A Better VM HA Solution:
Split-brain Solving & Host Network Fault Awareness

Jiang WU

Nov 14th, 2018

Agenda

- *Introduction*
    - *Basic Principles*
        - *Key Features*
            - *More Concerns*
                - *Tests & Others*

Introduction

Introduction

- VM HA(*High Available*) is **still** an important feature, especially for legacy services
  - Still unreformed/cannot be reformed in short term

- The **disadvantages** of traditional HA solutions
  - Rely on IPMI
  - Can only handle single scenes
  - Almost no solution to the "split-brain" problem

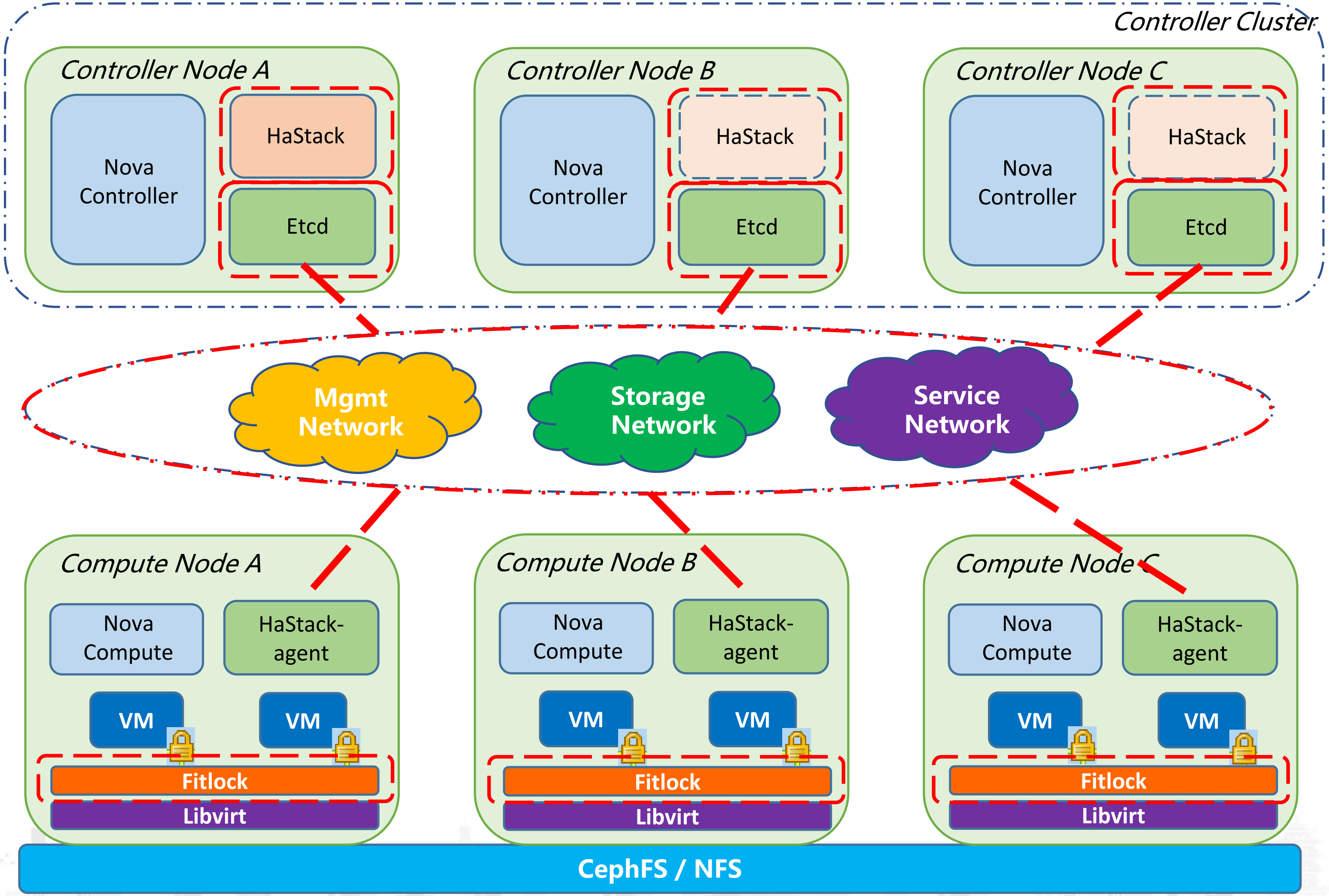
Objectives

- ## Design Requirements
  - **Integrate** with FitOS v3.3
    - *Fiberhome IaaS Cloud Platform based on OpenStack since 2015*
  - **Independent** of OpenStack
  - **Try not to modify native codes** AMAP
  - *Easy to use, easy to maintain*

- ## Feature Requirements
  - Solve the "split-brain" problem
  - Flexible handling of some complex faults

