



## **User Experiences with Implementing RedHat OpenShift Container Platform (RHOCP) under z/VM:**

**Multiple Nodes with Assisted Installer  
Single Node with Assisted Installer**

Samuel D. Cohen  
Levi, Ray & Shoup, Inc.  
[sam.cohen@lrs.com](mailto:sam.cohen@lrs.com)  
(217) 862-9227



## What I was trying to do

- Understand Openshift better
  - Learn more about the technology and terminology
- See what it takes to setup a standalone demo environment under z/VM
- Try to understand the Redhat-provided documentation
- Provide “lessons learned” for others that want to try this at home
- Discover the differences between Automated Installer and the “old way”
- I can’t answer what you would do with this, other than impress your Kubernetes-centric colleagues
  - Lots of continuously-running processes add a lot of overhead (more than z/OS)



## Credits

- S. Michael Benson, *Docker Swarm or Kubernetes?*, Enterprise Tech Journal, 2019 Issue 6.
- Filipe Miranda, IBM, published articles in LinkedIn on his experience and provided sample settings for DNS, Load Balancer. Also provided 1:1 assistance
  - <https://www.linkedin.com/pulse/red-hat-openshift-installation-process-experiences-ibm-filipe-miranda/>
  - <https://www.linkedin.com/pulse/understanding-network-definitions-from-openshift-4-ibm-filipe-miranda/> describes network settings used internally by Openshift
- Redhat Technical Support for background on new installation steps



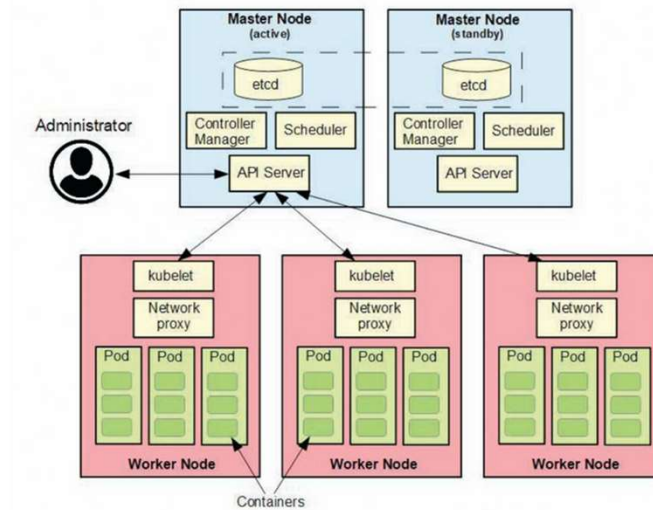
## Terminology

- RedHat uses Kubernetes as a deployment and management engine for “pods”
- Terminology
  - Container: Think of it like a stand-alone application with static links
  - Pod: a business application, made up of one or more containers
  - Service: multiple identical pods distributed throughout the cluster for load balancing and higher availability
  - Node: a virtual machine (in our case) where multiple pods can be deployed
  - Control (Master) Node: schedules and controls pods across multiple compute nodes
  - Compute (Worker) Node: where pods are actually executing
  - Kubelet: an agent in each compute node for communication with the master nodes

Reference: S. Michael Benson, *Docker Swarm or Kubernetes?*, Enterprise Tech Journal, 2019 Issue 6. Used with permission



## Better shown in a picture



Reference: S. Michael Benson, *Docker Swarm or Kubernetes?*, Enterprise Tech Journal, 2019 Issue 6. Used with permission

In later RHOCP releases, master nodes have been renamed control nodes; worker nodes have been renamed compute nodes



## External Services

- Layer 2 TCP/IP Network
  - Used z/VM Virtual Switch
- Domain Name Services (named)
  - Used for external access to OpenShift Services
- Web Server (apache)
  - Used for initial loading of RHCOS root filesystem (operating system)
- Traffic Load Balancer (haproxy) for multiple node configuration
  - Used to rotate incoming requests between pods
- Defined range of IP addresses for externally-visible services
  - Used with Domain Name Services for specific deployed workloads
- Zoned SAN access to disk if using LUNs larger than 3390 can handle



## Deployment process

- Control nodes and compute nodes use CoreOS as the operating system
  - CoreOS (RHCOS) is lighter weight than RHEL (even in minimal installation)
- Create at least one server (called a “bastion” server by some) as the source of initial RHCOS rootfs via http (e.g. through a web server) and to provide a point of entry to your RHOCF cluster with a load balancer to your nodes (only in a multi-node environment)
- Multiple nodes using Assisted Installer: One defined control node acts as a “boot” node for the other nodes, then is reconfigured as a control node
- Single node using Assisted Installer: The single node is used as a boot node for initial definitions, then continues to load compute and control node info

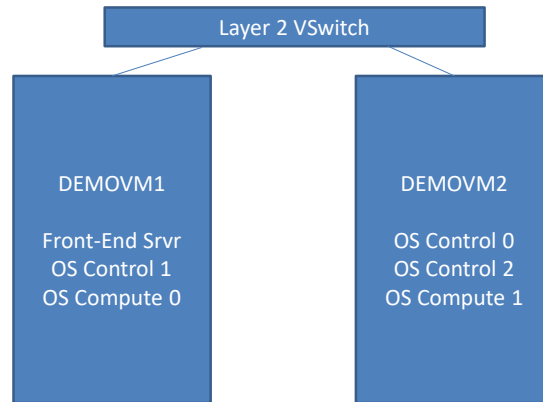


## PoC Environment – Multiple Nodes

- I defined 3 Control nodes and 2 Compute nodes
- I split nodes between 2 different SSI cluster members
- I used FB disks attached via 2 FCP channels
  - NPIV is active so subchannel traffic only goes to the appropriate VM guest
    - Otherwise, all guests see all FCP traffic
  - I added multipath connections for both SSI cluster members in each node
- I created a single RHEL 9 guest to provide required front-end interfaces:
  - http server for reading installation images
  - haproxy for load balancing incoming messages to multiple nodes
- I used my existing DNS server for resolving cluster addresses



