Security + Improvements in Ceph and Rook

Federico Lucifredi, Sage McTaggart – IBM

me! me! me!

Things I worked on

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

me! me! me!

Things I worked on

Ceph Storage
OpenShift Data Foundation
Incident Response
Computer Security
Theoretical and practical

me! me! me!

Things I worked on

Red Hat Ceph Storage
Hero of Ceph support
Growing tomatoes
...

FOSDEM

CEPH IN BRIEF

Ceph

- Still the future of storage!
- File, block and object storage
- Highly resilient
- No single-point-of-failure
- Highly available
- Scale out

...absolutely awesome.

ROOK IN BRIEF

Rook is a community-initiated project that is:

- Cloud-native storage for k8s
- Ceph-based: hyperscale
- Storage on top of compute: hyper-converge…

  …or optionally external storage

- Automated resource management w/operators
- Automated upgrade and rollback

Why do we care about storage security?

- Data is expensive to back up.
- Even more expensive to collect
- Breaches, and the resultant lawsuits, are not fun. Especially for PPI
- Can wreck customer trust even after the fact
    - See Last Pass
- The scenario of Ransomware on your storage hard drive is hard to protect against
    - Knowing how to keep systems up to date is easy to not be low hanging fruit

Today’s talk

- We will go over some recent improvements
- We will also go over how we secure Ceph by design
- Agenda
  - Thread modeling
  - Network Security Zones and Dashboard improvements
  - Open Source PSIRT with IBM
  - EO 14028 reflections
  - Encryption and key management
  - Control Plane
  - Cryptography, identity and access
  - Data Retention and Hardening choices

THREAT MODEL

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

THREAT MODEL

- Moving beyond compliance
  - Thinking of threats to software
  - Beyond a simple inventory and hardening guide
  - Underlies our principles with secure design
- Consistency with existing protocols
  - Cryptographic algorithms should follow community standards-we had a recent case with AES implementation and decided to be consistent with Kerberos' model
  - Developers' model is that we are not as good at crypto as cryptographers

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

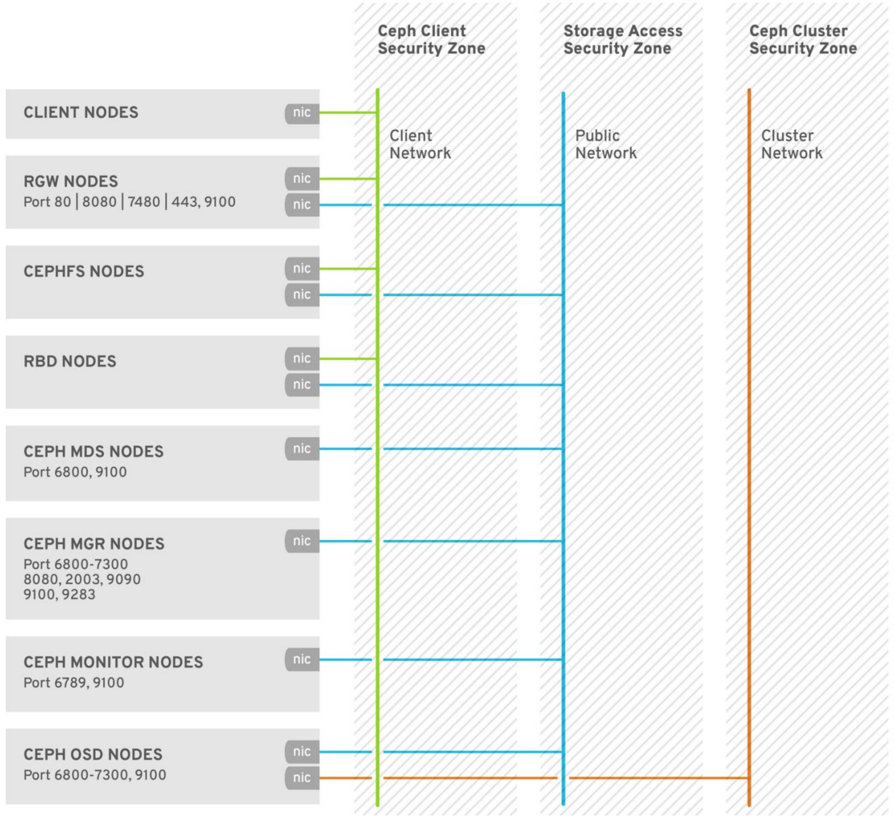
- Public Zone
  - **not** the public_network in Ceph
- Ceph Client Zone
- Storage Access Zone
  - public_network in Ceph
- Ceph Cluster zone

