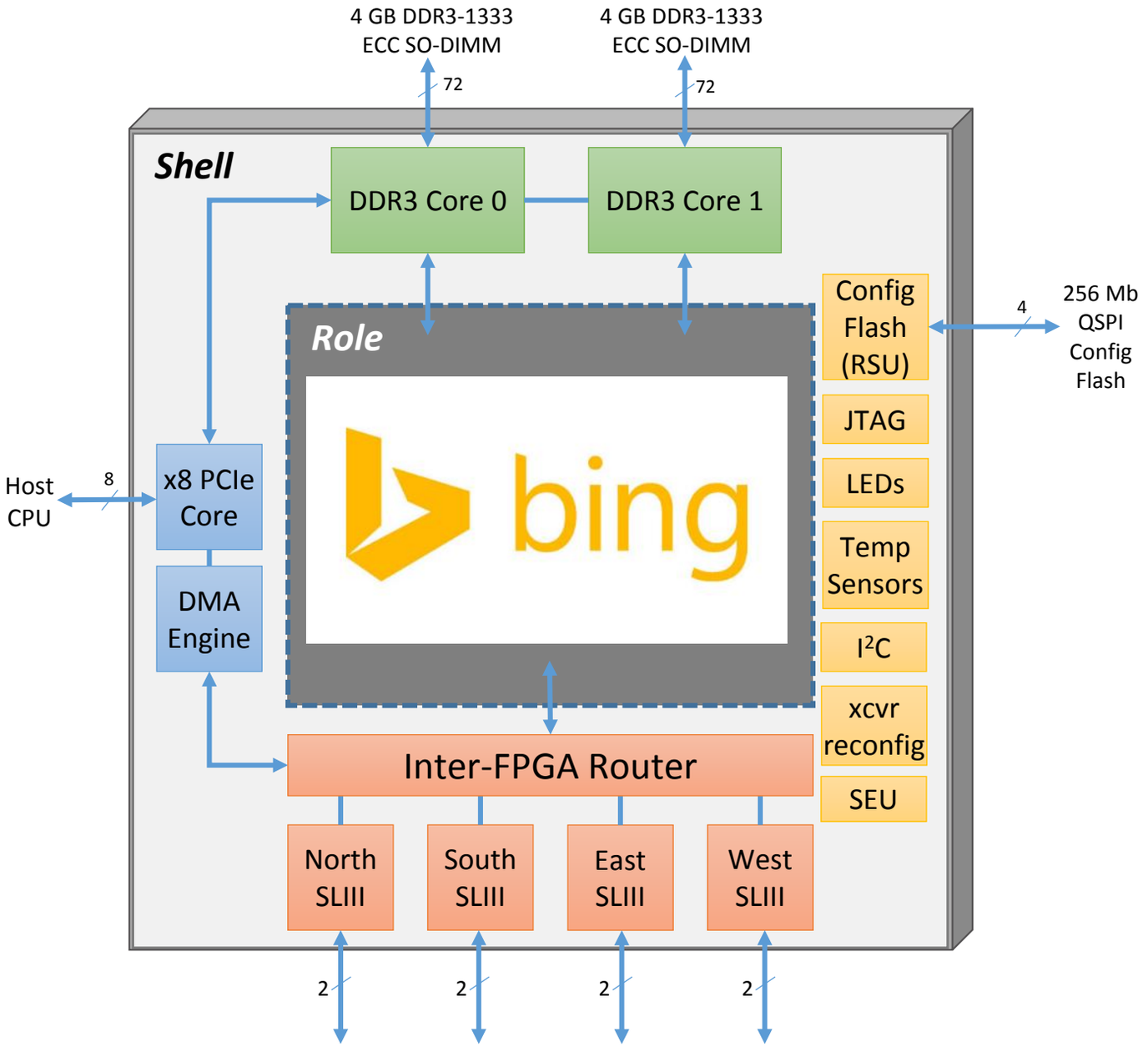
[2]https://github.com/soumith/convnet-benchmarks

Programming

- So far, all applications have been written in SystemVerilog
  - Bing – Domain specific program
  - ML, Compression, SmartNIC – Function Libraries
- OpenCL, MPI, Delite, Vivado HLS can (hopefully) help
- Not many datacenter applications look like matrix multiply / loop unrolling
- Remember – datacenter applications change *daily*.
  - Compile time / design space exploration are critical

Shell & Role

- *Shell* handles all I/O & management tasks
- *Role* is only application logic
- Shell exposes simple FIFOs
- Flight data recorder for scale-out debug
- Role is Partial Reconfig boundary



Easy Application Use & Migration

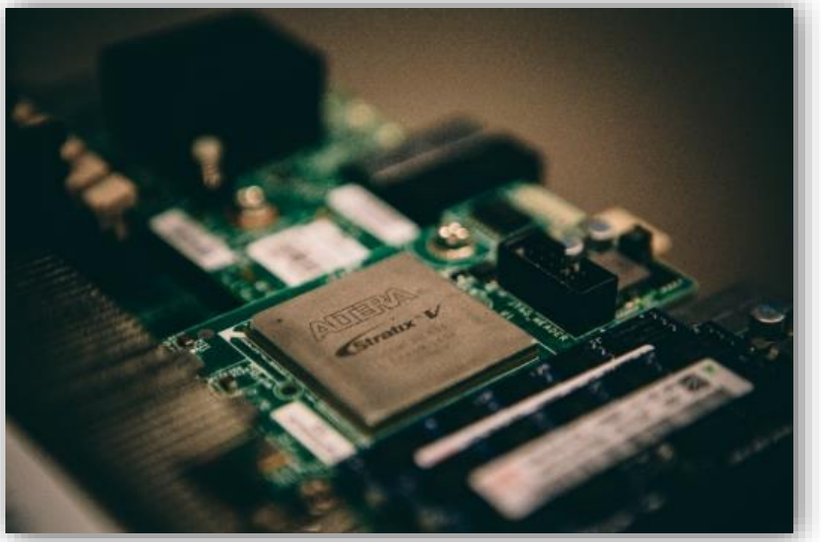
- Shell/Role + API enable developers to focus on their application, not on board specifics
- Build for variable numbers of I/O (PCIe, DRAM, Net)
  - Use shims if applications require more than physically available
- Performance may differ, but it's a clock-domain crossing FIFO, so applications remain independent
- Easy, stable target for HLS tools

Where do GPUs fit in?

