

# Caribou: Intelligent Distributed Storage

**Zsolt István, David Sidler, Gustavo Alonso**

Systems Group, Department of Computer Science, ETH Zurich

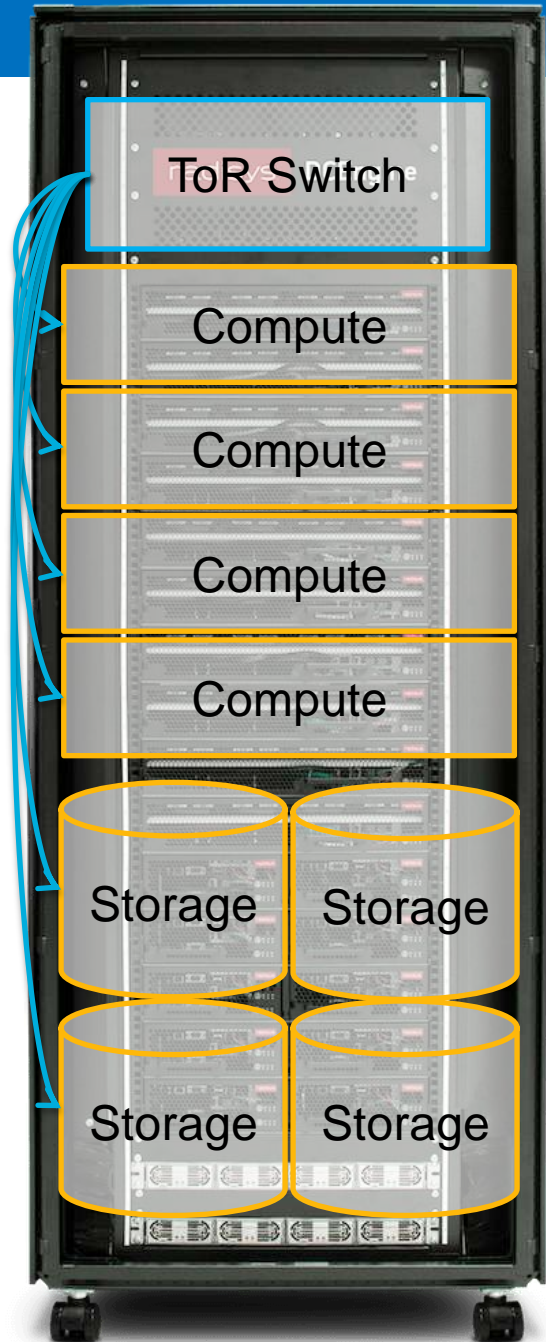
# Rack-scale thinking

In the Cloud



- + Provisioning
- + Independent Scalability
- Data movement bottleneck

in an Appliance



# Storage Design Options

Compute > Bandwidth



Oracle Exadata  
 IBM PureData  
 Deuteronomy  
 ...

- + Full-fledged
- SW+HW overhead
- Large footprint



Compute ~ Bandwidth

Features similar to software  
 Balanced design

Compute < Bandwidth



Samsung YourSQL  
 Winsconsin SmartSSD  
 Kinetic Drives  
 BlueCache  
 ...

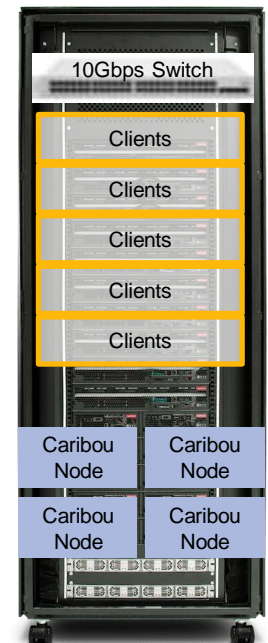
- Outside management
- + No-overhead access
- + Small footprint

