International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161
©2014 INPRESSCO®, All Rights Reserved
Available at http://inpressco.com/category/ijcet

General Article

NEBULA: Cloud Computer for Universe of Big Data

Saumya Kumari[Á*], Supriya Jha[Á], Lahar Singh Nishad[B̀], Sangita Kumari[Ć]

[Á]Computer Science and Engineering, [B̀]Information and Communication Technology, [Ć]Department of Electronics and Communication Engineering, Jayoti Vidyapeeth Women's University, Jaipur, Rajasthan, India

Accepted 20 Nov 2014, Available online 01 Dec 2014, **Vol.4, No.6 (Dec 2014)**

Abstract

Through this review, I present theories of Cloud Computing and its framework for implementing cloud computing technology across NASA. This review include case study with issues and complexities of using cloud computing platform NEBULA in NASA. It is open source platform for building cloud services (offered by Google, Amazon) using OpenStack (a Linux for cloud computing) for providing instant access to virtual machines for storing huge digital data. For operating Nebula cloud, OpenStack offers 3 pillars for computing, storing and networking which includes services like image management, identity and web interface. Nebula has limited beta and NASA is looking for beta testers to test driving cloud.

Keywords*: Cloud computing, Nebula, NIST, OpenStack, Service-oriented architecture.*

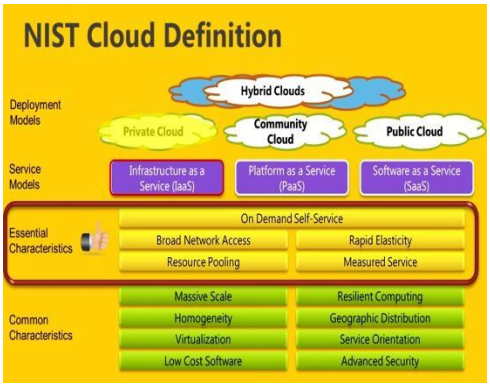
Introduction

It is a date powerhouse which combines cloud computing and data centre containers to compute on-demand power for disrupting and democratizing cloud computing to NASA researchers. Nebula is a self-service platform built from Open source (Open Stack) software providing high capacity computing, storage and network connectivity.

Theoretical Review

Cloud computing is an expansion and extension of already existing disruptive technology taking revolutionary shift from traditional computing delivered as service. Nebula cloud computing model is defined as:

According to NIST cloud computing model



On-Demand Self Services means provision of computing capabilities as needed automatically without requiring human interaction with service's provide. **Broad Network Access** means capabilities are available over network and accessed through mechanism promoting heterogeneous client platforms. Service Oriented Architecture (SOA) approach is adopted to provide an interoperable interface for all devices. **Resource Pooling** provides computing resources that are pooled to serve multiple consumers. Nebula virtual machine offer remote and distributed capabilities without disclosing physical location. Nebula node can also be created supporting parallel processing. **Rapid Elasticity** means ability of provisioning capabilities to scale out quickly and rapidly releasing them to scale in. To compose with other VMs, Nebula VMs needs to be SOA-enabled. **Measured Service** means implementation of algorithm for controlling and optimising resource for a service. A embedded service inside Nebula Node needs to be developed for monitoring physical and logical resource

IaaS, Saas, Paas service data-centric application for real tim

A Plumbing i.e. layering foundation is done in Infrastructure as a service (IaaS) for application running on cloud. These Iaas service used by PaaS to develop SaaS instance i.e. Application on Nebula such as federated data management service, a streaming service with filtering and aggregation operators, temperature, a interoperability service.

Elastic use of Nebula cloud resource by adhoc networks

Ex-user picks data from sensor network and later deallocate VMs on fixed cloud infrastructure using routing scheme (ex-Epidemic routing where path is unknown from source to destination).Some data streaming from isolated

*Corresponding author: **Saumya Kumari**

sensor network connected to cloud periodically by satellite connectionn.

Elastic incorporation of ad-hoc resource in cloud

Sharing of Virtual machine hosted on adhoc nodes are participated by cloud service. A first device computational resource is pooled by secondary device.

Nebula built on DTN (Delay Tolerant Network) application layer

A set of protocol based on store and forward communication ex-Bundle protocol extension is required at session and application layer.

