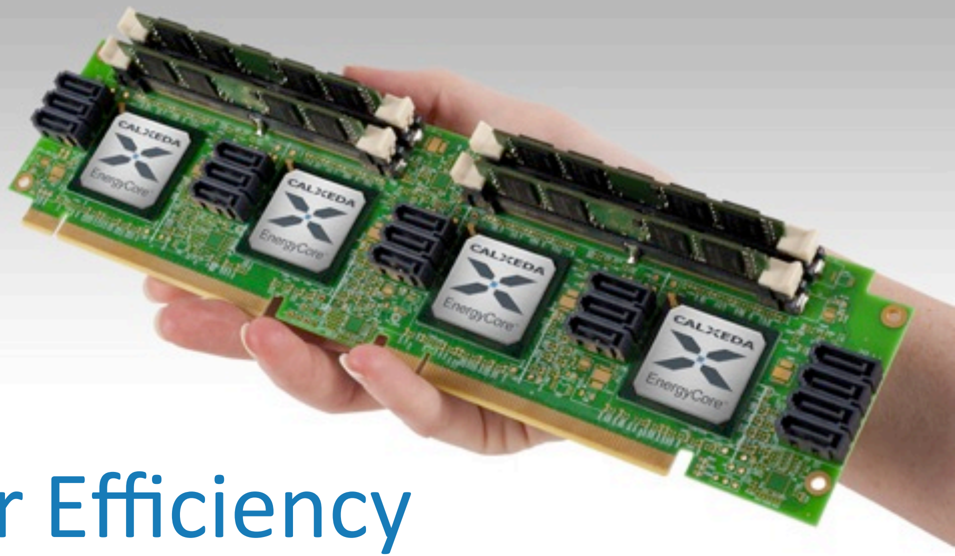


Redefining Datacenter Efficiency

An Overview of Calxeda's architecture and early performance measurements

Karl Freund
November 12, 2012

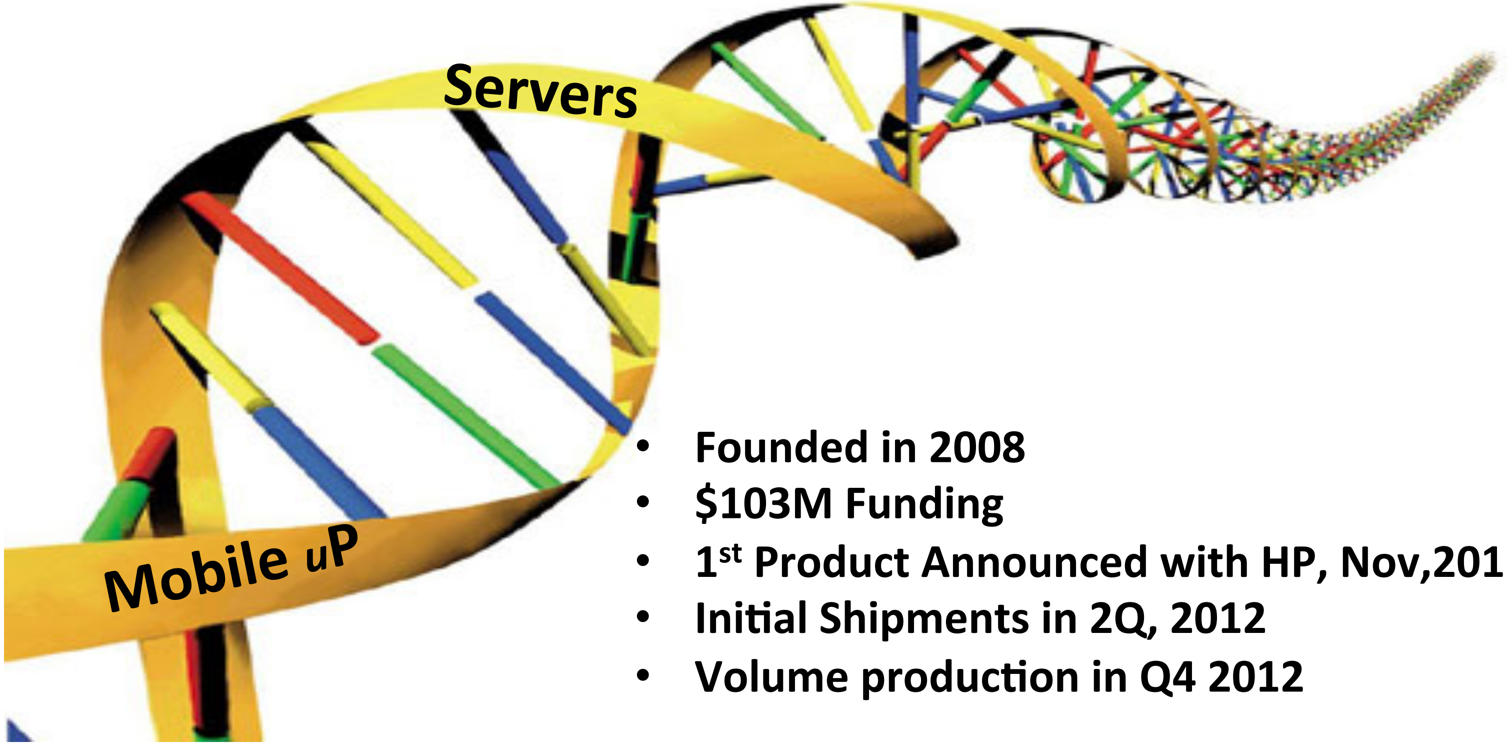
Imagine…

3000 servers, in a single rack

- Delivering 12,000 cores of computing power and 12 TB DRAM
- Reducing direct power requirements by 90%
- With integrated networking and management
- Eliminating 9 miles of cabling
- Eliminating 125 Ethernet Switches

… And that's just the beginning

Calxeda: Datacenter performance, cell phone power



- Founded in 2008
- $103M Funding
- 1st Product Announced with HP, Nov,2011
- Initial Shipments in 2Q, 2012
- Volume production in Q4 2012

