

# Generalization vs Specialization in cloud computing

**Gustavo Alonso**

**Systems Group**

**Department of Computer Science**

**ETH Zurich, Switzerland**

# ETH Systems Group

[www.systems.ethz.ch](http://www.systems.ethz.ch)



# What is the cloud?

Primarily a change in business model driven  
by computing demand

-

A shift from manufacturing to services

# What is big data?

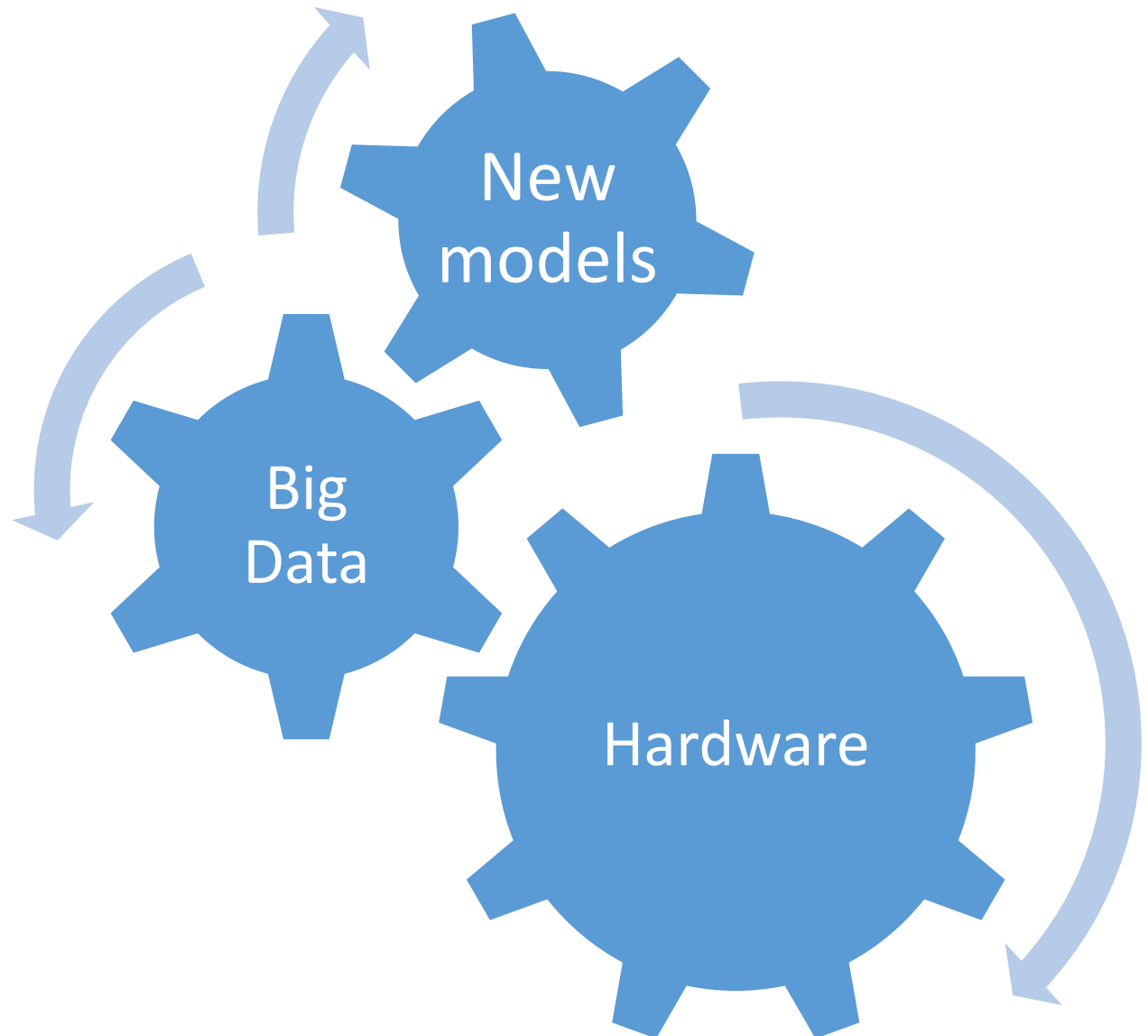
Primarily a change in business model driven by the ability to process large data collections

-

A shift toward customization and personalization of “services”  
(from services to information & human services)