# What is Caribou?

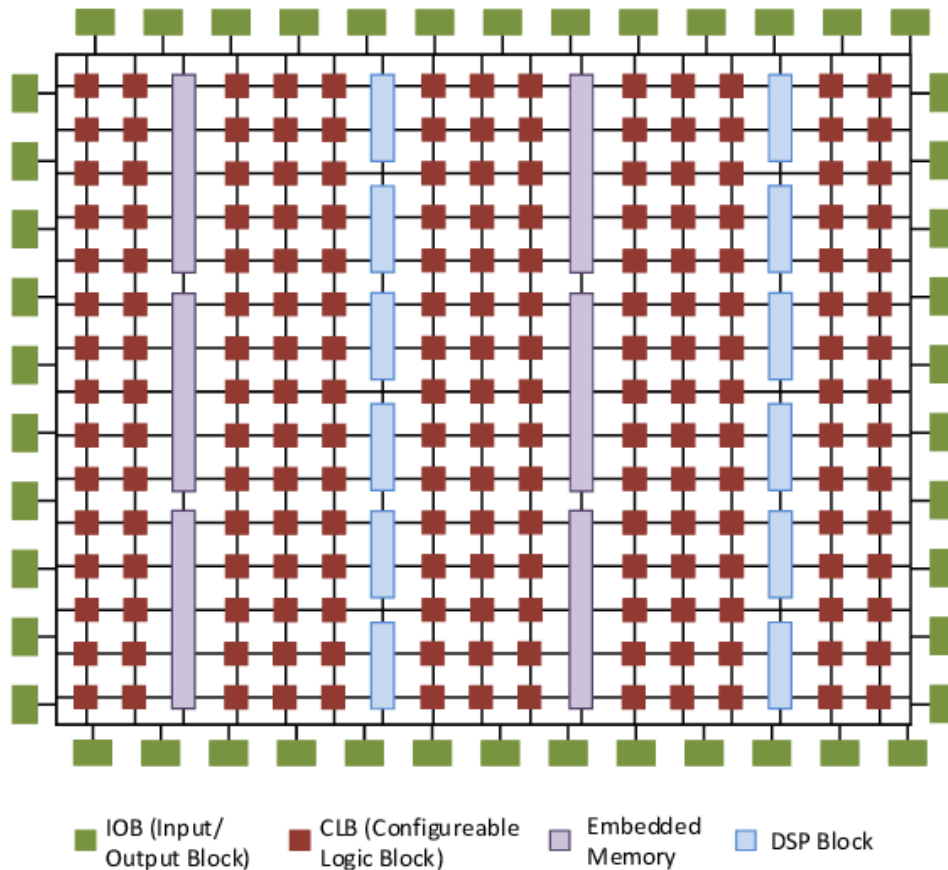
- Intelligent Distributed Storage with FPGAs
- Easy integration on commodity network
- Random access to tuples & in-storage scans
- Selection predicate pushdown
- Data replicated consistently to nodes
- Extensible (open-source) design



