SCALITY | ARTESCA

WHITE PAPER

Scality ARTESCA: Simple S3 object storage for immutable, cyber-resilient backups

2024

Content

I. ARTESCA Introduction

Your last line of defense against any threat to data recoverability? A cyber-resilient storage solution.

Immutable storage has become a cornerstone of cybersecurity. But in the era of AI-powered ransomware, it's no longer enough. Protecting data against the widest possible range of current and future threats requires a new way of thinking — one where immutability is just one piece of the puzzle.

Designed to provide the strongest form of immutability plus end-to-end cyber resilience, Scality ARTESCA is the only solution that safeguards data at five core levels for unbreakable protection against the cyberthreats of today and tomorrow.

With backup repositories targeted in 93% of ransomware attacks, organizations are seeking surefire ways to mitigate risk, avoid paying extortionate ransoms, and maintain uninterrupted business operations amidst inevitable assaults on their data.

Raising the stakes even higher is a new era of AI-supercharged cyberthreats that's not just on the horizon — it's here now. To protect data against increasingly sophisticated and evasive threats, your solution must meet a new standard of protection.

Scality ARTESCA is built on the strongest zero-trust data security principles to meet the evolving real-world data protection needs of cybersecurity-obsessed and cost-conscious organizations. Offering an optimal balance of security, performance, and ease of use, it's the most secure, efficient, and simple backup target on the market.

II. Functionality

Functionally, ARTESCA provides comprehensive, enterprise-class object storage software capabilities:

- **Object protocols**
  - Amazon S3-compatible APIs, including standard bucket and object APIs, multi-part upload (MPU), bucket versioning, lifecycle management, replication and object locking
  - Security through AWS compatible identity and access management (IAM) capability. This includes AWS-style v4 authentication and policy-based access control
- **Lightweight, simplified deployment**
  - The system can be deployed starting as small as one single server (physical or VM)
  - Capacity can be expanded incrementally, to provide increasing levels of storage capacity and performance
- **Data durability**
  - Local repair codes speed up repairs of disk failures on a server by eliminating the need for network access during the rebuild process
  - Multi-server deployments utilize dual-level erasure coding with distributed network codes and local repair codes on each server
  - Replication (mirroring) of data to a remote ARTESCA for disaster protection
- **Multi-cloud data management**
  - Global namespace across ARTESCA, RING and public cloud storage (AWS S3, Azure Blob Storage, Google Cloud Storage) with a comprehensive UI
  - Built-in data management through lifecycle and replication policies to public clouds
  - Metadata search across ARTESCA, RING and public cloud storage

- **Flexible platform support**
  - Designed to support x64-compatible (Intel and AMD-based) physical servers or virtual machines (VMs)
  - High-performance all-flash or traditional hybrid SSD/HDD-based storage servers
  - Available as a dedicated hardware appliance for Veeam

III. Microservices, Kubernetes Architecture

ARTESCA has been architected and implemented using Kubernetes design principles and methods to fit natively into the deployment models and consumption patterns of these new applications. The system is designed as a set of distributed microservices delivered as containers at all layers of the software stack.

Most of these services are stateless to enable graceful scale-out and address increasingly higher workload demands and capacities. A limited set of services maintain the system state for the namespace, IAM entities, and in-progress data workflows — and in all cases, system state is stored in reliable fault-tolerant, scale out databases. ARTESCA services are deployed and orchestrated on Kubernetes, to provide automation capabilities across deployment environments, such as data centers and edge locations.

System services can be organized logically into several categories, as depicted on the stack diagram. These can be summarized as follows, with detailed explanations and diagrams in the linked sections:

Data services: A set of stateless services that provide object data access (Amazon S3 API) to applications and future object storage service provisioning operators for Kubernetes.

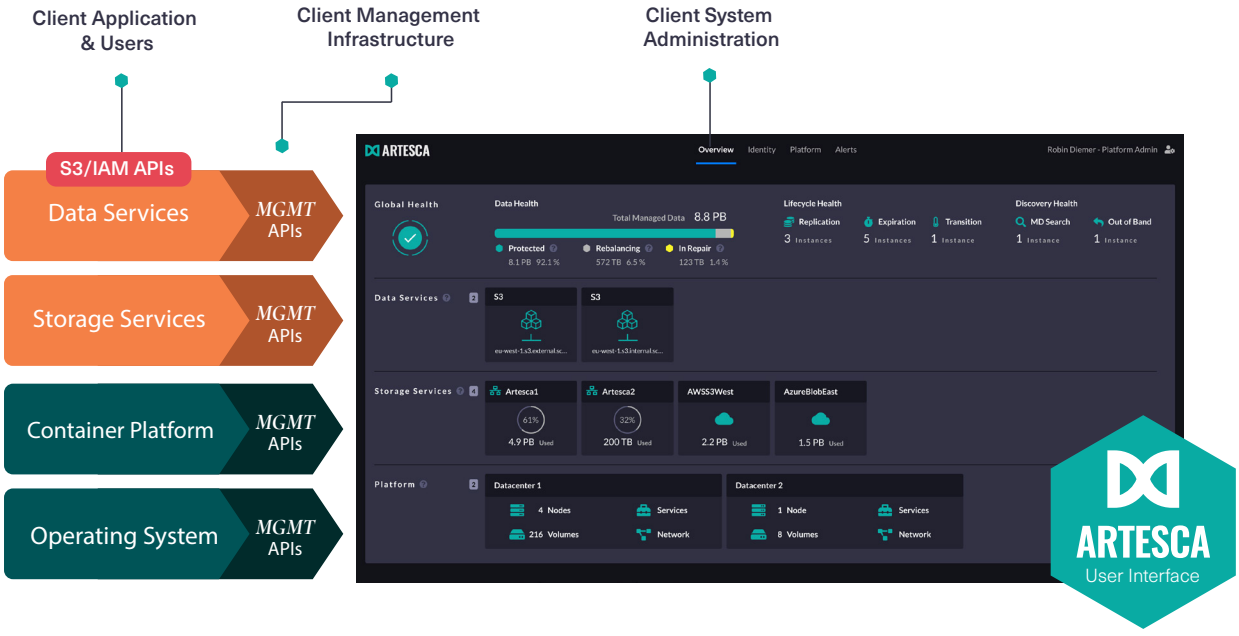
Common services: A set of services that are used across the system for functional areas such as identity and access management (IAM), metadata management and the management control plane (including system monitoring). These are the key stateful services in the system, providing persistent and consistent repositories for the system namespace, storage access and metadata search, managing user identities (IAM) and system health (control plane).

Multi-cloud services: For multi-cloud data data management policies, including communication and data movement between the local ARTESCA instance and remote target storage locations (other ARTESCA instances, RING instances) and public cloud targets. Multi-cloud functions here include lifecycle management and cross-region replication (CRR).

Storage services: Provides data durability and protection capabilities in the form of data replication and erasure coding, and interfacing to the native object storage layer through the underlying operating system. A key aspect for data durability on high-density platforms is a new dual-level data protection capability that provides network and local (server) level erasure-codes and object replication.

Platform services: The entire services stack is deployed and managed as containers on Kubernetes. The system is hardware-agnostic and can be deployed on x64 machines, but is supported for production on platforms that adhere to the reference architecture specifications (including minimum numbers of disk drives, SSD sizing for metadata) as described in the product documentation.

Management services and UI portal: ARTESCA provides comprehensive management and monitoring of the software and hardware platform stack. The system can be managed through the administration UI portal, a command line interface (CLI), and a RESTful API.



ARTESCA architecture stack

IV. Object Storage Capabilities

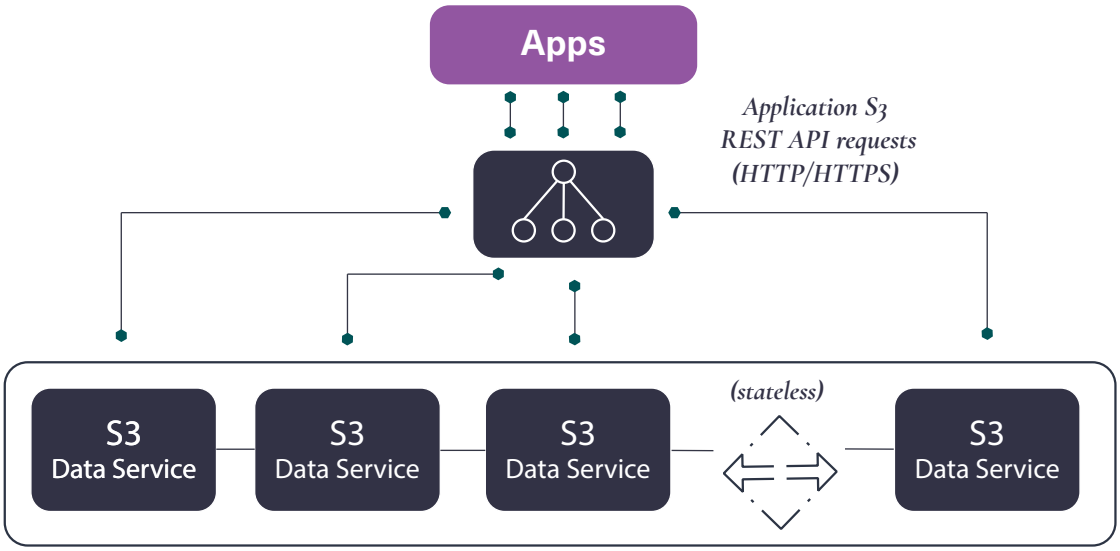
ARTESCA provides comprehensive object storage access capabilities over the popular Amazon S3 API. ARTESCA supports a wide range of S3 APIs including:

- Bucket and object operations — the foundational APIs of the S3 object model
- Multi-part upload (MPU) operations for efficient upload of very large data objects
- Advanced S3 APIs for object versioning, bucket lifecycle management, replication, data immutability (object locking), CORS, website and more.

