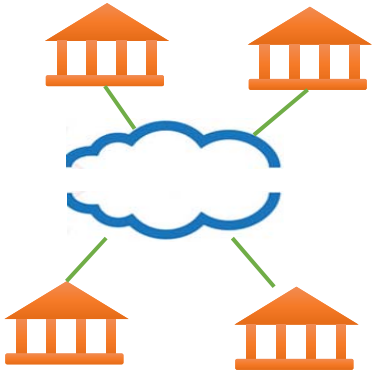
Is Government Ready for Public Cloud?



Adisak Srinakarin

Executive Vice President

Electronic Government Agency (Public Organization)

22 September 2016

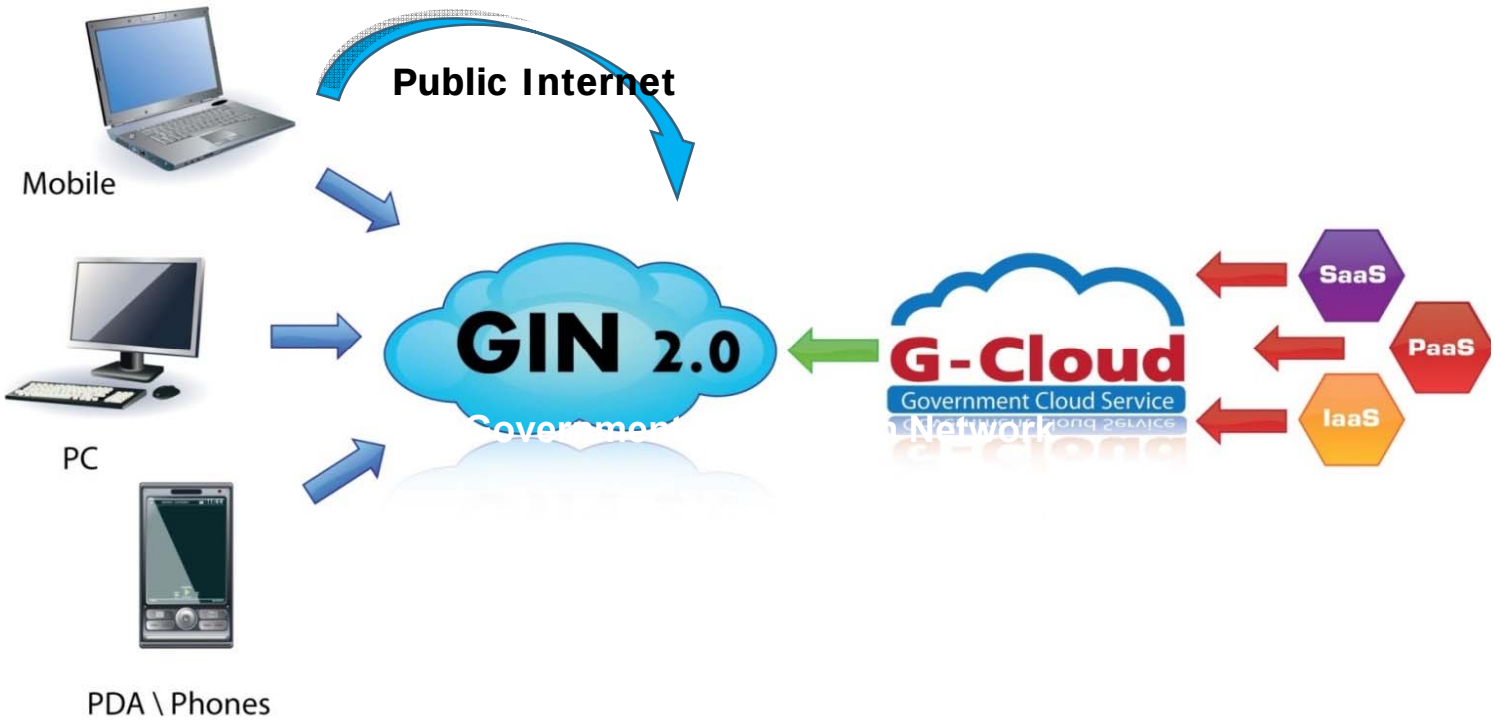
EGA
e-Government Agency

Government Cloud (G-Cloud)

Objective

1) To reduce the redundancy in government budget proposals.
2) To ensure the security system meets all current IT-related laws and standards.
3) To ensure the system is adequately treated with proper management and maintenance from specialized experts for better levels of reliability and service availability.

