Cloud Native Security
CONFERENCE
NORTH AMERICA

Data Security and Storage Hardening In Rook and Ceph

Federico Lucifredi, Ana McTaggart, Michael Hackett

me! me! me!

Things I worked on

Red Hat Ceph Storage
Ubuntu Server
Landscape
SUSE Studio
SLES
SMT
Ximian Red Carpet
Man (I)

Rook in brief

Rook

- Cloud-native storage for k8s
- Ceph-based: hyperscale
- File, Block and Object
- Storage on top of compute: hyper-converge
- ...or optionally external storage
- Highly resilient
- Highly available
- Automated resource management w/operators

Threat Model

- Identify threat actors
  - Nation states
  - Organized crime
  - Hacker groups
  - Motivated individuals
  - Privileged insiders
  - Script kiddies
  - ...

Network Security zones

- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone

Network Security zones

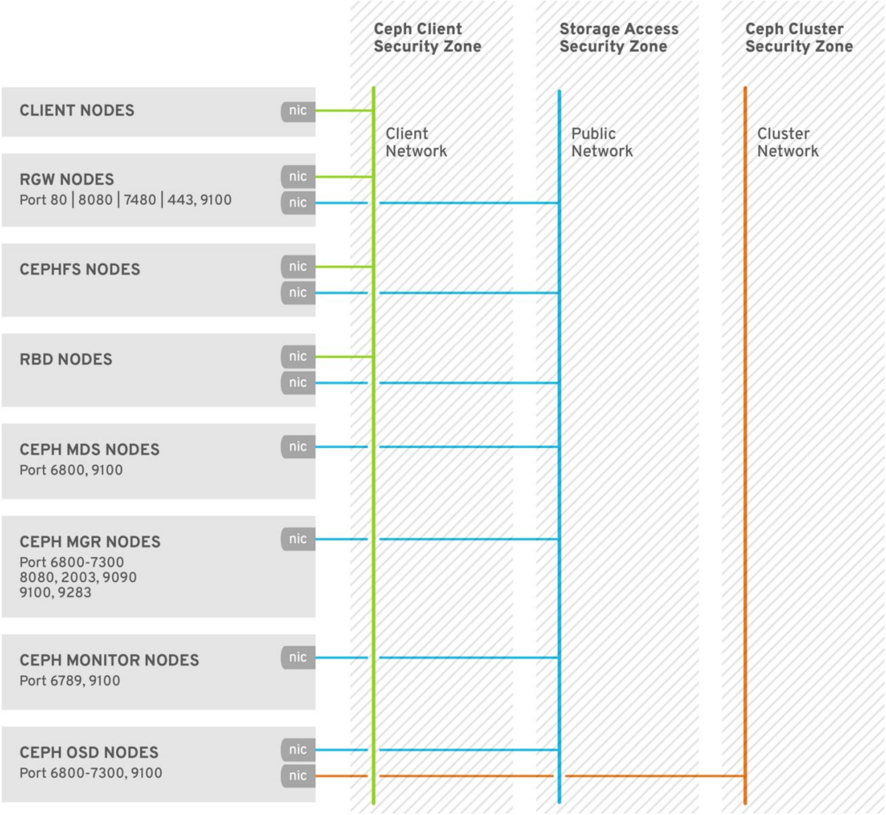
- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone

Network Security zones

- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone
  - cluster_network in Ceph

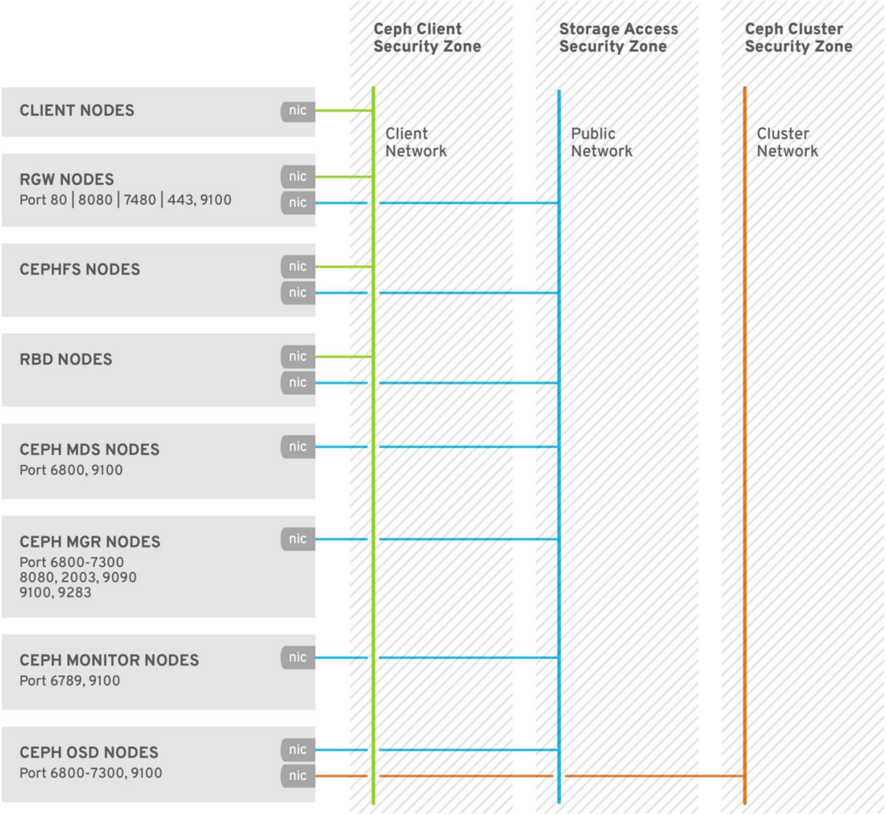
Connecting Security zones

- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone
  - cluster_network in Ceph



