

The multi-release, multi-project road to multi-attach volumes

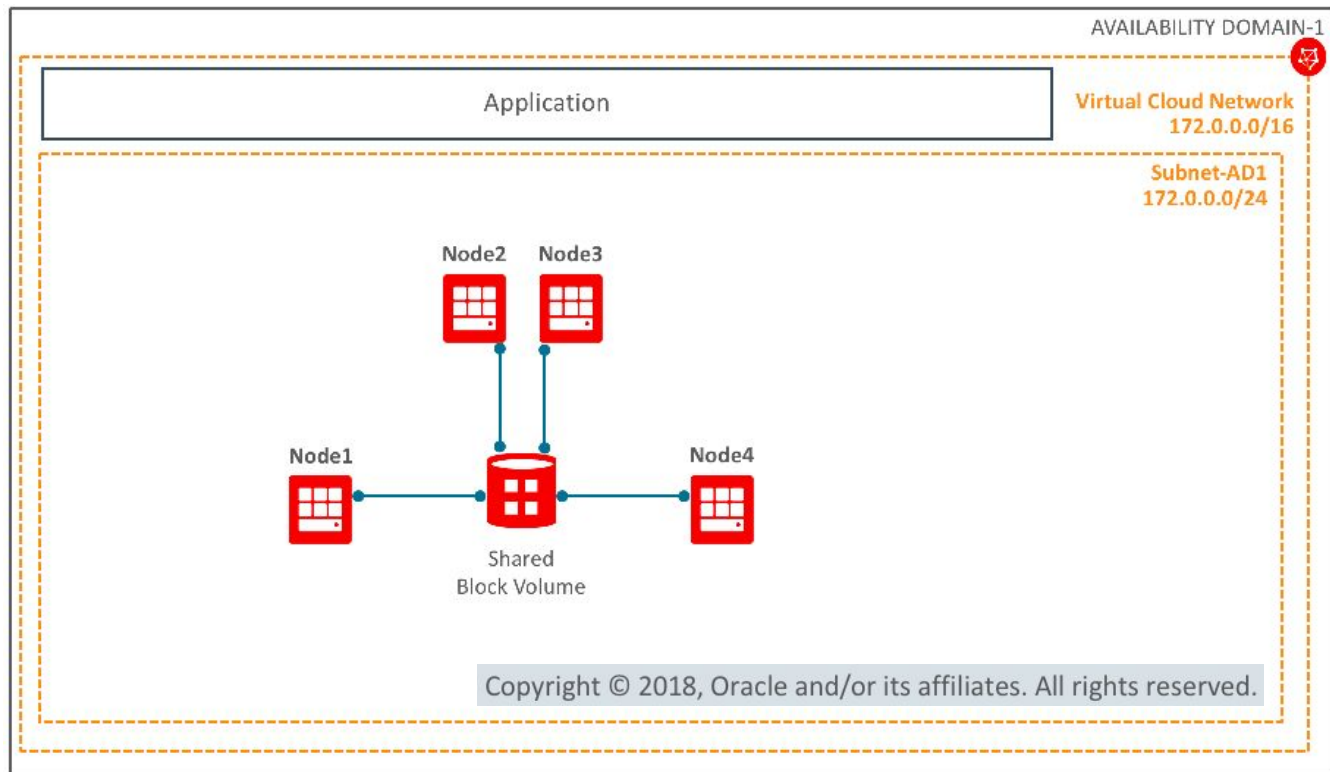
What took you so long!?!


Some of you have been waiting a while



Remind me what this is and why I wanted it?

- Oracle RAC
- High Availability
- Clustered FS



Some history and background

- **Grizzly**: Cinder proposal
- **Kilo**: Cinder merged volume.multiattach API change
- **Mitaka**: Nova/Cinder discussions about integration
- **Newton**: new Cinder API ideas
- **Ocata**: coming together; Cinder 3.27 API
- **Pike**: swimming in the same direction; working out edge cases
- **Queens**: delivered; Nova 2.60 API

Technical Debt and Stalling Efforts

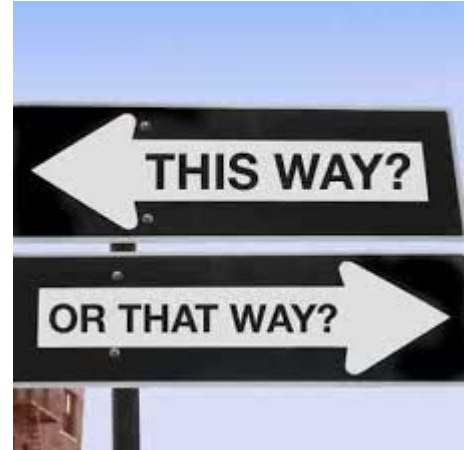
- Remember nova-volume?
 - Split brain
 - Technical debt
 - Tight coupling
- Bolt-ons are bad
- Need to invest in decoupled APIs



Turned into MUCH more than just “implement multi-attach”

We **joke about** this effort **taking “years”**, honestly though the work up until the last year was what happens when you **try** and **do** a multi-project change **without agreement/help** from the other projects.

