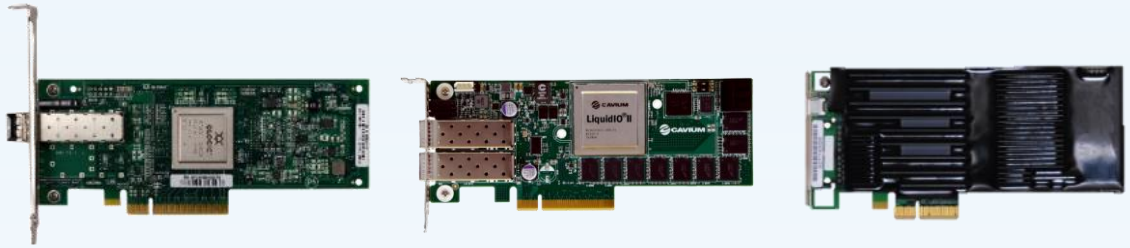


Driving HPC and Compute Workloads on ThunderX

Corporate Overview



- Founded: 2001
- NASDAQ IPO (CAVM): 2007
- Locations: Silicon Valley, MA, India, China, Israel, Taiwan, Singapore, Ireland
- CAGR since IPO = ~29%
- Q3'16 Revenue: $168.1M
- Profitable, Strong Cash Flow

Enterprise | **Mobile Infrastructure** | **Data Center and Cloud** | **Service Provider Cloud**

Multi-Core MIPS, ARM Processors, Security, SDN Switch and Server/Storage Connectivity ~$10B TAM

Products for Data Center and Cloud

ARM Server, SDN Switch, Ethernet/FC Adapters, Security Adapters & Appliances

XPliant®

Ethernet SDN TOR and Spine Switches

THUNDERX®

COTS and OCP Servers

QLogic®/LiquidIO®

Server Adapters, Ethernet/FC, OVS, IPSec Offload

THUNDERX®

Storage Controllers, Ceph

LIQUID SECURITY™

FIPS HSM Appliance
Key management, Authentication as-a-service

Nitrox® Security & FIPS Adapters

CAVIUM

THUNDERX® World's True Server Class ARM Processor

ARM Leadership – ThunderX FIRSTS for ARM Processors

- 24 to 48 custom ARMv8 cores
- 2 Socket (96 core) configurations
- High memory bandwidth
- 1 TB system memory in 2S configuration

- vSOC – Full virtualization from Core to IO
- Configurable Network & Storage IO
- Accelerators for data protection, data integrity, compression, security

ATX

½ SSI

CAVIUM

ThunderX® ARMv8 State of Union

BEST-IN-CLASS WORKLOAD OPTIMIZED™ ARM SERVER PROCESSOR
ROBUST ECO-SYSTEM

Target Markets

- Cloud Computing
- Big Data
- Cloud Storage
- Secure
- HPC
- NFV

Optimized Workloads

redis, APACHE HTTP SERVER, HAPROXY, NGINX, VARNISH CACHE, MySQL, ceph, MEMCACHED, Cassandra, hadoop, OpenACC Directives for Accelerators, OPNFV

ODM/OEMs

Inventec, GIGABYTE™, wistron, Lanner, PENGUIN COMPUTING, FOXCONN, E4 COMPUTER ENGINEERING, CRAY THE SUPERCOMPUTER COMPANY, Lenovo, inspur