fpgasystems



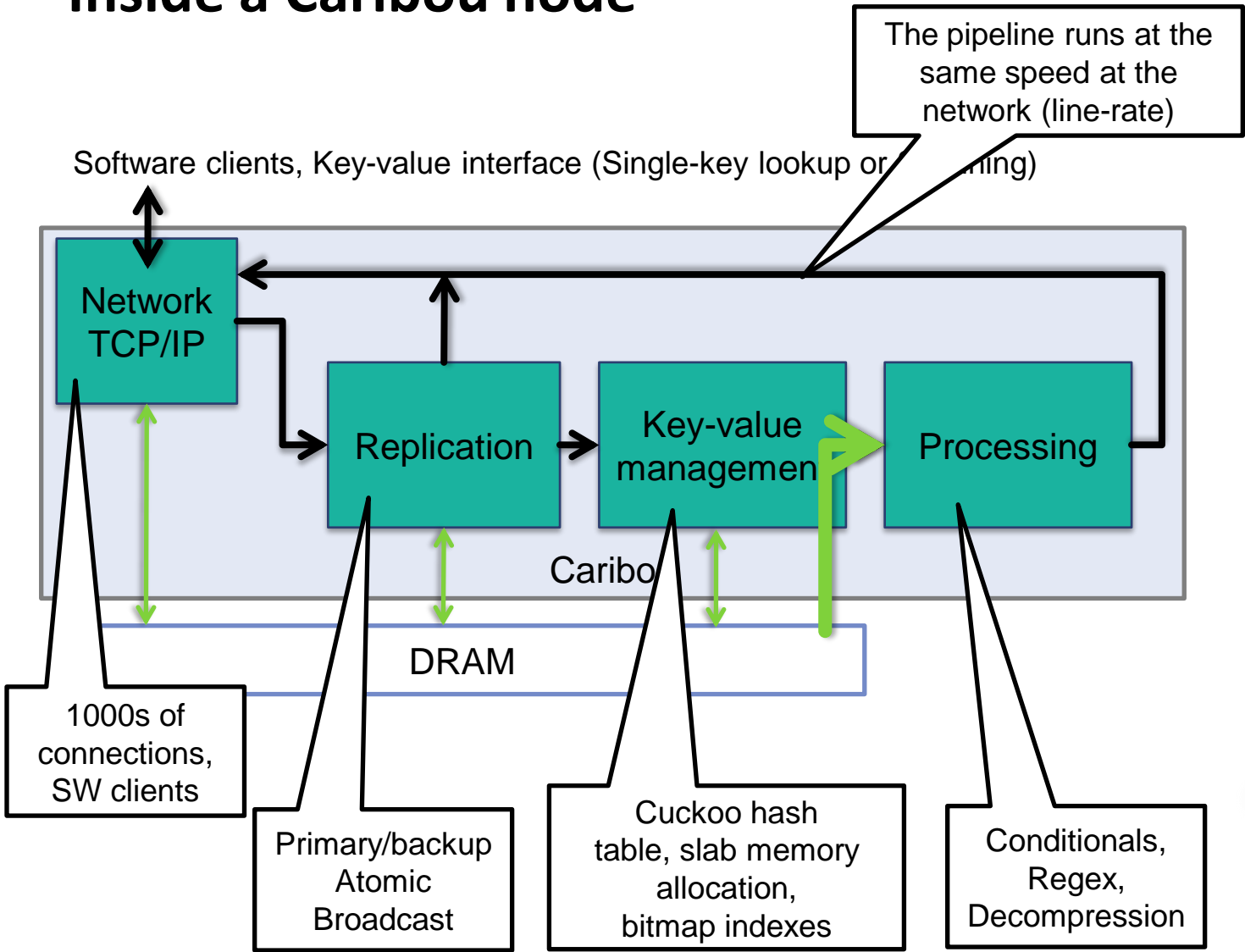
# FPGA 101



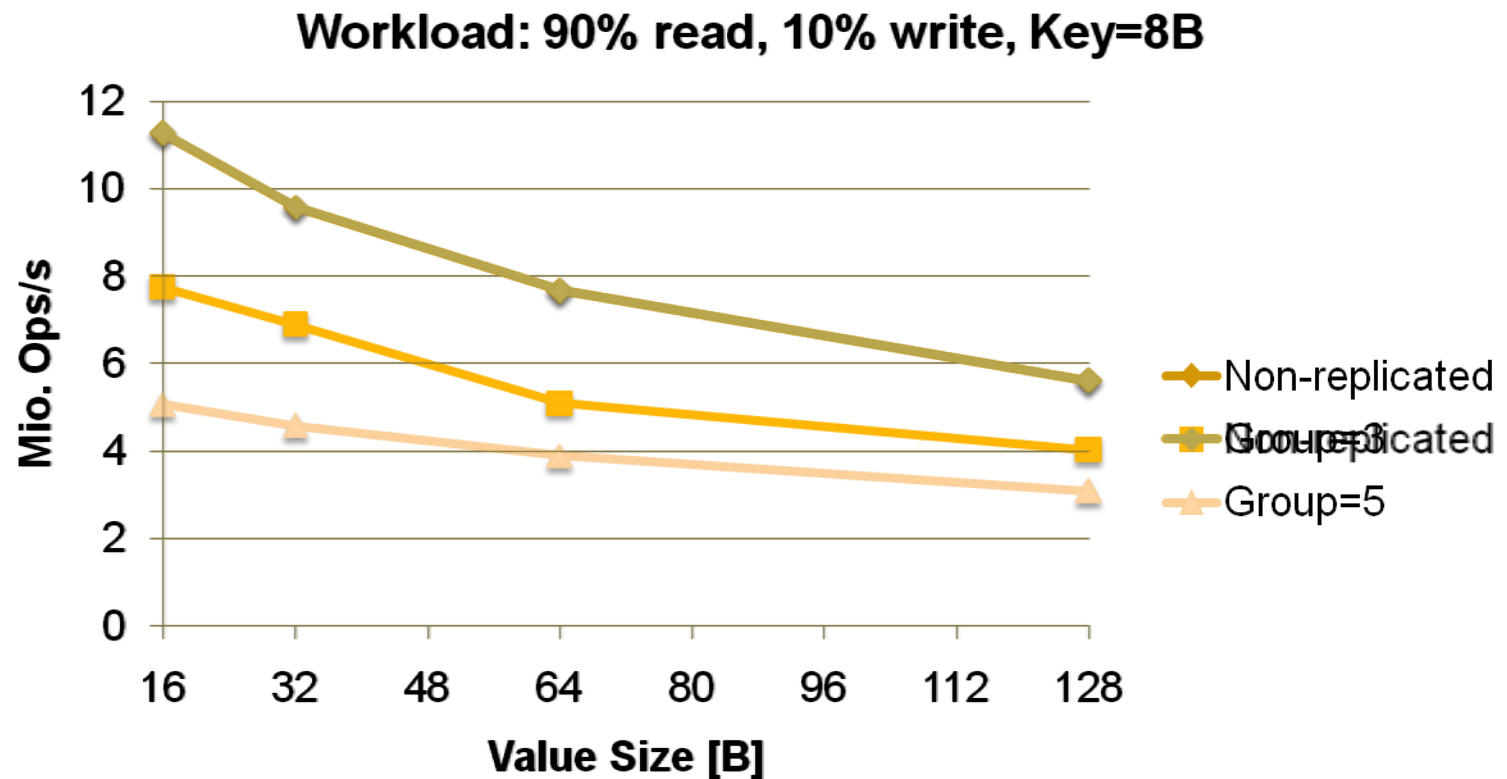
## Field Programmable Gate Array

- Reprogrammable hardware
- Large number of configurable logic blocks
- Tight integration, massive parallelism
- Network/App Co-design
- Innovation...