## OpenShift Demo Environment



Cluster name:  
DNS Domain:  
Machine network:  
Cluster network:  
Kubernetes "service" network:

democluster2  
demo.lrsinc.org  
10.96.64.192/26  
10.128.0.0/14  
172.30.0.0/16



## Getting Started

- Prepared z/VM system (Layer 2 VSwitch)
- Prepared existing DNS server
- Prepared a front-end virtual machine (load balancer, web server, code)
- Defined OpenShift virtual machines
  - Each Control machine has 6 virtual IFL Processors, 16G of Memory, VNIC to a Layer 2 Vswitch
  - Each Compute machine has 2 virtual IFL Processors, 8G of Memory, VNIC to a Layer 2 Vswitch
- Registered with Redhat so I can download the code
- Used Redhat Hybrid Cloud Console to define my cluster:
  - (<https://console.redhat.com/openshift/assisted-installer/clusters/~new>)
  - Also download a file called “pull-secret”, generated by your userid/password
- Kept documentation handy
  - [https://docs.openshift.com/container-platform/latest/installing/installing\\_ibm\\_z/installing-ibm-z.html](https://docs.openshift.com/container-platform/latest/installing/installing_ibm_z/installing-ibm-z.html)
  - [https://docs.redhat.com/en/documentation/openshift\\_container\\_platform/latest/html/installing\\_on-premise\\_with\\_assisted\\_installer](https://docs.redhat.com/en/documentation/openshift_container_platform/latest/html/installing_on-premise_with_assisted_installer)



## **z/VM Configuration Entries**

- MACADDR Specifications in SYSTEM CONFIG
  - Defaults may be sufficient
- Vswitch Definition
  - VLAN Awareness is not necessary
- Control and Compute Node Directory Profiles
- Control and Compute Node USER Definitions
  - Need unique final octets of MACADDRs for each virtual NIC when using Assisted Installer



## SYSTEM CONFIG MACADDR Specification

```
DEMOVM1: VMLAN Mac 020001 UserPre 020000  
DEMOVM2: VMLAN Mac 020002 UserPre 020000
```

In an SSI configuration, the Vswitch MACADDR should be unique for each z/VM SSI member, while the user virtual MACADDR should be the same in case you want to relocate the Linux guest from one SSI member to another



## Vswitch Definition (Layer 2)

```
Rdevice 0200-022F EQID ETH00200 Type OSA  
Define VSwitch OSVSW Rdev 0223.P00 Ethernet VLAN Unaware
```

Note this is a Layer 2 Vswitch.



## z/VM Directory Profiles

### PROFILE OSMASTR

CLASS GL  
STORAGE **16G**  
MAXSTORAGE **24G**  
ACCOUNT OPNSHIFT  
ACIGROUP OPNSHIFT  
MACHINE ESA **6**  
COMMAND SET VCONFIG MODE LINUX  
COMMAND DEFINE CPU 00 TYPE IFL  
COMMAND DEFINE CPU 01 TYPE IFL  
...  
COMMAND DEFINE CPU 05 TYPE IFL  
COMMAND DEFINE STORAGE INITIAL STANDBY REMAINDER  
CRYPTO APVIRT  
IPL CMS  
IUCV ALLOW  
IUCV \*IDENT RESANY GLOBAL  
OPTION TODENABLE APPLMON  
CONSOLE 0009 3215 T  
SPOOL 000C 2540 READER \*  
SPOOL 000D 2540 PUNCH A  
SPOOL 000E 1403 A  
(CMS Minidisk Links)

### PROFILE OSWRKR

CLASS GL  
STORAGE **8G**  
MAXSTORAGE **12G**  
ACCOUNT OPNSHIFT  
ACIGROUP OPNSHIFT  
MACHINE ESA **2**  
COMMAND SET VCONFIG MODE LINUX  
COMMAND DEFINE CPU 00 TYPE IFL  
COMMAND DEFINE CPU 01 TYPE IFL  
COMMAND DEFINE STORAGE INITIAL STANDBY REMAINDER  
CRYPTO APVIRT  
IPL CMS  
IUCV ALLOW  
IUCV \*IDENT RESANY GLOBAL  
OPTION TODENABLE APPLMON  
CONSOLE 0009 3215 T  
SPOOL 000C 2540 READER \*  
SPOOL 000D 2540 PUNCH A  
SPOOL 000E 1403 A  
(CMS Minidisk Links)



## z/VM Directory Entries

- Front-End Services  
**USER OSDHCPD LBYONLY**
  - IP Address = 10.96.64.193/24INCLUDE LNXGUEST  
MACHINE ESA 1  
DEFINE STORAGE 2G  
SHARE REL 500  
NICDEF 1000 TYPE QDIO LAN SYSTEM VSWITCH1  
(FCP channels)
- Control Node(s)  
**USER MASTER[0-2] LBYONLY**
  - IP Addresses 10.96.64.196-198/26INCLUDE OSMASSTR  
NICDEF 1000 TYPE QDIO LAN SYSTEM OSVSW  
MACADDR xx  
(FCP channels)
- Compute Node(s)  
**USER WORKER[0-1] LBYONLY**
  - IP Addresses 10.96.64.199-200/26INCLUDE OSWRKR  
NICDEF 1000 TYPE QDIO LAN SYSTEM OSVSW  
MACADDR xx  
(FCP channels)

- NICDEFs are not included in the PROFILE entries, since each Virtual NIC must have a unique MACADDR for configuring the Installation Assistant.
- Note the different hostnames depending on the IP address in the front-end services machine; this is necessary to separate RHOCP traffic when loadbalancing with multiple clusters in the same load balancer
- The Front-End Services do not have to be in the same network as the RHOCP machines, but must be reachable by the end-users and the RHOCP machines



## Prepare for Installation – DNS Updates

- Updated existing DNS server with zone demo.Irsinc.org for RHOCP machines
  - Enabled Dynamic DNS in demo.Irsinc.org
  - Added new entries for external network:
    - Virtual machine addresses in range 10.96.64.192/26 to include existing default gateway
      - » Used existing demo systems' network 10.96.64.0/24 with gateway 10.96.64.254
    - \*.democluster2.demo.Irsinc.org entries were added as CNAME entries to front-end server for load balancing



