XDF XILINX DEVELOPER FORUM

Accelerating databases with FPGAs

Presented By

rENIAC

Prasanna Sukumar
Head of Engineering
December 10, 2018

XILINX

Our DNA – Hardware, Software, Systems & Data

Management

Prasanna Sundararajan, CEO & Founder
Ex-Computing Market Architect, Xilinx

Chidamber Kulkarni, CTO & Founder
Ex-Staff Engineer, Network Solutions, Xilinx

Thomas Jorgensen, VP, Customer Success
Co-founder & ex-VP, Napatech (IPO)

Prasanna Sukumar, Head of Engineering
Ex-Technical Director, Riverbed

Nishant Vyas, VP, Products
First DB Engineer @ LinkedIn; Former CTO & Head of Product @ MariaDB.

Investors

intel Capital
Menlo
BULLPEN CAPITAL
Divergent
First Star Ventures

Key team members past companies

XILINX
amazon web services
LinkedIn
IBM
SAMSUNG
napatech
riverbed
MariaDB
MARVELL

THE PROBLEM

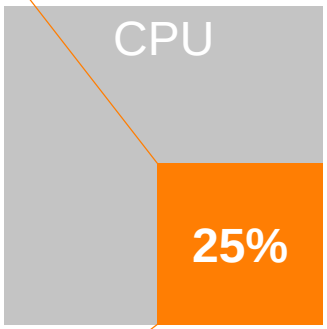
Today's performance requirements are inhibited by CPU bottlenecks

REQUIREMENTS

Low Latency
Predictable low-latency SLAs at scale

High Throughput
Handle billions of transactions

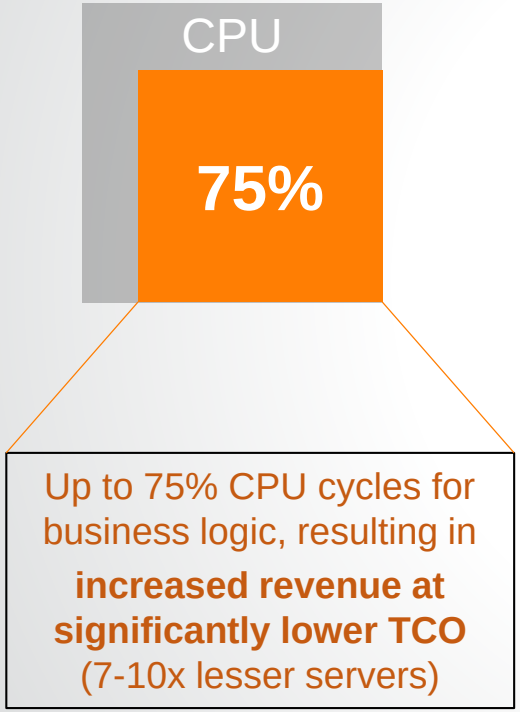
Scale
Serve massive data & transactions while limiting cost



When running an Open Source DB on a standard CPU-based system, ~75% CPU cycles are spent on system compute & I/O

Leaving only 25% devoted to business logic. Not enough to meet increasingly complex requirements of Distributed Databases and AI systems

rENIAC Software solves system + I/O bottlenecks



rENIAC

Distributed Data Engine

- Decouples data and application layers, simultaneously acting as an I/O accelerator to resolve any bottlenecks
- Unique ability to accelerate AI inference algorithms close to the data store and speeding up analytics
- Tightly couples storage class memory to a low latency network stack
- Up to 30x increase in performance
- Leverages COTS servers/CPU + FPGA + SSD
- Deployed as a network service with **no software change required**

rENIAC Solution using FPGA

- Move Compute & IO from SW (CPU) to FPGA,
  - TCP Engine
  - Cluster, Control & Consensus
  - Storage Engine

