doppioDB – A Hardware Accelerated Database

David Sidler, Muhsen Owaida, Zsolt István, Kaan Kara, Gustavo Alonso
Systems Group, Department of Computer Science, ETH Zürich

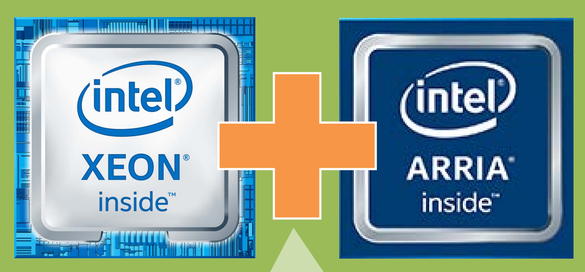
Accelerating novel workloads using hybrid architectures

Novel workloads:

- Analyzing unstructured and user-generated text data
- Gain insights from data stored in databases through machine learning algorithms

Analytical operators incur significant compute load

Intel Xeon+FPGA



CPU and FPGA in same package

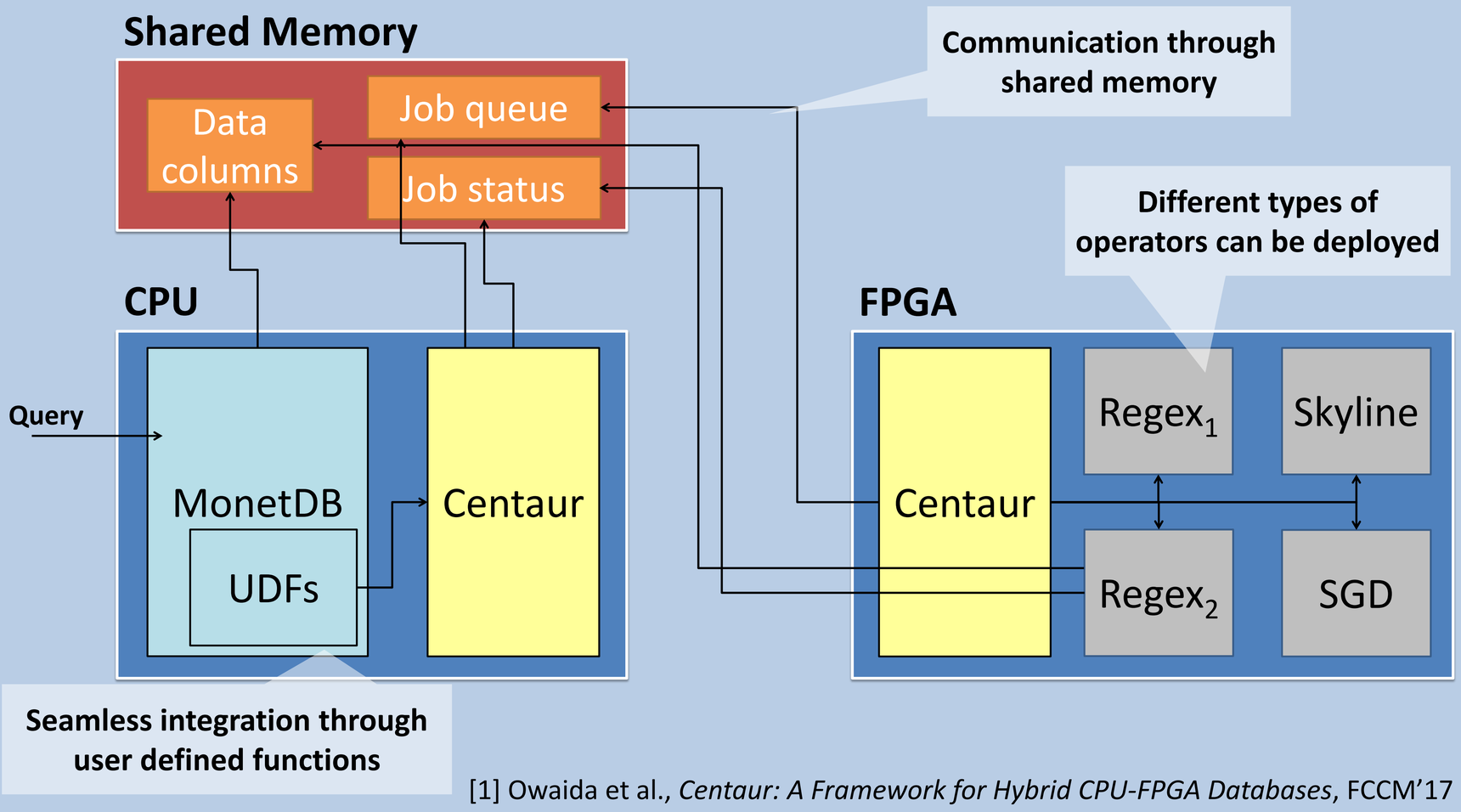
Hybrid CPU-FPGA platforms:

- FPGA has coherent access to main memory => critical for databases
- No need for data partitioning, copying or reformatting
- FPGA is a **specialized** core

Hybrid database

Integration:

- Database can create and monitor jobs on the FPGA through Centaur[1]
- Operators on the FPGA are represented as hardware threads
- Concurrent execution of hardware operators

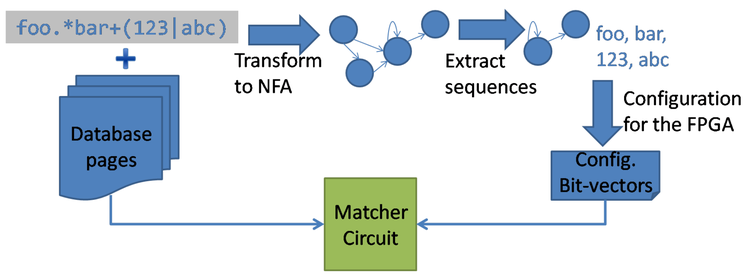


[1] Owaida et al., *Centaur: A Framework for Hybrid CPU-FPGA Databases*, FCCM'17

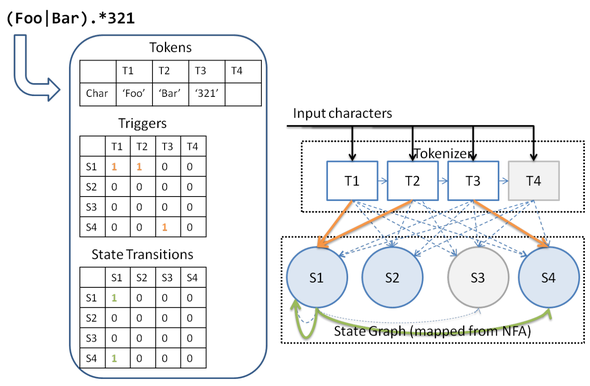
Hardware operators

Compute Intensive

Regular Expression [2]

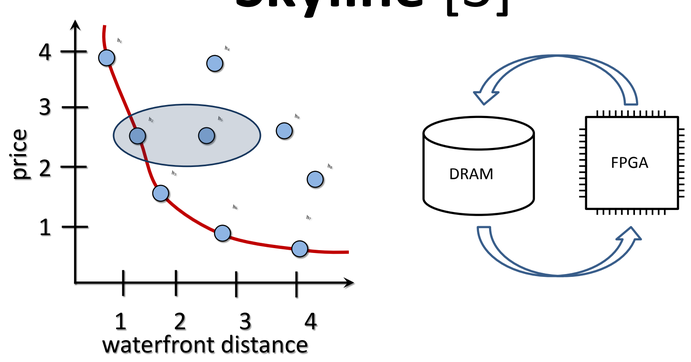


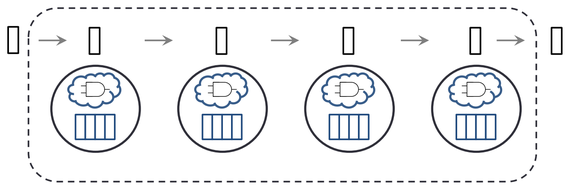
- NFA skeleton parameterized at run-time
- Expression translated into config. vector
- High throughput by deploying parallel NFAs



Deep Pipelining

Skyline [3]

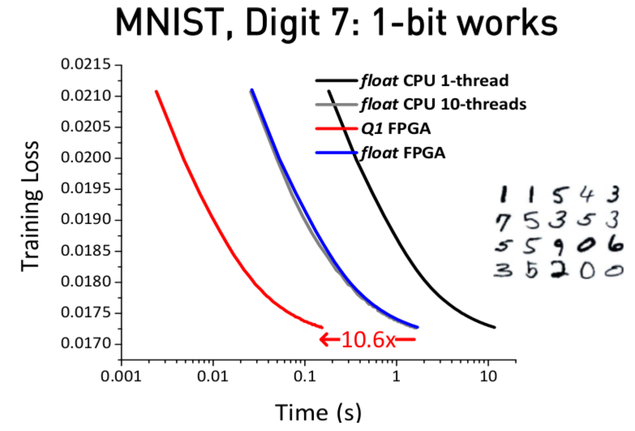




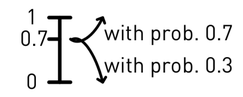
- Keeps candidate set in on-chip pipeline composed of memory cells and comparison logic
- Depending on result size performs multiple iterations

Custom Precision

SGD [4]



- Works on compressed data (Probabilistic rounding to <32 bits)



=> More computation per data moved

- Exploits MIMD parallelism
- Implements custom data types

[2] Sidler et al., *Accelerating Pattern Matching Queries in Hybrid CPU-FPGA Architectures*, SIGMOD'17
[3] Woods et al., *Parallel Computation of Skyline Queries*, FCCM'13
[4] Kara et al., *FPGA accelerated Dense Linear Machine Learning: A Precision-Convergence Trade-off*, FCCM'17

Acknowledgements: We would like to thank Intel for their generous donation of the HARP prototype. This work is funded in part through the Microsoft Joint Research Center MSR-ETHZ-EPFL.

systems.ethz.ch/fpga
github.com/fpgasystems

Open Source