## DNS Config File extract

**/var/lib/named/dyn/demo.lrsinc.org**

```
$TTL 2d
@           IN SOA      lnxutils.lrsinc.org.  root.lnxutils.lrsinc.org. (
                                2025121500      ; serial
                                3h              ; refresh
                                1h              ; retry
                                1w              ; expiry
                                1d )           ; minimum

$ORIGIN demo.lrsinc.org.
demo.lrsinc.org.      IN NS      lnxutils.demo.lrsinc.org.
...
osdhcpd              IN A      10.96.64.193
master0              IN A      10.96.64.196
master1              IN A      10.96.64.197
master2              IN A      10.96.64.198
worker0              IN A      10.96.64.199
worker1              IN A      10.96.64.200
api.democluster2    IN CNAME  osdhcpd
*.apps.democluster2 IN CNAME  osdhcpd
api-int.democluster2 IN CNAME  osdhcpd
```



## DNS Config File extract

**/var/lib/named/master/64.96.10.in-addr.arpa**

```
$TTL 2D
@           IN SOA      lnxutils.demo.lrsinc.org.
root.lnxutils.demo.lrsinc.org. (
    2025121800    ; serial
    3H           ; refresh
    1H           ; retry
    1W           ; expiry
    1D )         ; minimum
IN NS      lnxutils.demo.lrsinc.org.

$ORIGIN 64.96.10.in-addr.arpa.
193       IN PTR      osdhcpd.demo.lrsinc.org.
196       IN PTR      master0.demo.lrsinc.org.
197       IN PTR      master1.demo.lrsinc.org.
198       IN PTR      master2.demo.lrsinc.org.
199       IN PTR      worker0.demo.lrsinc.org.
200       IN PTR      worker1.demo.lrsinc.org.
```



## Prepare for Installation – Front-End Services

- Defined Front-End Server (RHEL 9.7) with Apache webserver, haproxy load balancer
  - Created filesystem at mountpoint /mnt to hold installation utilities
  - Defined new mountpoint from anonymous user entry point to Apache webserver
    - /var/www/html/mnt
  - Added bind mount to /etc/fstab to make contents of /mnt visible to anonymous web server clients
    - /etc/fstab: /mnt /var/www/html/mnt none bind,nofail 00
  - Made contents of /mnt world-readable:
    - chmod +r /var/www/html/mnt
  - Defined user “core” with password “corepass”
    - Defined ssh key for this user for copying “public” key to ignition file
    - Commands used:
      - » su – core
      - » ssh-keygen -t rsa -b 4096 -N '' -f /home/core/.ssh/id\_rsab
  - Setup Apache server to listen on port 81 and enable anonymous read-only access
  - Setup haproxy to point to the correct nodes (see example)

Note: RHEL’s provided anonymous entry point is not the same as SUSE’s anonymous entry point.



## Prepare the Installation Assistant

- Login to <https://console.redhat.com>
  - select “RedHat OpenShift”
  - Find “RedHat OpenShift Container Platform” and select “Create Cluster”
  - From the “Select an OpenShift cluster type to create” window, select the “Datacenter” Tab, then find “IBM Z”
  - Choose “Interactive” installation
  - Initial Panel:
    - Enter your cluster name and DNS domain name
    - Be sure to select IBM Z architecture
    - List the number of control nodes
    - Select Static IP
  - Network-Wide Configurations Panel:
    - Enter the DNS IP address along with cluster IP network/mask
  - Host-Specific Configurations Panel:
    - Enter the IP address and MACADDR for each node of the cluster (MAC bytes are colon-separated)
    - Press “Next” when you have added all the IP/MAC Addresses for your cluster



## Prepare the Installation Assistant

- Do not select any Operators
- Host-Discovery Panel:
  - The panel is waiting for each initially-booted machine to contact it.
  - You don't have the correct configuration yet to boot the virtual machines
  - To get that configuration, choose the "Add Hosts" button just below the "Host Discovery" title
  - Select an iPXE file and include the `/home/core/.ssh/id_rsab` contents from your front-end machine
  - Generate the ISO and get the file contained in the iPXE script file URL (via your browser)
  - Using the contents of the iPXE script file, download the contents of the individual URLs for the kernel, initrd and rootfs. Load those into `/mnt` on the front-end services machine and make them world-readable
  - Copy the kernel and initrd to a z/VM disk to be punched to the virtual readers of each node.
  - Create a PARMFILE that provides the disk and network hardware information for boot along with a URL to the rootfs via http
  - On each RHOCN node, punch the kernel, parmfile and initrd to the reader and boot from the reader. The initrd will set the IP addresses based on the MACADDR in the z/VM Directory for each node.
  - You should see the nodes appear in the "host inventory" section of the RHOCN console. When all hosts are there, you can proceed.



## Parameter File for Installations

```
ro rd.neednet=1 console=ttysclp0 fips=0
ignition.firstboot ignition.platform.id=metal
coreos.live.rootfs_url=http://10.96.64.193:81/mnt/discovery.rootfs
rd.znet=qeth,0.0.1000,0.0.1001,0.0.1002,layer2=1,portno=0
rd.zfcp=0.0.2000,0x5005076812151BCB,0x0000000000000000
rd.zfcp=0.0.2000,0x5005076812151BCC,0x0000000000000000
rd.zfcp=0.0.2100,0x5005076812161BCB,0x0000000000000000
rd.zfcp=0.0.2100,0x5005076812161BCC,0x0000000000000000
zfcp.allow_lun_scan=0
```

← fips flag was required

← IP address of front-end

← Network hardware

← 100GB LUN (multipath)



# Clips from the installation

Red Hat Hybrid Cloud Console

OpenShift Cluster List

Filter by name or ID... Cluster type... [Create cluster](#) [Register cluster](#) [View cluster archives](#) [View cluster reports](#)  View only my clusters

| Name                           | Status | Type | Created     | Version | Provider (Region) |
|--------------------------------|--------|------|-------------|---------|-------------------|
| <a href="#">demo-cluster-1</a> | Draft  | OCP  |             | 4.30.8  | N/A               |
| <a href="#">demo-cluster-2</a> | Ready  | OCP  | 03 Jan 2026 | 4.30.8  | NONE              |

1-2 of 2



# Clips from the installation

OpenShift

- 1. Cluster details
- 2. Static network configurations
- 3. Operators
- 4. Host discovery
- 5. Storage
- 6. Networking
- 7. Review and create

### Cluster details

Technology Preview  I'm installing on a disconnected/air-gapped/secure environment

Cluster name \*

Base domain \*

Enter the name of your domain (demo.k8s.org). The cannot be changed after cluster install. All DNS records must include the fully-qualified subdomains of the base you enter. The full cluster address will be **democluster2.demo.k8s.org**.