The case for general  
purpose =  
Economies of scale

# The case for specialization



## Big data

- Social graphs
- Page rank
- Science data
- HPC
- Genomic

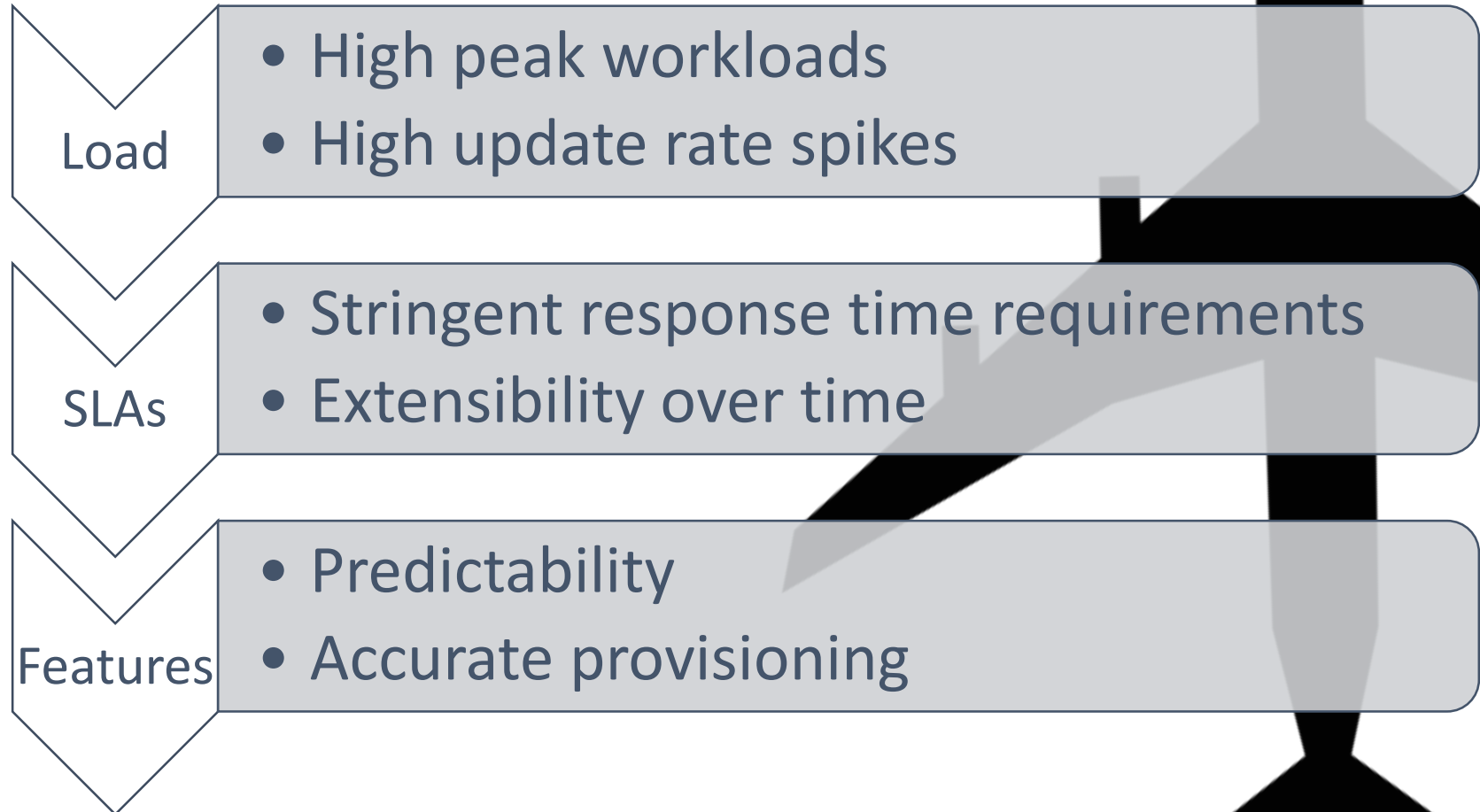
## Hardware

- Multicore
- GPUs
- FPGAs
- RDMA
- Smart NICs
- Active memory

## New Models

- Data center
- Appliances
- Cloud computing

# Big Data Example

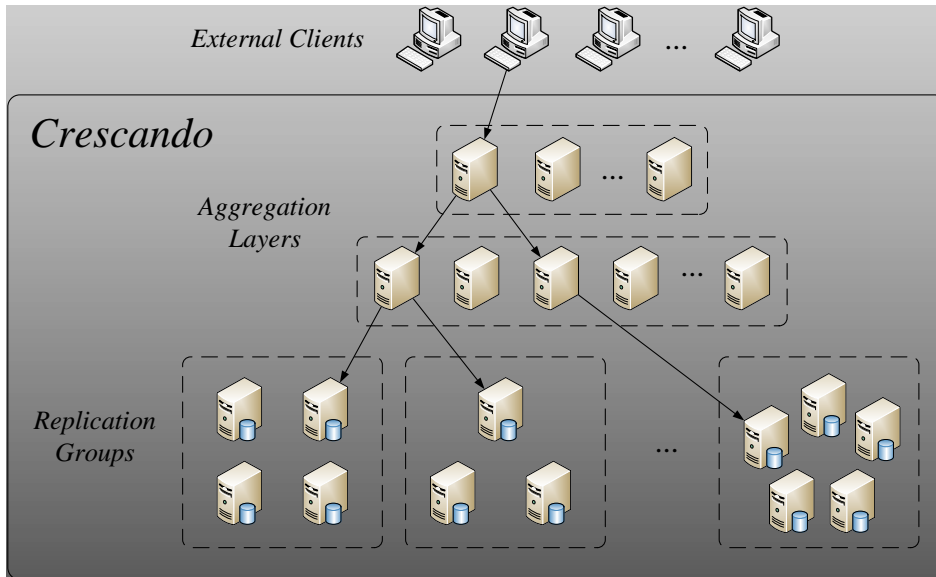
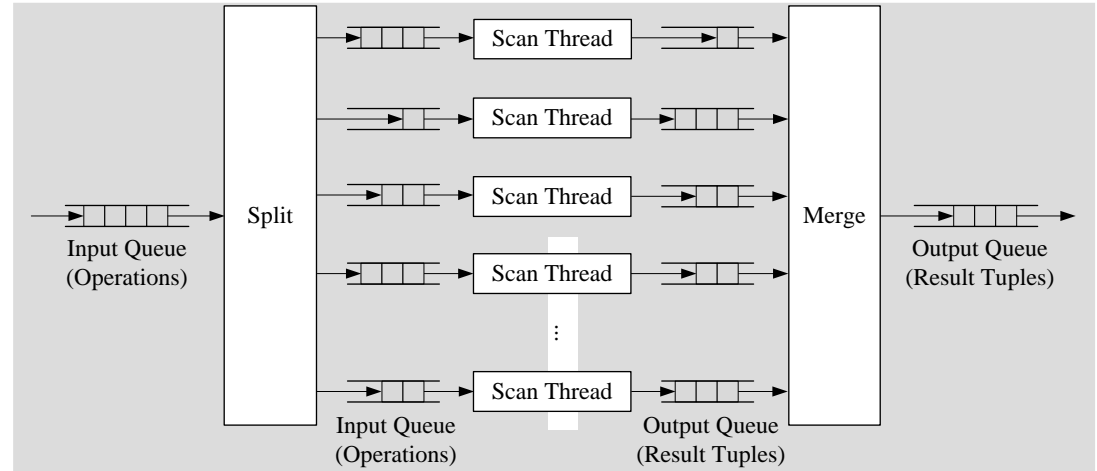