DTN environment includes

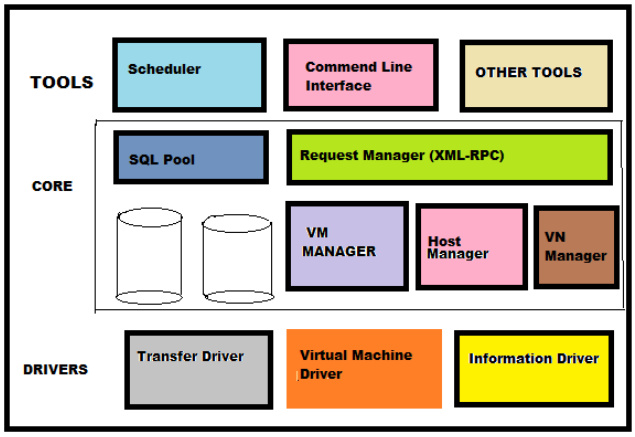
- Low-propagation delay-DTN bundle agent utilise Internet protocol for negotiating connectivity in real time. Ex- in planetary surface environment.
- High-propagation delay-To enables connectivity between 2 agents DTN bundle agent use some form of scheduling. Ex-in deep space.

NASA's Cloud Computing platform Nebula supporting NASA for viewing and exploration of moon and mars by uploading hundred's of high resolution images over 100 terabytes data. Nebula is designed to port data sets and code. It saves time and labour. Nebula's services allow flexibility for NASA in mission stages and needs of extension, timelines with delays and cancellation. Nebula is supporting Federal government websites for storage as data grows.

About Nebula

Nebula deploys private cloud computing infrastructures. Nebula has developed a hardware appliance which allows business to build a private computing cloud from computers. Nebula was founded in April 2011 by Chris C. Kemp at NASA Ames Research Centre. Nebula's mission is to ignite a new era of global innovation by laying the foundation of the coming industrial revolution of big data.

Nebula Architecture



Nebula requirements

Characteristics	
Type of Environment	Cloud
Main Memory	96 GB
Network Interconnect	Cisco Nexus 7000 , 10 GigE switch
Network Topology	Cisco proprietary
Number of Sockets	2
Cores per sockets	6
Cores per node	12
Compiler	Intel 11.1
Processor Type	Intel Westmere (Xenon X5660)
Processor Speed	2.80 GHz
Hypervisor	Kernel virtual machine(KVM)

Some Case Studies

Jet Propulsion Lab (JPL) focus on Mars

The portal BeAMartian.jpl.nasa.gov developed by MS Azure using API to connect visitors of website with pictures of Mars without any additional data storage on JPL computers. User can see pictures, videos, post questions read response and send message.

Enterprise data centre strategy

NASA's re-evaluation of enterprise data centre strategy for outsourced data centre services by delivering the use of NEBULA. Ex-Message Passing Interface (MPI) by Flight Vehicle Research and Technology Division.

Data Sources for Nebula

(a)Internal Data-Nessus data for producing detailed vulnerability information about data hosts on network. It is a basis of CVSS scores.Patch link data for producing detailed information about patch status for hosts on network. It is a basis for patch status scores.
(b)External Data-Nessus external scanner dumps DNS tree and scan all ports on all hosts from external posture for self-discovery for exposure to outside world. TCP/Netflow data (SYN/ACK), Google Search results(search API)
(c)Operational Data-It dumps asset database, DHCP log files and MAC association.
(d)Intel Data- It determines threat sources and IR (incident response) tools to look for hostname and sys-admin which lead to increase risk factor.

Uses

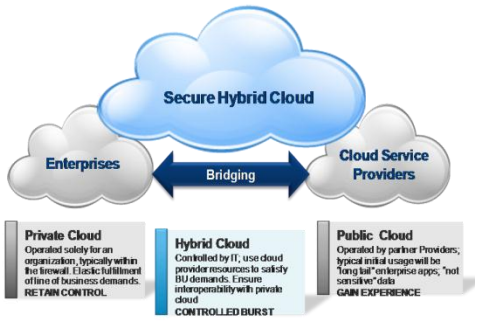
- Nebula is used to create virtual workstations to give software developer more control over developed environment for sharing modules and library over cloud.
- Nebula is used for collaboration with non-NASA partner (Microsoft, Amazon) via FTP and running web based application which helps in analyzing data produced by NASA's Airspace concept evaluation system.

