

Millicomputing – The Future in Your Pocket and Your Datacenter

Open Hardware by the milliWatt

<http://www.millicomputing.com>
acockcroft@netflix.com

June 27, 2008 Usenix Boston

Although the author is employed by Netflix Inc. these are the personal opinions of the author and no endorsement by Netflix Inc. is implied.

Content published under Creative Commons Attribution Share-Alike 2.5

<http://creativecommons.org/licenses/by-sa/2.5/>

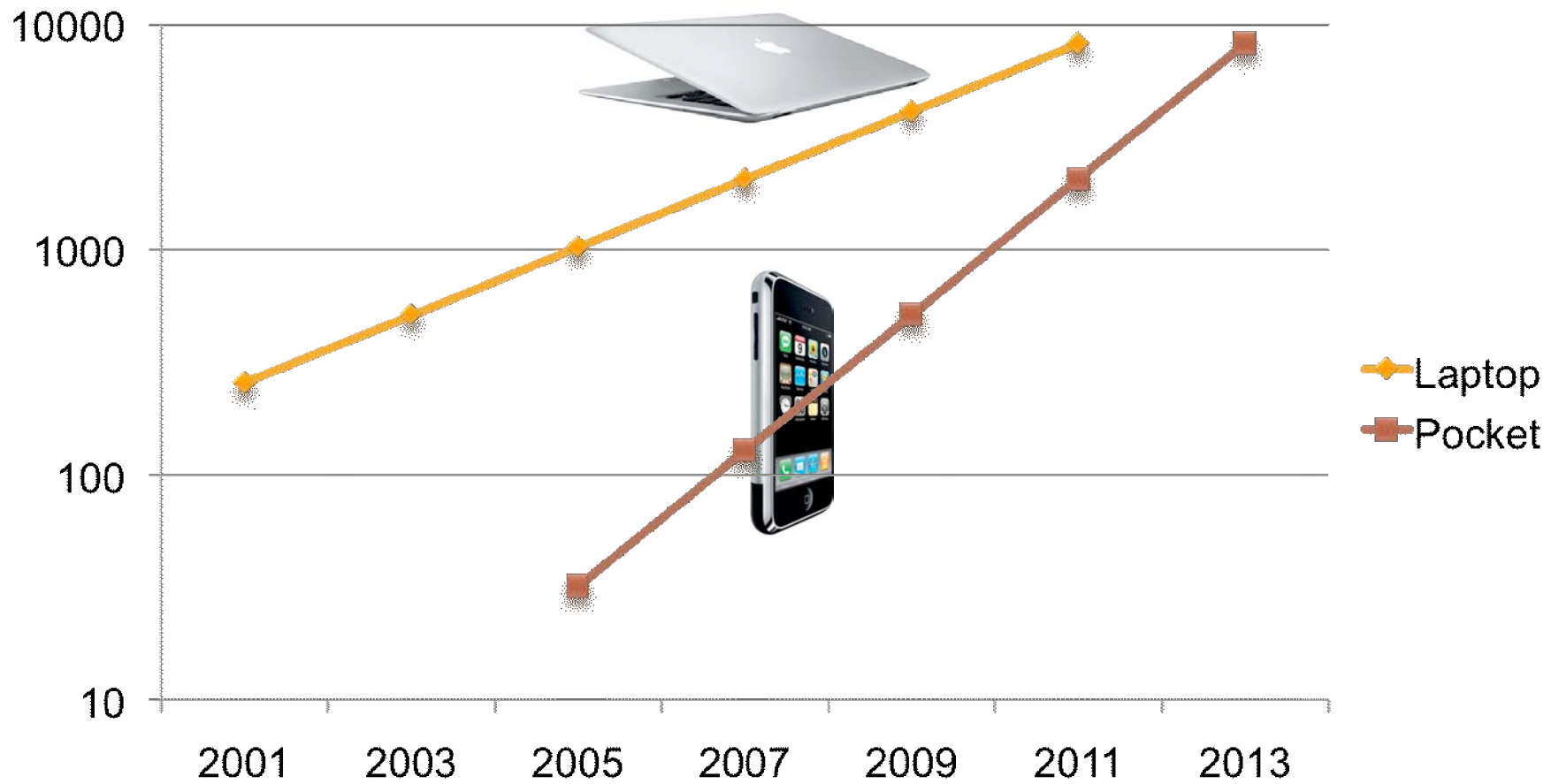
Millicomputers

- Battery powered portable games, phones, cameras etc. have increasingly powerful CPUs inside
- Very low power use, very low cost 32bit RISC with Linux support
- Tiny and reliable system on a chip
- Millicomputer definition

A computer that uses less than one Watt

Put it in your pocket without burning your leg!

Memory Trends (MB)