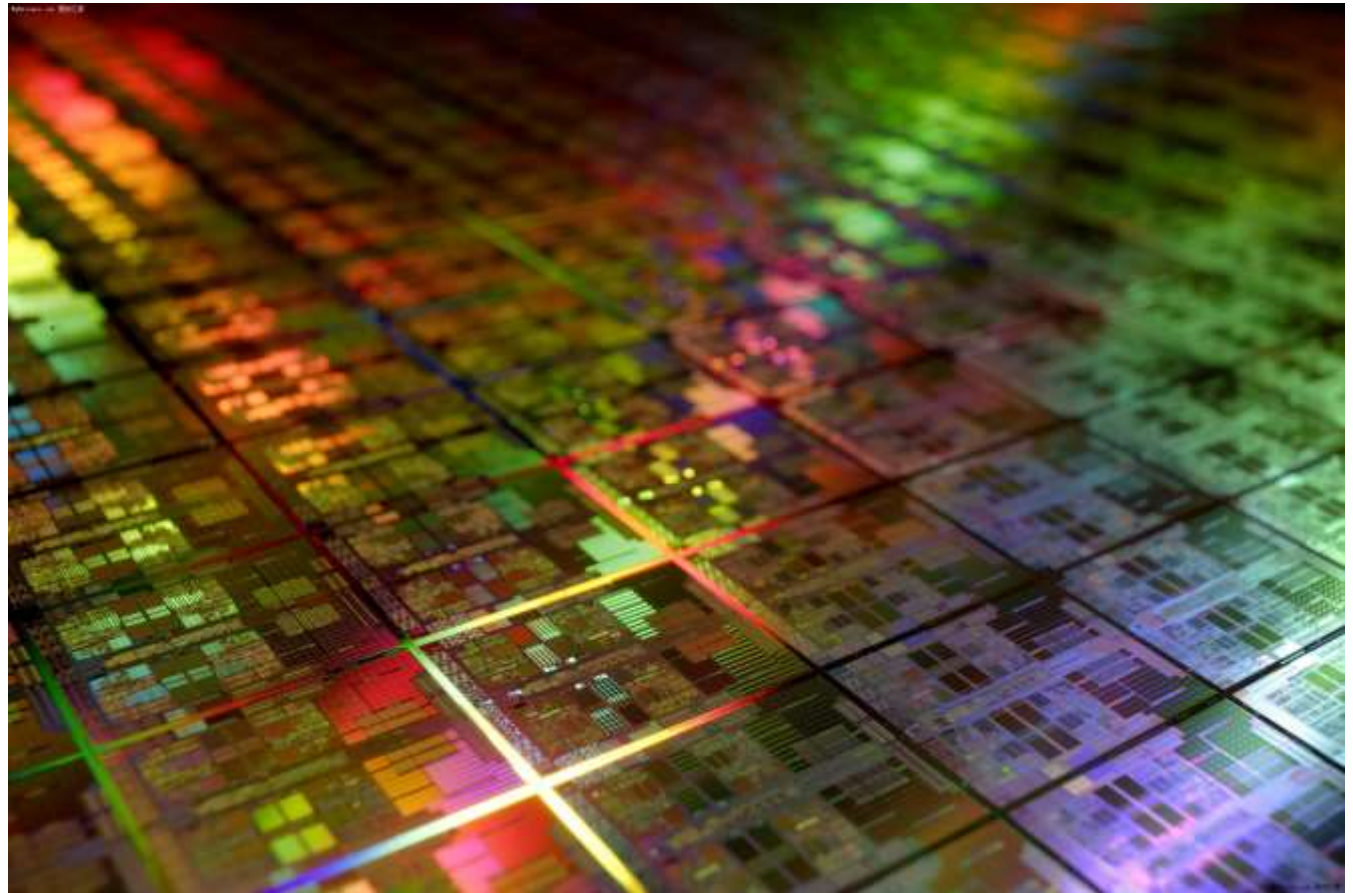
# CRESCANDO

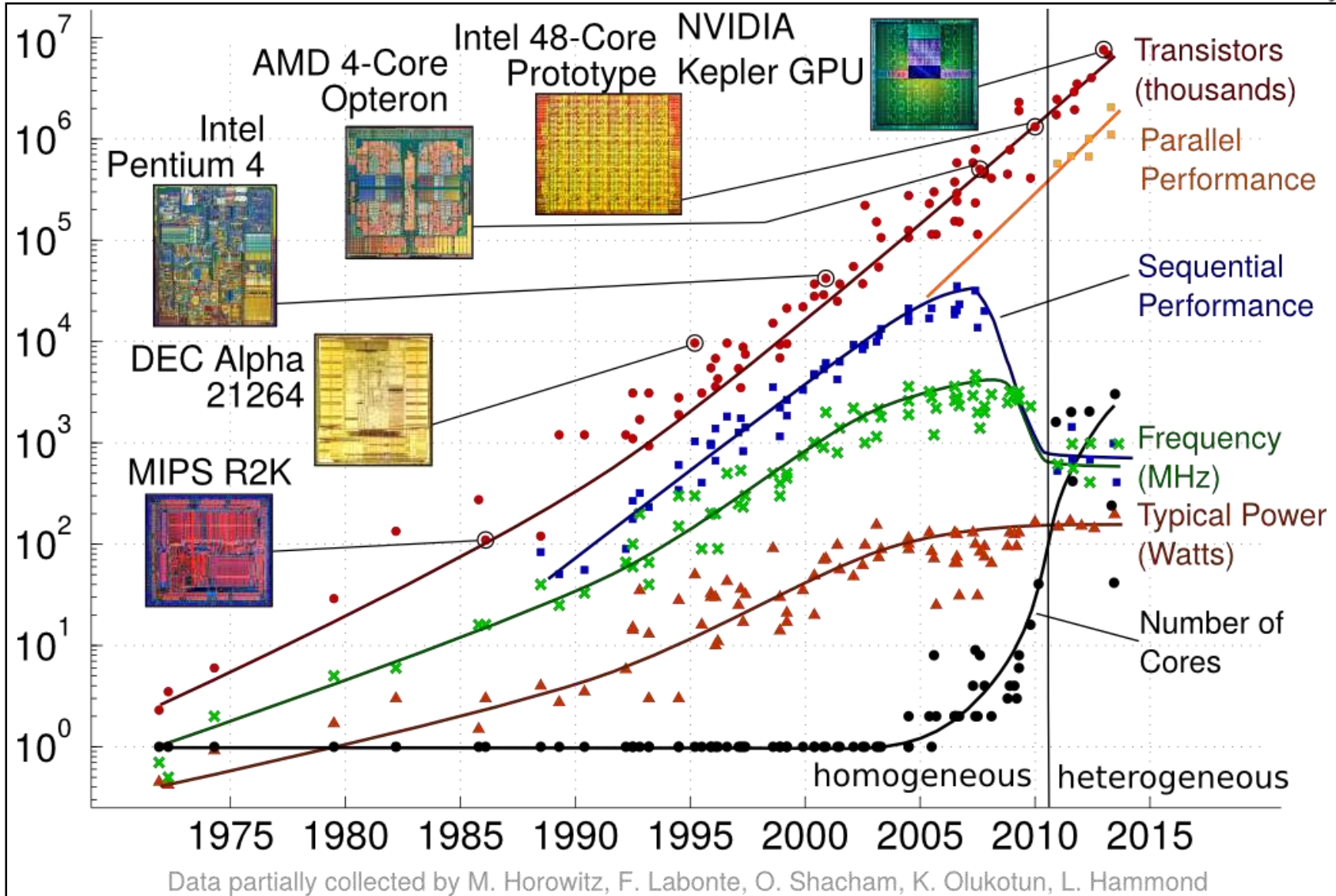
QUERIES  
UPDATES

BUILD QUERY  
INDEX FOR  
NEXT SCAN



# Modern Hardware



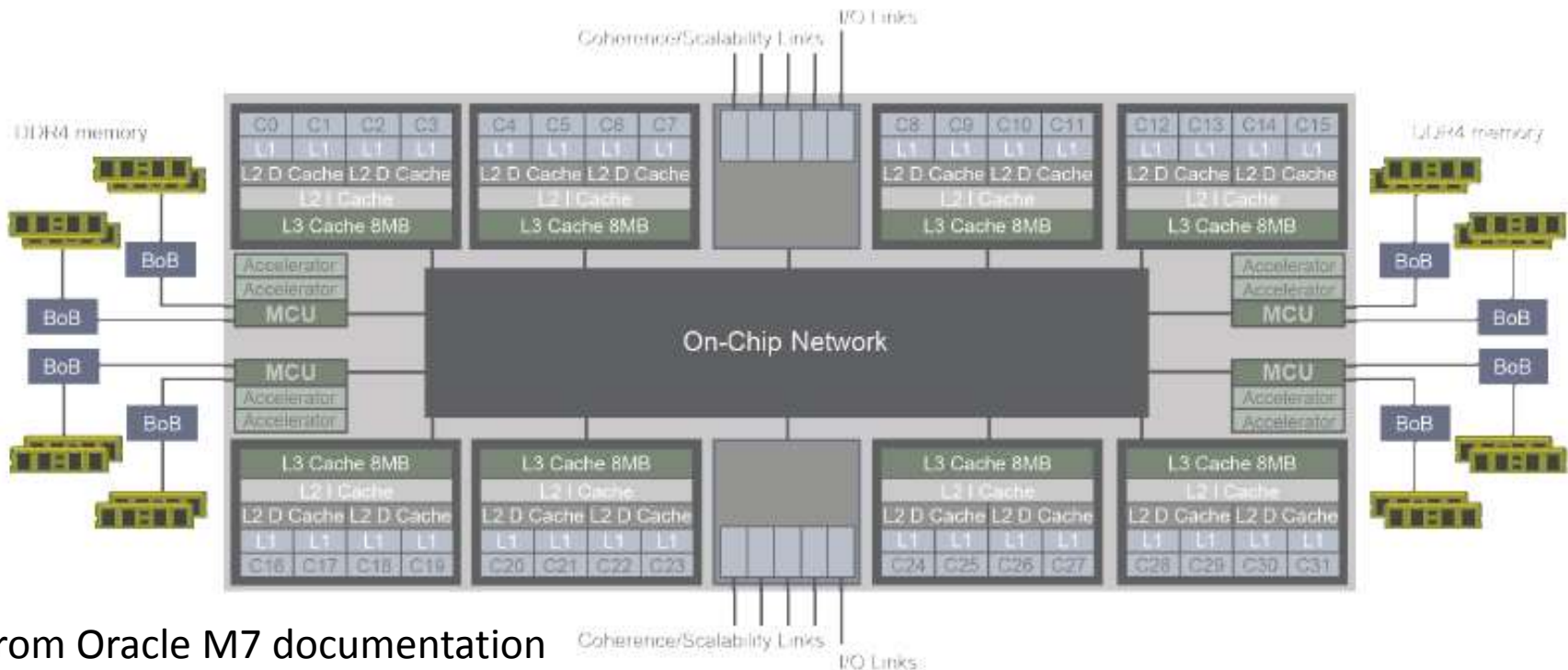


Slide courtesy of Torsten Hoefler (Systems Group, ETH Zürich)

# Example 1: Processors

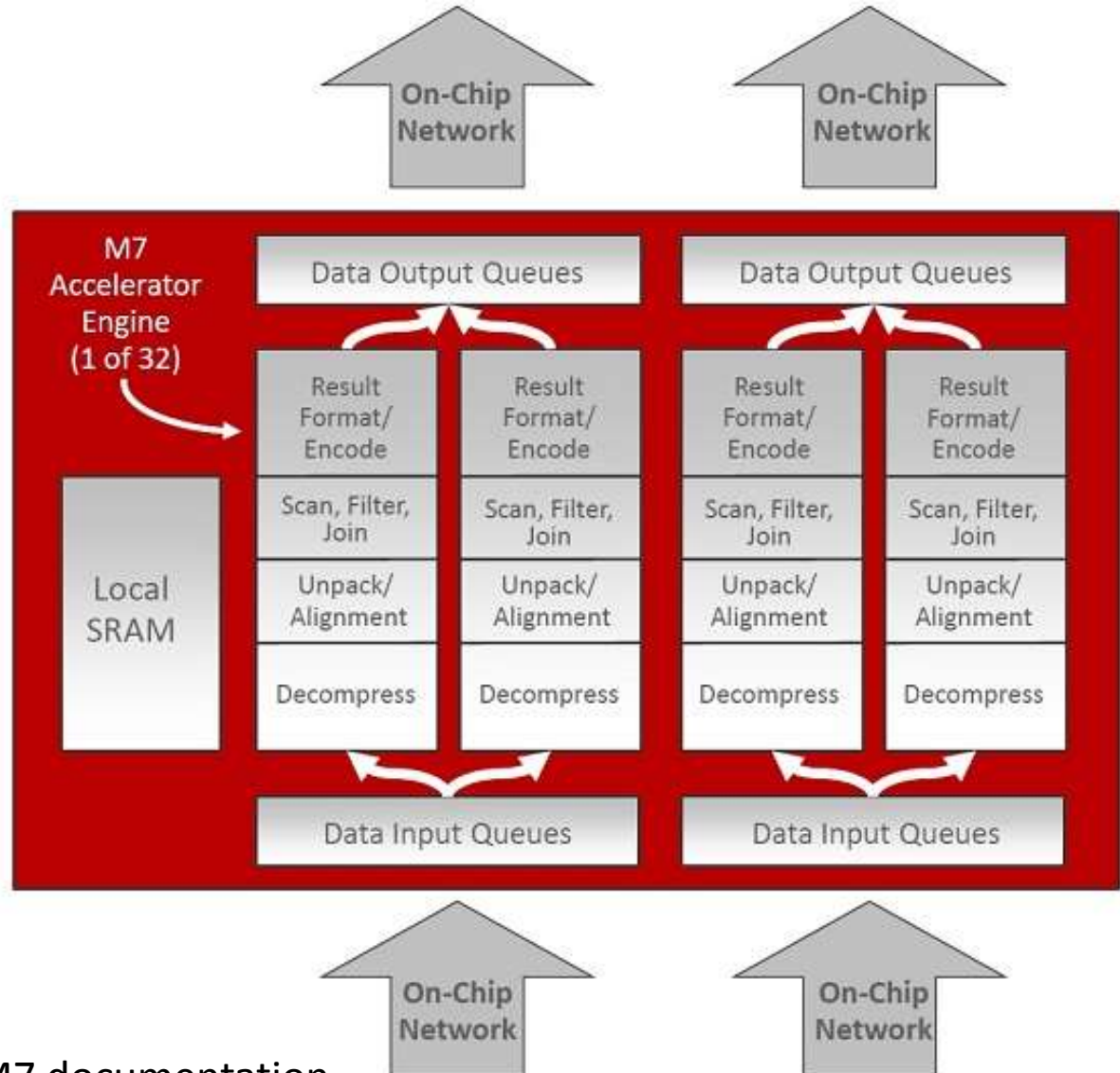
Oracle's SPARC M7 processor: "SQL in silicon" accelerators processing streams of data from memory:

Decompress, Scan, Select, Translate



From Oracle M7 documentation

# Accelerators to come



From Oracle M7 documentation

# Example 2: Architectures

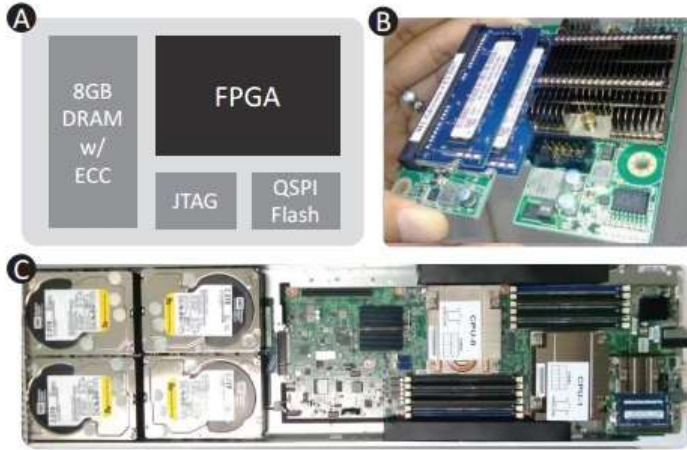


Figure 1: (a) A block diagram of the FPGA board. (b) A picture of the manufactured board. (c) A diagram of the 1 U, half-width server that hosts the FPGA board. The air flows from the left to the right, leaving the FPGA in the exhaust of both CPUs.

