

OPEN Compute Summit January 28–29, 2014San Jose



Microsoft cloud server specification Blade overview – Compute and Storage

Martin Goldstein, Principal Systems Architect Microsoft Cloud Server Infrastructure Engineering





Chassis



Microsoft cloud server features

EIA 19" Standard Rack Compatibility

Chassis 12U

- Highly efficient design with shared power, cooling, and management
- Cable-free architecture enables simplified installation and repair
- High density: 24 blades / chassis, 96 blades / rack

Flexible Blade Support

- Compute blades Dual socket, 4 HDD, 2 SSD
- JBOD Blade scales from 10 to 80 HDDs

Scale-Optimized Chassis Management

- Secure REST API for out-of-band controls
- Hard-wired interfaces to OOB blade management



Chassis components



Sidewall (right)



Compute Blade



Blade design: guiding principles

Simplicity

- No hot plug or redundant components
- Low cost, simplified management

Flexibility

- LFF SATA HDD and SFF SATA SSD

Serviceability

- Blind-mate connectors simplify server insertion and removal
- Cable-free design minimizes cable-based NTF issues

Total Cost of Ownership Density optimized for IT-PAC (container)

- deployments
- across 24 servers



Three IO card options (LAN, SAS, PCIe)

Shared chassis infrastructure amortized

Key compute blade specifications

Label -

Processor	
CPU	Dual Intel® Xeon® E5-2400 v
Core QTY	Up to 10 cores / CPU, 20 / Bla
TDP Wattage	Up to 95W
Memory	
Memory Busses and DIMM	3X memory bus / CPU, 6 / Bla
Slots	6 DIMM slots
DIMM Type / Speed	16GB, 2Rx4, 1333MHz, 1.35\
Max Capacity	192 GB / Blade
On-Board Devices	
Storage Controller	Intel® C602 PCH
SATA (HDD)	4 ports @ 3.0 Gb/s
SATA (SSD)	2 ports @ 6.0 Gb/s
Server Management	
Chipset	BMC-Lite serial thru Chassis I
Interface	REST API, CLI thru Chassis N
System Firmware	
Version	UEFI 2.3.1
Security	TPM 1.2, Secure Boot
Blade I/O	
PCI-Express Slots	One Gen3 X16 Riser
Networking	Single or Dual 10G Mezzanin
SAS	Dual 4X SAS @ 6G Mezzanir

Engineering Workshop

e Card ne Card

Manager Manager

ade

2 family ade

Compute blade highlights



Compute blade BMC-Lite

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

Command name	Reference	Туре	Fn	Cmd	Compute blade	JBOD blade
Get Device ID	20.1	Арр	OGh	01h	м	м
Set ACPI Power State	20.6	Арр	OGh	06h	м	N/A
Get ACPI Power State	20.7	Арр	06h	07h	м	N/A
Get System GUID	22.14	Арр	06h	37h	м	м
Get Channel Authentication Capabilities	22.13	Арр	06h	38h	м	м
Get Session Challenge	22.16	Арр	OGh	39h	м	N/A
Activate Session	22.17	Арр	OGh	3Ah	м	N/A
Set Session Privilege Level	22.18	Арр	06h	3Bh	м	N/A
Close Session	22.19	Ap	OGh	3Ch	м	N/A
Get Session Info	22.7	Аре	OGh	3DI	м	N/A
Set Channel Access	17.21	Арр	06h	400	м	N/A
Get Channel Access	22.2	Арр	OEh	41h	м	N/A
Get Channel Info Command	22.24	Арр	2 tih	42h	м	N/A
Set User Access Command	22.26	App	OGh	43	м	N/A
Get User Access Command	22.27	Арр	OGh	441	м	N/A
Set User Name	AN 75	Арр	OGh	45h	м	N/A
Get User Name Command	22.29	Арр	16h	46h	м	N/A
Set User Password Command	22.3	Арр	06h	47h	м	N/A
Get System Interface Capabilities	22.9	Арр	O6h	57h	м	N/A
Get Chassis Capabilities	28.1	Chassis	ODh	00h	N/A	N/A
Get Chassis Status	28.2	Chassis	ooh	01h	м	м
Chassis Control	28.3	Chassis	ooh	02h	м	N/A
Chassis Reset	28.4	Chassis	ooh	03h	N/A	N/A
Chassis Identify	28.5	Chassis	ooh	04h	м	N/A