G-Cloud Scenario



G-Cloud Standard

Security
- Virtual Firewall
- Antivirus

SLA Uptime 99.5% ⟵ Guarantee

Contact Center 24x7
- NOC Monitoring

Reliability

- SLAs 99.5%
- Security
- Environment
- EGA Contact Center
- Data Protecting
- Vulnerability Scan
- Backup System

EGA
e-Government Agency

Type of Service

Operating System

Windows Server 2012

Windows Server 2008 R2

CentOS

ubuntu

Database

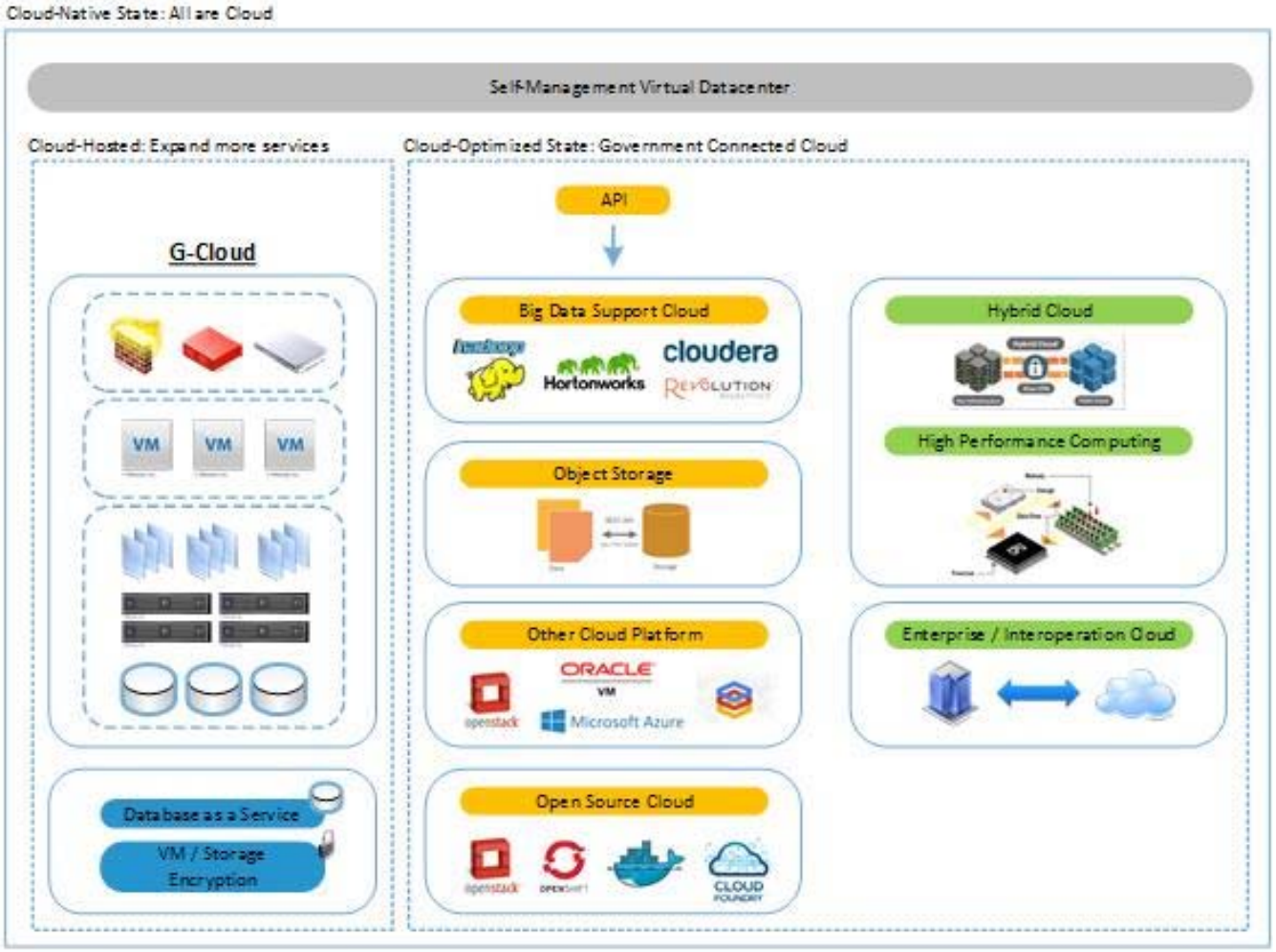
Microsoft SQL Server

Standard Package
- 2 vCPU, 4GB vMem, HDD 200GB

Customize Package
- As you request


Is Government Ready for OpenStatck??

