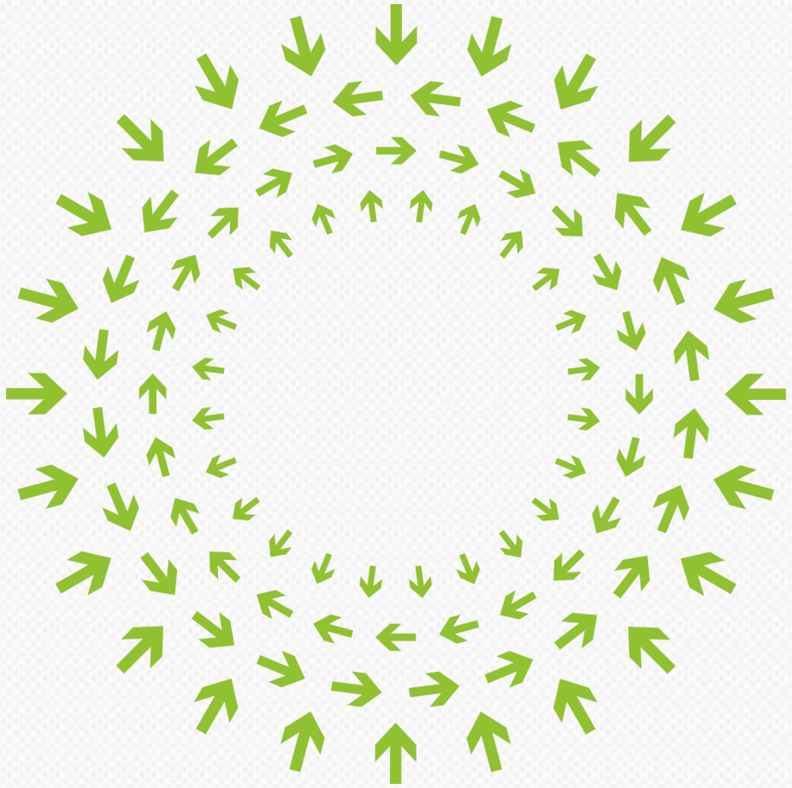


OPEN Compute Summit

January 28–29, 2014 San Jose

OPEN Compute Summit

Microsoft's cloud server specification

Chassis Manager Hardware Overview

Bryan Kelly,
Senior Platform Software Engineer
Microsoft Cloud Server Firmware Development

Microsoft

Microsoft cloud server spec features

EIA 19” Standard Rack Compatibility

Chassis

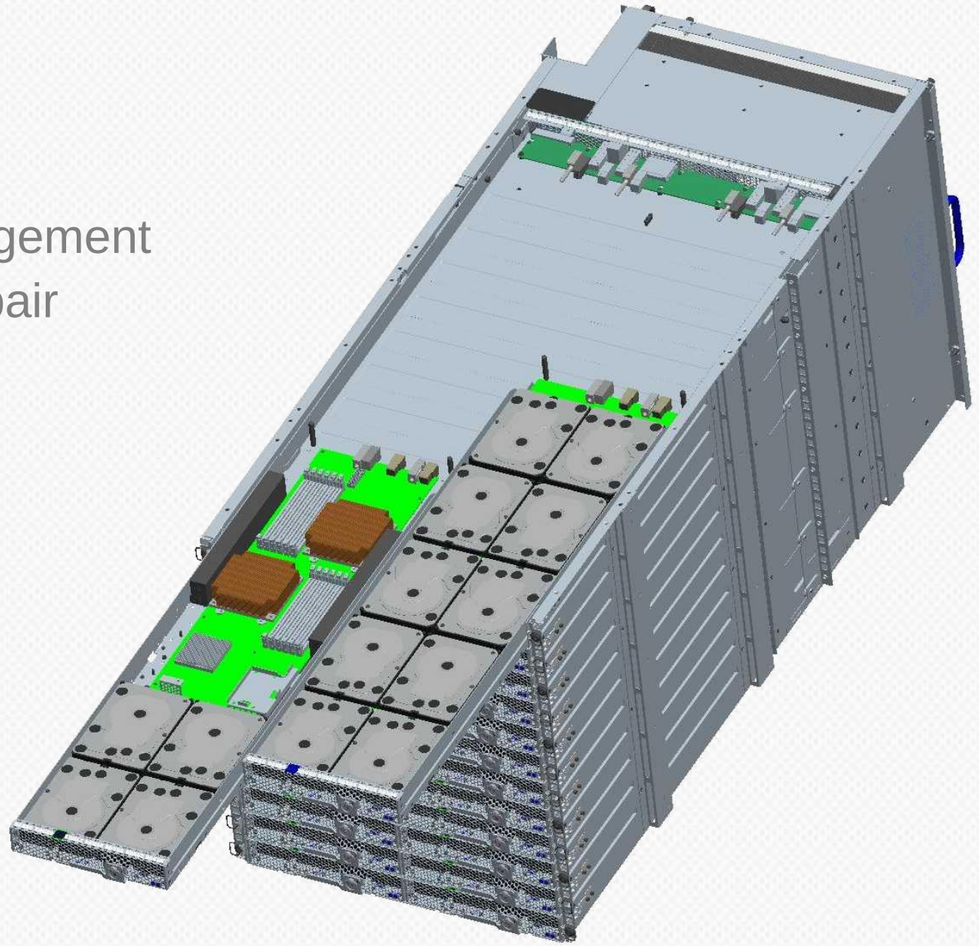
- Highly efficient design with shared power, cooling, and management
- Cable-free architecture enables simplified installation and repair
- High density: 24 blades / chassis, 96 blades / rack

Flexible Blade Support

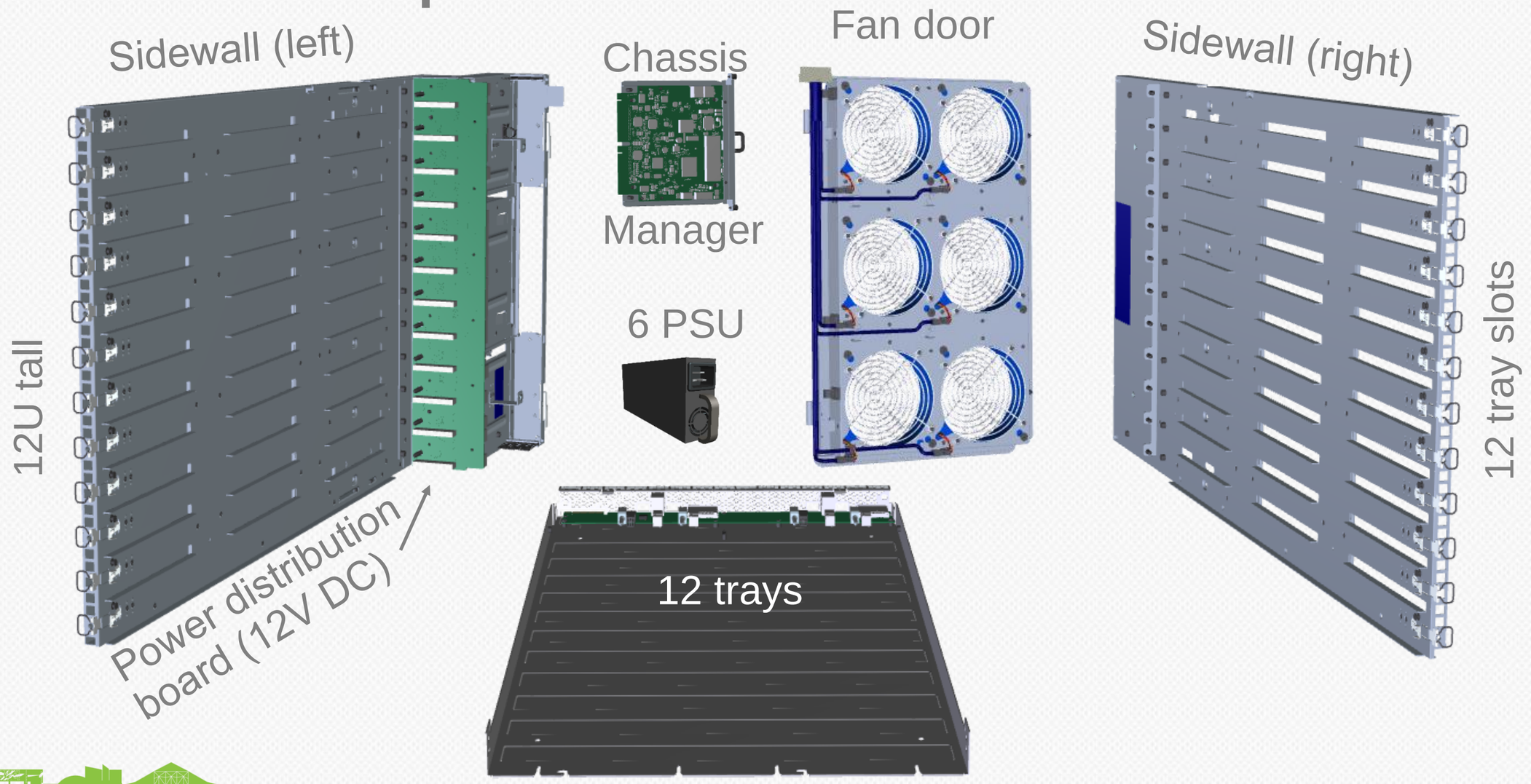
- Compute blades – Dual socket, 4 HDD, 2 SSD
- JBOD Blade – scales from 10 to 80 HDDs