Basic Principles

Architecture



Composition

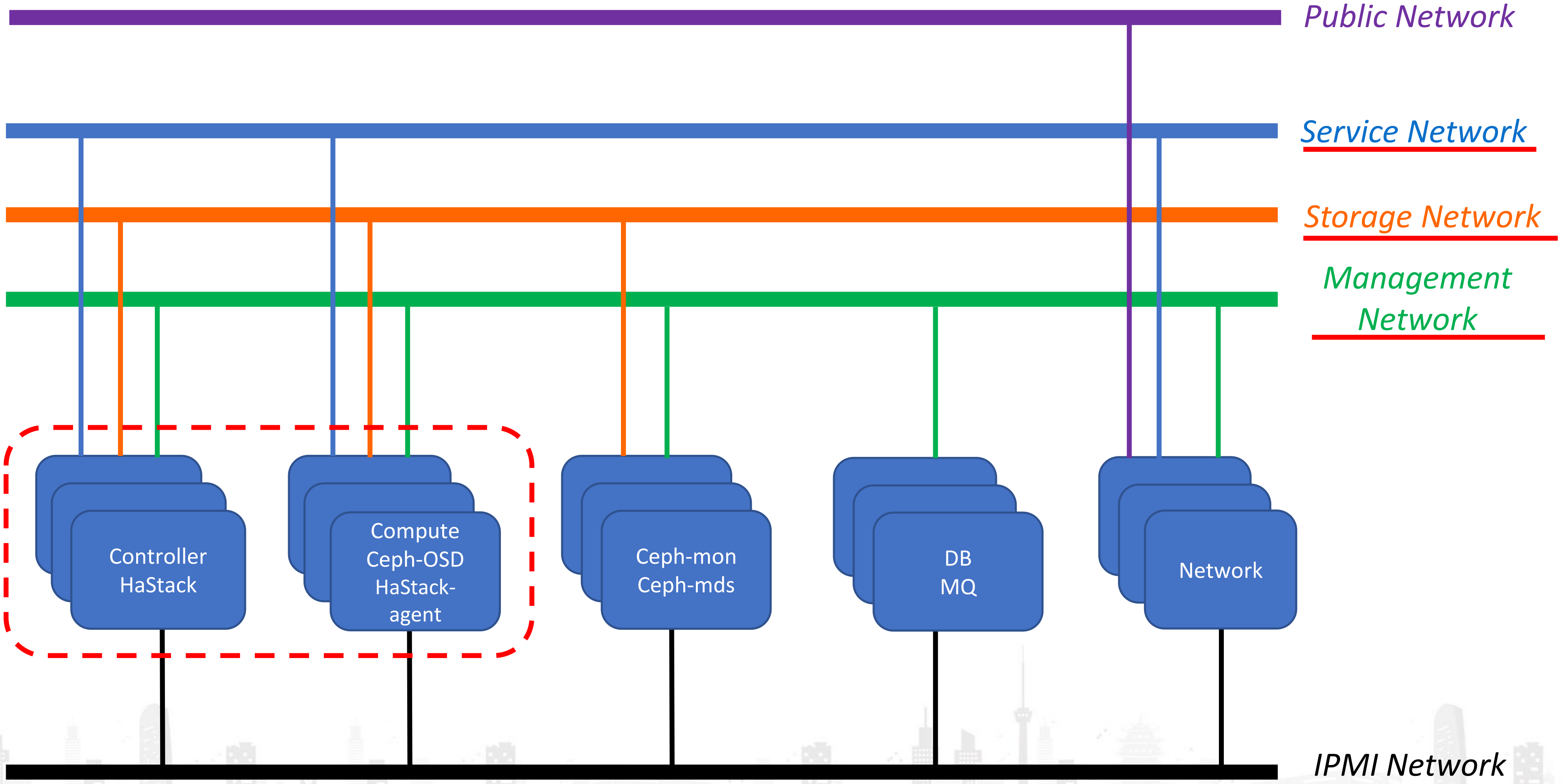
- CentOS 7.4 + OpenStack
- Shared Storage: **CephFS/NFS**

- **HaStack**: HA controller
  - Controller Node
  - *A-S*
- **Fitlock**:
  - A lock-manager, for "split-brain" protection
  - *Compute Node*
- **Etcd**:
  - To provide 3 network plain health detection
  - Controller Node
  - *A-A*

New Components

Components Name	Positions	Deployment requirements	Reliability Requirements	Components Description
HaStack	Controller Node	3 nodes by defaults, A-A	Not allowed to fail at the same time	To control the entire HA system
Etcd	Controller Node	3 nodes by defaults, A-A 3 clusters	Two processes are not allowed to fail at the same time in the same cluster	1. To establish a 3 network plane cluster and sense the global health status for HA decision 2. As a message bridge between HaStack and HaStack-agent
HaStack-agent	Compute Node	Single node	-	To complete partial HA management with HaStack: 1. Mount the shareable folder 2. Report the heartbeat status of the node and the VM Fencing event
Libvirt	Compute Node	Single Node	-	-
Fitlock	Compute Node	Single Node	-	A lock manager like Sanlock, work with Libvirt to complete registration and heartbeat updates for each lock resource on shared storage
CephFS	Storage Node	1. Sharable folder: mount to each compute nodes 2. Ceph-mon: 3 nodes by defaults, A-A 3. Ceph-mds: 3 nodes by defaults with ceph-mon, A-S	1. Shareable folder: Ceph 3 copies； 2. Ceph-mon: 2 processes are not allowed to fail at the same time 3. Ceph-mds: not allowed to fail at the same time	Ceph components(ceph-mds, ceph-mds), to provide shared file system storage for storing lock files

