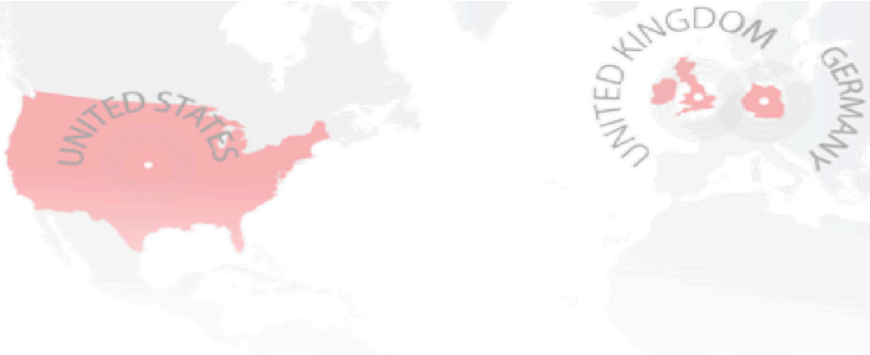

VIRIDIS
Microserver
THE BOSTON VIRIDIS ARM SERVER:
Addressing the Power Challenges of Exascale
In partnership with
ARM
CALXEDA
Power Your Tomorrow

Agenda

- Introduction to the Boston Virids
- Boston Viridis Architecture
- Benchmarks and Customer Case Studies



Who are Boston?

- **Founded in 1992** – 20+ years of innovation
- Global HQ in London with branch offices in London, Munich, Mumbai, Bangalore and New York. Further global expansion is planned.
- Calxeda's **global launch partner** with Viridis, at ISC12
- **First company to ship** production ready ARM servers
- Shipping Viridis servers for the past 12 months

Viridis in the media

VIRIDIS Microserver

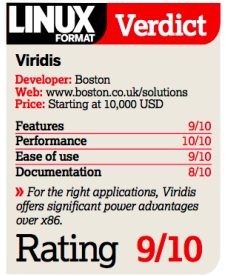
*Viridis server is **"…a serious threat to x86 server domination"** – ExtremeTech*

EXTREMETECH

Green I.T. AWARDS 2013 WINNER

ANANDTECH

*The current A9 based Calxeda EC 1.4 GHz is about **40% faster and consumes half the power of the Atom S1260** - Anandtech*



*For the right applications, Viridis offers **significant power advantages over x86** - LinuxFormat*

LINUX FORMAT The #1 source for Linux

CRN

CRN – Viridis is one of '***The 10 Coolest Servers Of 2012***'

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

The Boston Viridis

Our Mission: **to enable our partners and customers to increase IT efficiency by an order of magnitude.**

How we will achieve it: **by integrating efficient processors from ARM® Holdings with Calxeda's fabric, management, and I/O innovations.**

The result:

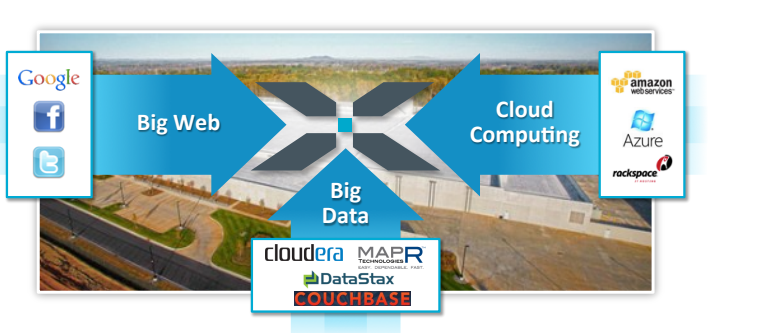
- 1⁄10th the energy
- 1/10th the space
- 1⁄2 the TCO

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

The Data Center Landscape: Massive Changes

Reshaping the Data Center



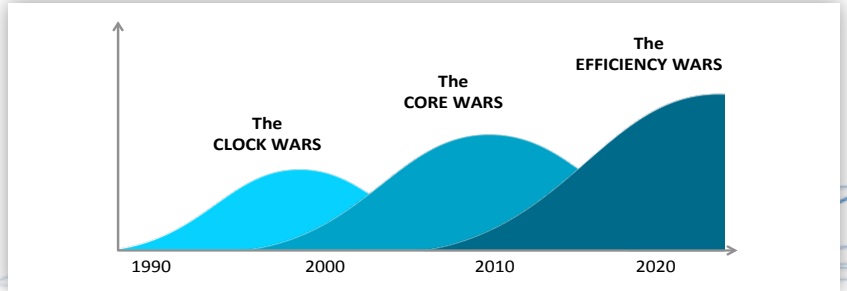
Creating Huge Challenges, and Opportunities