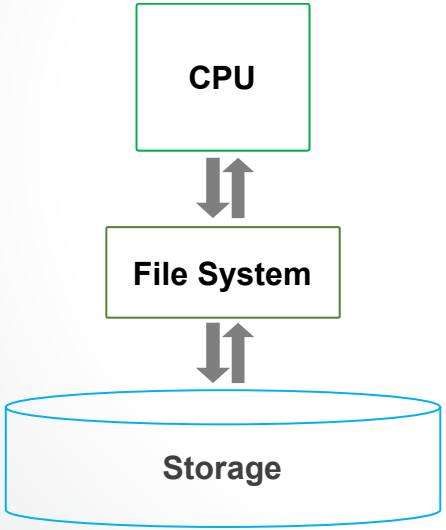


Figure 1. DB on CPU with Traditional IO

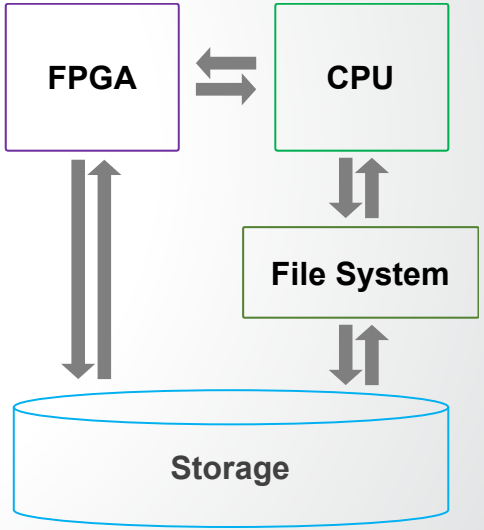
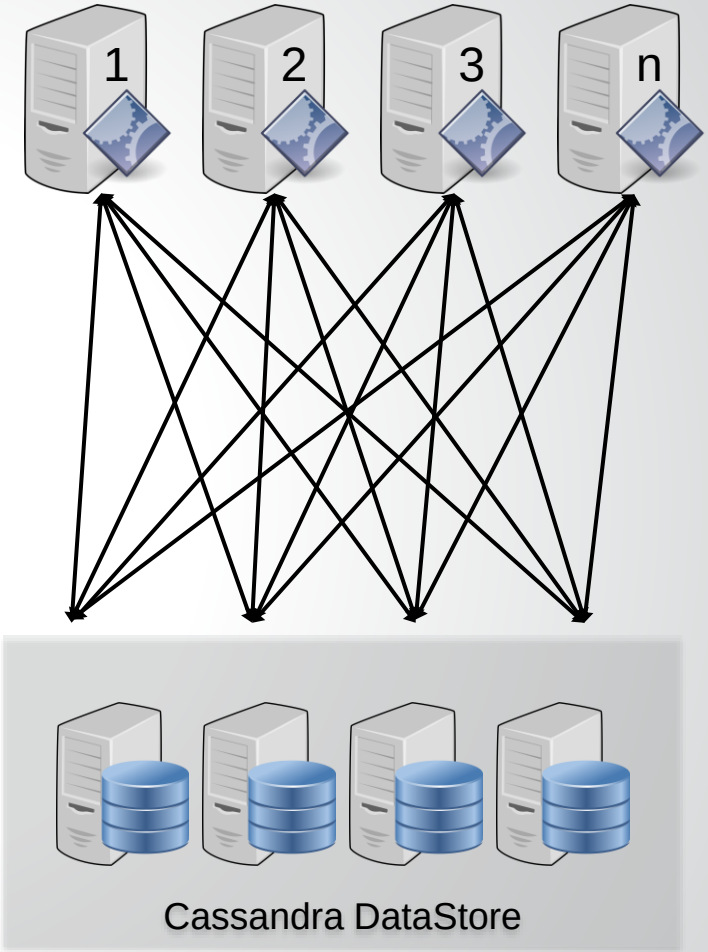