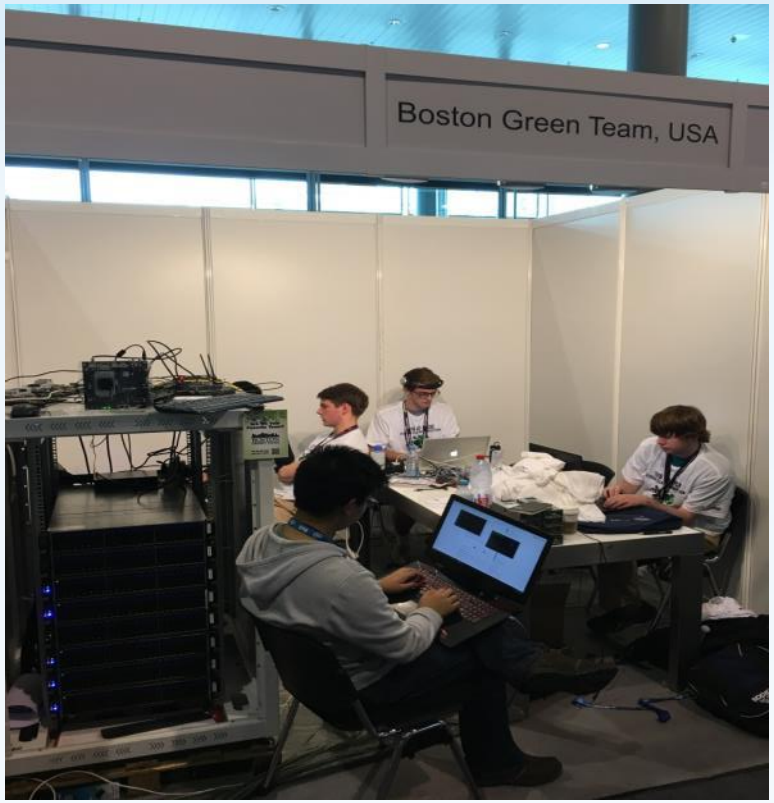
Eco-System

allinea, ORACLE, PathScale High Performance Compilers, redhat, MontaVista, CANONICAL, SUSE We adapt. You succeed., AptioV, American Megatrends, MEGARAC®

CAVIUM

Student Cluster Competition

Thunderstruck - Spain



Boston Green Team

INTRODUCING ROCM SOFTWARE PLATFORM

A New Fully Open Source Foundation for HPC Class GPU computing

AMD

Graphics Core Next Headless Linux® 64-bit Driver

- Multi-GPU Shared Virtual Memory
- Large Memory Single Allocation
- Peer to Peer Multi-GPU
- Peer to Peer with RDMA
- Systems Management API and Tools

HSA drives rich capabilities into the ROCm hardware and software

- User Mode Queues
- Architected Queuing Language
- Flat memory Addressing
- Atomic Memory Transactions
- Process Concurrency & Preemption

HSA™ FOUNDATION

Rich Compiler Foundation for HPC Developer

- LLVM Native GCN ISA Code Generation
- Offline Compilation Support
- Standardized loader and Code Object Fo
- GCN ISA Assembler and Disassembler

Open Source Tools and Libraries

- Rich Set of Open Source Math Libraries
- Tuned Deep Learning Library
- Optimized Parallel Programing Frameworks
- CodeXL Profiler and GDB Debugging

GPUOpen

IBM® Spectrum Computing Community Editions

Free-to-Use HPC Workload Scheduling & Cluster Management

- **IBM Spectrum LSF Community Edition**
  - **Industry leading workload scheduler**
  - IBM Spectrum LSF Express
  - IBM Spectrum LSF Application Center – Basic Edition*
  - Supports up to 10 nodes, Linux on ARM v8 (64-bit), IBM POWER8 Little Endian, and x86-64
- **IBM Spectrum Cluster Foundation Community Edition***
  - **Cluster deployment and management** (Replaces existing IBM Platform Cluster Manager Standard Edition)
- **IBM Platform MPI Community Edition**
  - Supports up to 4K ranks**

ARM

OpenPOWER™

POWER8

x86-64

http://ibm.biz/Bd4vxR

@IBMSDI

*IBM POWER8 LE & x86, **x86

ThunderX and OpenHPC – ARMv8 Technology Preview

- OpenHPC is broad array of HPC OpenSource Software
- OpenHPC V1.2 is now supported on ARMv8
- ThunderX is reference platform for development and test on ARMv8