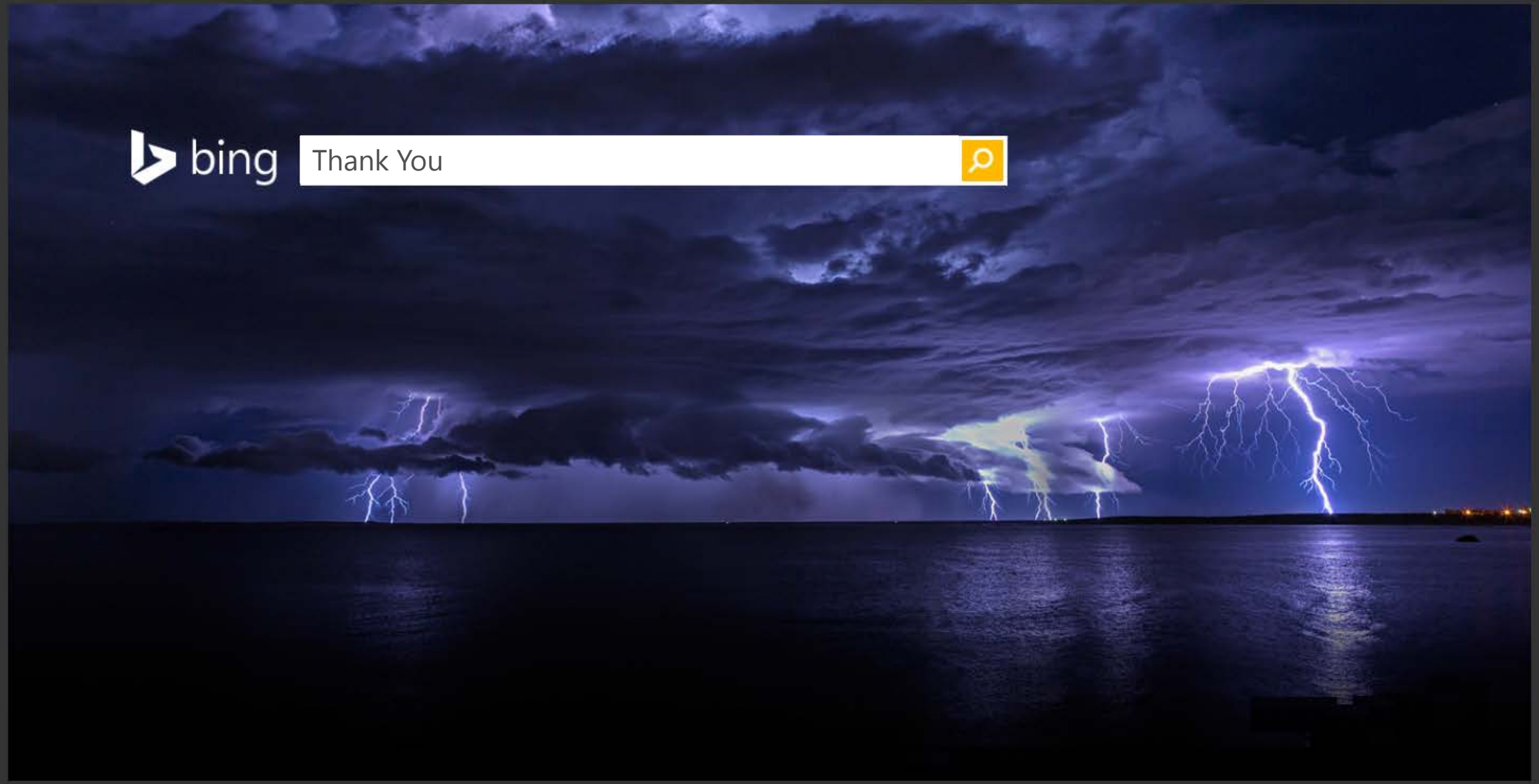
- Ratio of GPUs to CPUs for machine learning is ~8:1
- 8:1 boxes do not fit most datacenter workloads
- Mobile GPU performance not much better (if any) than two big Xeon processors
- Expect a few clusters of big 8:1 machines, but widespread GPU integration is an open question

Conclusions

- Catapult FPGA architecture allows specialization with homogeneity for major datacenter workloads
  - Search, Machine Learning, SDNs, encryption, compression… many more
- Power efficiency allows growth through scaling
- Greatest Challenges:
  - Programmability – All applications so far are HDL
  - Compile Time ∝ Development time & agility
  - Debugability
  - Forward Compatibility



- Just the start for FPGAs computing


IMAGES
VIDEOS
MAPS
NEWS
SEARCH HISTORY
MORE
MSN
OUTLOOK.COM
Make Bing my homepage
2
250
Sign in
bing
Thank You
© Image Credits
© 2014 Microsoft | Privacy and Cookies | Legal | Advertise | About our ads | Help | Feedback