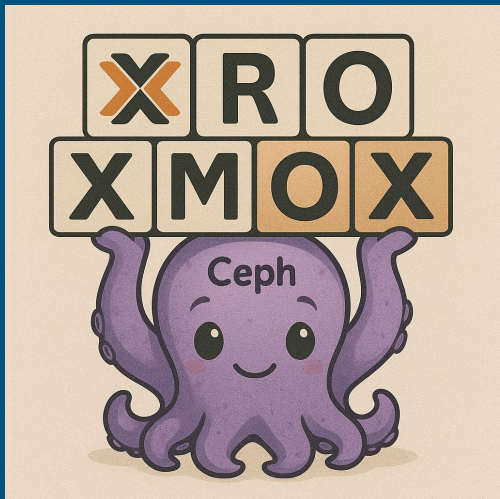# Ceph in Proxmox Virtual Environment

Alex Gorbachev
Intelligent Systems Services, Inc.
alex@iss-integration.com

Who the Ceph am I?

Started ISS in 1994, many large projects in firm's experience deploying and operating large federal, commercial, and nonprofit environments

Looked for "Apache" of SAN in 2013, found Sage/Ceph/Inktank

Built a lab in 2014 to test Ceph at SSD speeds, and at midrange scale

Integrated Ceph into high end mission critical solutions using ISS Storcium brand name

Certified Ceph with iSCSI/NFS for VMWare in 2016, still on HCL

Got disillusioned with T10 based protocols

Proxmox to the Rescue

Open Source Hypervisor solution (KVM Based), started 2005

No-Subscription Repository is free as in completely

Debian platform using Proxmox kernel and distro

Mature VM/LXC and clustering platform, using Corosync as messenger and pmxfs as Cluster sysfs

Integrates their own Ceph orchestrator, and keeps up with versions

Business model is subscription and partner based (ISS is a US Gold Partner)

Proxmox Ceph Integration

Simple and Practical, follow wiki and forums (lots of youtube videos as well)

RBD and CephFS available directly from UI

External Ceph clusters connect via keyring and caps

Small labs can be set up with mini PCs and removable drives

VM Snapshots integrated with Ceph

Backup Server (PBS) supports QEMU dirty bit mapping

Documentation | Create VM | Create CT | root@pam

Server View

Datacenter

Help

- Datacenter (rockgarden)
  - rock1
  - rock2
  - rock3
  - rock4

- Search
- Summary
- Notes
- Cluster
- Ceph
- Options
- Storage
- Backup
- Replication
- Permissions
  - Users
  - API Tokens
  - Two Factor
  - Groups
  - Pools
  - Roles
  - Realms
- HA
- SDN
  - Zones
  - VNets
  - Options

Health

Status	Nodes		Ceph
✓	✓ Online	4	✓
	✗ Offline	0	
Cluster: rockgarden, Quorate: Yes			HEALTH_OK

Guests

Virtual Machines		LXC Container	
● Running	6	● Running	0
○ Stopped	19	○ Stopped	0
○ Templates	3		

Resources

CPU	Memory	Storage
1%	11%	4%
of 128 CPU(s)	111.43 GiB of 1006.16 GiB	2.57 TiB of 70.07 TiB

Nodes

Name	ID	Online	Support	Server Address	CPU usage	Memory usage	Uptime
rock1	1	✓	-	192.168.60.106	1%	8%	28 days 01...
rock2	2	✓	-	192.168.60.107	1%	11%	32 days 22...
rock3	3	✓	-	192.168.60.108	1%	15%	28 days 01...
rock4	4	✓	-	192.168.60.109	1%	10%	32 days 23...

Subscriptions

No Subscription

✗

Tasks	Cluster log

Start Time ↓	End Time	Node	User name	Description	Status	
May 12 05:10:41	May 12 05:10:46	rock4	root@pam	Update package database	OK	>
May 12 04:56:04	May 12 04:56:08	rock2	root@pam	Update package database	OK	>

Documentation | Create VM | Create CT | root@pam

Server View

Node 'rock1'   Reboot | Shutdown | >_ Shell | BulkActions | Help

- Datacenter (rockgarden)
  - rock1
  - rock2
  - rock3
  - rock4

Search
Summary
Notes
>_ Shell
⚙ System
  ⇄ Network
  ✳ Certificates
  ◉ DNS
  ◉ Hosts
  ⚙ Options
  ◷ Time
  ☰ System Log
  ⟳ Updates
  ⧉ Repositories
  🛡 Firewall
  ⊟ Disks
  ▪ LVM
  ☐ LVM-Thin
  ▪ Directory
  ▦ ZFS
  ◉ Ceph
  ⚙ Configuration

Status

✓

HEALTH_OK

Ceph Version: 17.2.8

		Summary ↑	
		No Warnings/Errors	

OSDs

	● In	○ Out
⊕ Up	28	0
⊖ Down	0	0

Total: 28

PGs

● active+clean: 417

Services

Monitors			Managers			Metadata Servers		
rock1: ✔	rock2: ✔	rock3: ✔	rock1: ✔	**rock2:** ✔	rock3: ✔	rock1: ✔	**rock2:** ✔	rock3: ✔