Cloud Roadmap



2560	Cloud-Hosted: Expand more services
2561	Cloud-Optimized State: Reduced Backward compatible Agile
2562	Cloud-Optimized State: Government Connected Cloud
2563	Cloud-Native State: All are Cloud

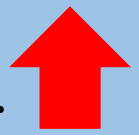
EGA
e-Government Agency

Benefits

- Open-source and much more flexible and vendor-neutral cloud environment. As a result, we can lower our costs, avoid the risks of vendor lock-in, and add new capabilities and approaches much more quickly and easily.
- A complete solution. OpenStack already includes computing, networking, storage, and other essential cloud elements, already integrated and interoperable.
- The best developers and engineers are already familiar with OpenStack. We just have to pick the staff.
- Hardened, enterprise-class OpenStack solutions are now available from open-source leaders like RedHat, Mirantis, etc.

Concerns

Finding Experienced Software system engineers is still a challenge in Thailand.

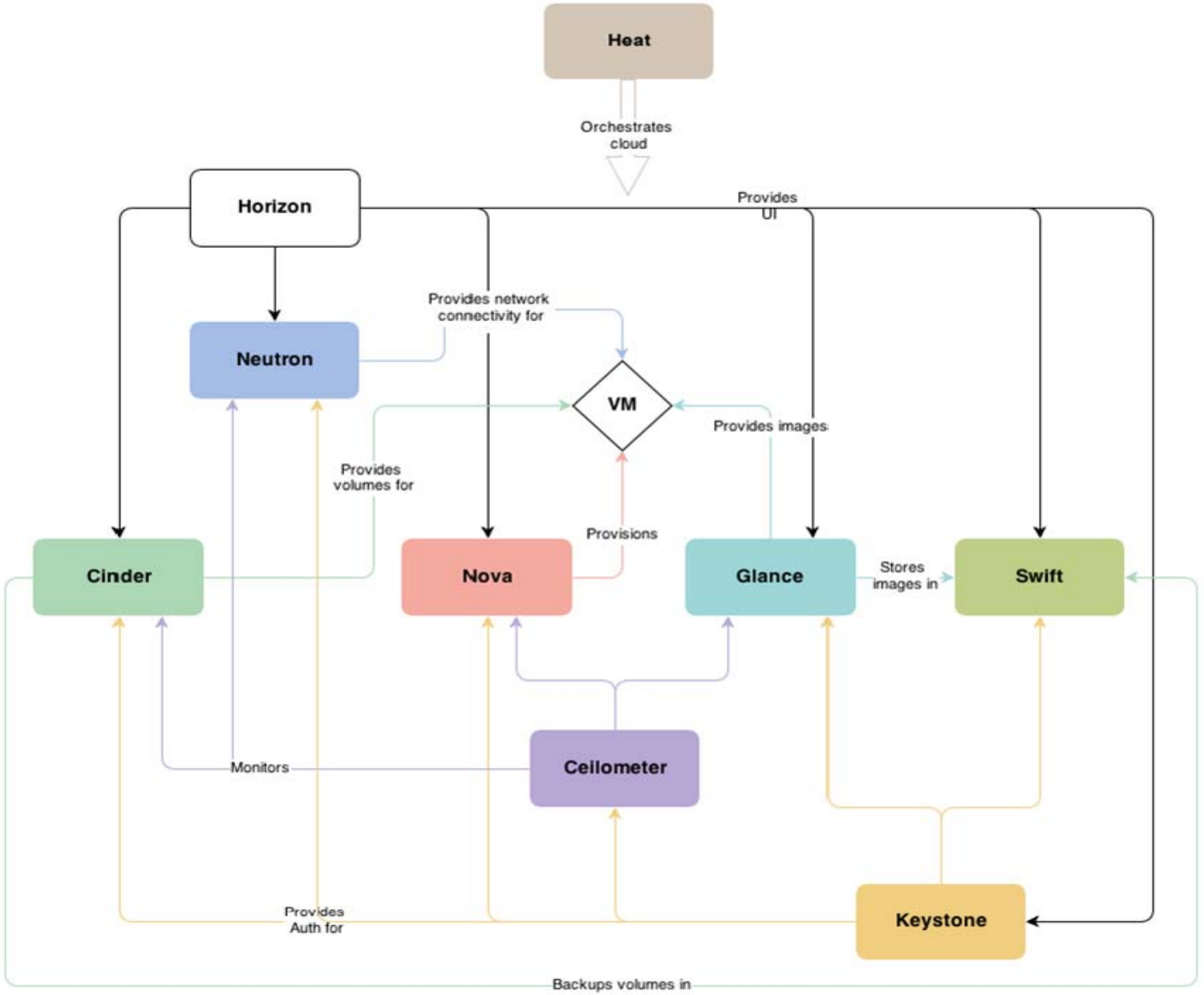
High Turnover rate of staff. 

