Big Data for the Private Cloud

Manzoor S. Brar
Solutions Marketing Manager

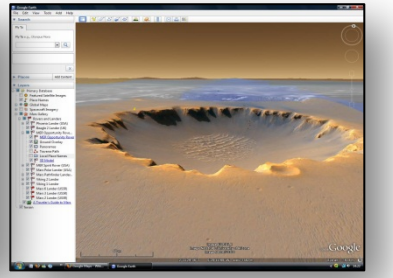
May 12, 2014

nebula®

(cloud) Computing for the Enterprise

About Nebula

- Nebula, Inc. continues on the legacy of work started at NASA Ames Research Center in 2006
  - **Goal:** Develop a compute and storage platform to process petabytes of space exploration data from NASA satellites and other research projects



- Company Launched in July 2011
- **Product:** Nebula One Private Cloud
  - Scale-out “Infrastructure-as-a-Service”

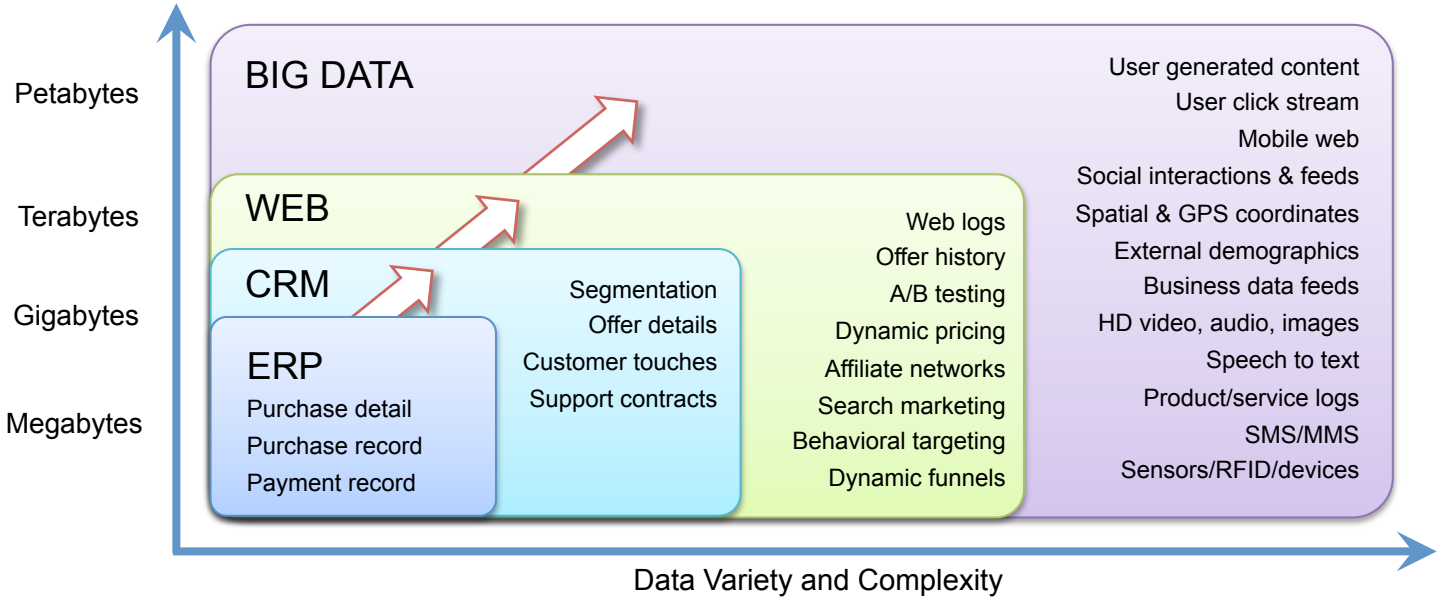
Introducing Nebula One

- **Appliance Model Approach:** Simple, predictable, and secure performance for the private cloud
- **Converged Systems Approach:** Optimized hardware and software enabling 'scale-out' computing
- **On-demand and Self-service:** Accelerates multi-tenancy behind the enterprise firewall
- **Built on OpenStack:** Open-standards for 'pooling' compute, storage, and network resources
- **Turnkey Package:** Orchestrate infrastructure cost-effectively

What's Big Data?

