Storage Services - Objects (Object Services)

The Nutanix Objects feature provides highly scalable and durable object services via an S3 compliant API (More Information on S3: LINK). Given Nutanix Objects is deployed on top of the Nutanix platform, it can take advantage of AOS features like compression, erasure coding, replication and more. Objects was introduced in AOS 5.11.

Supported Configurations

The solution is applicable to the configurations below (list may be incomplete, refer to documentation for a fully supported list):

Core Use Case(s):

- Backup target with compliance-level WORM
- Archival target for PACS / VNA, Email, long term backups
- Big Data including Data Lakes, AI/ML workloads, analytics, machine loggins, SIEM, Pub/Sub
- DevOps, cloud-native, content repos, K8s

Management interfaces(s):

- Prism Central (PC)

Hypervisor(s):

- N/A - Runs on Nutanix MSP (Dependent on MSP supported Hypervisors)

Upgrades:

- LCM

Key Features:

- WORM
- Versioning
- Legal Hold
- Multiprotocol
- Streaming Replication
- Multi-part Upload
- Fast Copy
- S3 Select

Supported Protocols:

- S3 (version 4)
- NFSv3

Nutanix Microservices Platform (MSP)

Nutanix Objects leverages the Nutanix Microservices Platform (MSP) and is one of the first core services to do so.

Nutanix MSP provides a common framework and services to deploy the Objects component's associated containers and platform services like Identity and Access Management (IAM) and Load Balancing (LB).

Key terms

The following key terms are used throughout this section and defined in the following:

- Bucket
  - An organization unit exposed to users and contains the objects (think share to a file on a file server). A deployment can, and typically will, have multiple buckets (e.g. departmental, compartmental, etc.)

- Object
  - The actual unit (blob) of storage and item interfaced with via the API (GET/PUT).

- S3
  - The term used to describe the original object service Amazon Web Services (AWS) introduced and is now used synonymously for an object service. S3 also is used to define the object API which is highly leveraged throughout projects.

- Worker VMs
  - Virtual machines created during object store deployment that host various containerized Objects services. The are also referred to as Objects nodes.

- Objects Browser
  - A user interface (UI) that allows the users to directly launch the object store instance in a web browser and perform bucket and object level operations.

The figure shows the high-level mapping of the conceptual structure:
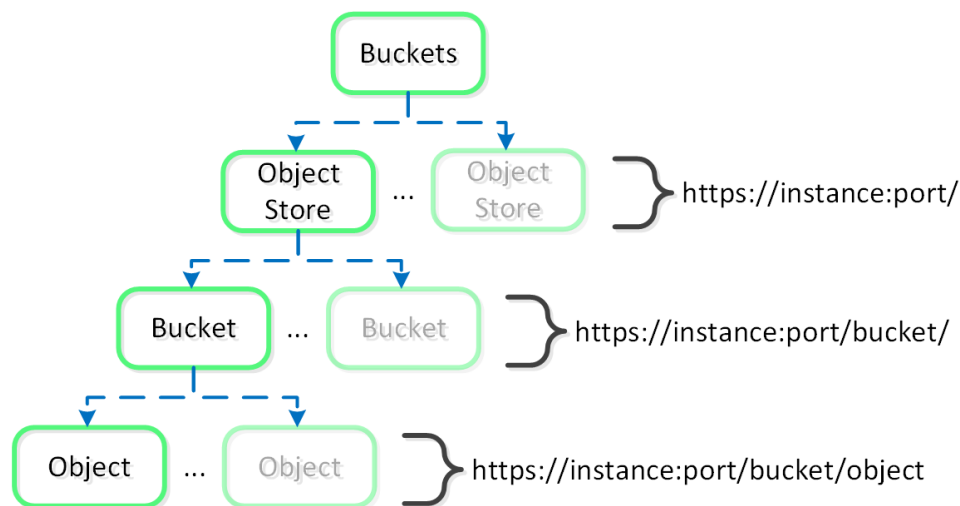

Objects - Hierarchy

Objects Constructs

This feature is composed of a few high-level constructs:

- Load Balancer
  - The load balancer is part of the Nutanix MSP and serves as a proxy for service and data requests. This ensures high-availability for the service and load balancing among the Objects containers.

- Service Manager
  - The service manager serves as the endpoint for all UI requests and manages object store instances. It is also responsible for collecting stats from instances.

- Metadata Server
  - The metadata server contains all the metadata around a Nutanix Objects deployment (e.g. buckets, objects, etc.). To bolster performance of the metadata server, Nutanix developed ChakrDB, leveraging a RocksDB based Key-Value store which consumes Nutanix Volumes for storage.

- Object Controller
  - Built in-house, this is the data path engine that is responsible for managing object data and coordinating metadata updates with the Metadata Server. It interfaces with Stargate via the Storage Proxy API.

- Region Manager
  - The Region Manager is responsible to managing all of the object storage information (e.g. Region) on DSF.

- Region
  - A region provides the high-level mapping between an object and the corresponding locations on Nutanix vDisk(s). Similar to a vDisk ID, offset and length.

- Atlas Service
  - The Atlas Service is based on the MapReduce framework and is responsible for object lifecycle policy enforcement, garbage collection, tiering and more.

- S3 Adapter
  - Built in-house and all new in Objects 3.2, the S3 adapter provides the S3 API, handles authorization as well as API requests from S3 clients. These are then translated to the Object Controller and processed.