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

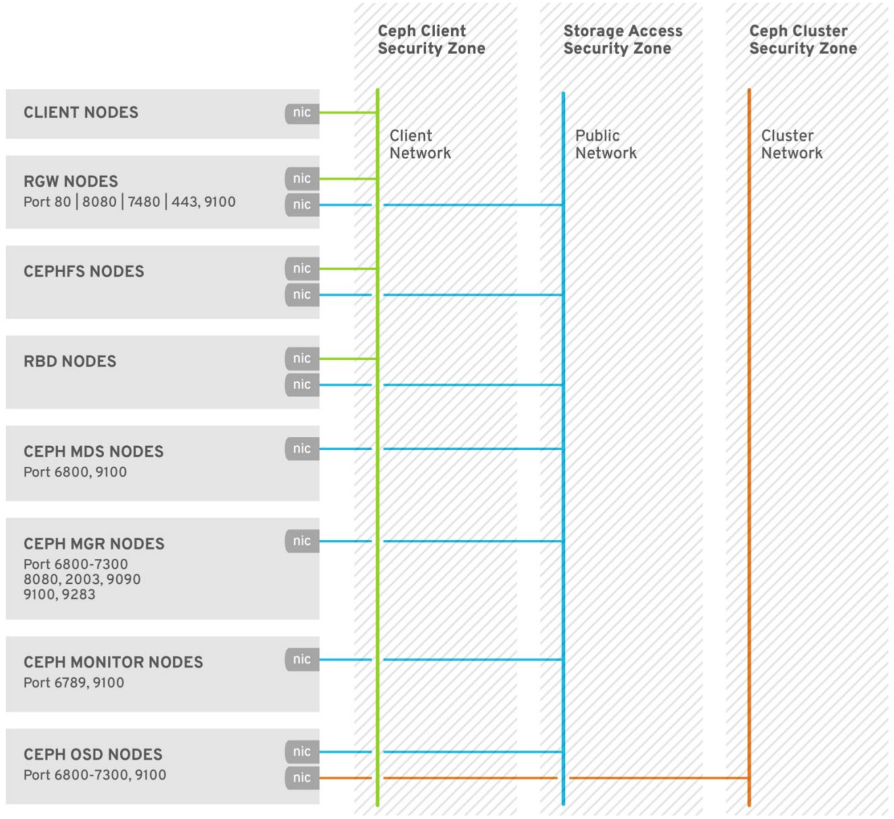
CONNECTING SECURITY ZONES

- 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



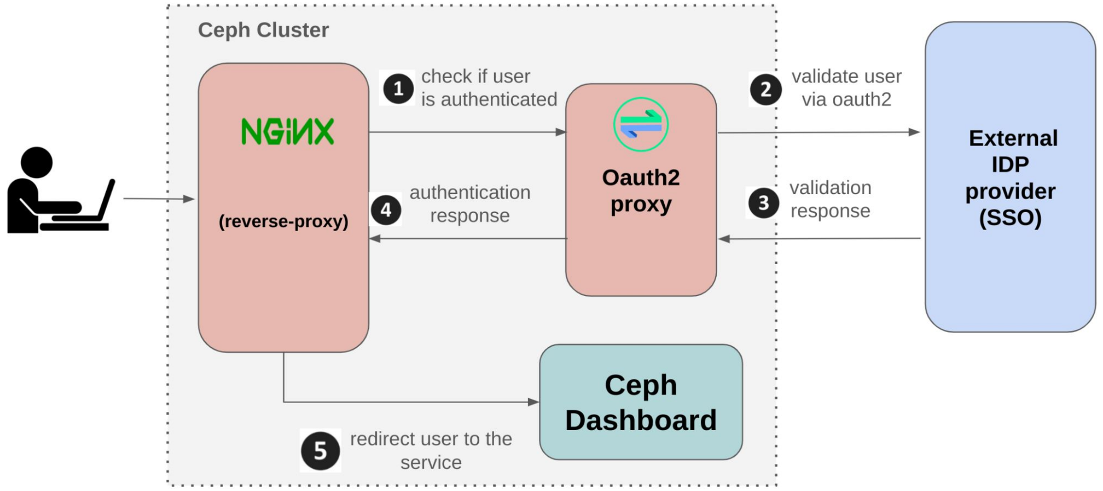
CONNECTING SECURITY ZONES

- 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



Dashboard improvements

All webhooks now covered by an authentication layer. Ensures Dashboard monitoring is within Ceph Cluster Zone. Secure by design, better for multi-tenant


Ceph Cluster
NGINX
(reverse-proxy)
1 check if user is authenticated
Oauth2 proxy
2 validate user via oauth2
External IDP provider (SSO)
3 validation response
4 authentication response
Ceph Dashboard
5 redirect user to the service

