



CLOUD COMPUTING AND AI TOTAL SOLUTIONS PROVIDER

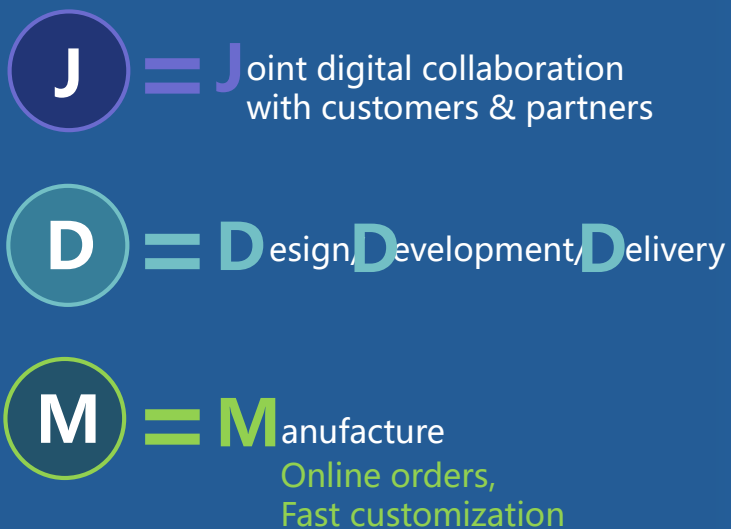
Powering The Open Data Center Revolution

John Hu
CTO, Inspur Group

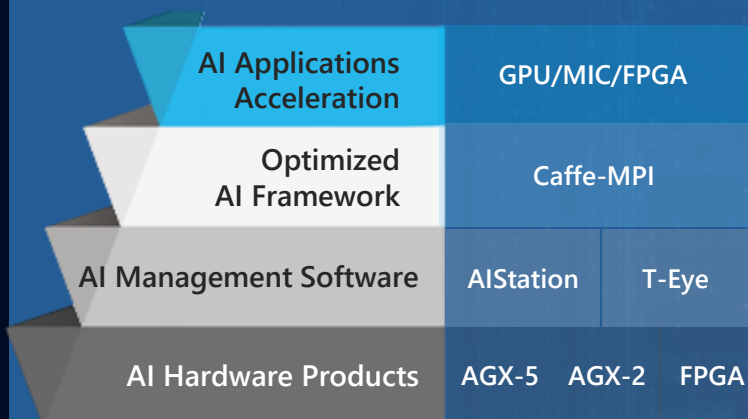
About Inspur

- **Top 3 Server Vender Worldwide**
- **Leading Full Stack AI Solution Provider**
- **Leading Open Data Center Products and Solutions Provider**

JDM Model



AI Full Stack Solution



Platinum Member of OCP



ODCC

OCP

Open Platform: Making Ecosystem More Open

Embrace
Open

Open Hardware

Member: Facebook, Inspur, Google, Microsoft



Over billion deployment

Members: Baidu, Tencent, Alibaba, Intel, Inspur



Millions of nodes deployed



OpenPower

Open Software



Rack Level Management

Deliver
Value

Open Design Platform





Accelerate Innovation



Decrease TCO



Manage Power Consumption



Lower Failure Rate



Easy to Join

Open to
Customers





Goal for Open Platform

Contribute Designs

Contribute Designs to the OCP Community

JDM

Jointly Design with Customers



Design Base

Architecture

Compute + Storage + AI
Schematic Design

Interconnection
& Pooling

Rack Management

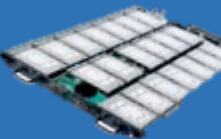
Use Cases & Technology Experience



No.1 Server supplier for Alibaba

Alibaba e-Commerce Platform is running on
Inspur ODCC Rack

Peak of system transaction reached 175,000 orders per second



Over **80%** of Baidu's hyperscale
rack-level servers Deploying **10,000
nodes per day**

