Hardware design for Cloud-scale datacenters

Kushagra Vaid
GM, Server Hardware Engineering, Microsoft
kvaid@microsoft.com

USENIX LISA14
# Public Cloud

## Common Use Case Scenarios *

<table>
<thead>
<tr>
<th>Cloud Native Applications</th>
<th>Web Based Business Applications</th>
<th>e-business Hosting and Applications</th>
<th>Test Development Prototyping</th>
<th>Disaster Recovery / Business Continuity</th>
<th>Batch Computing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Architected specifically for Cloud</td>
<td>Custom and commercial, general purpose</td>
<td>SaaS Apps, Commerce Sites</td>
<td>Project oriented resourcing, timing</td>
<td>Storage &amp; DB replication, failover sites</td>
<td>HPC, data analytics, modeling</td>
</tr>
<tr>
<td>Growing fast</td>
<td>Predictable bursting, usage, seasonality</td>
<td>Unpredictable bursting, peak intervals</td>
<td>Unpredictable bursting, demand</td>
<td>Provisioning for growth/scale</td>
<td>Large on-demand capacity</td>
</tr>
</tbody>
</table>

* Potential for mixed mode PaaS and IaaS optimization of Applications

* Source: Gartner, Lydia Leong, March 2012
5.8+ billion worldwide queries each month

2.4+ million emails per day

8.6+ trillion objects in Microsoft Azure storage

1 in 4 enterprise customers

250+ million active users

400+ million active accounts

48+ million users in 41 markets

50+ million active users

50+ billion minutes of connections handled each month

8.6+ trillion objects in Microsoft Azure storage

200+ Cloud Services

1+ billion customers · 20+ million businesses · 90+ markets worldwide
## Server scale implications: Design

<table>
<thead>
<tr>
<th></th>
<th>&lt;10K SMB/Enterprise</th>
<th>100K Hosters</th>
<th>1M Cloud-Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong># SKUs</strong></td>
<td>Several</td>
<td>Limited</td>
<td>Extremely limited</td>
</tr>
<tr>
<td><strong>Redundancy model</strong></td>
<td>Hardware based (Hot-*)</td>
<td>Software based (Local datacenter)</td>
<td>Software based (Geo-distributed)</td>
</tr>
<tr>
<td><strong>HW availability</strong></td>
<td>99.999% or higher</td>
<td>99.9% - 99.999%</td>
<td>99% - 99.9%</td>
</tr>
<tr>
<td><strong>HW type</strong></td>
<td>Enterprise SKU</td>
<td>Off-the-shelf design, custom integration</td>
<td>Custom designs, custom integration</td>
</tr>
<tr>
<td><strong>Infrastructure co-design</strong></td>
<td>None</td>
<td>Limited integration with Datacenter and Network</td>
<td>OS, Datacenter, Server and Network tightly integrated</td>
</tr>
</tbody>
</table>
## Server scale implications: Operations

<table>
<thead>
<tr>
<th>Break/fix support</th>
<th>&lt;10K SMB/Enterprise</th>
<th>100K Hosters</th>
<th>1M Cloud-Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>24 hours x 7 days</td>
<td>8 hours x 5 days</td>
<td>Up to 1-2 weeks</td>
</tr>
<tr>
<td>Issue triage model</td>
<td>IT admin</td>
<td>Some automation, Admin support</td>
<td>Fully automated, Machine learning</td>
</tr>
<tr>
<td>OOB HW management</td>
<td>Full command set, BMC required</td>
<td>Basic feature set, BMC required</td>
<td>Power On/Off only, No BMC</td>
</tr>
<tr>
<td>Management domain scale</td>
<td>100’s of servers</td>
<td>1000’s of servers</td>
<td>10’s of 1000’s of servers</td>
</tr>
<tr>
<td>FRU granularity</td>
<td>Hot-swappable components</td>
<td>Component replacement</td>
<td>Entire server replacement</td>
</tr>
</tbody>
</table>
Cloud storage overview

Append-only distributed file system
Replication within a stamp

Extent:
- Unit of replication
- Sequence of blocks
- Size 3GB
Cloud storage design tradeoffs

A single network I/O for a write entails:
• 3 copy I/Os for initial commit (with responses)
• 18 writes for 14+4 erasure code
• Garbage collection

A single read can entail reading from multiple erasure coded extents
Query performance is measured as an aggregate of ALL compute nodes
Machine learning techniques used for ranking algorithms
Search app is very CPU intensive

Source: “Web search using mobile cores, ISCA 2010”
Cloud workloads are different!

Examples:

• Several opportunities for offloading compute intensive tasks
• Read-mostly, write-once per day
• Sequential write streams for object stores
• Synchronous replication for data durability; No RAID
Goldilocks problem applied to Cloud servers

Performance
Customization
Uniformity
Cost
Simplicity
Reliability
Agility
Power

Infrastructure design challenges
Other considerations ...