Is OpenStack mature enough to compete with some commercial products like VmWare?

EGA OpenStack (Reloaded)

- ✓ Specifically dedicated team
- ✓ Kick-start with the highly experienced team (engineers from the OpenStack companies) so that we can evaluate this approach quickly (configuration, migration, automation, etc.)

OpenStack for Infrastructure as a Service



EGA
e-Government Agency

OpenStack for Big Data

Big Data

Using data processing model to process the data and transfer it become high value.

Cloud Computing

A shared resources infrastructure to support a flexible IT environment and fulfill the requirement on demand.

OpenStack meets Hadoop

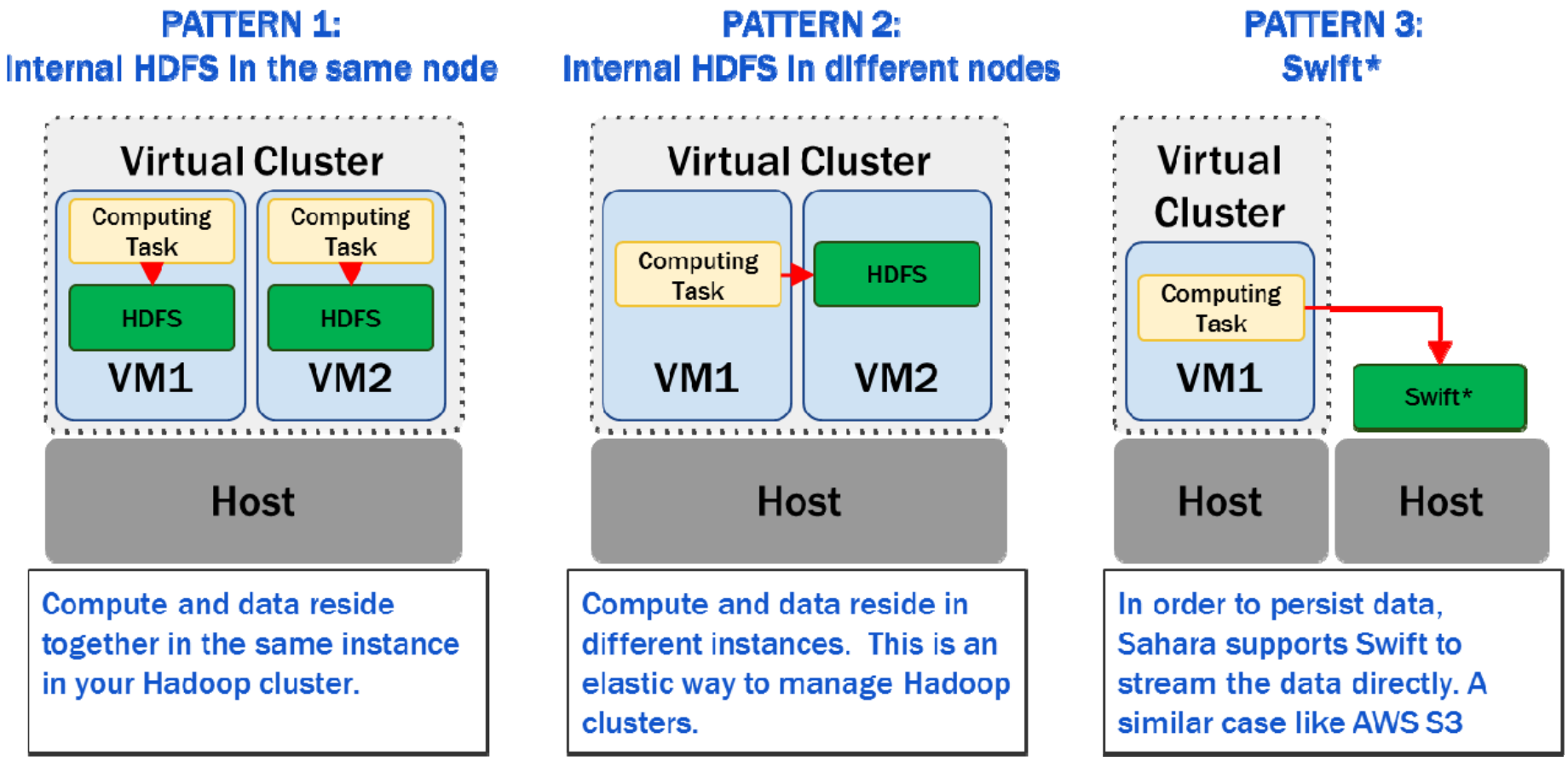
- *Most Companies using OpenStack cluster in their IT environment are also preparing another Hadoop cluster for Big Data analytics.*
- *Sahara is the antidote to bring Hadoop into OpenStack.*
- *Sahara background*
  - ✓ Basic Idea comes from Amazon Elastic MapReduce (EMR)
  - ✓ To provide users easily provisioning Hadoop clusters by specifying several parameters
  - ✓ Analytics as a Service for data scientist or analyst