What's New - 2007

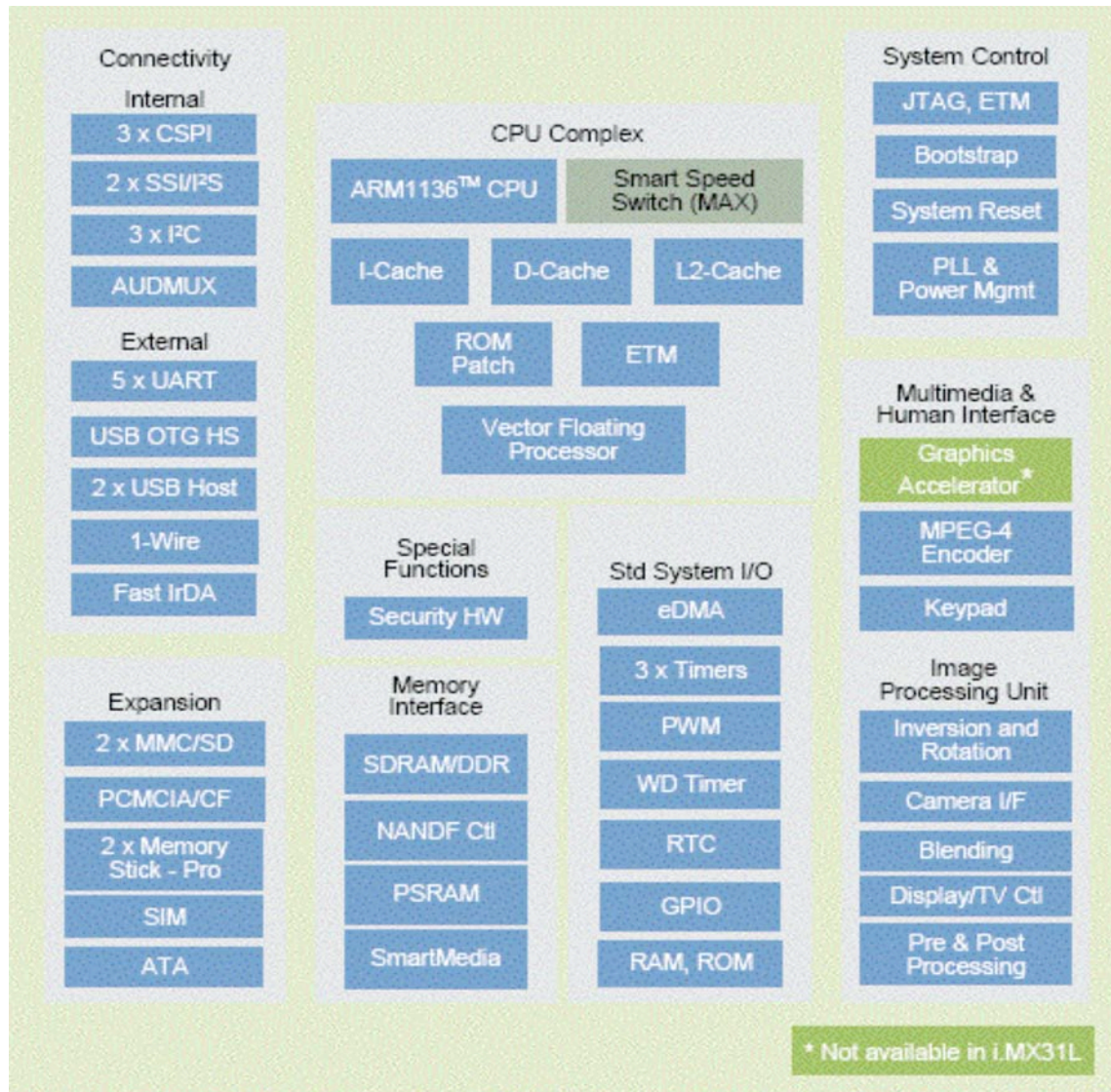


Under the Hood
620MHz CPU
128MBytes RAM
8GBytes storage
Runs cold, no fan
Recharge daily

Apple iPhone and iPod Touch

Full function web browser, Multitouch
Powerful, usable, networked applications
Music and video playback
MacOS in your pocket
Applications (hacked and official)

Freescale i.MX31 System on a Chip



Example of
2007 era on-chip
capability

Max 250mW
Idle 2mW

Commercial Millicomputer Modules



Freescale SoM 76x59mm i.MX31

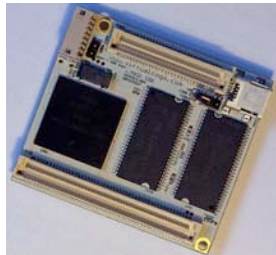


Compulab 68x58mm PXA270



Colibri 68x37mm PXA320

Triton 68x26mm PXA320



Virtual Cogs 50x44mm i.MX21



ADELAIDE 85x54mm i.MX31



Gumstix 80x20mm PXA270



(Most of these support up to 128MB RAM and cost ~\$100)
Specifications and pictures subject to owners copyright

Homebrew Mobile Club and Open Hardware

- Lets make hardware open as well...
 - Fully customizable so you can build your own phone!
- Open specifications for components
 - No NDA required for access to the full device specification
 - Allows open source Linux device drivers to be released
- Openly published schematics and PCB layouts
 - Free access to circuit designs for open components
 - Supports incremental innovation and design improvements
 - Easy re-purposing, custom PCB shapes, added devices
- Openly published mechanical packaging
 - 3D CAD files for components and assemblies on hbmobile.com wiki
 - Free access to a range of mutable layouts and case designs

myPhone Mobile Millicomputer



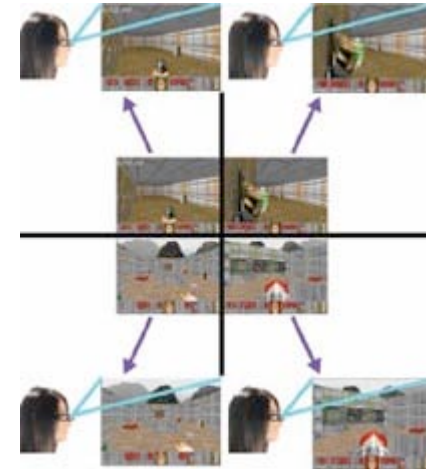
CAD design shared under creative commons on gumstix site
ABS plastic case manufactured one-at-a-time using 3D Printer
Gumstix millicomputer module mounted on phone-specific I/O PCB
Gumstix "Goliath" homebrew phone board, 4.3" LCD with touch screen
GSM/EDGE, GPS, 3-Axis Accel, 620MHz/128MB, USB Master, WiFi

Interfaces Today

MyVu



Vuzix



Head Mounted Cameras

CAPCAM



NeuroSky
Brainwave input

What's New - 2008?