Moving companies as an OSS product

- The security of a product is good, but what about the support of a product?
  - An abandoned product can be forked, but do you really want to maintain it and patch it?
- We moved to IBM approximately three years ago from Red Hat
  - Definitely a change, but overall going well
- I will discuss some aspects of IBM Product Security, and how that's been going, with a focus on PSIRT and compliance as open source!
- I will also discuss some of our accomplishments, and what we are planning going forward in the next year

PRODUCT SECURITY

- Product Security at IBM performs SDL activities across the life-cycle of each release
  - Goal to reduce risk and improve the security of Ceph
  - Actively creating better threat models, iterating, etc
  - More regular Pen Tests, with edits made based on them
  - Can’t release with known criticals without significant mitigation
- Still working with upstream on secure design and with individual engineering teams to focus on key improvements
  - Have also worked with IBM for things like Call Home
- Worked with IBM research to investigate how to implement TCMs and where quantum resistant algorithms will go.

IBM PSIRT

- We have manifested and documented all dependencies within both IBM and RH's PSIRTs
- Automated many security scans of our code
  - Implemented scans of all containers
  - Working with upstream communities to ensure fixes are made for all projects
    - Can feel like we've just got to update all dependencies…
- We are working on automatically updating all/most dependencies
  - Updating Grafana with every release, as one of our problem libraries
    - Hard to automate, ran into major build issues
    - Human-automaton of regular processes helps, less mental overhead
  - Improved security of dashboard with re-design, to reduce exposure.

EO 14028

- Scanners often only detect CVEs from dependencies
  - Discourages many people from using open source
  - Updated dependencies fix this
- Emphasis on scanning prevents differentiation between vulnerable and affected
  - Have historically focused on vulnerable
  - Focusing redesign on vulnerable, to reduce attack surface, especially with exploits
  - VEX scores can be helpful here
  - Updating a dependency isn’t that hard tho! Reduces burden later
- By knowing our vulnerabilities, we are able to redesign and fix them
  - Call home was not open source, and the security model originally required closed source code. We were able to use our incident response skills to create a more secure one, advocating for open source in the process!
  - This just brings to light the issues with outdated and vulnerable code! Great ultimately for OS

EO 14028

- As part of the 2021 EO by President Biden, we had to manifest and document all dependencies, and document all known vulns (CVEs)
  - We are actively fixing all now, and were able to get to an attestation ready state this summer, with all CVEs documented and/or on track for fixes at IBM
- The only way to achieve this as a routine practice is to automatically update dependencies, in upstream and downstream
  - Had to implement container scanning with tools that were not designed for it
  - Had to redesign upstream CVE reporting to account for limited capacity and timelines
- Why does this matter for Open Source? Government customers need this, and corporate environments do not want to use poorly maintained software. No one should

IBM PSIRT

- We will continue to review existing vulns, and release regular security updates, and improve code security preemptively
  - Have gotten more upstream engagement, hundreds of vulns, within past year
  - Have implemented changes, e.g, to container updates, based on requests
- We will eventually follow IBM standards to fix vulns and ensure compliance
  - Currently we’re fixing all bugs, and aim to be fully within IBM SLA/regs by EOY

Automation?

- Historically, much of our security work has been a manual, in depth, process
  - This is very difficult for a very small security and build team
- What does automation look like?
  - Automatically bumping versions
    - Causes build issues, especially if in a major release
  - Building it into processes automatically
    - What we’re doing, offers less of a burden to engineers
    - Challenges are transitive dependencies (do we fix golang in grafana?)
    - Can miss key steps
    - Best if implemented with only worst case libraries

PRODUCT SECURITY INCIDENT RESPONSE

- Responds to incidents and all reported vulnerabilities
- Triages to release fixes in a timely manner
- Works with upstream to review flaws
- Reviews all minor releases, coordinates publication dates with engineering and upstream
- Coordinates release when exploits, not just vulnerabilities, are discovered

PRODUCT SECURITY ARCHITECT

- Ensures correct manifesting, we know (all of) what is in our code and (mostly) where it is used
- Reviews major releases, approves any exceptions to security policy
- Works with compliance team to validate security policy, federal standards
- Works to suggest design improvements prior to release
    - Cryptographic algorithms, in particular, improvements to msgr2
    - Design choices, key storage improvements, RGW design

https://docs.google.com/document/d/1U0i_MYu9fVuTH7Q2sjI82oa5vcZGefNJn1S02z3QEpk/edit?tab=t.0#heading=h.wrf39576czg0

NEW WORK AT IBM

- All of the CVE fixes will be continue to be ported to upstream Ceph. We aim to keep all versions of Ceph, be that Red Hat, IBM or Upstream, equally secure
- New collaboration produces new challenges, and lots of goals. Going well!
- Worked to improve Call Home functionality and security for IBM support by applying Open Source principles
    - If it stinks, other people will smell it. Visible bugs are a boone
- Working on use as backend for AI,
- Reach out to talk about what you want to see with our collaboration with IBM, and feel free to discuss any concerns with us