Scale-Optimized Chassis Management

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



Chassis components



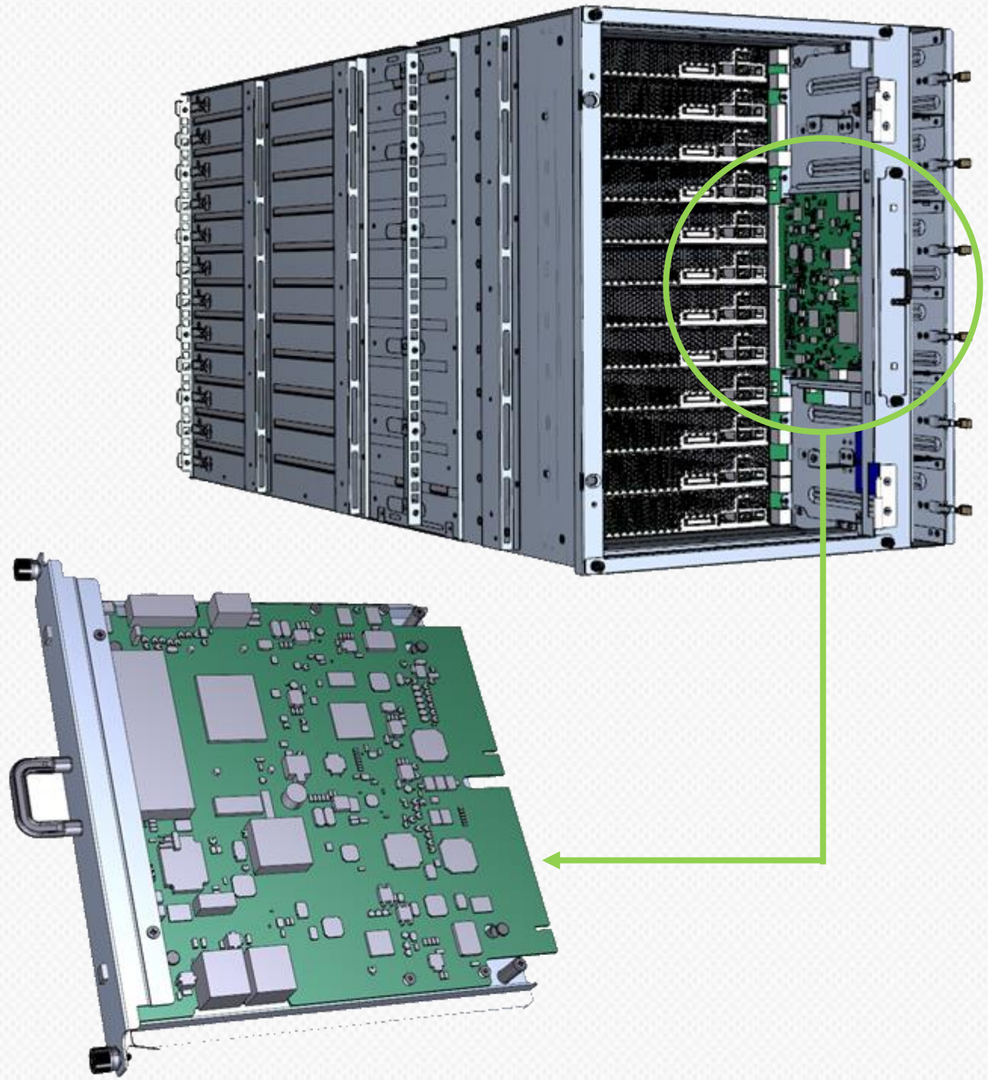
Chassis manager

Integrates into the chassis, occupies zero u-space.

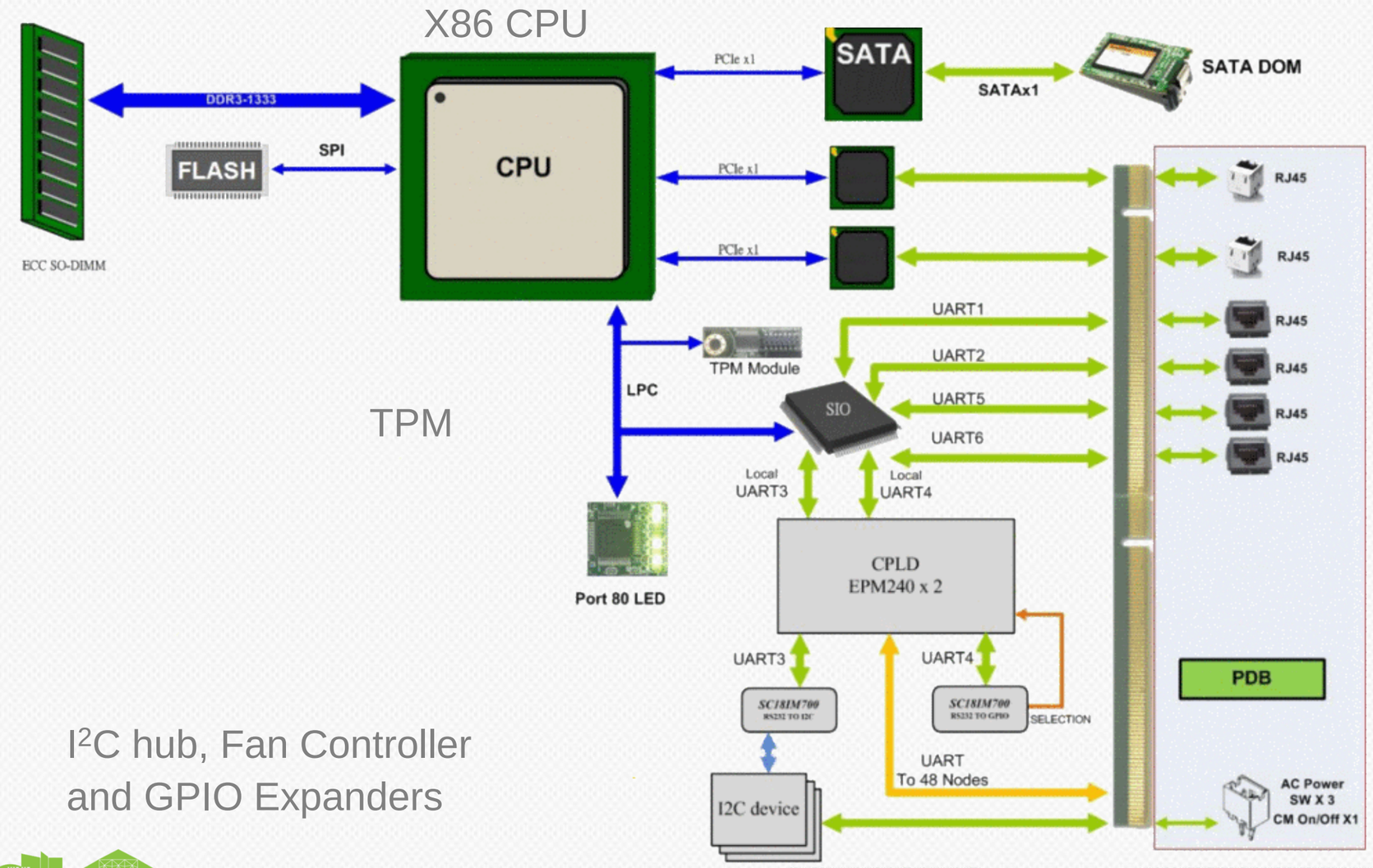
- Low powered stateless device.
- Hot pluggable
- Always On
- Embedded x86 board.