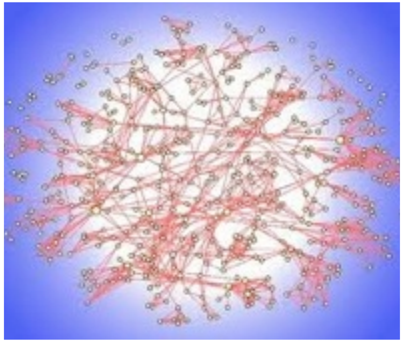
The Calxeda EnergyCore™ SOC

Efficient

Integration

A Complete Server-on-a-Chip:
90% less energy
90% less space
50% lower costs

Scalable



Interconnect

Connect thousands of server nodes into an integrated cluster solution

Smart

Management

Cluster-level power and system optimization

EnergyCore architecture at a glance

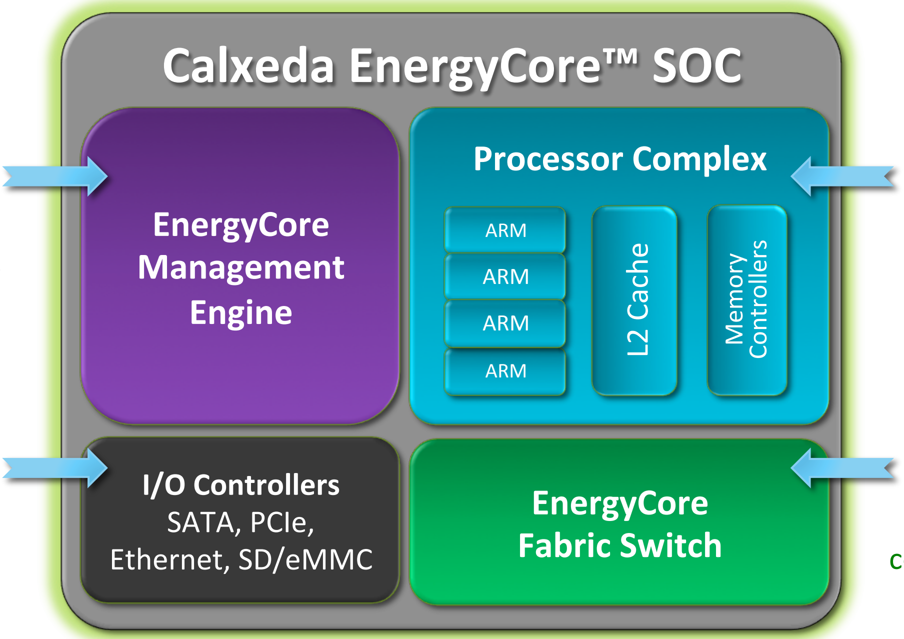
A complete building block for hyper-efficient computing

EnergyCore™ Management Engine

Advanced system, power and fabric optimization and management for energy-proportional computing

I/O Controllers

Standard drivers, standard interfaces. No surprises.



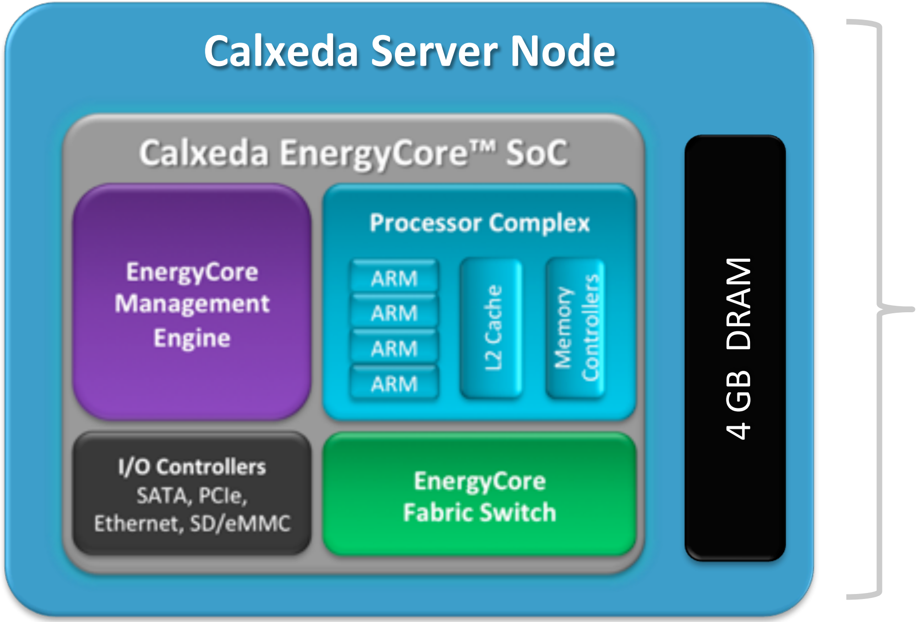
Processor Complex

Multi-core ARM® processors integrated with high bandwidth memory controllers

EnergyCore™ Fabric Switch

Integrated high-performance fabric converges internode communication, I/O, and storage networking

A Complete Server, only 5 Watts



Typical* Max Power: 5 Watts

Power at Idle: < ½ Watt

* The power consumed under normal operating conditions under full application load (ie, 100% CPU utilization)

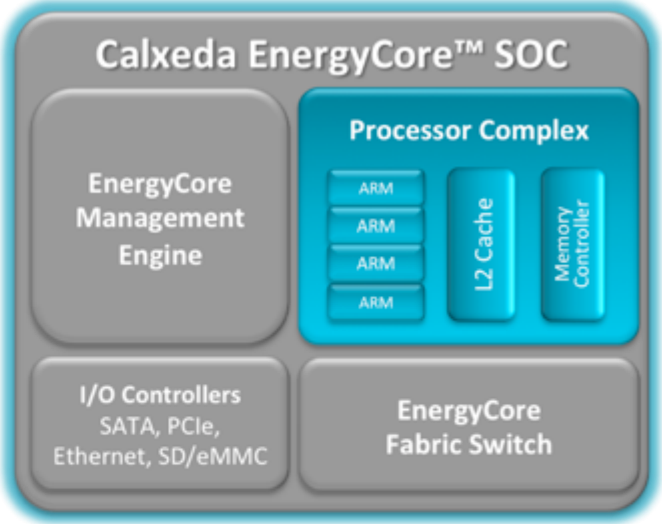
Calxeda is Efficient: it starts at the core

Multi-core ARM® Cortex™ processor

"Calxeda deserves kudos for following its energy-efficiency vision, eliciting the interest of significant partners, and exciting the imagination of the marketplace." *Charles King, PundIT*

  - 1.1GHz, up to 1.4 GHz
  - Supports FPU (scalar) and NEON (SIMD) Floating Point
    - FPU: full IEEE-754 compliant, single & double precision FP
    - NEON: 64-bit and 128-bit registers supporting SIMD operations
- On-board 4MB shared L2 cache
- Integrated Memory Controller
  - 72-bit datapath, with ECC
  - DDR3/3L: 800, 1066, and 1333 MT/sec

