
Chulho Kim & Miguel Terol
Lenovo Datacenter Group - HPC
Lenovo HPC Innovation Center Stuttgart

Industry Leaders

- Intel – Processor/Accelerator Innovation
- Mellanox – High Speed Fabrics
- Nvidia – Graphics Processing
- IBM – Files System and Work Load

Clients

- LRZ – Energy Efficient Systems and Software
- Daresbury – HPC Software and Optimization
- BSC – Extreme Application Scaling
- CINECA – Big Data and Many Cores
- RZG – Advancing Material Science
- FZJ – Active Storage

Visionary Client Partners

ISVs

Bring together the newest technology and skills

Focused knowledge and deep skills advance the science of HPC
Deep Commitment to HPC Industry Standards

• **MPI 3.1/4.0 Standards**
  • Fully participating member since December of 2014 (1\textsuperscript{st} meeting after becoming Lenovo)
  • 4 meetings a year. Helped approve MPI 3.1 standard and now starting to look at MPI 4.0 efforts.

• **InfiniBand Trade Association (IBTA)**
  • In process of finalizing registration

• **OpenFabric Alliance (OFA)**
  • Registered at “Supporters” level and active participant
  • Participating in OFA workshop in March

• **OpenHPC initiative**
  • Founding member
Open Software Solutions for HPC

- Founding member of OpenHPC
- Collaborating with Oxford University to create an Open System Management Web Console (OSMWC)
- Architecture owner for Open Source system management projects: Confluent and OSMWC
- Continued contribution to xCAT
- Allow customers to add/replace components as needed
- Professional “Services” through to help customers deploy and optimize their clusters
Open Source HPC Collaborations

- **Barcelona Super Computing Center**
  - Energy aware MPI runtime
  - HPC Virtual environment research (Lenovo Job Capsules, Docker, KVM)
- **Oxford University**
  - Open System Management Web Console collaboration
# xCAT Open Source Toolkit

## Lights Out Hardware Management
- Everything can be done remotely
- Discovery hardware on the network and configure it
- Update firmware and ASU settings
- Power control, remote console, energy mgmt, switch mgmt

## Extreme Scalability
- Up to 100,000 nodes
- Fully distributed architecture
- Custom IPMI stack
- Custom bootloader

## Operating System Deployment Choices
- Stateful – OS on disk
- Stateless – ease of mgmt
- Statelite – unique hybrid
- Capture image from node
- Apply software updates to nodes

## Broad Hardware and OS Support
- NextScale, iDataPlex, system x rack mount
- Other IPMI-controlled svrs
- Pure Flex and more...
- RHEL, CentOS, SL, OL
- SLES, Debian, Ubuntu
- Windows, AIX

## Broadest Virtualization Support in Industry
- KVM
- VMWare, RHEV-M/H
- PowerVM, zVM
- Stateless hypervisor deployment
- VM life-cycle mgmt

---

**Cloud/Cluster**

- Eclipse Public License
- Extensible plug-in architecture
- Active community contributions, mailing list

---

2016 Lenovo
Confluent – xCAT Enhancement

- Lenovo led project to improve upon xCAT heritage
- Faster performance and high concurrent request count
- Enabled for web developers
- Easier to learn and use
  - Developers must use a common messaging layer, improving consistency for user
  - Command names are consistent and map to familiar concepts for non-xCAT users
  - Simpler, yet more powerful configuration expressions
  - All functionality mapped to a filesystem-like discoverable structure
  - More robust local CLI access channel
- xCAT-style noderanges and commands
- Native RESTful interface
- Interactive CLI
- Improved security
  - Symmetric credential encryption
  - Safer configuration expression implementation

2016 Lenovo
Open System Management Web Console (OSMWC)

**GOAL:** Make HPC Open Source More User Friendly

- Provide easy access to new HPC customers using Open Source HPC Infrastructure
  - Low cost and simplified entry into HPC
  - Easy to deploy, manage and submit jobs

- Visual summary views to help understand cluster usage
  - Admin Console – User management, Cluster Monitoring
  - User Console – Jobs submission, Job/Cluster Monitoring

- Initial target and Proof of Concept trials – China Market
  - Focus on China Market first – A lot of customers are just coming into HPC workloads
  - Collaborating with customers to understand their usage models and future requirements
  - Very positive feedback and market acceptance
  - LiCO – Lenovo Intelligent Computing Orchestration was released to China market

- WW Market – Create English version and work with collaborators to release the English version as “Open Source” project
  - Oxford University has signed on to work on this collaboration
**Initial OSMWC COMPONENTS**

**WEB CONSOLE GUI**
- Ganglia
- Torque/MAUI
- xCAT/Confluent
- OpenMPI, MVAPICH
  - MPICH, Intel Parallel studio
- CentOS/RHEL
- Lustre
- OFED
- Server
- Storage
- Network

**Installation guide / scripts**
**Admin guide / scripts**

**What is Web Console:**
An Unified GUI
- User Portal (dashboard, submit job, monitor job)
- Admin Portal (dashboard, user/account management)

**Future Work Items:**
- SLURM integration
- ICINGA integration
- Intel OPA integration
- LDAP integration

Lenovo components
Open Source/3rd party
Lenovo Hardware

Main components listed here would be part of OpenHPC project
What are your important Open Source components that we should add?

2016 Lenovo
END USER CONSOLE
ADMIN CONSOLE

Cluster Name

Jobs

Nodes

Network Traffic

CPU Usage

Memory Usage

Disk Usage

Jobs

All queues

In the past one hour

Uncompleted

Node Occupancy

Management Nodes

Login Nodes

Compute Nodes

I/O Nodes

2016 Lenovo
HPC Virtualization - Job Capsules

- Somewhere between a typical ‘container’ and traditional Linux system
- Avoid network/fabric latency penalties
- Mask certain capability as appropriate (A job can ‘disable’ hyperthreading)
- Offer users ability to self-service selection from a menu of administrator blessed runtime environments
- Compatible with docker image build/maintenance procedures
- Does not require docker
- Run under a traditional Linux install, stateless, or ‘Boot capsule’ platforms
Job Capsules

Base
- Ubuntu LTS
- Ubuntu
- EL7
- EL6
- SLES11
- SLES12

Expansion
- Math Libraries
- JRE
- Different python versions
- More.....
THANK YOU

DAKUJEM  DANK  BEDANKT  MERCI  謝謝  ありがとう  СПАСИБО  GRACIAS  DZIĘKUJĘ  DANKE  OBRIGADO  БЛАГОДАРЯ  GRAZIE  धन्यवाद  GRACIAS