ARTESCA adheres to published Amazon S3 API specifications — including headers and response codes — and extends the specification only in areas where additional value-added functionality can be provided. A full list of supported S3 APIs is provided in the product documentation.

S3 data services

The architecture implements a set of data services that provide S3 API endpoints for storage access by object-enabled applications. S3 data endpoints are stateless services accessed over HTTP/HTTPS through Kubernetes Ingress. On a multi-server distributed system, data services can be scaled-out to deliver higher throughput. For multi-server clusters, a standard setup can use DNS to round-robin requests across the S3 endpoints using a virtual IP address (VIP). Because the system can internally distribute traffic to available backends and because VIPs are always hosted on multiple servers, the system can move connections to provide high-availability, making an external load balancer optional. S3 data services receive S3 API requests and forward them to the required backend services, before responding to the application, as will be described in the following sections.



Stateless S3 Data Services & Load Balancing

Scale-out, stateless S3 data services

In the case of a single-server ARTESCA deployment, all S3 data services (and all system services in general) are co-located on a single machine host. More generally, as a distributed multi-server system, ARTESCA services — including S3 data services — are deployed on separate physical machines to provide fault domain tolerance in the event of host machine failures or outages. This stateless redundancy provides another element to ensure the system is resilient to failures, and maintains service availability.

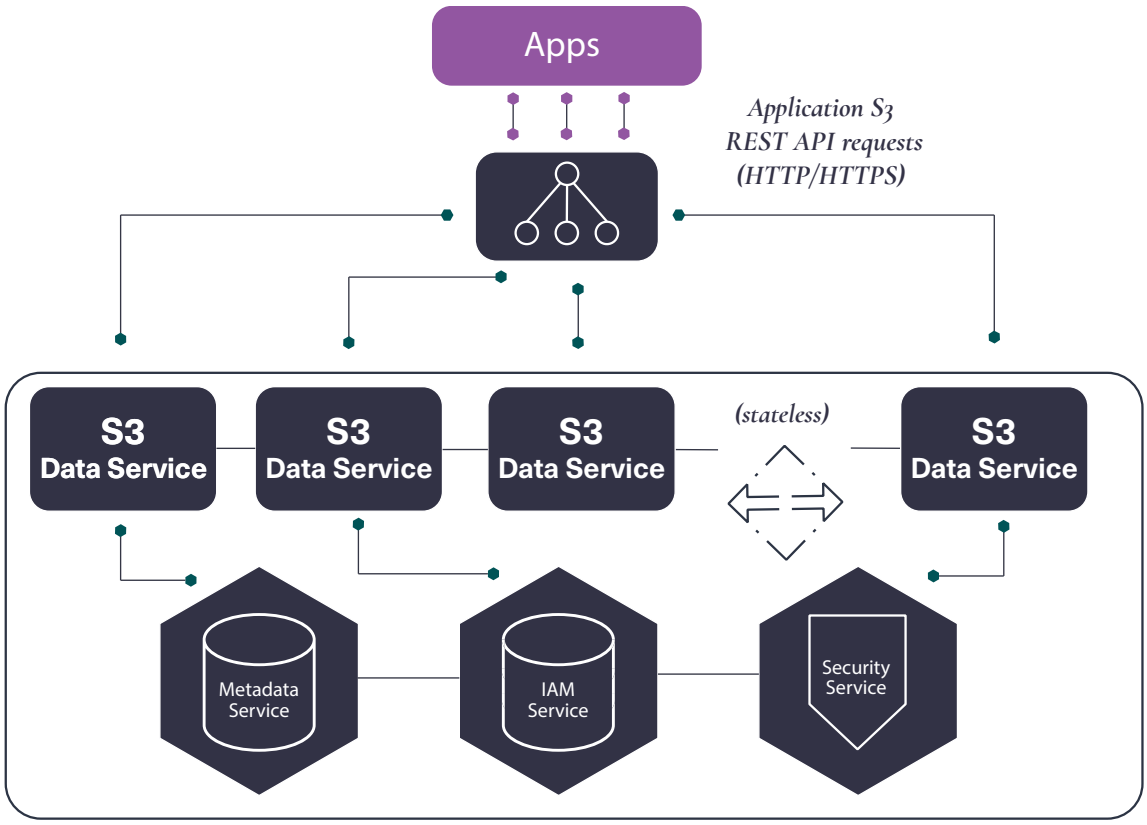
S3 data services interact with multiple other ARTESCA services as follows:

Metadata service: For all S3 operations that require bucket or object metadata for all S3 operations that require bucket or object metadata, this service provides a stateful repository for storing information on new objects, locating existing objects and metadata search.

IAM service: Stateful service for IAM objects such as Accounts, Users and Policies, for v4/v2 authentication, access control.

Storage service: Provides durable object storage for S3 PUT, GET and DELETE operations.

Security service: Data encryption leveraging industry-standard AES-256 bit encryption algorithms.



Stateless S3 Data Services & Common Services

ARTESCA S3 data services interaction with metadata, IAM & security

Data encryption

To ensure data privacy for highly sensitive data, ARTESCA supports data encryption at rest using industry-standard AES-256 bit encryption algorithms.This assures data confidentiality and authenticity are always maintained.

S3 data services manages data encryption to enable encryption/ decryption for buckets/objects on which it is configured. The encryption module is based on industry-standard AES-256 (256 bit) libraries, which can leverage special processor instruction sets to make the performance impact of encryption nearly negligible to the system and application. In order to maximize performance, the encryption keys are securely stored in the system metadata cluster.

V. Identity and Access Management (IAM)

In addition to support for the S3 API, ARTESCA also implements an AWS compatible identity and access management (IAM) service. This provides the ability to manage users and access control through the following mechanisms:

Accounts: One or more accounts can be created in a single ARTESCA instance to model multiple tenants, applications or use case workloads.

Users: Accounts can define one or more IAM users to represent applications. Identities have AWS-compatible access/secret keys (credentials) authenticated through the AWS signature v4 (HMAC) method. This is fully compliant with the Amazon S3 model of authentication to help preserve application portability and interoperability.

Groups: Users can be assigned to one or more groups to simplify access control through policies.

Roles: Can be defined, and identities coming from an identity provider (IDP or LDAP) can access the S3 API through assumeRoleWithWebIdentity.

Policies: AWS-compatible user, group and bucket policies provide powerful and flexible access control capabilities to allow/deny access on particular data, or enable full or partial (read-only) access to data.

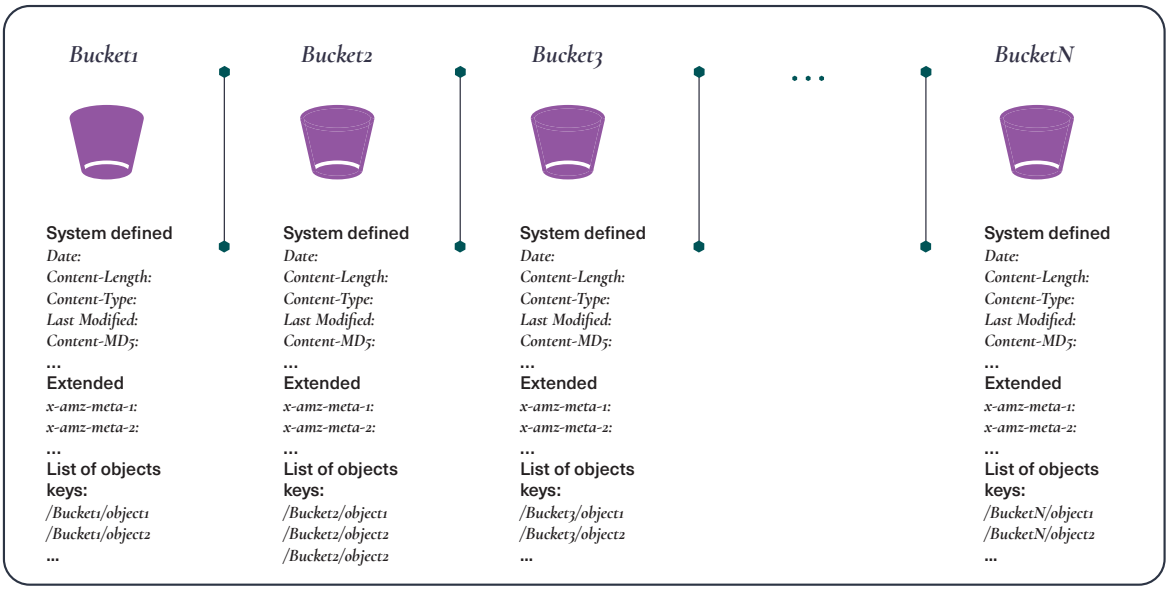
IAM state pertaining to users and other objects described above is stored in a reliable, replicated database that ensures high availability and consistency of the IAM stateful data. S3 data services make requests to the IAM service for operations, such as creating new IAM objects, and also for authentication of API requests using AWS v4 authentication based on the user's access keys. In addition, the IAM service is responsible for processing access control requests through stored user and bucket policies, which describe the level of access that should be provided.

In the case of a multi-server ARTESCA deployment, the IAM service itself is distributed and spread across multiple machine hosts for redundancy and fault tolerance. The database itself is replicated across the same machine hosts, with each copy instance providing access to the IAM objects. This ensures that failures or outages at the machine or services level do not impact availability for requests to the IAM service. The database is a distributed consensus cluster, with algorithms to ensure all instances of the database remain in-sync and a consistent view of the database is always used to service IAM requests.

