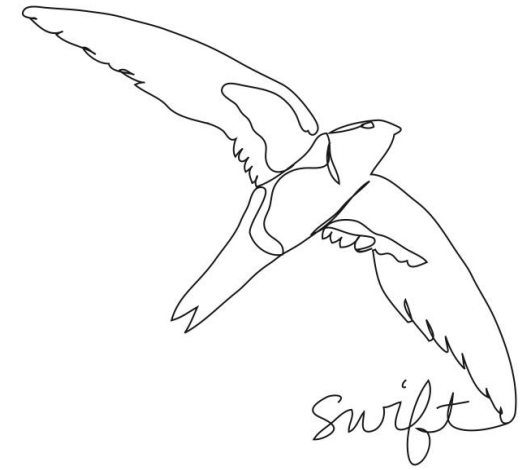# OPENSTACK SWIFT

RAMIT SURANA

WHAT IS OPENSTACK SWIFT ?

- Object storage based system

- Designed to store files, videos, analytics data, web content, backups, images, virtual machine snapshots.

- Designed from the ground up to scale horizontally without any single point of failure.

- Scale to thousands of machines providing hundreds of Petabytes of storage distributed in geographically distant regions.
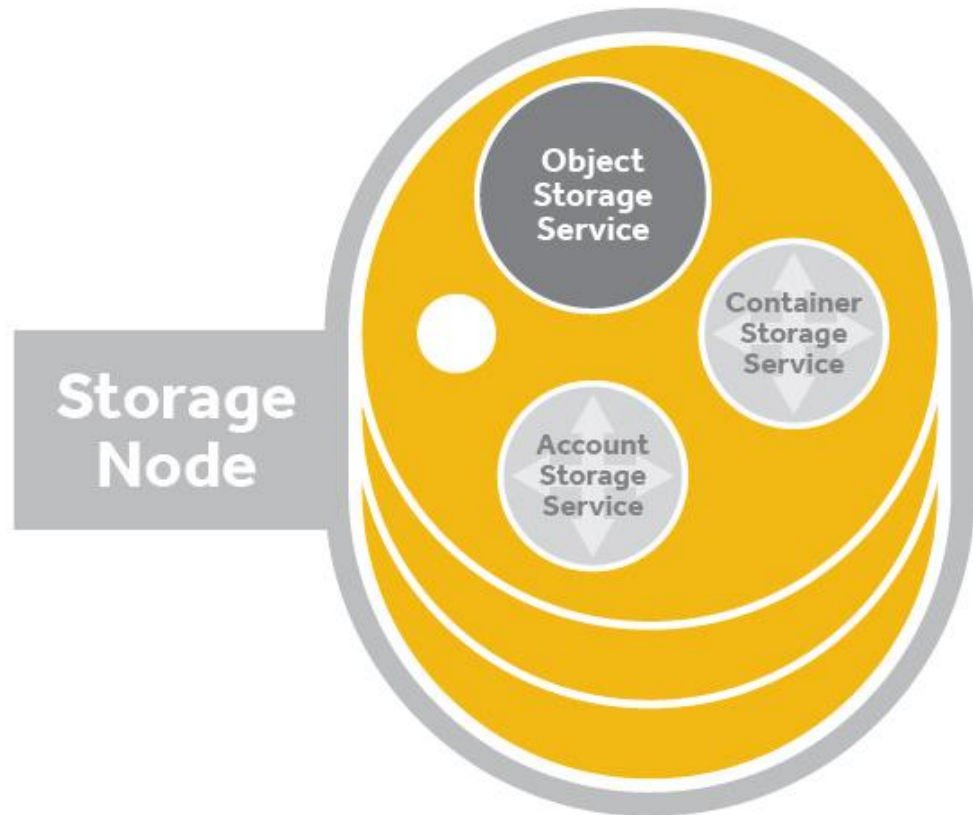
SWIFT CHARACTERSTICS

- **All objects (data) stored in Swift have a URL.**

- **Failed nodes & drives can be swapped out while the cluster is running with no downtime.**

- **All objects are stored with multiple copies and are replicated in as-unique-as-possible availability zones and/or regions .**

- **Adding or replacing hardware, data does not have to be migrated to a new storage system.**

