White Paper

Red Hat Ceph Storage: Object Based Storage Ready for Primetime

Sponsored by: Red Hat

Ashish Nadkarni Amita Potnis

June 2016

IDC OPINION

As businesses embrace next-generation applications (NGAs) as a vehicle to take them on the digital transformation journey, they need a storage platform that can service the data persistence needs of the next-generation application stack, which is increasingly relying on RESTful APIs for data access. Key/value based storage platforms (aka object-based storage platforms or OBS platforms) are the solution to such needs. However, OBS platforms cannot add another silo to an already overly burdened infrastructure – in the digital economy, businesses cannot afford to deploy and manage their infrastructure in silos. They need a common storage platform that can singularly service the various data persistence needs of the application stack, not just data access via RESTful APIs. This platform has to be enterprise grade not just from a feature/function perspective but also from an operational perspective.

This is where platforms like Ceph come into play. The role of Ceph in the OpenStack ecosystem – as an open source community support storage platform that supports Cinder, the block storage service – has been well publicized. Ceph fits nicely between ephemeral storage (Logical Volume Manager, LVM) and expensive commercial storage, and services a wide variety of use cases that require block storage. While this is indeed an important piece of functionality offered by Ceph, it underserves Ceph's core architecture as an object-based storage (OBS) platform. Red Hat Ceph Storage marries the best of both worlds. It is a commercially supported open source storage platform that can be implemented as a standalone solution or as part of an OpenStack initiative to deploy any type of cloud. In either case, Ceph enables businesses to implement an agile, scalable and cost-effective infrastructure:

- Ceph is software-defined storage. It runs on server-based storage with industry-standard flash and HDD as the persistent components. This approach means that it can be implemented in a capex-friendly manner, with no expensive initial outlay. Businesses can add storage capacity when needed, and pay only incremental maintenance costs.
- Ceph can simplify the storage infrastructure. Ceph's object-based layout lends itself to Ceph being a truly scalable unified storage platform that supports object, block and file access. Resources are managed through a single pane of glass and provisioned either in a standalone fashion or via OpenStack.
- Ceph is future-proof. Ceph is friendly toward current-generation applications as well as next-generation applications, deployed on bare metal or virtual machines. Ceph (via the appropriate access protocol) serves as the storage layer for these applications.

Red Hat's open source DNA and credibility in the Linux community speaks for itself. And by way of the Inktank acquisition, Red Hat has committed itself to the Ceph community. Red Hat Ceph Storage is a corresponding, but staggered release of the community version, not a fork. This ensures that customers will always get a stable, fully supported platform.

SITUATION OVERVIEW

As businesses aggressively embark on their digital transformation journey, they are increasingly embracing next-generation applications. Most NGAs rely on very different design principles, one of which is to treat the infrastructure as a collection of software-defined resources. Infrastructure provisioning and management is not an activity that can occur in a silo, rather it is an extension of the application development, deployment, and operations paradigm. Data persistence - a core component of this infrastructure – should therefore have the following characteristics:

- **Simple to deploy and manage.** Components of this infrastructure (i.e. the data persistence platform, the data access layer, and the services provisioning layer) should be simple to install, manage, and monitor. This infrastructure should operate in a standalone fashion or fit seamlessly into frameworks such as OpenStack.
- **Software defined and open.** The storage platform should support industry-standard hardware (server-based storage), and can be controlled and accessed via standard REST APIs. This is one of the key ingredients of a capex friendly infrastructure that has low initial outlays, and reduced operating costs.
- **Distributed and highly scalable and supports newer compute models.** The storage platform supporting the infrastructure needs to operate in unison with NGAs that are distributed and shared nothing and are friendly toward newer compute models like containers.

The Case for Object-Based Storage

As ironic as it sounds, the higher the level of storage service (i.e. the more abstracted the storage platform), the more capable, scalable and flexible it is (i.e. the more data access protocols and services it can support). Tightly consistent key/value pair-based storage platforms (aka object-based storage platforms) are the perfect example of platforms that can offer object, block, and file data access protocols. Such platforms utilize a flat tenant-account-container-object namespace that is not plagued by file size or number limitations. In recent times, they have gained in popularity for three primary use cases:

- **As a unified data repository for large multi-petabyte-scale data sets.** Such data sets may require nearline or online access, with varied performance or access requirements. They cannot be stored on distributed or monolithic file systems because of the sheer number of files and/or folders.
- **To host multi-dimensional data sets in which "protocol conversion" is not always possible.** These types of data sets are not necessarily large but are often ingested and/or accessed via protocols that are closest to that data. For example, data that is Hadoop friendly requires data ingest and/or access via HDFS and not NFS or SMB. Many such data sets have additional metadata associated with them that cannot be stored in a file system.
- **As a unified data platform for block, file, and object access.** The storage platform appears with a distinct persona to the same application or computing platform. Even though the data is stored in a flat namespace, it is partitioned and set up in a manner that allows each "repository" to be tuned to the type of protocol used to access it. With support for standard

APIs and data access protocols, they can be implemented in a standalone fashion or as a component of a framework like OpenStack – for block (Cinder), object (Swift), or file (Manila).

