OpenStack Storage: Maturity & Choice

Scalable & flexible storage options help drive cloud adoption on OpenStack

Executive Summary

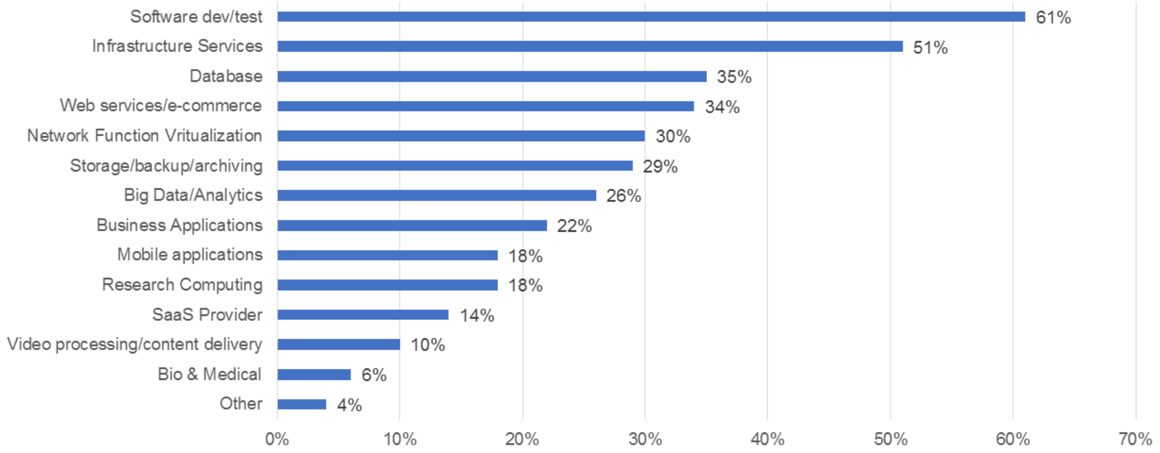
With the growing popularity of cloud computing, there is still a need for alternatives to the major public offerings. OpenStack presents an option that enables organizations to build elastic and scalable cloud services that deliver more control and choice for their IT environments. The rapid growth of data consumption, combined with the elasticity that cloud solutions can offer, creates a compelling reason for businesses to begin migrating away from traditional architectures and towards cloud solutions. OpenStack represents a new choice for both enterprises and service providers, helping balance performance, scalability, and cost for clouds. With mature storage components as well as third-party options for storage, OpenStack is now viewed as a more robust solution for cloud deployment.

Today's Customer needs

IT organizations are moving to cloud technologies rapidly to gain efficiencies, time-to-market advantages, and cost savings over traditional IT infrastructure. Many are turning to the open source OpenStack as a mechanism for IT organizations to build their own on-premise private clouds or for service providers to build their public cloud infrastructure. OpenStack provides a framework to deliver compute, storage, and networking resources that can be consumed as a service. The strong ecosystem of vendors who deliver solutions and services enable enterprise support and viability for public or private cloud solutions built on OpenStack technology.

Over the last several years, OpenStack has evolved from a platform typically used by early adopters for test/dev environments into one now used by enterprises for production workloads. According to the Fall 2016 OpenStack User Survey, the share of OpenStack deployments in production is 20% greater than a year ago, with 71% of clouds in production or full operational use. The latest results also showed that 72% of those surveyed said their top business driver for deploying OpenStack is to save money over alternative infrastructure choices. While many organizations are using OpenStack for cloud-native workloads, there has also been an increase in the number of traditional workloads deployed in an OpenStack environment.

FIGURE 1: WORKLOADS RUNNING ON OPENSTACK



Source: https://www.openstack.org/analytics November 2016

GROWTH & MATURITY OF OPENSTACK STORAGE

When OpenStack was first released in 2010, Rackspace contributed the source code of its Cloud Files product as the initial code for the OpenStack Object Storage component (now Swift), which was designed to be an open source alternative to Amazon S3. While the first release only supported object storage, other storage capabilities have been gradually added over time. The OpenStack Nova project originally had a component called “nova-volume” which then broke off into the independent Cinder project that brought block-level storage to OpenStack in 2012. After that, OpenStack saw the addition of Manila as an independent experimental project as a file sharing component with full support under the Liberty release in the fall of 2015.