➔ **Maniacal focus on Performance/Watt/$**



Calxeda is Scalable: An integrated fabric

The EnergyCore Fabric Switch

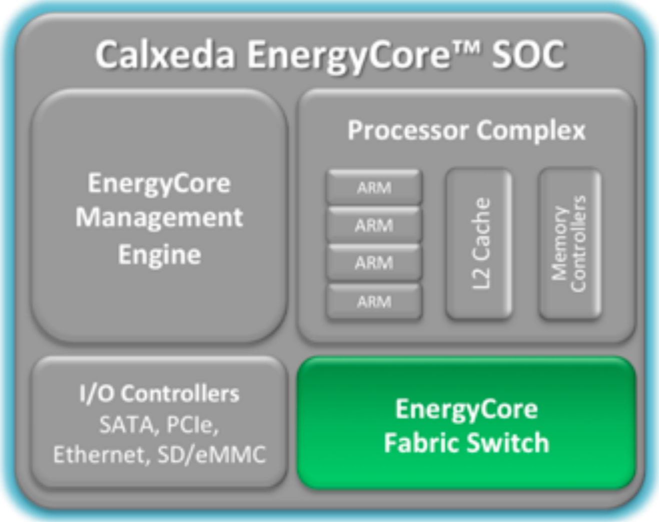
“Calxeda’s fabric scales magnificently." *Roger Kay, Forbes.com*

High-bandwidth fabric converges inter-node communication, I/O, storage, and networking

- Up to 5 channels:
  - Dynamic bandwidth: 1 Gb to 10 Gb per channel
  - < 200 Nano-Seconds latency, node to node
- Topology agnostic
- Fabric is transparent to OS and software
  - Presents 2 Ethernet ports to the OS

→ **Eliminates Top-of-Rack-Switch ports & cabling**

→ **Enables extreme density, lowers cost and power**



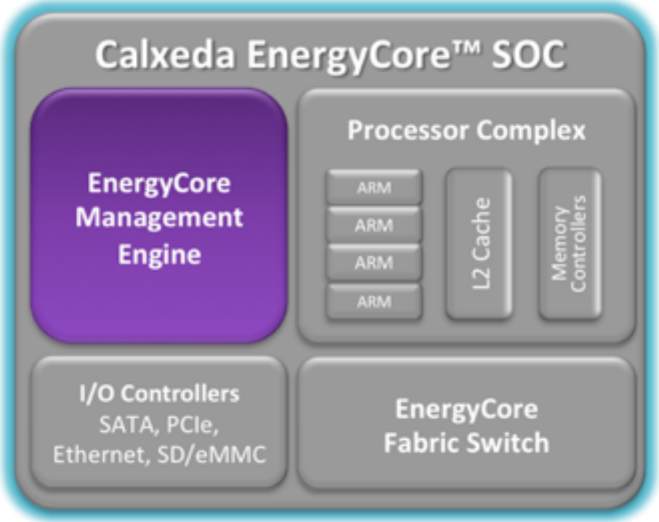
Calxeda is Smart: Integrated Management

The EnergyCore Management Engine:

"The magic is that they really understand the data center side and not just the low-power side (of the processor design). It is the right blend of what you need, and that is impressive." *Carl Claunch, Gartner*

It's like a free BMC, including software. **PLUS:**

- Autonomic SOC power management
- Autonomic Fabric routing and power optimization
- Industry standards systems management Interfaces
  - IPMI, DCMI
- Dynamically updates & configures management stack
- Eliminates ~$28 **per node** in BMC, FW, and port costs*
- Enables OEM added-value for management offerings



Calxeda EnergyCore™ SOC

EnergyCore Management Engine

Processor Complex

ARM

ARM

ARM

ARM

L2 Cache

Memory Controllers

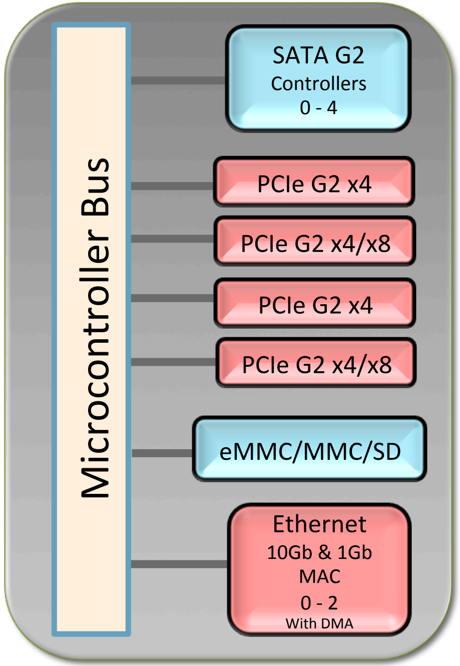
I/O Controllers
SATA, PCIe,
Ethernet, SD/eMMC

EnergyCore Fabric Switch

➔ Built-in Management lowers costs & optimizes cluster

* From IDF presentation by Alex Renzin, Facebook: http:www.intel.com/go/idfsessions

Optimizing I/O for a balanced system



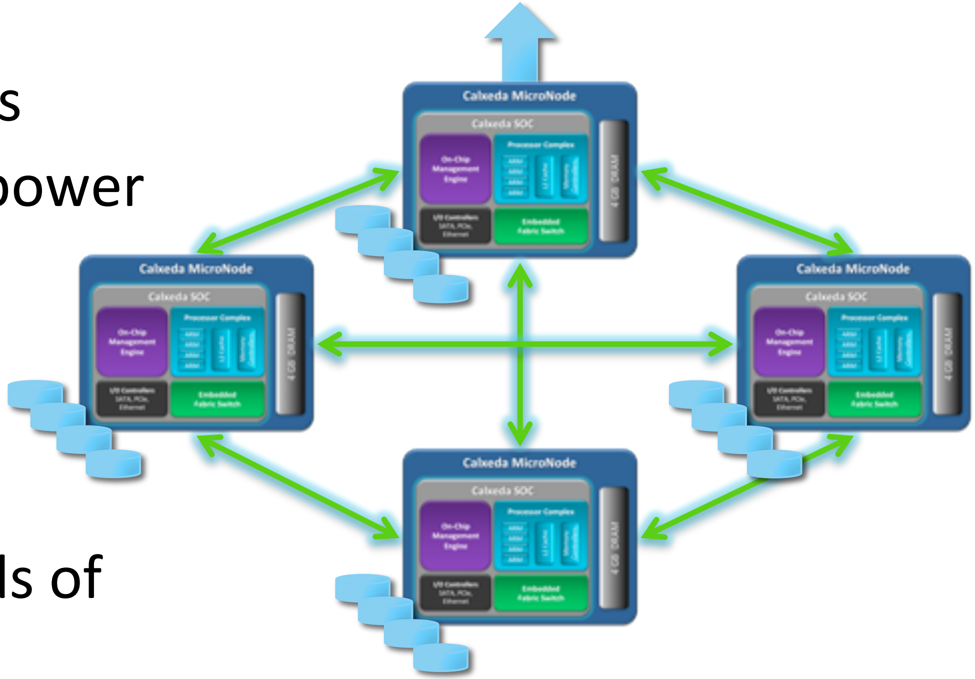
- SATA Drives (3Gb/s)
  - Up to 5 disk drives
- Configurable PCIe for expansion
  - Four x4/x2/x1 *-or-*
  - Two x8
- Integrated eMMC/MMC/SD controller
  - Card or device support
- Three 10Gbs Ethernet Controllers
  - With DMA to the Quad Cortex A9