Functional Areas	Components
Base OS	RHEL/CentOS 7.1, SLES 12
Administrative Tools	Conman, Ganglia, Lmod, LosF, ORCM, Nagios, pdsh, prun
Provisioning	Warewulf
Resource Mgmt.	SLURM, Munge. Altair PBS Pro*
I/O Services	Lustre client (community version)
Numerical/Scientific Libraries	Boost, GSL, FFTW, Metis, PETSc, Trilinos, Hypre, SuperLU, Mumps
I/O Libraries	HDF5 (pHDF5), NetCDF (including C++ and Fortran interfaces), Adios
Compiler Families	GNU (gcc, g++, gfortran)
MPI Families	OpenMPI, MVAPICH2
Development Tools	Autotools (autoconf, automake, libtool), Valgrind,R, SciPy/NumPy
Performance Tools	PAPI, Intel IMB, mpiP, pdtoolkit TAU

Lenovo, Hartree and Cavium Paper Presentation

Quantifying Energy Use in Dense Shared Memory HPC Node

Miloš Puzović*, Srilatha Manne†, Shay GalOn† and Makoto Ono‡
*The Hartree Centre, STFC Daresbury Laboratory, Sci-Tech Daresbury, Cheshire, WA4 4AD, UK
milos.puzovic@stfc.ac.uk
†Cavium, Inc., 2315 N. First Street, San Jose, CA 95131, USA
{bobbie.manne, shay.galon}@cavium.com
‡Lenovo (United States), Inc., 7001 Development Drive, Morrisville, NC 27560, USA
mono2@lenovo.com

***Abstract*—In this paper we introduce a novel, dense, system-on-chip many-core Lenovo NeXtScale System® server based on the Cavium THUNDERX® ARMv8 processor that was designed for performance, energy efficiency and programmability. THUNDERX processor was designed to scale up to 96 cores in a cache coherent, shared memory architecture. Furthermore, this hardware system has a *power interface board* (PIB) that measures with high accuracy power draw across the server board in the NeXtScale™ chassis. We use data obtainable from PIB to measure the energy use of PARSEC and Splash-2 benchmarks and demonstrate how to use available hardware counters from THUNDERX processor in order to quantify the amount of energy that is used by different aspects of shared memory programming, such as cache coherent communication. We show that energy used required to keep caches coherent is negligible and demonstrate that shared memory programming paradigm is viable candidate for future energy aware HPC designs.

***Keywords*—system-on-chip, manycore, energy efficiency, high density, high performance computing, power monitoring, accuracy, measurement

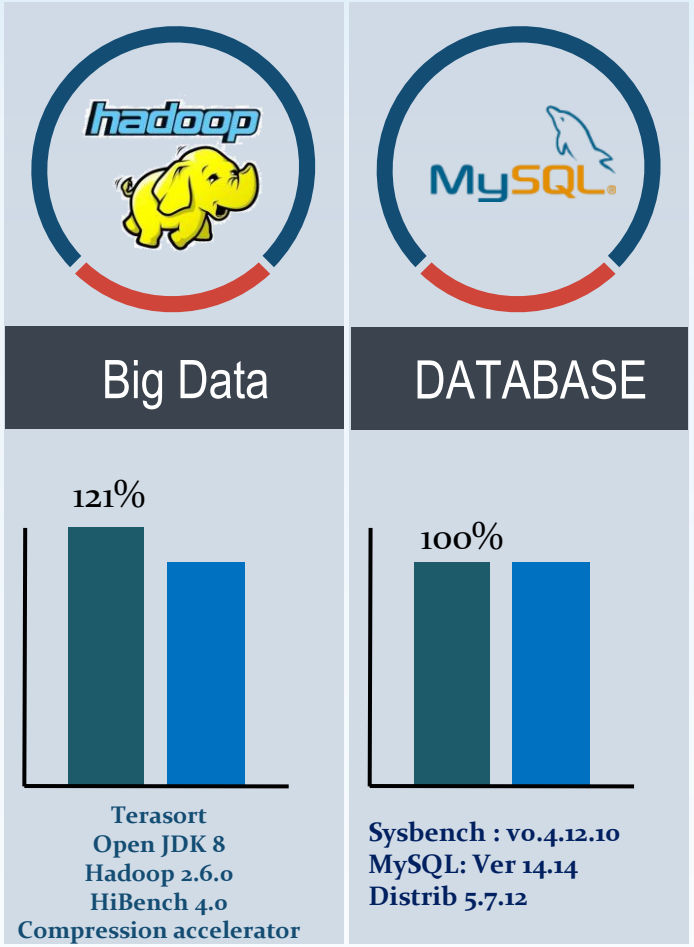
dual socket system consisting of two THUNDERX® processors with a total of 96 ARMv8 cores on a single node. Additionally, it integrates memory, storage and interconnects within a node in order to minimise glue logic and reduce power. Unlike other HPC systems that sacrifice programmability by heterogeneous computational elements such as GPUs to achieve energy efficiency [2] ASIANCAT delivers energy efficiency via integration and programmability via homogeneous compute.