VI. Metadata Service

The metadata service is the stateful heart of the ARTESCA system. It provides the namespace for data in the form of an Amazon S3-compatible bucket/ object view. Using S3 terminology, the namespace consists of buckets (logical containers for data objects), the objects themselves (the data payloads), system and user-defined extended metadata attributes. System-defined metadata attributes are assigned by default to objects by ARTESCA, and include:

- System-defined S3 object metadata: These include common attributes such as date/time, object size, content type, last modified date, MD5 checksum, version ID, and more.
- Optional extended (user or application-defined) attributes: Up to 2KB of key/value attribute pairs can be assigned to each object through the “x-amz-meta-”headers on a PUT Object S3 API call, and through the PUT Object tagging S3 API, respectively.



Logical view of metadata service database

Just as in AWS S3, ARTESCA supports various configuration options on buckets. For example, any bucket may be configured for website hosting, versioning, lifecycle management of objects in the bucket, asynchronous replication (CRR), object locking and more.

ARTESCA supports AWS S3 subresources for storing and managing bucket configuration information (a list of subresources supported by ARTESCA are listed in the product documentation). As described below (in the section Multi Cloud Data Services), ARTESCA extends the S3 concept of bucket location to define where a bucket's object data will be stored for configuration of lifecycle transition and replication policy target destinations.

ARTESCA uses the metadata service as a key accelerator for nearly all S3 command requests, to provide rapid lookups of object data through the database, and to avoid scanning through large amounts of data. Since S3 endpoint services make frequent requests to the metadata service for all S3 API command requests received from applications, best practice recommendations are to store metadata on flash (SSD/NVMe) media, for fast, low-latency response times. This accelerates common Bucket/Object PUT, GET, DELETE operations since they entail an update of the metadata database for a new or deleted object, or a lookup of the object location for a read access. Other operations requiring access to metadata include bucket listings, metadata searches through GET Bucket with a search criteria, and metadata accesses.

Extensible metadata and metadata search

ARTESCA supports metadata search on both system and extended attributes as described previously. Searches can be accessed by end-users through the ARTESCA UI, or programmatically through an extended GET Bucket S3 API that allows a search criteria on one or more attribute values.

Search conditions can include a range of predicates, such as standard predicates — equality/inequalities/greater/less than (=, <, <=, >, =>, !=) — plus compound conditions using AND and OR clauses (as with SQL), and supports standard data types including numerics, dates, characters and strings. The system processes metadata searches and returns the set of object keys within a bucket that match the search condition back to the requesting client application.

Both system-defined and extended attributes are searchable via the S3 API and the ARTESCA UI. See the ARTESCA API Guide for full details.

The ARTESCA UI also provides a user-friendly interface for adding/editing metadata to objects through the S3 data browser, as described in the management and UI section later in this document.

VII. Storage Services

ARTESCA implements advanced data durability protection through the capabilities in the storage services layer. ARTESCA provides the highest levels of durability by protecting data against a wide range of common and less common failures. This includes protecting data against loss or corruption due to common error and failure scenarios such as bit rot errors, disk drive failures, server failures, power outages and other component-level errors.

Storage services are designed to optimize durability on today's ultra-dense storage servers and for increasingly high-density flash media. Deployment is supported on standard x64 storage servers, including hybrid flash (SSD) / hard disk (HDD) configurations, as well as all-flash (SSD/NVMe/QLC) servers. Note that ARTESCA is also supported on virtual machine (VM) deployments with a similar amount of CPU, RAM and disk resources.

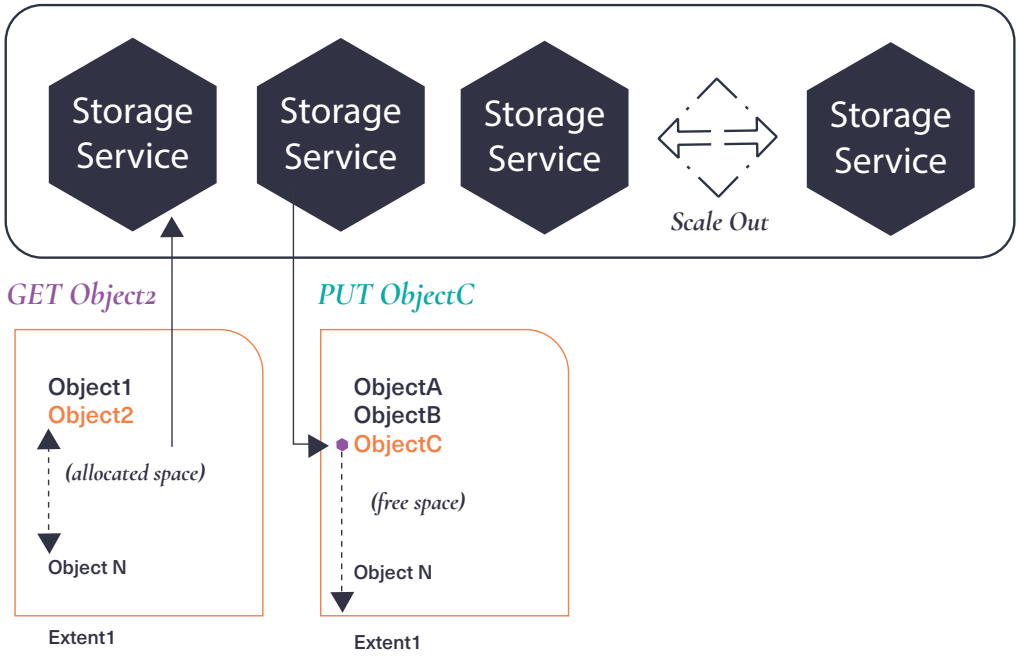
Innovations in ARTESCA storage services enable extreme data durability with special optimizations for high-density media (HDD & SSD). Fast repair times (low mean-time-to-repair) after common disk failures is a key benefit of ARTESCA storage services, thereby providing super high data durability. ARTESCA storage services provide the following mechanisms to ensure data is stored efficiently and protected:

- **Data extents:** Data is stored on disk in fixed-size files referred to as extents. By grouping many smaller objects into extents, this avoids fragmenting the disk file system and speeds IO performance. Extents are written to in an append-only mode, which increases overall performance by minimizing random IO.
- **Data replication:** Multiple copies of objects stored across multiple different disk drives (and physical machines if available). This is optimal for smaller data objects, where the overhead of multiple copies is a better tradeoff than erasure coding. Objects below 60 KB are by default stored as replicas by default.
- **Dual-level erasure coding (EC):** For multi-server deployments, one of the main techniques to maximize data durability is an innovative combination of network (distributed) erasure codes and local repair codes (within a server). This provides both additional protection against failures, as well as accelerated local (non-network) repair times for disk failures within a server.

- **Checksums for data integrity assurance:** Extents and objects on disk are stored with a corresponding checksum to ensure that any data that is read is the same as when it was stored, ensuring data integrity.
- **Background disk scrubbing:** Extent level checksums are scanned ("scrubbed") by a background process to ensure that data is not changed due to unrecoverable bit errors (sometimes termed bit-rot). The system also validates object-level checksums upon reading the data to verify that data integrity is maintained. In the event of a checksum mismatch, the system can restore the original data from other data/parity stripes or replicas in the system.
- **Self-healing:** Background processes monitor system and component health, and can respond in the event of failure conditions to restore/ repair the system automatically.

Disk file system and extents

ARTESCA uses fixed-size (128MB) extents in the disk file system to store object data payloads. As objects are created, they are allocated space in the current open extent and written sequentially to the available space. As extents are filled, they are closed — and new extents are allocated. Small objects will mainly fit into the available space in the open extent, whereas larger files may be split and stored across multiple extents.



Storage layer with disk extents (container files) for data objects

Write caching

In storage servers with spinning disks (HDDs) for data, ARTESCA uses write caching to accelerate writes for small objects (files). The use of a write cache has benefits for response time latency for small payloads, as it defers writes to the HDD layer until the extent is filled.

Single-server data protection

ARTESCA can be deployed on a single server (machine host), which makes it suitable for small capacity configurations, such as in edge applications. To provide durability, ARTESCA utilizes local codes to protect data against loss from disk failures. In the case of a single machine deployment, local repair codes are used alone without the network repair codes. In this scenario, additional parity stripes protect against two simultaneous disk failures, using a default 8+2 EC scheme. To ensure stripes are stored on separate disk drives, ARTESCA requires at least 12 drives to be available on the server to accommodate data and system requirements.

In the event of a disk failure, the system will self-heal by rebuilding the stripes that were on the failed drive to other available disks in the system, using the surviving copies (replicas) or EC stripes still stored on those drives. There is no concept of an idle “spare” disk drive needed (such as in a RAID system). Instead, ARTESCA can use any available disk capacity in the server for rebuilds (or more generally across the cluster in a multi-server deployment).

Note that a single machine deployment can be geo-replicated using CRR to provide a remote disaster recovery copy of the data on another ARTESCA instance, RING, or public cloud target.

Multi-server deployments with dual-level data protection

An ARTESCA system can be deployed across multiple servers to provide increasing capacity and performance. By distributing services and data across servers, a multi-server deployment offers additional benefits in terms of data durability and availability than a single system. This provides fault-tolerance and eliminates the server itself as a single point of failure.

In a multi-server deployment, ARTESCA services employs a dual-level data protection scheme consisting of:

- **Network erasure-codes:** The well-known M+N data protection model (M=data, N=parity) of stripes distributed across physical servers within the multi-server cluster. Network codes provide very high durability in many server and cluster configurations by protecting against a defined number (N) of server failures. For smaller objects, erasure coding is a less efficient scheme. In that case, ARTESCA stores replicated copies of the object across servers to ensure that the loss or unavailability of a server still provides access to another copy of the object.