OpenShift version \*  
OpenShift 4.20.8

CPU architecture \*  
ARMv2 (a30c)

Integrate with external partner platforms. \*  
No platform integration.

Number of control plane nodes

Include custom manifests   
Additional manifests will be applied at the install time for advanced configuration of the cluster.

Host's network configuration  
 DHCP only  Static IP (bridge, veth, and bond)

Encryption of installation disks

- Control plane nodes
- Workers
- Admin



# Clips from the installation

Questions

- Cluster details
- Static network configurations
- Network-wide configurations
- Host specific configurations
- Operators
- Host discovery
- Storage
- Networking
- Review and create

### Static network configurations

Network configuration can be done using either the form view or YAML view. Configurations done in this step are for discovering hosts.

Configure via:  Form view  YAML view

Form view supports basic configurations. Select YAML view for advanced configurations.

**⚠ To add new hosts that will use the new or edited configurations, you'll need to regenerate the Discovery ISO in the 'Host discovery' step and boot your new hosts from it.**

### Network-wide configurations

The following configurations are applicable to all the hosts.

**⚠ If you edit the values and click Next, the earlier configurations will be replaced with the new values for all of the hosts.**

Networking stack type \*

IPv4  Dual-stack  IPv6

Use VLAN

DNS \*

10.96.64.10

IPv4

Subnet \*  /  (0/10.96.64.0 - 10.96.64.255)

Default gateway \*



# Clips from the installation

Overview Access central Cluster history Support

1: Cluster details

2: **Static network configurations**

3: Operations

4: Host discovery

5: Storage

6: Networking

7: Review and create

### Static network configurations

Network configurations can be done using either the form view or YAML view. Configurations done in this step are for discovering hosts.

**To add new hosts that will use the new or edited configurations, you'll need to regenerate the Discovery ISO in the 'Host discovery' step and boot your new hosts from it.**

#### Host specific configurations

Host 1  
MAC to IP mapping: 02:00:00:00:00:00 → 10.26.64.96

Host 2  
MAC to IP mapping: 02:00:00:00:00:00 → 10.26.64.97

Host 3  
MAC to IP mapping: 02:00:00:00:00:00 → 10.26.64.98

Host 4  
MAC to IP mapping: 02:00:00:00:00:00 → 10.26.64.99

Host 5  
MAC to IP mapping: 02:00:00:00:00:00 → 10.26.64.200

[Add another host configuration](#)

[Next](#) [Back](#) [Cancel](#) [View cluster events](#)



# Clips from the installation

Overview Access control Cluster history Support

- 1 Cluster details
- 2 Static network configurations >
- 3 Operators
- 4 Host discovery
- 5 [Finraan](#)

### Operators

Find bundles or operators

> Single Operators (28 | 0 selected)

Next Back Cancel View cluster events

#### Subscription settings ⚠

|  |                                      |
|--|--------------------------------------|
| <b>Subscription type</b><br>Not set                                    | <b>Cluster usage</b><br>Not set      |
| <b>Service level agreement (SLA)</b><br>Self-Support 60-day evaluation | <b>Subscription units</b><br>Not set |
| <b>Support type</b><br>Not set   |                                      |

[Edit subscription settings](#)



# Clips from the installation

Overview | Access control | Cluster history | Support

- 1 Cluster details
- 2 Static network configurations
- 3 Operators
- 4 Host discovery**
- 5 Storage
- 6 Networking
- 7 Review and create

### Host discovery

[Add hosts](#)

Run workloads on control plane nodes

**Information & Troubleshooting**  
[Missing hardware requirements](#) [Hosts not showing up?](#)

### Host Inventory

0 selected [Actions](#)

| Hostname  | Role | Status | Discovered on | CPU Cores | Memory | Total storage |
|---|------|--------|---------------|-----------|--------|---------------|
| <b>(0)</b><br>Waiting for hosts...<br>Hosts might take a few minutes to appear here after booting.<br><a href="#">Hosts not showing up?</a> |      |        |               |           |        |               |

**Cluster is not ready yet.**  
The following requirement must be met:  
The cluster must have exactly 3 dedicated control plane nodes. Add or remove hosts, or change their roles configurations to meet the requirement.

[Next](#) [Back](#) [Cancel](#) [View cluster events](#)







## Clips from the installation

OpenShift  
democluster2

Alerts and recommendations 1

Overview Access control Cluster history Support Add Hosts

1 Cluster details  
2 Static network configurations  
3 Operators  
4 Host discovery  
5 Storage  
6 Networking  
7 Review and create

### Storage

| Hostname               | Role        | Status | Total storage | Number of disks |  |
|------------------------|-------------|--------|---------------|-----------------|--|
| master0.demos.bosn.org | Auto-assign | Ready  | -             | 5               |  |
| master1.demos.bosn.org | Auto-assign | Ready  | -             | 5               |  |
| master2.demos.bosn.org | Auto-assign | Ready  | -             | 5               |  |
| worker0.demos.bosn.org | Auto-assign | Ready  | 10737 GB      | 5               |  |
| worker1.demos.bosn.org | Auto-assign | Ready  | -             | 5               |  |

**Installation disk limitations**

- The installation disk selected for host master0.demos.bosn.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk selected for host master1.demos.bosn.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk selected for host master2.demos.bosn.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk selected for host worker0.demos.bosn.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk selected for host worker1.demos.bosn.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.

All bootable disks, except for read-only disks, will be formatted during installation. Make sure to back up any critical data before proceeding.

Next Back Cancel View cluster events

As each virtual machine comes up, the discovery.rootfs maps the MACADDR to the assigned IP address and reports the findings up to the RHOC console. This is the point where you want to change the boot information under each hostname to use multipath addresses instead of individual disk paths.



