SuMegha Cloud Kit

Create your own private scientific cloud

Presentation by B.Arunachalam

Centre for Development of Advanced Computing (C-DAC) Bangalore, India

Authors : **Vineeth S Arackal, Arunachalam B, Kalasagar B, Sumit Kumar, Mangala N, Sarat Chandra Babu, Prahlada Rao, Sukeshini @ C-DAC Bangalore,India**

Presentation Outline

- Overview
- Scientific Cloud
- Architecture
- Features
- Distribution & Installation
- Scientific Cloud portal and other tools
- Challenges faced
- Conclusion

Overview

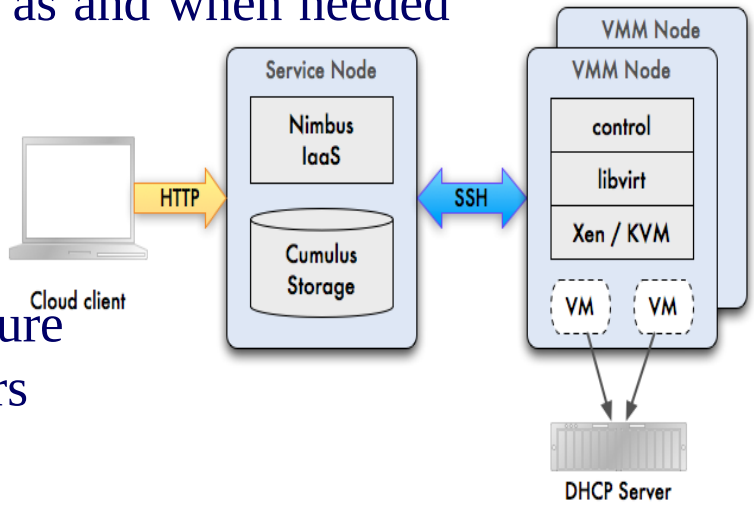
- Enables the user to setup their own private scientific cloud
- Offers IaaS, PaaS and SaaS for scientific community to work on problems of modelling and simulation
- Effectively utilize the idle computers and provides MPI or Hadoop clusters on demand to solve compute-data intensive applications
- Cloud can be set up in a single desktop and designed in a modular fashion using a judicious mix of open source and in-house developed components
  - Nimbus cloud middleware, Xen hypervisor and SuMegha Portal

Overview…

- Storage as a service is provided by the in-house developed CloudVault solution integrated into the lab kit.
- Supplemented with Software as a Service such as Seasonal Forecast Model (SFM), Next Generation Sequencing (NGS), etc
- Provides a comprehensive environment for scientific computing and Useful academic institutions, to easily set up the Cloud Lab

Scientific Cloud

Scientific Cloud is a model for on-demand access to a shared pool of HPC resources (ex: Servers, Storage, Networks, Applications) that can be easily provisioned as and when needed by the researchers/ scientists



Benefits of Scientific Cloud

- On demand access to HPC resources
- Ease of access to the available infrastructure
- Virtual ownership of resources to the users
- Ease of deployment