The OpenStack Newton release, made available in October 2016, is one of the newer and more stable releases. The Newton release includes at-rest data encryption, performance improvements, as well as scalability enhancements for all storage services (Swift, Cinder, and Manila). MI&S describes OpenStack’s growth and enterprise readiness in an earlier whitepaper.

The OpenStack project has done an excellent job publishing information to understand the maturity of each project included in a release. OpenStack defines maturity as a metric (1 through 8) based on a set of distinct stability criteria. These include the existence of an install guide, support by 7 or more SDKs, an adoption percentage greater than 75%, project achievement of corporate diversity, and whether the project has stable branches.

TABLE 1: ADOPTION & MATURITY FOR STORAGE SERVICES

	Storage Type	Package Type	Adoption	Maturity	Age
Cinder	Block Storage	Core Storage	88%	8 of 8	5 years
Swift	Object Storage	Core Storage	47%	7 of 8	7 years
Manila	Shared File System	Optional Storage	11%	5 of 8	3 years

Source: http://www.openstack.org/software/project-navigator

Cinder has the highest adoption of the storage services at 88%, with full maturity in all eight categories. Swift misses full maturity, because its 47% adoption is below the 75% requirement. Object storage may be less commonly used than block storage, and other maturing object storage alternatives (namely Ceph which provides block, file, and object storage) are available and well established. Manila is the youngest of the three storage projects, and its current adoption rate of 11% is significantly lower than the other two. Manila does not have any supporting SDKs and does not support rolling upgrades at this time. While actual adoption is low, it appears Manila interest is growing in the community as 36% of those surveyed are interested in Manila for future deployments.

CINDER: BLOCK-LEVEL STORAGE

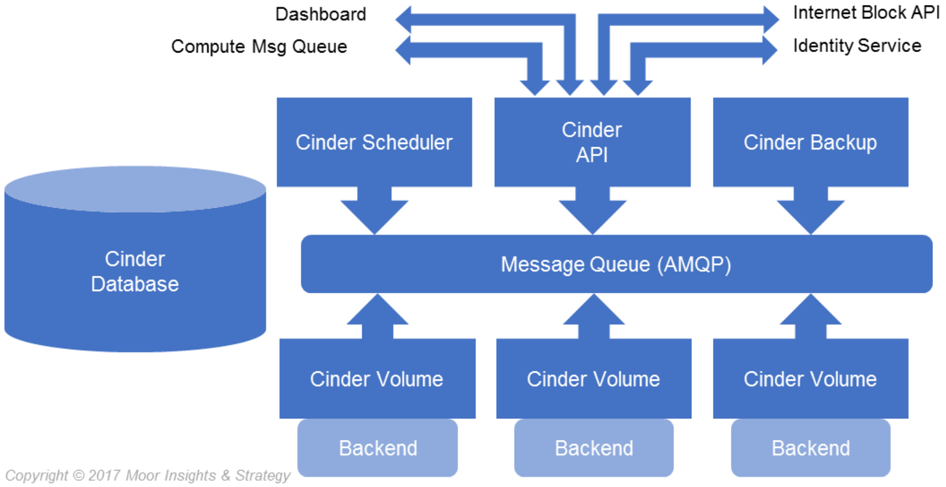
The Cinder service is an on-demand, software-defined block storage API for OpenStack. The service is designed to present self-managed block storage to end users either by a reference implementation via the Local Volumes Managed (LVM) service or through plugin drivers that work with a variety of backend storage devices.

Cinder is deeply integrated into OpenStack. In fact, it is difficult, though not impossible, to use Cinder as a block storage service outside of OpenStack. Because of this tight linkage between Cinder and OpenStack, the profile of a Cinder user is identical to the general user profile that is created for OpenStack

The message API and command-line interface are recent additions to Cinder that enable users to obtain the results of an asynchronous task. This feature is valuable for troubleshooting failed storage requests.