Use Cases for Serving Hyper-scaler

❑ Comprehensive server customization solutions

❑ Major contributor to China Scorpio rack-scale server standards

❑ No.1 in China Internet industry, 50%+ market share

❑ 150,000+ rack-scale server shipment, 70%+ market share



No.1 Strategic Server Supplier of Alibaba

- Sustain No.1 market share in Alibaba
- 50% market share up to now
- The Best Service Provider Award 2016
- Won \$735Million order in 2017H1
- Supplying Intel® Scalable® Processor based servers since May 2017
- Applied Inspur FPGA successfully in Alibaba Cloud



No.1 Strategic Server Supplier of Baidu

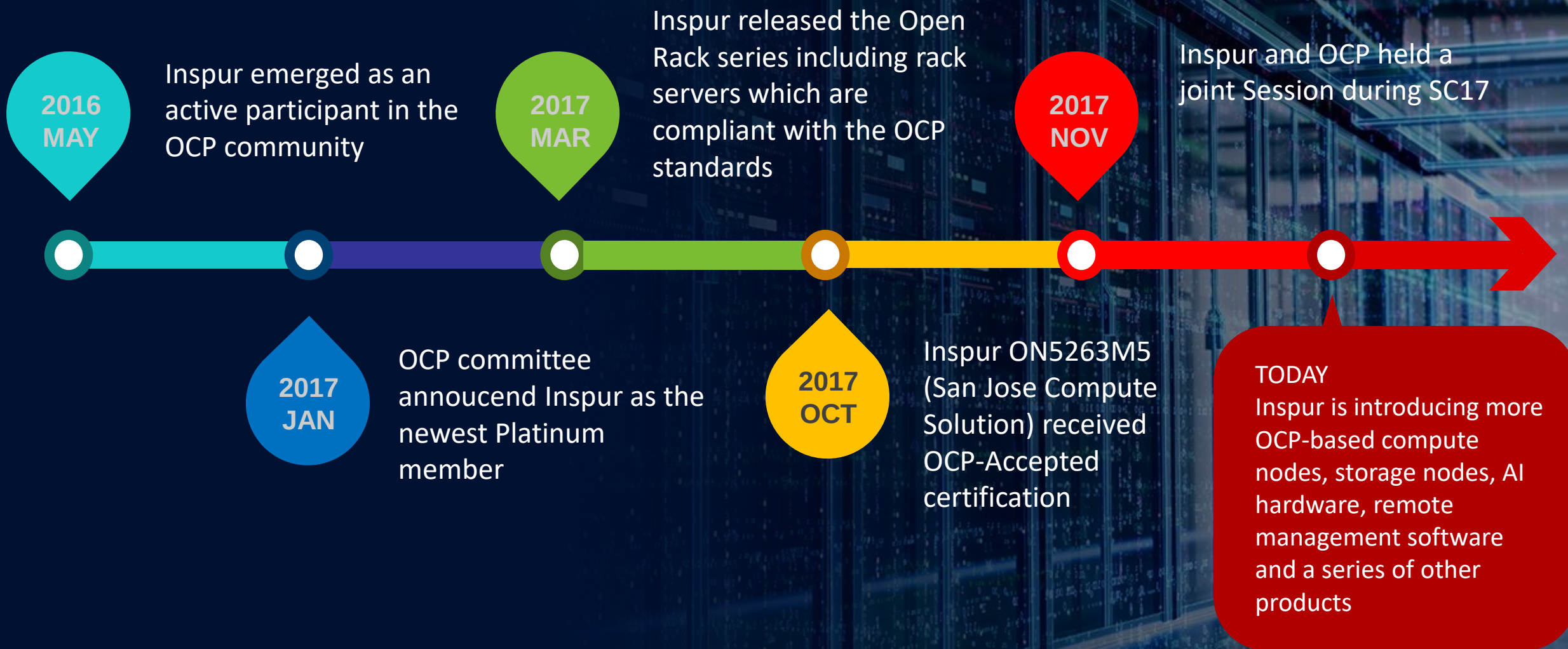
- 60% market share up to now
- 100% GPU server market share
- Won the largest bidding share in 1H'2017
- The only server vendor won the Technology Innovation Award 2016
- Joint development and release of the first 4U16 GPU BOX (SR-AI) which has been applied to Baidu Driverless Car Project.