Scientific Cloud Offerings

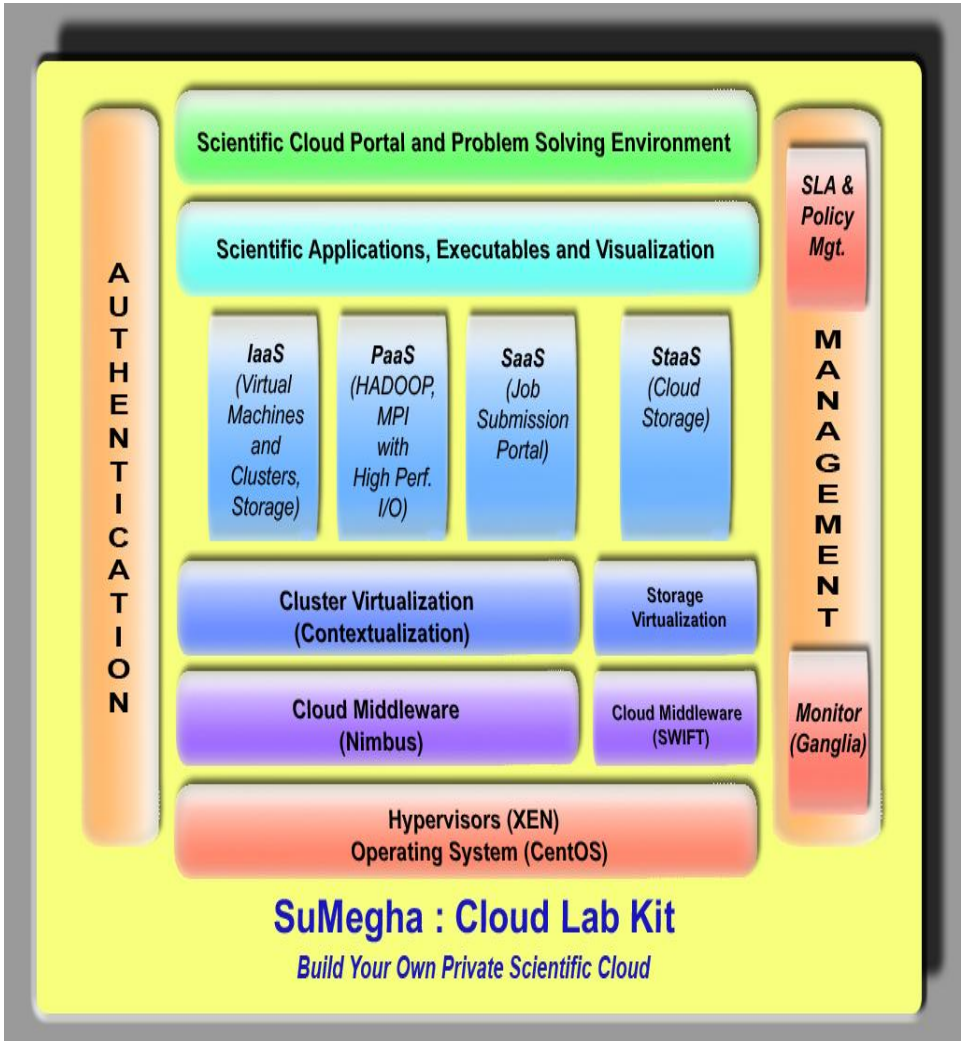
Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Software-as-a-Service (SaaS), Storage-as-a-Service (StaaS)

Related Work

- Open Source cloud solutions : Nimbus, Xen Hypervisor, Stratus Lab, Cumulus, Open Cirrus and GridGain
- *Scientific cloud infrastructure facility - Objectives*
  - Easy to access the HPC resources for academic researchers & research communities
  - On demand provisioning of challenging parallel environments by providing virtual clusters
  - Enablement of scientific applications in the areas of climate modelling and bioinformatics.

SuMegha – Architecture

- Cluster as an Infrastructure Service (IaaS)
  - Provisioning of on demand HPC clusters
  - Map Reduce based Hadoop clusters to process large data sets
  - Support for MPI for tightly-coupled applications
- Software as a Service (SaaS)
  - Job Submission Portal (C_JSP), Cloud Access Portals
  - Monitoring tools
- Platform as a Service (PaaS)
  - Program development and Analysis Tools
- Storage as a Service (StaaS)
- Applications as Service (AaaS) PSEs
  - PSE for Climate Modelling
  - PSE's for Bio Informatics



SuMegha Components

COMPONENTS

- Hypervisor
  - ✓ Xen
- Cloud Middleware
  - ✓ Nimbus, Openstack Swift
- Storage
  - ✓ Glusterfs
- Portal
  - ✓ SuMegha Portal
  - ✓ Job Submission Portal
  - ✓ Storage Portal
- PSE
  - ✓ PSE for SFM
  - ✓ PSE for NGS Pipeline

- **Packaged software for automated deployment of private cloud**
- **Suitable for academic & research organizations to setup Cloud labs**

SUPPORTED IMAGES

- **Virtual Machines**
  - ✓CentOS (Small, Medium, Large)
  - ✓Job Submission Portal (JSP)
- **Virtual Clusters**
  - ✓MPICH enabled clusters with JSP
  - ✓HADOOP Clusters
- **PSE enabled images**
  - ✓SFM
  - ✓ NGS Pipeline

Salient Features

- **Automated deployment software** for installation and configuration of private cloud
- **Golden images:** Virtual machine images with centos can be provided to create virtual machines
- **Parallel programming environment**
- **Hadoop data intensive applications**
- **Creation of Virtual Machines/Clusters**
- **CDAC in house developed tools**
- **Problem solving environments for SFM and NGS Scientific applications**

Types of Virtual Machines

Images sizes	Description
Small	1 vCPU & 1GB RAM
Medium	2vCPU & 2GB RAM
Large	4vCPU & 4GB RAM

Distribution & Installation

- SuMegha lab kit Stack consists of :
  - Hypervisor
  - Cloud Middleware
  - Cloud portal
  - Golden images with CentOS, OpenMP, MPI & HADOOP environments
  - Preloaded HPC applications
  - Manuals – Installation Manual and User Manual