Performance

Reads: 0 B/s

Usage

Writes: 21.40 KiB/s

Tasks | Cluster log

Start Time ↓	End Time	Node	User name	Description	Status	
May 12 05:10:41	May 12 05:10:46	rock4	root@pam	Update package database	OK	>
May 12 04:56:04	May 12 04:56:08	rock2	root@pam	Update package database	OK	>

Server View

Node 'rock1'   🔄 Reboot ⏻ Shutdown >_ Shell ⋮ Bulk Actions ❓ Help

- Datacenter (rockgarden)
  - rock1
  - rock2
  - rock3
  - rock4

System
- ⇄ Network
- ✱ Certificates
- 🌐 DNS
- 🌐 Hosts
- ⚙ Options
- ⏲ Time
- System Log
- ⟳ Updates
- Repositories
- 🛡 Firewall
- 💾 Disks
  - LVM
  - LVM-Thin
  - Directory
  - ZFS
- Ceph
  - ⚙ Configuration
  - 🖥 Monitor
  - OSD
  - CephFS
  - Pools

🔄 Reload   Create: OSD   Manage Global Flags   No OSD selected ⓘ Details ▶ Start ■ Stop 🔄 Restart ○ Out ⬛ In ≡ More

Name	Class	OSD Type	Status	Version	weight	reweight	Used (%)	Total	Apply/Commit Latency (ms)	PGs
default				17.2.8						
rock4				17.2.8						
osd.26	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.84	3.49 TiB	0 / 0	62
osd.23	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.94	3.49 TiB	0 / 0	61
osd.21	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.78	3.49 TiB	0 / 0	53
osd.27	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	8.94	1.09 TiB	0 / 0	36
osd.25	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	8.45	1.09 TiB	0 / 0	28
osd.24	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	6.72	1.09 TiB	0 / 0	27
osd.22	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	8.19	1.09 TiB	0 / 0	35
rock3				17.2.8						
osd.16	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.97	3.49 TiB	0 / 0	65
osd.15	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.84	3.49 TiB	0 / 0	60
osd.14	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.91	3.49 TiB	0 / 0	62
osd.20	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	7.47	1.09 TiB	0 / 0	29
osd.19	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	7.51	1.09 TiB	0 / 0	30
osd.18	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	9.18	1.09 TiB	0 / 0	35
osd.17	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	9.54	1.09 TiB	0 / 0	35
rock2				17.2.8						
osd.9	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.96	3.49 TiB	0 / 0	62
osd.8	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.88	3.49 TiB	0 / 0	63
osd.7	nvme4t	bluestore	up / in	17.2.8	3.49309	1.00	0.84	3.49 TiB	0 / 0	58
osd.13	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	7.47	1.09 TiB	0 / 0	30
osd.12	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	9.16	1.09 TiB	0 / 0	35
osd.11	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	9.56	1.09 TiB	0 / 0	37
osd.10	nvme1t	bluestore	up / in	17.2.8	1.0916	1.00	8.85	1.09 TiB	0 / 0	34

Tasks   Cluster log

Start Time ↓	End Time	Node	User name	Description	Status	
May 12 05:10:41	May 12 05:10:46	rock4	root@pam	Update package database	OK	>
May 12 04:56:04	May 12 04:56:08	rock2	root@pam	Update package database	OK	>
May 12 04:22:56	May 12 04:23:01	rock1	root@pam	Update package database	OK	>
May 12 01:35:05	May 12 01:35:10	rock3	root@pam	Update package database	OK	>

Server View

- Datacenter (rockgarden)
  - rock1
  - rock2
  - rock3
  - rock4

Node 'rock1'

Reboot | Shutdown | >_ Shell | BulkActions | Help

System
- Network
- Certificates
- DNS
- Hosts
- Options
- Time
- System Log
- Updates
- Repositories
- Firewall
- Disks
- LVM
- LVM-Thin
- Directory
- ZFS
- Ceph
- Configuration
- Monitor
- OSD
- CephFS
- Pools

Create | Edit | Destroy

Pool #	Name	Size/min	# of Placement Groups	Optimal # of PGs	Autoscaler Mode	CRUSH Rule (ID)	Used (%)
1	.mgr	3/2	1	n/a	on	nvme1t_rule (1)	87.39 MiB (0.00%)
2	nvme1t	3/2	128	n/a	off	nvme1t_rule (1)	1.44 TiB (8.81%)
3	nvme4t	3/2	128	32	off	nvme4t_rule (2)	329.89 GiB (0.81%)
4	kubfs1_data	3/1	128	n/a	on	replicated_rule (0)	52.55 GiB (0.10%)
5	kubfs1_metadata	3/2	32	n/a	on	replicated_rule (0)	1.45 GiB (0.00%)
							1.81 TiB

Tasks | Cluster log

Start Time ↓	End Time	Node	User name	Description	Status	
May 12 05:10:41	May 12 05:10:46	rock4	root@pam	Update package database	OK	>
May 12 04:56:04	May 12 04:56:08	rock2	root@pam	Update package database	OK	>
May 12 04:22:56	May 12 04:23:01	rock1	root@pam	Update package database	OK	>

References

https://pve.proxmox.com/wiki/Storage:_RBD

https://pve.proxmox.com/wiki/Storage:_CephFS

https://iss-integration.com#proxmox

https://www.proxmox.com/en/products/proxmox-virtual-environment/get-started

ask anything: alex@iss-integration.com