A small Calxeda Cluster

A Simple Example:

- Start with four ServerNodes
- Consumes only 20W total power
- Connected via distributed fabric switches
- Connect up to 4 SATA drives per node
- Then scale this to thousands of ServerNodes

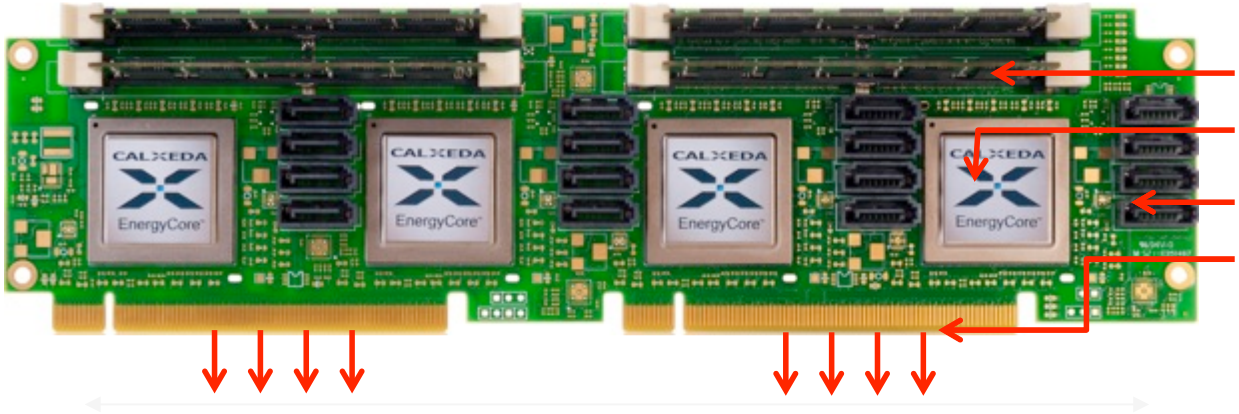
Extend the fabric, or connect to Ethernet (1-10Gb)



EnergyCard: a Quad-Node Reference Design

- Four-node reference platform from Calxeda
- Available as product and/or design
- Plugs into OEM system board with passive fabric, **no** additional switch HW

  EnergyCard delivers 80Gb Bandwidth to the system board. (8 x 10Gb links)



4 GB DRAM ECC mini-DIMMS

Quad-core servers

4 SATA / Node (flexibility!)

Power, SATA, & Fabric

Approximately 10”

4 Servers.
Complete.
Only 20W.

Example EnergyCore ECX-1000 System Configurations

Redstone Development Platform

89% Less Energy

94% Less Space

64% Less Cost

288 EnergyCore Server Nodes in 4U

Viridis Project

BOSTON
Servers · Storage · Solutions

Ultra Dense, Ultra Low Power Computing Platform

2U SuperMicro Chassis with:

- 48 EnergyCore Server Nodes
- 24 x 2.5" HDD's

Penguin Computing Server

Ideal for Hadoop and Cloud storage

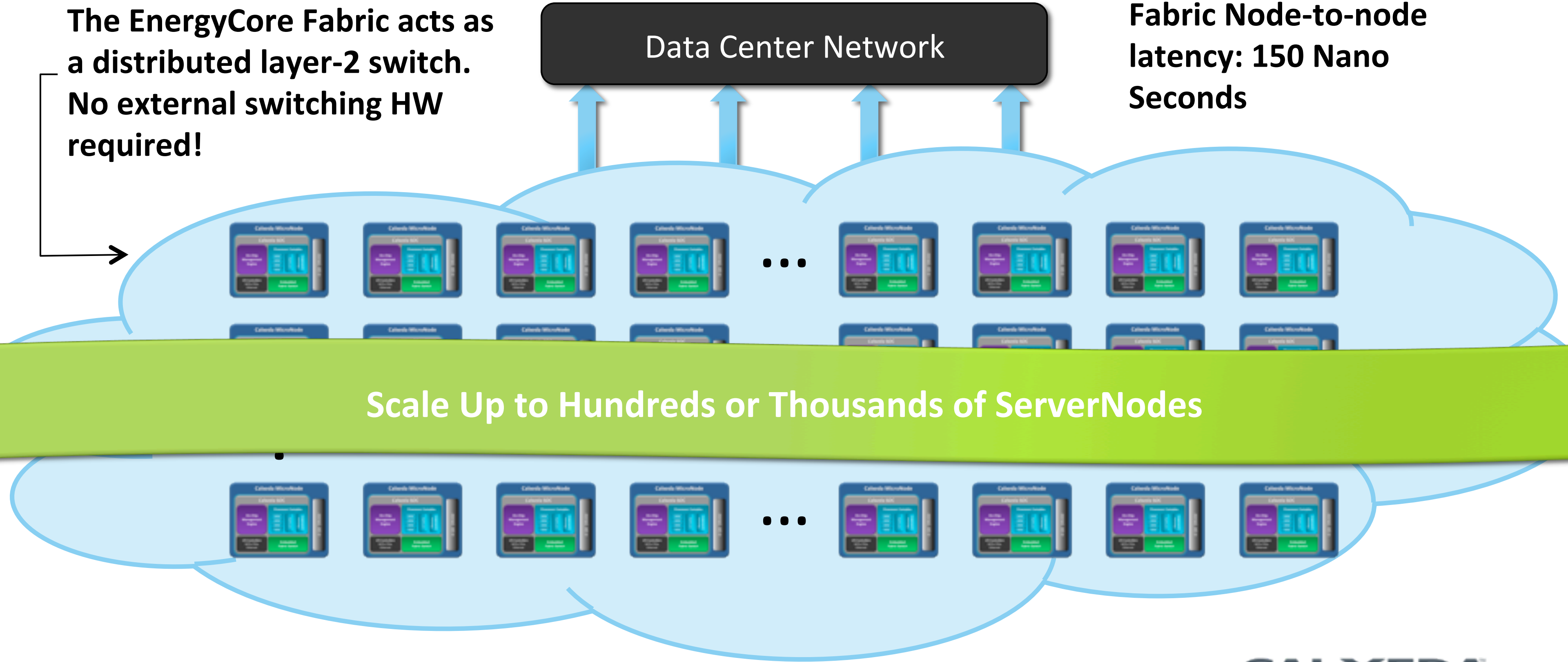
- Up to 48 quadcore ECX-1000 nodes
- Up to 36 3 ½ " Disks