Function:

- Chassis control: OOB blades, PSU, fans and auxiliary devices



Chassis manager: features



Dual 1G Ethernet

4x External Serial Ports

2x Internal Serial Ports

3 x External 12v Power Switches

I²C hub, Fan Controller and GPIO Expanders

Chassis manager: hardware

6 x UART SuperIO

- 2 Internal
- 4 External

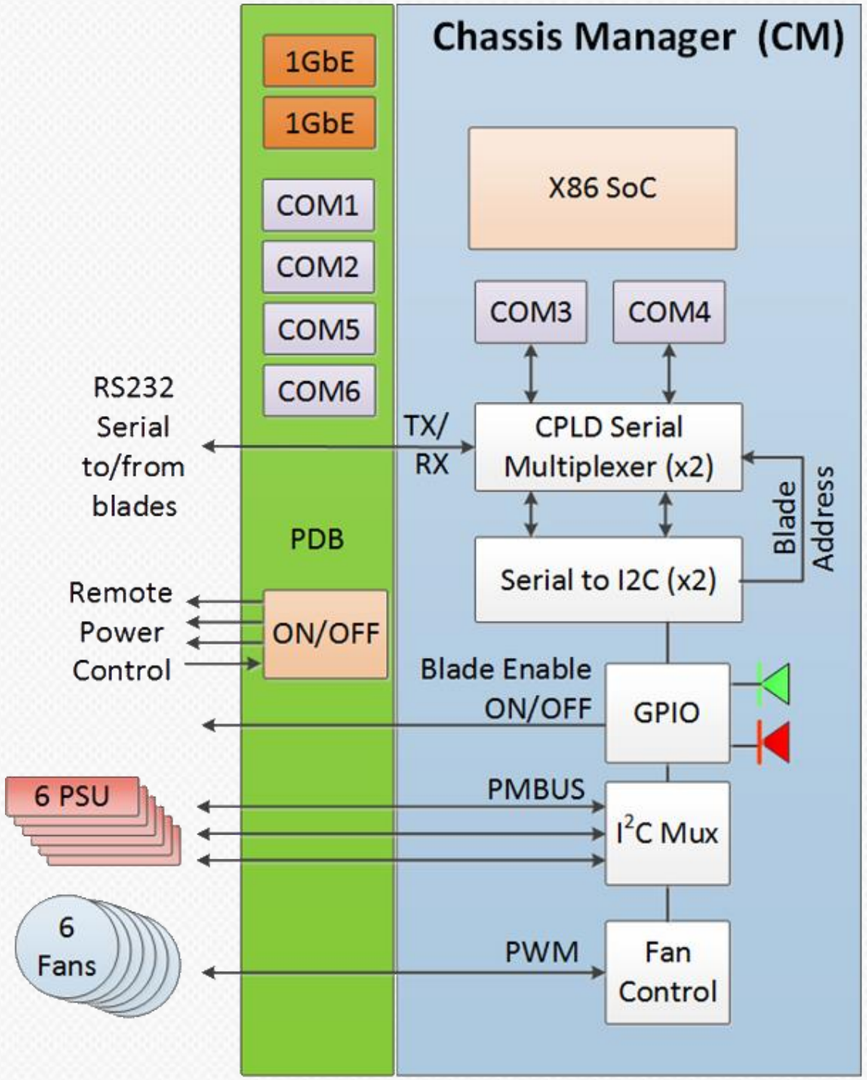
CPLD and GPIO for blade selection

GPIO extender for blade inrush controller On/Off

I2C Mux for PSU Control

Fan Controllers

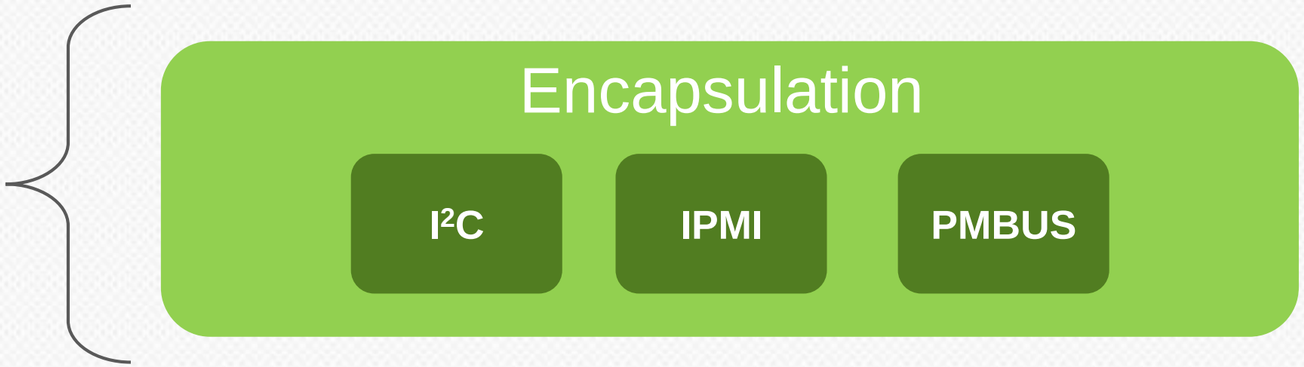
All chips UART controlled



Chassis manager: UART control

Single Driver Framework

- Encapsulates multiple protocols
- Extensible Framework
- Multi-Threaded



Simplified HW Driver Model

- Common Transport Mgmt
- Chip initialization
- Resource allocation
- Communication error control

Chassis manager: blade communication

Serial Communication

- Position defined target
- No addressing
- No custom configuration

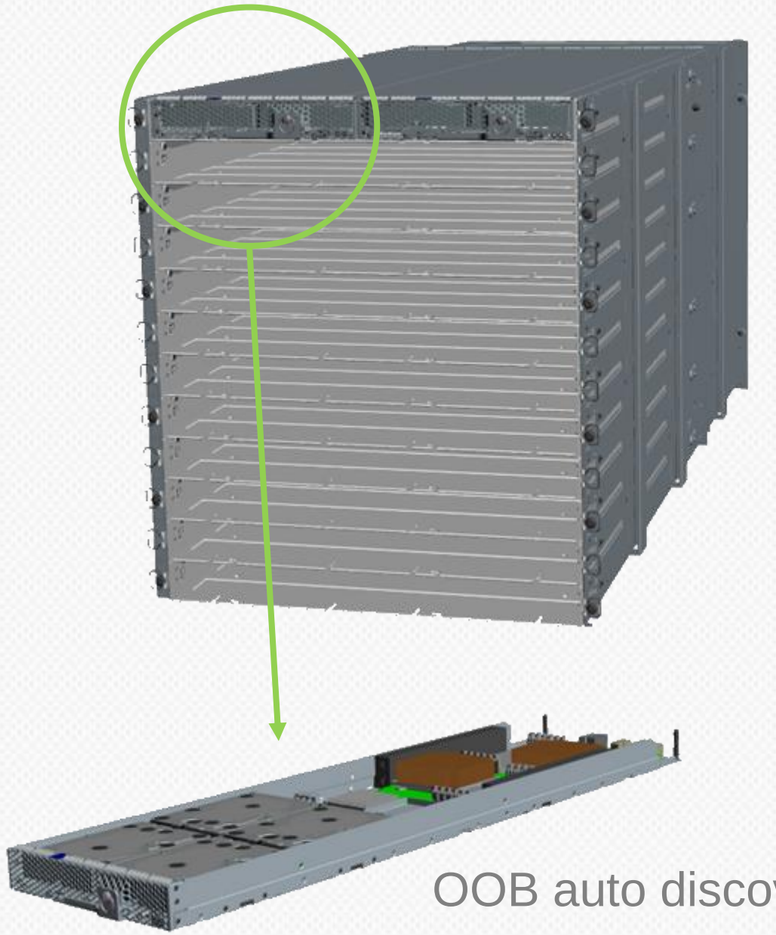
Common configuration set during manufacturing

All blades are automatically discoverable

Blade in slot 1 is always blade 1.