STORAGE INTELLIGENCE

Physical Hardware

CONTROLLER

openstack™
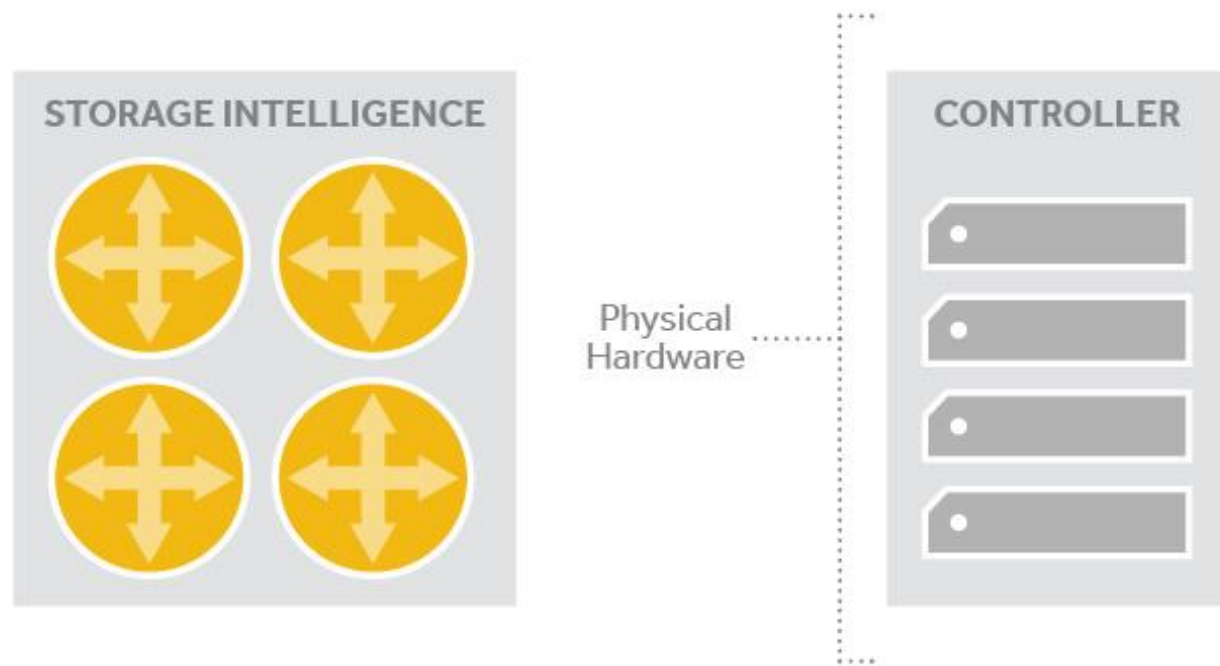
SWIFT CLUSTER ARCHITECTURE

- Nodes :Machine that is running one or Swift processes.It belongs to two logical groups.

- Regions : User-defined and a method to indicate when parts of the cluster are physically separate --usually a geographical boundary. Minimum of one region and there are many single region clusters as a result.

- Zones: Distinct set of physical hardware whose failure would be isolated from other zones.
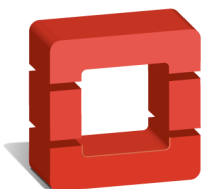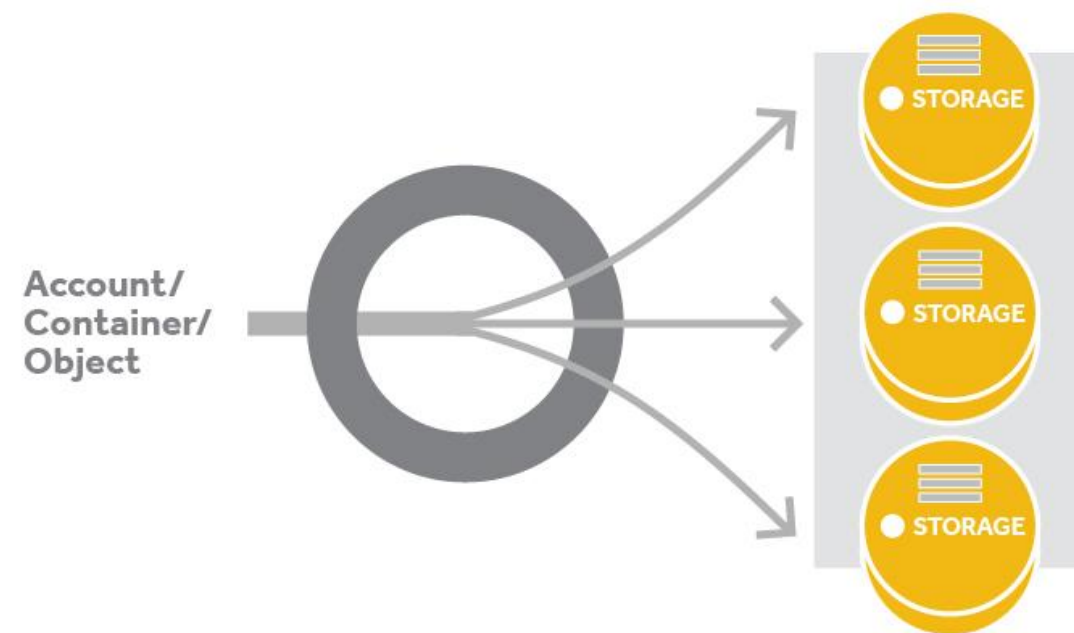
SWIFT PARTITIONS