Amazon EMR vs OpenStack Sahara

Items	Amazon EMR	Sahara
Cluster Provisioning	Use EC2 (Create, Terminate, Clone, Resize)	Use Heat&Nova (Create, Terminate, Clone, Resize)
Transient Cluster	Yes	Yes
Hadoop Distribution	• Amazon Hadoop v1.0.3 and v2.4.0 • MapR v2.1.3, v3.0.2, v3.0.3, v3.1.1, v4.0.2	• Apache Hadoop v1.2.1 and v2.6.0 • MapR v4.0.2 • CDH v5.3, CDH v5.4(Liberty) • HDP v1.3.2, v2.0.6
Application Installed*	Hive, Pig, Hue, HBase, Spark, Impala, Ganglia, Hank	Hive, Pig, Hue, HBase, Spark, Impala, Key-value store indexer, Solr, Sqoop
Data Source	HDFS, Object Storage(Amazon S3)	HDFS, Object Storage(Swift), Manila(Liberty)
Job Management	Yes(Using Steps)	Yes(Using Sahara EDP)
Job Scheduling	Yes(Using Amazon Data Pipeline)	Yes(Liberty)
Data Flow Control	Yes(Using Amazon Data Pipeline)	No(Planning)
Hadoop HA	HDFS HA	HDFS HA(Liberty), Yarn HA(Liberty)
Health Monitoring	Yes(Including bandwidth, throughput, ...etc)	No(Planning)

Sahara use cases (1/5)

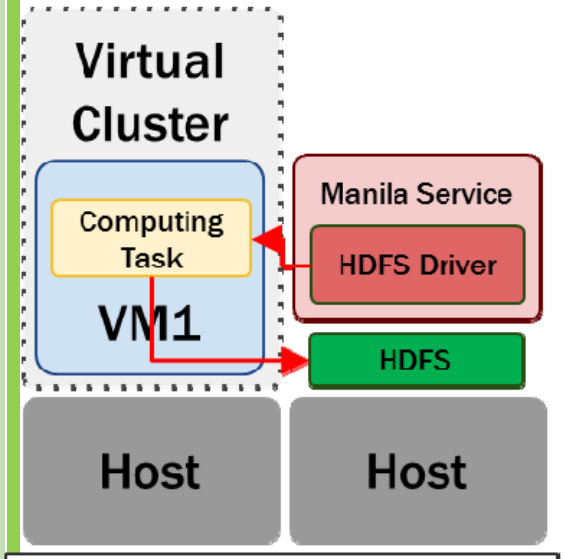
Data processing model (1/2)



