OpenStack High Availability Technologies:

A framework to test High Availability architectures

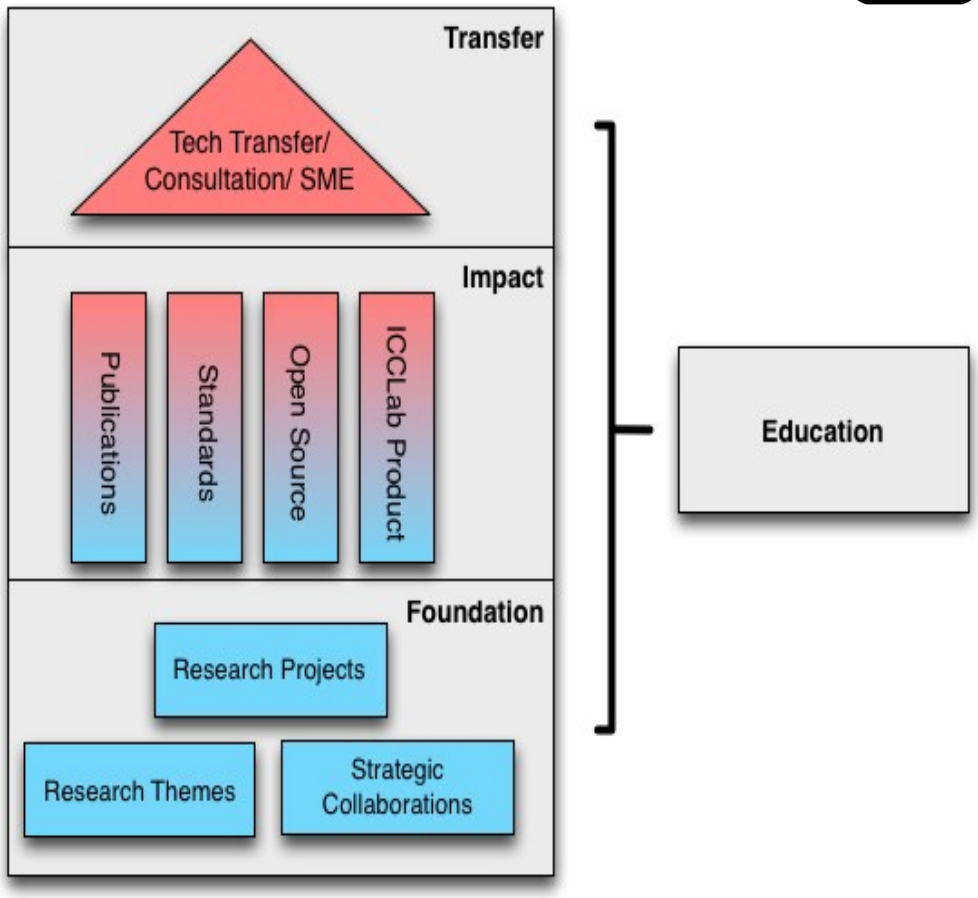
Konstantin Benz, Thomas M. Bohnert

Conference on Future Internet Communications
University of Coimbra, May 2013

ICCLab
www.cloudcomp.ch, @ICC_Lab, #icclab

The ICCLab

Research Themes
- **IaaS**
- **PaaS**
- **MobileCloud**
- **Quality of Cloud Services:**
  - **Data Privacy**
  - **Security**
  - **Interoperability**
  - **High Availability**
  - **...**
- **SDN**



www.cloudcomp.ch

Cloud Computing

No worries ...

… it's in the cloud

No worries …

Reddit And Netflix Down? Amazon Network Issue Causes Downtime For Multiple Sites (UPDATE)

Posted: 10/22/2012 2:55 pm EDT Updated: 10/22/2012 5:56 pm EDT

An apparent issue with Amazon Elastic Cloud Compute network caused partial outages of several popular sites on Monday afternoon.

Social news site Reddit was inaccessible for some users. The site confirmed the downtime in a tweet sent around 1:30 p.m. ET on Monday: "The site is down right now. It appears to be a network-related issue. We are investigating."

(At the time of writing, the site appeared to be at least partly up, but pages were slow to load when they loaded at all.)

Diagnostics site DownRightNow also reported around the same time that Netflix was experiencing a "likely service disruption."