## Clips from the installation

OpenShift

Overview Access control Cluster history Support Add hosts

1 Cluster details

2 Static network configurations

3 Operators

4 Host discovery

**5 Storage**

6 Networking

7 Review and create

### Storage

| Hostname                | Role        | Status | Total storage | Number of disks | ... |
|-------------------------|-------------|--------|---------------|-----------------|-----|
| master0.demos.bosch.org | Auto-assign | Ready  | -             | 5               | ... |

### 5 Disks

| Name            | Role              | Limit | Form.                               | Drive type | Size      | Serial           | Model | Wear                             |
|-----------------|-------------------|-------|-------------------------------------|------------|-----------|------------------|-------|----------------------------------|
| sda (bootable)  | Name              | -     | <input type="checkbox"/>            | FC         | 107.37 GB | 0354A04037956000 | 2445  | 04605076803800458000000000000000 |
| sdb (bootable)  | Name              | -     | <input type="checkbox"/>            | FC         | 107.37 GB | 0354A04037956000 | 2445  | 04605076803800458000000000000000 |
| sdc (bootable)  | Name              | -     | <input type="checkbox"/>            | FC         | 107.37 GB | 0354A04037956000 | 2445  | 04605076803800458000000000000000 |
| sdd (bootable)  | Name              | -     | <input type="checkbox"/>            | FC         | 107.37 GB | 0354A04037956000 | 2445  | 04605076803800458000000000000000 |
| dm-0 (bootable) | Installation-disk | IS    | <input checked="" type="checkbox"/> | Multipath  | 107.37 GB |                  |       | 04605076803800458000000000000000 |

|                         |             |       |           |   |     |
|-------------------------|-------------|-------|-----------|---|-----|
| master0.demos.bosch.org | Auto-assign | Ready | -         | 5 | ... |
| master1.demos.bosch.org | Auto-assign | Ready | -         | 5 | ... |
| worker0.demos.bosch.org | Auto-assign | Ready | 107.37 GB | 5 | ... |
| worker1.demos.bosch.org | Auto-assign | Ready | -         | 5 | ... |

**Installation disk limitations**

- The installation disk sda selected for host master0.demos.bosch.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk sdb selected for host master1.demos.bosch.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.
- The installation disk sdd selected for host worker1.demos.bosch.org is managed by multipath. We strongly recommend using the multipath device dm-0 to improve reliability.

All bootable disks, except for read-only disks, will be formatted during installation. Make sure to back up any critical data before proceeding.

Note that by opening the “twisty”, you can select the multipath device instead of an individual disk path. As you do that, the installation note at the bottom is updated to remove the recommendation for that particular virtual machine.



# Clips from the installation

OpenShift

Cluster List > democluster2

democluster2

> Alerts and recommendations 13

Overview Access control Cluster history Support Add Hosts

- 1 Cluster details
- 2 Static network configurations
- 3 Operators
- 4 Host discovery
- 5 Storage
- 6 Networking
- 7 Review and create

## Storage

| Hostname                              | Role        | Status | Total storage | Number of disks |  |
|---------------------------------------|-------------|--------|---------------|-----------------|--|
| <a href="#">master0.demo.banc.org</a> | Auto-assign | Ready  | -             | 5               |  |
| <a href="#">master1.demo.banc.org</a> | Auto-assign | Ready  | -             | 5               |  |
| <a href="#">master2.demo.banc.org</a> | Auto-assign | Ready  | -             | 5               |  |
| <a href="#">worker0.demo.banc.org</a> | Auto-assign | Ready  | 10737 GB      | 5               |  |
| <a href="#">worker1.demo.banc.org</a> | Auto-assign | Ready  | -             | 5               |  |

All bootable disks, except for read-only disks, will be formatted during installation. Make sure to back up any critical data before proceeding.

Next Back Cancel

View cluster events



# Clips from the installation

Search

democluster2

Network and configuration

Overview Access control Cluster history Support Add Nodes

- Cluster details
- Static network configurations
- Operators
- Host discovery
- Storage
- Networking**
- Review and update

### Networking

Network Management

Cluster Management Network Management

Refer to the [Cluster Management/Network Management](#) for configuring your cluster's networking, including:

- DHCP or static IP Addresses
- Load balancers
- Network ports
- DNS

Use advanced networking  
Configure advanced networking (experimental, TDD changes)

Host SSH Public Key for troubleshooting after installation

Use the same host discovery SSH key

#### Host inventory

| Hostname                  | Role        | Status | Active MC | IPv4 addr. | IPv6 addr. | MAC addr. | ... | ... |
|---------------------------|-------------|--------|-----------|------------|------------|-----------|-----|-----|
| master01.democluster2.com | Auto assign | Ready  | -         | -          | -          | -         | -   | -   |
| master02.democluster2.com | Auto assign | Ready  | -         | -          | -          | -         | -   | -   |
| master03.democluster2.com | Auto assign | Ready  | -         | -          | -          | -         | -   | -   |
| worker01.democluster2.com | Auto assign | Ready  | -         | -          | -          | -         | -   | -   |
| worker02.democluster2.com | Auto assign | Ready  | -         | -          | -          | -         | -   | -   |

Next Back Cancel View cluster events



# Clips from the installation

Overview

democuster2

Cluster details

Cluster configuration

Operations

Host discovery

Storage

Networking

Review and create

Review and create

Cluster summary

- Cluster details
  - Cluster address: 10.10.10.10/24
  - Operator version: 4.2.0.0
  - CPU architecture: x86\_64
  - Host network configuration: Scale 0
- Host inventory
  - Hosts: 1
  - Total CPUs: 4
  - Total memory: 16.00 GB
  - Total storage: 107.37 GB
- Networking
  - Administration type: User Manager (Network)
  - Block type: IPv4

Advanced networking settings

- Cluster network CIDR: 10.10.10.0/24
- Cluster network host prefix: /24
- Service network CIDR: 172.16.0.0/16
- Relay network type: Open vSwitch Network (OVN)