- **Local repair codes:** To optimize data durability on the latest high-density (PB+ capacity) servers, a new innovation provides a lightweight (very low storage overhead) repair code stored on each physical server machine and distributed across several disks. These lightweight local codes provide sufficient information to repair local disk failures without needing to access the network codes across the network. This accelerates repair times by avoiding network IO and localizing the repair tasks.

In general, ARTESCA employs a dual-level data protection strategy using two schemes, based on the size of the object being stored:

- Erasure coding for protection of large data objects (the default configuration is for objects over 60KB in size)
- Replicated copies for smaller data objects (objects below 60KB in size)

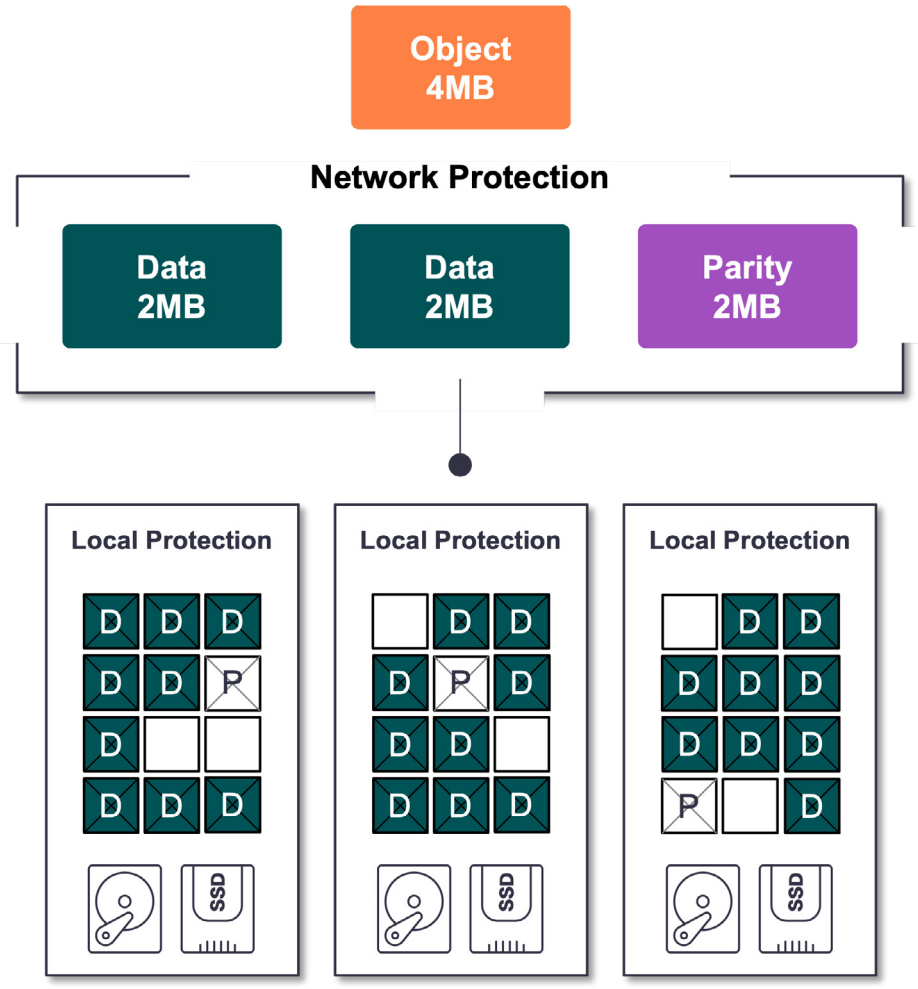
In either case, the system intelligently distributes erasure coding stripes or replicated copies across servers in the cluster, assuming multiple servers have been deployed. This ensures that a failure, outage or unavailability of a server — and the stripes or replicas it stores — does not affect the overall data availability or durability.

In the case of erasure coding, objects are first split and then encoded using standard encoding libraries for computing parity stripes. Erasure coding libraries take advantage of ISA-L and AVX-512 instruction sets in Intel® and AMD® processors to accelerate these calculations. The system intelligently distributes the EC data and parity stripes across physical servers and disks to ensure that a single failure does not impact multiple stripes.

When a failure occurs requiring this network EC information, the system utilizes distributed background services to perform the repair process in parallel. For example, if a complete storage server failure occurs (an increasingly rare event but one that must be planned for), the full performance of the system is used to repair the lost data stripes as fast as possible in a distributed manner. A complete server repair requires the system to find sufficient unused capacity in the cluster on the remaining servers. This process uses multiple distributed background processes running in parallel — and also uses space across the entire cluster to avoid creating IO bottlenecks on specific system resources. This helps to restore data protection to its original (nominal) state as fast as possible, thereby maximizing data durability. Note that a similar capability applies to the case of rebuilding distributed copies (replicas) across servers for small object data durability.

To maintain space efficiency in a multi-server scenario, local repair codes are reduced in overhead as compared to the single-server scenario. Effectively, the local repair codes are computed with a reduced-overhead scheme to maintain a reasonable tradeoff between data durability and storage space. The total combined overhead of local data protection, plus the overhead of the network data protection, is maintained at a comparable level to standard data protection schemes. In larger system configurations the total overhead is actually lower than standard 8+4 (data+parity) network erasure coding schemes (specific capacity overheads are detailed in the ARTESCA product documentation).

The innovation in ARTESCA's local repair codes is its ability to protect against a very high percentage of local disk failures with a very low space overhead. In the rare case that the local repair codes do not have sufficient information to localize data rebuilds, the system will use standard network rebuild techniques. In that scenario, the system defaults to a normal distributed rebuild mode, and never sacrifices data durability or the ability to self-heal after a disk failure.



Local protection (local EC) and network protection (distributed EC)

Self-healing services

ARTESCA provides background services to ensure that even if failures occur, the system will revert to its nominal protected state (assuming sufficient resources are available). In contrast to legacy RAID schemes, which depend on idle “spare” disk drives for repair, the ARTESCA system can utilize any available space in the cluster pool to perform a repair. As long as there is sufficient unused disk capacity in the system, a failed server or failed disk drive does not need to be immediately replaced and can be replaced at the next service interval. The system will automatically restore the lost information to other parts of the cluster, as will be described next.

VIII. Platform Services

ARTESCA software can be deployed on bare-metal servers or VMs, with an integrated Kubernetes and Linux Operating System. Kubernetes is leveraged for deployment automation, cluster management and runtime orchestration of ARTESCA services. An ARTESCA system can grow capacity through the addition of disk drives (in existing servers), or the addition of new storage servers. New server or storage resources are added to the Kubernetes cluster that can then be used by ARTESCA services.

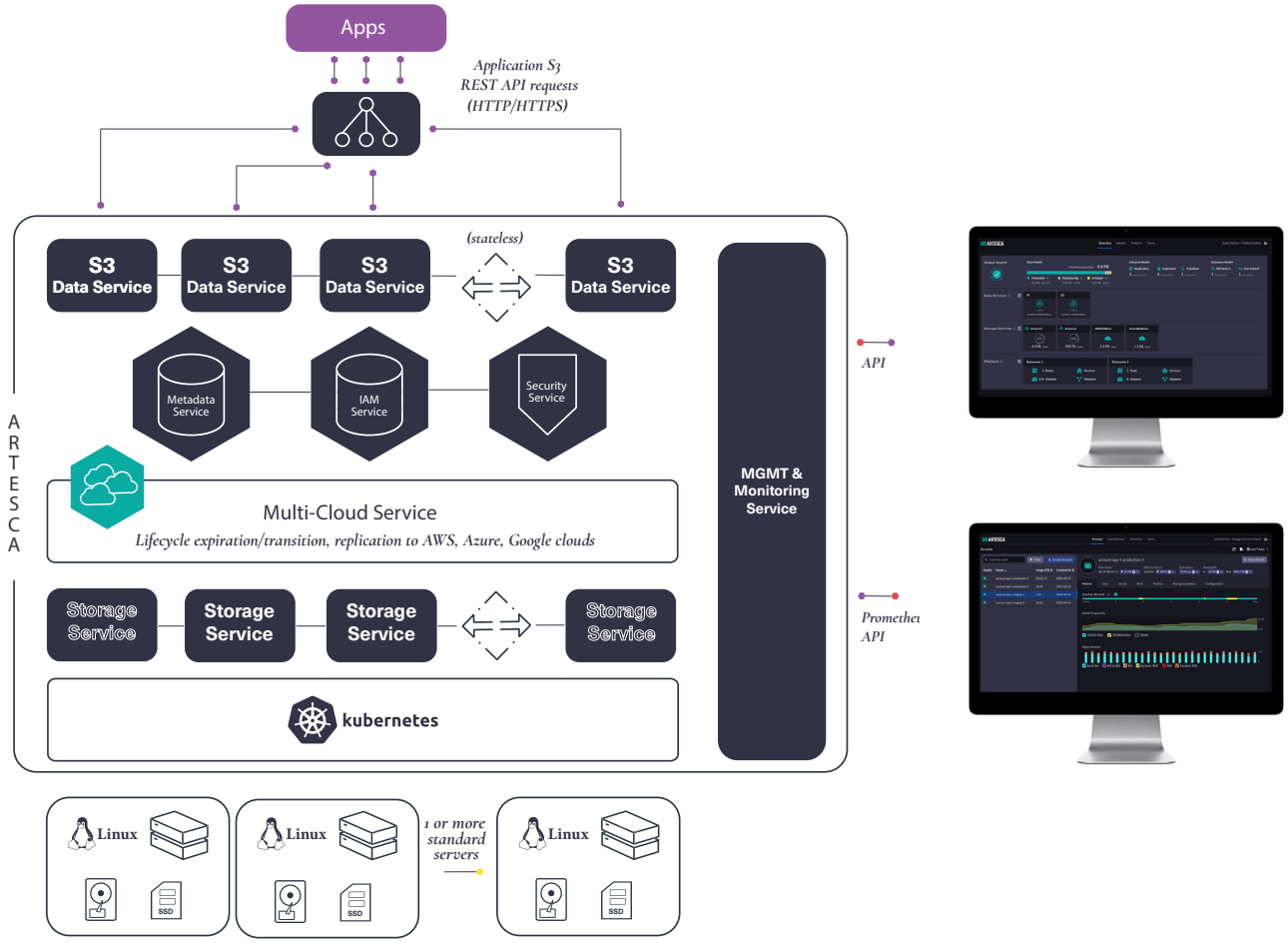
ARTESCA provides an integrated Kubernetes distribution. Detailed installation and deployment descriptions are available in the ARTESCA product documentation.