Installation…

Virtual Machine Manager Node

Operating System	CentOS Version > 6.2
CPU	One or more 64-bit x86 CPU(s), 1.5 GHz or above, 2 GHz or faster multi-core CPU recommended
RAM	Minimum 4 GB
Disk Space	Minimum 60 GB; Minimum 2GB for /boot partition
Network	Internet Connectivity

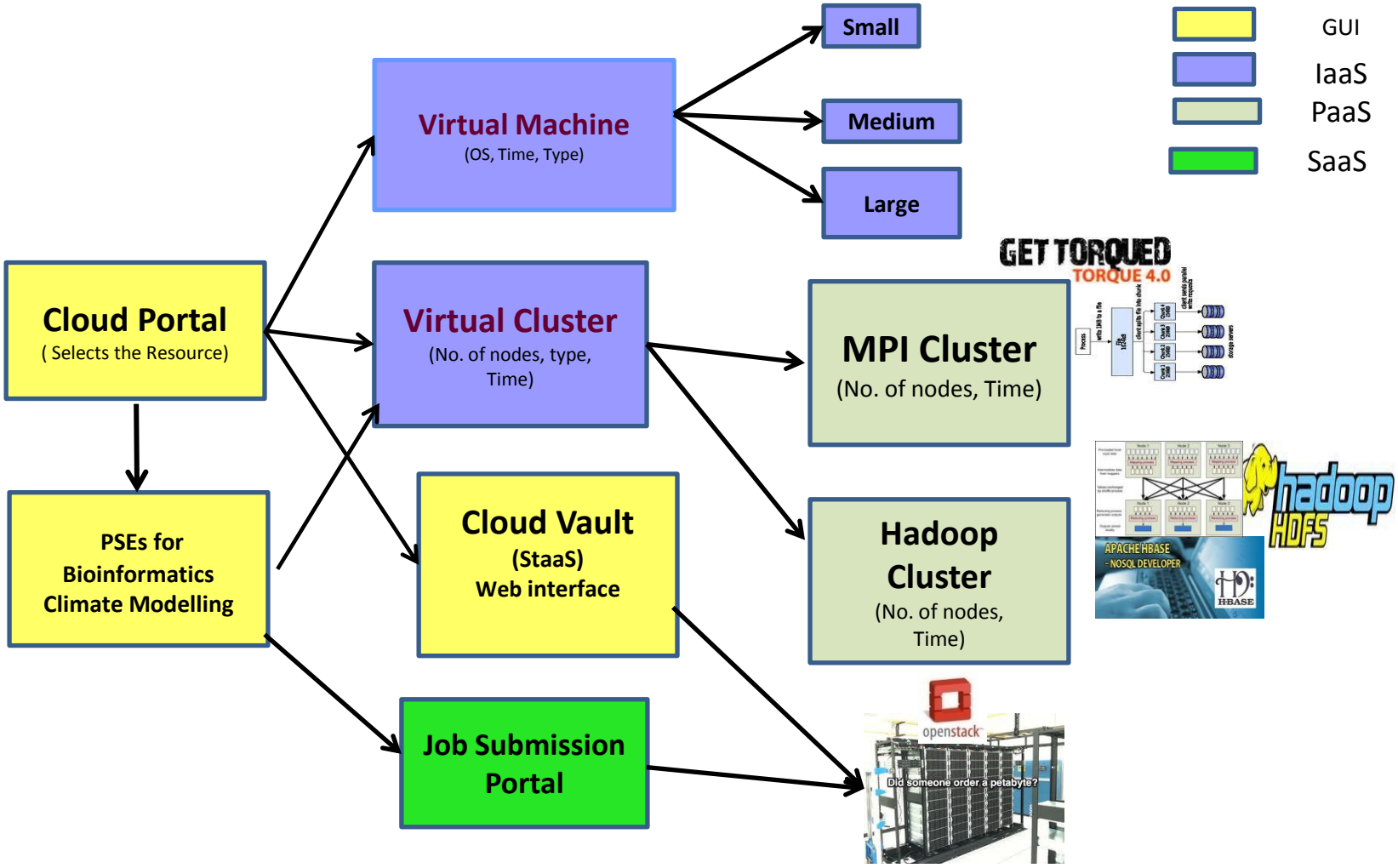
Service Node

Operating System	CentOS Version > 6.2
RAM	4 GB
Disk Space	Minimum 100 GB
Network	Internet Connectivity
Software	Oracle JAVA 1.7, Python (2.6 – 3.0)