FOSDEM

ENCRYPTION AND KEY MANAGEMENT

- Data at rest (OSD)
  - OSDs can be encrypted with dm-crypt 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
  - Working on implementing post quantum cyphers as they are certified
  - Stored in the Monitor daemon (MON) – or KMS, or K8s Secrets
- Object Gateway (RGW)
  - Data is encrypted at rest relying on OSD strategy
  - Alternatively, data can be encrypted at ingestion with client managed keys
  - Keys can be managed externally with HashiCorp Vault KMS
  - OpenStack Barbican and KMIP-compatible KMS support is also available

HashiCorp Vault

LUKS

FOSDEM

ENCRYPTION IN TRANSIT

- Data in transit
  - Ceph’s internal protocol can be encrypted as a Messenger v.2.1 protocol option
  - Working on potential re-design for quantum resistance
  - 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 between 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
  - Working on potential re-design for quantum resistance
  - 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 between 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
  - Allows principle of least privilege to easily be implemented
- Rook provides at-rest data encryption as discussed
  - Setup of Msgr v2 in-flight encryption exists as of v1.9
  - Use software-defined cloud network fabric to segregate traffic
- Standard k8s user permissions apply to persistent volumes claims
  - Nothing Rook needs to do here
- CSI driver supports KMS
  - PVs can be encrypted with individual keys, limiting key scope

CONTROL PLANE

- SSH
  - Cephadm, ceph-ansible and other tools
  - User (cephadm or ceph) with password-less root access can be used
  - Access is secured with SSH keys
  - Port 22
- Management Dashboard
  - HTTPs access from storage access zone
  - Storage access zone often tailored by operators to suit local threat model, default now (for dashboard metrics) is in Ceph Cluster Zone
  - Option for SAML/SSO authorization, Kubernetes native authorization
- Manager (MGR)
  - Ceph protocol on port range 6800-7300 (storage access zone)

New Cryptography Goals

- Working with IBM research, discussing quantum and confidential computing
    - Currently determining viability, but has been productive to see where we use cryptography
    - Open source research code implemented into an open source storage system is the goal!
- Documenting all places we use cryptography and if it is quantum safe(er)
    - Via architectural diagrams, working to see best places for TCMs
- Priority is to have all work on this be open source and use open source libraries here
- Working with upstream to discuss correct use of cryptographic modules such as AES.
    - Active discussions on Kerberos Model and best ways to avoid threats.
    - Ties in with need to upgrade deps automatically, this applies to cryptography too!

IDENTITY AND ACCESS

- Cephx
  - Shared secret keys are in use for authentication
  - Mechanism protects cluster from MITM attacks
  - Authentication and authorization are on by default
    - If user is not supplied, provide client.admin as user, and restricted accordingly
- Object Gateway (RGW)
  - S3 user: access key and secret model, option for bucket policy
  - 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, continually improving our process here
    - Identity is an IAM API, consistent and secure
    - Token based authentication with STS API
  - Can couple with OIDC providers (Keycloak, etc), backed by organizational IdP (FreeIPA) for granular role or attribute access with SSO

IDENTITY AND ACCESS

- LDAP and Active Directory users can be used as identity services for the dashboard
    - Secure LDAP is highly recommended
- Actively supporting modern SSOs in recent versions of Ceph
- 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 destroying 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 8.4 is the most recent certified version
- Hardened binaries
  - `-D_FORTIFY_SOURCE=2`
  - `-D_GLIBCXX_ASSERTIONS`
  - `-fstack-protector-strong`
  - `-fcf-protection`
  - `-fstack-clash-protection`
- Additional Kernel or OS-supplied hardening
  - `SECCOMP`
  - `PIE`
  - `RELRO`
  - `BIND_NOW`
  - `ASLR (all varieties)`

FOSDEM

Thank you!

RESOURCES

- Managing and Securing Kubernetes Secrets
  - Rani Osnat - Aquia Security
- Hacking Kubernetes – chapter 6: Storage
  - Andrew Martin and Michael Hausenblas (O'Reilly)
- Data Security and Hardening Guide
  - Red Hat Ceph Storage 5 documentation
- Encrypting Secret Data at Rest
  - Kubernetes documentation
- Recommended Compiler and Linker flags for GCC
  - Survey of Kernel and Userspace hardening options

RESOURCES

- Ubuntu Kernel and Userspace Security Features
  - Ubuntu Security Team - Canonical
- Production-ready Ceph Object Gateway with New Caphadm Features
  - Daniel Parkes - IBM

CREDITS

Federico Lucifredi
Sage McTaggart
Michael Hackett
John Wilkins
J.C. Lopez
Travis Nielsen
Sébastien Han
Ken Dreyer
Kyle Bader
Gabriella Roman
Chris Blum