Deployment Model



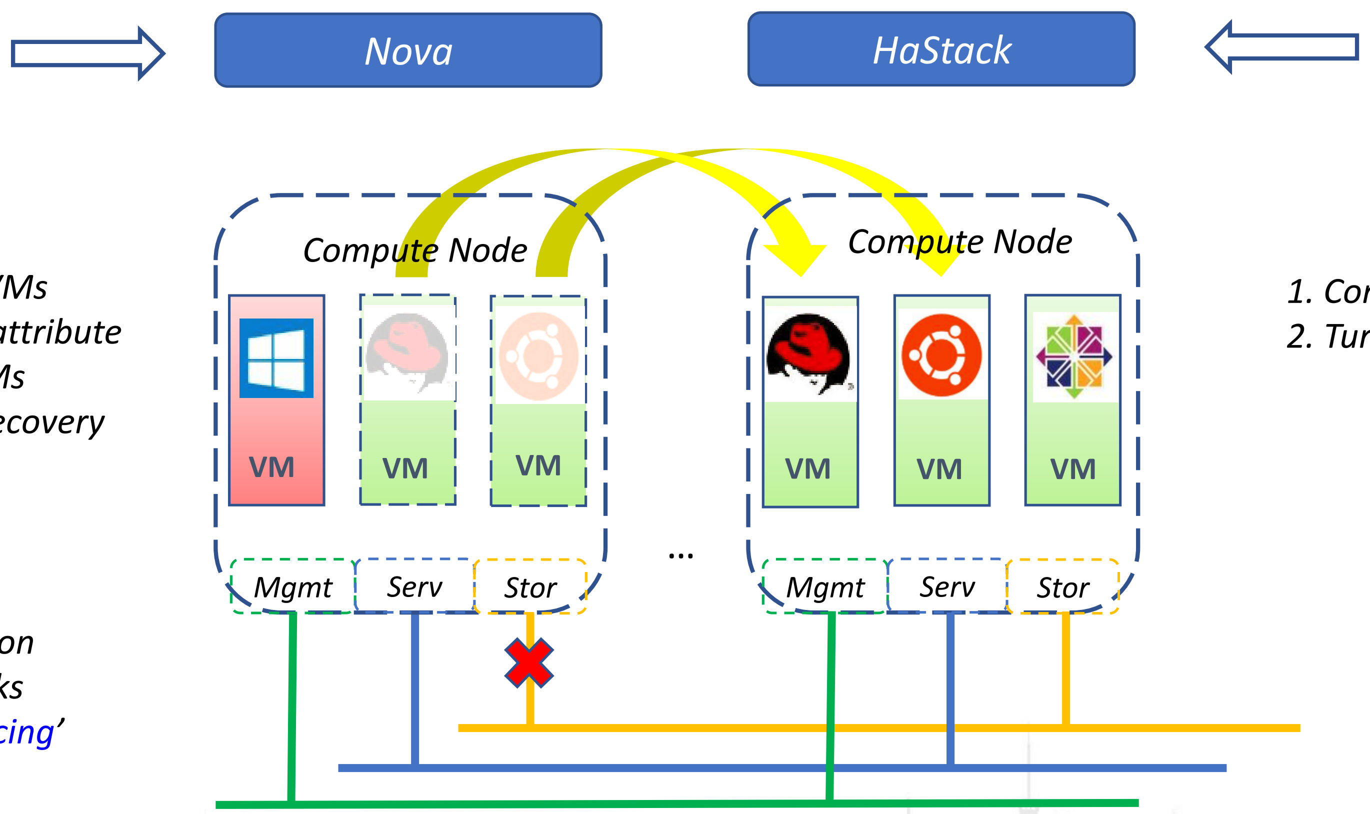
Use Cases

User

1. *Create HA VMs*
2. *Modify HA attribute*
3. *Fault HA VMs automatic recovery*

HaStack

1. *Fault Detection*
2. *Track HA tasks*
3. *Execute ‘Fencing’*



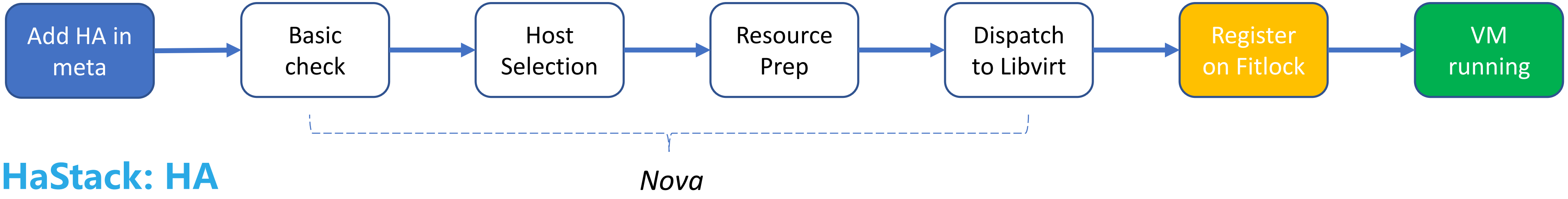
Admin

1. *Config HA strategy*
2. *Turn On/Off HA ability*

Fencing: The process of locking resources(VMs) away from a node whose status is uncertain --> Stop related VMs

HA Workflows

1. Nova: Create a HA VM



Add HA in meta

Basic check

Host Selection

Resource Prep

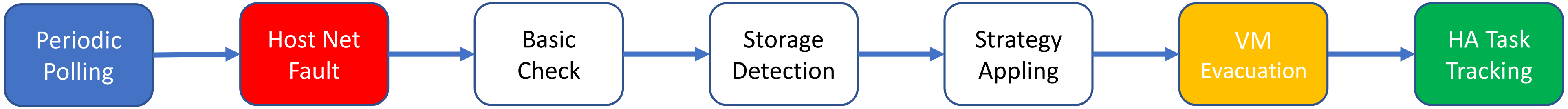
Dispatch to Libvirt

Register on Fitlock

VM running

Nova

2. HaStack: HA



Periodic Polling

Host Net Fault

Basic Check

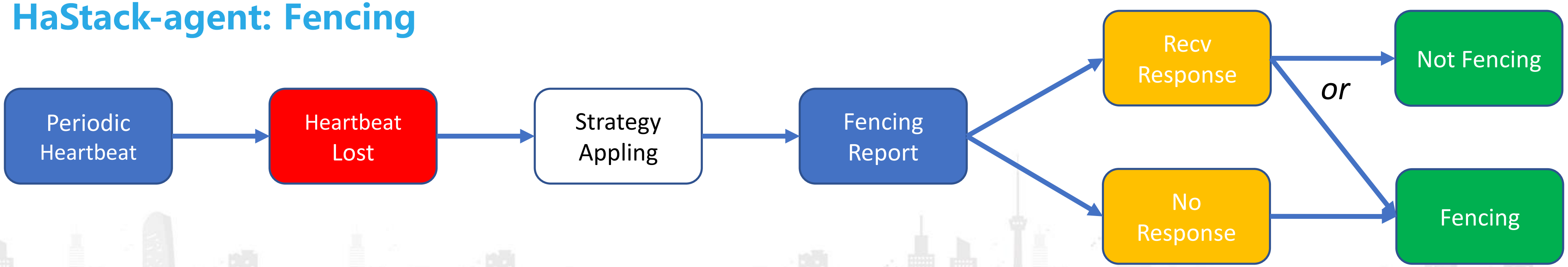
Storage Detection

Strategy Appling

VM Evacuation

HA Task Tracking

3. HaStack-agent: Fencing



Periodic Heartbeat

Heartbeat Lost

Strategy Appling

Fencing Report

Recv Response

No Response

or

Not Fencing

Fencing

HA Detection

When will it trigger HA?