Figure 2. DB using rENIAC SE on FPGA

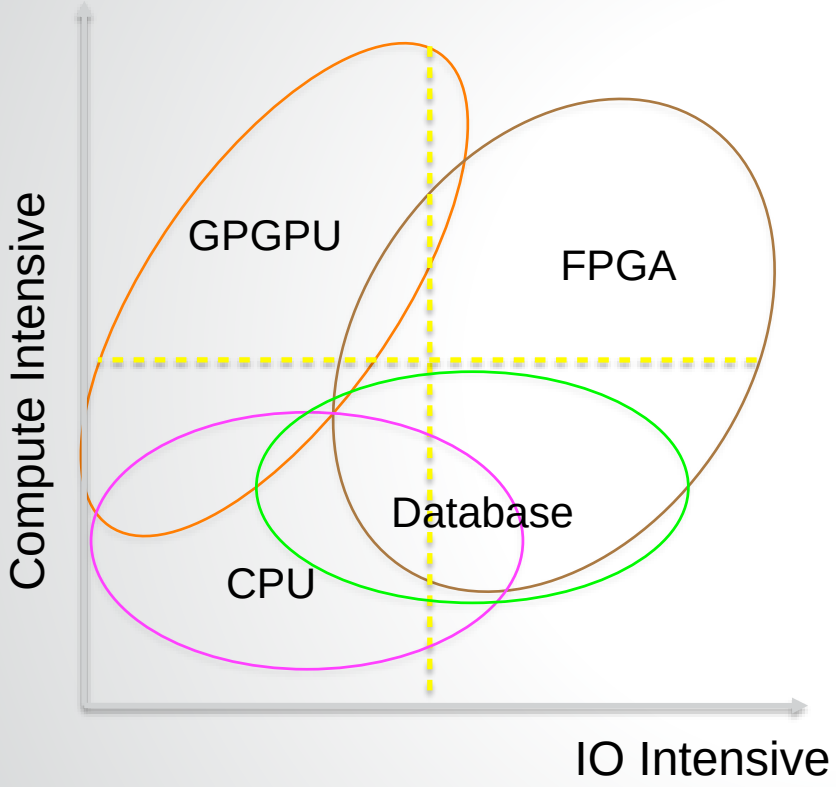
Deployment and Architecture challenges in C*

- Being a write-optimized NoSQL store (using LSMT) Cassandra has to manage memtables and SStables, forcing further IO amplification/bottlenecks by
  - compaction
  - repairs and
  - JVM garbage collection
- Additionally, multi-threaded SW is not scaling well with multi-core HWs.

This impacts 99th percentile,
Tput per node and TCO

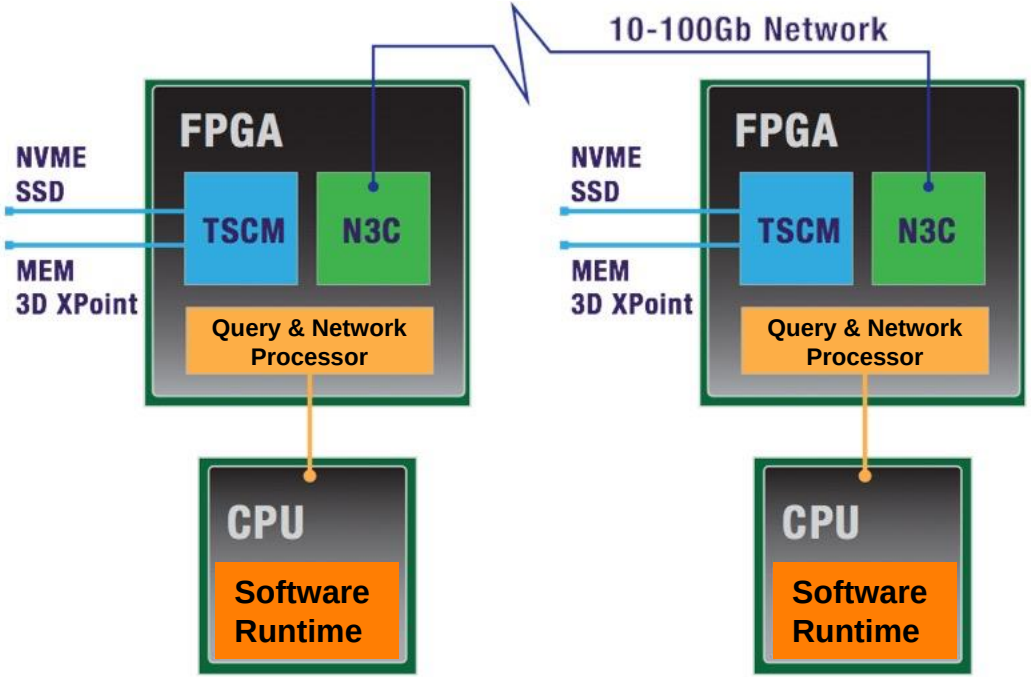


FPGAs for Data Acceleration



- Key parameters for accelerator choice:
  - Compute intensive
  - IO intensive
    - Network
    - Storage
  - Cost (f($, power/TCO))
- Databases are:
  - IO intensive
  - Moderately compute intensive
    - Exceptions like, read repairs, compression, encryption, OLAP queries, etc

