Powering The Open Data Center Revolution

John Hu
CTO, Inspur Group
About Inspur

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

**JDM Model**

- **J** = Joint digital collaboration with customers & partners
- **D** = Design, Development, Delivery
- **M** = Manufacture, Online orders, Fast customization

**AI Full Stack Solution**

- **AI Applications Acceleration**
- **GPU/MIC/FPGA**
- **Optimized AI Framework**
- **Caffe-MPI**
- **AI Management Software**
- **AIStation, T-Eye**
- **AI Hardware Products**
- **AGX-5, AGX-2, FPGA**

**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

OpenBMC

Redfish

Rack Level Management

Deliver Value

Open Design Platform

Accelerate Innovation

Decrease TCO

Manage Power Consumption

Lower Failure Rate

Easy to Join

Open to Customers

Tencent

ODCC

OpenPower

Open to

Customers
Goal for Open Platform

**Contribute Designs**
Contribute Designs to the OCP Community

**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
- No.1 in China Internet industry, 50%+ 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 $735 Million order in 2017H1
- Supplying Intel® Scalable® Processer 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

- Major contributor to China Scorpio rack-scale server standards
- 150,000+ rack-scale server shipment, 70%+ market share
Momentum of Inspur with OCP

- **2016 MAY**: Inspur emerged as an active participant in the OCP community.
- **2017 MAR**: Inspur released the Open Rack series including rack servers which are compliant with the OCP standards.
- **2017 JAN**: OCP committee announced Inspur as the newest Platinum member.
- **2017 OCT**: Inspur ON5263M5 (San Jose Compute Solution) received OCP-Accepted certification.
- **2017 NOV**: Inspur and OCP held a joint Session during SC17.
- **TODAY**: Inspur is introducing more OCP-based compute nodes, storage nodes, AI hardware, remote management software and a series of other products.
Inspur OCP Rack Scale Product Overview

Compute

New

I/O Balance
- 2x 2.5”HDD or NVMe
- 2x M.2
- 1x FHHL(x16) : CPU0
- 1x FHHL(x16) : CPU1

New

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

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

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

Compute

New

5

Energy Efficiency
- 2x M.2

Expansion Slot:
- 1x FHHL(x16) : CPU0

Storage

New

Solution

Ethernet Switch

Server

Server

Server

SAS Switch

Server

MP

New

AI

GPU Box
- 40U form factor
- 16 PCIE GPU cards
**Inspur OCP JBOD Storage**

**High Density and Cost Efficiency**

**High Density**
- 2OU 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
- 4OU 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

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.

**Redfish/Restful Function**, support Redfish OCP Baseline Profile
Inspur OCP Standard Rack Scale Total Solution
General Purpose to OCP Standard

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

**Inspur NF5180M5**
- 2 x SKL CPUs, TDP up to 205W
- Support CPU with Omni-path
- Support CPU with FPGA
- 2x DDR4 DIMM
- Support RDIMM/LRDIMM/NVDIMM and AEP
- 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 x 2.5
- No integrated Ethernet port; OCP Ethernet card Optional
- Up to 3 standard PCIe and 2 OCP/10CP+PHY Ethernet card
- 2 x micro SD card

**Inspur NF5280M5**
- 2 x SKL CPU, TDP up to 205W
- Chipset C622
- 24 x DDR4 DIMM
- Support RDIMM/LRDIMM/NVDIMM and AEP
- 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
- No integrated Ethernet port; OCP/PHY Ethernet card Optional
- Up to 10 standard PCIe and 1 OCP/PHY Ethernet card
- Up to 4 PCIe x16 slot
- 500w/800w/1300w/1600w
- Chassis 435mm(W) x 87mm(H) x 798/780mm(D)
- 2 x Micro SD card
THANK YOU

Learn more about Inspur at www.inspursystems.com