IX. Management Service & ARTESCA UI

ARTESCA provides management and comprehensive observability services, which provide a superset of monitoring capabilities for the entire set of distributed software and platform services as well as the underlying hardware platform:

- Alerting
- Logging
- Dashboards

The management service is accessible through a set of RESTful APIs, including a Prometheus-compatible monitoring API. These APIs are utilized by the ARTESCA portal, a web-based graphical user interface (the ARTESCA UI).



Management service, UI & APIs

An overview of the ARTESCA UI is provided below (a complete tour and user perspective of the UI is provided in the ARTESCA product documentation). In addition to management and observability capabilities outlined above, the UI also supports the multi-cloud namespace capabilities including a dashboard view of managed backend storage capacity across deployments of ARTESCA, RING, and public cloud storage services.

The UI is enabled for role-based access control (RBAC) with predefined administrative identities, each with default permissions for certain tasks (note that users can have one or more of these roles assigned):

Storage Manager: Effectively a super admin, this role is responsible for managing top-level accounts within a single or multi-tenant system, and creating account owner admins.

Platform Admin: Responsible for managing and monitoring hardware resources, such as servers, disk drives, networks.

In addition, there are predefined IAM roles (and additional custom roles/ policies can be defined):

Storage Account Owner: Manages an account, with the ability to create account-level objects such as users, groups and policies within the account. Controls who has access to different buckets and data objects for the account.

Data Consumer: A role for accessing storage via the S3 API either as an individual user or an application identity. Data consumers can create and delete buckets, change bucket attributes, read, write and delete objects within those buckets.

Multi-factor authentication (MFA) adds an extra layer of security to ARTESCA users logging into the UI. It is a security process which requires a second form of verification before gaining access to the system via the use of a one-time code generated by third party authentication apps.

MFA reduces the risk of unauthorized access from password theft, phishing attempts, and keyloggers because they would still need the additional factor to gain access. MFA also protects against man in the middle attacks and credential stuffing where lists of usernames and passwords from previous breaches are used.

Once enabled, the next time that user logs in they are presented with a one time authenticator setup screen. From here they enter their selected third party app and scan the QR code provided on the ARTESCA screen, once confirmed by entering the displayed code, this user's account is now protected with MFA.

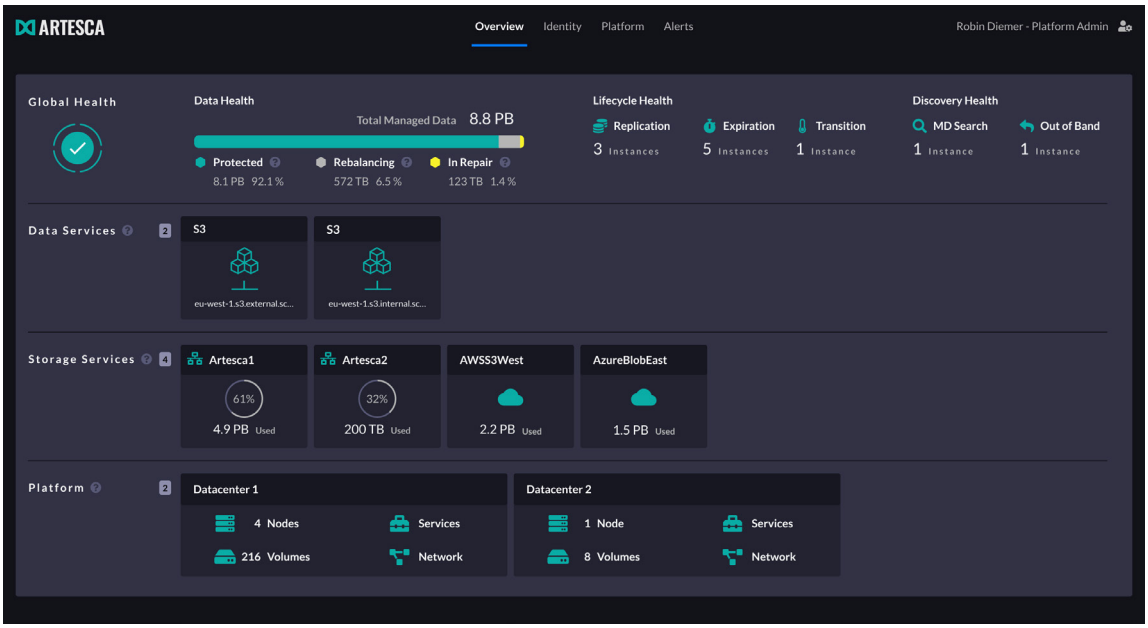
On next login the user enters their email and password followed by the code displayed on their authenticator app.

MFA can be enforced at either the individual user or at a system wide level for all users, from the Global Settings menu. In this case, the MFA feature cannot be disabled or removed at the user's level. Users can also be forced to register a new device on the next login.

To increase overall system security MFA is recommended to be enforced globally.

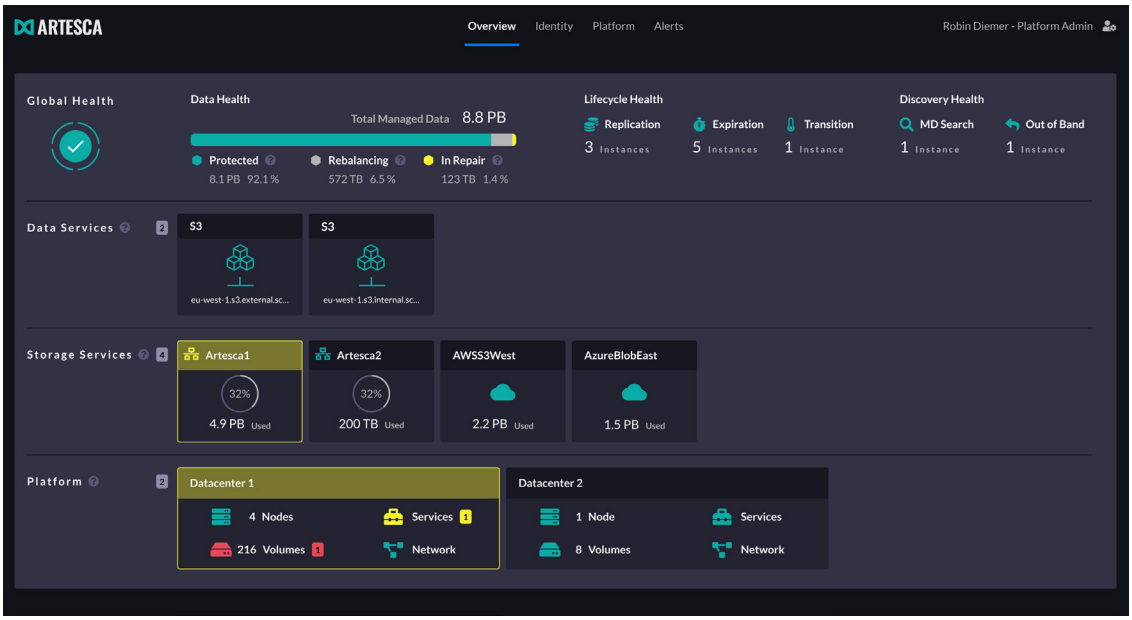
The ARTESCA UI visualizes the system in three logical layers:

- **Data access services:** The S3 API endpoints described previously, configured into service domains.
- **Backend services:** The multi-location and multi-cloud layer, showing managed capacity across storage locations and public clouds. The example image above shows two ARTESCA instances (one local and the other remote) and two public cloud storage locations, one in AWS and the other in Azure.
- **Platform:** The physical infrastructure layer, depicting the system structure across servers, including disk resources. This also provides views of key performance indicators (KPIs) of networks, servers and disk drives.



ARTESCA UI overview screen

Alerts and warnings are flagged on the dashboard through color codes on the various layers of the UI, including a detailed drill-down view of all alerts.