Image-sharing site Imgur was also down for some, as was vacation-booking site Aribnb, Mashable reports.

Reddit-watcher blog the Daily Dot was affected and tweeted that the problem was on the Amazon network's end. "Amazon has broken the Internet and taken the Daily Dot with it. We'll be back just as soon as possible," read a tweet from the site's official feed.

According to TechCrunch, Amazon's AWS servers (in addition to EC2 servers) in its north Virginia data center were the source of the problem: "[Amazon's] Elastic Block Store part seems to be down. All services hosted on Heroku's virtual private servers are still down at the time of writing." Twitter users reported that GitHub, Pinterest, Gamespot and other services were also down as a result.

Have you had a problem with any of these sites today? Let us know in the comments or tweet us @HuffPostTech.

UPDATE: 5:55 p.m. Amazon is working to restore normal service to the servers affected, per the company's Web Services Service Health Dashboard.

High Availability

Why is High Availability so important?

- Consumers have extremely high expectations:
  - 24 hours per day / 7 days per week availability
  - IT services should always be up-to-date / highly secure etc. (intensive maintenance)
- Unplanned IT downtime can cost companies up to 5'000 $ per minute (according to Uptime Institute Report 2011)
- Companies can cease to exist due to outage of IT services

High Availability

What is «High Availability»?

- Availability:
  Ability of end users to access a system and perform required tasks
- Availability Measurement:
  - Availability = (Uptime / Total Operating Time) x 100

    Alternative calculation:
    ((Total Operating Time – Downtime) / Total Operating Time) x 100
  - Downtime: 1 day per year
    Operating Time: 365 days

    Availability = (364 / 365) x 100 = 99.73 %
- **«High Availability» > 99.99 %**

High Availability

High Availability - Classifications

- Several Nines:
  - According to Downtime / Operating Time ratio

Yearly Availability	Downtime per Year	Availability Class
90.00 %	36.50 d	
95.00 %	18.25 d	
98.00 %	7.30 d	
99.00 %	**3.65 d**	**2 – stable**
99.50 %	1.83 d	
99.80 %	17.52 h	
99.90 %	**8.76 h**	**3 – available**
99.95 %	4.38 h	
99.99 %	**52.60 m**	**4 – high availability**
99.999 %	**5.26 m**	**5 – fault resilient**
99.9999 %	**31.50 s**	**6 – fault tolerant**
99.99999 %	**3.00 s**	**7 – fault resistant**

High Availability

High Availability - Classifications

- Availability Environment Classification AEC (Harvard Research Group):
  - Classification based on allowed **impact** of interruptions

Class	Title	Business Impact
AEC - 0	Conventional	IT service is allowed to be interrupted. Data integrity is not essential.
AEC - 1	Highly Reliable	IT service might be interrupted as long as data integrity is preserved.
AEC - 2	High Availability	Only planned or short interruptions are allowed. Data must not get lost, but transaction losses are acceptable.
AEC - 3	Fault Resilient	IT service must be interruption free. No data or transaction loss allowed. Performance reduction is acceptable.
AEC - 4	Fault Tolerant	IT service must be interruption free. No data or transaction loss allowed. No performance reduction allowed.
AEC - 5	Disaster Tolerant	IT service must be free of interruptions, data or transaction loss or performance reductions even in case of disasters and destruction of physical assets (like e. g. fire, earthquake, vandalism etc.).

High Availability

High Availability - Strategy

- What factors decrease availability?
  - Planned unavailability:
    - System maintenance
  - Unplanned unavailability:
    - Complex system interactions
    - Bad configuration
    - Many user interactions (load, traffic etc.)
    - …
- **Complexity** is often the main reason, why an IT service becomes unavailable

High Availability

High Availability - Strategy

- What factors increase availability?
  - Recovery from outage:
    - Rollback scripts
    - Data backups
    - ...
  - Avoid outages:
    - Redundant systems
    - Balanced control flow between systems
    - Recovery is transparent / invisible to end user
    - …
- **Redundancy** generally increases availability, but:
  - **Redundancy also increases complexity**

High Availability

DRBD

- **D**istributed **R**eplicated **B**lock **D**evice
- Works on top of block devices (hard disk partitions, logical volumes etc.)
- Mirroring of a whole block device via an assigned network to a distant node
- After an outage DRBD resynchronizes unavailable node to latest available version of data
- Often referred to as “network based RAID-1”