- Not Enough Power
- Not Enough Space
- Complex Infrastructure

The Opportunity:

- Improve Efficiency by 10X
- Lower Cost, Lower Power
- Improve Service Levels

Entering a New Era in Computing



VIRIDIS Microserver

Boston Ltd

BOSTON Servers · Storage · Solutions

Targeted Applications and Benefits

Web Serving

40% Lower TCO
65% Reduction in Power
Reduced Failure Domain

MiddleWare

45% Lower TCO
70% Reduction in Power

Analytics

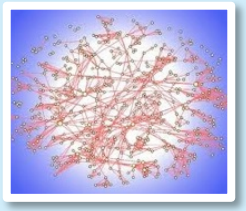
4X throughput per rack
70% lower power

Storage & File Serving

3-4X Throughput per $
Side-Band Replication

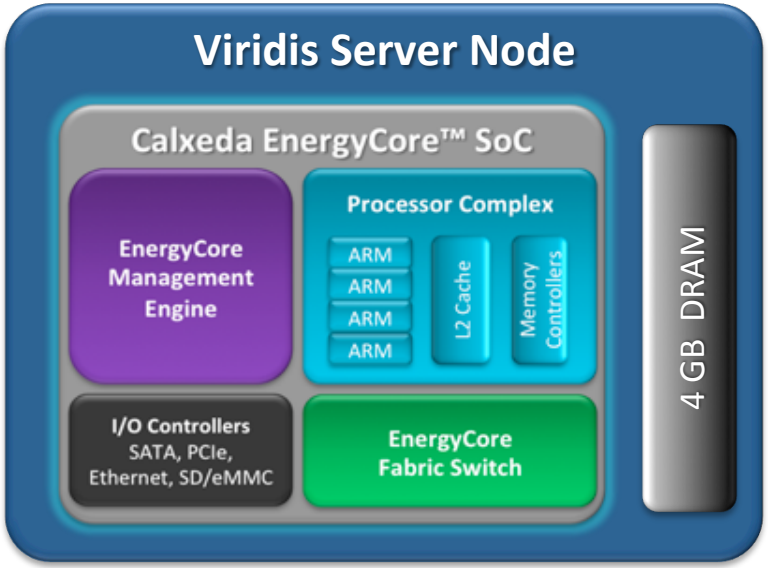
CLOUD Infrastructure as a Service:
50% Lower TCO

Enabling an Efficient Software-Defined Data Center

	Efficient	Scalable	Smart
			
	Integrated Server SOC	**Integrated Fabric**	**Mgt. & Offload Engines**
Today Optimize a Rack	Complete Server-on-a-Chip	Hundreds of nodes 10 Gb fabric	Cluster-level power and system optimization
Future Optimizing a Fleet	64-bit multi-core SOC Server-grade I/O & features	Hundreds of racks @40Gb Shared Resource Pools	Fleet Application Services Offload Engines
Customer Pain Point	Improve Performance / Watt	Lower Infrastructure Costs Optimize Perf/Watt/$	Optimize hyperscale application performance and latency

The starting point:

A Complete Server for 5W



Typical* max power per node:

5 Watts

* The power consumed under normal operating conditions under full application load (ie, 100% CPU utilization) with one 1Gb link active. Adding 4 x 10 Gb links adds 1 watt. Does not include shared power overhead of fans and power supply