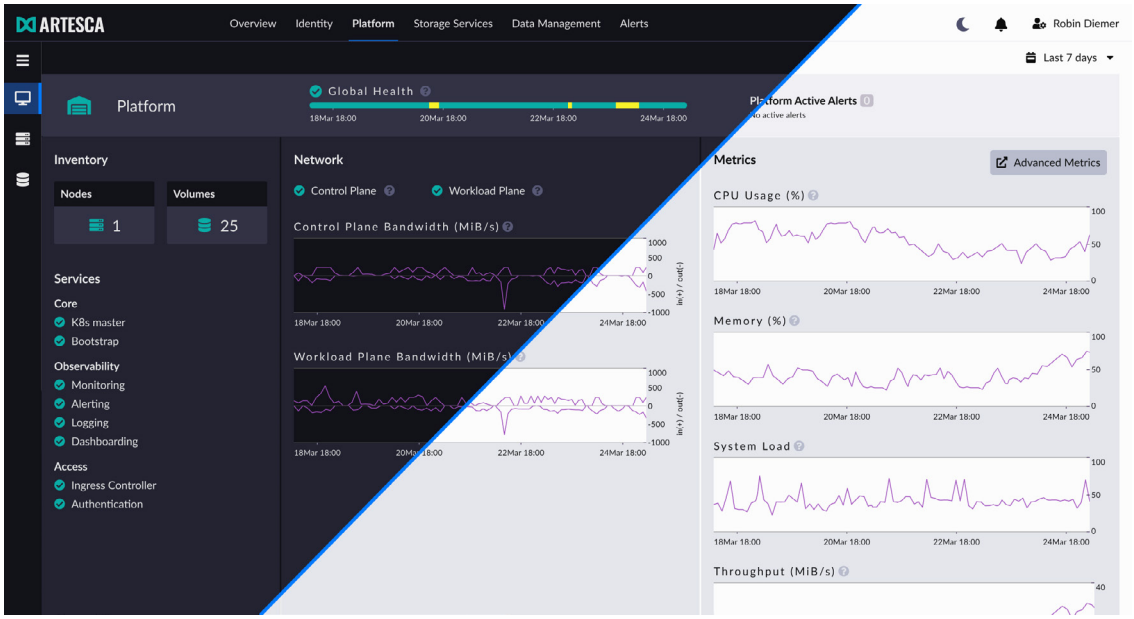


ARTESCA UI displaying alerts

All areas of the top level dashboard provide further access to details through point-and-click to the relevant sections. Below the top-level dashboard, the UI provides graphical access for the following ARTESCA functionality:

- **IAM Service Management:** For management of IAM objects such as users, groups, roles & policies, as well as viewing storage utilization analytics (KPIs).
- **S3 Data Browser:** An end-user browser for creating user buckets, upload/download of data objects, metadata inspection and editing, and interactive metadata search.

The ARTESCA user interface is available in both dark and light themes. It is simple to switch between the two to match your personal preference by clicking either the sun or moon context sensitive icon on at the top of the UI next to the notification center.



ARTESCA UI showing both dark and light themes

ARTESCA will notify you if there is a new version of software available, this is displayed in both the alerts notification center and by clicking the' about' section which will show the current version, end of support date and the new version number. There will be a link to the release notes containing more details of the upgrade's new features and showing the simple, streamlined upgrade procedure which updates the operating system and ARTESCA software.

Data management policies

The UI provides management and monitoring of lifecycle and replication policies. As described earlier, ARTESCA supports policies for data management across all supported storage systems and public cloud storage services. An ARTESCA deployment can have multiple simultaneous policies defined on different buckets.

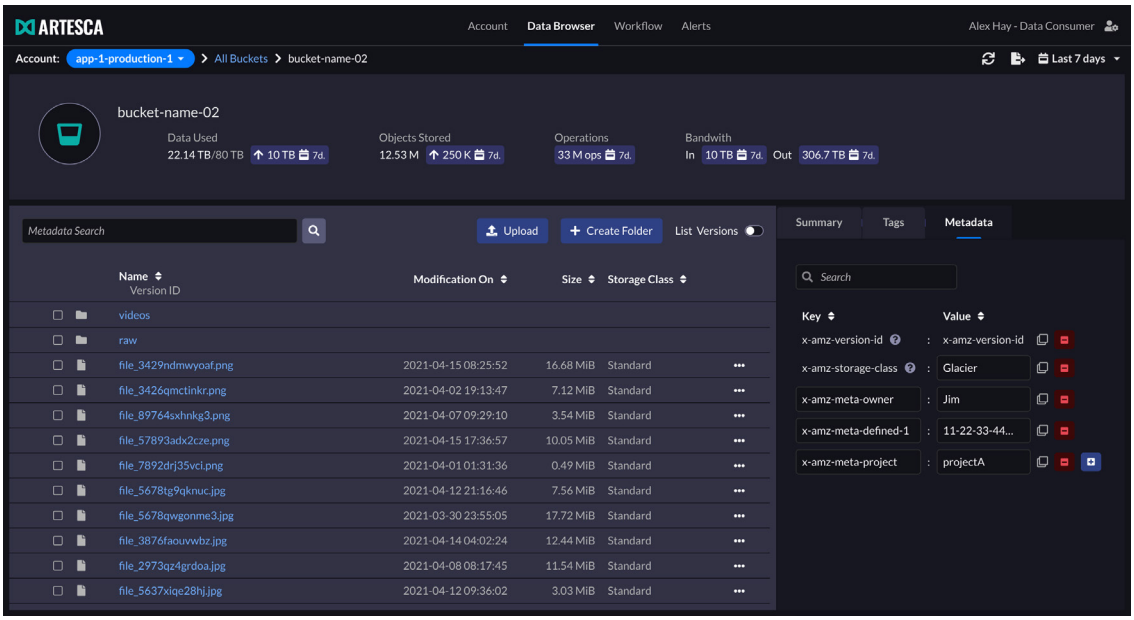
The UI also provides the capability to define new policies across the various storage locations, and for monitoring the status and performance of live policies. The system displays the replication status, rate, speed and any errors reported for an ongoing replication policy.

Analytics and capacity planning

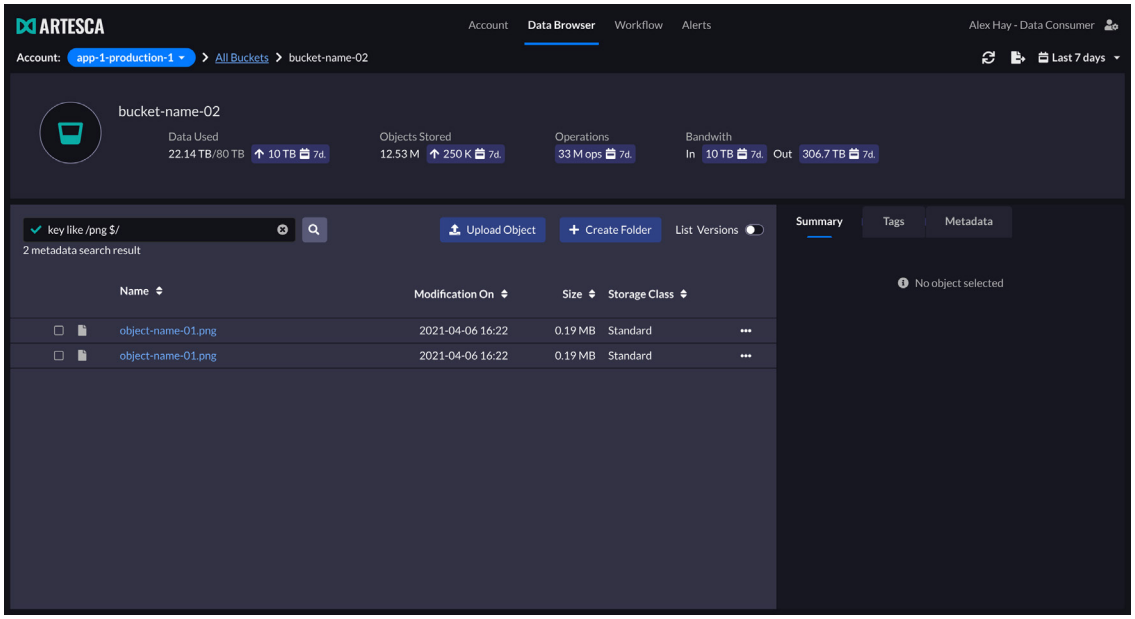
The UI provides utilization analytics and capacity planning metrics that are tracked by ARTESCA. Utilization data includes capacity-based (number of objects, storage consumed) and performance-based metrics (S3 ops/sec, bandwidth in/out). The UI also provides time-based trending graphs and visual meters for these metrics, which can aid in overall system health monitoring and capacity planning.

S3 data browser

For simple data browsing of user object data, the UI provides an integrated S3 data browser. This allows permitted users within an account to view existing buckets, create new buckets, and upload/download data objects into buckets. Moreover, the UI enables browsing of system-defined metadata attributes on each object (to define new metadata attributes and tags on objects) and provides a user-friendly interface for interactive search based on metadata attributes. This interface provides a simple human-friendly interface on top of the API-based access to metadata search described earlier.