Some key focuses include stability and expanded functional testing that includes third-party CI(s), Active / Active High Availability, and other capabilities designed to make Cinder a more feature-rich storage management platform. As Cinder matures, the community will continue to face the question of whether it should embrace more cloud-centric thinking or pursue a more enterprise-centric direction. By any measure, this project is a success as more drivers for storage systems are being added, enabling Cinder to serve in enterprise roles beyond cloud usage.

FIGURE 2: CINDER ARCHITECTURE



Source: Moor Insights & Strategy

SWIFT: OBJECT-BASED STORAGE

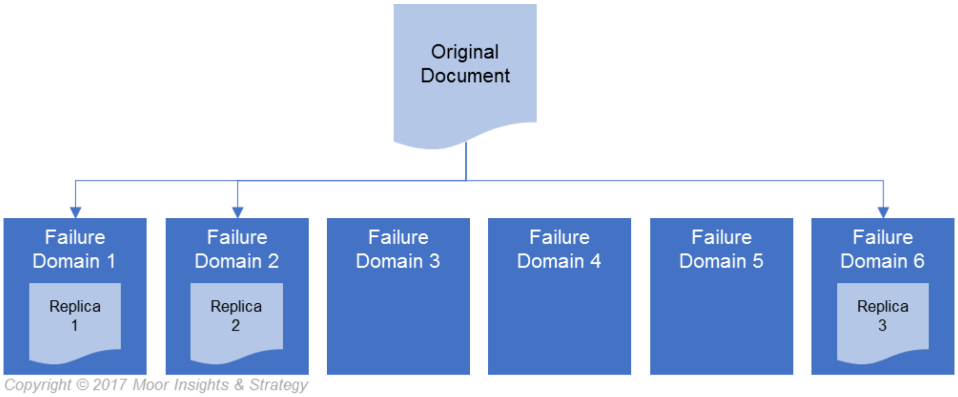
The Swift service is an on-demand, software-defined object storage service that enables cost-effective storage which can be implemented on industry-standard servers. Swift allows users to create, modify, and get both objects and metadata by using a simple and scalable RESTful HTTP-based API. The service can scale from a single server to very large clusters, all based on a highly fault tolerant, distributed, and consistent architecture. Swift is not a file system and has no mount points or mountable directories; items that are stored using Swift are in a binary large object (BLOB) format.

While the predominate use of Swift is as an object storage service for OpenStack, it has a fully defined set of interfaces and APIs that make it an excellent service for applications using extremely large objects, including the use outside of OpenStack.

Swift receives software-based encryption enhancements along with general performance and scalability improvements related to global clusters and concurrency / throughput in the latest version. Future releases will include additional performance and scalability improvements, automatic data tiering, and global erasure-coded clusters that address some of the limitations of the current implementation.

Swift servers achieve high availability via replication with placement controlled by hash rings. Typically, three different copies of the file will be stored on different failure domains (configurable to higher numbers for greater availability at the cost of storage capacity). The ring concept is based on the premise that no replicated copies of an object should be placed in the same failure domain. In the example below, there are 6 failure domains and each of the three replicated copies is placed in a distinct domain.