PENGUIN COMPUTING

Calxeda is Scalable

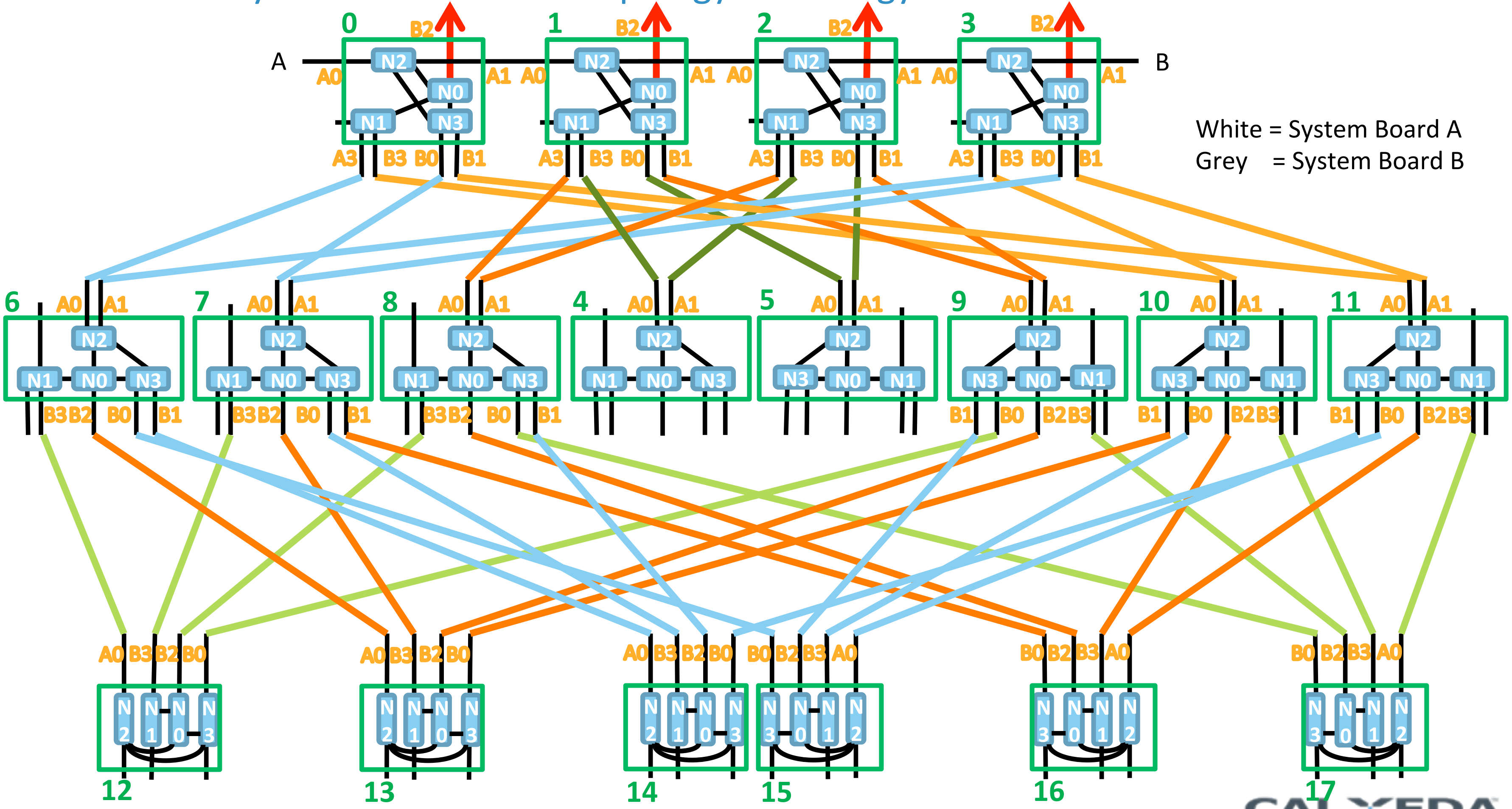
The EnergyCore Fabric acts as a distributed layer-2 switch. No external switching HW required!