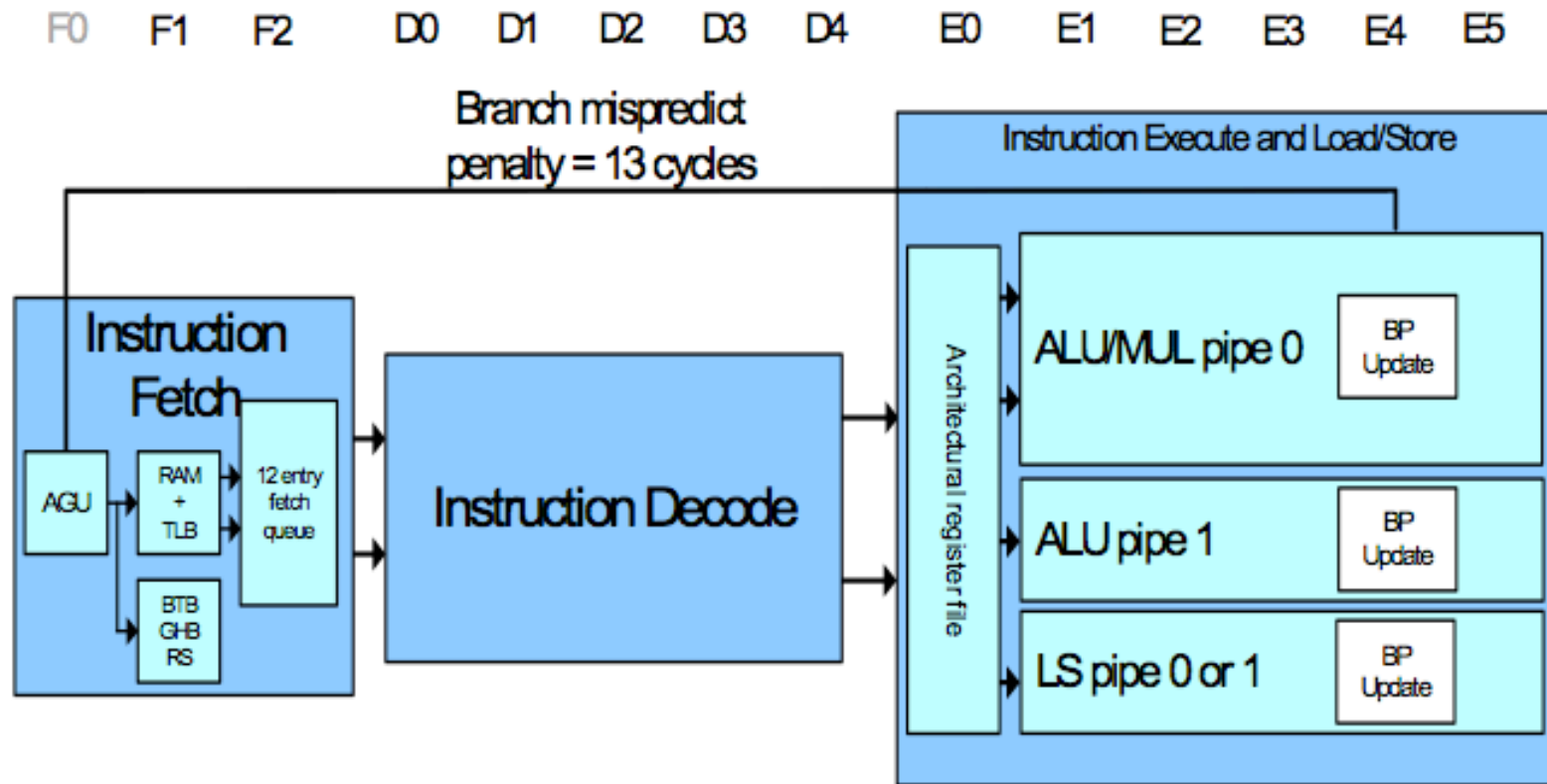
Under the Hood

Double CPU speed
Graphics Accelerator
256MB RAM
16-64GB storage
Faster networking

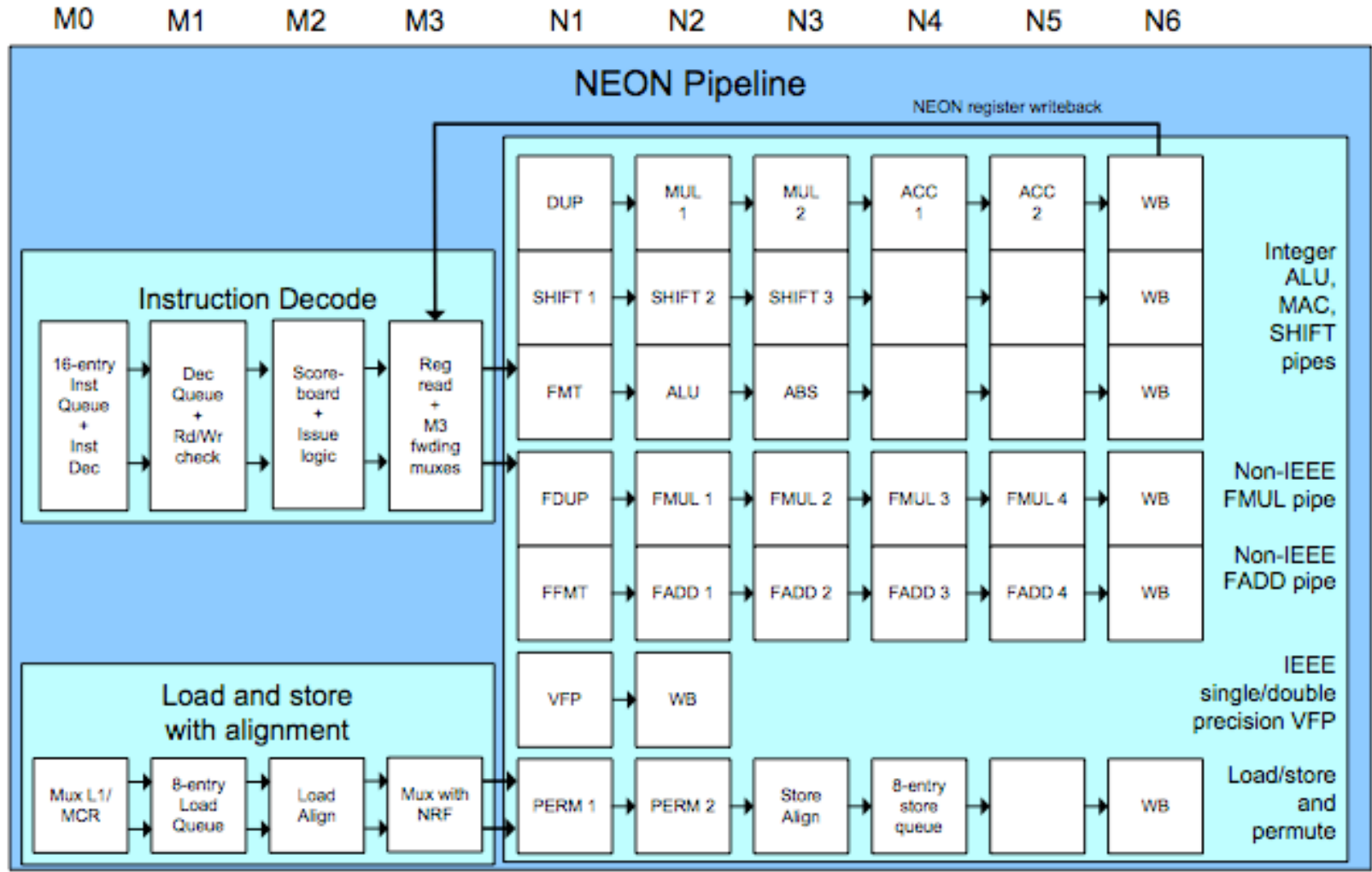
Google Android Takes on iPhone

Similar feature set, flexible implementation
Linux in your pocket
Applications (Open Software and Hardware)
Multiple phone vendors and formats

ARM Cortex A8 Pipeline



ARM Cortex A8 NEON Accelerator



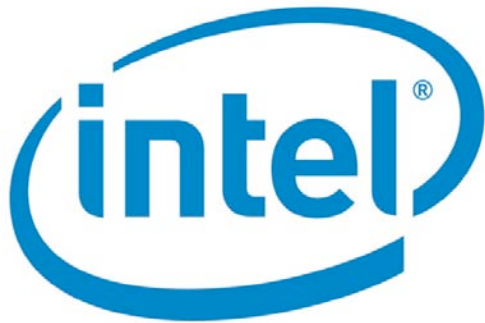
Networking Roadmap

- Cellular Network Speeds
 - ATT stated 20Mbit/s 3G HSPA in 2009 for the USA
 - 4G Long Term Evolution (LTE) in 2010 steps up to 100Mbit/s
- Wireless USB – UltraWideBand
 - The next generation – starting to appear in laptops in 2008
 - 480Mbit/s, 10 meter range
 - Wideband network, very low power, secure, low interference
 - Perfect for mobile, video etc.