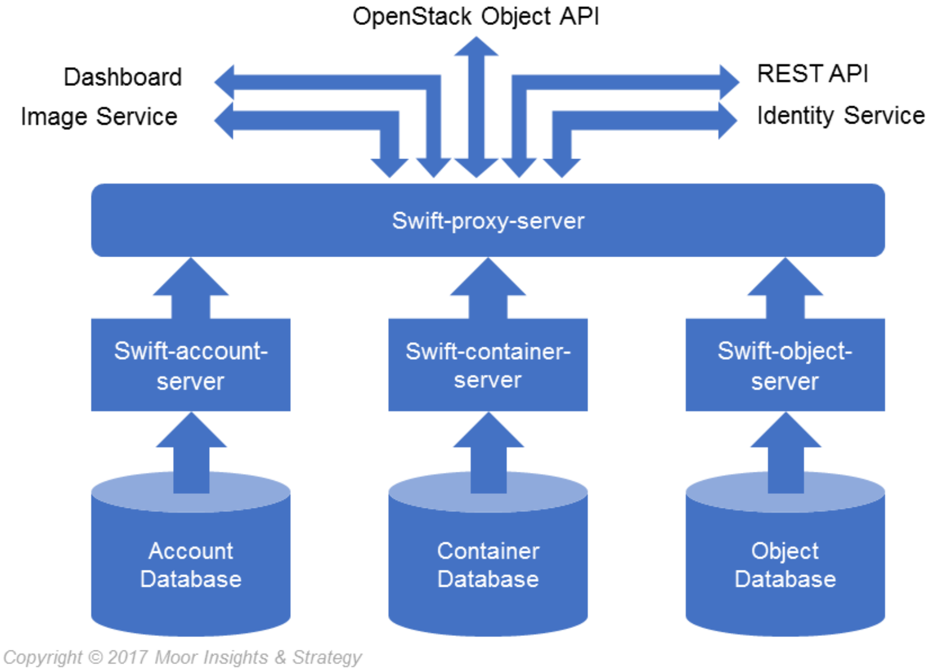
FIGURE 3: FAILURE DOMAINS



Source: Moor Insights & Strategy

Failure domains can be as localized as an individual storage device or server, or more locationally diverse as individual racks, datacenters, or even countries.

FIGURE 4: SWIFT ARCHITECTURE



Source: Moor Insights & Strategy

MANILA: FILE SHARING SERVICE

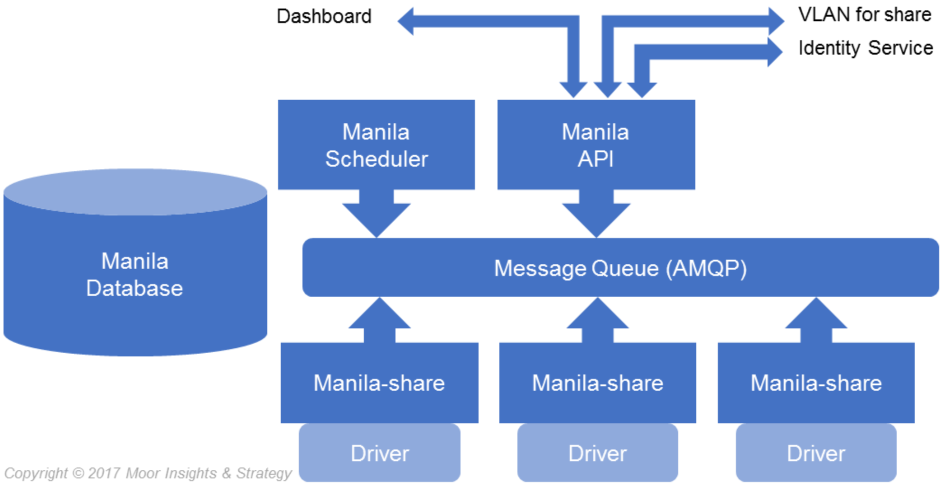
The Manila service is an OpenStack project that provides a shared file system as a service; it is intended to provide equivalent value as Cinder for block storage and Swift for object storage. Manila works to overcome some of the limitations of Cinder and Swift, bringing more functionality to OpenStack. Because of the lack of locking and synchronization, a Cinder volume / LUN could historically only be accessed by a single Nova guest. And while Swift objects are excellent for storing and managing large objects, they are not well-suited for a transactional environment. The Manila sharing service fills these gaps and provides options to specify any major network file system. Additionally, Manila provides automation for provisioning file shares to Nova hosts to help with the automation of file sharing.

In Manila parlance, a share refers to a mountable file system that can be accessed by several hosts and users at the same time. When the share is created, it must be associated with a network so that it can be listed, queried for information, updated, or deleted. Upon creation of the share, it can be associated with any of the major network file systems including Network File System (NFS), Common Internet File System (CIFS), Gluster file system (GlusterFS), Hadoop Distributed File System (HDFS), or Ceph File System (CephFS). Once specified, the user can then access the share using the customary techniques and commands appropriate for the selected protocol.