Compute blade I/O

Flexible options to meet networking and storage requirements







Specified for 1x or 2x 10GbE •

LAN mezzanine

- PCI-Express Gen3 x8
- Single or Dual 10Gbit Ethernet
- Direct Attach Copper Cables, up to 5 meters





52 mm (2.44")

[reference designs shown]

Image: State of the state

83 mm (3.27")

SAS mezzanine

- PCI-Express Gen3 x8 •
- Eight lanes of 6Gb/s SAS •
- Direct Attach Copper Cables, 0.5 meters •





83 mm (3.55")

[reference designs shown]

105 mm (4.49")

Safety and compliance

Ready for data centers world-wide

- Microsoft requires full compliance
- Containment at blade and tray
- Chassis is contained for use in EIA racks

Safety is Microsoft top priority

UL, IEC, CSA standards among others

EMI Compliance is important

CISPR, ANSI, IEC standards to start with







Additional features

Status LEDs

- Health LED in the front
- Attention LED in the front and back

Blade Insertion and Removal

- Front access, tool-less blade extraction
- Rotate latch before engaging release lever
- Two-phase release enables in-rack shipments



Green Health: Blade OK



Amber Health: **Blade Fault**



Red Attention: Identify Blade



Manufacturing blade build-out



Bare metal

W/ motherboard

W/ HDD backplane

W/ 4x3.5" HDDs

Expansion JBOD



JBOD design: guiding principles

Simplicity

- Scale in 10 HDD blocks
- Direct attach cabling

Flexibility

- Support 14 to 84 HDD per server

Serviceability

- Blind-mate connectors simplify JBOD insertion and removal
- Cable-free design minimizes cable-based NTF issues

Total Cost of Ownership

- Short cables save cost and weight
- Shared chassis infrastructure is amortized across 24 blades



Eight SAS lanes in or Four in / Four out

- Density optimized, up to 800 HDDs / rack

Key JBOD specifications

LFF 7200 RPM 3.5" SATA				
20 TB, 10 x 2TB HDD				
30 TB, 10 x 3TB HDD				
40 TB, 10 x 4TB HDD				
50 TB, 10 x 5TB HDD				
60 TB, 10 x 6TB HDD				
8 Channel SAS @ 6G				
LSI 20-port LSISAS2X20				
Storage Server Configuration Capacities				
Compute with 4 HDDs plus				
14 HDDs				
24 HDDs				
44 HDDs				
64 HDDs				
84 HDDs				

Expansion JBOD reference design



20-lane SAS expander

- 10 internal lanes connect to LFF SATA HDDs
- 8 external lanes connect to tray backplane

Expander connects to chassis manager via RS-232 port

Managed with the same command set as the compute blade

Expander

8 Lanes SAS

Expander board details





Storage expander board



Storage HDD backplane

- Blind-mate to tray backplane (SAS, management)
- Direct connect to two 3.5" SATA HDDs
- Cable connect to two storage HDD backplanes 1

[reference designs shown]

t.t.

Cable-free attach simplifies drive replacement and eliminates NTFs caused by cable connection issues



14 HDD per server

84 HDD per server

Manufacturing JBOD build-out





W/ expander backplane

Bare metal

Engineering Workshop

W/ 10x3.5" HDDs



More Information: Technical breakouts

Technical Workshop	Presenter
Hardware Overview	Mark Shaw, Directo
Management Software Overview	Badriddine Khessib
Chassis Manager Hardware Overview	Bryan Kelly, Senior



r

, Director

Platform SW Engineer

Microsoft cloud server spec: OCP contribution

Source Code

Chassis management source code through Open Source

/// <summary> /// Gets Fan speed in RPM /// </summary> /// <param name="fanId">target fan Id</param> /// <returns>Fan speed in RPM</returns> internal FanSpeedResponse GetFanSpeed(byte fanId) {

Specifications

Chassis, Blade, Chassis Manager, Chassis, Blade, Chassis Manager, Mezzanines, Management APIs Mezzanines





Board Files & Gerbers Chassis Manager, Tray Backplane, Power Distribution Backplane

Contribution Mechanical CAD Models Chassis, Blade, Chassis Manager, Mezzanines



Microsoft datacenter resources

Microsoft Datacenters Web Site & Team Blogs

www.microsoft.com/datacenters

Windows Azure

http://www.windowsazure.com

Office 365

http://www.office365.com







Q&A





Microsoft

© 2014 Microsoft Corporation. All rights reserved. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.

