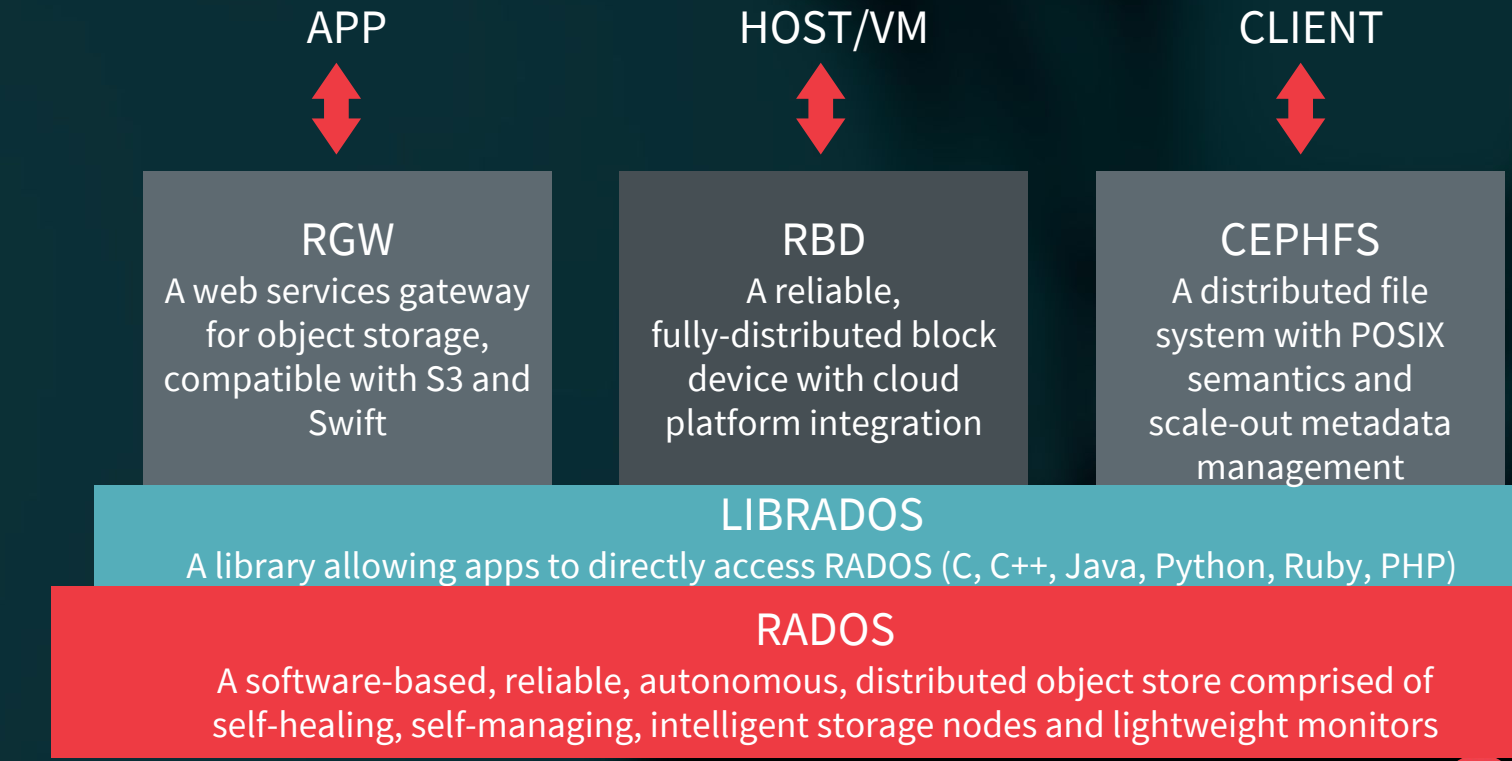
WHAT'S NEW IN CEPH

Jtech CEPH - 27.11.2018

Sébastien Han - Red Hat
seb@redhat.com
@leseb_

ceph

ARCHITECTURAL COMPONENTS



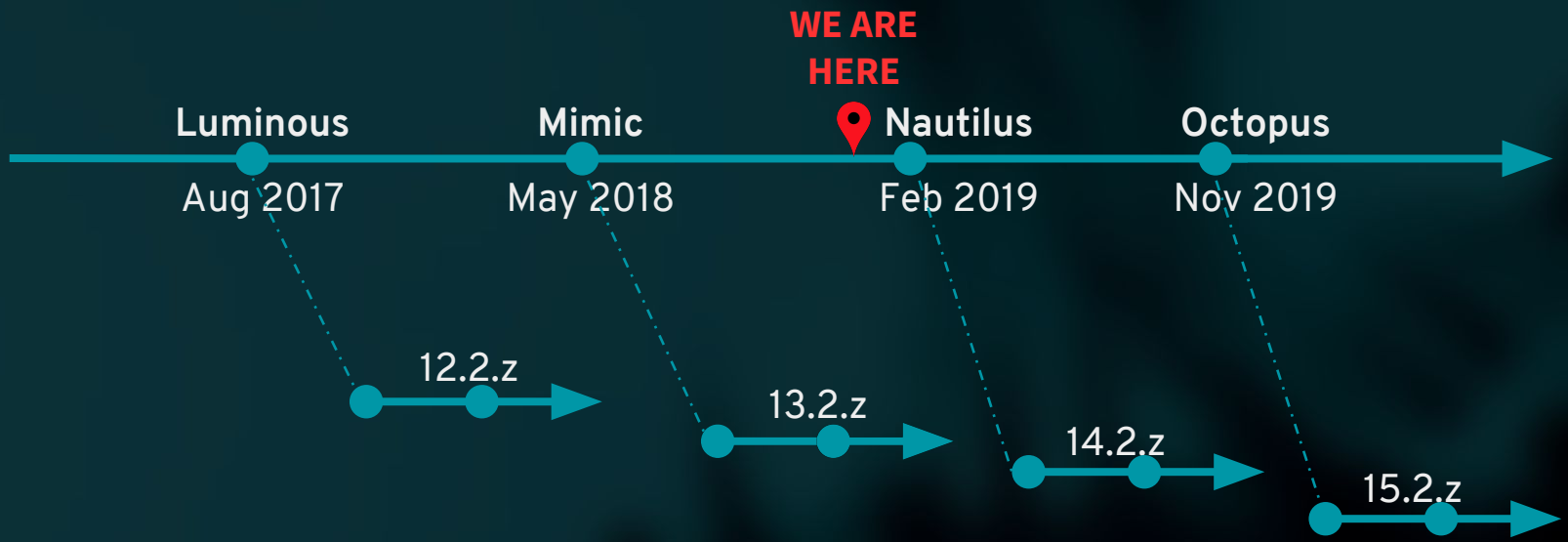
ceph

CEPH FOUNDATION

https://ceph.com/foundation/

ceph

RELEASE SCHEDULE



- Stable, named release every 9 months
- Backports for 2 releases
- Upgrade up to 2 releases at a time
  - (e.g., Luminous → Nautilus, Mimic → Octopus)

ceph

TOP COMMUNITY PRIORITIES

MANAGEMENT AND USABILITY

AUTOMATION AND MANAGEMENT

- Simplified configuration
  - ceph config stored on mons
  - Simplified memory size tuning (e.g., osd_target_memory)
- Focus on “hands off” operation
  - Conditional defaults of performance-related functions based on device types
- Hidden/automated pg_num selection
  - Enable pg_num decreases as well as increases
  - Automated, hands-off management of pool pg_num based on utilization, workload, etc.
- Additional guard rails
  - ‘ceph osd safe-to-destroy’, ‘ok-to-stop’ checks
  - Include safe-to-destroy check in ‘ceph osd destroy/purge’

ceph

TELEMETRY

- Phone home via telemetry module
  - Opt-in (of course)
  - Non-identify cluster metrics: size, version, protocols used, features enabled
- Centralized collection of crash reports
  - Alerting for transient failures (daemon crash + restart)
  - Phoned home to track failures and bugs in the wild, prioritize bugs, etc.
- Enablement for proactive/preemptive support

ceph

PHYSICAL DEVICES

- Tracking of physical devices consumed by OSDs, Mons
  - `'ceph device …'`
  - Built-in collection of health metrics (e.g., SMART)
  - Outreachy project by Yaarit Hatuka
- Disk failure prediction
  - Track "life expectancy" of physical devices
  - Self-contained prediction model ('local' mode)
  - Cloud-based SaaS prediction mode ('cloud' mode)
- Preemptive evacuation of unhealthy devices
  - Improve data safety and overall reliability of the system