1. An interruption occurred on the host network plane, *and*
2. This interruption conforms to the HA strategy

When will not trigger HA?

- VM status is **not in**: ACTIVE, STOPPED, ERROR
- VM **internal** exceptions (*blue screen/crush..)*
- The VM **virtual network** is abnormal
- The **core components**(*Etcd, Ceph..*) of the platform are abnormal

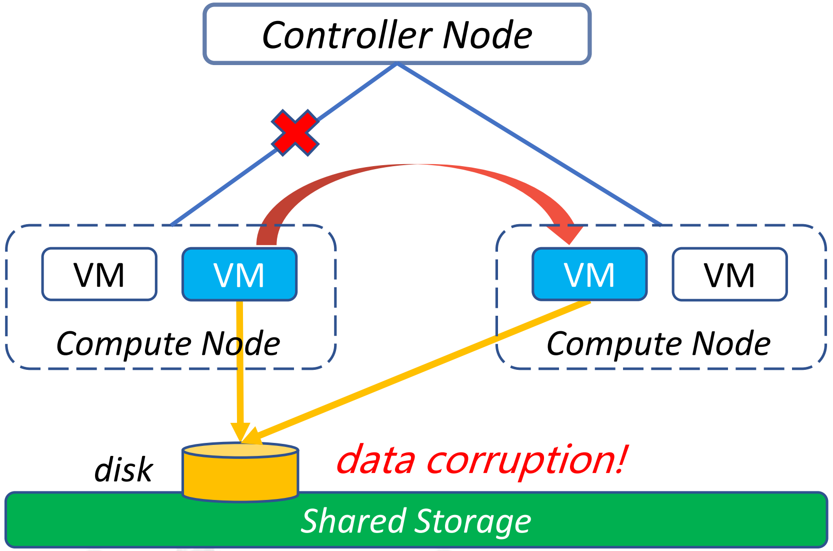
Key Features

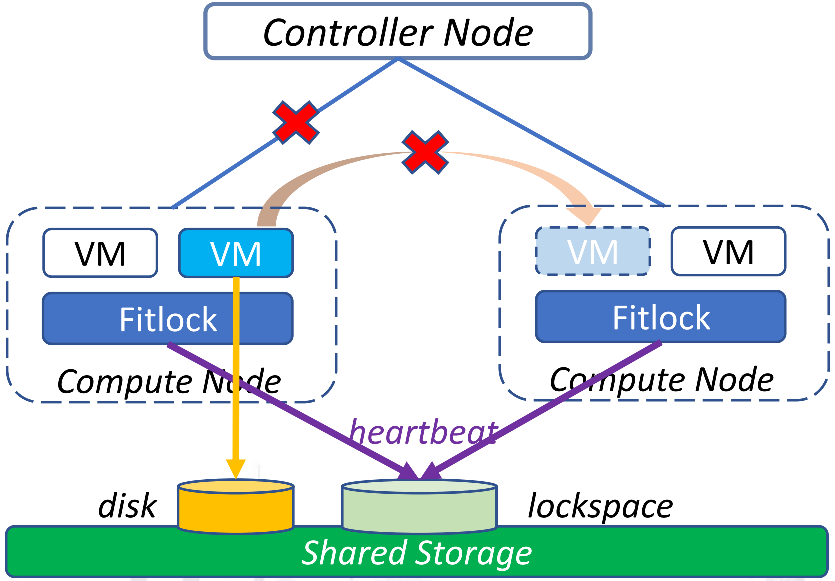
Split-brain Solving (1)

- What' s the "split-brain"?

- What' s the influence on the system?

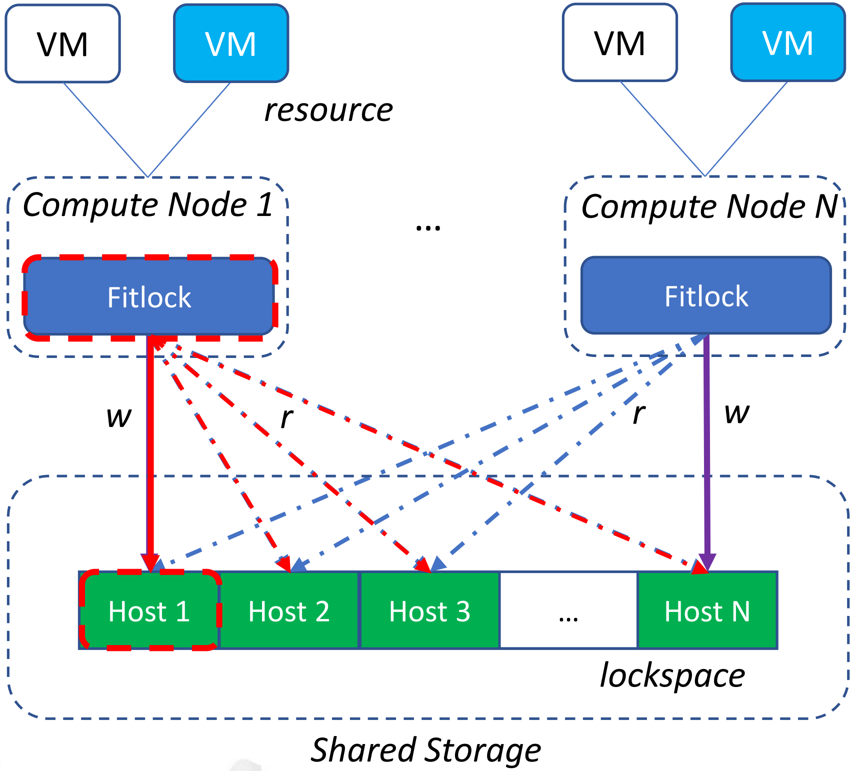
- Our Proposal
  - **Fitlock**: A "split-brain" protection read-write lock manager like Sanlock
  - **Fencing Protection**: Avoid unnecessary VM Fencing





Ensure the write-rights are unique!