Nonvolatile Roadmap

- Flash microSDHC - thumbnail sized
 - 2007 - 4-8GB, C4 speed (minimum write speed 4MB/s, max 13MB/s)
 - 2008 - 16-32GB, C6 speed (min 6Mb/s, max 20MB/s)
- Mainstream Flash
 - Faster and larger CF format alternatives
 - Solid state flash based disk alternatives with SATA interfaces
- NRAM - NanoRAM
 - New technology using metal nano-filament migration
 - Promises RAM speeds, lower power than Flash
 - A few years out, lots of current investment
- PRAM and PMC
 - Magnetic phase based technologies
 - Promises RAM speeds, lower power than Flash
 - A few years out, lots of current investment

Guesses for 2009?



Under the Hood

64bit Intel vs ARM

Faster 3D Graphics

512MB RAM

64-128GB storage

Skype video phone?



Intel x86/x64 In Pocket Phone Format

Low power Intel matches ARM Cortex A8

No fans, no burnt leg syndrome

Multiple vendors and formats

Android open to early x64 adoption

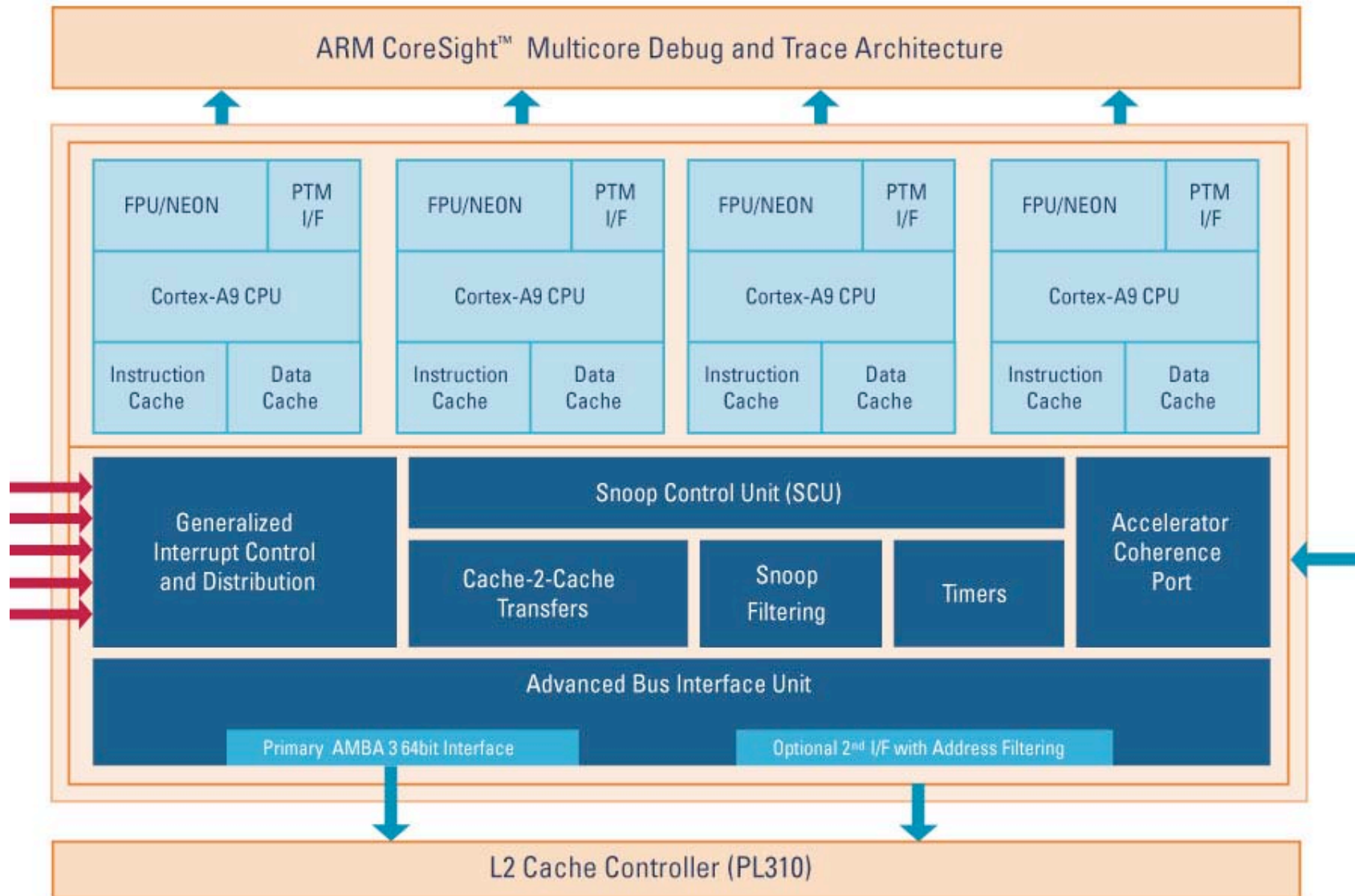
Palm, WinCE, Symbian fading...

Applications 2009

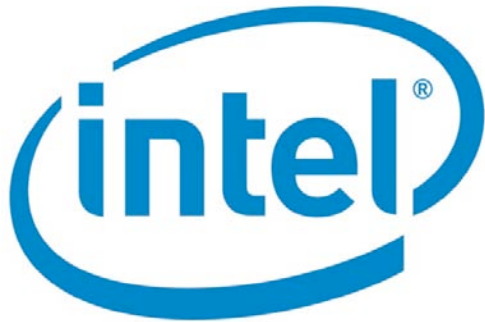
- Networked video out from your pocket over WiFi or UWB
- Wireless display anywhere
 - Play shows to WiFi-TV
 - Present to WiFi-projector
 - Wireless dock desk and pocket
 - Video-chat with friends
 - Head-up display in car



ARM Cortex A9 Multicore for 2010



Guesses for 2010??



Under the Hood
64bit Intel vs. ARM
Multi-Core CPUs
1GB+ RAM
128GB+ storage
Networked display



Intel and Arm Battle it Out
ARM Cortex A9 “8x today’s iPhone CPU”
Pocket device replaces laptop
iPhone vs. Android is the choice

Ambient Presence

- Always on 3D processed audio (OpenAL)
- Ambient “conferencing” in the back of your head
- Video presence
- Virtual world/Second Life integration
- Intimate social interaction – life sharing

Computer Assisted Telepathy!

Interfaces 2010?



MyVu Crystal

- Combine Everything
- Wireless USB
 - 3D Stereo Audio
 - Stereo Microphones
 - HD Video display
 - Split screen
 - Head Camera
 - Accelerometer
 - Compass
 - Brainwave Reader

Your Turn

What would you do with an always on, always connected laptop capacity millicomputer in your pocket?