Next > Back Cancel



# Clips from the installation

OpenShift  
Cluster List > democluster2

democluster2

> Alerts and recommendations 1

Overview Access control Cluster history Support Add Hosts

Installation progress

Started on 1/20/2026, 1:42:38 PM

Preparing for installation 0%

[Abort installation](#) [Download kubeconfig](#) [View cluster events](#) [Download Installation Logs](#)

| Hostname               | Role                           | Status                                     | Discovered on          | CPU Cores | Memory   | Total...  |
|------------------------|--------------------------------|--|------------------------|-----------|----------|-----------|
| master0.demo.ksinc.org | Control plane node             | <a href="#">Preparing for installation</a> | 1/20/2026, 10:54:43 AM | 4         | 16.00 GB | -         |
| master1.demo.ksinc.org | Control plane node (bootstrap) | <a href="#">Preparing for installation</a> | 1/20/2026, 10:47:32 AM | 4         | 16.00 GB | -         |
| master2.demo.ksinc.org | Control plane node             | <a href="#">Preparing for installation</a> | 1/20/2026, 10:55:50 AM | 4         | 16.00 GB | -         |
| worker0.demo.ksinc.org | Worker                         | <a href="#">Preparing for installation</a> | 1/20/2026, 10:52:25 AM | 2         | 8.00 GB  | 107.37 GB |
| worker1.demo.ksinc.org | Worker                         | <a href="#">Preparing for installation</a> | 1/20/2026, 10:56:53 AM | 2         | 8.00 GB  | -         |





## PoC Environment – Single Node

- I defined a Single virtual machine for both Control and Compute functions
- I used FB disks attached via 2 FCP channels
  - NPIV is active so subchannel traffic only goes to the appropriate VM guest
    - Otherwise, all guests see all FCP traffic
  - I added multipath connections for both SSI cluster members in each node
- More memory/engines are needed for a single node than for each multi-node v.m.
- I used the same single RHEL 9 guest to provide required front-end interfaces:
  - http server for reading installation images
  - Load balancing is not required
- I used my existing DNS server for resolving cluster addresses



## z/VM Directory Profiles

```
PROFILE OSSINGLE
CLASS GLB
STORAGE 20G
MAXSTORAGE 24G
COMMAND SET VCONFIG MODE LINUX
COMMAND DEFINE CPU 00 TYPE IFL
COMMAND DEFINE CPU 01 TYPE IFL
COMMAND DEFINE CPU 02 TYPE IFL
COMMAND DEFINE CPU 03 TYPE IFL
COMMAND DEFINE CPU 04 TYPE IFL
COMMAND DEFINE CPU 05 TYPE IFL
COMMAND DEFINE CPU 06 TYPE IFL
COMMAND DEFINE CPU 07 TYPE IFL
COMMAND DEFINE STORAGE INITIAL STANDBY REMAINDER
CRYPTO APVIRT
MACHINE ESA 8
IPL CMS
IUCV ALLOW
IUCV *IDENT RESANY GLOBAL
OPTION TODENABLE APPLMON
SHARE RELATIVE 350
CONSOLE 0009 3215 T OPMGRM1
NICDEF 1000 TYPE QDIO LAN SYSTEM OSVSW MACID 00003F
(Spool and CMS Mdisk statements)
```



## z/VM Directory Entries

- Front-End Services (same as before)
  - USER OSDHCPD LBYONLY**
  - IP Address = 10.96.64.193/24
    - Hostname: osdhcpd
  - INCLUDE LNXGUEST
  - MACHINE ESA 1
  - DEFINE STORAGE 2G
  - SHARE REL 500
  - NICDEF 1000 TYPE QDIO LAN SYSTEM VSWITCH1  
(FCP channels)
- Single Node
  - USER RHOCPLBYONLY**
  - IP Addresses 10.96.64.194/26
  - INCLUDE OSSINGLE  
(FCP channels)
  - NICDEF 1000 TYPE QDIO LAN SYSTEM OSVSW MACID 3F

- NICDEFs are not included in the PROFILE entries, since each Virtual NIC must have a unique MACADDR for configuring the Installation Assistant.
- Note the different hostnames depending on the IP address in the front-end services machine; this is necessary to separate RHOCPL traffic when loadbalancing with multiple clusters in the same load balancer
- The Front-End Services do not have to be in the same network as the RHOCPL machines, but must be reachable by the end-users and the RHOCPL machines



## DNS Config File extract

**/var/lib/named/dyn/demo.lrsinc.org**

```
$TTL 2d
@           IN SOA          lnxutils.lrsinc.org.
root.lnxutils.lrsinc.org. (
                                2025121500      ; serial
                                3h                ; refresh
                                1h                ; retry
                                1w                ; expiry
                                1d )              ; minimum

$ORIGIN demo.lrsinc.org.
demo.lrsinc.org.      IN NS          lnxutils.demo.lrsinc.org.
...
osdhcpd              IN A          10.96.64.193
rhocp                 IN A          10.96.64.194
api.democluster      IN CNAME     rhocp
*.apps.democluster   IN CNAME     rhocp
api-int.democluster  IN CNAME     rhocp
```

Note that the cluster DNS entries point directly to the single node, not to the location with a load balancer.



## DNS Config File extract

**/var/lib/named/master/64.96.10.in-addr.arpa**

```
$TTL 2D
@           IN SOA      lnxutils.demo.lrsinc.org.
root.lnxutils.demo.lrsinc.org. (
                                2025121800    ; serial
                                3H             ; refresh
                                1H             ; retry
                                1W             ; expiry
                                1D )           ; minimum
                                IN NS         lnxutils.demo.lrsinc.org.

$ORIGIN 64.96.10.in-addr.arpa.
193                IN PTR      osdhcpd.demo.lrsinc.org.
194                IN PTR      rhocp.demo.lrsinc.org.
```



## Parameter File for Initial Installation

```
ro rd.neednet=1 console=ttysclp0 fips=0
ignition.firstboot ignition.platform.id=metal
coreos.live.rootfs_url=http://10.96.64.193:81/mnt/discover.rootfs
rd.znet=qeth,0.0.1000,0.0.1001,0.0.1002,layer2=1,portno=0
rd.zfcp=0.0.2000,0x5005076812151BCB,0x0000000000000000
rd.zfcp=0.0.2000,0x5005076812151BCC,0x0000000000000000
rd.zfcp=0.0.2100,0x5005076812161BCB,0x0000000000000000
rd.zfcp=0.0.2100,0x5005076812161BCC,0x0000000000000000
zfcp.allow_lun_scan=0
```

← fips flag was required

← Downloaded

← Network hardware

← 100GB LUN (multipath)

Note the same name for the rootfs as used for multi-node. The key during initial installation is a different initrd for each cluster configuration. The initrd maps the MACADDR to the IP address. If you are building clusters at different release levels, each release level will have a different kernel and rootfs, while the initrd is unique to each cluster (and must match the release level it was built for).