No intermediate switch or aggregator

BMC-Lite, no network stack.



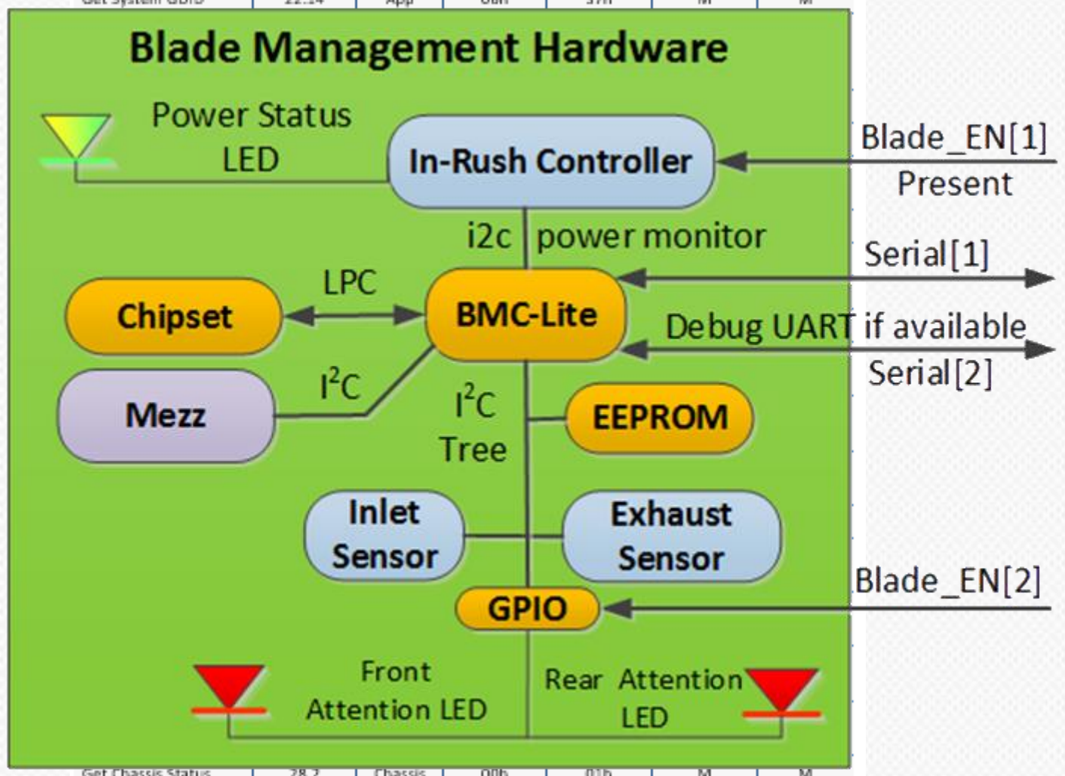
OOB auto discovery

Compute blade BMC-Lite

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~~

Command name	Reference	Type	Fn	Cmd	Compute blade	JBOD blade
Get Device ID	20.1	App	06h	01h	M	M
Set ACPI Power State	20.6	App	06h	06h	M	N/A
Get ACPI Power State	20.7	App	06h	07h	M	N/A
Get System GUID	22.14	App	06h	37h	M	M
Get Chassis Status	28.2	Chassis	00h	01h	M	M
Chassis Control	28.3	Chassis	00h	02h	M	N/A
Chassis Reset	28.4	Chassis	00h	03h	N/A	N/A
Chassis Identify	28.5	Chassis	00h	04h	M	N/A



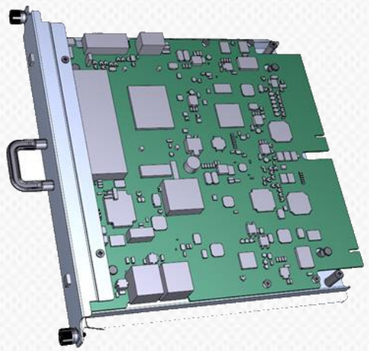
Blade communication

Communication over internal UART 4

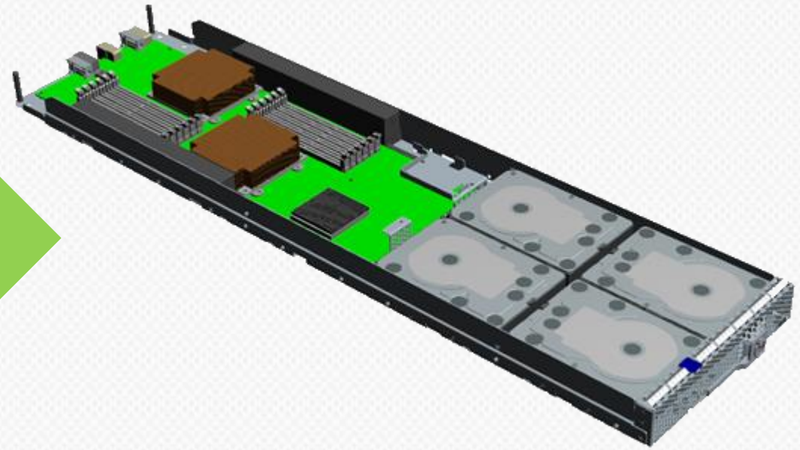
Serial OOB through the PDB (no cables) to blade BMC-Lite