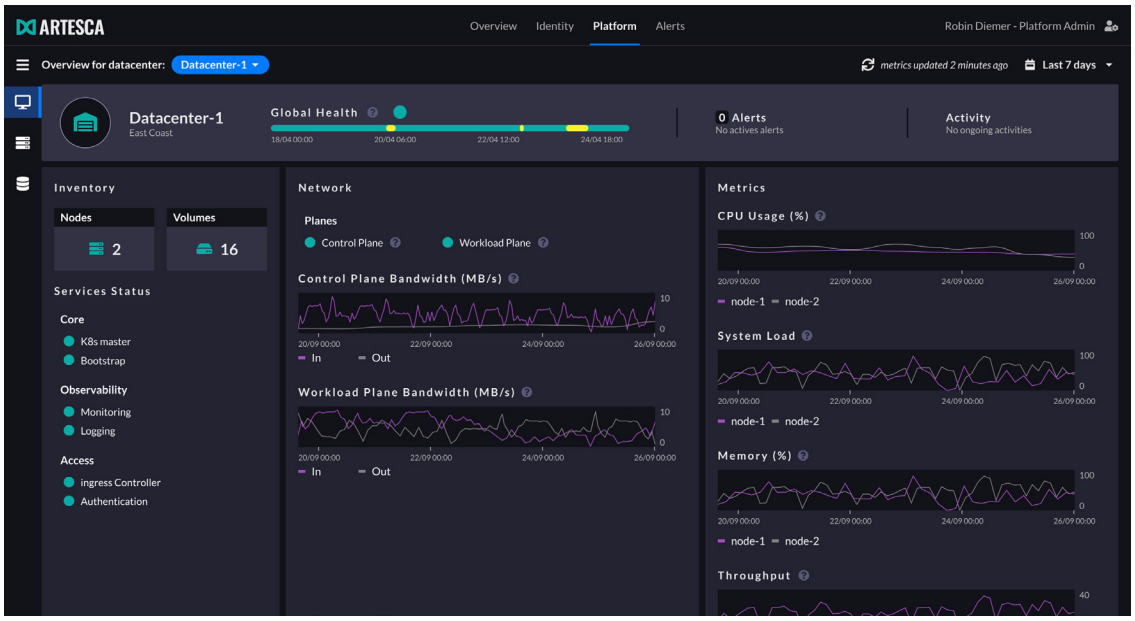
S3 data browser metadata editing



S3 data browser metadata search

Platform monitoring

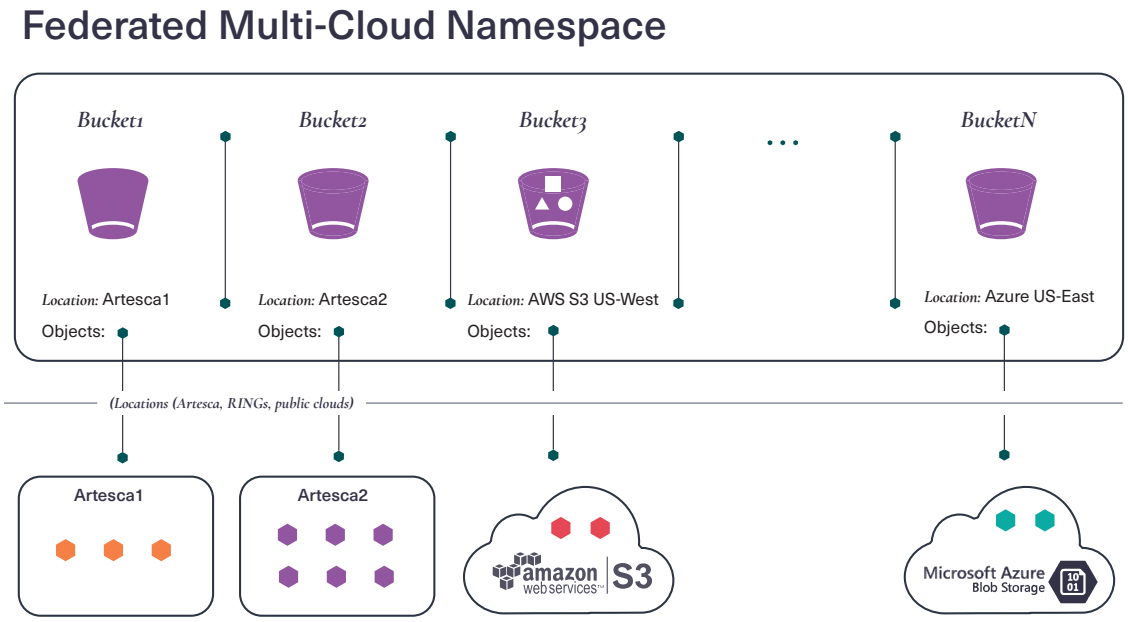
The UI supplies detailed monitoring of the underlying platform layer, both at the Kubernetes and physical hardware levels. This provides utilization metrics on the platform (CPU and memory), health checks and performance metrics. Monitoring views are provided per data center, per server and per disk — and also include the status of network interfaces. Below is a data center view of the hardware platform, including bandwidth, CPU, system load and memory utilization metrics.



Platform monitoring view with health and utilization metrics

X. Multi-Cloud Data Management Services

A set of services enable multi-cloud (more generally, multi-location) data capabilities. Starting with the metadata namespace described previously, ARTESCA provides the ability of that namespace to span across multiple “location” configurations. ARTESCA provides a multi-cloud namespace, data management policies, and metadata search — all automated by an asynchronous workflow processing engine.



Multi-cloud namespace with buckets in multiple locations

ARTESCA is multi-cloud aware through the concept of S3 Bucket storage locations. In ARTESCA, a storage location describes a combination of endpoint information (such as a URL for cloud storage) and access credentials to the specific account/bucket/container (if required) to allow access to the system or service. ARTESCA supports storage locations to ARTESCA, RING (over S3), AWS S3, Azure Blob Storage, Google Cloud Storage.

The ARTESCA namespace becomes a multi-location/multi-cloud global namespace through these location descriptors. It also enables policies for lifecycle management and cross-region replication (CRR) between ARTESCA and public clouds (or other combinations of sources and target locations).

Note that ARTESCA always writes data to storage systems and cloud services using the native API and data models, to ensure that data is directly accessible and consumable from services or applications using that storage. ARTESCA essentially translates and maps application S3 API commands to data services into corresponding commands for the backend storage (for example, using the Azure Blob API for the Azure Blob storage service). That is, there is no proprietary data format that prevents access to the data without using ARTESCA, or creates any form of lock-in.

In addition, assignment of extended metadata attributes and metadata search is supported on external locations, including all supported public cloud services (whether or not they support extended metadata). This means an application can perform a search on any ARTESCA bucket independent of its location, and without dealing with differences in data or format based on where the data is stored.

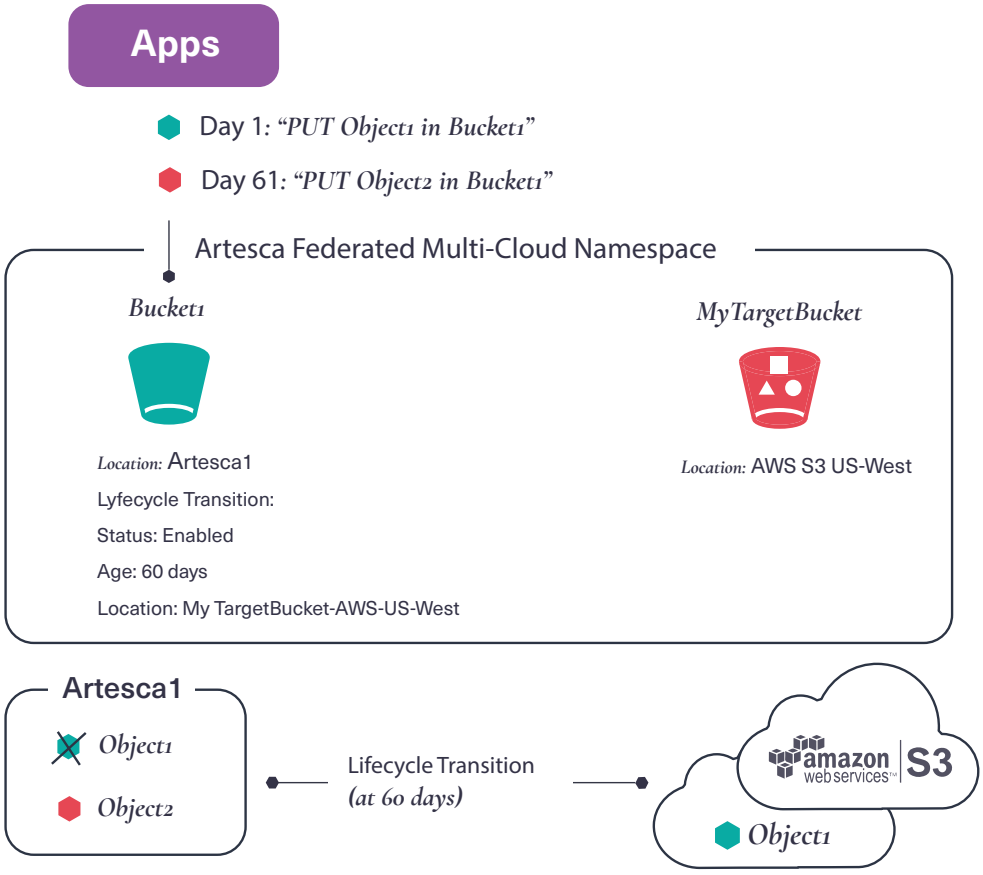
Lifecycle management

ARTESCA supports lifecycle management as per the AWS S3 Bucket Lifecycle API. This is a per-bucket policy that can be configured by the application or system administrator, with rules for managing lifecycle actions for expiration and transition of objects. A single system can have buckets that are lifecycle enabled and others that do not have lifecycle enabled. A bucket that is lifecycle enabled must also have bucket versioning enabled, as per the S3 specification.

Expiration follows the policy rules set by the AWS S3 API specification, based object tags, prefix and age. More specifically, lifecycle expiration enables automatic deletion of objects within the bucket based on the age of the object or reaching a specific configured date, or if tags are used in the configuration then the rules apply to only objects that match the tag values.

Lifecycle transition is defined as a “move” of the object data from a source location to a target location. This enables objects to be physically transitioned between the two locations, for example, from the original ARTESCA instance to another ARTESCA instance, or to a target public cloud service. Note that when an object is transitioned, the data is moved but the Object URL remains unchanged to maintain transparent access.

The reasons for managing transitions can be driven by application or business requirements to move data to a different geographic location, a lower-cost platform (ARTESCA instances can be deployed on higher or lower density/cost platforms, for example), or to move the data to storage in the cloud regions. ARTESCA supports different source and target locations for each bucket, making it possible to have different lifecycle rules for different data sets, to support different business needs.



Lifecycle transition from ARTESCA to AWS bucket

By supporting the concept of generalized locations, ARTESCA makes lifecycle management more powerful, and enables users to take advantage of multi-cloud to optimize data across regions, services, or in terms of price/performance. Lifecycle management on buckets can be configured and monitored through the ARTESCA UI, as described later in this paper.