Viridis architecture at a glance

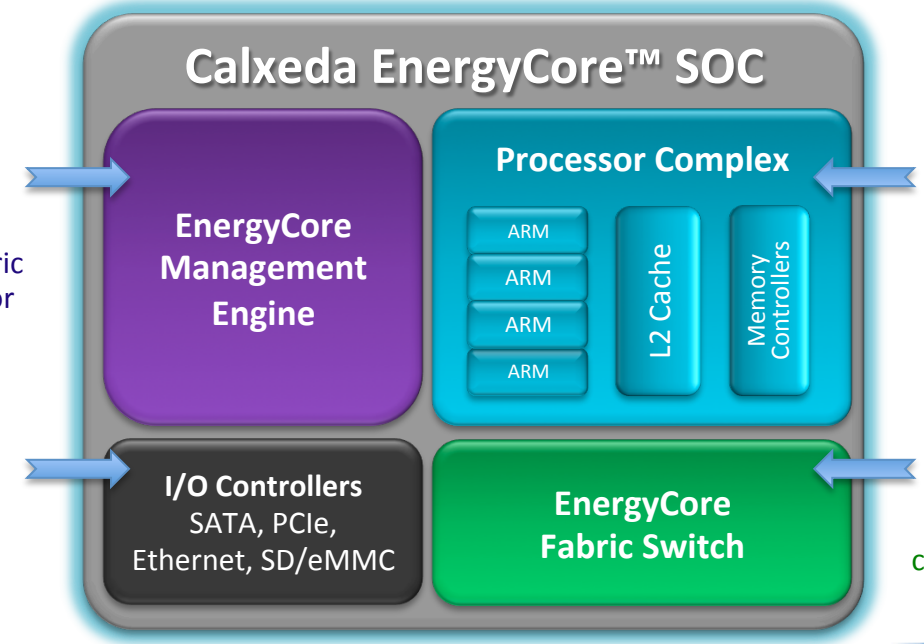
A complete building block for hyper-efficient computing

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

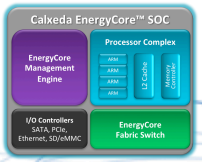
VIRIDIS Microserver

Boston Ltd

BOSTON Servers · Storage · Solutions

Integrated Server Management Infrastructure

Rack Level
Scalable
Control & Mgt



- Fabric Management
  - Dynamic discovery
  - Cluster partitioning
  - Bandwidth Optimization
- Power Control
  - Thermal monitoring
  - Fine grain power optimization
- Health Status
  - Error tracking & event logging
  - Detailed metrics for operation and performance targets
- Industry Standard OOB Interfaces
  - Standard IPMI 2.0 support
  - Extensions for Calxeda Value Add features
- Cluster Wide Management
  - Cluster aware control for configuration
  - Policy based control
- Extensible
  - Opportunity for tighter partner integration

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

The Viridis EnergyCore® Management Advantage

Built-in “Management Engine” provides BIOS, BMC, and Power Management features

Reduces cost and Power	Replaces separate BMC chips and FPGAs with on-chip management processor & firmware, with dedicated Ethernet MAC for secure management.
Boot Management	uBoot from SATA or PXE-boot from network device
Server Management	Read and modify a wide range of server settings and status through industry-standard IPMI and DCMI interfaces
Power Management	A wide variety of power settings enable policy-based power optimization for the SOC and Fabric.
Fabric Optimization	Creates and maintains fabric routing information. Provides automatic fault routing to improve resilience and availability. Also provides adaptive routing for better dynamic performance & hotspots reduction.

The Viridis EnergyCore™ Fabric Advantage

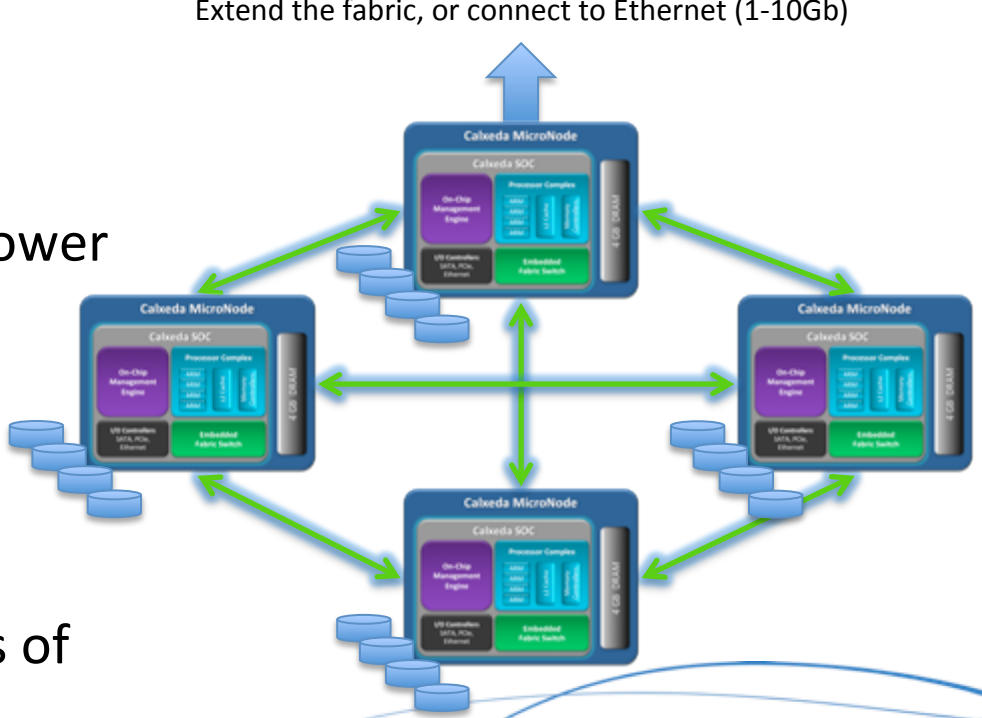
Simple inter-node connectivity reduces expensive networking gear while maintaining standard Ethernet software and datacenter interfaces .