Like other mainstream HPC systems that address the programmability constraint, ASIANCAT uses a shared memory paradigm for low-level communication. To reduce the average memory latency and memory traffic, each core has a private cache which can lead to cache incoherence across the cores. The problem of incoherence in ASIANCAT is solved using a hardware cache coherence algorithm where single-writer, multiple-reader invariant is maintained. The accepted wisdom [3]–[5] is that hardware cache coherence is not scalabl

- Innovative energy study with ThunderX
- Detailed Many Core performance analysis
- Research completed on Lenovo NeXtScale system

ThunderX Benefits: Data Management

Example applications: MySQL, Hadoop, Spark…..



Workload Characteristics

- Highly parallelizable
- Network & Storage IO bound
- Big Data & Database needs compute capacity

ThunderX benefits

- Large core count ➔ Higher throughput
- Integrated low latency, high BW networking
- Integrated, dedicated multiple SATA controllers
- SoC delivers lower cost & power
- Accelerators benefit storage workloads

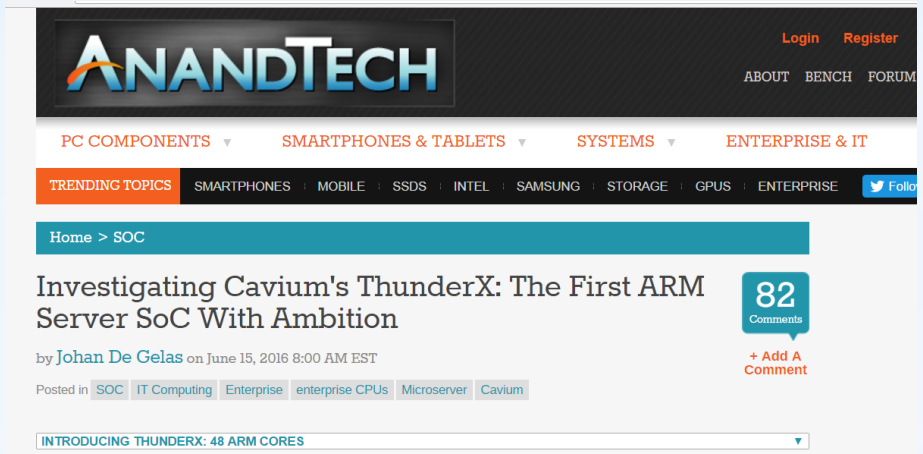
ThunderX 48 cores, 2S, 2.0 GHz

Intel: E5-2690v3, 1S, 12 cores/HT enabled, 2.6 GHz

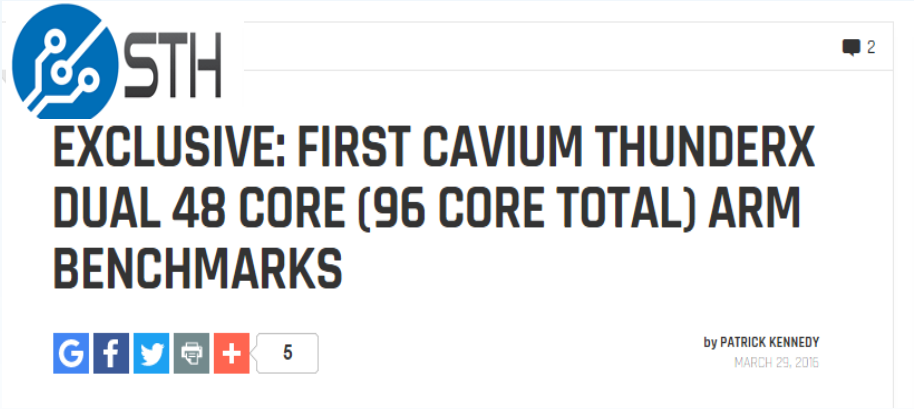
CAVIUM

ThunderX Third Party Benchmarks Available

Performance being compared against competitors by industry experts