- Advantages:
  - Technologically simple solution
  - Great to cluster data objects with fixed size: VM instances, VM images, Volumes...
  - Especially useful for OpenStack Glance (volume management) service

- Drawbacks:
  - DRBD uses fixed size blocks to store data: not suitable to store variably sized data objects

High Availability

Ceph / RADOS

- **R**eliable **A**utonomic **D**istributed **O**bject **S**tore
- Ceph relies on clusterable object storage component: RADOS
  - Technology-specific block device: Ceph RBD
  - Technology-specific filesystem: Ceph FS filesystem
  - Technology-specific network gateway: Ceph RADOS GW (RESTful Gateway)
- LIBRADOS library allows applications to access RADOS
- Variably sized objects

- Advantages:
  - Ceph can cluster almost anything: VM images and instances, VM volumes, application data...
  - Useful for all OpenStack services, but especially for OpenStack Swift (object storage) service: Ceph uses Swift API
- Drawbacks:
  - Rather complex solution: configuration is very difficult

HA technologies

MySQL Galera cluster

- Synchronous multi-master cluster for MySQL/InnoDB database
- Database replication is not simply replication of data objects:
  - Lots of (concurrent) transactions
  - Outage leads to inconsistent data (lost transactions during outage etc.)
- Proprietary group communication system layer
- **W**rite**S**et **Rep**lication (wsrep API):
  - Transaction writesets are replicated over several nodes before they are commited
  - Global Transaction IDs to uniquely identify transactions
  - (Virtually) synchronous replication
- Advantages:
  - No lost transactions when they are commited before outage
  - Useful to make OpenStack MySQL DB highly available
- Drawbacks:
  - Additional memory is consumed for uncommitted writesets: memory management necessary

HA technologies

Pacemaker

- Open Source HA resource manager for clusters
- Automatic detection and recovery of machine and resource-level failures
- Presence of resource in cluster propagated by cluster membership daemons (e. g. Heartbeat, Corosync)
- Compatible to many different clustering technologies:
  - Cluster Abstraction Layer to support different cluster membership management technologies
  - Cluster membership and resource information stored in Cluster Information Base (CIB)
  - Cluster Resource Management Daemon to manage resources according to CIB
  - “Fencing” between concurrent primary nodes managed by fencing management subsystem

- Advantages:
  - Replaces manual failure detection and recovery procedures
  - Compatible to many HA technologies
- Drawbacks:
  - Efficiency of detection and recovery mechanisms heavily depends on correct cluster information base configuration

HA technologies

HA technologies for OpenStack

- Technologies to **increase redundancy**:
  - Clustering of MySQL: MySQL Galera
  - Clustering of Hypervisor Service (Nova): Pacemaker resource agent
  - Clustering of Dashboard Service (Horizon): Pacemaker resource agent
  - Clustering of Block Storage Service (Cinder): DRBD
  - Clustering of Object Storage Service (Swift): Ceph Object Storage
  - Clustering of Image Service (Glance): Use Swift and Ceph Object Storage
  - Clustering of Network Service (Quantum): Pacemaker resource agent
  - Clustering of Identity Management Service (Keystone): Pacemaker resource agent
- Technologies to **balance/control replication**:
  - Pacemaker including Corosync for OpenStack

HA test framework

Testing the technologies

- Which technology fits best?
  - Many different architectural solutions possible:
    - Redundant OpenStack “all-in-one” installations
    - Redundant Compute-, Controller- and Network nodes
    - Redundant nodes for DRBD, Ceph, storage...
  - HA solutions **increase complexity**:
    - Unintended reboot scripts
    - Concurrency between nodes after failure: STONITH
    - ...
  - Method to “test” quality of an architecture:
    - In practice: trial and error
    - Is there a better solution?