# Inside a Caribou node

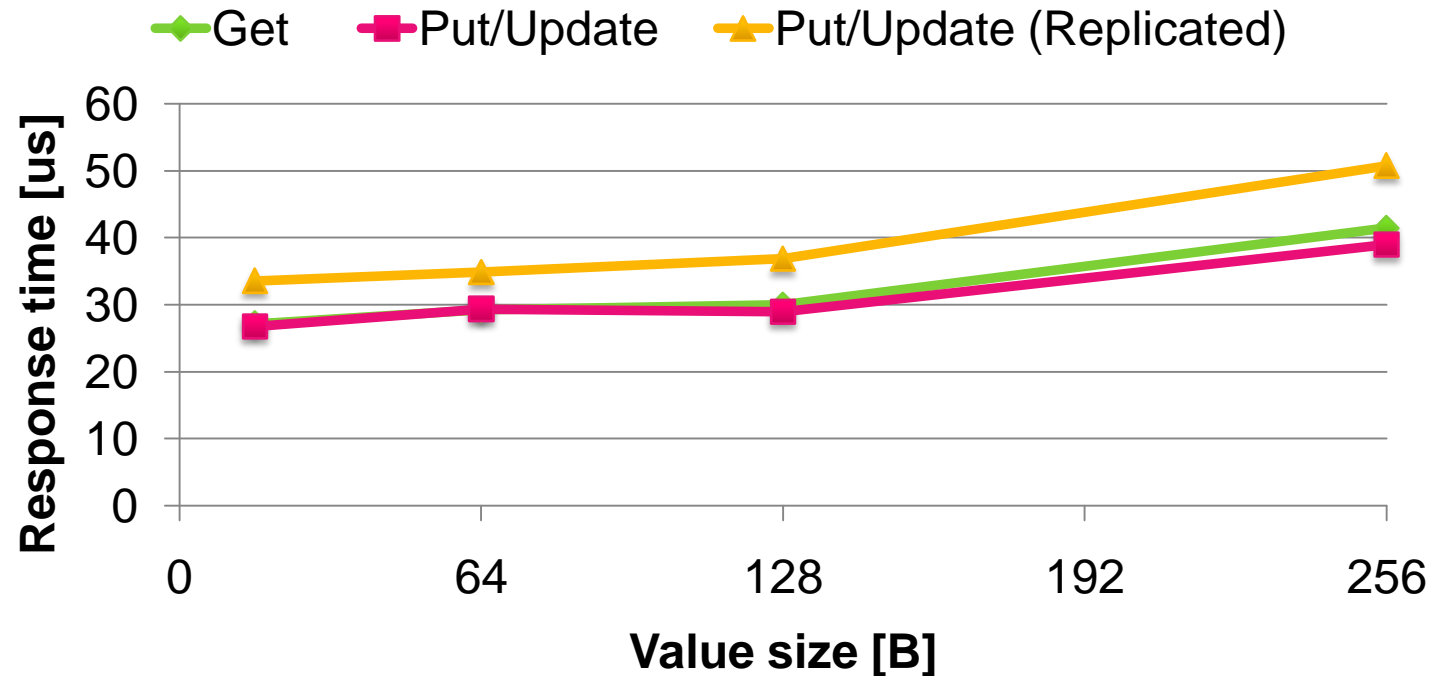


# Throughput of random access to storage



# Random access response times

- Response times comparable to SW on Infiniband, but Caribou uses commodity networking