- Partitions are represented as directories on each disk.

- Hashing is the key to the data locations. When a process, like a proxy server process, needs to find where data is stored for a request, it will call on the appropriate ring to get a value that it needs to correctly hash the storage location.

- The hash value of the storage location will map to a partition value.

- As a ring is built the partitions are assigned to drives in the cluster.

The Ring

STORAGE

STORAGE

STORAGE

STORAGE

openstack™
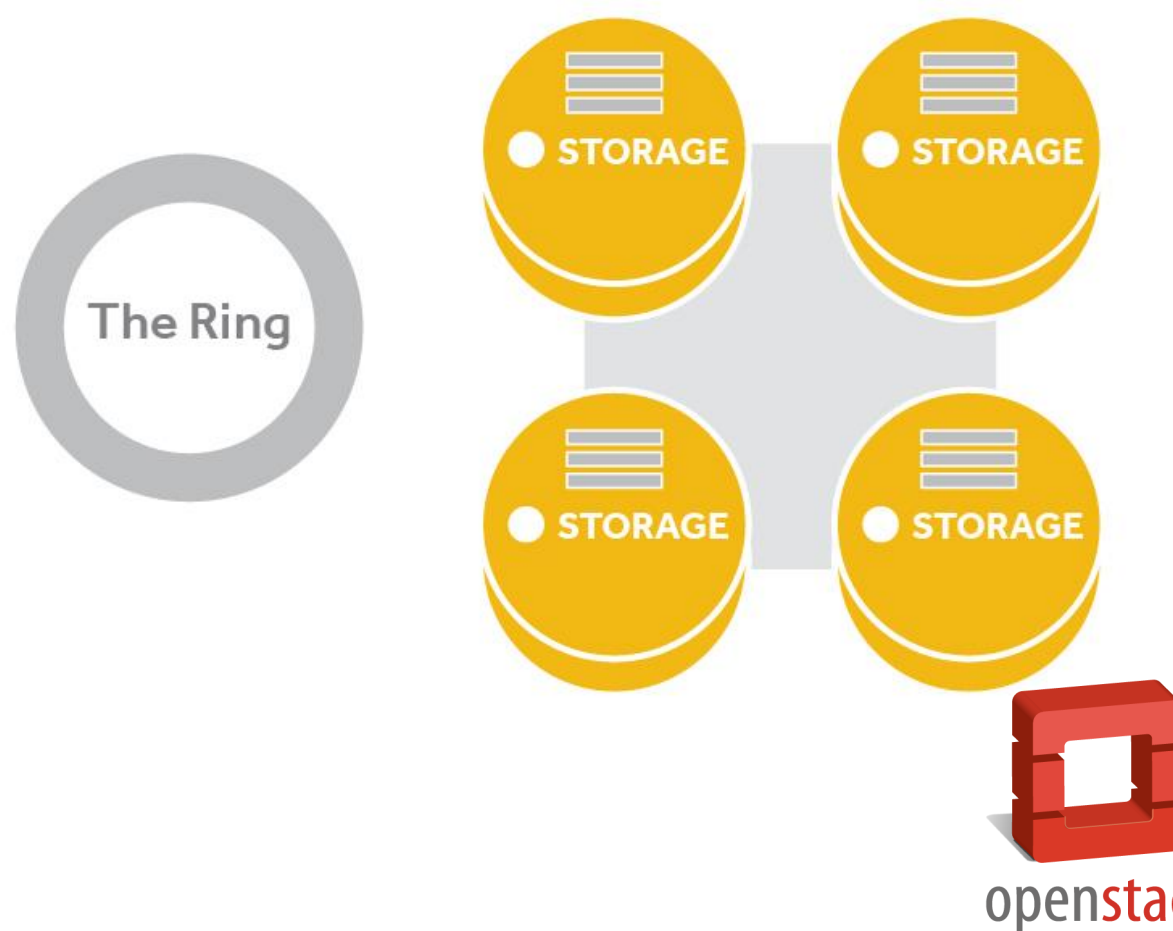
SWIFT REPLICAS

- It means Partitions are replicated.