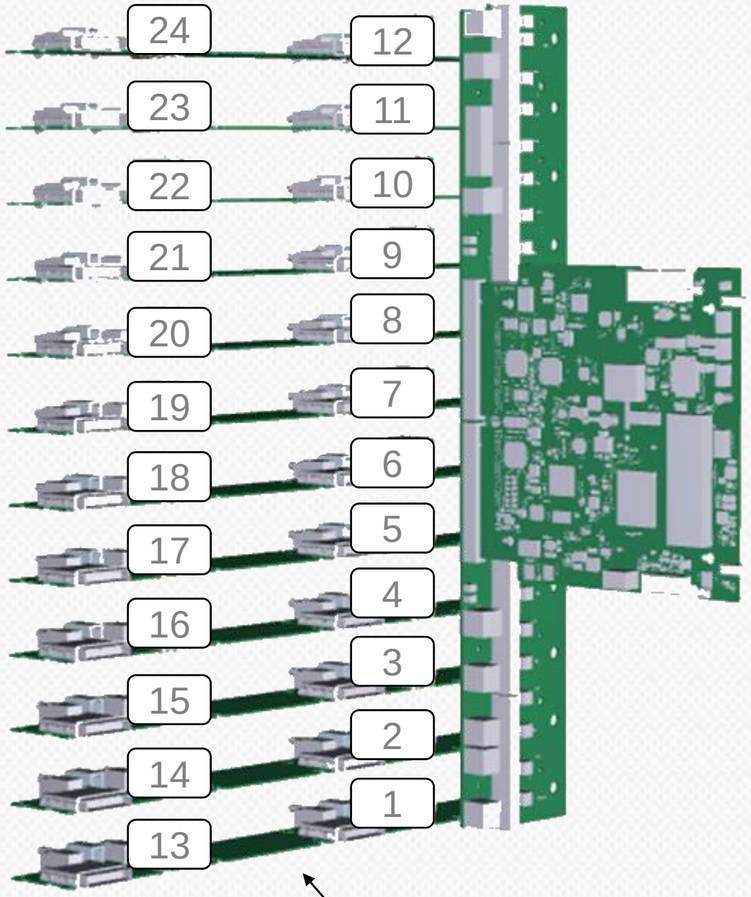
Transport Protocol: IPMI Basic Mode

Blade selection using muxing mechanism



IPMI
Serial



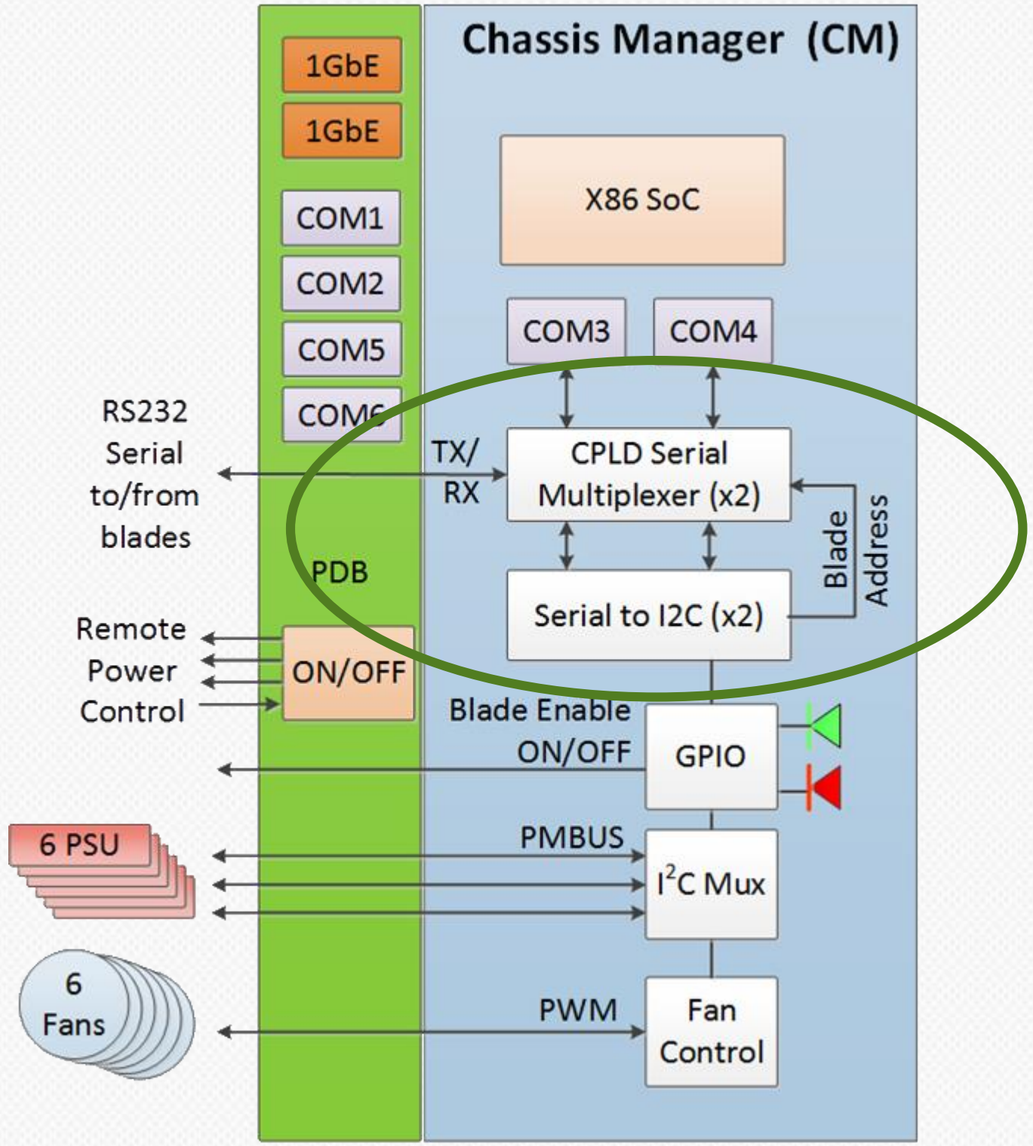


Blades 1-24

Chassis manager:
Blade selection muxing (COM4)

Mux Sequence:

1. UART Buffer Flush
2. DTR -> CLPD serial TX/RX to the GPIO
3. Blade address is written to GPIO (CPLD changes TX/RX blade selection)
4. DTR -> CPLD serial TX/RX to blades
5. IPMI request command is sent
6. BMC-Lite will respond within 100ms or a timeout occurs



OOB serial console redirection

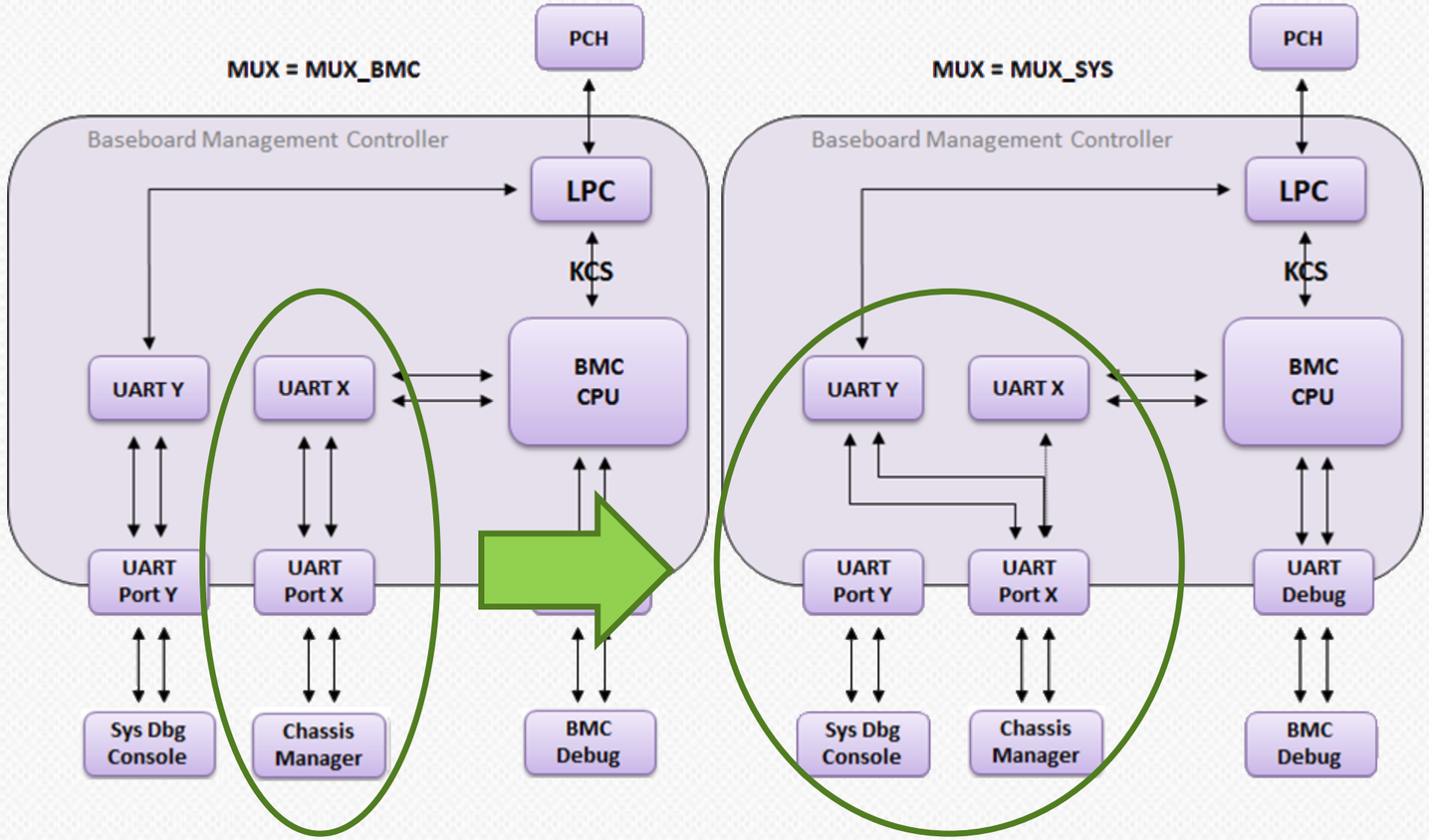
Muxing mechanism for Request/Response messaging