## Operator push-down

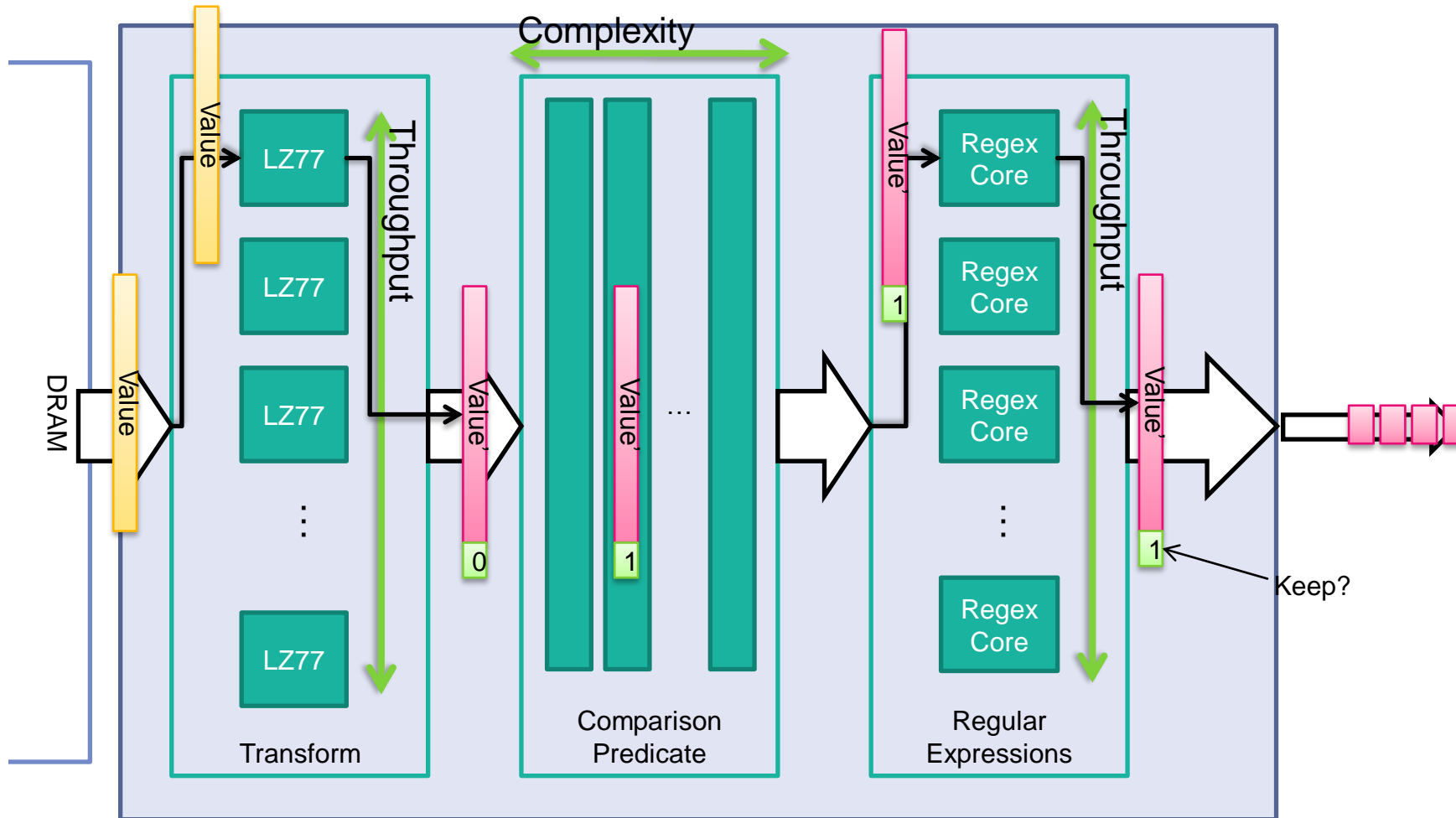
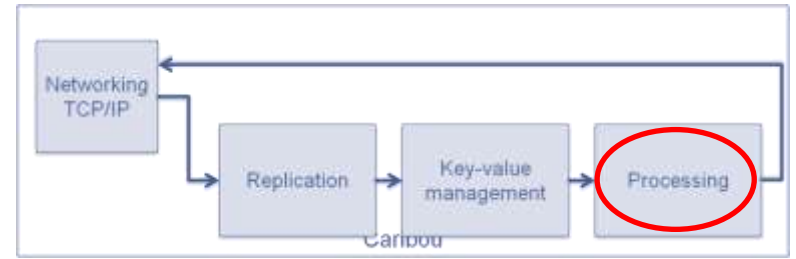
```
SELECT ... FROM customer  
WHERE age<35 AND purchases>2  
AND address LIKE "%Luzern%CH%"
```

The filtering circuits are parameterized at runtime, with no overhead.

- Multiple comparisons to constants (conjunction)
- Substrings or regular expression matching [1]
- Can filter compressed data (LZ77)
- Extensible pipeline design