Manila as a file sharing service is an extremely important step supporting enterprise migration to OpenStack. The ability to create a familiar environment that supports this type of transaction processing will make Manila highly valuable for the transition. The Newton release provides several new drivers, expands backend support, and improves stability of existing drivers.

Manila now also includes driver-assisted data migration, which provides efficient migrations between common vendor storage backends. In the future, the Manila community plans to improve migration and / or replication, add stability to scalable deployments, enable user error messaging for failed requests, and improve snapshots and share groups.

FIGURE 5: MANILA ARCHITECTURE



Source: Moor Insights & Strategy

CEPH: A THIRD-PARTY ALTERNATIVE OPEN SOURCE SOLUTION

Ceph is an open source, server-based storage product that is highly scalable and redundant, providing object, block, and file system storage to a single storage cluster. In a single product, Ceph can deliver similar functionality to Swift, Cinder, and Manila combined. Ceph is based on commodity hardware and provides distributed operation, enabling it to run with no single point of failure. Scalable to the exabyte level and based on Reliable Autonomic Distributed Object Store (RADOS), a Ceph storage cluster can be tightly integrated with applications for a variety of storage needs.

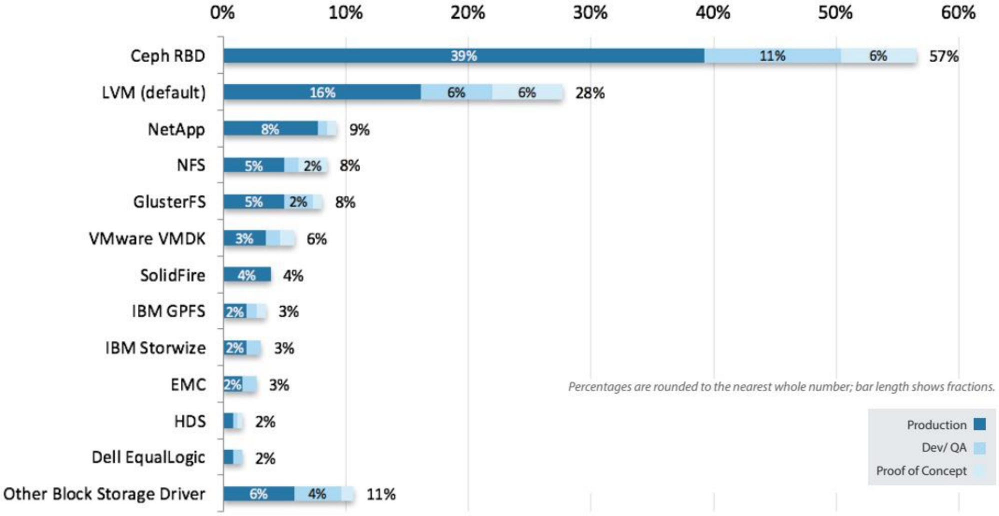
The Controlled Replication Under Scalable Hashing (CRUSH) algorithm helps optimize data placement and supports high availability. A Ceph storage cluster is self-managing, self-healing, and allows for the fine-tuning of many parameters including CRUSH, the specification that sets rules for data placement in the cluster.

Ceph has three major components.

- **Object Storage Daemon (OSD)** is responsible for the storage location of the data. A typical deployment creates one OSD for each hard drive or SSD drive.
- **Meta-Data Server (MDS)** stores the metadata used for the CephFS, building a POSIX file system on top of objects for Ceph clients.
- **Monitor (MON)** is a management monitor that acts as the proxy for communication with external applications and clients.

While Swift is integrated directly into the OpenStack project, Ceph (as a third-party alternative) is the most widely deployed block storage for OpenStack use cases according to the April 2016 OpenStack user survey.