Architecture should be adapt to variety of cloud workloads

- E.g. Dense virtualization, Big data, HPC, Mailbox, Gaming, Media Streaming
- Varying requirements for Compute, Storage and Networking
- Allow for both generality and specialization in the same design

Support for global datacenter operating environments

- Seamless integration with installed IT footprint, e.g. EIA 310D 19” rack compatibility
- EMI compliance (e.g. CISPR, ANSI, IEC), Safety Standards (e.g. UL, IEC, CSA)
- Electrical distribution and cooling
Cloud scale servers: **Design Principles**

- Standardization & Modularization
- Design Simplicity
- Operations Excellence
OpenCompute Project contribution

Open CloudServer (OCS) design

Open Source Code
Chassis management
Operations Toolkit

Specifications
Chassis, Blade, Mezzanines
Management APIs
Certification Requirements

Mechanical CAD Models
Chassis, Blade, Mezzanines

Board Files & Gerbers
Power Distribution Backplane
Tray Backplane

http://www.opencompute.org/wiki/Server/SpecsAndDesigns
OCS key features

Shared infrastructure for efficiency and TCO optimization

- Power delivery, mechanics, thermals/cooling, management
- Optimized for mass contract manufacturing and assembly
Chassis features

Blind-mated signal connectivity

- Simplified installation and repair
- Cable free design for significantly fewer operator errors during servicing
- Reduces need for cabling reseats

Network Repairs

- 25% H/W Replaced
- 75% Reseated

Blind-mated connectors (12V Power, Ethernet, SAS, Management)
HDDs are #1 failure item
- AFR increases with temperature

Simplified fan control cools HDDs
- HDDs in front of hot motherboard
- Closed loop fan moderates temperatures

1 DSN 2011: Impact of Temperature on Hard Disk Drive Reliability in Large Datacenters
Secure OOB management
- Low-cost embedded x86 SoC
- REST API for machine management
- CLI interface for human operations

Hard-wired management
- On/Off to blade power cut-off circuit
- IPMI-over-serial out of band communication
- Fan and PSU control and monitoring
- Remote switch and CM power control
Security: defense in depth

Security at all layers
Hardware, UEFI, APIs, User Management

Trusted Platform Module v1.2
Blades and Chassis manager

UEFI Firmware v2.3.2
Secure BIOS and Boot

Chassis manager interfaces
TLS (SSL) and IPsec for communication encryption

User Management
Active Directory integration and authentication
BMC-Lite

✓ IPMI basic mode over Serial
✓ I2C Master (SDR)
✓ UART I/O
✓ System Event Log
✓ Power Control
× KVM, Video drivers
× Ethernet, Network Stack or SOL
× USB
× Full IPMI Command Set
OCS open source Operations Toolkit

Targeted for deployment and production support

- Collection of PowerShell scripts, applications, and utilities
- 3rd party utilities and applications can be integrated with the scripts
- Runs under Windows Dekstop and Server OS and WinPE

Features

- **Diagnostics** to reduce repair times
- **Stress tools** for higher quality validation
- Automated **updates** to eliminate human errors

http://github.com/MSOpenTech/OCSOperationsToolKit
Diagnostics example - Defect Identification

Identify defective components by physical location
- Read BMC SEL and SSD/HDD status to determine failed components
- Runs in-band on compute blade
- Identifies disks, DIMM, motherboard, adapters reporting errors

Summarize data for quick repairs
- Provide physical location of component (internal lookup tables)
- Automatically run during operation and add info to repair requests (tickets)

```
SYSTEM HEALTH DEGRADED - FOUND ERRORS...

<table>
<thead>
<tr>
<th>Location</th>
<th># Errors</th>
<th>Last Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIMM A1</td>
<td>2</td>
<td>Uncorrectable ECC - [10/15/2014 12:30:57 PM]</td>
</tr>
</tbody>
</table>
```
Diagnostics example – View Configuration

View configuration command - View-WcsConfig

```
PS C:\WcsTest> view-wcsconfig

System Info
-----------------------------------------------
Computer: WCSAZ31SUSI001
TotalMemory: 137403125760 (128.0 GiB)
TotalProcessors: 20

Software Info
-----------------------------------------------
BIOS Version: T6M_3C05
BMC Version: 4.05
OS Name: Microsoft Windows Server 2012 R2 Datacenter (Version 6.3.9600)

FRU Info
-----------------------------------------------
Chassis Part Number: X873021-001
Chassis Serial Number: QTFCTM4250001
Board Manufacturer: Microsoft
Board Name: C1020
```
## Diagnostics example – View Configuration