The timer was reset in Barcelona (Ocata design summit).



The Hard parts

- Getting people to work on and review it
- Finding all the nooks and crannies
- Supporting existing operators
- Backward compatibility
- Smooth transition
- (Mistakenly) writing Cinder APIs in isolation

Forward and Backward compatibility is hard

- Microversions
- Online data migrations



Non disruptive API changes are really hard

- Attach / detach
- Boot from volume
- Swap volume (retype)
- Shelve
- Rebuild
- Live migrate
- Cold migrate

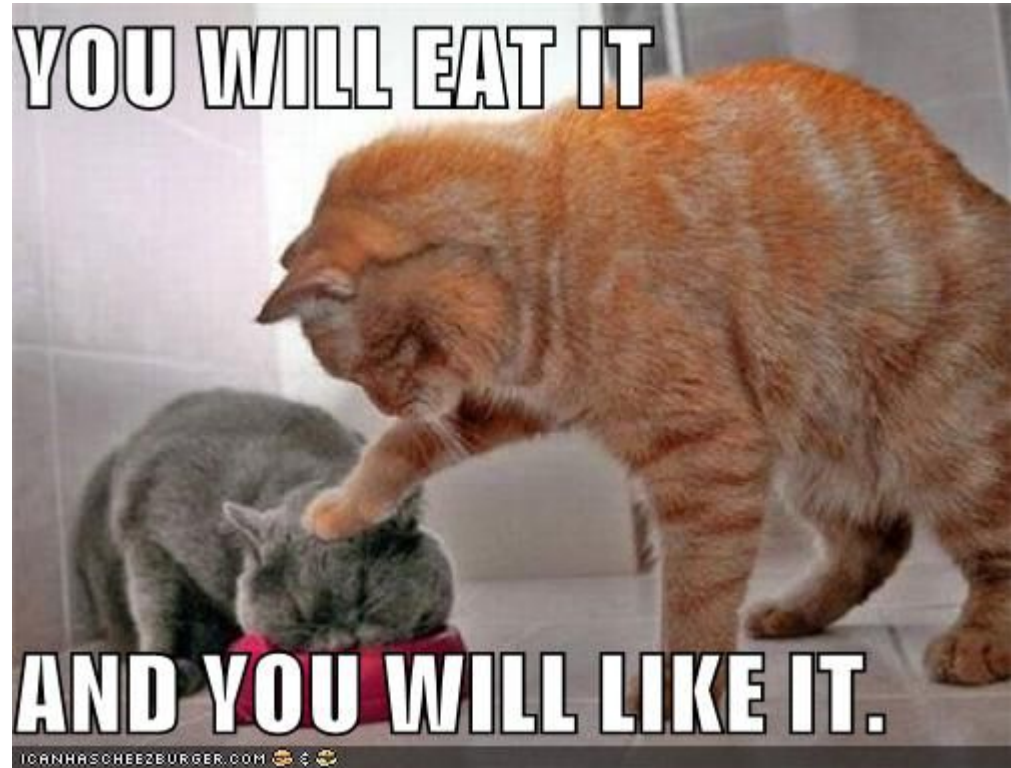
Figuring out all the obscure little things that have been added in over the years is hard

- Split brain
 - `check_attach()`
- `BlockDeviceMapping.connection_info`
 - Lemme just stash this right....here
- “shared_targets” Cinder 3.48 API
 - Volume driver differences



Community can be HARD

- Democracy vs dictatorship
- Slow but powerful



Summary of things we learned that we can apply to “next time”

- Big efforts need a champion
- QA is key
- Regular interlocks



So.... where did we end up?

- Nova libvirt driver (libvirt \geq 3.10)
- Cinder LVM, NetApp/SolidFire, Oracle ZFSSA drivers
- Cinder volume attachment APIs
- No retype to/from multiattach with an in-use volume
- No multiattach to shelved offloaded servers
- Cinder RBAC policy rules
- Read/write attachments only

Is it done? Is there more to come in the future?

- More driver support
- Compute API changes
 - Boot multiple servers from the same volume at the same time
 - Attach volumes in read-only mode
- Migrate old attachment records
- [Forum session Tuesday at 11:50am Room 221](#)

More info

- Demo: <https://www.youtube.com/watch?v=hZg6wqxdEHk>
- Docs
 - <https://docs.openstack.org/cinder/latest/admin/blockstorage-volume-multiattach.html>
 - <https://docs.openstack.org/nova/latest/admin/manage-volumes.html#volume-multi-attach>

Q & A