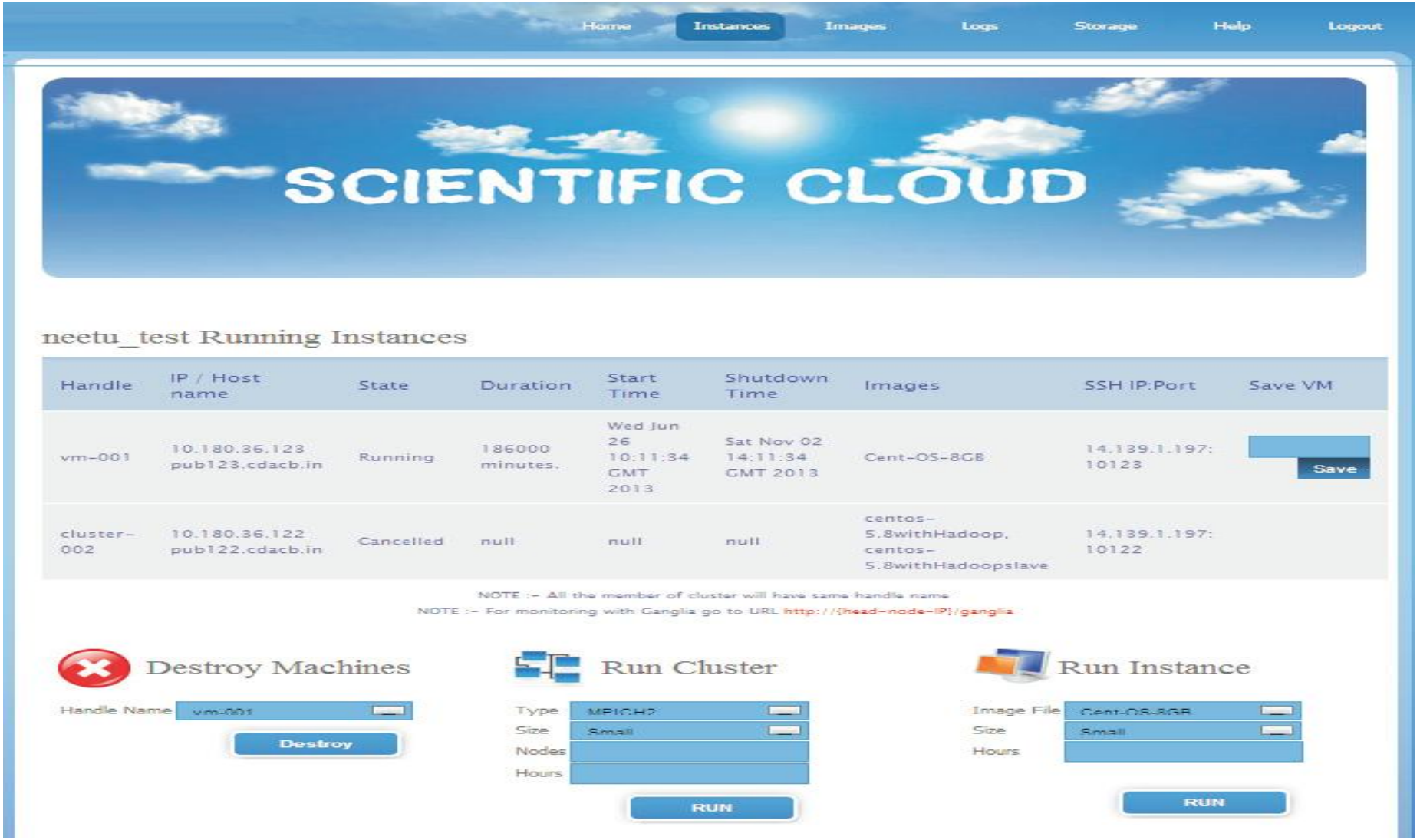
```
Successfully Installed Bridge Utils Package.......
Do you want to create Default Xen Bridge..Please confirm (yes or no) :yes
You Have entered yes
Enter the Ip Address value : 10.180.32.8
Enter the Gateway Ip Address : 10.180.32.1
Enter the DNS Server IpAddress : 10.180.0.11
The values Enters are : 10.180.32.8 ,     10.180.32.1     10.180.0.11
In xen_bridge function The values Enters are : 10.180.32.8 ,     10.180.32.1     10.180.0.11

Enter the Interface name : eth1
Shutting down interface eth1:  bridge xenbr0 does not exist!
                                                           [  OK  ]
Shutting down loopback interface:                          [  OK  ]
Bringing up loopback interface:                            [  OK  ]
Bringing up interface eth1:                                [  OK  ]
Bringing up interface xenbr0:                              [  OK  ]
Network Service Started Successfully.........
Press Enter to Continue
```