Fabric Node-to-node latency: 150 Nano Seconds



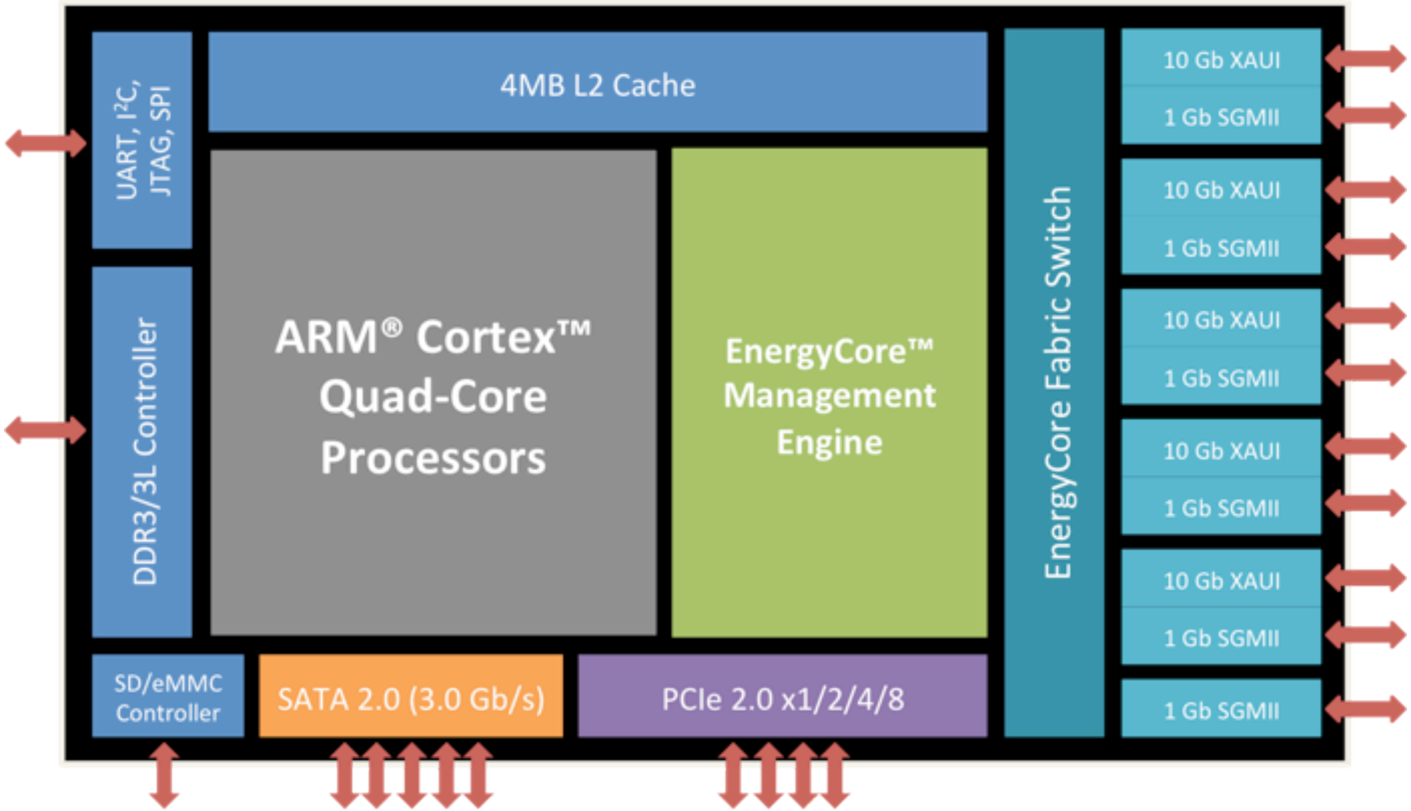
Scale Up to Hundreds or Thousands of ServerNodes

System Board Fabric Topology: 18 Energy Cards = 72 nodes



White = System Board A
Grey = System Board B

Calxeda EnergyCore™ Block Diagram



Calxeda Software Ecosystem – Base Packages

Linux Kernel v3.2

ubuntu Server 12.04 LTS

Compilers/Languages

- GCC/gFortran 4.6.3
- PHP 5.3.10
- Perl 5.14.2
- Python 2.7.3
- Ruby 1.8.7
- Erlang r14

Debuggers/Profilers

- GDB 7.4
- GProf 2.15
- OProfile 0.9.6

Java

- Oracle JVM SEv7u6
- OpenJDK 6b24

Applications

- Apache 2.2.22
- Tomcat 6.0.35
- MySQL 5.5.22
- PostgreSQL 9.1

fedora v17+

Compilers/Languages

- GCC/gFortran 4.7.0
- PHP 5.4.0
- Perl 5.14.2
- Python 2.7.2, 3.2.2
- Ruby 1.8.7
- Erlang r14B

Debuggers/Profilers

- GDB 7.4
- GProf 2.13
- OProfile 0.9.6

Java

- ~~Oracle JVM SEv7u6~~
- OpenJDK 6b24

Applications

- Apache 2.2.21
- Tomcat 7.0.25
- MySQL 5.5.20
- PostgreSQL 9.1.2

* Version numbers subject to change and are highly dependent on Linux distribution

Calxeda Software Ecosystem – HPC Packages

Linux Kernel v3.2

ubuntu Server 12.04 LTS

MPI

- MPICH 1.2.7
- OpenMPI 1.4.3
- MPICH2 1.4.1
- Open-MX 1.5.2

Checkpoint

- DMTCP 1.2.1
- Condor 7.2.4

Libraries

- BLAS 1.2
- FFTW 2.1.5
- ScaLAPACK 1.8.0

Monitoring

- Ganglia 3.1.7

fedora v17+

MPI

- ~~MPICH 1.2.7~~
- OpenMPI 1.5+
- MPICH2 1.4.1+
- Open-MX 1.5.2

Checkpoint

- ~~DMTCP 1.2.1~~
- Condor 7.4.2+

Libraries

- ~~BLAS 1.2~~
- FFTW 3.3
- ScaLAPACK 1.7.5+

Monitoring

- Ganglia 3.1.7

Struck thru items are not yet available for ARM in this version of Fedora but Calxeda is proposing they be added.

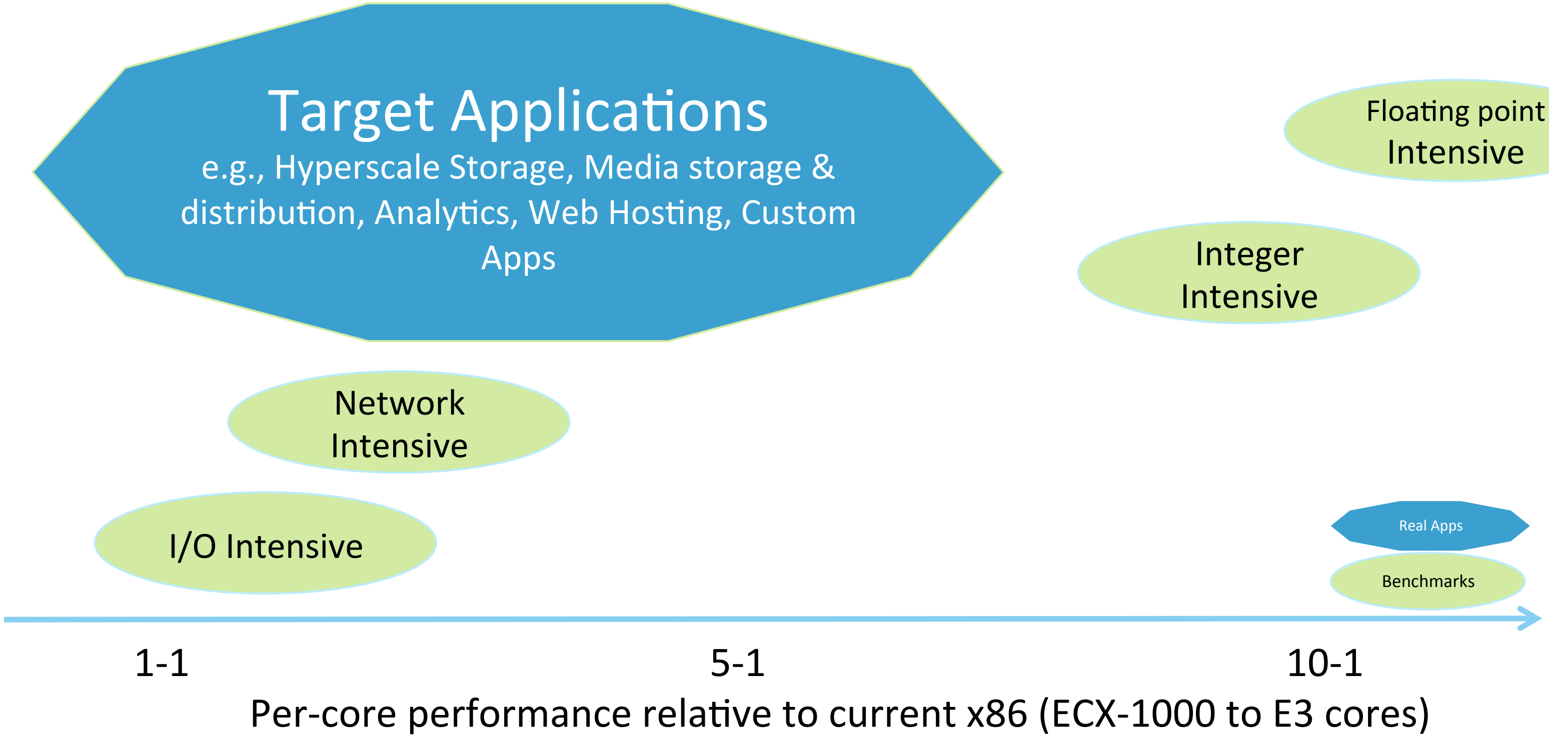
* Version numbers subject to change and are highly dependent on Linux distribution