## Microsoft Catapult

Data center nodes with FPGA boards

FPGAs connected through their own network (Torus topology)

Search and deep learning applications

## Microsoft Cypherbase

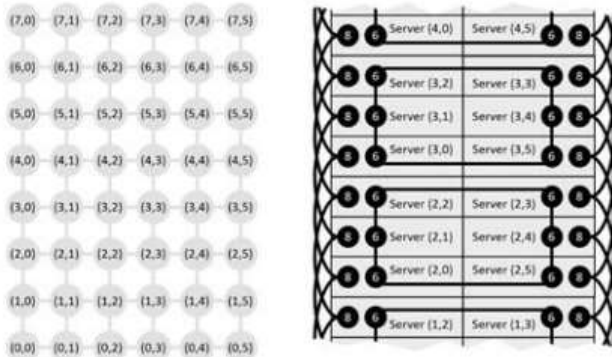
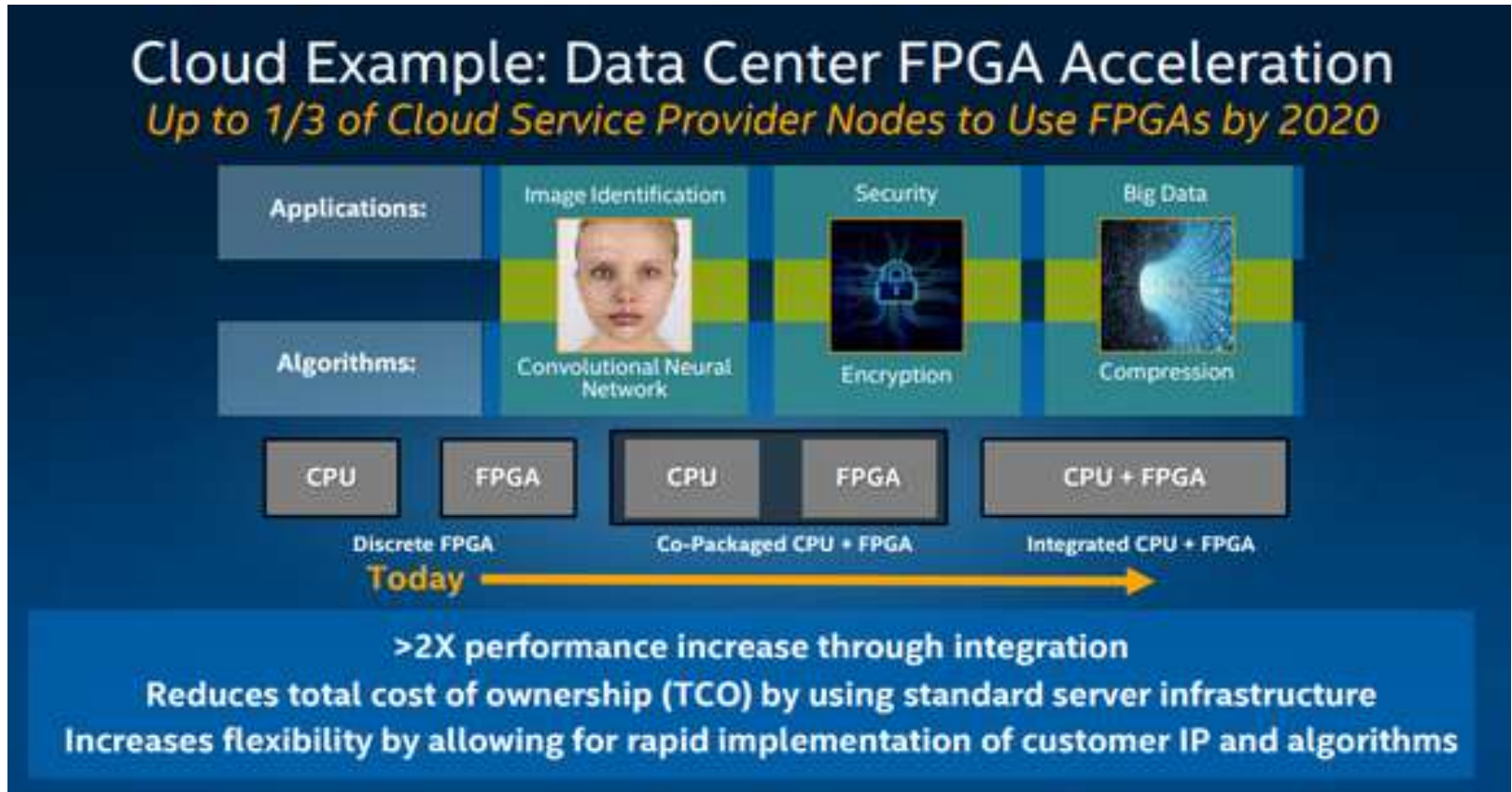


Figure 2: The logical mapping of the torus network, and the physical wiring on a pod of 2 x 24 servers.

# Example 3: both!!



From INTEL

The computer you know  
no longer exists  
(at least, the interesting ones)



# Nobody ever got fired for using Hadoop on a Cluster

A. Rowstron, D. Narayanan, A. Donnelly, G. O'Shea, A. Douglas  
HotCDP 2012, Bern, Switzerland

## Analysis of MapReduce workloads:

Microsoft: median job size < 14 GB

Yahoo: median job size < 12.5 GB

Facebook: 90% of jobs less than 100 GB

Fit in main memory

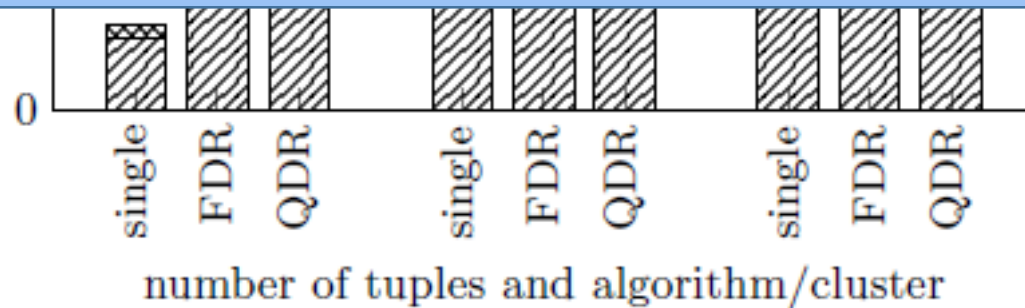
One server more efficient than a cluster

Adding memory to a big server better than using a cluster

# Multicore vs cluster



The predominant architecture will not be one computer (multicore) but a networked set of processing elements, possibly with limited or no cache coherency