SuMegha Cloud Services



SuMegha –Resource Creation



SuMegha –Image Repository

Repository Images

Name	Mode	Modified Date	Size
CentOS-3C8-x86_64	Read only	Oct 9 2015 @ 10:27	3221225472 bytes (~3072 MB)
MPICH2-Master	Read only	Oct 9 2015 @ 12:19	8589934592 bytes (~8192 MB)
MPICH2-Slave	Read only	Oct 9 2015 @ 12:16	8589934592 bytes (~8192 MB)
dsdba-master	Read only	Oct 14 2015 @ 10:49	21172846592 bytes (~20192 MB)
dsdba-worker	Read only	Oct 14 2015 @ 10:58	21172846592 bytes (~20192 MB)

NOTE :- These are the images provided by the Administrator to create Virtual Machine

Challenges faced

- Installation, Configuration, Launching web applications from the VM and usage of this private scientific cloud
- Configuring of Network Bridge,;
- Assigning the Dynamic IP to access the local network
- Configuring the DHCP server generating the new IPs.
- Virtual cluster creation there is a need to assign a domain name for the Server node.
- Enhancement of either disk space or shared file system and configuring the extra space was not simple

Conclusions

- Many scientists predict that Cloud Computing may replace the traditional ‘own and use’ paradigm of computing.
- It presents the need for cloud lab kit, and its features, architecture and components.

- Supports development and execution of very simple sequential applications, as well as complicated OpenMP/MPI/Hadoop applications.

- This environment can accelerate the adoption of Cloud computing and easy-to-use platform for application development in the cloud.

- Designed to support a wide variety of open source components and useful to achieve different objectives of the users.
-
- Future Work : OpenStack as cloud middleware and more hypervisors including KVM

Contact Details

- Web site : www.sumegha.in
- Email sumeghacloudlabkit@cdac.in

References

Aminatul Solehah Idris, et al. *The Readiness Of Cloud Computing: A Case Study In Politeknik Sultan Salahuddin Abdul Aziz Shah, Shah Alam*- Paper published at International Conference on Computational Science and Technology – 2014 (ICCST'14) Pages: 1 – 5

Yanuarizki Amanatulla, et al. *Toward Cloud Computing Reference Architecture: Cloud Service Management Perspective* - Paper published at International Conference on ICT for Smart Society (ICISS), 2013 Pages: 1 – 4

Vineeth Simon Arackal, et al. SciInterface: *A Web-Based Job Submission Mechanism for Scientific Cloud Computing* Published at Second IEEE International Conference on Cloud Computing in Emerging Markets (IEEE CCEM 2013), October 2013, Bangalore

Shivay Veer Sharma, et al. *Accessing E-Infrastructures using CDAC Scientific Cloud (CSC) Services*, Published at IEEE Cloud Computing for Emerging Markets,16-19th Oct 2013

Chaker El Amrani, et al. *A Comparative Study of Cloud Computing middleware* - Paper published at 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing Pages: 690 – 693

John Bresnahan, et al. *Cumulus: An Open Source Storage Cloud for Science* http://www.nimbusproject.org/files/bresnahan_sciencecloud2011.pdf

Deepanshu S, et al. *Problem Solving Environment for Seasonal Forecast Model on CDAC Scientific Cloud* Second International Conference on Advances in Cloud Computing (ACC 2013), September 2013, Bangalore

Aman A, et al. *Cloud NGS Tool: An online pipeline for NGS datasets* Second International Conference on Advances in Cloud Computing (ACC 2013), September 2013,Bangalore

Payal Saluja, et al. *CDAC Scientific Cloud: On demand provisioning of HPC resources for Scientific Applications*- Paper published and presented at 18th International conference of parallel and distributed processing techniques and applications(PDPTA 2012) 6th-19th July 2012 at Las Vegas, Nevada, USA

C.L.Biji, et al. *NGS read data compression using parallel computing algorithm* - Paper published at International Conference on Bioinformatics and Biomedicine (BIBM), 2015 Pages: 1456 – 1460

Ramesh Naidu Laveti, et al. *Seasonal Forecast Modeling application on GARUDA Grid Infrastructure* at the International Symposium for Grids and Clouds (ISGC-2012), Academia Sinica, Taipei, 26th Feb – 2nd March, 2012.

Thank You!