Benefits of working with Huge Data

(a)Security-Hybrid cloud offer protection by security services such as Intrusion Prevention, Web Application Firewalls, File Integrity Monitoring, and Event Management. This environment allows adding layers of security.

(b)Performance-varying nature of big data requires infrastructure flexibility and elasticity. The main drawback here is Cloud Bursting. It allows spinning up new workloads when information from system signals the need for additional resource avoiding jeopardizing workload performance. Hybrid cloud allows cloud bursting on any scale and offer adaptable cloud solution for managing and storing big data.

(c)Saving-Hybrid cloud environment allow for adding resource. Public cloud component of hybrid environment allows financial flexibility of spinning up additional resource and results in financial savings. Private cloud component provide resources required for data processing.



Embrace the benefits of Cloud Computing via Private Clouds, then Hybrid Clouds

Thus, this architecture of hybrid cloud provides security through layers of security services, optimal performance through cloud bursting capabilities and financial savings through flexible resource offerings.

Some advantages and disadvantages of Nebula Cloud computing platform

	Advantages	Disadvantages
1.	**Scalability**-on-demand provisioning of computing resources.	**Security**-loss of control over sensitive data.
2.	**Accessibility**-location and device independence.	**Integration**-difficult integration with other system
3.	**Redundancy**-Redundancy of sites is easier to implement.	**Dependency**-tied to cloud service provider
4.	**Multi-tenancy**-several customer sharing same infrastructure.	**Cost**-opaque cost structure.
5.	**Maintenance**-Upgrades are applied centrally by IT experts.	**Knowledge**-most knowledge is about cloud service provider.
6.	**Cost**-Transformation of capital expenditure for servers into and operating expense.	**Flexibility**-special customization of computing resource is not possible.

Nebula provide 3 class of storage

(a)Local Storage- Nebula use swappable commodity drives in a hardware RAID configuration. Virtual machine use local storage to run application.

(b)Persistent Block Drive (iSCSI)-Nebula use iSCSI to provide persistent network storage block device used by conventional application and decouples the storage from connected server.

(c)Object Store-Easy storage of petabytes of data and million files. Open-source implementation of object stores used with custom code adds in Access Control Layer (ALC).

Research Challenges

- **Universal Adoption**-Stakeholders object to openness of tool. Everyone is not open to the idea of open.
- **Fairness**-Stakeholders must be convinced there is no bias in scoring system.
- **False Positive**-A robust system is needed for dealing with this cloud.
- **DHCP/NAT** (Dynamic Host Configuration Protocol/Network Address Translation)-The consistent attribution of hosts across various IPs is required on various dates.
- **IPv6-**To do discovery scans across hosts. Infrastructure MUST include network monitoring and aggregating the V6 auto-configuration logs.
  - Scan on demand-using API
  - Scan on Connect- Tying DHCP/IPv6 auto-configuration logs to scan initiation.
  - Score-based and status-based situational gaming for sysadmins.

Conclusion

Nebula cloud support middleware i.e. design and implementation of virtual Nebula Node(NN) and lightweight Nebula Node and integration of Nebula Node with SOA technology such as OSGI(Open Service Gateway Initiative), integration of Nebula Node with event technology such as JMS(Java Message Service).

Nebula promotes machine- to- machine intelligence, location based and personalised services. My findings includes 2 folds-

(a) Utilisation of virtualisation layer in cloud computing platform to support on-demand access.

(b) Low performance of 10GigE network used in cloud computing systems to low latency high bandwidth inter-connects used in Supercomputers

Current Technology suffers from resource poverty and lack of maturity.

References

The Internet of Things- http://www.buildyourbestcloud.com /three-things-you-might-not-know-about-openstack

NASA:http://open.nasa.gov/blog/2012/06/04/nebula-nasa-and-openstack/, http://www.nasa.gov/open/plan/nebula.html, p://www.nasa.gov/multimedia/imagegallery/image_feature_2526.html#.VAasEGP5O74

IEEE:http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6167432&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6167432

http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=6332192&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D6332192

Smart Grid-http://smartgrid.ulitzer.com/node/975079

Wired-http://www.wired.com/2013/04/nebula-one-chris-kemp/

http://orgviz.ulitzer.com/node/97507

Steve Heistand, Subhash Saini , Rupak Biswas (NASA Supercomputer divisor): Performance Evaluation of NASA's NEBULA cloud computing Platform,