3 Vs: Volume, variety, and velocity

Big Data = Transactions + Interactions + Observations *



*** Source of graphic: Hortonworks and Teradata, Inc.**

Big Data Pain Points for IT & Business

Bare Metal

- Bare metal big data clusters are expensive
- Days or weeks to orchestrate/provision pilot POC
- Time/Cost leads to shadow IT groups to circumvent IT

Shadow IT

- *Risk of security breach*
- *Potential data loss and loss of overall IT control*
- *Overall system inconsistency*
- *Errors by non-technical staff managing it (HR, mktg, etc.)*
- *Possible regulatory compliance violations, and more*

Public Cloud

- Connected to public internet
- Shared among many users from many companies
- Not predictable – “noisy neighbor” issue
- No way to monitor, WAN bandwidth as chokepoint

IT & Business Needs

Agility, Scale

- Efficiently and quickly deliver services
- Turnkey solution for distributed compute, storage and network services
- Cost-effective, unified system

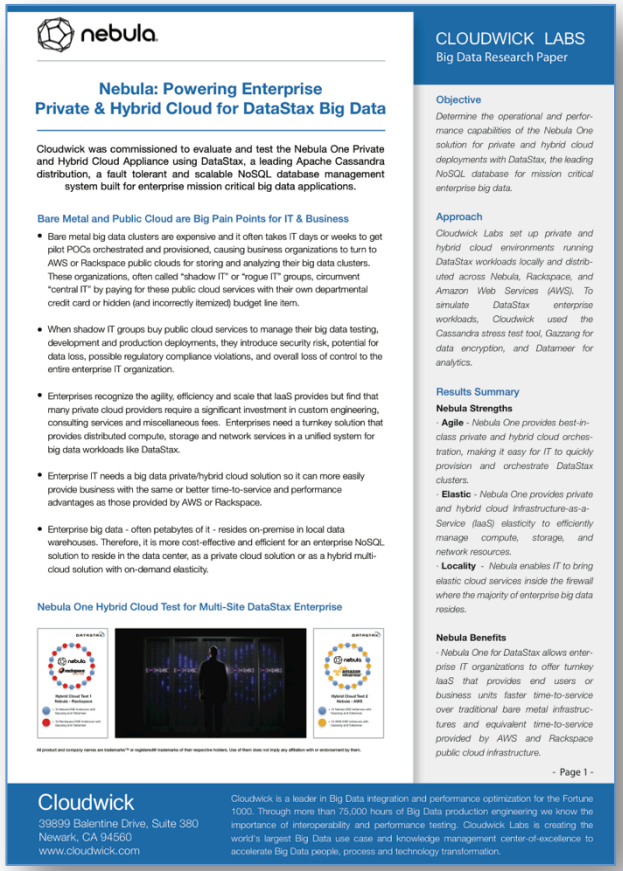
Efficiency

- Quickly & easily provide business units with same or better time-to-service and performance advantages of AWS
- Petabytes of big data
  - On-premise NoSQL solution more cost-effective and efficient than off-premise

Security

- Private big data solution on-premise is most secure
- Local data warehouse behind enterprise firewall
- Remove the temptation of rogue/shadow IT – make it easy

Tested and Validated: Multi-cloud Interoperability



nebula

Nebula: Powering Enterprise Private & Hybrid Cloud for DataStax Big Data

Cloudwick was commissioned to evaluate and test the Nebula One Private and Hybrid Cloud Appliance using DataStax, a leading Apache Cassandra distribution, a fault tolerant and scalable NoSQL database management system built for enterprise mission critical big data applications.

Bare Metal and Public Cloud are Big Pain Points for IT & Business

- Bare metal big data clusters are expensive and it often takes IT days or weeks to get pilot POCs orchestrated and provisioned, causing business organizations to turn to AWS or Rackspace public clouds for storing and analyzing their big data clusters. These organizations, often called "shadow IT" or "rogue IT" groups, circumvent "central IT" by paying for these public cloud services with their own departmental credit card or hidden (and incorrectly itemized) budget line item.
- When shadow IT groups buy public cloud services to manage their big data testing, development and production deployments, they introduce security risk, potential for data loss, possible regulatory compliance violations, and overall loss of control to the entire enterprise IT organization.
- Enterprises recognize the agility, efficiency and scale that IaaS provides but find that many private cloud providers require a significant investment in custom engineering, consulting services and miscellaneous fees. Enterprises need a turnkey solution that provides distributed compute, storage and network services in a unified system for big data workloads like DataStax.
- Enterprise IT needs a big data private/hybrid cloud solution so it can more easily provide business with the same or better time-to-service and performance advantages as those provided by AWS or Rackspace.
- Enterprise big data - often petabytes of it - resides on-premise in local data warehouses. Therefore, it is more cost-effective and efficient for an enterprise NoSQL solution to reside in the data center, as a private cloud solution or as a hybrid multi-cloud solution with on-demand elasticity.