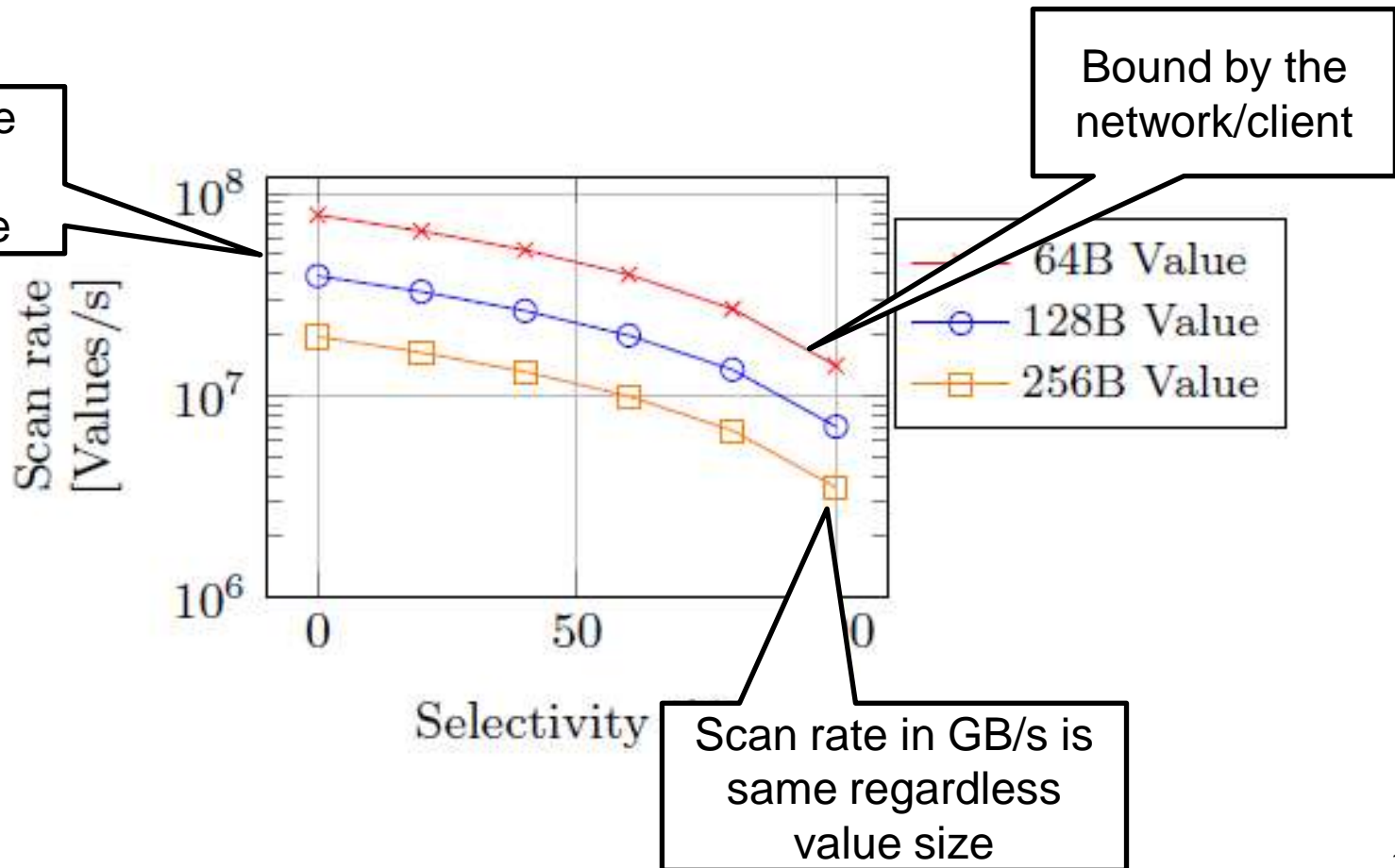
[1] **Accelerating Pattern Matching Queries in Hybrid CPU-FPGA Architectures.** D. Sidler, Zs. Istvan, M. Ewaida, G. Alonso. *2017 ACM SIGMOD/PODS Conference (SIGMOD'17)*

# Exploiting Parallelism



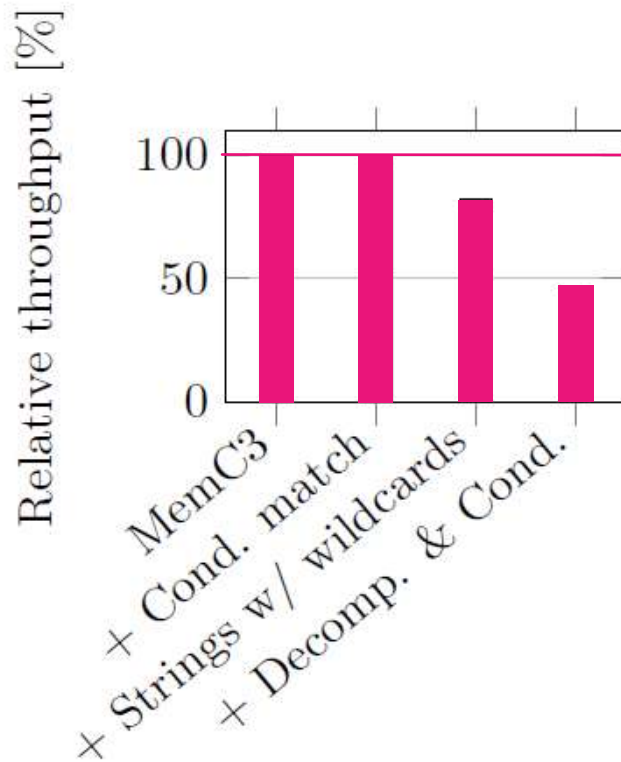
# Scan and filter

- Choice of filter and value size do not impact scan rate.