Serial Console Redirection is supported for debug only

During redirection CM locks the CPLD selection

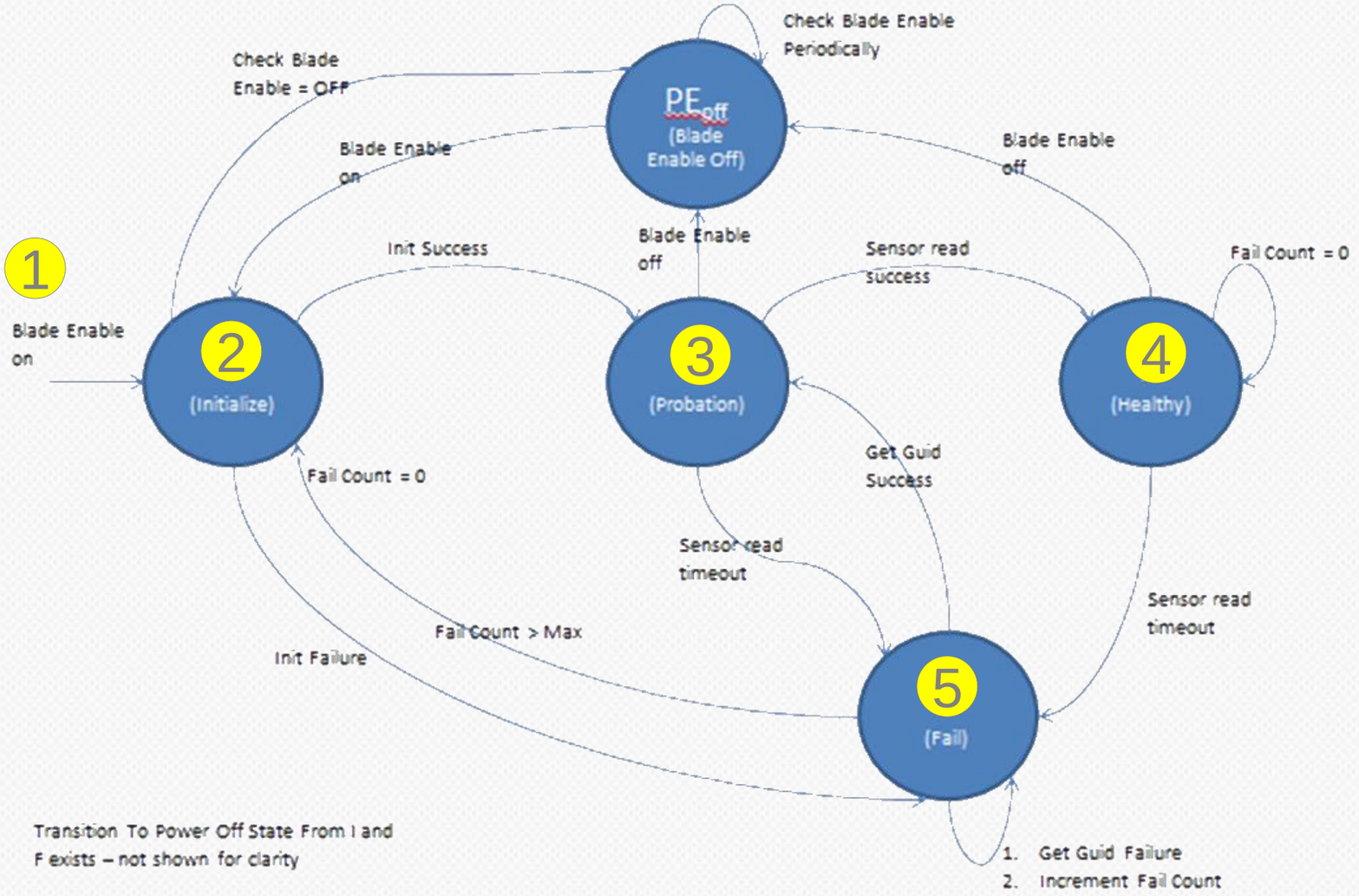
Blade polling is suspended

Fan speed increases to compensate temperature excursions



Chassis manager: blade state management

1. Power On
2. Init()
3. Probation
4. Healthy
5. Fail

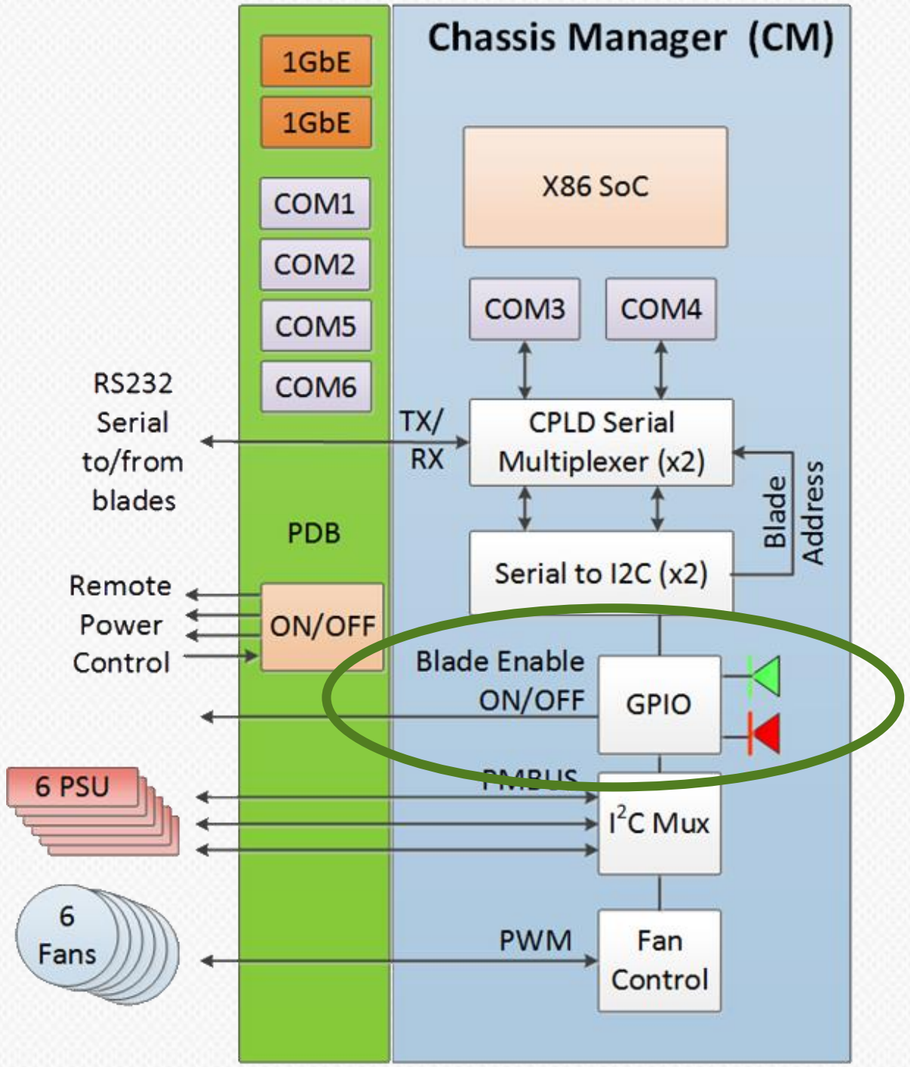


Cloud server spec: power control

GPIO expander on UART 3 provides power control to blade inrush controllers using blade enable pins.

