Zsolt István, David Sidler, Gustavo Alonso
Department of Computer Science, ETH Zürich

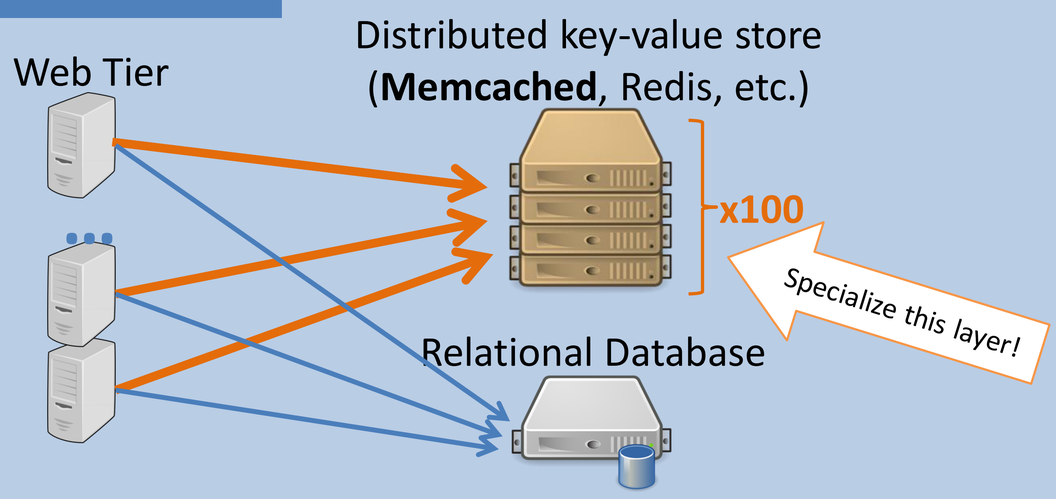
Specialized μservers for the data center

MICROSERVER *n.* An application-specific server implemented as a small appliance

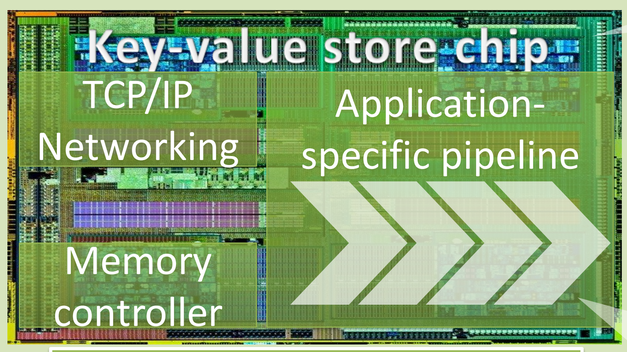
1 Motivation – Optimizing scale-out workloads in the datacenter

Main-memory key-value stores deployed in most datacenters
- Network-heavy: little computation, high concurrency
- Random access: multi-level cache poorly utilized

Modern CPUs and architectures are optimized for the opposite!



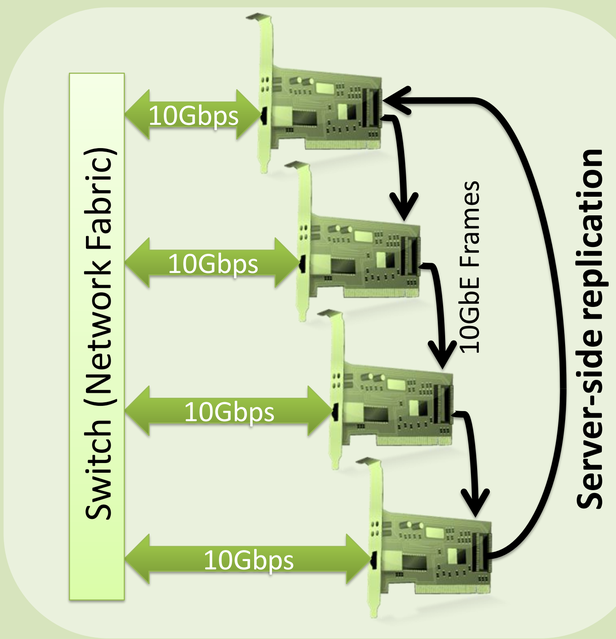
2 Implementation – From single instance to distributed service