Port Consolidation	Minimize # of Ethernet Phys to save power, costs, cabling cost / complexity / failures, ports on Ethernet switches, …
Bandwidth & Latency	Dramatically increase bandwidth / reduce latency in fabric vs. traditional rack deployment. Provides minimum latency to standard datacenter Ethernet network.
Power	Static and dynamic adaptive link management reduces power to match workload needs.
Flexibility	Range of deployments: from single node, o a small cluster to maximum density rack, up to 4096 server nodes. Facilitates dynamic node provisioning, all while providing standard Ethernet rack connectivity and existing datacenter & software infrastructure.
Fault Resilience	Automatic fault routing to improve resilience and availability. Also provides adaptive routing for better dynamic performance & hotspots reduction.
Security	Support for secure “tunnels” across fabric for multi-tenancy and isolated management networks.

A Small Viridis Cluster

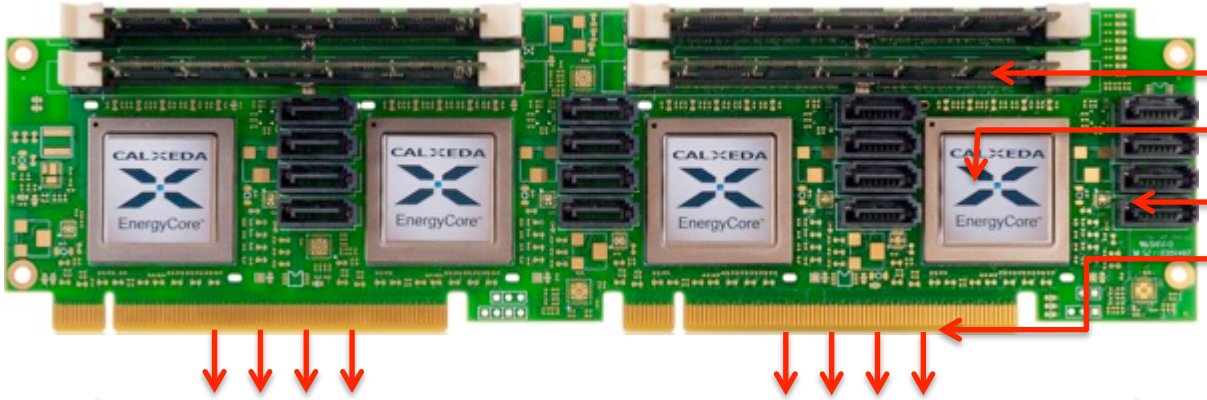
A Simple Example:

- 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



EnergyCard: Quad-Node Design

- Four-node reference platform from Calxeda
- Plugs into Boston Viridis 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.

Viridis is Scalable

The EnergyCore Fabric acts as a distributed layer-2 switch. Reduces external switch port count, cost, and power.