Split-brain Solving (2)

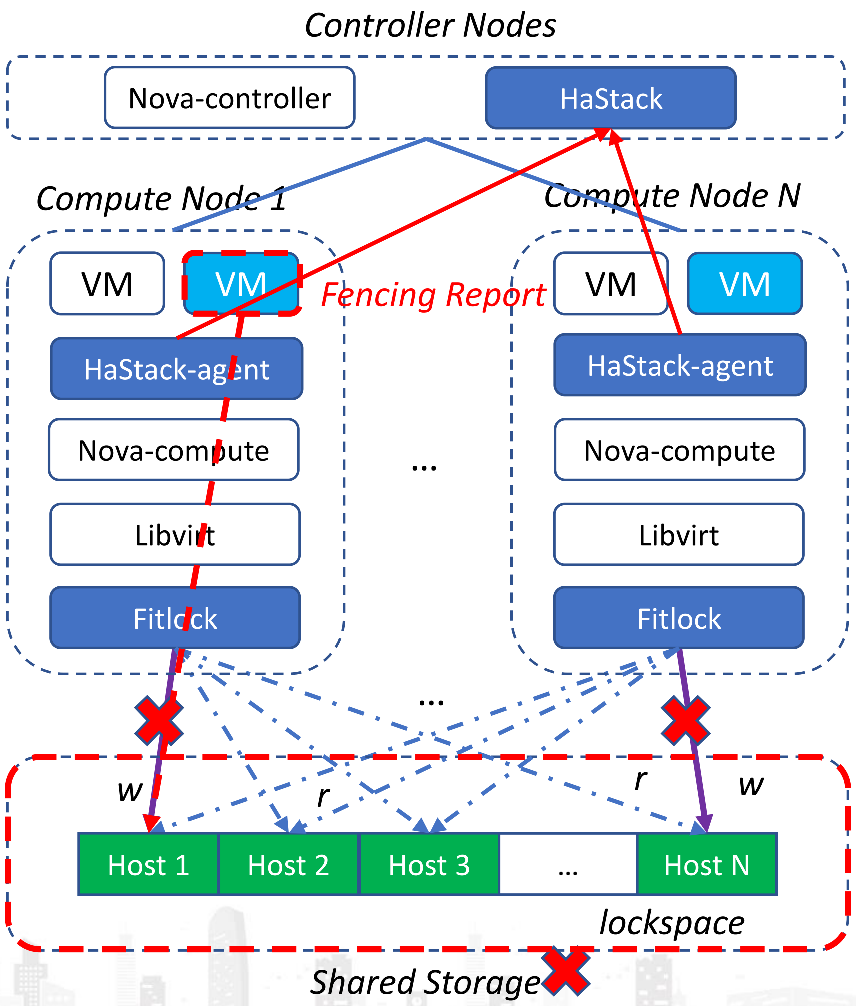


- **Fitlock**
  - **A lock manager** built on shared storage using *Delta Lease* & *Paxos Lease* like Sanlock
  - The **host lease renewal** = All that **host' s VM leases renewal**
  - ***Key point:*** If host lease is being renewed, then the VM lease is owned cannot be acquired, until it has expired:
    - A VM that is already running on one node cannot start simultaneously on another node
    - There won' t be two identical VMs in the system!

- **What' s the differences between Fitlock and Sanlock?**

Items	Sanlock	Fitlock
Lock granularity	VM disk	VM
When heartbeat lost	wait & kill VMs	ask & wait *(via socket)*
When process restart	lockspace will lost!	add Fencing Protection

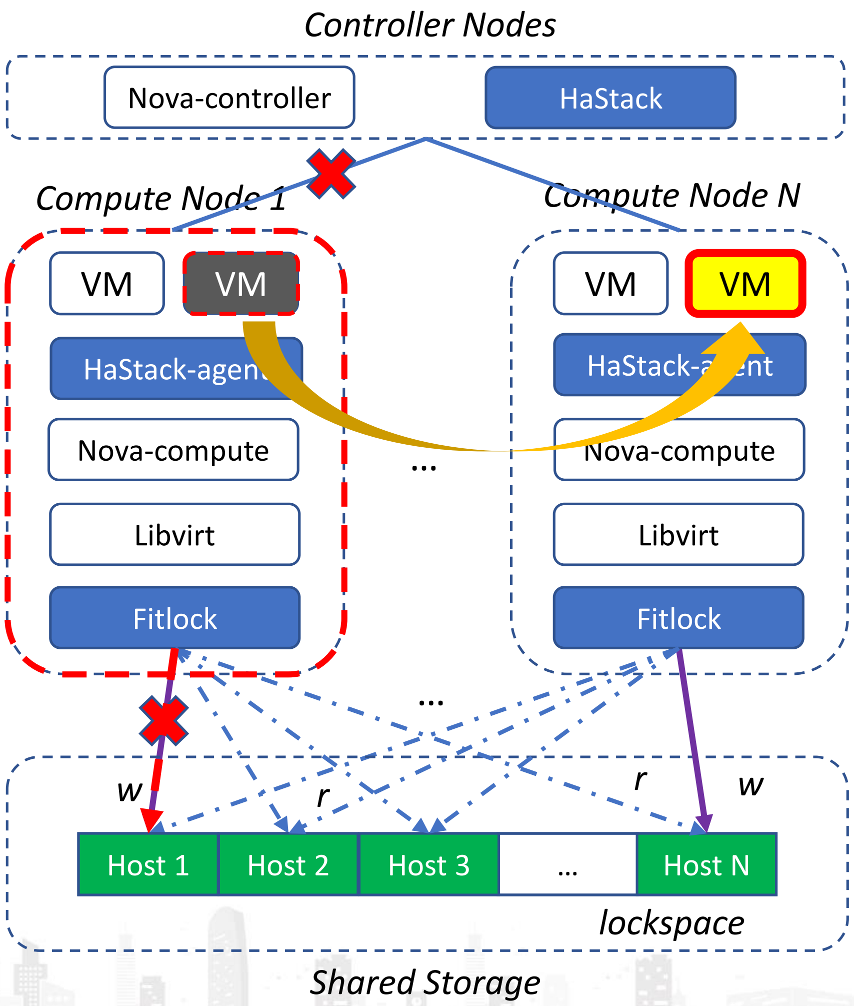
Split-brain Solving (3)