Nebula One Hybrid Cloud Test for Multi-Site DataStax Enterprise

CLOUDWICK LABS
Big Data Research Paper

Objective

Determine the operational and performance capabilities of the Nebula One solution for private and hybrid cloud deployments with DataStax, the leading NoSQL database for mission critical enterprise big data.

Approach

Cloudwick Labs set up private and hybrid cloud environments running DataStax workloads locally and distributed across Nebula, Rackspace, and Amazon Web Services (AWS). To simulate DataStax enterprise workloads, Cloudwick used the Cassandra stress test tool, Gazzang for data encryption, and Datameer for analytics.

Results Summary

Nebula Strengths

· **Agile** - *Nebula One provides best-in-class private and hybrid cloud orchestration, making it easy for IT to quickly provision and orchestrate DataStax clusters.*

· **Elastic** - *Nebula One provides private and hybrid cloud Infrastructure-as-a-Service (IaaS) elasticity to efficiently manage compute, storage, and network resources.*

· **Locality** - *Nebula enables IT to bring elastic cloud services inside the firewall where the majority of enterprise big data resides.*

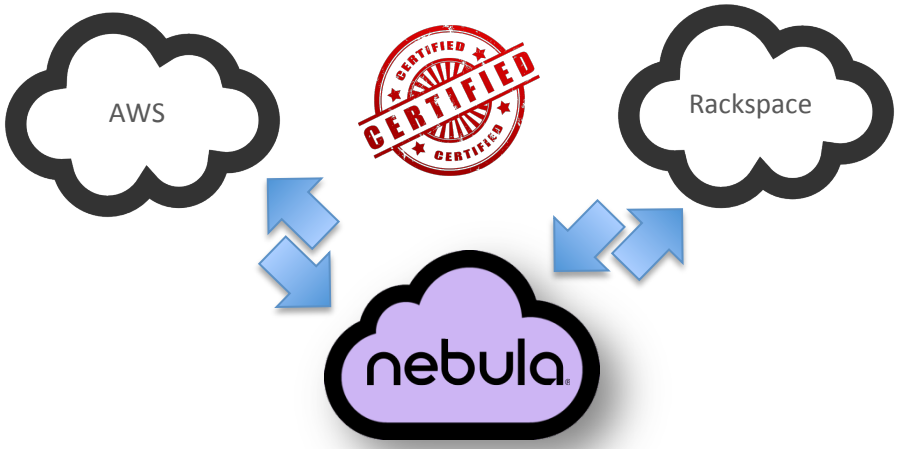
Nebula Benefits

· *Nebula One for DataStax allows enterprise IT organizations to offer turnkey IaaS that provides end users or business units faster time-to-service over traditional bare metal infrastructures and equivalent time-to-service provided by AWS and Rackspace public cloud infrastructure.*

- Page 1 -

Cloudwick
39899 Balentine Drive, Suite 380
Newark, CA 94560
www.cloudwick.com

Cloudwick is a leader in Big Data integration and performance optimization for the Fortune 1000. Through more than 75,000 hours of Big Data production engineering we know the importance of interoperability and performance testing. Cloudwick Labs is creating the world's largest Big Data use case and knowledge management center-of-excellence to accelerate Big Data people, process and technology transformation.



- Cloudwick Labs, an independent test lab, confirmed that Nebula delivers Cassandra workloads to span multiple data centers and validates multi-site hybrid cloud interoperability with Amazon Web Services (AWS) and Rackspace.