Data Center Network

Fabric Node-to-node latency: 150 Nano Seconds

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

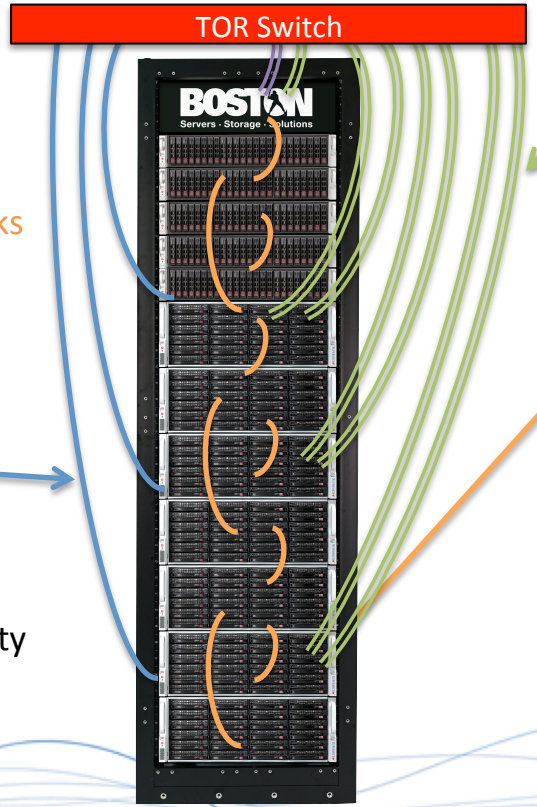
Using Viridis Fabric to Optimize for Hyperscale

Top of Rack Options

- Reduce number of TOR switches – Fabric Links used to share fewer Uplinks
- Eliminate TOR switches – Fabric Links used for all in-rack communication, Uplinks go directly to End of Row switches

Management traffic

- Can use dedicated Management Uplinks to TOR/EOR switches
- Can be routed on Network Uplinks
  - Viridis hardware fabric security ensures traffic can no be intercepted



Boston Ltd

Network Uplinks

- Connect Viridis systems to datacenter switches
- For traffic to other racks/rows