Top Server Supplier of Tencent

- 30% market share up to now
- Deep cooperation and launching auto project operation/maintenance system to secure automatic and dynamic online error report of components
- First FPGA applied in Tencent Cloud

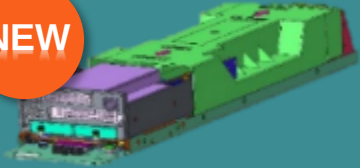
Momentum of Inspur with OCP



Inspur OCP Rack Scale Product Overview

Compute

NEW



I/O Balance

- 2x 2.5" HDD or NVMe
- 2x M.2
- 1x FHFL(x16) : CPU0
- 1x FHFL(x16) : CPU1

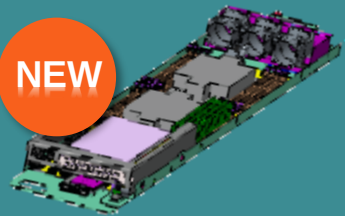
NEW



High Expansion

- 4x 2.5" HDD or NVMe
- 1x M.2
- 3x HHHL(x8) : CPU0

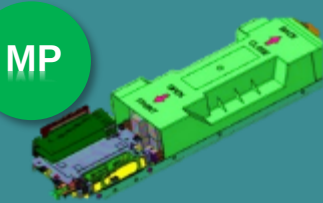
NEW



Density Computing

- 2x M.2
- Expansion Slot:
- 1x FHHL(x16) : CPU0

MP



Energy Efficiency

- 1x 3.5" HDD
- 1x M.2
- 1x FHHL(x16) : CPU0
- 1x FHHL(x16) : CPU0



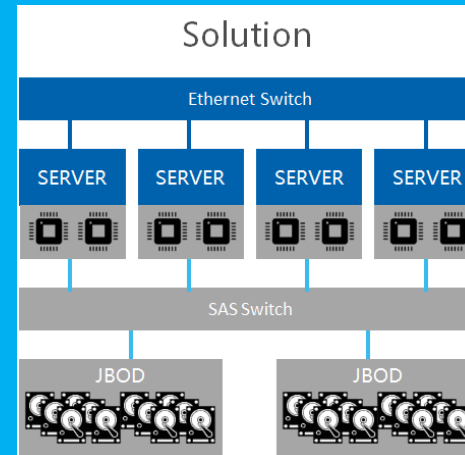
Storage

NEW



JBOD

- 20U form factor
- Two drive tray design
- 34*HDDs or SSD
- Each drive supports hot-plug