# What is a computing node?

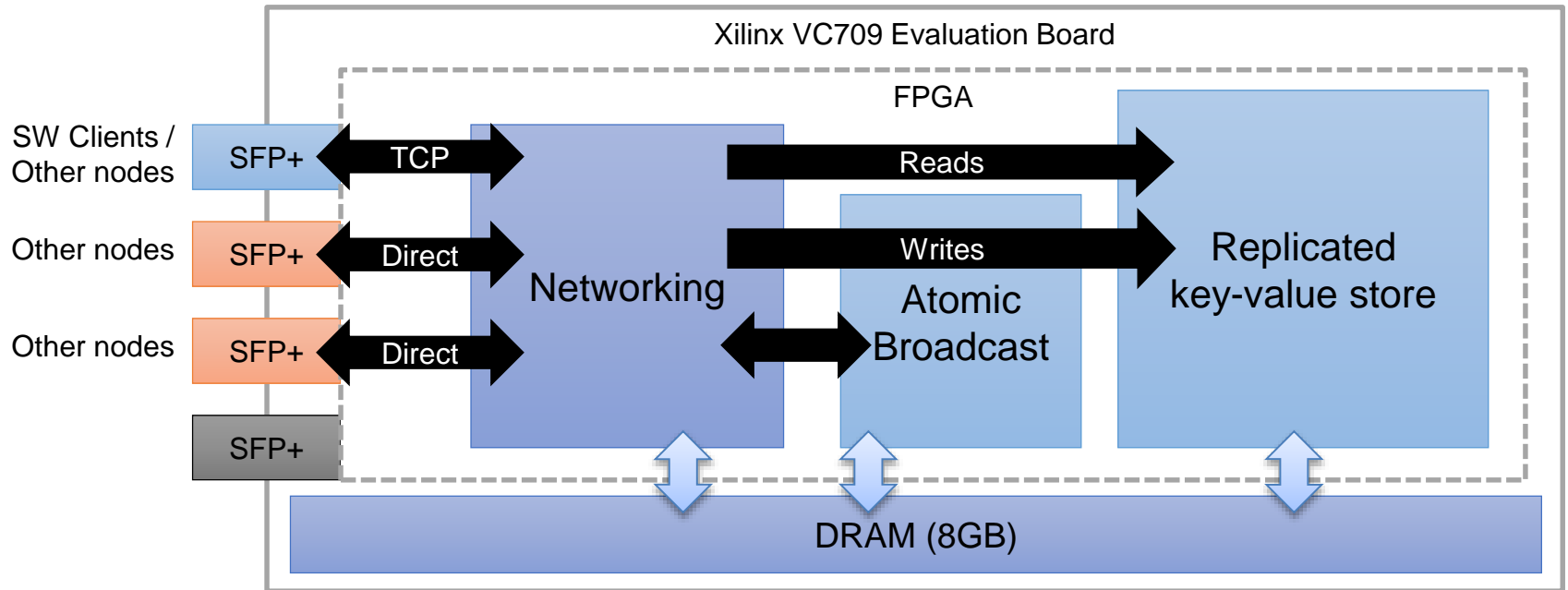


Knights Landing Die (INTEL)  
up to 72 cores  
2 ports Omni-Path  
Bootable host

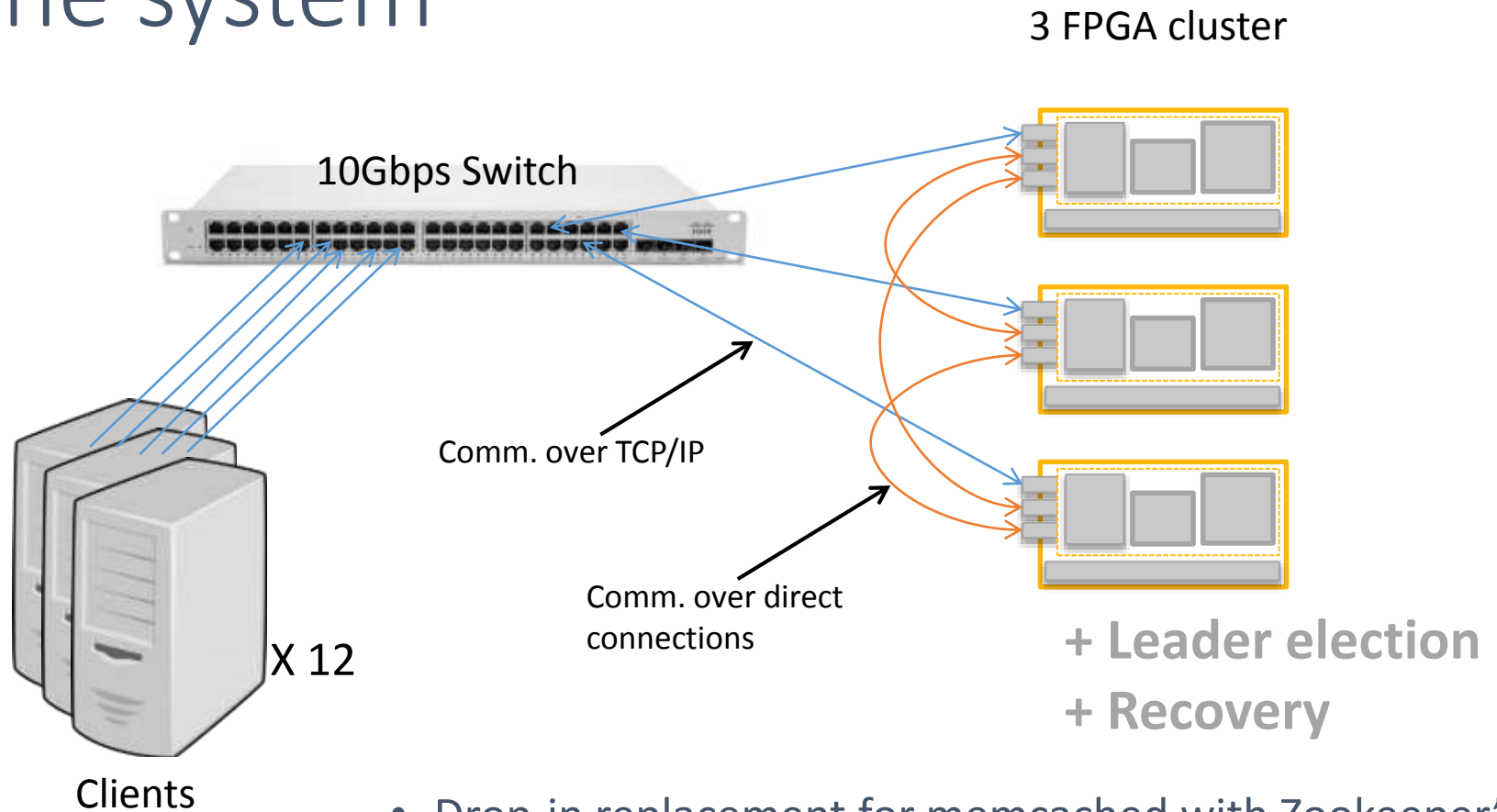
The form factor for a computing node in a data center is about to change

Disaggregation  
(HP's The Machine)