## Single Node vs. Multiple-Nodes

- Setup via Assisted Installer is simpler due to only one virtual machine
  - Larger (CPU, memory) single v.m. needed due to combined workload
- Not high-availability
  - Better for development, unit testing and integration testing
  - Use multi-node for load-testing and deployment
  - Spreading multi-node between SSI clusters allows for VM maintenance with reduced end-user impact



## Assisted Installation Results

- Much easier to use Assisted Installer than performing manual installation
  - Assuming you can take the defaults and limited modifications
- Each Cluster was successfully built
- z/VM Live Guest Relocation:
  - Front-end machine could be relocated
  - Multi-nodes could not be relocated
  - Single-nodes could be relocated
- Graphical interfaces worked (once I pointed to the DNS server in my VPN network settings)
- Someone else gets to figure out what to do with it



Questions?





## Manual Installation (The old way)



## Create installation ignition files

- Unpack oc.tar and openshift-install.tar in /mnt
- Verify the pull-secret file
  - Go to a website like *jsonlint.com* to verify the JSON in this file.
  - Verify you aren't missing any delimiters or curly braces (easily missed in a cut-and-paste)
- Create an “install-config.yaml” file
  - Position sensitive
  - Ensure “architecture: s390x” under computer and controlPlane tags
  - Imbed the pull-secret file
  - Imbed the public key from /home/core/.ssh/id\_rsa.pub
  - Note: the YAML file will disappear after it is used to create the manifests and ignition files
    - Openshift-install program uses the term “consumed”
    - Make a copy so you can repeat the process more easily
  - Don't change the cluster network and service network parameters



## Install-config.yaml sample file

```
apiVersion: v1
baseDomain: demo.lrsinc.org
compute:
- architecture: s390x
  hyperthreading: Enabled
  name: worker
  replicas: 0
controlPlane:
  architecture: s390x
  hyperthreading: Enabled
  name: master
  replicas: 1
metadata:
  name: democluster
networking:
  clusterNetwork:
  - cidr: 10.96.64.192/26
    hostPrefix: 194
  networkType: OpenShiftSDN
  serviceNetwork:
  - 172.30.0.0/16
platform:
  none: {}
fips: false
pullSecret: 'contents of pull-secret file'
sshkey: 'contents of /home/core/.ssh/id_rsa.pub'
```



## 4. Create installation ignition files

- Execute `openshift-install create manifests --dir=/mnt`
- If defining worker guests, change manifests/cluster-scheduler-02-config.yml file
  - Set `mastersSchedulable` to False
- Execute `openshift-install create ignition-configs -dir=/mnt`
  - 3 ignition files created: bootstrap.ign, master.ign, worker.ign
- Enable access to files via http
  - enable world-readable for created files
    - `chmod +r *.ign`
    - `chmod +r -R auth/`
  - mount `/mnt mount-point --bind`
    - `/srv/www` in SLES
    - `/var/www/html` in RHEL
- Note: security certificates used in this process expire after 24 hours
  - Don't get too distracted



## Build “bootstrap” machine

- Determine proper parameters vs. “ignition” file
  - Parm file used to define hardware (FCP) and network, points to ignition file
  - Ignition file used to customize virtual machine
  - Review the doc, not everything you need is in one place
- Parmfile that worked:

```
rd.neednet=1 console=ttysclp0 !conddef zfcf.allow_lun_scan=0 \  
rd.znet=qeth,0.0.1000,0.0.1001,0.0.1002,layer2=1,portno=0 \  
ip=10.96.64.194::10.96.64.254:255.255.255.192:osboot.demo.lrsinc.org::none \  
nameserver=10.96.64.138 \  
coreos.inst.ignition_url=http://10.96.64.193:81/mnt/bootstrap.ign \  
coreos.live.rootfs_url=http://10.96.64.193:81/mnt/rhcos-live-rootfs.s390x.img \  
coreos.inst.install_dev=sda \  
rd.zfcf=0.0.2100,0x500507680d791bdb,0x0000000000000000 \  
rd.zfcf=0.0.2100,0x500507680d791bda,0x0000000000000000 \  
rd.zfcf=0.0.2200,0x500507680d751bda,0x0000000000000000 \  
rd.zfcf=0.0.2200,0x500507680d751bdb,0x0000000000000000
```
- Use of a separate CONFIG file to hold parameters in RHEL doesn’t work with RHCOS



## Build “master” machines

- Create a unique Parmfile for each machine:

```
rd.neednet=1 console=ttysclp0 coreos.inst=yes zfcplib.allow_lun_scan=0
rd.znet=qeth,0.0.1000,0.0.1001,0.0.1002,layer2=1,portno=0
ip=10.96.64.195::10.96.64.254:255.255.255.192:master0.demo.lrsinc.org::none
nameserver=10.96.64.138
coreos.inst.ignition_url=http://10.96.64.193:81/mnt/master.ign
coreos.live.rootfs_url=http://10.96.64.193:81/mnt/rhcos-live-rootfs.s390x.img
coreos.inst.install_dev=sda
rd.zfcplib=0.0.2100,0x500507680d791bdb,0x0000000000000000
rd.zfcplib=0.0.2100,0x500507680d791bda,0x0000000000000000
rd.zfcplib=0.0.2200,0x500507680d751bda,0x0000000000000000
rd.zfcplib=0.0.2200,0x500507680d751bdb,0x0000000000000000
```
- IP address is different for each master, other parameters stay the same
- Real FCP subchannels are unique for each v.m. but are defined with the same virtual subchannels for ease of definition
  - Be careful of WWPNs, since you need to enter target WWPNs that are visible to all relocation domains