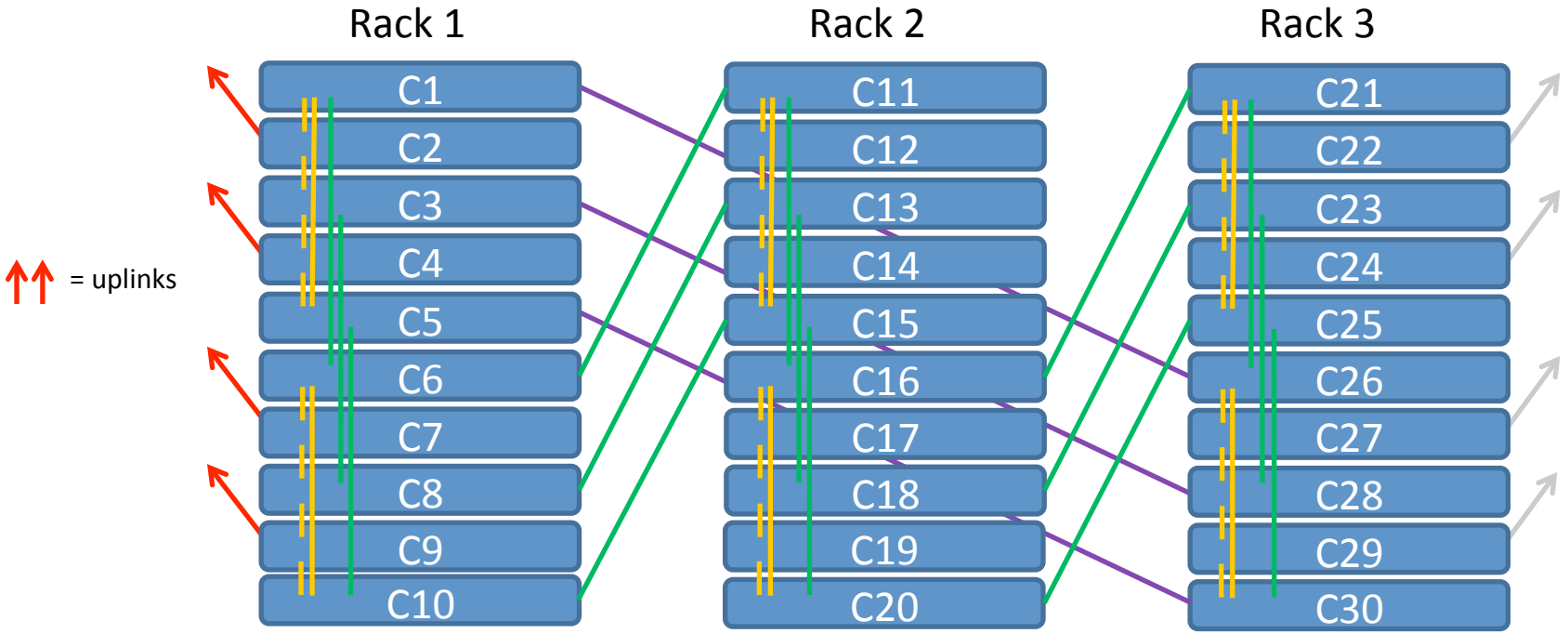
Fabric Links

- Links Viridis 10Gb fabric between chassis
- Can be used to share Network Uplinks and provide uplink redundancy
- For "East-West" traffic
  - Storage replication/ synchronization
  - Communication between tiers in the same rack

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

2D Torus: Multi-Rack Cabling View



- Rack 1 and 2 are powered from different AC circuit breakers
- Rack 3 is usually further away with independent power and cooling
- Data and metadata replicas in different power and failure domains

Boston 2U Viridis

12 Energy Cards each Energy Card has 4x SoCs

Each SoC : 4C/4GB RAM

Integrated Fabric

Each SoC has 2x 10GB low-latency NICs

5.5w per server on Idle

48 Servers (SoC) in 2U

Runs Ubuntu or Fedora

Worlds First production ARM Server!