Calxeda Software Ecosystem – Application Packages

Linux Kernel v3.2	
ubuntu Server 12.04 LTS	fedora v17+
• **Apache Cassandra 1.1.1+** Packages to be provided by DataStax • **Apache Hadoop 1.0.0+** Packaged to be provided by Cloudera • **Memcached v1.4.13+**	• **Apache Cassandra 1.1.1+** Packages to be provided by DataStax • **Apache Hadoop 1.0.0+** Packaged to be provided by Cloudera • **Memcached v1.4.13+**

* Version numbers subject to change and are highly dependent on Linux distribution

Workloads and Benchmarks: EnergyCore ECX-1000



Target Applications

e.g., Hyperscale Storage, Media storage & distribution, Analytics, Web Hosting, Custom Apps

Floating point Intensive

Integer Intensive

Network Intensive

I/O Intensive

Real Apps

Benchmarks

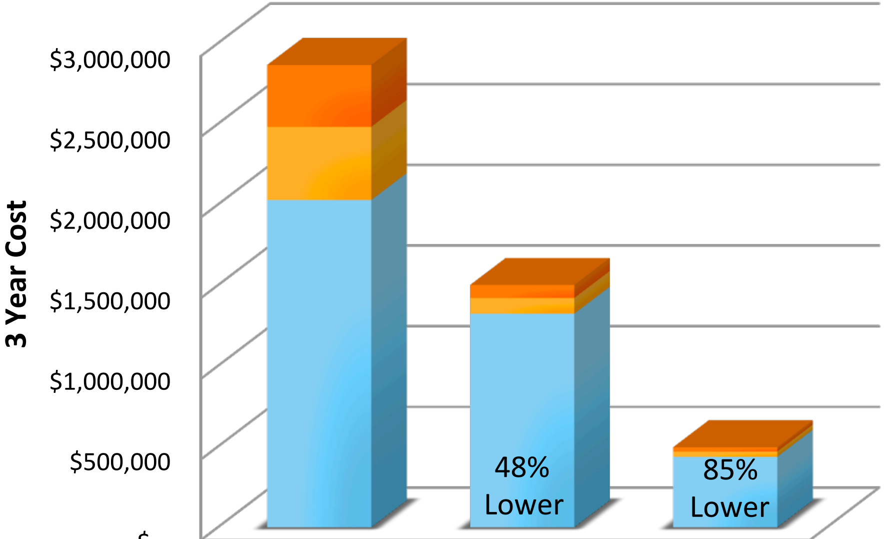
1-1

5-1

10-1

Per-core performance relative to current x86 (ECX-1000 to E3 cores)

TCO Benefits: Two Examples



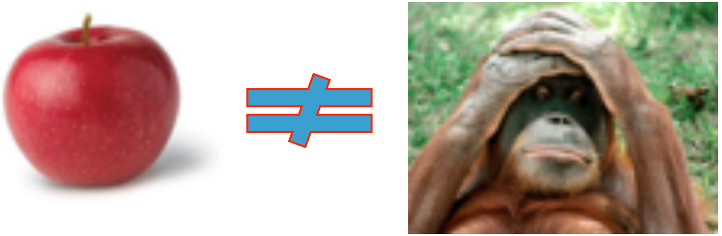
	X86 Sandybridge	CX: 8:1	CX: 2.6:1
🟧 Space	$383,200	$80,733	$26,663
🟨 Power	$454,381	$98,116	$32,404
🟦 TCA	$2,028,000	$1,323,500	$437,142

These two models reflect the 3 year cost of acquisition, power, and space of servers and networking gear for 2 different workloads, characterized by the ratio of Calxeda nodes needed to equal a **dual-socket X86 server**

CX: 8:1 shows TCO for workloads that are a mix of compute and I/

CX: 2.6:1 shows TCO for workloads that are I/O intensive
Note: “ApacheBench” ratio is 2.6:1)

Don’t compare Apples and Orangutans!



Power of:	EnergyCore ECX-1000 (4 cores)	E3-1220L v2 (2 Core Ivybridge)
Processor	2W (½W/CPU)	17W[1]
SoC or Chipset	3.8W	~27W
(SOC or Chipset) + 4GB DRAM	5W	~31W
System Power @ Wall (with Disk)	5.4W-8.5W	???

Don’t confuse a processor with an SOC when comparing power. The Calxeda SOC includes most or all of the functions of an entire server chipset.

EnergyCore ECX-1000 Power Measured Under Load

Workload (on 24 nodes & SSDs)	**Total System* Power (Today!)**	**~Power per ECX-1000 Node (with disk @Wall)**
Linux at Rest	130W	5.4W
phpbench	155W	6.5W
Coremark (4 threads per SOC)	169W	7.0 W
Website @ 70% Utilization	172W	7.2W
LINPACK	191W	7.9W
STREAM	205W	8.5W

*All measurements done on a 24-node system @1.1GHz, with 24 SSDs and 96 GB DRAM in the Calxeda Lab.