Millicomputing in the Datacenter

- Enterprise Computing Market Patterns
- Millicomputers
- Milliclusters
- Flash Storage
- Packaging
- Application Implications
- Management Implications
- Next Steps

Enterprise Computing

- A repeating pattern
 - Mainframes replaced by Minicomputers
 - Minicomputers replaced by RISC servers
 - RISC servers replaced by PC servers
- The same objections every time...
 - “It’s a toy, not enterprise-ready”
 - “It can’t do big I/O”, “It doesn’t have big memory”
 - “Its more efficient to manage fewer bigger machines”
- What replaces the PC server?
- How do we build “green” datacenters?

How to Get \$Billion Revenue

- Commodities and \$Billion server products
 - 1970: Sell 100 Mainframes @ \$10M each
 - 1980: Sell 1,000 Minicomputers @ \$1M
 - 1990: Sell 10,000 RISC servers @ \$100K
 - 2000: Sell 100,000 PC servers @ \$10K
 - 2008: Racked Blades: 50,000 Chassis @ \$20K
 - 10 Blades per chassis @ \$2K each
 - 2008: Mobile: 10,000,000 Millicomputers @ \$100
 - OK for consumer cellphone market, not for enterprise
 - Not economical to sell individually

The Millicomputing Questions

- Do these CPUs have enough capacity to be useful for general purpose enterprise computing tasks?
- What is the growth trend for millicomputers?
- What is the price/performance, Watts/performance, rack density?
- How hard is it to make a millicomputer?
- How can vendors package tiny cheap machines into products?

What is Open Hardware? No vendor!

- Which vendor “owns” Linux? Lets make hardware open as well...
 - Fully customizable to fit whatever needs you have!
- Open specifications for components
 - No NDA required for access to the full device specification
 - Allows open source Linux device drivers to be released
- Openly published schematics
 - Free access to circuit designs for open components
 - Supports incremental innovation and design improvements
- Openly published Printed Circuit Board (PCB) layouts
 - Free access to a range of PCB designs and cost tradeoffs
 - Supports incremental innovation and design improvements
 - Easy re-purposing, custom PCB shapes, added devices
- Openly published mechanical packaging
 - 3D CAD files for components and assemblies openly published
 - Free access to a range of mutable layouts and case designs

Millicomputer's - The Coolest CPUs

- ARM processor architecture – 32bit RISC
 - Power consumption when idle - a few milliwatts
 - Maximum power consumption - 250mW to 900mW
- Vendors – Marvell, Freescale, Samsung
 - Intel XScale business was sold to Marvell
 - Very common use in mobile devices
 - e.g. iPhone, Treo, Zune, iPod, most mobile phones
 - Annual worldwide sales of billions of units
- High End Millicomputer System-on-a-Chip's for ~\$20
 - Freescale i.MX31 – 532MHz
 - Includes FPU, Multimedia and 3D acceleration for peak 250mW
 - Marvell PXA270 – 624MHz
 - Marvell PXA320 – 806MHz

Open Millicomputers

- Gumstix – PXA270 based schematics and PCB layout for large range of IO device modules provided at www.gumstix.com under creative commons license
- Gumstix Goliath - Embeddable Open Phone Motherboard
 - GSM/GPS/Touchscreen/Linux 2.6
 - Currently porting OpenEmbedded and OpenMoko Applications
- OPiuM – custom ARM based design with 256MB RAM
 - Specified as open hardware for use by Silicon Valley Homebrew Mobile Phone Club - slow progress so far...
- Benchmarks
 - <http://docwiki.gumstix.org/Benchmarks>

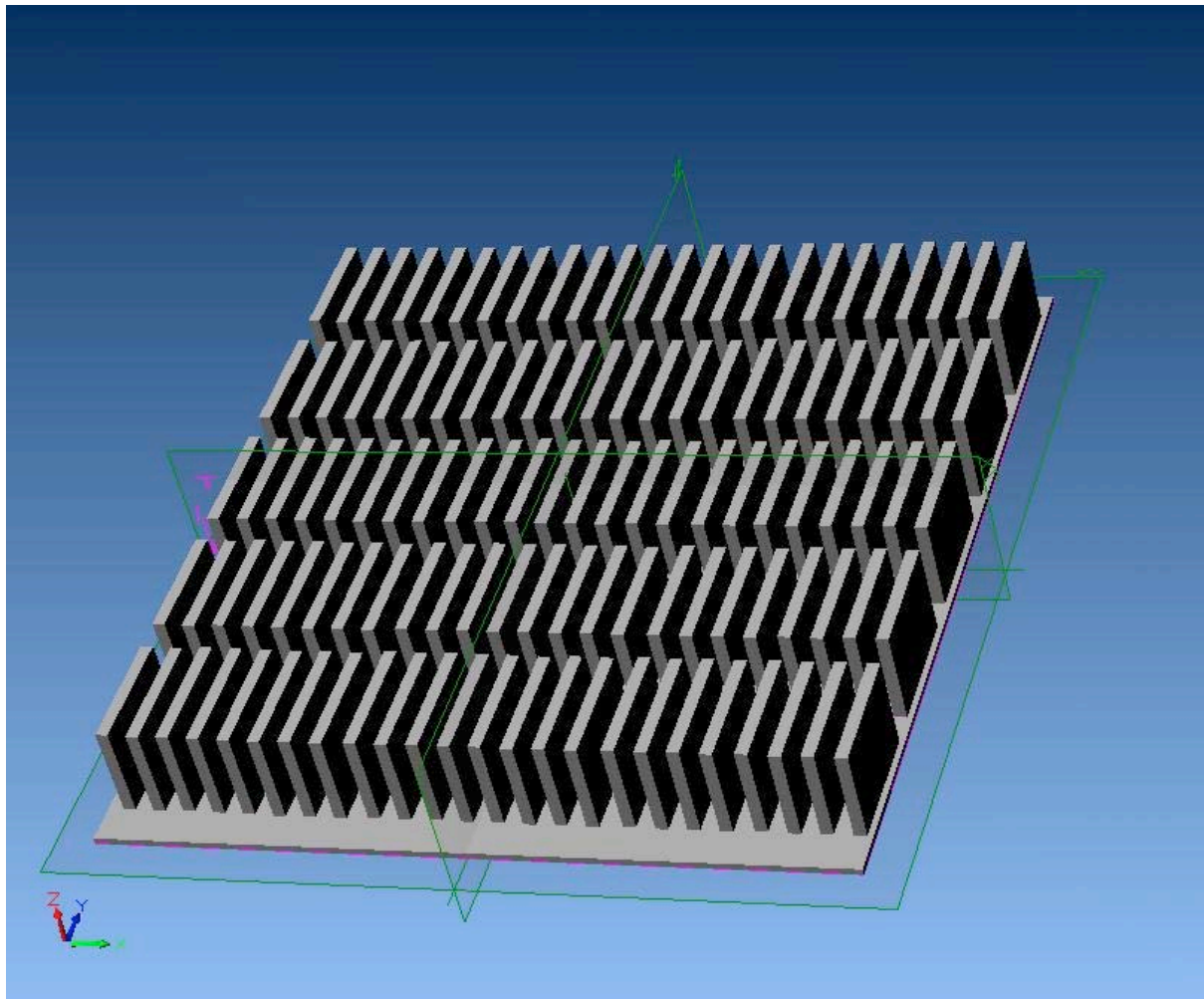
Roadmap

- All System on a Chip Designs under 250mW
- 2007 ARM Single Issue
 - 534-624MHz Common, 806MHz
 - iPhone is Samsung ~600MHz 128MB RAM
- 2008 ARM SuperScalar Cortex A8
 - Qualcomm Scorpion or TI OMAP 1GHz
- 2009 Intel x86 based competition?
- 2010 ARM Four Core SuperScalar Cortex A9
 - Performance claim “8x iPhone”