What makes such OBS platforms worthy of being labeled as flexible is that the namespace and the manner in which the object is stored (and referenced) are completely decoupled from the characteristics of the object – the latter stored in a separate data structure known as a metadata database. Objects are stored in containers, which are "buckets" that provide a certain association to similar objects. Containers are associated with an account – which is usually a structure for managing credentials. A higher-level account management structure known as tenants may or may not be used. Objects are referenced via a universal identifier known as a UUID and accessed via a referenceable network path that can be translated depending on the access protocol being used.

OBS platforms feature an extensive policy management framework. This framework can operate on an object, container, account, or tenant level. Policies can thus be applied on a granular basis. Several OBS platforms use NoSQL databases to store metadata and policy data.

For data resiliency, OBS platforms use either a simple "shard and replicate" scheme or, in many cases, a more complex erasure coding scheme in which data is sharded, erasure coded, and then distributed. With security as a top priority, most OBS platforms now employ an intermediate scheme that encodes the data and distributes the data and the keys in an asymmetric fashion (i.e., keys are stored separately from the data shards). For data consistency and availability, OBS platforms utilize the CAP model. Vendors thus make a choice between eventual versus tight consistency, RAIN, RING and CRUSH master-slave models; and local versus geo-distributed failure domains.

Most OBS platforms are software-defined storage platforms, meaning that they do not require special hardware to perform any of the data persistence, resiliency, or availability operations. With a shared nothing architecture and server-based storage, such platforms have to incorporate a fundamental design principle: Assume that hardware will fail, and therefore don't make any assumptions about its resiliency.

Ceph — Not Just Another Storage Platform

Ceph is a unified, software-defined storage distributed platform that is maintained by a worldwide developer community that makes its source code available via LGPL 2.1 [LGPL is an acronym for GNU Lesser General Public License, which is a free software license published by the Free Software Foundation]. Ceph utilizes a scale-out, peer-to-peer shared nothing, node-based storage architecture. Ceph uses an algorithm called CRUSH (which stands for Controlled Replication Under Scalable Hashing) to determine optimum data placement. CRUSH enables the Ceph Cluster to scale, rebalance, and recover.

Ceph's foundation is the Reliable Autonomous Distributed Object Store (RADOS). It consists of two primary components:

- Ceph OSD Daemon (OSD) that stores data as objects on a storage node
- Ceph Monitor (Mon) that maintains a master copy of the cluster map

Ceph can be deployed standalone or as part of OpenStack. As a standalone platform, Ceph is:

- **A tightly consistent platform for which consistency domain is limited to a single datacenter.** Multi-datacenters are supported by way of asynchronous replication (multi-geo access is newly introduced in Red Hat Ceph Storage v2.0).

- **A unified platform that supports file and block in addition to object.** File access is supported via CephFS, block access via RBD (RADOS Block Device), and object access via the RADOS Gateway (RGW), which exposes the object store as a RESTful API that is compatible with Amazon S3 and OpenStack Swift.

As part of an OpenStack deployment, Ceph integrates well with commonly deployed components:

- **Cinder**: RBD used as the backend for persistent storage
- **Nova**: RBD used as the backend for ephemeral storage
- **Keystone**: Authentication for account and tenant credentials
- **Glance**: RBD used for image hosting with CoW (Copy on Write) cloning to Nova or Cinder
- **Swift**: Replacement/alternative for OpenStack Swift
- **Manila**: File interface for CephFS

Over the past two to three years, Ceph has gained tremendous momentum in the enterprise. The global community of developers and vendors supporting Ceph has grown significantly. Ceph is now deployed in many of the Fortune 1000 enterprises, and supported by leading infrastructure vendors such as Red Hat, SUSE, SanDisk, Fujitsu and others.

Red Hat Ceph Storage

Since 1993, Red Hat's mission has been to enable enterprises to embrace open source platforms. Enterprises gravitate toward vendors like Red Hat for their ability to commercially support open source platforms. Without Red Hat, enterprises would have required a serious investment in additional human resources to deploy and maintain open source platforms in their environments. A key reason for the success of open source is vendors like Red Hat. And a key reason for Red Hat's success is the company's commitment and support to the community – both financial and code contributions. Red Hat is one of the largest employers of developers supporting all open source initiatives worldwide.