By switching off the hot swap controller, power is disconnected from the blade

Software rules enforce controlled off and on sequence



Chassis manager: PSU control

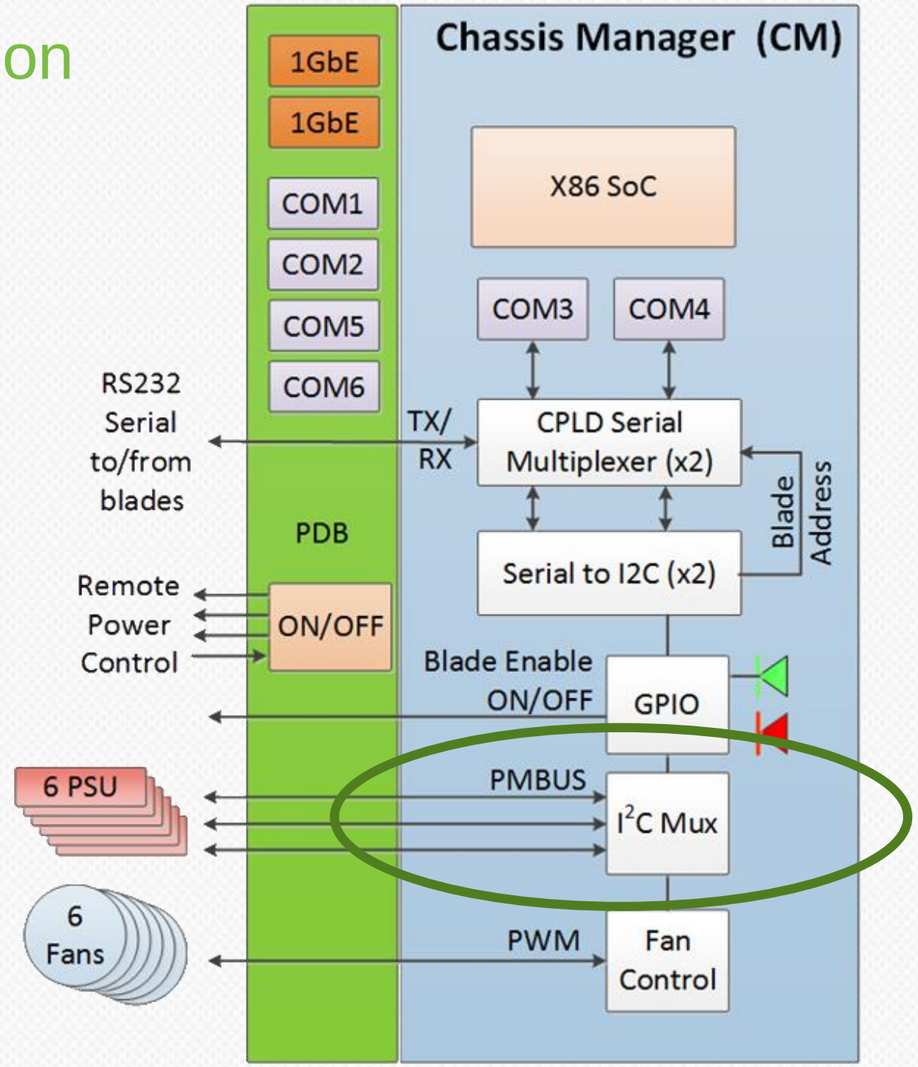
PSUs report health, status and power consumption

Soft faults automatically recovered

Faults reported and the attention LED is illuminated

Communication occurs over the UART 3 I2C hub using PMBUS protocol.

PSU off and on Control



Chassis manager: fan control

Variable fan speed based on blade requirements

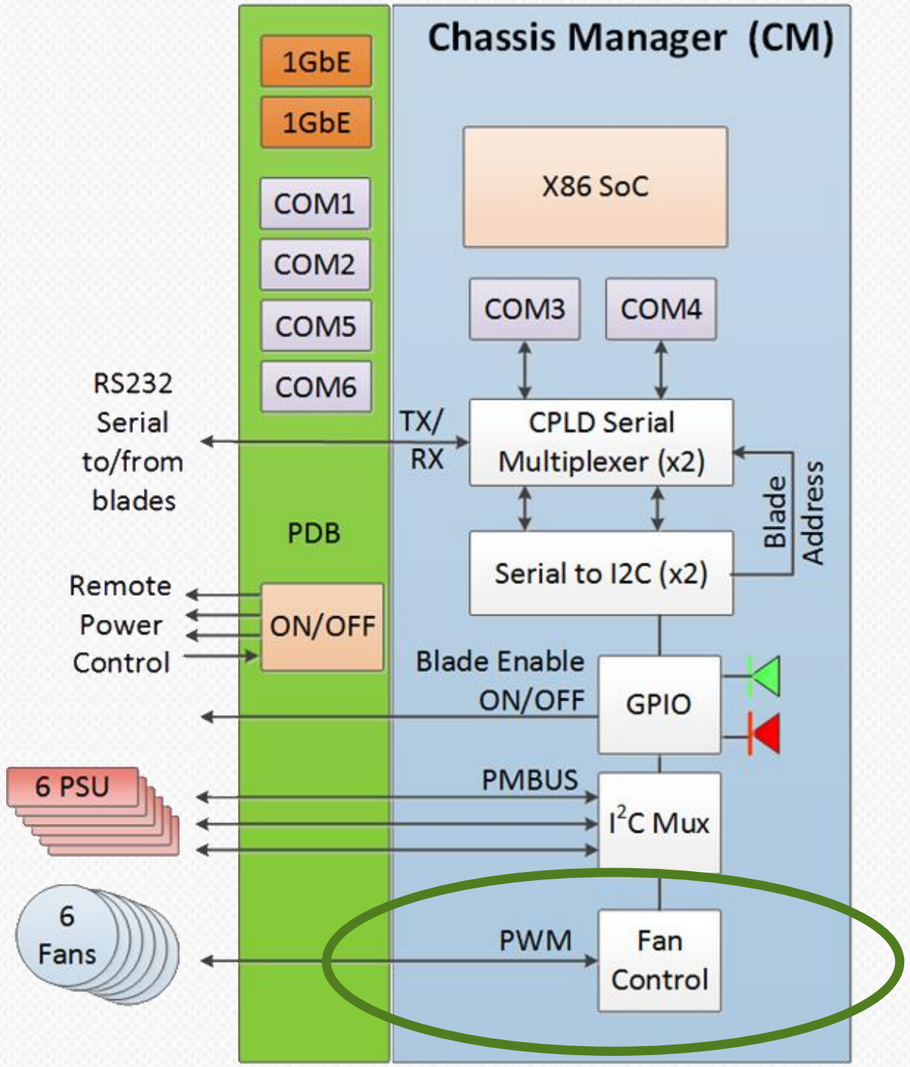
All blades polled for their PWM requirement

Fan Control performed by the fan controller (ADT7470)

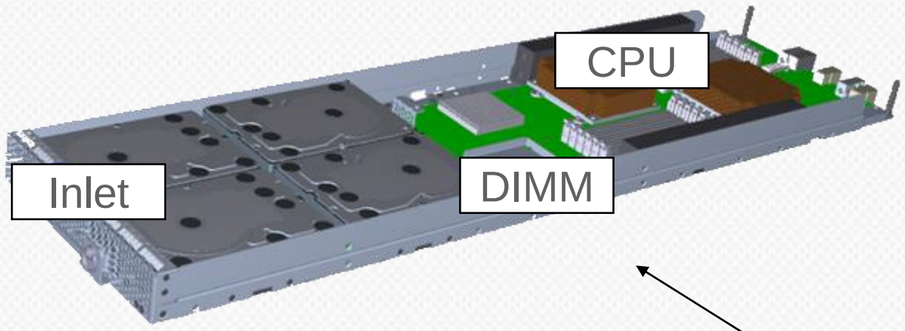
Watchdog ensures continuous operation

Fan speed and status is reported

Faults get reported and attention LED illuminated



Chassis manager: fan control algorithm



BMC-Lite

Monitors Thermally Critical Baseboard Components

Closed Loop Algorithm

Reports PWM (0 – 100)