VIRIDIS Microserver

Tech Blog: www.lowpowerservers.com

Boston Ltd

BOSTON Servers · Storage · Solutions

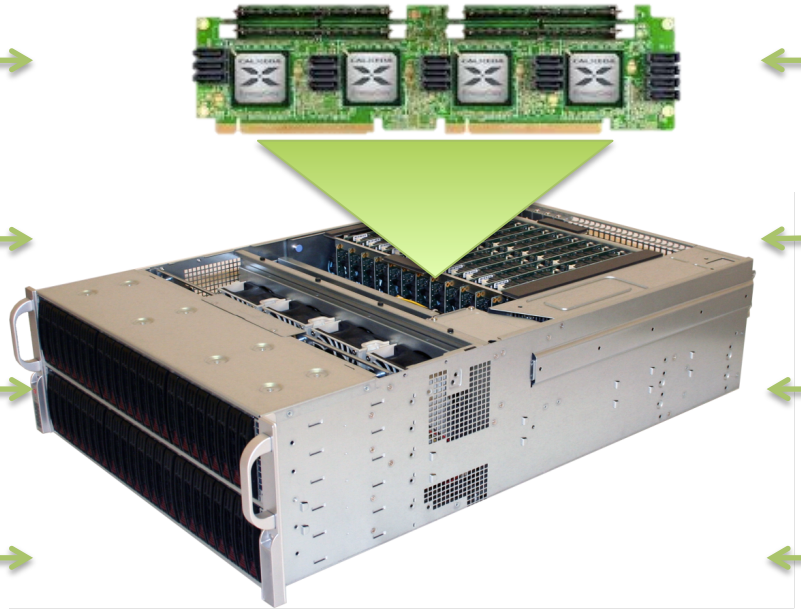
Boston 4U Viridis

12 Energy Cards each Energy Card has 4x SoCs

Each SoC : 4C/4GB RAM

Integrated Fabric

Each SoC has 2x 10GB low-latency NICs



5.5w per server on Idle

48 Servers (SoC) in 4U

Runs Ubuntu or Fedora

Up to 36/72 3.5”/2.5” drives

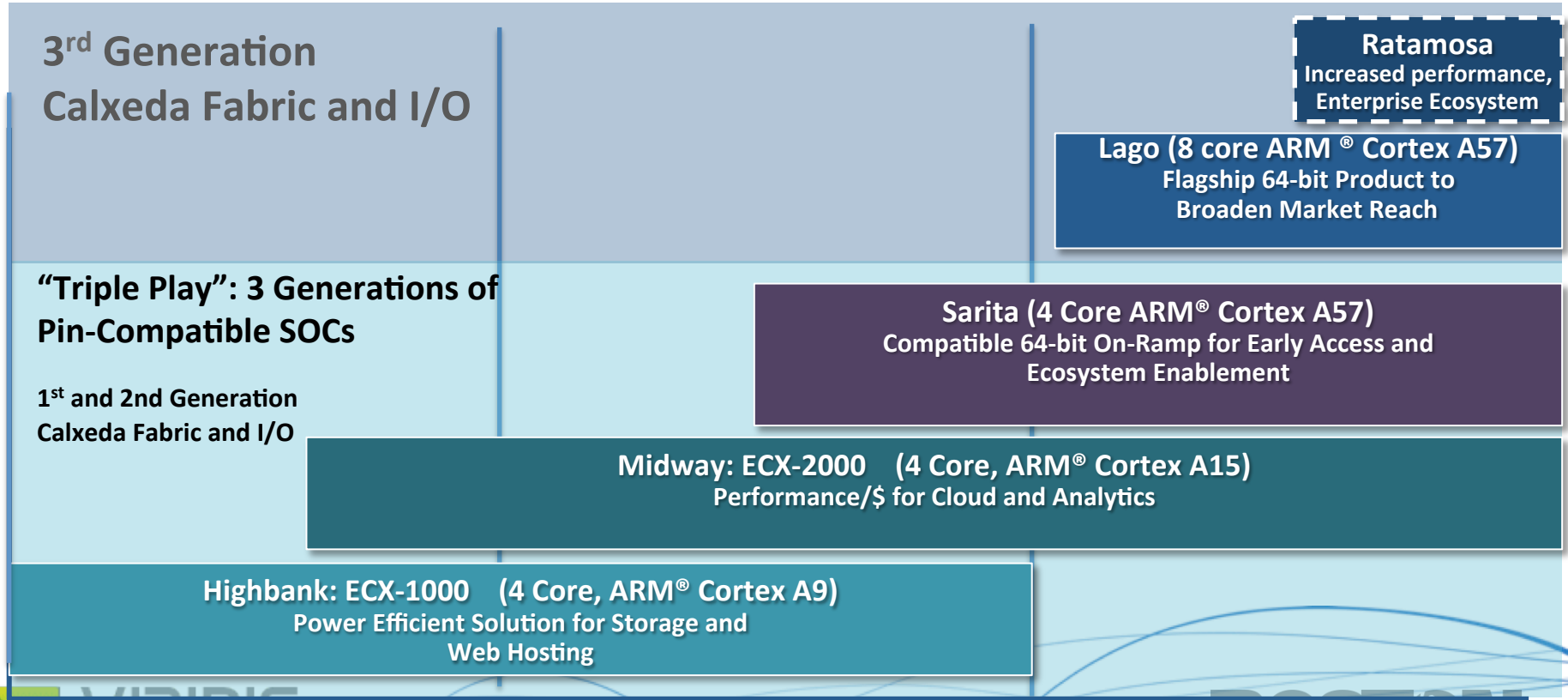
VIRIDIS Microserver

Tech Blog: www.lowpowerservers.com

Boston Ltd

BOSTON Servers · Storage · Solutions

Viridis SOC Roadmap



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

Viridis 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/Boston is proposing ~~they be added~~.

Some additional software..

- **Ellexus**; Tracing/Porting software
- **Eltechs**; Binary level translots to run x86 binaries on ARM (or open-source requiring proprietary modules)
- **Allinea**; HPC Debugging tools
- **FhGFS**; Fraunhoffer Parallel Filesystem (Lustre alternative)
- **CEPH**: Inktank Partnership

ellexus

ELTECHS

allinea

Fraunhofer FS

ceph

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

Viridis Power Measurements

Workload (on 24 nodes & SSDs)	Total System* Power (Today!)	~Power per Viridis 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

Video: www.lowpowerservers.com

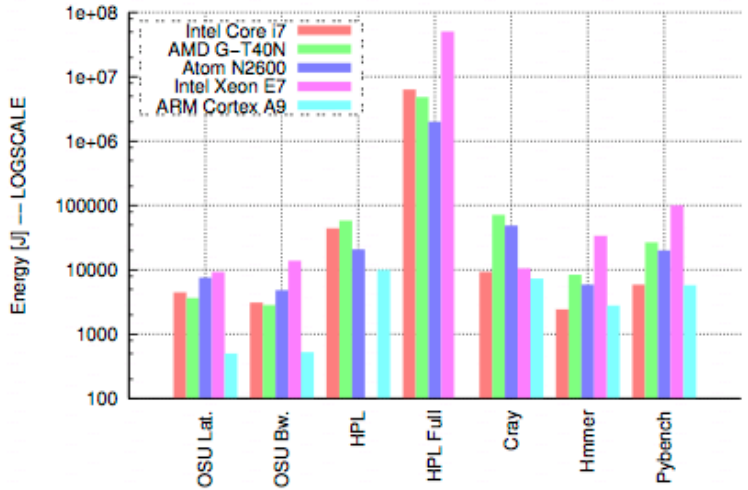
*All measurements done on a 24-node system @1.1GHz, with 24 SSDs and 96 GB DRAM in the Boston Lab.