Design Our Own Millicomputer

- The following presentation is speculative
 - It describes an Enterprise Millicomputer Architecture
 - A few people HomeBrewing in their spare time
 - No mainstream vendors are involved (or needed?)
 - Design specifications are subject to change
 - We may end up building nothing or a completely different design!
 - No timescales or commitments
- Additional Motivation?
 - Reduce global warming by accelerating move to millicomputing

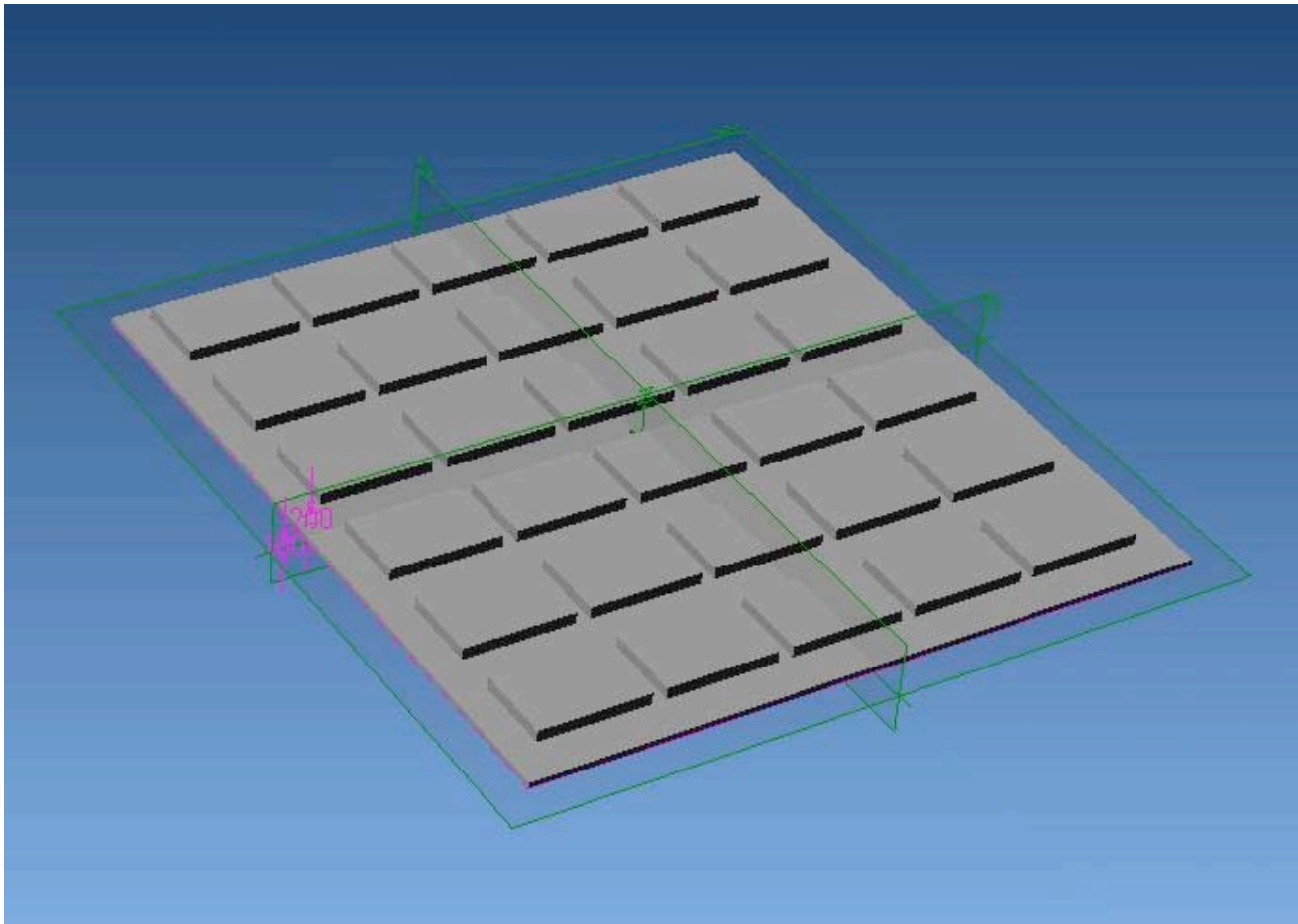
Enterprise Millicomputer Vertical Packaging