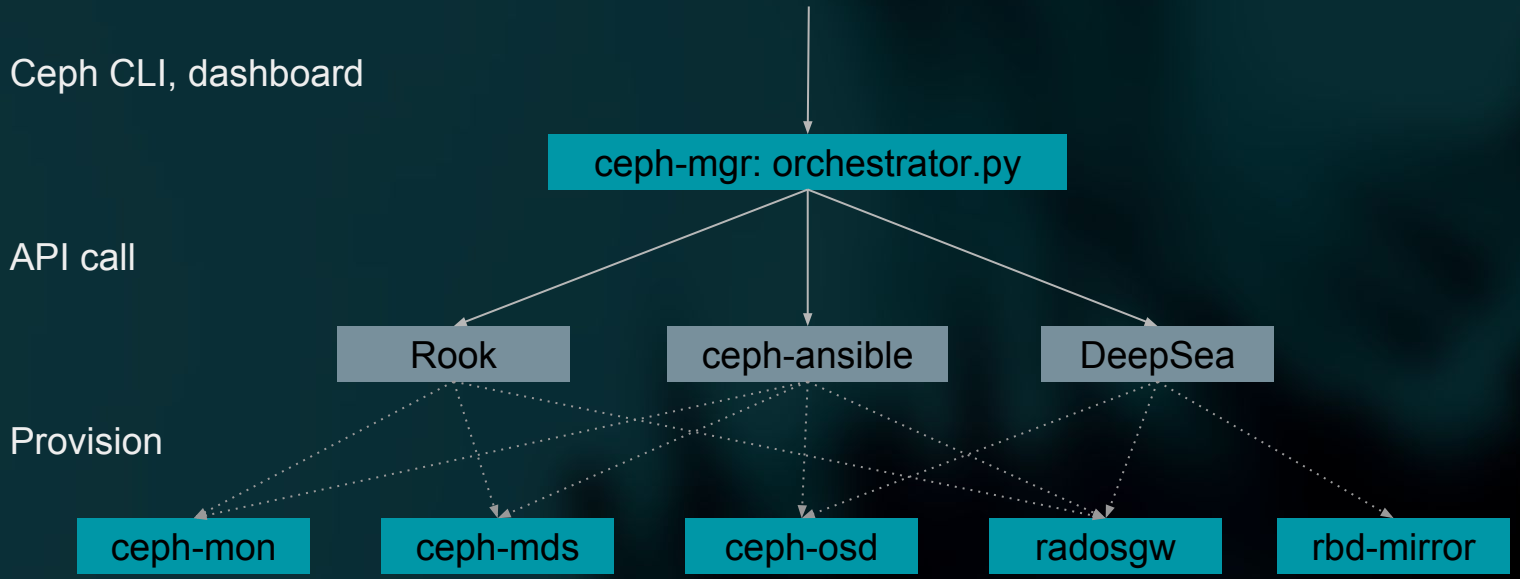
ceph

DASHBOARD

- Converged community investment on built-in Ceph dashboard
  - Hybrid of openATTIC and John Spray's original dashboard proof of concept
  - Self-hosted by ceph-mgr
  - Tight integration with other cluster management and automation functions
- Currently covers "logical" cluster functions
  - Management of Ceph services (pools, RBD images, file systems, configuration, etc.)
  - Subsumes ceph-metrics and openATTIC grafana metrics
- Indirectly laying foundation for stable and versioned management API