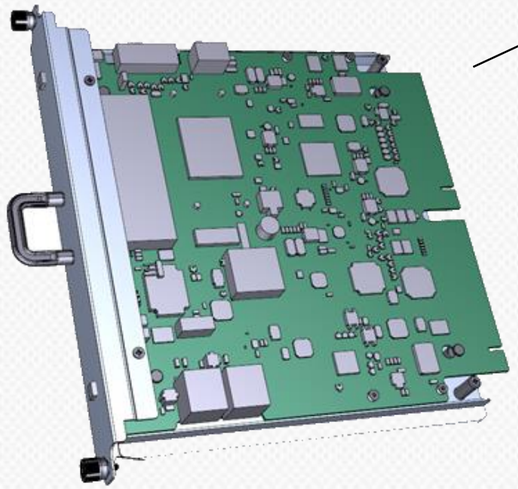
Chassis Manager

Polls 24x blades for PWM

Determines & Sets Fan Speed

Location based altitude correction

Fan failure compensation



Fans

6 x Variable Speed fans

Operate in unison

Report Speed and Status

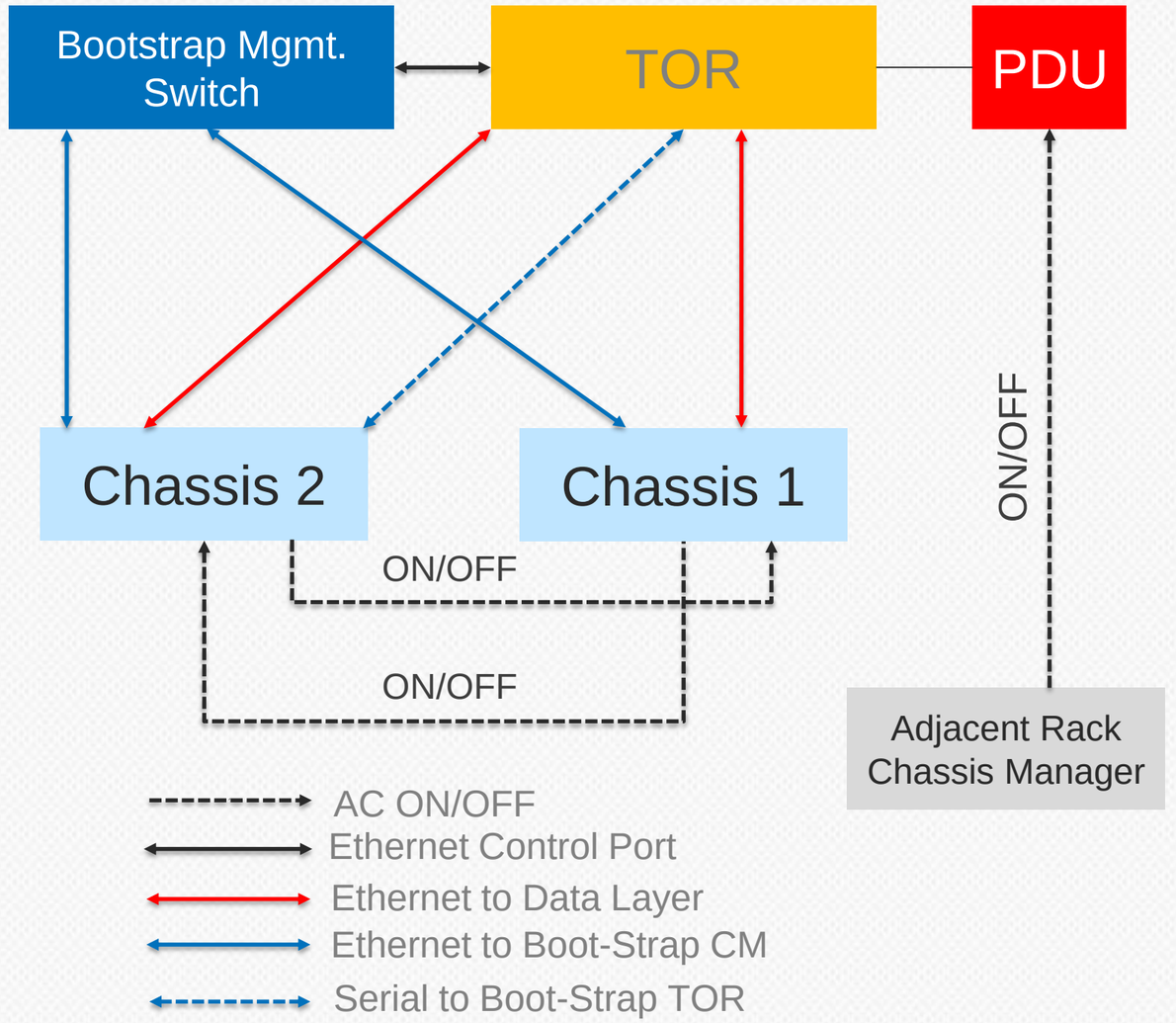
Chassis manager: power switch control

Auxiliary Power Switches:

- 3 x 12v power output control switches
- 1 x 12v power input control switch
- Relay or neighboring Chassis Manager remote power control

Power Control:

- Chassis Manager is hot pluggable, IO port physical connections on the PDB.
- Designed to be always on, only held off by the input power control switch

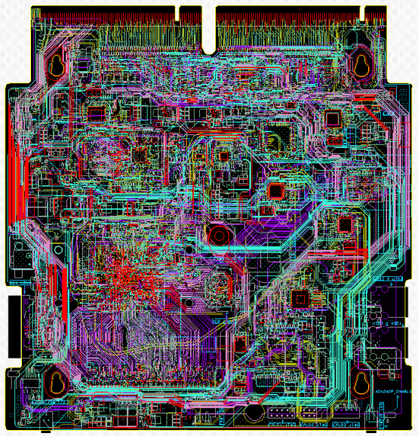


Microsoft cloud server spec: OCP contribution

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

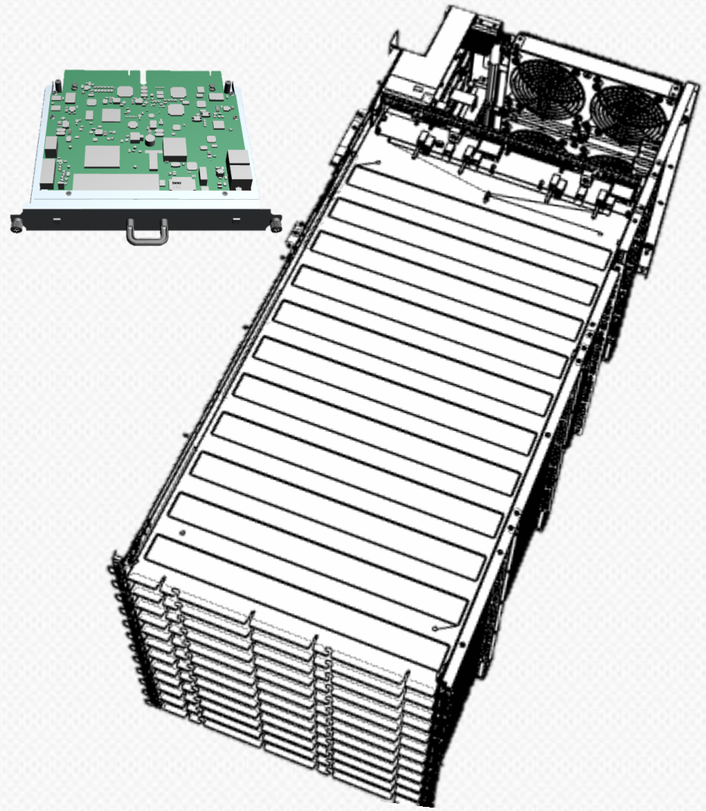
Chassis, Blade, Chassis Manager, Mezzanines, Management APIs

Board Files & Gerbers

Chassis Manager, Tray Backplane, Power Distribution Backplane

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