FIGURE 6: BLOCK STORAGE DEPLOYMENTS



Source: https://www.openstack.org/assets/survey/April-2016-User-Survey-Report.pdf

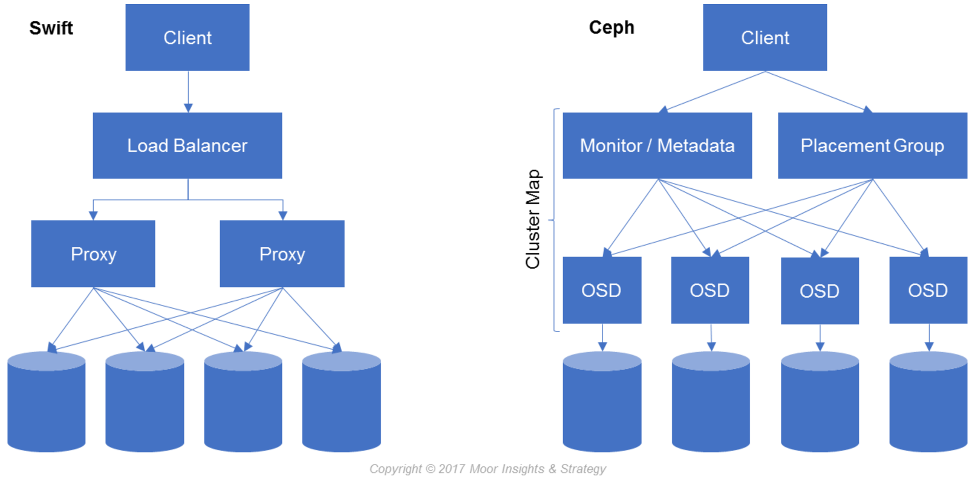
CONTRASTING CEPH & SWIFT

Swift and Ceph are the two most popular OpenStack storage solutions. Both are open source, distributed, scale-out storage solutions based on commodity hardware, but there are differences primarily in scope. Swift provides only object storage to a cluster; Ceph provides object, block, and file system storage, all from a single cluster.

Ceph clients communicate directly with the storage nodes through the Object Storage Daemon (OSD), while communications with a Swift cluster must go through a Swift proxy. If either does not have sufficient resources to meet workload demand, performance will suffer. Ideally, the Ceph architecture should produce lower latency, while the Swift architecture provides better abstraction and control of backend behavior.

Ceph is an excellent choice for organizations seeking a complete, flexible storage solution, primarily because it provides block, file, and object formats in a single solution. Ceph works well with OpenStack; Swift, Cinder, and Manila benefit from their deep OpenStack integration and the contributions of the OpenStack community. Careful consideration should be given to the specific business requirements to determine the best option for an OpenStack deployment.

FIGURE 7: SWIFT VS. CEPH ARCHITECTURAL VIEW



Source: Moor Insights & Strategy

INTEL'S CONTRIBUTION TO OPENSTACK & STORAGE

One of the key Intel contributions to OpenStack is though the creation of the OpenStack Innovation Center (OSIC) and its collaboration with Rackspace. The OSIC is uniquely equipped to allow community testing of the upstream code base by running it on two 1,000-node clusters. As of April 2016, the joint Intel / Rackspace team had submitted 174 patches directly and had been involved in the review of almost 1,000 other patches. These efforts underscore Intel's commitment to accelerating OpenStack adoption.

Intel has been and continues to be a key contributor to open source. Measuring the full number of OpenStack storage project code commits (as of December 2016), Intel was the #5 contributor to OpenStack Cinder, #14 contributor to OpenStack Manila, and the #5 contributor to OpenStack Swift. In addition to its work directly with the OpenStack project teams, Intel is one of the top 5 companies contributing to the Ceph code base.

MI&S'S VIEW OF THE OPENSTACK STORAGE MARKET

The industry's focus and progress in software-defined storage (SDS) is undeniable. While traditional enterprise storage companies continue to add innovative products and solutions to their portfolios, there is a strong embrace of OpenStack as a viable open source storage option. The speed of modern business is steadily increasing, and OpenStack is providing the agility and flexibility that innovative companies demand to keep them ahead of the competition. As we look at the market, we see clear trends that are driving decisions; these favor a greater use of OpenStack.