We use Xilinx VC709s for prototyping:
4x 10Gbps interfaces,
8GB memory

Distribute & Replicate
Performance & Fault-tolerance

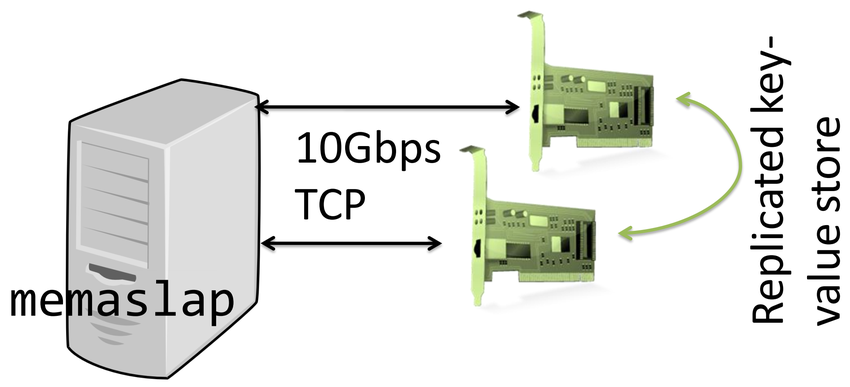
Microserver built from state-of-the-art blocks
[1] Istvan et al. A Flexible Hash Table Design For 10Gbps Key-value Stores on FPGAs. FPL'13
[2] Sidler et al. Scalable 10 Gbps TCP/IP Stack Architecture for Reconfigurable Hardware. FCCM'15



Implementation of the chain-replication protocol over raw Ethernet frames

3 Demo use-cases

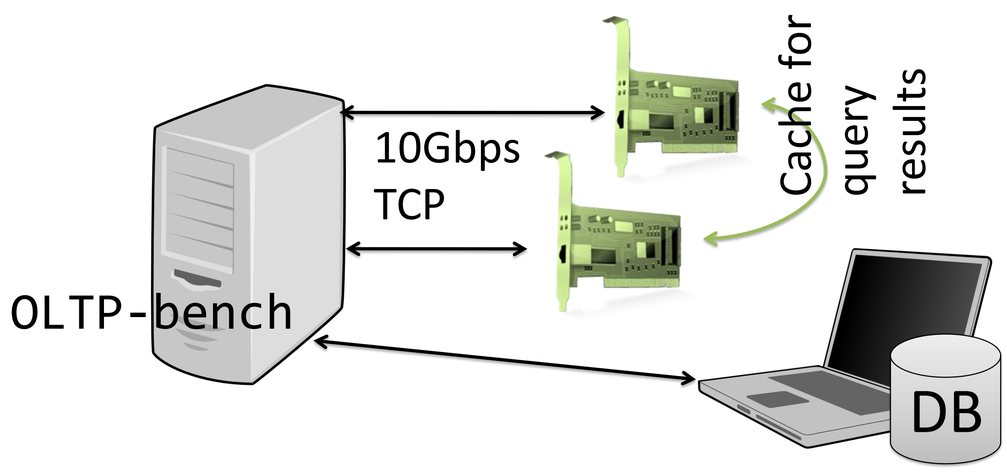
a) Benchmarking with memaslap



Memaslap is a C++ benchmarking client for memcached.
It offers many paramters for customized load generation:
- Fraction of *set* and *get* commands (write vs. read)
- Different key and value sizes
- Number of clients simulated on the machine

We demonstrate the functionality of our key-value store, and measure both throughput and round trip times. A single machine is not able to saturate the microservers (see results below).