# Consensus in a Box (Istvan et al, NSD'16)

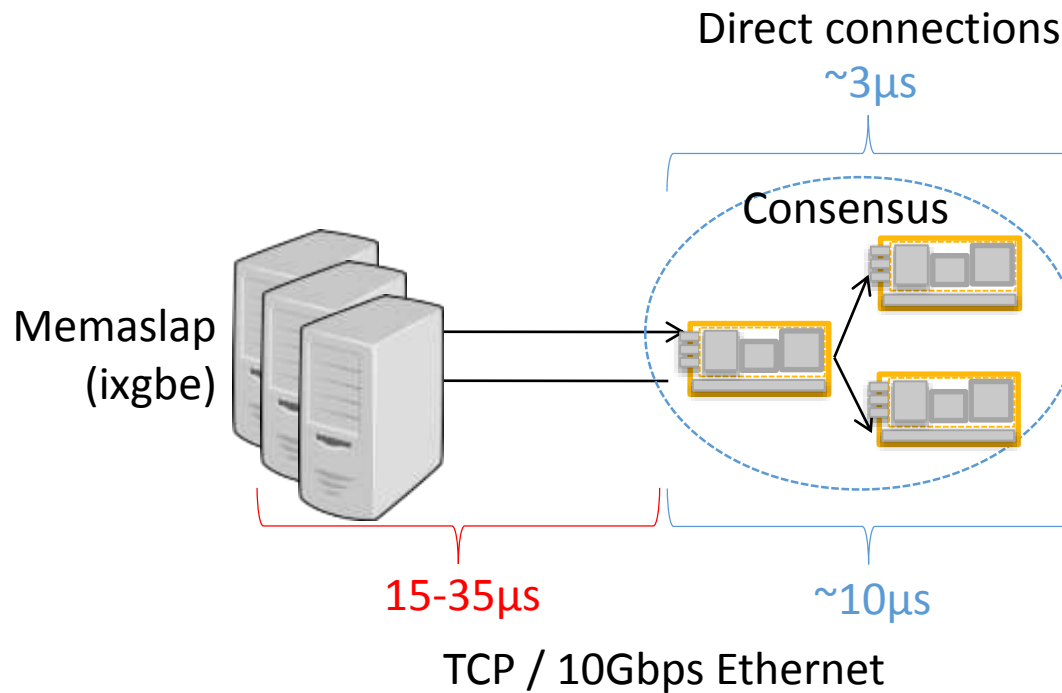


# The system

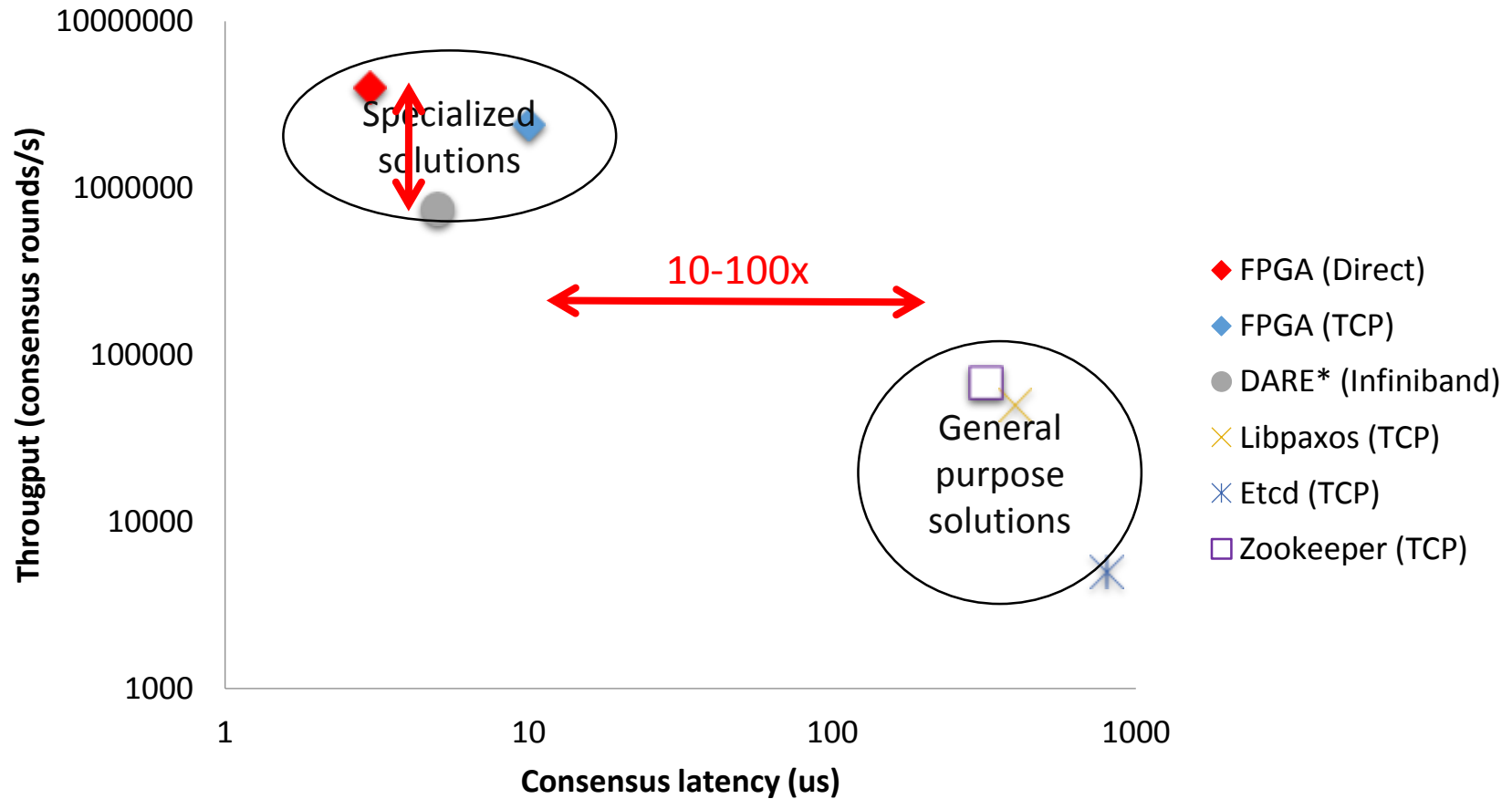


- Drop-in replacement for memcached with Zookeeper's replication
- Standard tools for benchmarking (libmemcached)
  - Simulating 100s of clients

# Latency of puts in a KVS



# The benefit of specialization...



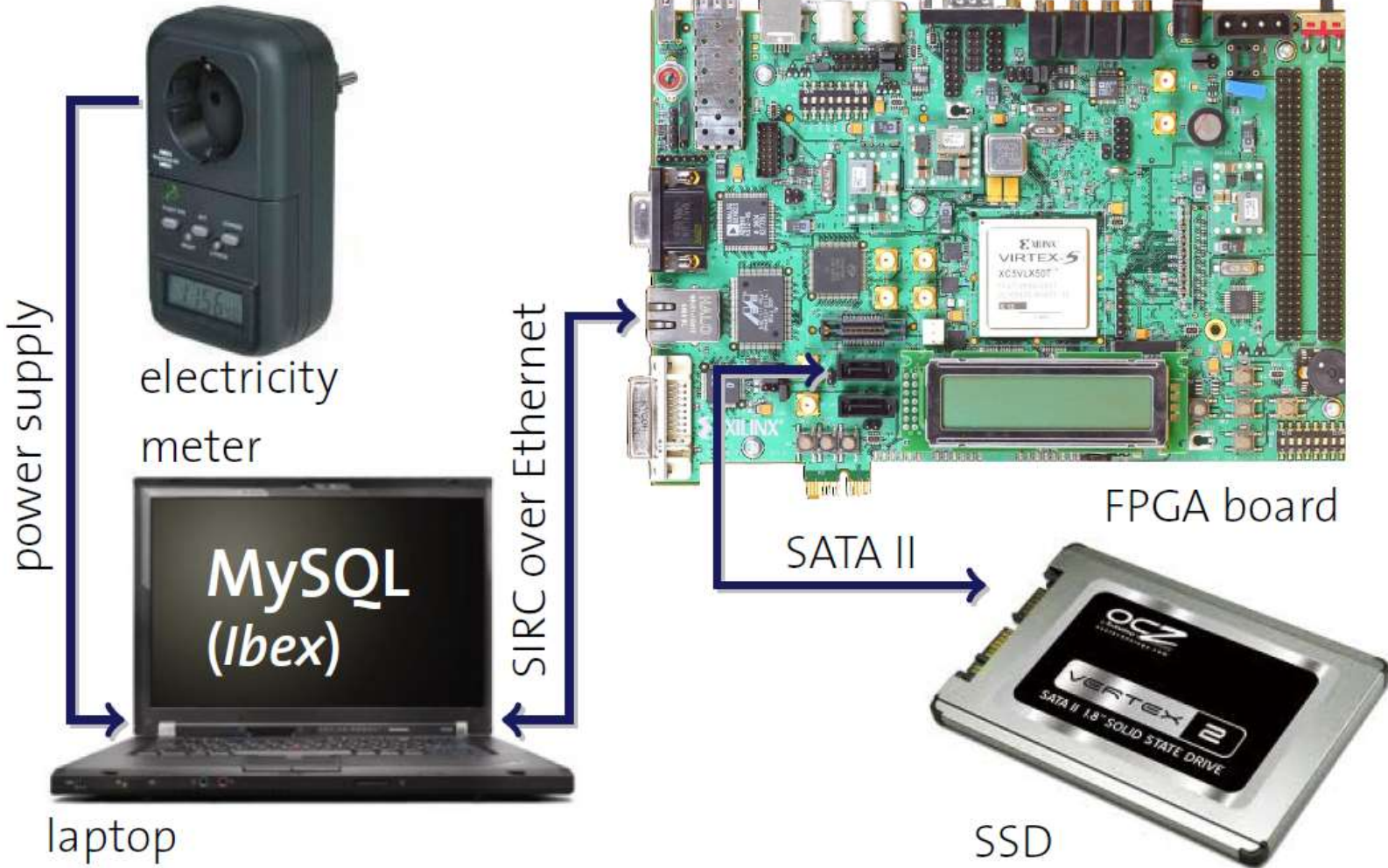
[1] Dragojevic et al. FaRM: Fast Remote Memory. In NSDI'14.