- **Increased Focus on Storage Cost**: With the explosion of data and its transformation into insight and action, companies are seeking open source alternatives to traditional enterprise storage models. This is primarily because open source solutions like OpenStack do a better job of scaling cost as storage needs escalate quickly.
- **Agility**: Companies are displaying an unprecedented focus on efficiency and the need to accelerate new service deployments in ways that would previously have been inconceivable. OpenStack, with its elastic capabilities, enables more agile scaling up and down of services in a rapid manner.
- **Multi-Cloud and Hybrid Cloud**: Between traditional IT, private cloud, and public cloud, businesses will need storage solutions that can span multiple domains and share information across application boundaries. OpenStack brings many of the same capabilities and structures found in public cloud services into the datacenter for private cloud, enabling more consistent deployment of technology.
- **Technology Transitions**: As forces like digital transformation reshape how businesses approach their IT to address new opportunities, open source solutions like OpenStack are well suited to building out that capability. The ability for OpenStack to scale storage capabilities is especially useful in these times of transition, because new technology approaches often underestimate the true amount of storage as all of the information within a company becomes digital. The scalability of OpenStack brings an elastic capability that can bend and shift as companies transition into new business models.

A singular focus on a specific need like storage can be necessary and essential for business evolution, but companies are cautioned to ensure they have a holistic vision of the desired outcome to ensure a comprehensive approach to their IT infrastructure.

CALL TO ACTION

OpenStack has reached critical mass and continues to build momentum. Storage is no longer simply just a part of OpenStack; it is a key focus area that will help enable businesses change their velocity. With the maturity of the integrated OpenStack components (Cinder, Swift, and Manila), organizations deploying OpenStack will find stable, supportable, and functional tools to optimize their storage. Additional contributions from the development community and vendors continue to drive better stability and reliability while infusing new functionality into the projects. Enterprises can now confidently deploy OpenStack, as it has the capability to handle even complex storage needs, rivaling the traditional (proprietary) options.

Additional open source alternatives like Ceph also deliver excellent opportunities to improve storage both as part of and separate from OpenStack. MI&S encourages businesses looking at OpenStack to investigate both the integrated options as well as the third-party alternatives to determine applicability to their specific needs.

Those who are considering OpenStack will find that becoming a project contributor and engaging with the community can be a source of guidance and best practices. Intel is also both a key contributor to the technology as well as a source of information and guidance. More information about Intel’s involvement in OpenStack can be found in the Intel and Rackspace sponsored OpenStack Innovation Center.

Important Information About This Paper

Authors

John Fruehe, Senior Analyst at Moor Insights & Strategy
Gina Longoria, Senior Analyst at Moor Insights & Strategy
Jimmy Pike, Chief Technical Officer at Moor Insights & Strategy

Publisher

Patrick Moorhead, Founder, President, & Principal Analyst at Moor Insights & Strategy

Editor / Design

Scott McCutcheon, Director of Research at Moor Insights & Strategy

Inquiries

Contact us if you would like to discuss this report, and Moor Insights & Strategy will respond promptly.

Citations

This paper can be cited by accredited press and analysts but must be cited in-context, displaying author’s name, author’s title, and “Moor Insights & Strategy”. Non-press and non-analysts must receive prior written permission by Moor Insights & Strategy for any citations.

Licensing

Disclosures

This paper was commissioned by Intel. Moor Insights & Strategy provides research, analysis, advising, and consulting to many high-tech companies mentioned in this paper. No employees at the firm hold any equity positions with any companies cited in this document.

Disclaimer

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. Moor Insights & Strategy disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for errors, omissions, or inadequacies in such information. This document consists of the opinions of Moor Insights & Strategy and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice.

Moor Insights & Strategy provides forecasts and forward-looking statements as directional indicators and not as precise predictions of future events. While our forecasts and forward-looking statements represent our current judgment on what the future holds, they are subject to risks and uncertainties that could cause actual results to differ materially. You are cautioned not to place undue reliance on these forecasts and forward-looking statements, which reflect our opinions only as of the date of publication for this document. Please keep in mind that we are not obligating ourselves to revise or publicly release the results of any revision to these forecasts and forward-looking statements in light of new information or future events.