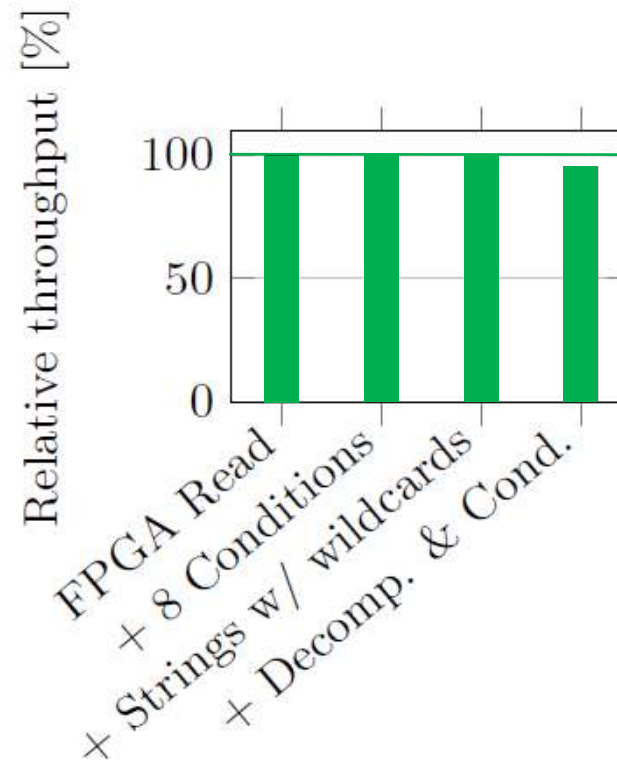


# Near Data Processing without Surprises

- Filtering can be combined with random access reads as well



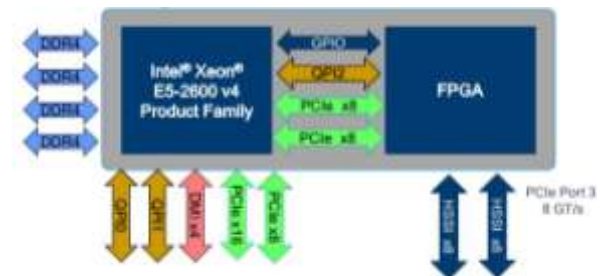
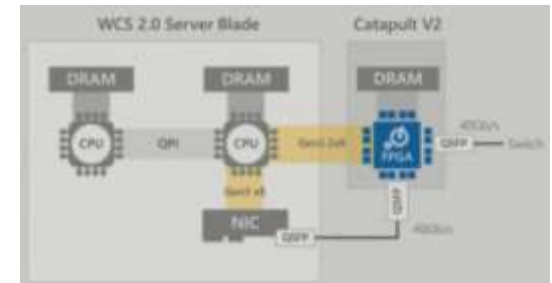
(a) Software



(b) Hardware

# “The Times They Are A-Changin’”

- In-Storage Processing
  - Stand-alone boards, MPSoC (ARM+FPGA)
  - Add NVMe flash, N.V. Memory
  - Explore different KVS (memcached, redis, ...)
  
- In-Network Processing
  - Microsoft Catapult NICs
  - Work on streaming data
  - Distributed service in the cloud
  
- Accelerator
  - Intel Xeon+FPGA
  - Offload computation without partitioning or copying data



# Time to Explore...

- Data movement bottleneck on many levels
- Caribou – Intelligent Distributed Storage
  - Software-like service in a small footprint
  - Balanced design with “right amount” of compute
- Caribou – Platform to Explore Near-data Processing
  - Open source, modular and portable
  - Data processing operators applicable on other HW platforms
  - <https://github.com/fpgasystems/caribou>