- When a drive fails, the replication/auditing processes notice and push the missing data to handoff locations.

- Each replica is placed as uniquely as possible across the cluster. Each subsequent rebuilding of the rings will calculate which, if any, of the replicated partitions need to be moved to a different drive.

Account/
Container/
Object

STORAGE

STORAGE

STORAGE

openstack™

SWIFT RING BUILDER

- Ring data structure includes the partition shift value which processes and services use to determine the hash of a storage location.

- Each entry for a drive includes its ID number, zone, weight, IP, port, and device name.

- When a ring is being built total number of partitions is calculated with a value called the partition power.

- The formula used is 2 raised to the partition power.

Partitions

		0	1	2	3	4	...
Replicas	0	7	0	1	4	22	
	1	12	4	8	10	18	
	2	1	21	10	0	3	

|	0	1	2	3	4	...
Devs	dict of dev 0 region : 1 zone : 3 weight : 1 ...	dict of dev 0 region : 1 zone : 3 weight : 1 ...	dict of dev 0 region : 1 zone : 3 weight : 1 ...	dict of dev 0 region : 1 zone : 3 weight : 1 ...	dict of dev 0 region : 1 zone : 3 weight : 1 ...	

openstack™

SWIFT REQUESTS

- Storage location is given in one of three formats :

- /account : Uniquely named storage area that contains the metadata about the account itself as well as the list of containers in the account.

- /account/container : Container storage location is the user-defined storage area within an account where metadata about the container itself and the list of objects in the container will be stored.

- /account/container/object: Object storage location is where the data object and its metadata will be stored.
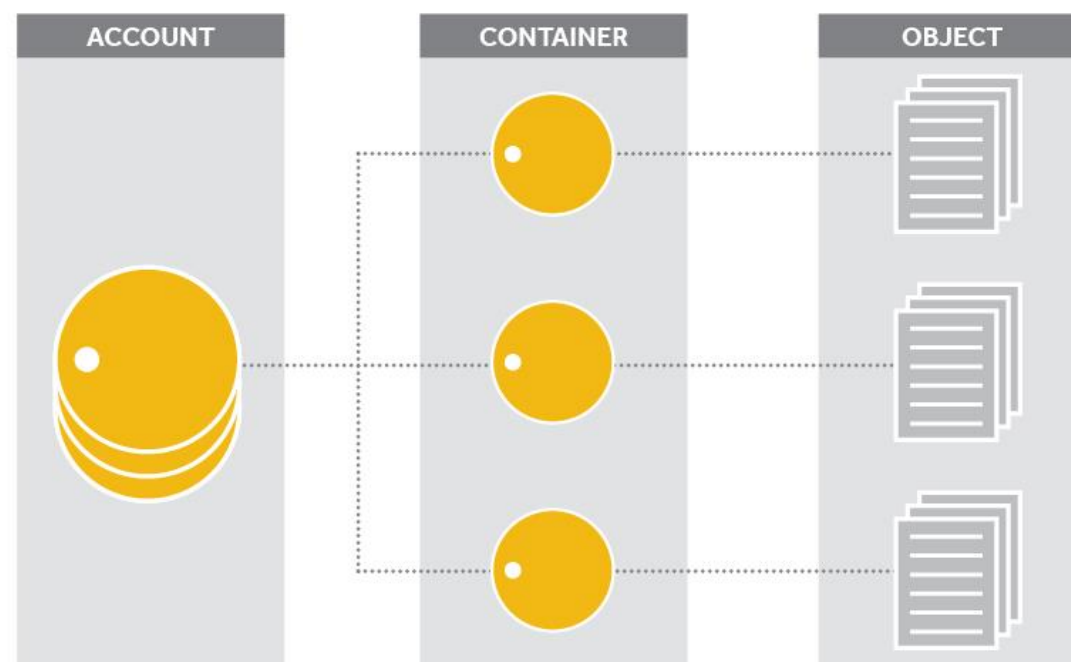
SWIFT HTTP REQUESTS

- The process uses the object ring to look up the IP and other information for those three devices.