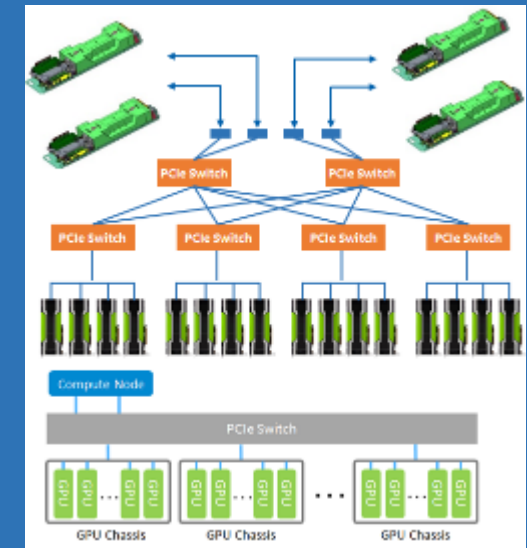
AI

NEW



GPU Box

- 40U form factor
- 16 PCIe GPU cards



Inspur OCP JBOD Storage

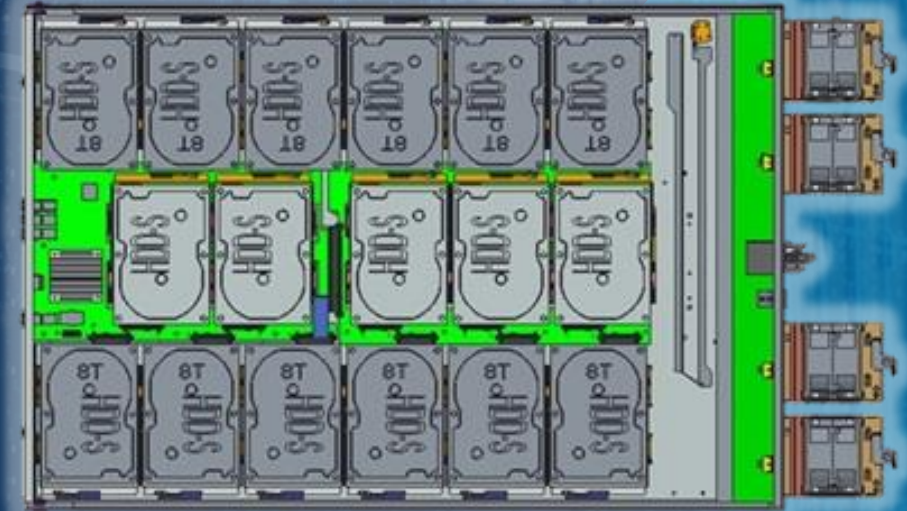
High Density and Cost Efficiency

High Density

- 20U 34*HDDs, 13% storage capacity improvement

Flexible Architecture

- Dual separate storage tray, 17 drives per tray
- Each tray supports SATA and SSD, maintained individually



Inspur OCP AI Product

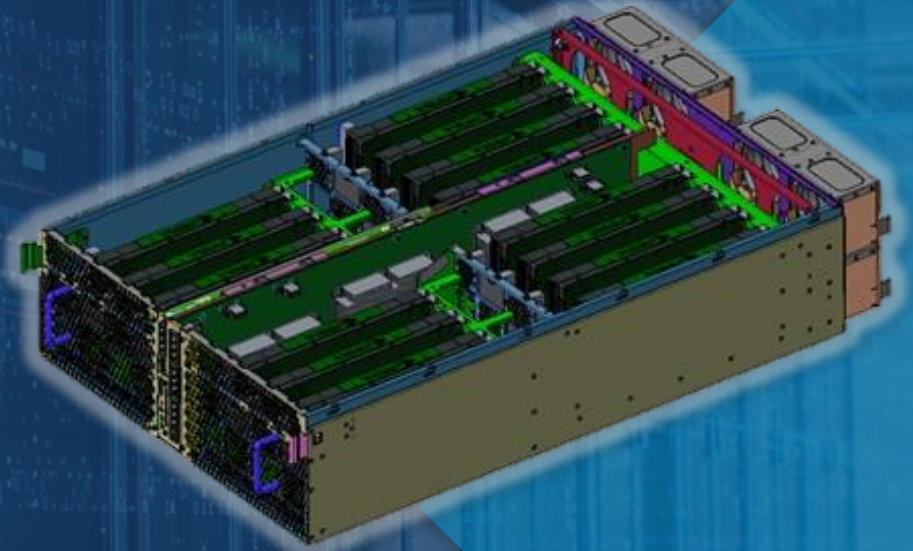
Scalable and Leading Compute Architecture

Ultral High Density GPU

- 40U 16*GPU for training or 16*FPGA for inference
- Improve efficiency of system's computation expansion

Optimized Architecture for AI

- Speed up the training process
- Flexible architecture to support multi-node based on application