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

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

Customer Report: University of Luxembourg



Name	Processor Type	MFlops/W	Green500 Rank*
viridis	ARM A9 Cortex	595.93	130
i7	Intel Core i7	565.13	133
bull-bcs	Intel Core E7	324.85	186
atom64	Intel Atom N2600	55.00	476
amdf	AMD Fusion G-T40N	65.45	467

* Based on November 2012 list http://www.green500.org/

- ARM Cortex A9 is almost always **the most energy-efficient** CPU
- Intel Xeon E7 requires much more energy to execute the same application

UNIVERSITÉ DU LUXEMBOURG

VIRIDIS Microserver

Ref: http://www.irit.fr/~Georges.Da-Costa/ee-lsds2013/slides/arm.pdf

BOSTON Servers · Storage · Solutions

Fedora Project selects Viridis

- Fedora project using Boston Viridis for their ARM distribution builds and porting/testing.
- Running in production today
- Expect F19 to be fully built and tested on Boston hardware!

*"The Fedora Project team's experience—**from install to deployment to production**—is a testament to compatibility of Linux code on Calxeda: it just works," Karl Freund, vice president of marketing for Calxeda**

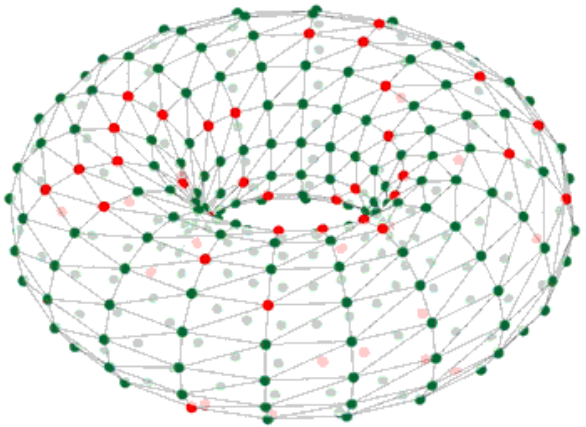
VIRIDIS Microserver

*Reference: http://www.lowpowerservers.com/?p=330

BOSTON Servers · Storage · Solutions

Manchester SpiNNaker

What is SpiNNaker? SpiNNaker is a novel massively-parallel computer architecture, inspired by the fundamental structure and function of the human brain



How is Viridis being used? The Boston Viridis is being used as a host system for launching jobs and workload on their custom ARM cluster

CRUK / EuroCloud

Mapping Biomedical HPC Workloads to Low Power SoC Environments

We expected that, like other typical Cloud environments, this architecture would have a problem with "Big Data". But a key feature of the EnergyCore is integrated 80Gb high performance networking;

Our proposed application architecture, with LAMP services, parallel scheduler and parallel file system in a single modular component may provide an ideal solution for many traditional data-centric HPC workloads.

Reference: CRUK Poster Abstract ISC13:
http://www.isc-events.com/isc13_ap/presentationdetails.php?t=contribution&o=2114&a=select&ra=eventdetails

Aaas {ARM-as-a-Service} – www.armasaservice.com

The world's first ARM® cloud service dedicated to facilitating software migration to the ARM® architecture

- ARM®-as-a-Service (AaaS) has been designed specifically to ease the transition of moving or migrating your software to the ARM architecture.
- The solution is based on the Viridis Microserver and provides developers with all the tools and services required to port and migrate software to the ARM platform.

ARM BOSTON Servers · Storage · Solutions CALXEDA Power Your Tomorrow ellexus ELTECHS

Video on using Breeze:
http://www.lowpowerservers.com/?p=268

VIRIDIS Microserver

BOSTON Servers · Storage · Solutions

More on our Blog:

www.lowpowerservers.com

- Viridis vs HP Moonshot (anandtech)
- OLTP Benchmarking
- ARM vs Atom, Phoronix Benchmarks
- Power Tests Video
- Docking Throughput Tests
- **x86 code running on ARM!**