"The fact that Cavium allows independent testing instead of just showing numbers on PowerPoint presentations tell us that they have confidence in the future of their ARMv8 endeavors." – Johan De Galas (Anandtech)



"The Cavium ThunderX is the only game in town when it comes to having a generally available data center ARM platform that not just meets Intel's platform in some price/ performance areas but exceeds them in several" - Patrick Kennedy (Servethehome)

Roadmap

THUNDERX2™ Delivers 2-3X the Performance

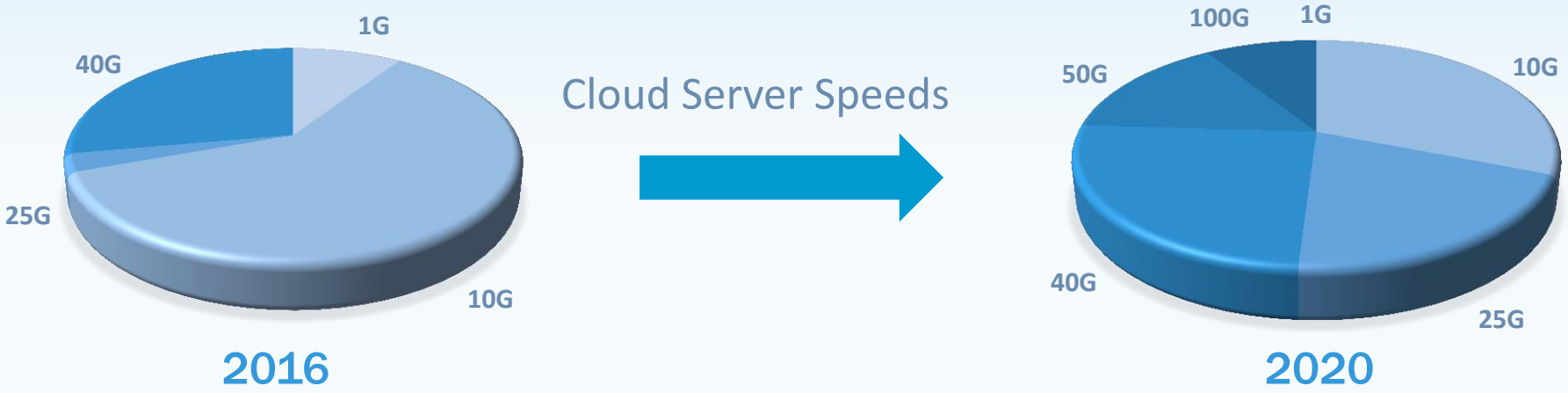
CORES	MEMORY	Next Gen IO
2.5x PERFORMANCE 54 CORES	2x BANDWIDTH 1.5X CAPACITY	Integrated 10/25/50/100G

CACHE	POWER	ACCELERATORS
64K I-cache 40K D-cache	WIDER RANGE LOW IDLE	Packet Process Storage Security

VIRTUALIZATION: virtSOC™ Gen2

CAVIUM

ThunderX2 Enables 25GE Datacenter Transition



- Cloud Data Center in early stages of 25GE adoption
- Today over 50% of installed servers in cloud use 10G links
- 25GE ramping up in 2018

ThunderX2

Integrated
10/25/50/100GE Ports

2.5x Higher Server Bandwidth

CAVIUM

ThunderX2 Expands Target Applications

CAVIUM

See Cavium This Week at SC16

- Tuesday, 10:30AM: Red Hat Booth Presentation
- Tuesday, 2PM: SUSE Booth Presentation
- Entire Show: AMD GPU Demonstrations

Please visit us at Cavium Booth # 4057

THANK YOU