Having successfully commercialized Linux, OpenShift, JBoss, KVM, and OpenStack to name a few, Red Hat is working to change the storage market with Gluster and Ceph. In fact, since the Inktank acquisition, Red Hat has made significant strides with Ceph Storage. The recent release of Red Hat Ceph Storage 2.0 brings many improvements that make Ceph truly enterprise worthy:

- Global clustering and namespace that enables multi-site federation
- Enhancements to the S3 and OpenStack Swift 3.0 APIs for additional security and compatibility
- Active directory/LDAP integration that makes it easier to deploy Ceph as a standalone platform (in addition to integration with Keystone, which enables OpenStack integration)
- Indexless buckets that allow storage of large data sets without managing the actual bucket indexes (should round out Big Data use cases).
- Support for additional object schematics like object expiration (to make Ceph more appealing to the cable provider/telecom industry for content distribution as a use case)
- HDFS access via a plug-in enabling the use of Ceph as a data lake for Big Data and business analytics
- Other improvements such as stability of the object store, security of individual Ceph processes, and small write performance on SSD with RHEL 7.2 – all geared towards enterprise use cases

With the release of Red Hat Ceph Storage 2.0, Red Hat is squarely targeting enterprise use cases that showcase the versatility of Ceph as an object store in the enterprise. Use cases include:

- **Active archiving**: Ceph as a standalone nearline object storage platform with online access
- **Content repository**: With performance tuning, Ceph as a platform for content streaming
- **Storage for next-gen apps**: Ceph's support for direct S3 or Swift access
- **In-place analytics platform**: Ceph as a nearline, simpler, and cheaper drop-in replacement for HDFS that supports in-place MapReduce operations
- **Unified storage for OpenStack**: Ceph continues its strong affinity to OpenStack environments – it works well as a stable block layer for Cinder and a viable alternative to OpenStack Swift

FUTURE OUTLOOK

Within a span of two to three years, open source storage has gained significant traction in the storage industry. Vendors like Red Hat have ushered in an era in which enterprises — much like how they embraced Linux — have begun to use Ceph for storing their primary data. When enterprises are developing infrastructure to support next-generation applications, platforms like Ceph (either as part of OpenStack or in a standalone capacity) feature prominently on their short list.

IDC believes that future infrastructure decisions will continue to be influenced heavily by open source technologies. Ceph is already winning in the enterprise. Slowly but surely, more enterprises will appreciate the benefits of a unified platform like Ceph and its applicability in use cases like active archives, content streaming, next-generation applications and Big Data and analytics.

CHALLENGES/OPPORTUNITIES

In spite of the huge uptick in adoption of Red Hat Ceph Storage, Red Hat is diligently working to gain mindshare in the enterprise. In that sense, it is battling on two fronts:

- Red Hat is vying with public cloud service providers, which are major competition for all storage players, and Red Hat is no exception. On paper, public cloud may appear cheaper than on-premises OBS deployments, but when other indirect costs such as WAN bandwidth are factored in, it is not so cheap. Red Hat therefore needs to continue to emphasize the point about the overall total cost of ownership (TCO) of adopting Red Hat Ceph Storage.
- Red Hat is taking on the traditional, proprietary commercial storage vendors — many of which have been in the industry for over three decades with mature storage platforms. As an upstart, Red Hat can continue to invest in rounding out the enterprise worthiness of Ceph.

IDC believes that Red Hat's success in the storage industry lies in a robust partner ecosystem that includes:

- Server manufacturers such as Lenovo, Dell, Cisco, HPE, and SuperMicro with jointly engineered solutions
- Application ISVs that certify Red Hat Ceph Storage for use with their products and jointly develop reference architectures for various use cases
- Global solutions and services vendors that can implement Red Hat Ceph Storage and integrate it with custom applications being developed by the business or being acquired from ISVs

CONCLUSION

As businesses embrace next-generation applications (NGAs) as a vehicle for their digital transformation journey, they will need to pay specific attention to the transformation of their infrastructure. This infrastructure transformation, which includes storage, will need to occur in lock-step with their application transformation, and effectively end up being a pool of software-defined resources that is provisioned, controlled and managed in an application centric, developer and operator friendly fashion.

With platforms like Red Hat Ceph Storage, businesses can aggressively build out their infrastructure to support use cases like active archiving, content repositories, unified data persistence for NGAs and in-place analytics in a standalone fashion, or as a core component of frameworks like OpenStack. With a partner like Red Hat, that has a long history in successfully commercializing open source projects and platforms, businesses are well equipped to embark on this journey.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com