## Build “worker” machines

- Create a unique Parmfile for each machine:

```
rd.neednet=1 console=ttysclp0 coreos.inst=yes zfcf.allow_lun_scan=0
rd.znet=qeth,0.0.1000,0.0.1001,0.0.1002,layer2=1,portno=0
ip=10.96.64.198::10.96.64.254:255.255.255.192:worker0.demo.lrsinc.org::none
nameserver=10.96.64.138
coreos.inst.ignition_url=http://10.96.64.193:81/mnt/worker.ign
coreos.live.rootfs_url=http://10.96.64.193:81/mnt/rhcos-live-rootfs.s390x.img
coreos.inst.install_dev=sda
rd.zfcf=0.0.2100,0x500507680d791bdb,0x0000000000000000
rd.zfcf=0.0.2100,0x500507680d791bda,0x0000000000000000
rd.zfcf=0.0.2200,0x500507680d751bda,0x0000000000000000
rd.zfcf=0.0.2200,0x500507680d751bdb,0x0000000000000000
```
- IP address is different for each worker, other parameters stay the same
- Real FCP subchannels are unique for each v.m. but are defined with the same virtual subchannels for ease of definition



## Add worker machines to the cluster

- New for OpenShift 4.6, the worker machines are not automatically added to the cluster
- Must issue commands to add them to the cluster
  - **oc get csr**
    - Note csr name associated with worker node
  - **oc set csr *csr-name* approve**
  - Wait a few minutes, then issue **oc get nodes** and see if the worker nodes are up and running



## Enable multipathing

- New for OpenShift 4.9, if you are using FB disks and predefined the paths during creation, you must enable the multipath daemon.
  - Updating /etc/multipath.conf doesn't work
  - Creating/updating /etc/zfcp.conf doesn't work
  - You must tell openshift to do it for you
- Build a YAML file named 99-master-kargs-mpath.yaml

```
apiVersion: machineconfiguration.openshift.io/v1
kind: MachineConfig
metadata:
  labels:
    machineconfiguration.openshift.io/role: "master"
  name: 99-master-kargs-mpath
spec:
  kernelArguments:
    - 'rd.multipath=default'
    - 'root=/dev/disk/by-label/dm-mpath-root'
```
- Worker nodes have a slightly different YAML file
  - Replace “master” in the YAML above with “worker”



## Manual Install Lessons Learned

- Couldn't login to RHCOS machines
  - Had to start ssh-agent from user "core" and load that agent every time I rebooted the DDNS/DHCP server:
    - `eval "$(ssh-agent -s)"`
    - `ssh-add path/filename of private key`
  - Then could ssh in; alternative was `ssh -i .ssh/id_rsab machine-IP-address`
  - Connected to machine and watched progress via command `journalctl -b -f -u bootkube.service`
    - Started to build master and worker machines once bootkube service settled down
    - My clue was to look for "SELinux: mount invalid" messages on the boot server after reboot
- To handle anonymous http, bind mount /mnt to /var/www/html/mnt on RHEL or /srv/www/htdocs/mnt on SLES
- Had to change \*.ign files to permissions 644 (added world-readable for anonymous http)
- Had to allow world-readable auth subdirectory to let userid "core" get the credentials



## Lessons Learned

- Using SCSI-disk means that IND USER won't show I/O counts
  - Harder to tell if you're stuck; had to use SCIF to monitor along with watching bootkube.service (which goes in fits and starts)
- SLES vs RHEL for DNS/Load balancer/Load source
  - haproxy (RHEL) vs. nginx (SLES), nginx didn't like to simply listen on ports 80/443
  - YaST DNS-Server dialog doesn't like "\*" in the DNS config file; had to manually edit the demo.lrsinc.org file
  - Ended up staying with RHEL due to issues with nginx for http/https
    - Probably could have gotten it to work with more knowledge of nginx



## Additional Configuration Files



## Additional File Settings

/etc/httpd/conf/httpd.conf

```
ServerRoot "/etc/httpd"
Listen 81
Include conf.modules.d/*.conf
User apache
Group apache
ServerAdmin root@localhost
ServerName osdhcpd.lrsinc.org:81
<Directory />
    AllowOverride none
    Require all denied
</Directory>
DocumentRoot "/var/www/html"
<Directory "/var/www">
    AllowOverride None
    # Allow open access:
    Require all granted
</Directory>

<Directory "/var/www/html">
    Options Indexes FollowSymLinks
    AllowOverride None
    Require all granted
</Directory>
<IfModule dir_module>
    DirectoryIndex index.html
</IfModule>
<Files ".ht*">
    Require all denied
</Files>
EnableSendfile on
```



## Additional File Settings

/etc/haproxy/haproxy.conf

```
global
    log                  127.0.0.1 local2
    chroot               /var/lib/haproxy
    pidfile              /var/run/haproxy.pid
    maxconn              4000
    user                 haproxy
    group                haproxy
    daemon
    stats socket /var/lib/haproxy/stats
    ssl-default-bind-ciphers PROFILE=SYSTEM
    ssl-default-server-ciphers PROFILE=SYSTEM
defaults
    mode                  http
    log                  global
    option                tcplog
    option                dontlognull
    option                redispatch
    retries               3
    timeout http-request 10s
    timeout queue        1m
    timeout connect     10s
    timeout client       1m
    timeout server       1m
    timeout http-keep-alive 10s
    timeout check        10s
    maxconn              3000

listen ingress-http
    bind *:80
    mode tcp
    balance source
    server worker0 10.96.64.199:80 check inter 10s
    server worker1 10.96.64.200:80 check inter 10s

listen ingress-https
    bind *:443
    mode tcp
    server master0 10.96.64.196:80 check inter 1s
    server master1 10.96.64.197:80 check inter 1s
    server master2 10.96.64.198:80 check inter 1s

listen api
    bind *:6443
    mode tcp
    option httpchk GET /readyz HTTP/1.0
    option log-health-checks
    balance roundrobin
    server master0 10.96.64.196:6443 weight 1 verify none check check-ssl
    server master1 10.96.64.197:6443 weight 1 verify none check check-ssl
    server master2 10.96.64.198:6443 weight 1 verify none check check-ssl

listen api-int
    bind *:22623
    mode tcp
    server master0 10.96.64.196:22623 check inter 10s
    server master1 10.96.64.197:22623 check inter 10s
    server master2 10.96.64.198:22623 check inter 10s
```