[2] Poke et al. DARE: High-Performance State Machine Replication on RDMA Networks. In HPDC'15.

\*=We extrapolated from the 5 node setup for a 3 node setup.

# Processing everywhere

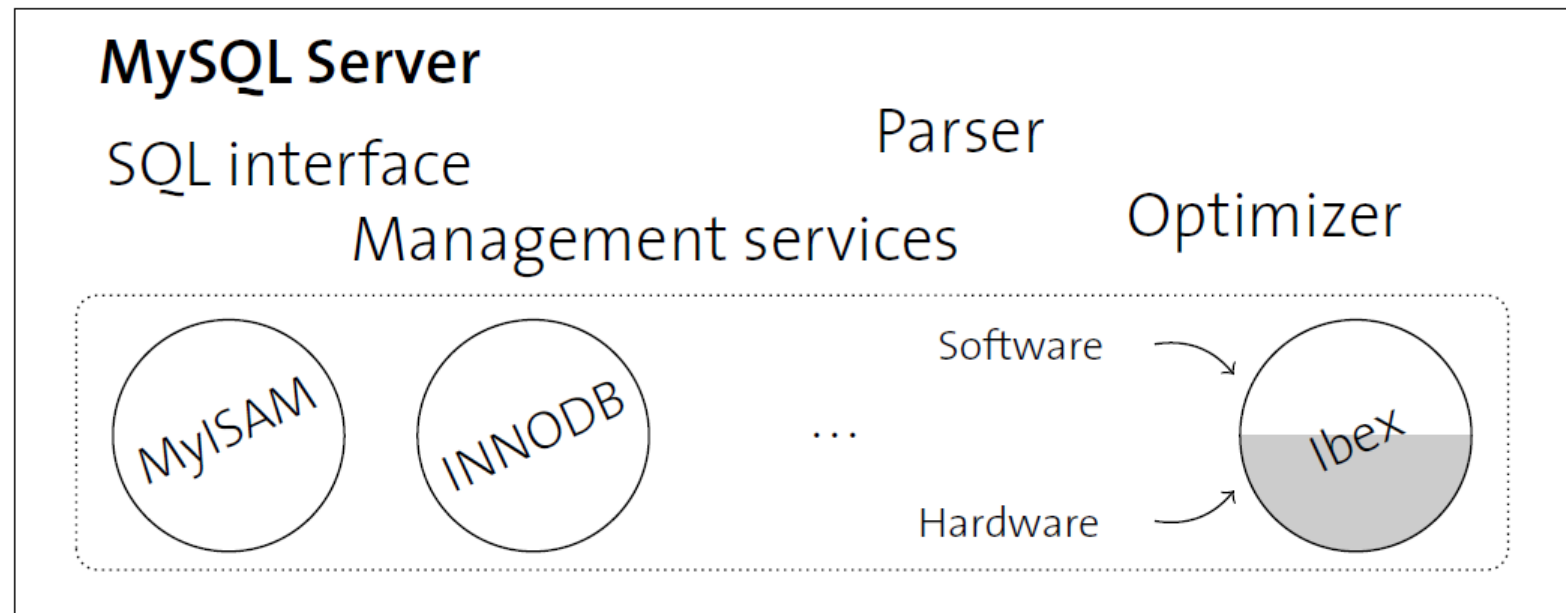
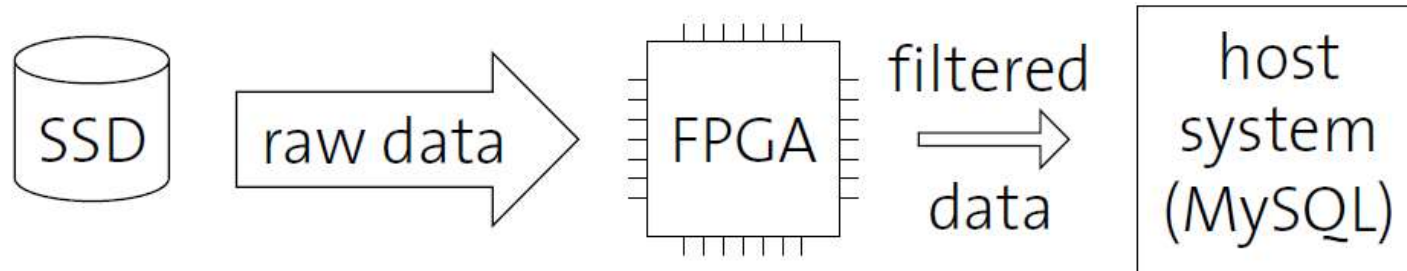




# IBEX

(Woods, VLDB'14;  
Istvan, SIGMOD'14)

# A processor on the data path



# Sounds good?

Imagine the same at all levels:

- Smart storage

- On the network switch (SDN like)

- On the network card (smart NIC)

- On the PCI express bus

- On the memory bus (active memory)

Every element in the system  
(a computer rack)  
will be a processing component

# In a cloud?

These developments will force rethinking many aspects of the cloud

- scheduling

- sharing and virtualization

- platforms for storing and processing data

Ignoring the specialization through hardware not a good idea

- performance / relevance

This is the end ...

# The agenda ahead of us

- Very interesting times
  - Many opportunities driven by hardware
  - Plenty of use cases justifying specialization
- Many challenges
  - Hardware changes affect the whole stack
  - How to program heterogeneous architectures
- A new era in computer science
  - Architecture
  - Data centers
  - Economic pressures and models