

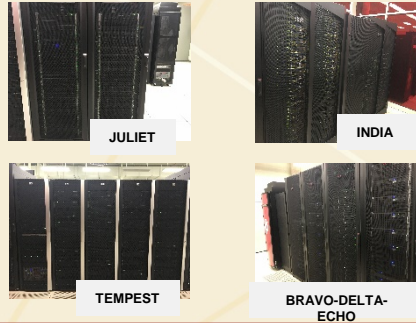
Digital Science Center

Geoffrey C. Fox, David Crandall, Judy Qiu, Gregor von Laszewski, Fugang Wang, Badi' Abdul-Wahid, Saliya Ekanayake, Supun Kamburugamuva, Jerome Mitchell, Bingjing Zhang
School of Informatics and Computing, Indiana University

DSC Computing Systems

- Just installed 128 node Haswell based system (Juliet)
 - 128 GB memory per node
 - Substantial conventional disk per node (8TB) plus PCI based SSD
 - Infiniband with SR-IOV
 - 24 and 36 core nodes (3456 total cores)
- Working with SDSC on NSF XSEDE Comet System (Haswell 47,776 cores)
- Older machines
 - India (128 nodes, 1024 cores), Bravo (16 nodes, 128 cores), Delta (16 nodes, 192 cores), Echo (16 nodes, 192 cores), Tempest (32 nodes, 768 cores) with large memory, large disk and GPU
- Optimized for Cloud research and Large scale Data analytics exploring storage models, algorithms
- Build technology to support high performance virtual clusters

FutureSystems



Use of FutureSystems India Cluster

Infrastructure

- HPC and cloud mixed environment
 - Had openstack in service since Cactus release back to 2011
 - Currently operating two openstack clouds: Juno and kilo, with ~60 compute nodes total.
- Usage Data of OpenStack
 - 1,060,442 hours of wall time (between 01/01/2015 - 10/31/2015)
 - 14,318 VM instances launched total (between 01/01/2015 - 10/31/2015)
 - Average 63 VMs launched per day
 - 70% of tiny or small server size and 30% of medium or (x)large server size were used

Courses

- Total Courses to date: 75 course projects from 31 distinct institutions since 2010
 - SP15 v594 (online) Topic: cloud computing, big data
 - Audience: #38 globally distributed GE employees
 - Projects: deployments of big data platforms and analysis of data sets
 - FA15, FA14 (online) [in progress]
 - Topic: big data
 - Audience: #139 students (on site & remote)
 - Projects: analysis of sports, Amazon movie reviews, stock market, twitter data
- Undergraduate Student Research
 - Developing Cloudmesh for cloud environment management
 - REU and on-site students

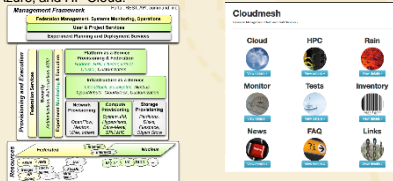
Rapid Prototyping HPC Environment for Deep Learning

We are developing an on-ramp to deep learning that utilizes HPC and GPU resources. It serves as an aggregate for modules related to deep learning. We address a multitude of issues including deployment, access, and integration into HPC environments. Simple interfaces are provided that allow easy reusability. We are working with our partners on a performance optimized convolution kernel that will be able to utilize state of the art GPUs including AMD. Data management will be available as part of deep learning workflows.



Cloudmesh for Managing Virtual Clusters

Cloudmesh is an important component designed to deliver a software-defined system – encompassing virtualized and bare-metal infrastructure, networks, application, systems and platform software – with a unifying goal of providing Cloud Testbeds as a Service (CTaaS). Cloudmesh federates a number of resources from academia and industry. This includes existing FutureSystems, Amazon Web Services, Azure, and HPC Cloud.



Virtual Clusters with SDSC Comet

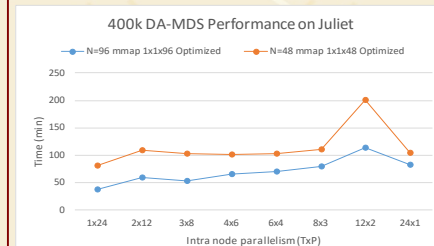
Comet, is a new petascale supercomputer designed to transform advanced scientific computing by expanding access and capacity among traditional as well as non-traditional research domains. Comet will be capable of an overall peak performance of nearly two petaflops, or two quadrillion operations per second.

We are working together with SDSC to deliver an easy to use service that allows users to request virtual clusters that are close to hardware. Advanced users will be able to request such clusters and allow the users to manage them and deploy their own software stacks on them.

Comet is the first virtualized HPC cluster, and delivers a significantly increased level of computing capacity and customizability to support data-enabled science and engineering at the campus, regional, and national levels, and in turn support the entire science and engineering enterprise, including education as well as research.

IU is currently working on delivering easy to use client interfaces and leverages experiences from our cloudmesh software.

Threads v. Processes on 24 core Juliet Nodes for 48 and 96 nodes



Threads v. Processes on 24 core Juliet Nodes for 24, 48 and 96 nodes