Connecting Security zones

- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone
  - cluster_network in Ceph



Encryption and Key Management

- Data at rest (OSD)
  - OSDs can be encrypted with dmcrypt at creation.
  - Write-ahead logs, journals and metadata stores can also be secured
  - LUKS provides a variety of cryptographic options
  - All data at rest is encrypted irrespective of access protocol
  - FIPS 140-2 certified cyphers can be used

- Encryption keys
  - Stored in the Monitor daemon (MON)

- Object Gateway (RGW)
  - Data is encrypted at rest relying on OSD strategy
  - Alternatively, data can be encrypted at ingestion with locally managed keys
  - Keys can be managed externally with HashiCorp Vault KMS
  - OpenStack Barbican and KMIP-compatible KMS support is also available

HashiCorp
Vault

LUKS

Encryption in transit

- Data in transit
  - Ceph’s internal protocol can be encrypted as a Messenger v.2.1 protocol option
  - Legacy cleartext protocol is still default for compatibility reasons
  - All data at rest is encrypted irrespective of access protocol
  - FIPS 140-2 certified cyphers can be used
- Client and public security zones
  - TLS security can be used from Object Gateway to S3 clients.
  - TLS termination at HAproxy a special case
- Network hygiene
  - Firewalld at individual nodes

Encryption in transit

- Data in transit
  - Ceph’s internal protocol can be encrypted as a Messenger v.2.1 protocol option
  - Legacy cleartext protocol is still default for compatibility reasons
  - All data at rest is encrypted irrespective of access protocol
  - FIPS 140-2 certified cyphers can be used
- Client and public security zones
  - TLS security can be used from Object Gateway to S3 clients.
  - TLS termination at HAproxy a special case
- Network hygiene
  - Firewalld at individual nodes

Rook specific

- CRDs can be used to encode security preferences
  - Example: client configuration
  - Example: RGW certificate
- Rook provides at-rest data encryption as discussed
  - Setup of Msgr v.2 in-flight encryption is still to come
  - Use software-defined cloud network fabric to segregate traffic
- Standard k8s user permissions apply to persistent volumes
  - Nothing Rook needs to do here
- CSI driver supports KMS
  - PVs can be encrypted with individual keys

Control Plane

- SSH
  - Cephadm, ceph-ansible and other tools
  - User (cephadm or ceph) with password-less root access is used
  - Access is secured with SSH keys
  - Port 22
- Management Dashboard
  - TLS on port 443 (operator facing (storage access zone)
  - Dashboard access zone often tailored by operators to suit local threat model
- Manager (MGR)
  - Ceph protocol on port range 6800-7300 (storage access zone)

Identity and access

- Cephx
  - Shared secret keys are in use for authentication
  - Mechanism protects cluster from MITM attacks
  - Authentication and authorization are aon by default
    - If user is not supplied, it is assumed to be client.admin

- Object Gateway (RGW)
  - S3 user: access key and secret model
  - Swift user: access key and secret model
    - Note that default Swift user is sub-user of S3 user, deleting S3 user will delete the Swift user as well
  - Administrative user: access key and secret with access to administrative API
  - User authentication is stored in Ceph pools

- LDAP and Active Directory users can be used as identity services
  - Secure LDAP is recommended

- OpenStack Keystone
  - Ceph supports using OpenStack Keystone to authenticate Object Gateway users

Auditing

- Operator actions
  - Stored in /var/log/ceph/ceph.audit.log

  For example:

```
2018-08-13 21:50:28.727176 mon.reesi001 mon.0 172.21.2.201:6789/0 2097902 :
audit [INF] from='client.348389421 -' entity='client.admin' cmd=[{"prefix": "osd set",
"key": "nodown"}]: dispatch
```

```
2018-08-13 21:50:28.872992 mon.reesi001 mon.0 172.21.2.201:6789/0 2097904 :
audit [INF] from='client.348389421 -' entity='client.admin' cmd='[{"prefix": "osd set",
"key": "nodown"}]': finished
```

  - In distributed systems, actions may start on one node (dispatch) and propagate to others (finished)

Data retention

- RADOS
  - End users generally do not have the ability to read, write or delete objects directly in a storage pool

- Ceph Block Device (RBD), Object Gateway (RGW), Filesystem (MDS)
  - Users can create, delete, modify volume images, objects or files
  - Deletion destroys corresponding RADOS object in unrecoverable manner
    - RBD pools may provide “trash bin” functionality with spare capacity
    - RGW bucket lifecycle supports versioning. Residual data artefacts may persist in storage medium

- Secure deletion
  - Sanitize retired media by encrypting the OSD contents at rest, and replacing the encryption key

Infrastructure hardening

- SELinux
  - Red Hat Ceph storage clusters default to SELinux in enforcing mode
- FIPS 140-2 support
  - Certified cryptography can be imported in RHEL “FIPS mode” setup
  - RHEL 8.2 is the most recent certified version
- Hardened binaries
  - SECCOMP
  - PIE
  - -D_FORTIFY_SOURCE=2
  - RELRO
  - BIND_NOW
  - ASLR (all varieties)
  - ...

Thank you!

CREDITS

Federico Lucifredi
Ana McTaggart
Michael Hackett
J.C. Lopez
Travis Nielsen
Sébastien Han

© 2021 by the authors
CC-BY-SA

Cloud Native Security
CONFERENCE
NORTH AMERICA

Cloud Native Security
CONFERENCE
NORTH AMERICA