Hardware Management For OCP Rack

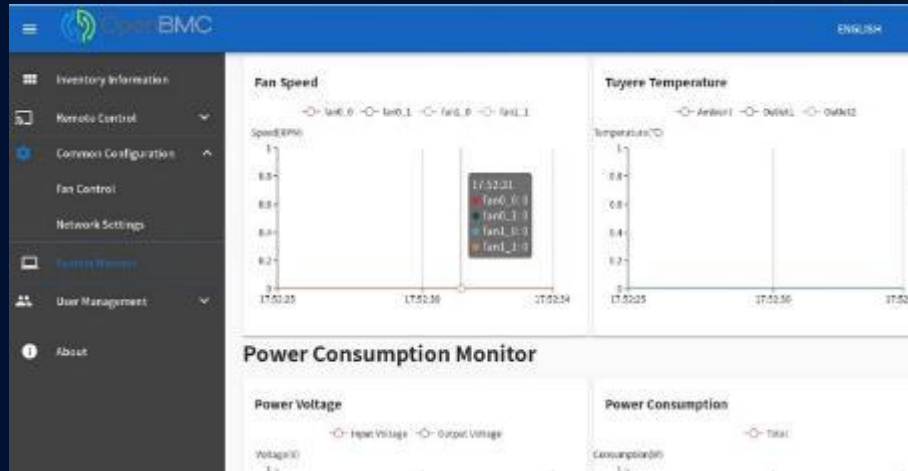
Runs OpenBMC on OCP Compute Node

Implement Redfish OCP Baseline for compute node and pass the dmtf redfish plugfest