Cloudwick Conclusions

- Cloudwick Labs determined that Nebula:
  - Delivers Cassandra workloads across multiple data centers and validates multi-site hybrid cloud interoperability with AWS and Rackspace
  - Offers ease of provisioning and orchestration of clusters
  - Enables IT to provide business with
    - **faster time-to-service** than bare metal
    - **at the same speed** as Amazon and Rackspace
  - brings on-demand compute and storage elasticity **behind the enterprise firewall**
  - Enables IT to rapidly provide internal stakeholders with critical infrastructure services, while maintaining control by **minimizing shadow IT operations and security risks**

Agility, Scale

Efficiency

Security

Turnkey Big Data Private Cloud

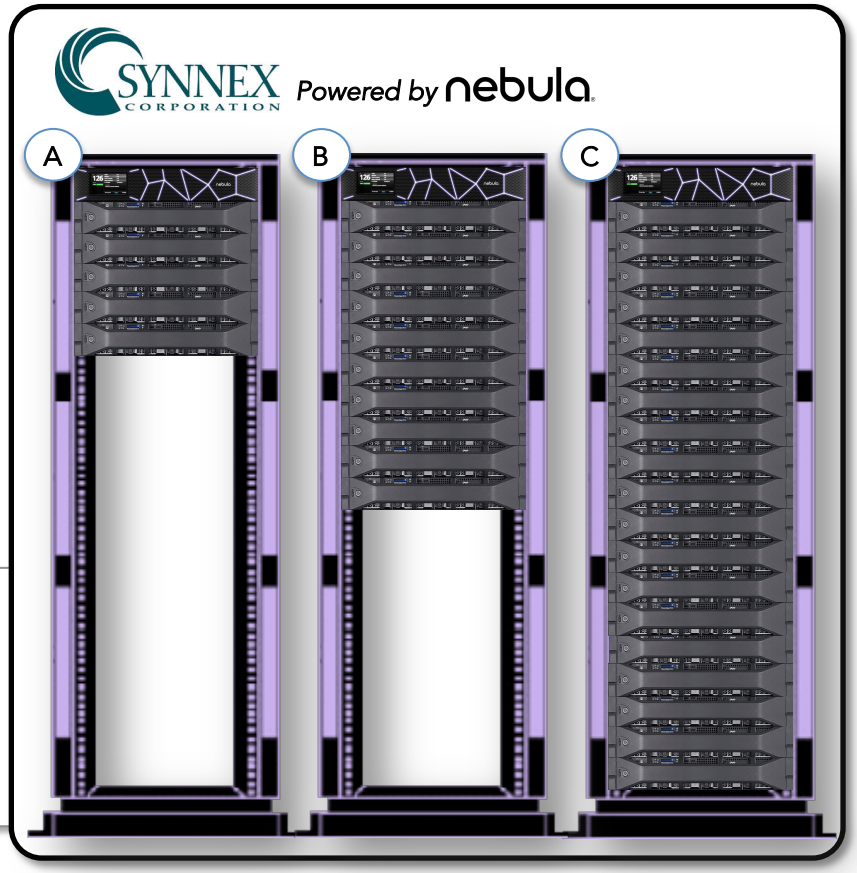
- Synnex Corporation brings to the channel a fully integrated Private Cloud Rack Appliance to tackle Big Data
  - Powered by Nebula
  - Bundled with Cassandra and MongoDB
    - Software integration by Cloudwick
  - Utilizing industry-standard servers

Turnkey Big Data Private Cloud

- Solutions provider can source three configurations
  - A Entry Level *(Nebula One Controller + 5 servers)*
  - B Half Rack *(Nebula One Controller + 10 servers)*
  - C Full Rack *(Nebula One Controller + 19 servers)*
  - *Larger configurations available upon request*
- Pre-configured with:

Cassandra mongoDB

About SYNNEX
SYNNEX Corporation (NYSE: SNX), a Fortune 500 corporation, is a leading business process services company, servicing resellers, retailers and original equipment manufacturers in multiple regions around the world. The Company provides services in IT distribution, supply chain management, contract assembly and global business services. Founded in 1980, SYNNEX employs approximately 49,000 full-time and part-time associates worldwide. Additional information about SYNNEX may be found online at www.synnex.com.



Nebula One

DevOPs/Test

Social Collaboration

Bio-Tech

Big Data + Analytics

Other …

Mobile Apps

Cassandra

mongoDB

Turnkey Infrastructure as a Service (IaaS)

Management/Orchestration/API Interface

Identity/Security

Compute

Storage

Block

Object

Network

openstack

1000+ companies contributing to OpenStack

Industry Standard Servers

Thank You!

- Please check out our breakout session
  - Speakers: Chris Kemp and Vish Ishaya
  - Topic: Introducing Nebula. Experience Turn-key Private Cloud
  - Time: Monday May 12 @ 2:50pm
  - Location: B314

- Visit us at Booth #C3 to learn more.

THANK YOU!