HA test framework

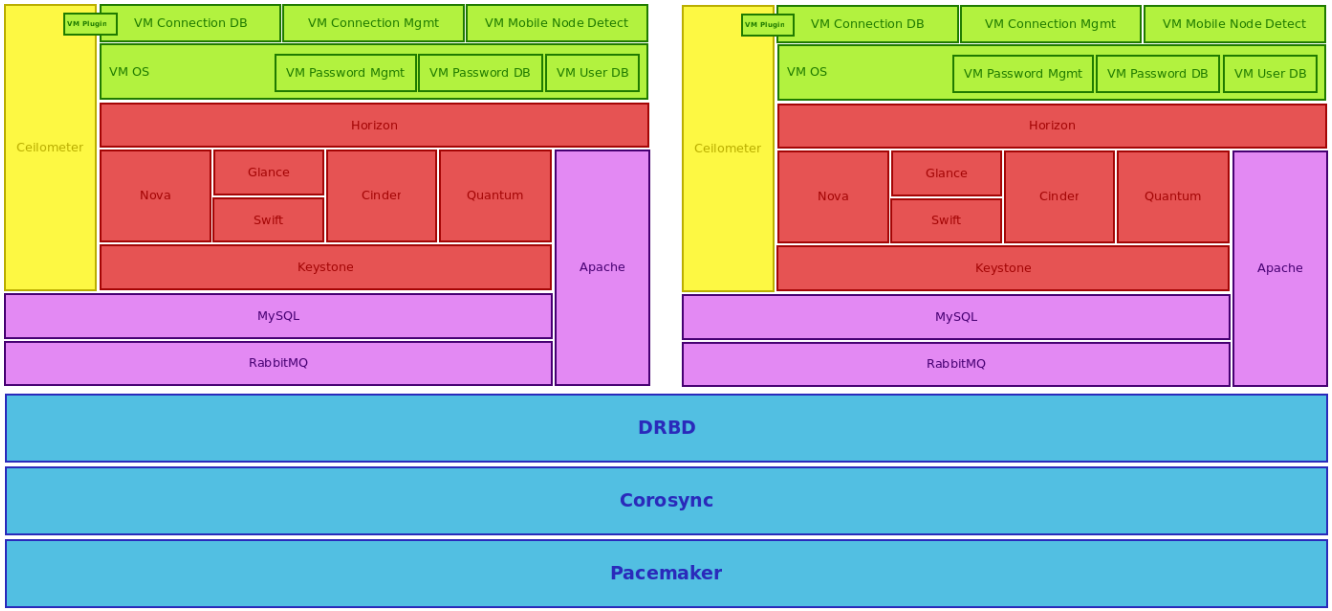
Test framework

- Systematic “trial and error”:
  1. Implement / configure an OpenStack architecture (including HA technologies)
  2. Simulate outage of components:
     - Random shutdown of services
     - Power off some nodes
     - Unplug network devices
     - **Use realistic probabilities**
  3. Check impact of outages from the end user perspective:
     - Which services are still usable?
     - Which tasks can be performed?
     - How important are the service outages to the end user?
     - **Rate the impact**
  4. **Store results** and undo changes after outage
  5. Repeat previous steps on multiple different architectural setups

HA test framework

Implement a HA architecture

- Basic principles:
  - Test non-redundant systems too (“Null-hypothesis”)
  - **Automate** installation and configuration to make your implementation **reusable for multiple test-runs**
  - Chosen architecture defines structure of database where test results are stored



HA test framework

Simulate outage of components

- Useful Tool:
  - "Chaos Monkey" (Netflix):
    - Experience with cloud outages
    - Tool which **randomly disables** services to **test impact of outages**
- Basic principles:
  - Run Chaos Monkey "attacks" on OpenStack nodes
  - Assign **probabilities** to Chaos Monkey attacks (simulate randomness of attacks)
  - **Daily outage risk:**