Example 1U Server package

5x24 array of modules the same volume as standard 1U enterprise motherboard

Enterprise Millicomputer Horizontal Packaging



Thin stackable 5x6 array of modules same shape as 1U enterprise motherboard

Networking

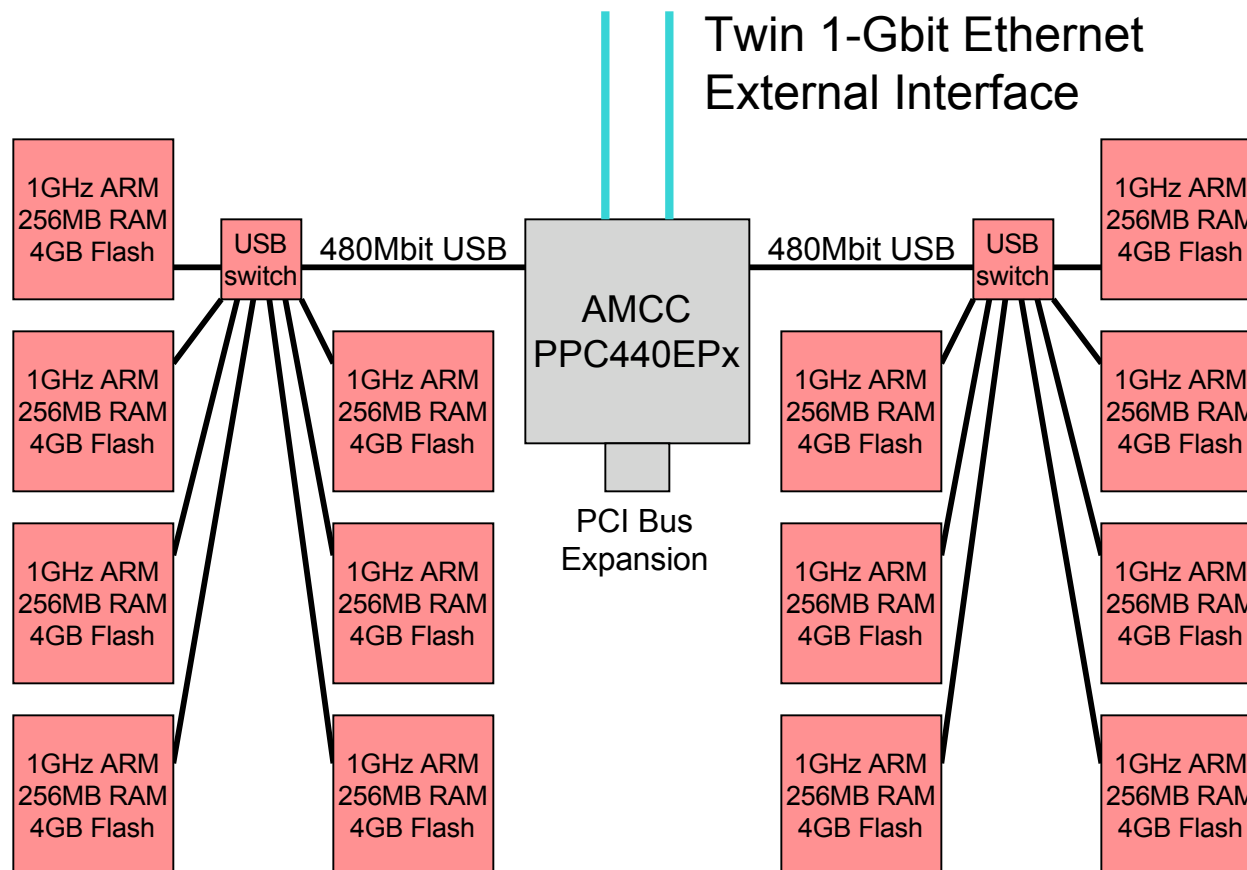
- Ethernet required for external connections
 - Power draw about 1W per 1 Gbit ethernet port
 - More than the CPU, too much per module
 - Configure an Ethernet gateway per cluster
 - Implement load-balancer functions in gateway
- CPU has on-chip USB 2.0 480Mbit/s
 - Use USBNet transport to route to Ethernet
 - Use High speed, low power 8-port 480Mbit/s USB switches

Gateway and Load Balancer

- Needs Ethernet, USB, perhaps PCI interfaces
- A few watts needed to drive Gbit Ethernet
- AMCC PPC440EPx
 - 400MHz PowerPC system on a chip @ 3W
 - Dual 1Gbit Ethernets
 - Dual 480Mbit USB2.0
 - PCI Bus Interface
- Linux Load Balancer Open Source choices
 - Haproxy, XLB, Balance, Ultra Monkey 3, vrrpd

Enterprise MilliCluster

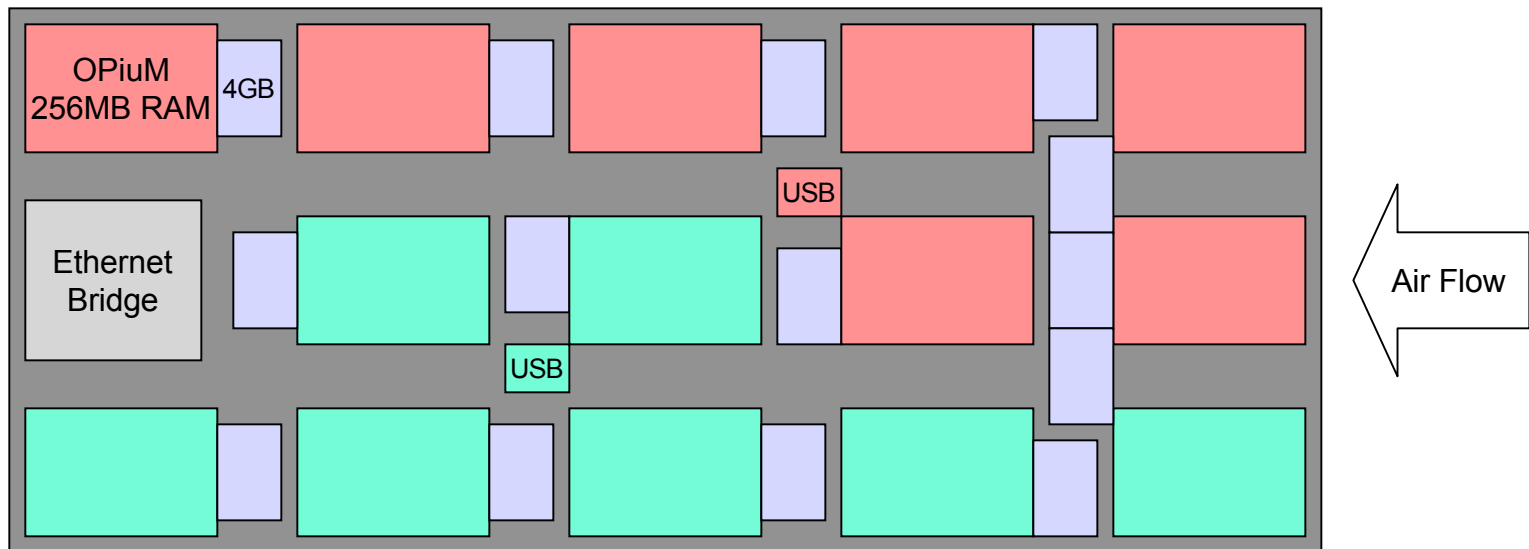
14 OPiuM Millicomputer modules behind Ethernet Bridge/Load Balancer
1 Gbit/sec redundant network, 7.5 GHz CPU, 3.5 GBytes RAM, 56 Gbytes Storage
5.5" Wide x 12" Deep x 0.4" High - 3 Watts Idle, 20 Watts Peak, no heat sinks!