- All requests sent to Swift are made up of at least three parts :

- PUT : A client uses the Swift API to make an HTTP request to PUT an object into an existing container.

- GET : A client uses the Swift API to make an HTTP request to GET an object from the cluster.

- Storage URL (swift.example.com/v1/account)

- Cluster location: swift.example.com/v1/

- Storage location : /account/container/object
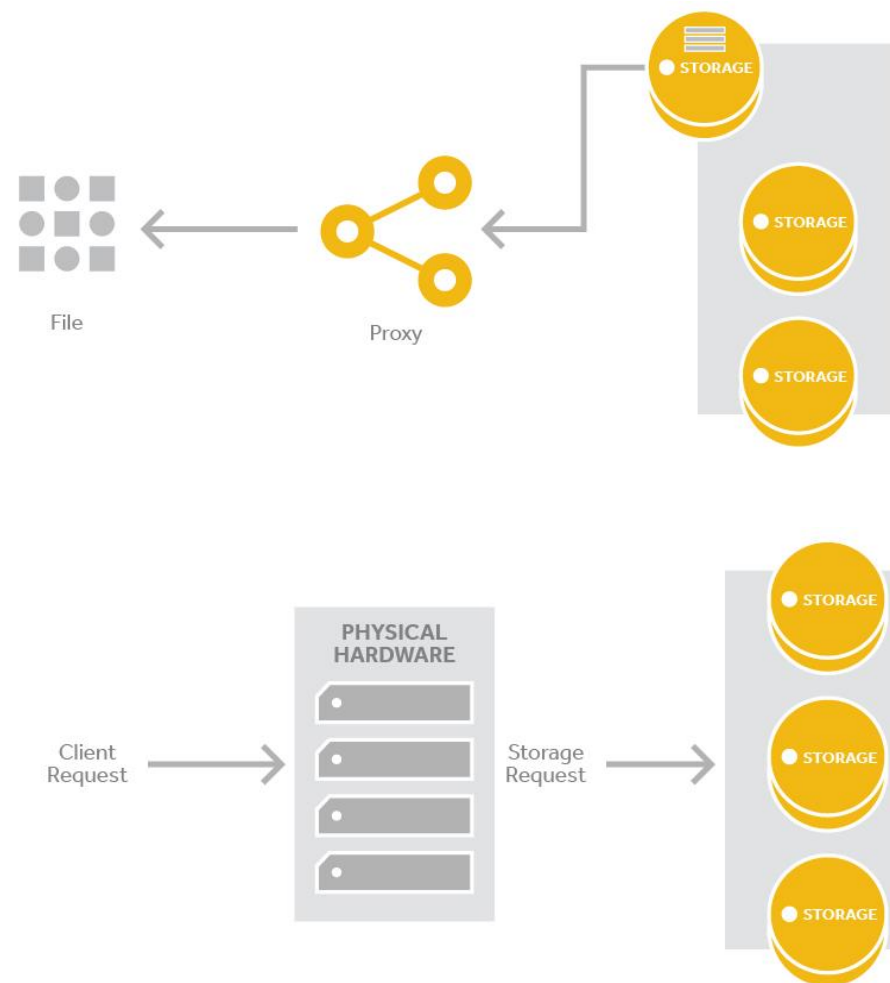
HAPPY CUSTOMERS !!
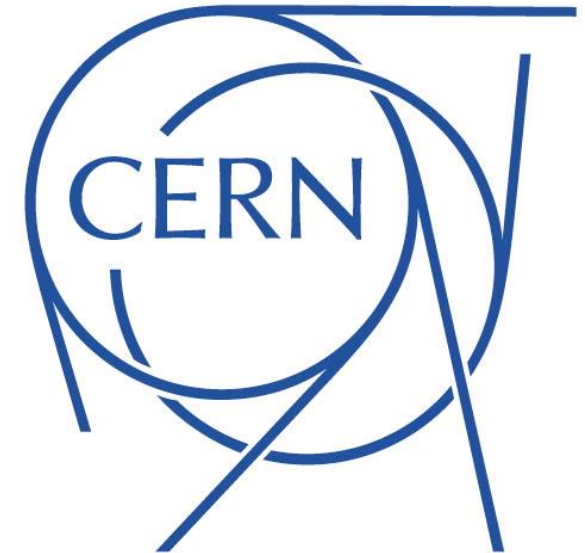
QUESTIONS ??

THANK YOU

Connect with me:
ramitsurana@gmail.com

Github : ramitsurana05
Twitter : ramitsurana