- Case 1: When a HA VM is spawned:
  - The VM lease will be registered in the lockspace

- Case 2: When shared storage is inaccessible:
  - HaStack-agent report Fencing event to HaStack and wait for a response:
    - If it get a response in time, follow the **instructions**
      - *Fencing*, or
      - *Not Fencing*
    - **Otherwise**, *Fencing*
  - In this situation:
    - HaStack will find the storage is abnormal
      - HaStack-agent will get *Not Fencing*
      - All HA VMs will remain

Split-brain Solving (4)



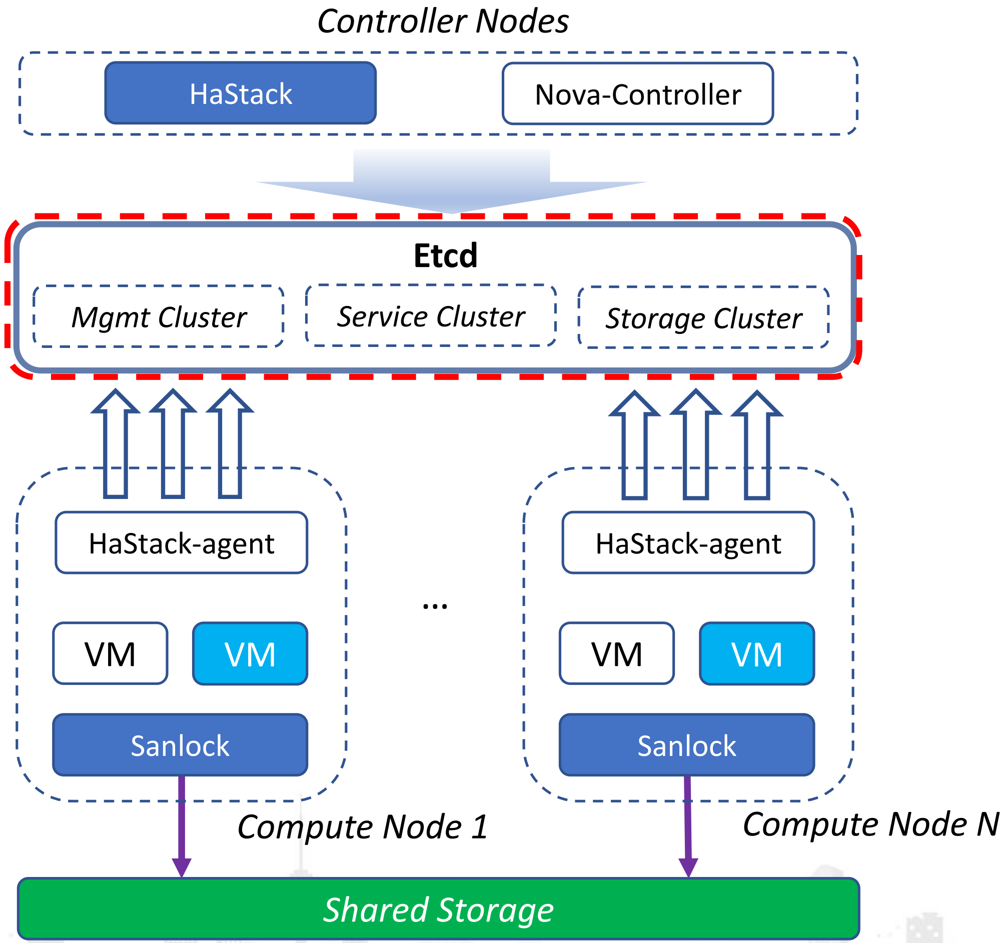
- **Case 3: When a compute node loses connections with controller nodes:**
  - The original host can **still** update heartbeat on shared storage, it **still has the lease**
  - The VM will still be running, and cannot be started on other hosts
  - The "split-brain" will not occur!

- **Case 4: When a compute node is isolated from all nodes:**
  - The compute node is disconnected from all nodes
  - The HA VMs will **be stopped** at original host due to *"split-brain" protection*
  - All HA VMs will be evacuated to other hosts at this time
    - If don' t Fencing:
      - Once the host communication **resumes**, all HA VMs will **continue to operate disks** in a short term
      - The "split-brain" may occur!

Host Network Fault Awareness (1)

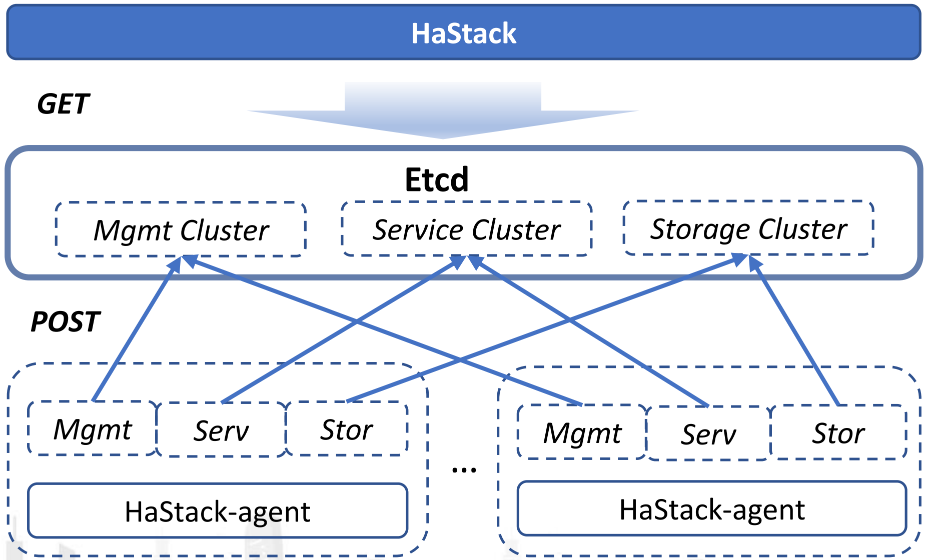
- **Build 3 Clusters on Etcd**
  - Corresponding to 3 physical network plains of host

