

Real Telco OpenStack upgrade lessons learned

Radosław Śmigielski

Gábor Sztás

2019.04.30

The Plan

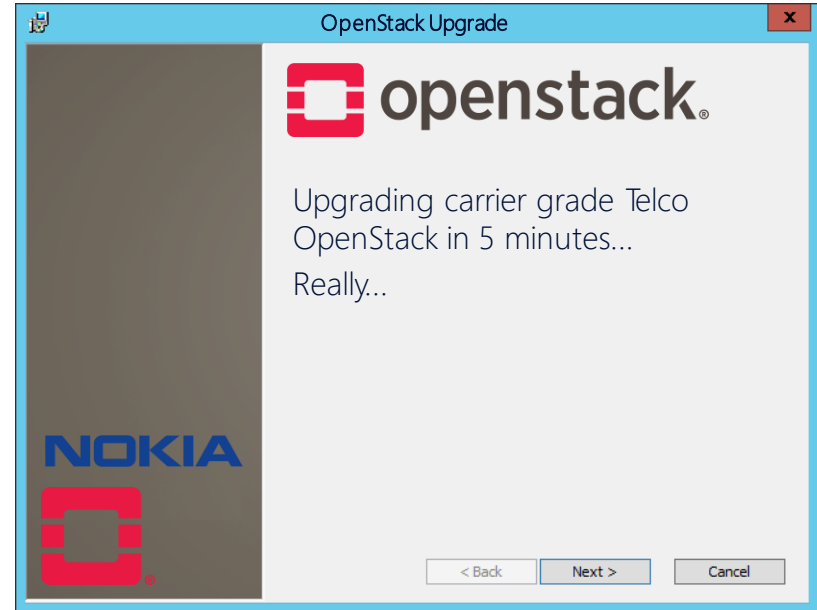
What are we going to talk about here?

1. Why are we here?
2. Define what is and what is not Telco upgrade?
3. The Upgrade
4. What could go wrong?

Why?

Why are we here?

1. OpenStack 6 months release cycle is too fast for many customers.
2. Upgrade every 2-4 years is what most of customers can do
3. Extended support requests
4. TripleO N+3 Fast Forward Upgrade



Telco upgrade

Definition

1. Upgrade with running VNFs and live traffic
2. Prevent/minimize traffic outage
3. Extended support: TripleO FFU N+3 OpenStack upgrade, Newton to Queens
4. Operating system upgrade
5. Third party NICs drivers upgrade
6. Ceph upgrade from Jewel latest to Luminous latest
7. Ceph data migration to BlueStore
8. Cluster reboot
9. Fit into maintenance windows
10. What is live traffic, carrier grade, MW and traffic offload?

OpenStack upgrade

FFU N+3 upgrade

1. OpenStack TripleO N+3 FFU upgrade
 - Undercloud
 - Overcloud
2. Services SQL schema upgrade, data migration
3. OpenStack services switch from classic systemd services to docker containers
4. New Pacemaker



Base system upgrade

Ordinary upgrade with nasty consequences

1. Ordinary base system upgrade
2. New kernel
3. Third party NICs drivers

Third party NICs drivers

SRIOV drivers

1. Relevant for all SRIOV NICs
2. New kernel brings (old) upstream drivers
3. SRIOV NICs upgrade
 - Never reduce number of VF to 0
 - Never uninstall existing drivers, do in-place upgrade instead
4. Problems
 - Nova lose track of VFs
 - Confuses network controllers if present

Live traffic

Running VNFs

1. Running VNFs with live traffic is a requirement
2. Multiple VNFs from different vendors onboard
3. Cooperation with VNFs
 1. Rebalance traffic before compute reboot
 2. Make sure services are operational after compute reboot
 3. Rebalance traffic after compute reboot
4. Significant time needed for multiple VNFs

Ceph upgrade

The easy part

1. Ceph upgrade Jewel to Luminous
2. Ceph services go to containers

Ceph data migration

The hard part

1. Ceph data format major (revolutionary) change
File Store -> BlueStore
2. No easy way of migrating data between FileStore and BlueStore
3. Essentially remove/add one or more OSDs and let Ceph recover
4. Scale
 1. Hundreds of OSDs
 2. Average OSD size 1TB (SSD), 1.4TB (NVMe) or 6-9TB (spinning)
 3. Cluster size hundreds of TBs, even PB
 4. Amount of data in Ceph cluster
5. Time...
 - Fastest recovery means bigger impact on running VMs (QoS of CEPH)
6. Alternative is to stay on FileStore with Ceph Luminous

Real example

True story

1. Multiple customers in the pipeline evaluation upgrade
2. Customer scale, data centers with 1500+ servers
3. Alternative scratch installation
4. Ceph FileStore to BlueStore migration time is a major issue
5. Upgrade evaluation, risk assessment and time estimation to help customers make decision

NOKIA