rENIAC Core Technology



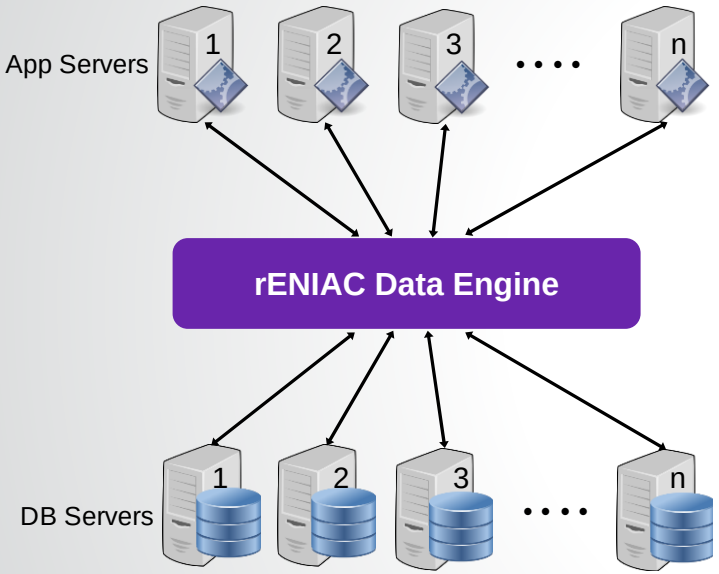
TSCM — **Persistent Data Engine based on** **T**iered **S**torage **C**lass **M**emory

Query Processor — **Network Attached** Low-Latency Transaction Processing

N3C — **Inter-node Routing Engine** **C**onsensus, **C**oordination and **C**ontrol

Software Runtime — **API Support** for any data store (NoSQL, SQL, Graph) and Search