- **Heartbeat Update**
  - HaStack-agent:
    - Every 20s
  - HaStack:
    - Obtain connectivity status every 20s
    - Execute HA after 2min when heartbeat lost



Host Network Fault Awareness (2)

- **Communication Method**
  - ➢ via Etcd API
    - ➢ ***Heartbeat***: via 1x network plane
    - ➢ ***Key messages***: via 3x network planes
      - ➢ Like *Fencing Event*..
      - ➢ HaStack removes redundancy



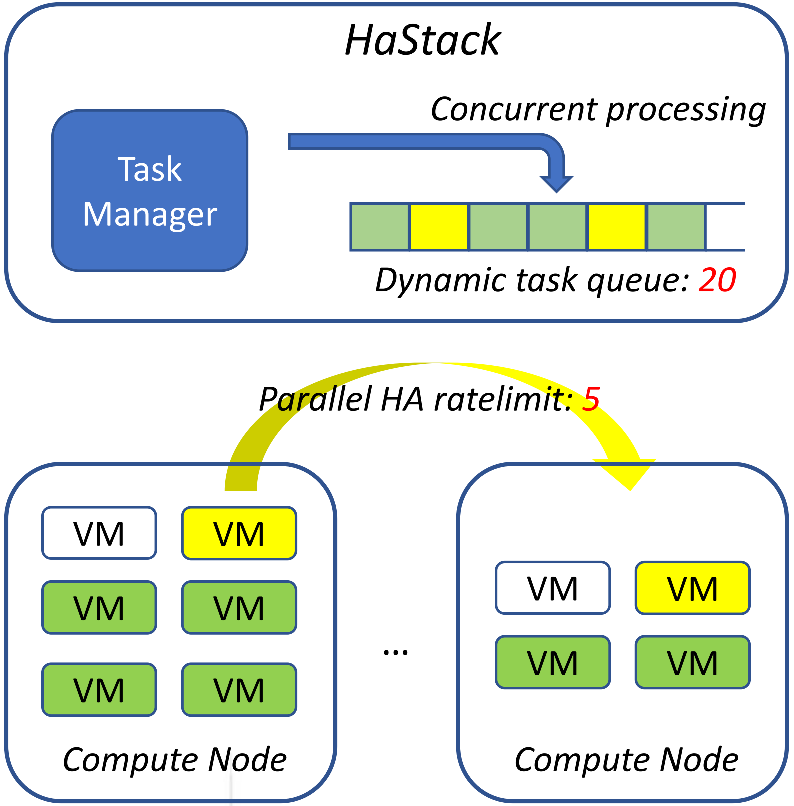
- **HA Strategy**
  - ➢ Flexible customization of HA recovery strategy
  - ➢ Configured by a *json* template

No.	*Mgmt*	*Service*	*Storage*	*HA?*
0	×	×	×	√
1	×	×	√	√
2	×	√	×	√
3	×	√	√	×
4	√	×	×	√
5	√	×	√	×
6	√	√	×	√
7	√	√	√	×

More Concerns

Task Tracking

- **Task Tracking**
  - All HA actions will be tracked
  - Failed task will be retried 5 times
- **HA Flow Control**
  - Global HA Ratelimit
    - A dynamic length-variable queue
      - Support runtime modification!
  - Host HA Ratelimit

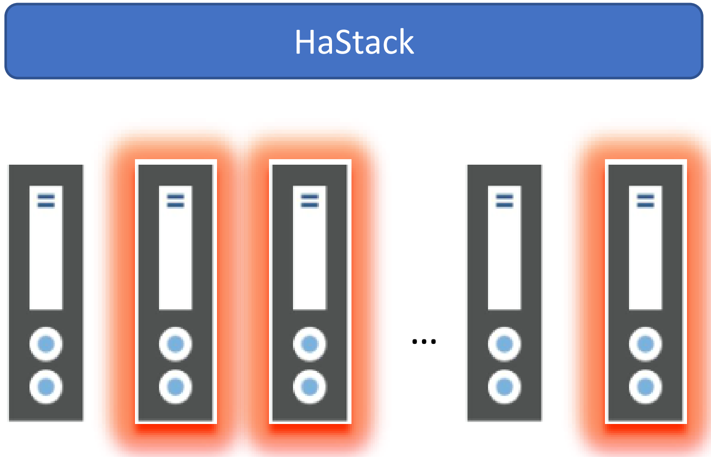


All parameters can be configured

Protection Mechanisms

- ## Process Protection
  - watchdog
- ## Self-Defense & Self-Recovery
  - Two protection mechanisms when large-scale failures occur
  - *Configurable*
- ## HA Maintenance
  - To turn off/on the HA function