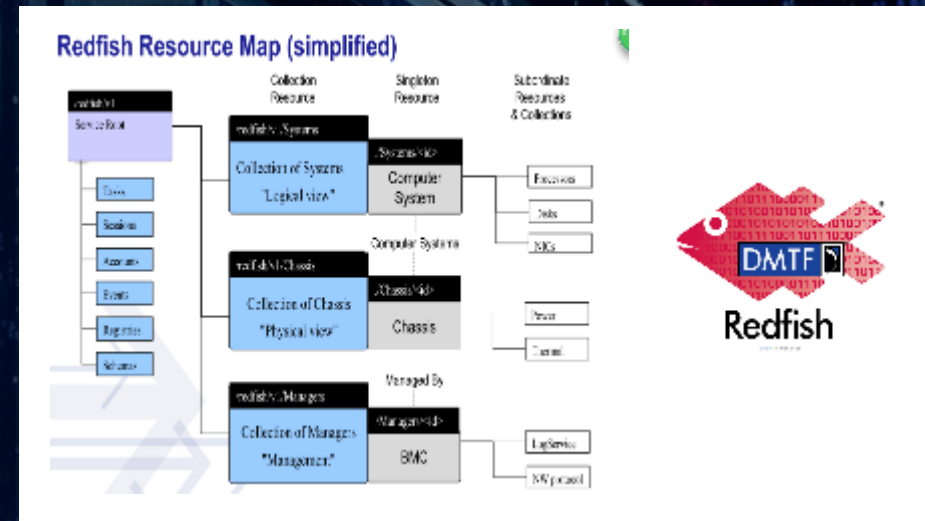
Optimized for dbus interface



Inspur OpenBMC Key Features



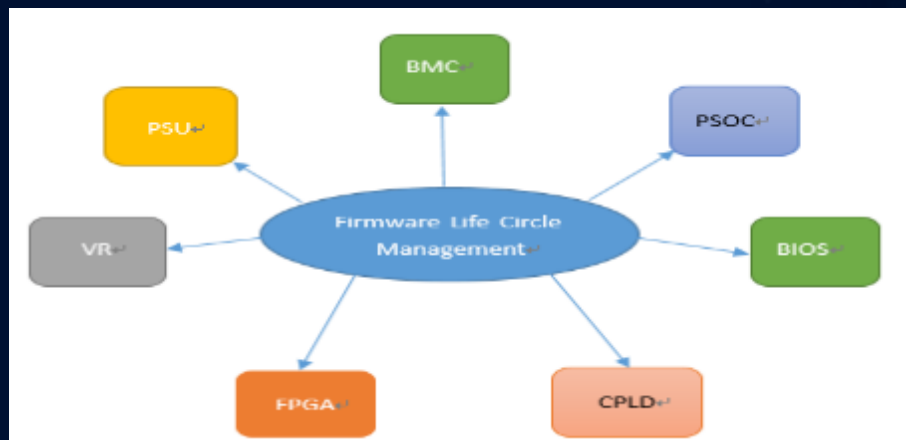
WebUI to monitor system status



Redfish/Restful Function, support Redfish OCP Baseline Profile

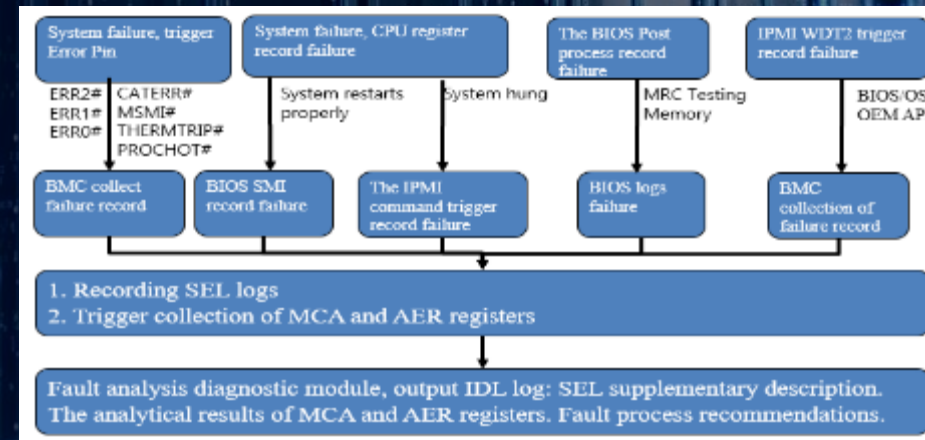
Component **Firmware Life cycle Management**

1. Version auto discovery
2. Intelligent update for BMC, BIOS, CPLD, FPGA etc.
3. Firmware rollback when error occurs.



Fault Diagnosis

1. Diagnosis system fault directly
2. Output the detailed fault records and recommendations
3. BMC subsystem fault diagnosis.



Inspur OCP Standard Rack Scale Total Solution



inspur

General Purpose to OCP Standard

NF5180M5



Inspur promote General 19
inch 1U and 2U server
Contribute in OCP

	Inspur NF5180M5
CPU	2 x SKL CPUs, TDP up to 205W Support CPU with Omni-path Support CPU with FPGA
Memory	24x DDR4 DIMM Support RDIMM/LRDMM/NVDIMM and AEP
Storage	Front: Up to 4 x 3.5 +2 x 2.5 or 10 x 2.5 Internal: 2 x M.2 SATA and 2 x PCIe Rear: 2 x2.5
Ethernet	No integrated Ethernet port; OCP Ethernet card Optional
PCIe	Up to 3 standard PCIe and 2 OCP/1OCP+PHY Ethernet card
Others	2 x micro SD card

NF5280M5



	NF5280M5
CPU	2 x SKL CPU, TDP up to 205W
Chipset	C622
Memory	24 x DDR4 DIMM Support RDIMM/LRDMM/NVDIMM and AEP
Storage	Front: Up to 12 x 3.5 or 24 x 2.5 HDD Internal: 2 x M.2 SATA/PCIe x1, 4 x 3.5 HDD Rear: Up to 4 x 3.5+4 x 2.5 HDD or 10 x 2.5 HDD
Ethernet	No integrated Ethernet port; OCP/PHY Ethernet card Optional
I/O	Up to 10 standard PCIe and 1 OCP/PHY Ethernet card Up to 4 PCIe x16 slot
PSU	500w/800w/1300w/1600w
Chassis	435mm(W) x 87mm(H) x798/780mm(D)
Others	2 x Micro SD card



THANK YOU

Learn more about Inspur at www.inspursystems.com

inspur