For targeted workloads, ECX-1000 can enable a complete 24-node cluster at similar power level as a 2 socket x86

Storage Solution Overview

Power and performance optimized storage solution for:

- Cloud Storage market
  - Amazon Web Services (AWS) S3
  - Rackspace CloudFiles
  - Dreamhost (DreamObject)
- Traditional NAS & SAN replacement
  - Enterprise & HPC NAS - traditional on-premise file archival, shared drives, etc
  - HPC shared storage market – backend storage for diskless compute nodes

More Performance/$ with Calxeda for I/O-Intensive Workloads

Traditional x86

10 servers
1.0PB HDD
20Gb Bandwidth

4X more IOPS per $

Calxeda

80 servers
1.0PB HDD
200Gb Bandwidth

- More CPU cores and SATA channels per rack
- Multiple additional built-in 10Gb interconnects
- “East-West” bandwidth for storage replication
- Fabric topology can be optimized for various storage strategies

ApacheBench Results and Power Comparison

	EnergyCore ECX-1000	Intel Xeon E3-1240[1]
Core Frequency	1.1 GHz	3.3 GHz
CPU Cores	4	4
Total Requests	1,000,000	1,000,000
Requests per Second	5500 → 6100!	6950
Latency (Average)	9 ms	7 ms
Power (Average)[2]	5.26 W	50-100 W[3]
Performance/Watt Advantage	7-15X	

Preliminary measurements provided by Calxeda. Running ApacheBench 2.3 against Apache v2.4.2 with 16k request size. Calxeda system running on one EnergyCore ECX-1000 SoC, with one 1Gb network link, and 4GB DDR3L-1066 memory. [1] Intel-based system running with one E3-1240, one 1Gb network link, and 16GB DDR3 memory. [2] Power measurements exclude disk and PSU overhead, with a sampling rate of 2-seconds. [3] Intel Xeon E3-1240 TDP: 80W; Intel C216 Chipset TDP: 6.7W; 16GB RAM: 16W

Sysbench OLTP (MySQL)-Preliminary Numbers

	EnergyCore ECX-1000	2 Socket Intel Xeon E3-2670 [1]
Core Frequency	1.1 GHz	2.6 GHz
CPU Cores	4	16
System Memory	4GB	32GB
Database	MySQL 5.5.24	MySQL 5.5.24
Transactions per second	634	3542
Estimated Performance/Watt Advantage	3-4X	

Calxeda rewrites the TCO equation

Traditional x86

$3.3M

400 servers
10 racks
20 switches
1,600 cables
91 kilowatts

89% less energy
94% less space
63% less cost
97% less complexity

HP ‘Redstone’

$1.2M

1,600 servers
1/2 rack
2 switches
41 cables
9.9 kilowatts

Calxeda rewrites the TCO equation vs. ATOM

Centerton Atom Based Servers

$1.54M

2400 servers
3 racks
243 cables
34 kilowatts

67% less energy
61% less space
36% less cost

Calxeda Based Servers

$0.99M

1800 servers
1 rack
80 cables
11 kilowatts

Calxeda Single Node Benchmark Preliminary Results

Calxeda EnergyCore 1.1GHz DDR-1066

LINPACK	**Performance GFLOPs**	**Node Power(W)**
Single Precision	4.83	5.5
Double Precision	2.78	5.8

*Preliminary System Power estimates would place Calxeda ~#160 on the June 2012 Green500 List

Calxeda EnergyCore 1.1GHz DDR-1066

Graph
SCALE: 21
32-bit Port
TEPS: 8,794,294.517148 (8.794294517148392e+06)

Currently modeled at 2-2.5X better than Sandybridge on Perf/Watt

Note that Midway will double Floating Point performance at same frequency

- All Tests run on Single Node with 1.1GHz Clock Frequency using Ubuntu 12.04 in a Calxeda Greenbox Reference Platform.
- Tests used 1066 memory

Preliminary Comparisons with You-Know-Who

Benchmark	CX	“YKW”	Perf Ratio	CX Power	YKW Power	Perf/Watt
FIO-Read MB/S	406	412	1.01	7.96	80.7	10.0
FIO-Write MB/S	365	363	0.99	7.96	80.7	10.2
Sysbench	490	1065	2.17	7.75	49	2.9
phpbench	10252	68497	6.68	5.9	94.1	2.4
Cloud Suite	21.77	152.67	7.01	7.1	97.9	2.0
coremark	11104.9	197522.7	17.79	6.6	191.5	1.6
stream	1492.7	18833.5	12.62	7.96	95.9	1.0

Distributed Storage Software & Partners

ceph

Ideal for:

- Cloud storage providers
  (ex: Dreamhost's DreamObject)
- Backend cloud compute storage
  (ex: Volume services for OpenStack)

Features:

- Object Storage
- Block Storage
- File System (POSIX)

Other Benefits:

- OpenStack SWIFT compatible
- Available on Ubuntu today
- Open-source licensing
- Service/Support through Inktank

GLUSTER

Ideal for:

- Cloud storage providers
  (ex: Dreamhost's DreamObject)
- Enterprise NAS replacement
  (ex: for internal file storage/archival)

Features:

- Object Storage
- File System (POSIX)

Other Benefits:

- OpenStack SWIFT compatible
- Available on Fedora/RHEL today
- Open-source licensing with large user community
- Service/Support through RedHat
  (through acquisition in 2012)

ScaleIO

Ideal for:

- SAN alternative for enterprise
  (ex: shared storage for diskless compute nodes)
- Backend cloud compute storage
  (ex: Volume services for OpenStack)

Features:

- High-performance, SAN storage for scale-out block storage

Other Benefits:

- Commercial license with support from ScaleIO
- Focus exclusively on block storage

Calxeda: Rewriting the TCO Equation

Calxeda increases compute efficiency by an order of magnitude.

1⁄10th the energy[1]
1/10th the space[2]
1⁄2 the TCO[3]
All the performance

[1] Calxeda's analysis of dual socket Intel 5620 @ 20% utilization = 135W vs. 2 Calxeda SoCs @ 10W
[2] Calxeda 120 node diskless compute server in 2U chassis compared to 20 dual socket Dell servers
[3] Based on James Hamilton's TCO tool, with Calxeda = 1/3 x86 performance, CX @ 5W http://perspectives.mvdirona.com/2010/09/18/OverallDataCenterCosts.aspx