The figure shows a detailed view of the Objects service architecture:
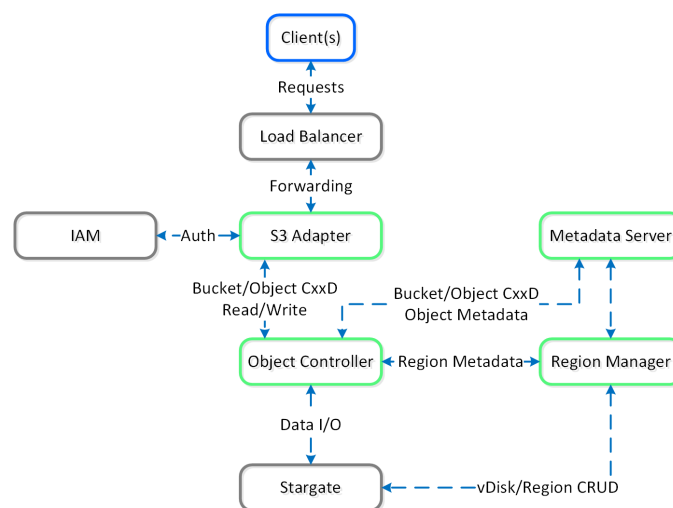
Objects - Architecture

The Objects specific components are highlighted in Nutanix Green. With objects there's no concept of an "overwrite" hence the CxxD vs. CRUD (Create/Replace/Update/Delete). The commonly employed method for an object "overwrite" is to create a new revision or create a new object and point to the new object.

Open Source Components

Nutanix Objects is built using a combination of organic innovation while also leveraging open source components. We believe this makes our products more robust and performant giving our customers the best solution for their business.

More Information on our Open Source License Disclosures: LINK

Object Storage and I/O

An object is stored in logical constructs called regions. A region is a fixed segment of space on a vDisk.

The figure shows an example of the relationship between a vDisk and region:
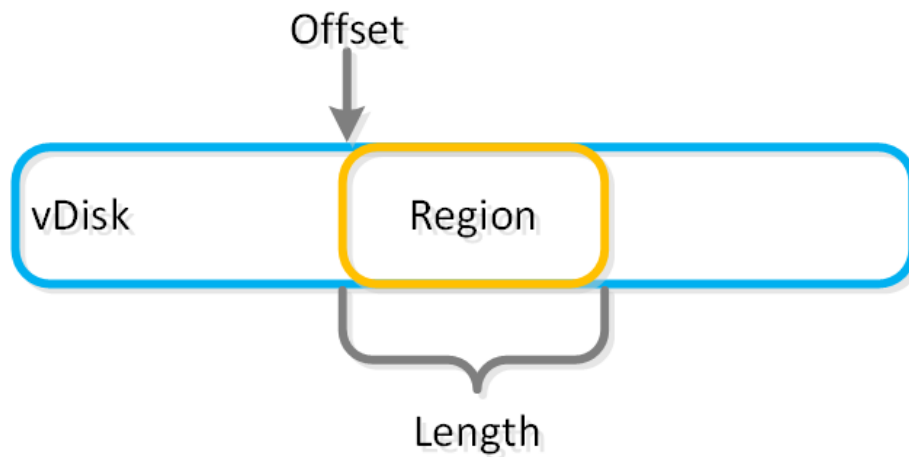
Objects - vDisk Region

Smaller objects may fit in a chunk of a single region (region id, offset, length), whereas larger objects may get striped across regions. When a large object is striped across multiple regions these regions can be hosted on multiple vDisks allowing multiple Stargates to be leveraged concurrently.

The figure shows an example of the relationship between a object, chunk and region:
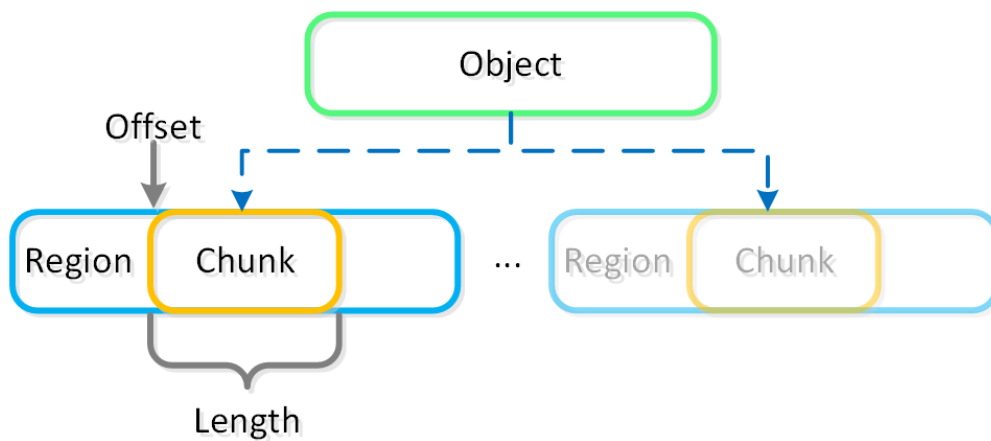
Objects - Object Chunk

The object services feature follows the same methodology for distribution as the Nutanix platform to ensure availability and scale. A minimum of 3 object VMs will be deployed as part of the Objects deployment.