rENIAC DB Acceleration Engine

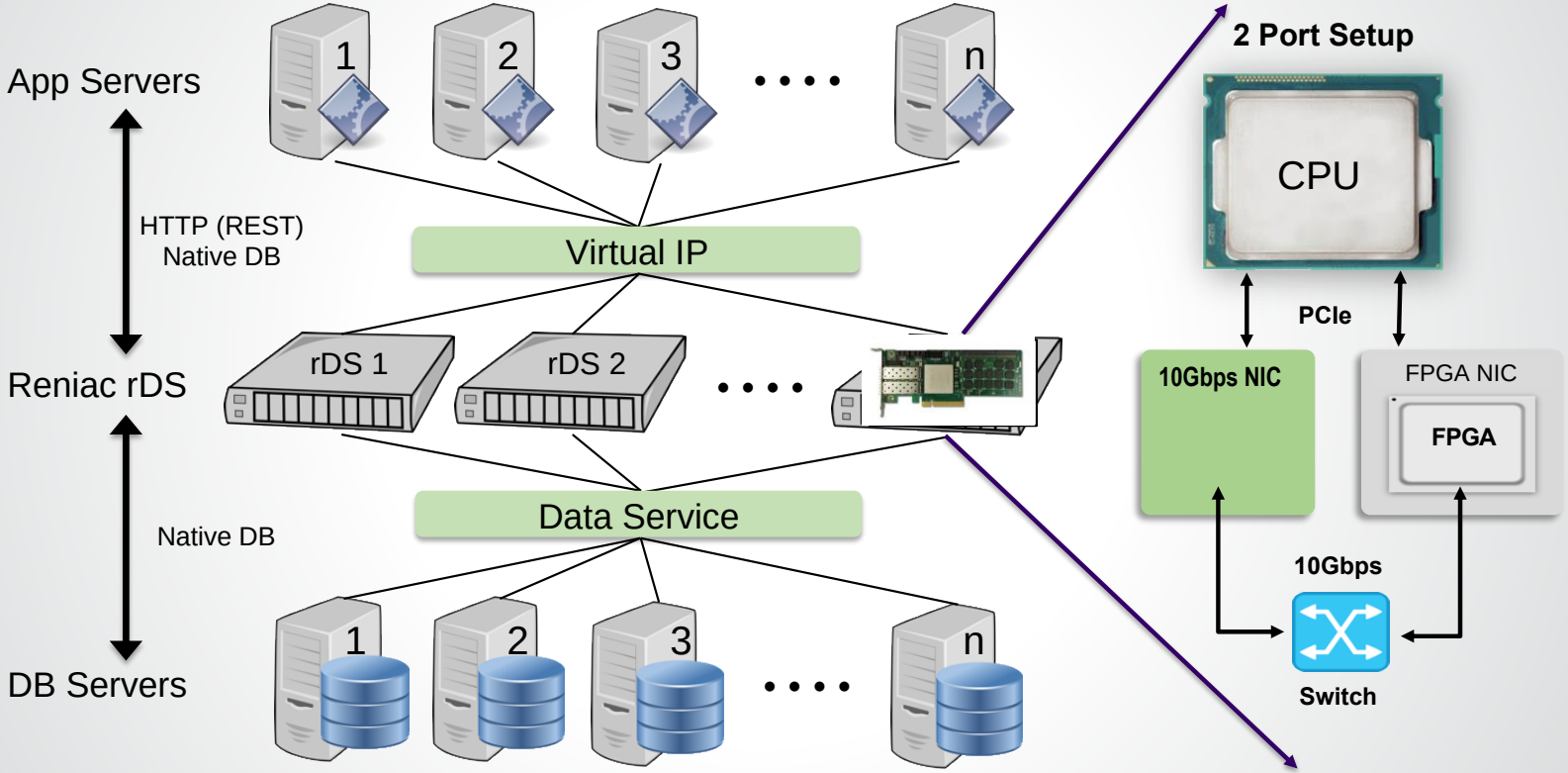


Advantages of rENIAC Engine with Open-Source Apache Cassandra

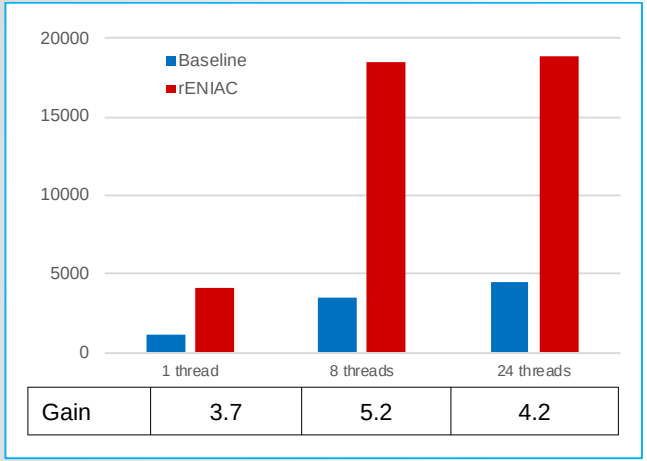
- Read scaling by facilitating extremely high read throughput per node.
- Predictable lower latency per read transaction (up to 99.99th percentile)
- Reduce compaction/garbage collection/thread concurrency inefficiencies

Throughput	4-10x increase
Latency	1/3-1/10 lower latency
Deployment	Plug & Play: No SW changes
Technology	Leveraging state of the art technology: FPGA, CPU, Memory and SSD

Deployment and Architecture



rENIAC Data Engine Benchmark Results



Gain	3.7	5.2	4.2

Latency (ms)	Baseline	rENIAC (rDS)	Perf Gain
95p latency	26.8	1.5	**18x**
99p latency	55.1	1.9	**29x**
99.9p latency	123.3	5	**25x**

Data Engine delivers up to 29x lower latency, over 5x Tput

- **Headroom for up to additional 12x Tput**

	rENIAC Data Engine - host server	Cassandra Client/Server
Processor	Intel Xeon 16C/32T	Intel Xeon 16C/32T
Memory	64GB DDR3	64-128GB RDIMM, 2666MT/s, Dual Rank
Hard drive/Boot	220GB SATA SSD	500GB-1TB SSD SATA/NVMe
NVME/Storage	1TB SSD NVMe	1TB SSD (DB Server)
OS and kernel	CentOS 7.3, Kernel 3.10	CentOS 7.3, Kernel 3.10
Software	rENIAC FPGA Data Engine & SW Connectors	Apache Cassandra v3.10 or later

Workload

- Read-only
- Partitions 5M
- Num trans 100K
- Dist Uniform
- Data size 4KB
- Cassandra v3.10

Roadmap

- Current design is running on a Virtex7 FPGA
- We have started porting the design to the Alveo card

Questions, Comments or Demo request:

- Prasanna Sukumar (Head of Engineering)
  - psukumar@reniac.com
- Thomas Jorgensen (VP Operations & Customer Success)
  - thomas@reniac.com


XDF
XILINX
DEVELOPER
FORUM