**View-Disk, View-Dimm, View-Nic, View-Fru, etc**

### DIMM Info

<table>
<thead>
<tr>
<th>DIMM</th>
<th>Type</th>
<th>Speed</th>
<th>Size</th>
<th>SN</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB4179</td>
</tr>
<tr>
<td>B2</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB4180</td>
</tr>
<tr>
<td>D1</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB417A</td>
</tr>
<tr>
<td>D2</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB417B</td>
</tr>
<tr>
<td>E1</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB48EA</td>
</tr>
<tr>
<td>E2</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB4918</td>
</tr>
<tr>
<td>F1</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB48E9</td>
</tr>
<tr>
<td>F2</td>
<td>Samsung M393B2G70Q0H-VK0</td>
<td>1333</td>
<td>16.0 GiB</td>
<td>37EB491E</td>
</tr>
</tbody>
</table>

### Disk Info

<table>
<thead>
<tr>
<th>Size</th>
<th>Type</th>
<th>Size</th>
<th>SN</th>
</tr>
</thead>
<tbody>
<tr>
<td>480.1 GB</td>
<td>SAMSUNG MZ7TD400HAGM-00003 FV: D8M7W3Q</td>
<td>480.1 GB</td>
<td>S16MNEADA06135</td>
</tr>
<tr>
<td>480.1 GB</td>
<td>ATA SAMSUNG MZ7TD400HAGM-00003 FV: D8M7W3Q</td>
<td>480.1 GB</td>
<td>S16MNEADA06205</td>
</tr>
<tr>
<td>480.1 GB</td>
<td>SAMSUNG MZ7TD400HAGM-00003 FV: D8M7W3Q</td>
<td>480.1 GB</td>
<td>S16MNEADA06136</td>
</tr>
<tr>
<td>480.1 GB</td>
<td>SAMSUNG MZ7TD400HAGM-00003 FV: D8M7W3Q</td>
<td>480.1 GB</td>
<td>S16MNEADA06134</td>
</tr>
<tr>
<td>4.0 TB</td>
<td>ATA WD40EJFR</td>
<td>4.0 TB</td>
<td>WD-WMD1004FZY0</td>
</tr>
</tbody>
</table>

### NIC Info

<table>
<thead>
<tr>
<th>Type</th>
<th>Speed</th>
<th>SN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mellanox ConnectX-3 Pro Ethernet Adapter</td>
<td>2 (10 gbit/s)</td>
<td>C4:54:44:56:E0:8C</td>
</tr>
</tbody>
</table>

### Mellanox Firmware Info

<table>
<thead>
<tr>
<th>DeviceID</th>
<th>FW</th>
<th>PCIe</th>
<th>UEFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>4103</td>
<td>2.30.5000</td>
<td>3.4.151</td>
<td>10.2.57</td>
</tr>
</tbody>
</table>
Diagnostics – Manage Error Logs

Check, clear, and log the Windows System Event Log and BMC SEL

- Check for hardware specific errors

View contents of BMC SEL

- View with decode of some hardware error entries

```
PS C:\WcsTest> view-wcssel
0001 [10/15/2014 12:29:15 PM] SEL cleared
0002 [10/15/2014 12:30:57 PM] DIMM A1 Correctable ECC
0004 [10/15/2014 12:31:35 PM] Voltage exceeded threshold. Sensor C4 EvtData(3-1) 0xE1D050
0005 [10/15/2014 12:31:37 PM] Voltage within threshold. Sensor C4 EvtData(3-1) 0xC3D050
```

- View without decode for raw data

```
PS C:\WcsTest> view-wcssel -NoDecode
0001 RecordType: 0x02 TimeStamp: 543E689B GenID: 2000 EvMRev: 04 SensorType: 10 Sensor: 8A Evt
0002 RecordType: 0x02 TimeStamp: 543E6901 GenID: 0001 EvMRev: 04 SensorType: 0C Sensor: 87 Evt
0003 RecordType: 0x02 TimeStamp: 543E6901 GenID: 0001 EvMRev: 04 SensorType: 0C Sensor: 87 Evt
```
System Updates

Commands...
- System Identification (Select update based on FRU/BIOS info)
- Dependency checking
- Logging
- Sequencing (can update from any version if possible)

Example: Update-WcsConfig Command
- Single command to update multiple components
- Example: Update blade BIOS, BMC, NIC and HBA FW
- Simplifies learning curve for repair technicians
Summary

- Cloud workloads are extremely diverse - need for balance uniformity with specialization

- Designing for cloud-scale based on simple principles: Modularity, Simplicity and Operations Excellence

- Open sourced server designs present opportunity to drive innovation on common platforms and leverage battle-tested tools
Q & A