	Time frame	31,536,000				
	Component	**Availability**	**Downtime**	**Avg. Recovery Time**	**Total Outages**	**Outage risk Per day**
1	Hardware of OpenStack Installation	99.90%	**31,536**	14,400	2.19	0.60%
2	OS of OpenStack Installation	99.90%	**31,536**	14,400	2.19	0.60%
3	Apache	90.00%	**3,153,600**	28,800	109.5	30.00%
4	Ceilometer	90.00%	**3,153,600**	28,800	109.5	30.00%
5	Cinder	90.00%	**3,153,600**	28,800	109.5	30.00%
6	VM internal Connection DB	90.00%	**3,153,600**	28,800	109.5	30.00%
7	VM internal Connection Management	90.00%	**3,153,600**	28,800	109.5	30.00%
8	Glance	90.00%	**3,153,600**	28,800	109.5	30.00%
9	Horizon	90.00%	**3,153,600**	28,800	109.5	30.00%
10	Keystone	90.00%	**3,153,600**	28,800	109.5	30.00%
11	VM internal Node Location Detection	90.00%	**3,153,600**	28,800	109.5	30.00%
12	MySQL	90.00%	**3,153,600**	28,800	109.5	30.00%
13	Nova	90.00%	**3,153,600**	28,800	109.5	30.00%
14	VM internal Operating System	90.00%	**3,153,600**	28,800	109.5	30.00%
15	VM internal Password DB	90.00%	**3,153,600**	28,800	109.5	30.00%
16	Quantum	90.00%	**3,153,600**	28,800	109.5	30.00%
17	RabbitMQ	90.00%	**3,153,600**	28,800	109.5	30.00%
18	VM internal Password Management	90.00%	**3,153,600**	28,800	109.5	30.00%
19	Swift	90.00%	**3,153,600**	28,800	109.5	30.00%
20	VM internal User DB	90.00%	**3,153,600**	28,800	109.5	30.00%
21	VM internal Ceilometer Plugin	90.00%	**3,153,600**	28,800	109.5	30.00%

HA test framework

Measure impact of outages

- Useful Tool:
  - "Edda" (Netflix):
    - Tool to **poll VMs** in cloud services to **check availability**
- Basic principles:
  - Check which services are available after Chaos Monkey "attacks"
  - Assign **weights** to outages
  - Calculate **impact of outages** from **weights of use cases**

Component	Use Case	Weight	Impact
Apache	Login/Logout to Dashboard	1	
	Manage Keypairs	3	4
Ceilometer	Measure SLAs	2	
	Meter usage of Telco service	3	
	Monitor VM and Infrastructure	3	8
Cinder	Create/Delete/Update VM/Instances	2	
	Create/Delete/Update Volumes	2	4
Connection DB	Telco Connect	3	3
Connection Management of VM	Telco Connect	3	3
Glance	Create/Delete/Update VM/Instances	2	
	Create/Delete/Update Images	2	4
Horizon	Login/Logout to Dashboard	1	
	Manage Keypairs	3	4
Keystone	Create/Delete/Update OpenStack Account	3	
	Create/Delete/Update policies	3	
	Login/Logout to Dashboard	1	
	Manage Keypairs	3	
	VM Admin Authenticate	3	13

Component	Outage Impact
Apache	4
Ceilometer	8
Cinder	4
VM internal Connection DB	3
VM internal Connection Management	3
Glance	4
Horizon	4
Keystone	13
VM internal Node Location Detection	2
MySQL	30
Nova	5
VM internal Operating System	3
VM internal Password DB	9
Quantum	2
RabbitMQ	27
VM internal Password Management	9
Swift	5
VM internal User DB	3
VM internal Ceilometer Plugin	8

HA test framework

Collect test results

- Basic principles:
  - Measure **outage impact**
  - Measure **architecture specifications**:
    - Number of nodes
    - Node configuration
    - Clustering technologies: DRBD, Ceph, MySQL Galera...
    - Reboot procedure: Pacemaker configuration
  - **Cleanup and re-run test**

Run_ID	Impact	#Nodes	Config_ID	Configuration	n1.ext_IP	n1.int_IP	n1.Apache	n1.Ceilometer	n1.Cinder
000180379	4	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180380	2	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180381	0	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180382	13	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180383	17	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180384	2	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180385	4	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180386	4	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180387	0	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE
000180388	0	2	000040567	2-node All-in-one	10.1.2.44	192.168.22.11	TRUE	TRUE	TRUE

Outlook

- Test different HA technologies with the Chaos Monkey
  - Collect statistical data about OpenStack HA technologies
  - Find technologies that generate the least impact of outages
  - Look for possible correlations between replication technology, number of nodes, node configuration and impact of outages

Closing

- High Availability
  1. Is important to build trust in Cloud services
- HA technologies
  1. Increase availability by adding redundancy
  2. Decrease availability by adding complexity
  3. Must be tested for suitability
- Test framework
  1. Implement OpenStack HA architecture
  2. Simulate random outages
  3. Measure impact
  4. Collect data to evaluate advantages/drawbacks

Thanks, questions?

ICCLab Community

Swiss OpenStack User Group!
Meetup soon to be announced


@openstackch

http://linkd.in/os-ug-ch

Join in!