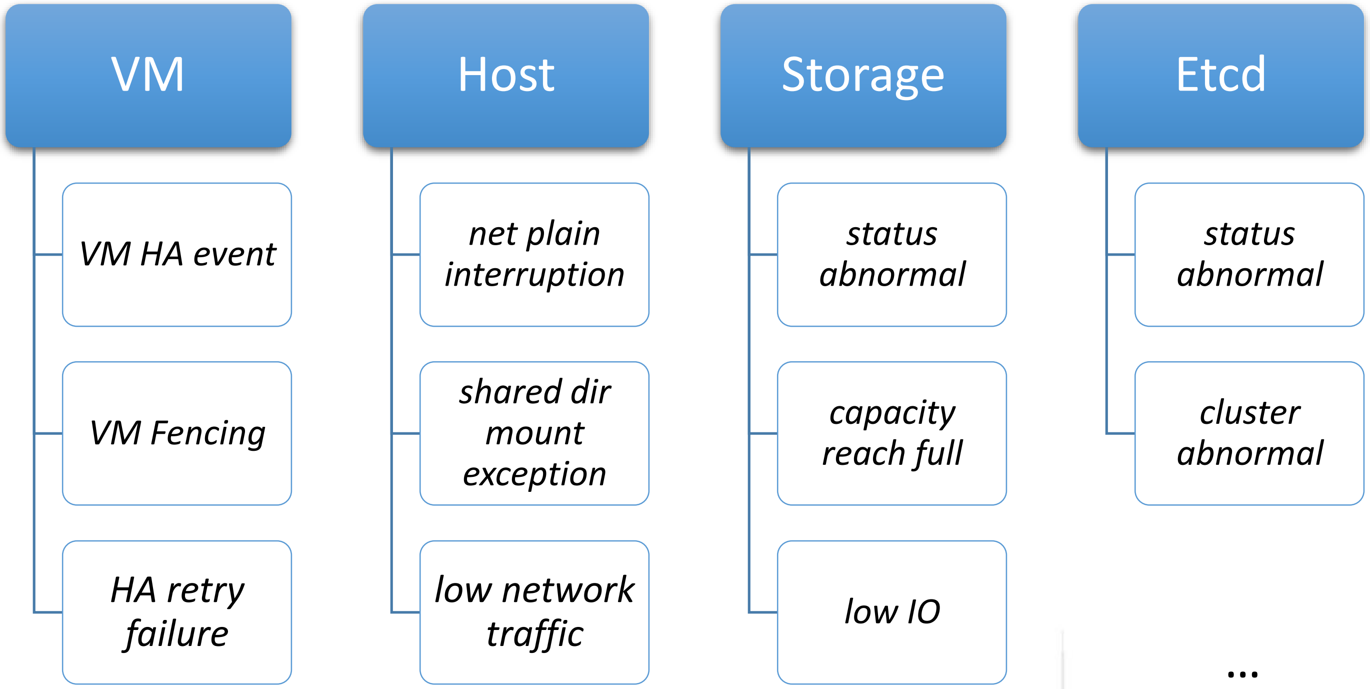
ha_maintenance	***HaStack:*** ***when host is powered off***	***HaStack-agent:*** ***when heartbeat lost***
ON	Not HA	Not Fencing
OFF *(default)*	HA	Report Fencing Event



- *Self-Defense: 50% hosts down, HaStack stop*
- *Self-Recovery: 70% hosts restore, HaStack recover*

Related Alarms

- Major Events & Alarms



- VM
  - *VM HA event*
  - *VM Fencing*
  - *HA retry failure*
- Host
  - *net plain interruption*
  - *shared dir mount exception*
  - *low network traffic*
- Storage
  - *status abnormal*
  - *capacity reach full*
  - *low IO*
- Etcd
  - *status abnormal*
  - *cluster abnormal*

…

Tests & Others

Tests

- Environments
  - 38x Compute Nodes
  - Tools
    - Rally, Heat, some scripts..
    - fio
    - Zabbix、Grafana

Count	CPU model	CPU	Mem	Disk	Ethernet
26x	E5-2658A v3 @ 2.20GHz	48	128	600G	2x 10GE 6x 1GE
12x	E5-2620 v3 @ 2.40GHz	24	64	500G	2x 10GE 4x 1GE

- Scenarios

- Results

Host Number: Down/Total	VM Number: HA/Total	Global HA Ratelimit	Storage pressurized	Single VM average Recovery Time	Total Recovery Time
20/38 (52.6%)	1000/1741	20	N	~1 min	41min
20/38 (52.6%)	1000/1741	100	N	1~2 min	20min
20/38 (52.6%)	1000/1741	20	Y	~2 min	1h 3min

Production Cases

FiberHome

Project_1:

楚天云

- **Scenario**: *Hybrid Cloud*
- **Scale**:
  - ***Region***: 5x
  - ***Servers***: 800x
  - ***VMs***: 2700x

Project_2:

中国航信 TravelSky

- **Scenario**: *Hybrid Cloud*
- **Scale**:
  - ***Region***: 1x
  - ***Servers***: 122x
  - ***VMs***: 1000x

Project_3:

云投集团 YUNNAN INVESTMENT GROUP

- **Scenario**: *Private Cloud*
- **Scale**:
  - ***Region***: 1x
  - ***Servers***: 131x
  - ***VMs***: 1000x

Future Works

- QGA Integration
- Visualized HA strategy template Selection
- Reduce HA recovery time

References

- About Split-brain:
  - [1] *https://en.wikipedia.org/wiki/Split-brain_(computing)*
  - [2] *http://linux-ha.org/wiki/Split_Brain*

- About Sanlock
  - [3] *https://www.ovirt.org/develop/developer-guide/vdsm/sanlock/*

- About CephFS
  - [4] *https://www.linux.com/news/converging-storage-cephfs-now-production-ready*
  - [5] *Benchmark from eBay*： *https://www.slideshare.net/XiaoxiChen3/cephfs-jewel-mds-performance-benchmark*

- About *<Distributed Health Checking for Compute Node High Availability>*:
  - [6] *https://www.openstack.org/videos/tokio-2015/distributed-health-checking-for-compute-node-high-availability*

Contact Us

Jiang WU
Architect, Fiberhome

[*wingwj@gmail.com*](mailto:wingwj@gmail.com)

- *Fiberhome*:
  - A globally information and communication network product and solution provider
    - *One of the world’ s 10 most competitive enterprises in optical communications*
    - *Ranked 1st in export among optical cable enterprises of China for 8 consecutive years*
  - Gold Member of the OSF since 2017

THANKS

FiberHome