Enterprise MilliCluster Packaging

Stack Side by Side Four Deep in 1U Package

Top view of One MilliCluster. Ethernet Bridge at Rear of Package
14 OPiuM 1GHz ARM Modules and microSD card mounts



Cross Section Through 1U Package Showing Eight MilliClusters, Rear Panel Has 16 x 1 Gbit Ethernet Ports



Enterprise Millicomputer Spec Overview

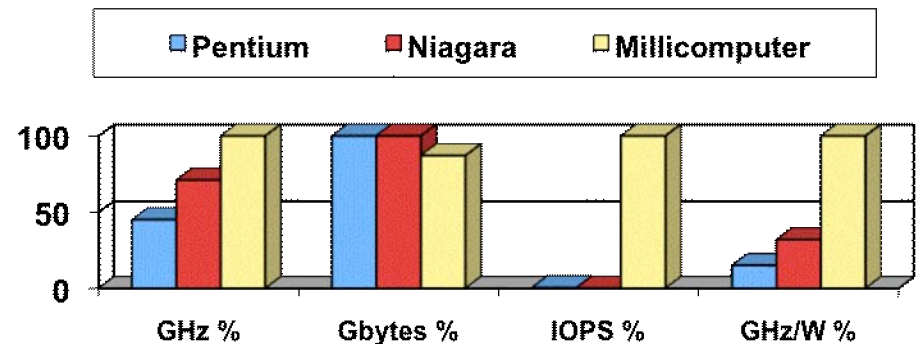
- Standard 1U Enterprise Server Package contains Eight MilliClusters
- Density – 112 OPIUM modules per RU, 4704 modules in 42RU rack
- Power – Peak 160 Watts/RU, Idle 24 Watts/RU, Peak ~6.7KW/Rack
- CPU - Performance total 112 GHz/RU, 28 GBytes/RU RAM
- Network - 8 Load balancer/bridge-routers per RU
 - 8 Gbits/RU module bandwidth on 16 redundant Gbit ports
 - Ethernet switch could be added to design to reduce port count
- Storage – microSD flash memory socket at each module
 - 4 GB microSD for very low cost, 8 GB for capacity, 16-32 GB in 2008
- Optional Extras
 - Disk – modules all include ATA disk controller if needed
 - Graphics – 1GHz ARM modules include OpenGL 3D graphics engine
 - Display – modules all include LCD display driver, touch screen
 - I/O – modules include multiple USB/serial interfaces etc.

The Flashiest Storage

- Per-module Flash in microSDHC format
 - <http://www.getflashmemory.info/category/microsdhc/>
 - 8-32 GBytes of NAND Flash in one microSDHC
 - Streaming read and write performance ~20MByte/s
 - NO SEEK! Random access at 1000's of IOPS
 - 112 x 8 GB = 896 Gbytes/RU, 37.6TB/Rack
 - 112 x 20 MB/s = 2240 MB/s/RU, 94 GB/s/Rack
- Per-module “Spinning Rust” Disks?
 - One ATA disk interface connector per module
 - Route one module per MilliCluster to connector
 - Connect four MilliClusters to disks or larger SSDs

Packaging Comparisons in 1U

- Sun x4150 Pentium1U - ~400W
 - CPU performance is probably double at the same GHz
 - 3.2 GHz eight cores x 2 = 50GHz, 32 GB RAM - \$14K list
- Sun T5120 Niagara 1U – ~250W
 - 1.2 GHz 8 core/64 threads = 80 GHz (optimistic), 32 GB RAM - \$24K list
- Enterprise Millicomputer 1U – 1GHZ ARM Cortex A8 based – ~160W
 - 1G Hz x 112 = 112 GHz, 28GB RAM - \$15K?
- Millicomputer Networking
 - Higher Network bandwidth
 - No external Load balancer
- Millicomputer Storage is No Contest!
 - 4-8x146GB disks >1000 IOPS vs. ~500000 IOPS, 896 GB Flash



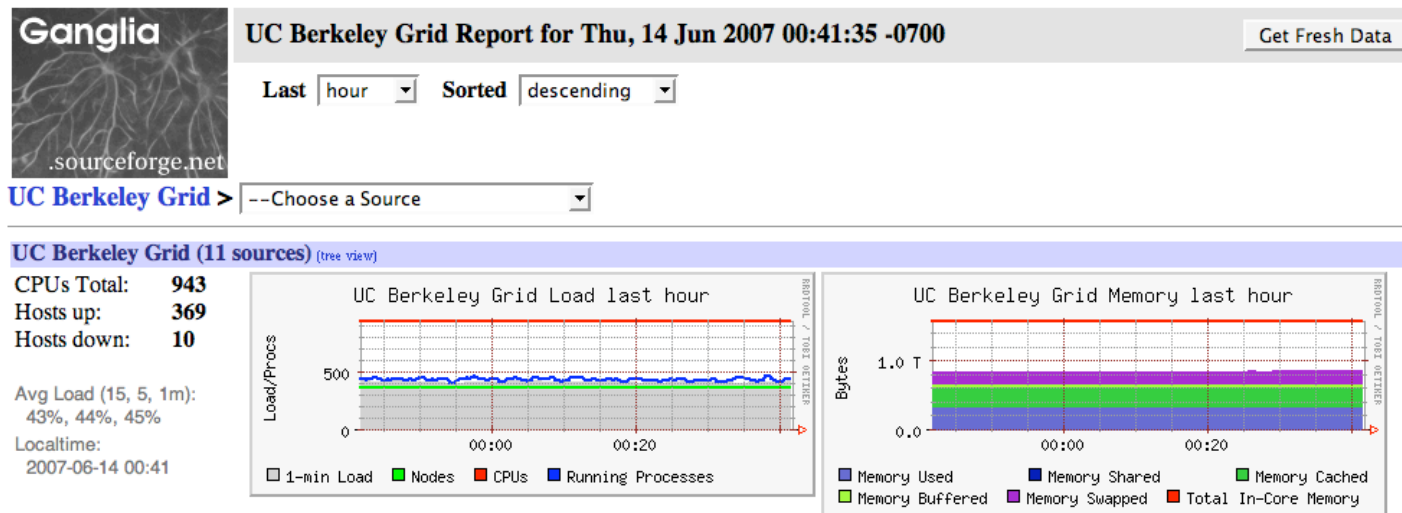
Prices from www.sun.com June 2008
Actual Performance benchmarks still need to be measured!

Software Implications

- Application memory size limit: 256MB
- Similar to mainstream systems from 2001
 - Sizes will catch up eventually
 - Classic “disruptive innovation” characteristic
 - Suitable for web content delivery
 - Especially intense random storage access
 - Static content and cache servers
 - Horizontally scaled MySQL services
 - Small Java applications
 - Video wall “Cave” display driver
 - Aggregate modules to drive tiled video outputs

Management Implications

- Large scale “grid” of small systems
 - Needs lightweight monitoring agent
 - Needs aggregation tools
 - Load balancer integration/awareness
- Ganglia? <http://ganglia.sourceforge.net/>



Summary

- For Similar 1U package, similar cost per package
 - Power less than a Niagara, less than half an Opteron system
 - Total RAM capacity similar
 - Raw CPU GHz higher, GHz per Watt much higher
 - Flash storage is 1000x faster for both random/sequential IOPS
- Applications that can be broken into small chunks
 - Small scale or horizontally scalable web workloads
 - Legacy applications that used to run on 5 year old machines
 - Graphical video walls and storage I/O intensive applications

Next steps

Performance and power benchmarking and validation

Seek out collaborators

acockcroft@netflix.com

<http://www.millicomputing.com>

Build prototypes...

Save Power!