ceph

ORCHESTRATOR SANDWICH



ceph

ORCHESTRATOR SANDWICH

- Abstract deployment functions
  - Fetching node inventory
  - Creating or destroying daemon deployments
  - Blinking HDD LEDs
- Unified CLI for managing Ceph daemons
  - `ceph orchestrator device ls [node]`
  - `ceph orchestrator osd create [flags] node device [device]`
  - `ceph orchestrator mon rm [name]`
  - ...
- Enable dashboard GUI for deploying and managing daemons
  - Likely to have very basic capabilities in Nautilus, but more to follow

ceph

CONTAINER PLATFORMS

CEPH-CONTAINER

- Publishing canonical upstream container images
  - Barebones ceph/ceph image: base OS + Ceph packages
  - Every minor release
  - Consumed by ceph-ansible, Rook
- Ceph daemon metadata to identify host/container/pod
  - Keep CRUSH tree aligned with hardware
  - `ceph osd find ...`

KUBERNETES ABOVE AND BELOW

- Expose Ceph storage to Kubernetes users
  - Scale-out application infrastructure needs scale-out storage
  - File and Block via Persistent Volumes (PVs)
  - Dynamic provisioning
- Run Ceph clusters in Kubernetes
  - Containers simplify/hide OS dependencies
  - Containers enable finer control over upgrades
  - Schedule deployment of various Ceph daemons across hardware nodes
    - OSDs are tied to physical storage devices, but
    - ceph-mgr, ceph-mds, radosgw, rbd-mirror, iSCSI gateways, nfs-ganesha, samba, …
- K8s as “distributed OS” for management of hardware in a data center

ceph

ROOK

- Native, robust operator for Kubernetes
  - Makes it extremely easy to get Ceph up and running
- Intelligent deployment of Ceph daemons
  - e.g., add/remove/move mon daemons while maintaining quorum
  - e.g., intelligently schedule RGW/iSCSI/NFS gateways across nodes
- Upgrade orchestration
  - Update Rook operator pod (triggered via CLI or dashboard?)
  - Rook updates Ceph daemons in prescribed order, gated with health, safety, availability checks
  - Rook manages any release-specific steps (like forcing scrubs, declaring upgrade “done”, etc.)
- Dynamic provisioning of file and block (via Persistent Volumes)
  - Coming soon: dynamic provisioning of Object, too!
- Enthusiastic user community, CNCF member project

ceph

CEPH-CSI

- Replace upstream Kubernetes and Rook flexvol with ceph-csi
- Development driven by CERN, Cisco, Red Hat
- RBD and CephFS-backed PVs
- Will replace Rook's flexvol and upstream Kubernetes drivers soon

ceph

Merci

BONNES PRATIQUE SUR CEPH

Jtech CEPH - 27.11.2018

David Casier- Aevoo
david.casier@aevoo.fr

ceph


ceph

MARCHÉ GROUPE LOGICIEL CNL

Jtech CEPH - 27.11.2018

Tony Botté - Red Hat
tbotte@redhat.com

ceph

MARCHÉ GROUPE LOGICIEL CNL

ceph

CONTENU DU MARCHÉ RED HAT

Double cas d'usages: Production et Pédagogie

Souscriptions logicielles	Formations	Accompagnement
L'ensemble des solutions Red Hat	**Elles couvrent l'ensemble des technologies Red Hat**	**Prestations de service**
Leur acquisition peut se faire au travers : - De souscriptions classiques - D'offres commerciales spécifiques Éducation nommées « Offres spéciales »	Chaque session, peut être délivrée : - En intra-entreprise, dans vos locaux ou ceux du titulaire - En inter-entreprise dans les locaux du titulaire - En auto-formation (e-learning)	Intervention via des Unités d'Oeuvres pour des missions d'assistance de niveau : simple, moyen ou complexe sur site ou à distance avec des délais de réalisation de 1 à 60 jours.

LES OFFRES SPÉCIALES

Souscription Site Infrastructure

Accélérer la transformation numérique

Ansible Tower

Automatisation

Souscription Site Linux

La standardisation des environnements

RED HAT ENTERPRISE LINUX

Offre spéciale CEPH 10 Po

Red Hat Ceph Storage

ceph

MERCI