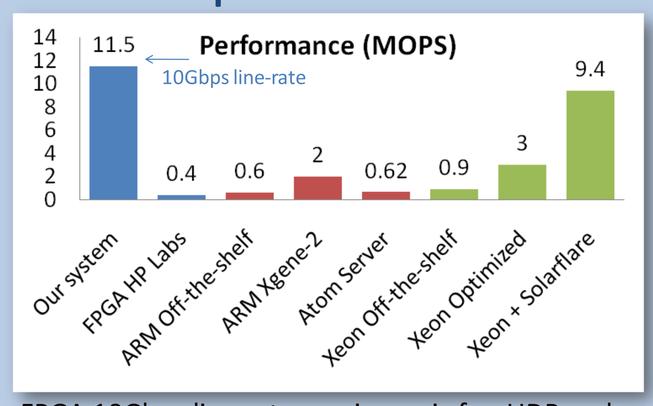
b) Integration with a database

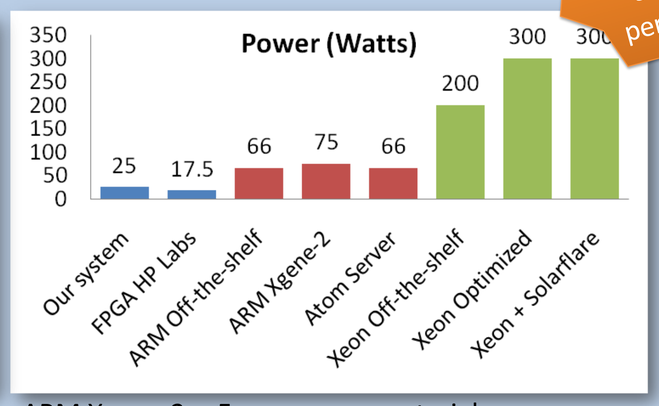


OLTP-bench is a collection of database benchmarks written in Java.
We use the *Twitter* benchmark to demonstrate the benefits of caching results in main memory → increased throughput.

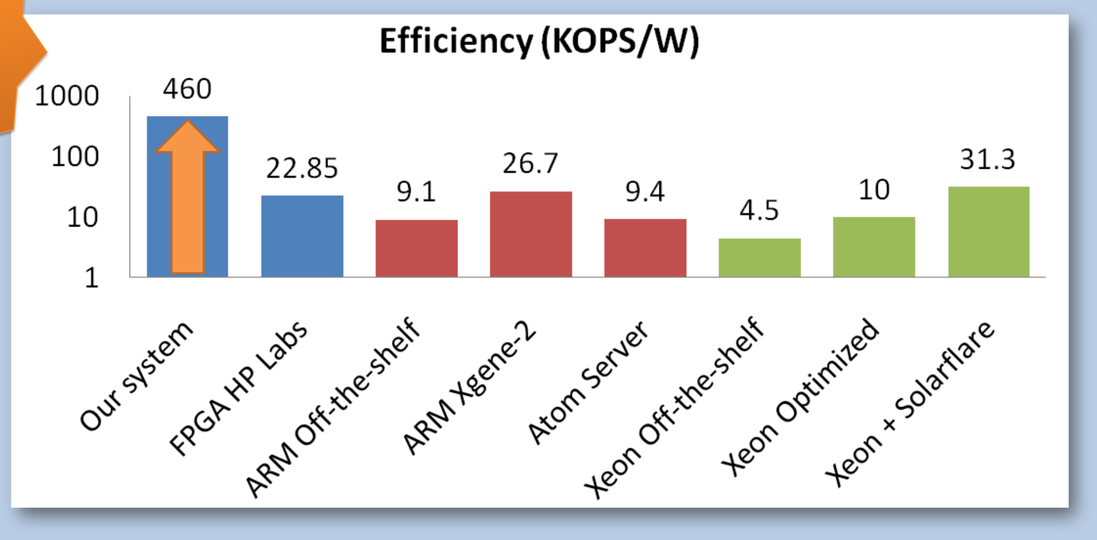
```
procedure get-tweets-of-user(userID)
    get(userID) from microserver
    if not found then
        run query on DB
        set(userID, tweets) on microserver
```

4 Performance of a node compared to related work





Highest efficiency, without sacrificing performance



FPGA 10Gbps line-rate maximum is for UDP and 16B key and 16B value, a common ground of other related work.

Related work & systems:
- FPGA HP Labs – Appliance optimized for 1Gbps networking (FPGA'13)
- ARM Off-the-shelf – 8x ARMV8 @2.4GHz
- ARM Xgene-2 – From promo material
- Atom server – 8x Atom C2750 @2GHz (Intel promo)
- Xeon Off-the-shelf – 8x Xeon E5 @2.4GHz
- Xeon Optimized – 16x Xeon E5 @2.2GHz (Intel whitepaper on optimizing memcached)
- Xeon + Solarflare – 20x Xeon E5 @2.2GHz (Solarflare whitepaper on optimizing memcached for its NICs)

Acknowledgements: This work is funded in part through the Microsoft Joint Research Center MSR-ETHZ-EPFL and an equipment donation from Xilinx.