



ceph

MANAGING OBJECTS WITHIN STORAGE

SAGE WEIL

LOC - 2016.09.19

WHAT IS CEPH



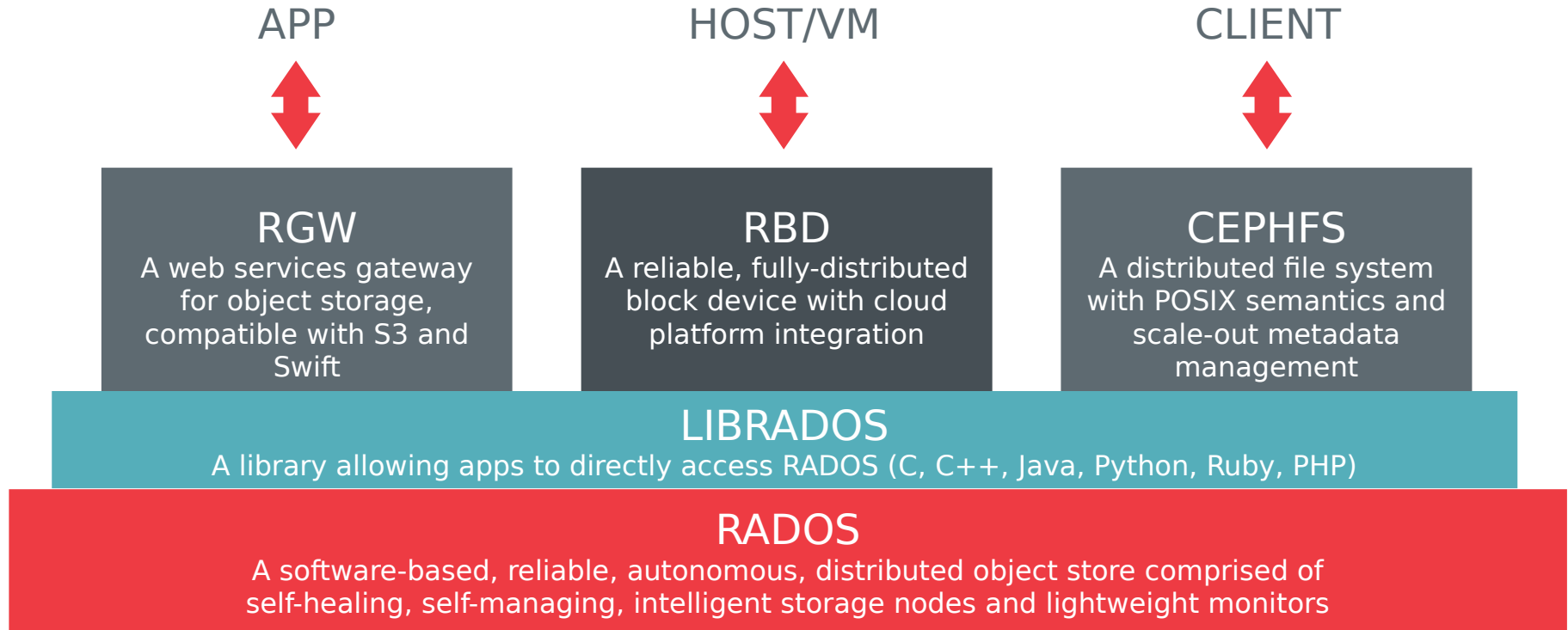
- Scale-out distributed storage
- Self manage whenever possible
- Fault tolerant – no single points of failure
- Storage hardware agnostic
- Commodity components
- Single cluster, multiple protocols
 - Object, Block, File
- Free and open source

VALUE OF OPEN SOURCE FOR ARCHIVES



- Cost at scale
- Hardware vendor independence
 - Drives down cost
 - Price vs performance vs robustness
- Software vendor independence
 - Data lifetime far exceeds vendor lifetime
- Transparency
 - How do you read your data in 10, 20, 50 years?
 - Data is not hostage to proprietary platform - source code is open
- Efficient investment of tax dollars
 - Technology investment benefits all users, not a single vendor

CEPH COMPONENTS



IS CEPH “UNIFIED STORAGE”?



- Yes, in the sense that...
 - Object (multiple protocols), Block, and File are presented from the same storage cluster
- Except...
 - RGW, RBD, and CephFS all store data differently
 - Each component is optimized for its workload and semantic structure
 - Can't write via one and read via another
- However, in object-land...
 - RGW provides S3 and Swift protocol access to the same objects with a unified security model
 - RGW now provides NFS (v4 and v3) access to objects

WHAT'S WRONG WITH POSIX



- Directory hierarchy
 - is only marginally useful for organization
 - allows directory renames, which are awkward to implement efficiently at scale
- Archives require index and search.
- Ceph Jewel included a rearchitected RGW federation design
 - Federate N clusters/sites into single object storage namespace
 - Sync/replication is being modularized to allow external integrations
 - Prototype export of object metadata into elasticsearch for index/query



Sage Weil
CEPH PRINCIPAL ARCHITECT



sage@redhat.com



[@liewegas](https://twitter.com/liewegas)