Sahara use cases (2/5)

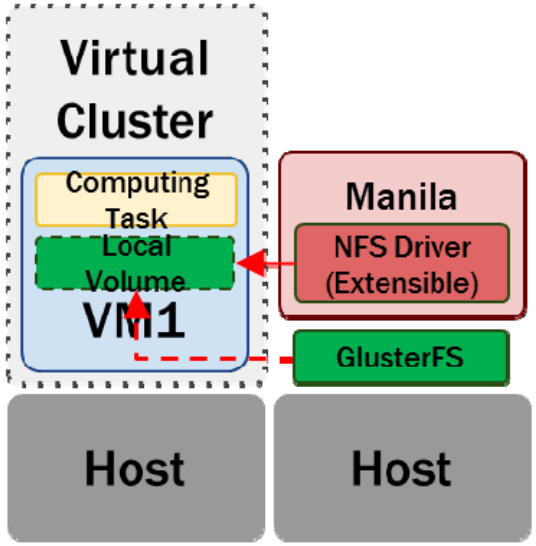
Data processing model (2/2)

PATTERN 4: External HDFS via Manila*



Sahara can support external HDFS by using the HDFS driver in Manila.

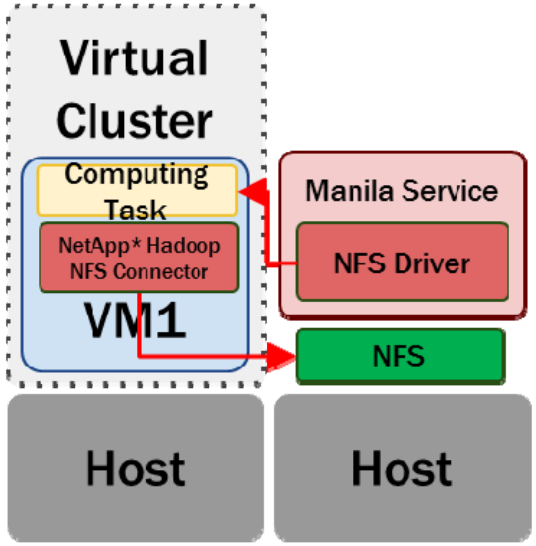
PATTERN 5: Local Storage with Diverse Storage Backend in Manila



Use local storage in Hadoop and remote mount any type of file storage in Manila.

PATTERN 6: NFS

This feature will be implemented in Mitaka



NetApp Hadoop NFS Connector can bring the NFS capability into Hadoop.

EGA
e-Government Agency

Sahara use cases (3/5)

Scenario 1: Set up data for persistent

- Requirements

  Data must be stored in a persist storage

- Solutions

  For Pattern1&2, use HDFS as a data source and use Cinder as a persist storage backend

  For Pattern 1&2, use HDFS as a data source with ephemeral and Swift as a back up storage

  For Pattern 3, use Swift as a data source and stream the data directly

  For Pattern 4~6, use Manila as a data source and integrate with individual storage backend

EGA
e-Government Agency

Sahara use cases (4/5)

Scenario 2: Running a small task

- Requirements

Customers may run a small task in secs, but to provision a cluster takes minutes

- Solutions

Running a long run cluster for small tasks. A long run cluster can use HDFS as a persist data store, but we still recommend to use Cinder or Swift as a backup storage

Sahara use cases (5/5)

Scenario 3: Cloud storage integration

- Requirements

  Customers would like to leverage current cloud storage architecture like Ceph, Cinder, or Swift.

- Solutions

  For Pattern 1&2, Cinder can be support as a HDFS data store

  For Pattern 3 to integrate with Swift

  Use Pattern 4~6 to integrate with Ceph or Other storage type

EGA is hiring

“Software Engineers”

for OpenStack development

www.ega.or.th

contact@ega.or.th

Openstack Thailand Chapter

(https://www.facebook.com/groups/204019566447627/?fref=ts)