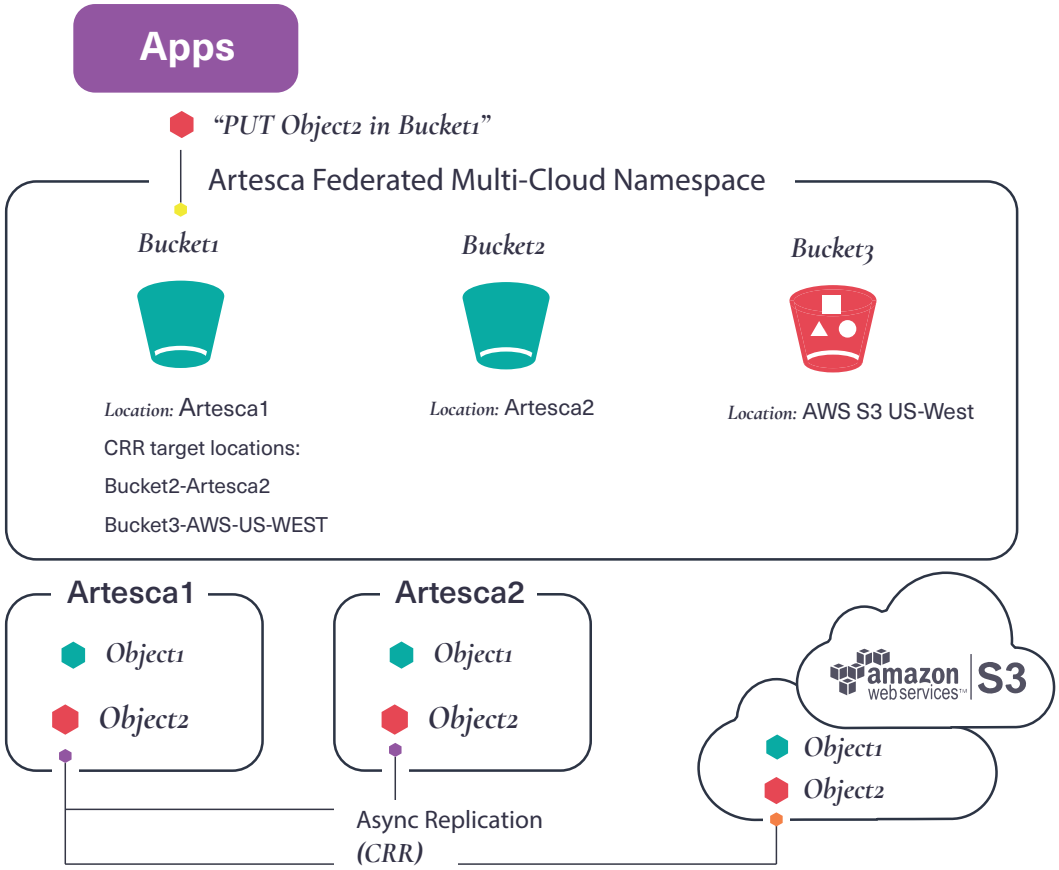
Asynchronous replication (CRR)

CRR is designed to support replication over higher-latency wide area networks (WANs) to provide business continuity and disaster recovery (D/R) solutions. To tolerate these higher network latencies, CRR is an asynchronous mechanism. That is, the system does not block the application — instead, it responds to the application's API request first before CRR processing continues. ARTESCA supports asynchronous replication in a model similar to the AWS S3 cross region replication (CRR) API. Asynchronous replication can be used to create a remote copy of data on the local ARTESCA system for disaster recovery purposes.

As described in the next section, replication events are queued and then processed asynchronously by the workflow engine. This allows the application to continue processing while replication continues in the background. ARTESCA also manages retries of failed replication attempts, in the case of network outages or other errors, and manages local caching of the object for multi-target replication in the event one or more of the replication streams fail. Asynchronous replication on buckets can be configured and monitored through the ARTESCA UI, as described later in this paper.

Replication is a per-bucket policy that can be enabled on selected buckets and not on others. CRR configuration on a source bucket can be done through the ARTESCA UI or programmatically through the S3 API. The target location of CRR is specified when configuring CRR on the bucket. Supported locations are the same as with lifecycle management — any ARTESCA instance, RING instance, or public cloud service regions are supported. One-to-many replication is supported whereby a single bucket can be replicated to multiple locations from a single policy.

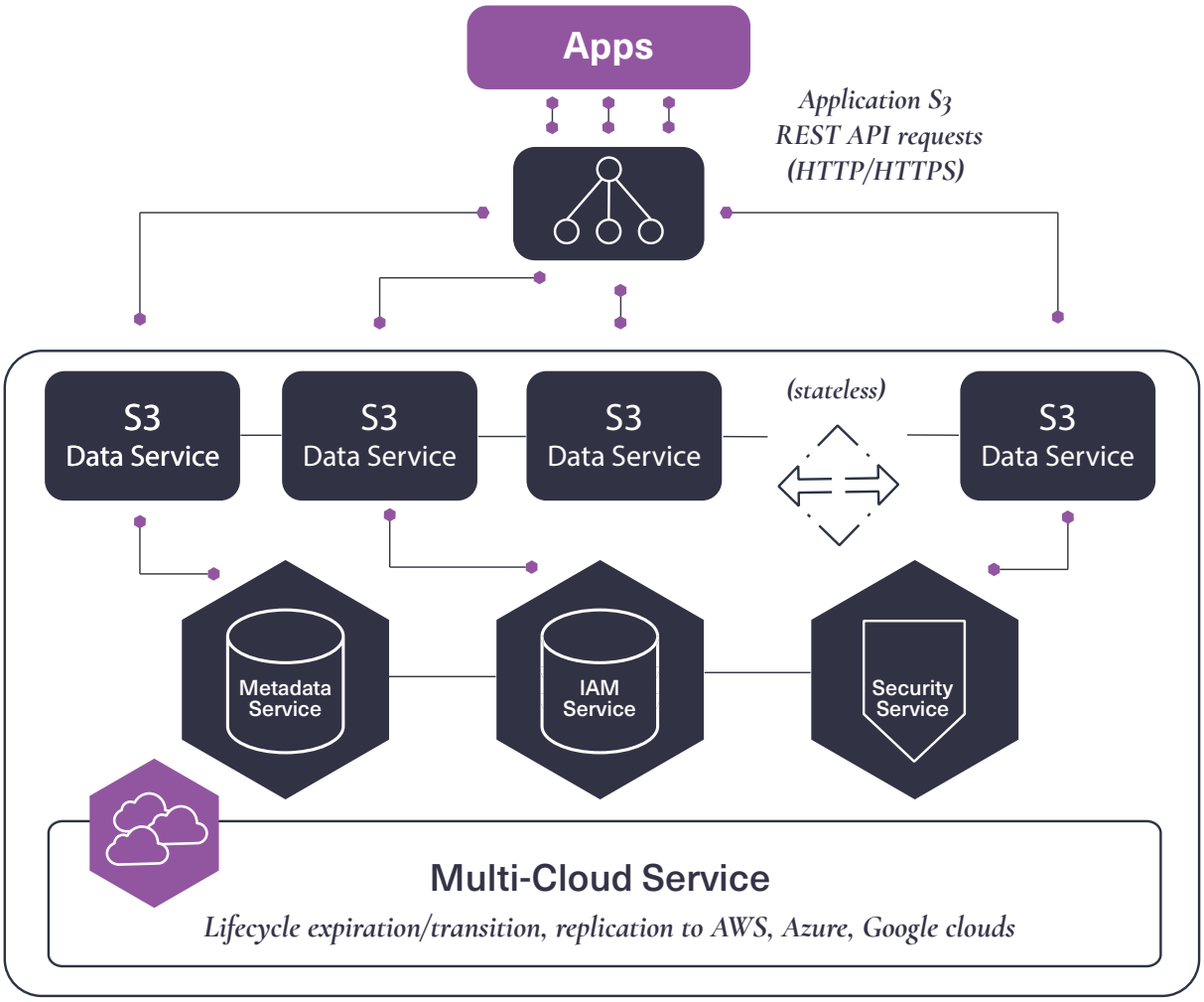
CRR can enable a variety of data management solutions for use cases such as content distribution across multiple data centers, data bursting from on premises to cloud, or data capture for edge applications with replication back to the core data center.



One-to-many CRR from local ARTESCA bucket to remote ARTESCA and AWS buckets

Multi-cloud workflow services

ARTESCA provides workflow processing services for managing data across multiple locations and clouds. As was described, ARTESCA currently supports lifecycle management and cross-region replication as integrated data management policies. Over time, data management policies may be extended in future releases, but the same general processing description provided here will apply.



Multi-cloud services

As a multi-location and multi-cloud manager, the system maintains information about the target locations (as described previously). To provide secure access to remote locations, ARTESCA requires access to the target's authentication credentials. Remote system and service credentials are captured and stored in encrypted and secure form using industry best practices. For example, private keys are not visible to the system administrator after being assigned. These access credentials allow ARTESCA to perform actions such as creating new objects on the target storage systems or clouds through secure HTTPS connections.

To manage workflow policies, multi-cloud services provision a series of queues that are associated with the corresponding buckets and events on them. The system also manages a set of background producer and consumer jobs to process events in the queues asynchronously. This parallelization of work across multiple queues makes the system scalable for processing on large numbers of buckets and policies and for high ingest workloads.

A key aspect of the workflow queues is that they store the command events but do not store the associated object data. Instead, the queue entries reference the actual state of the data in its corresponding bucket. This keeps overhead low by eliminating the need to duplicate the data in the queue, which would lead to higher resource consumption and performance impact.

Since all lifecycle and CRR-managed buckets must be versioned (this is a prerequisite of the S3 API), queues will always be able to reference both current and older version states of the data. Object data referenced by the queue will be used when the workers process the queue, for example, to replicate the actual data to a target location. By not storing object data into the queues, this also maximizes the ability for ARTESCA to buffer changes locally in the event of a network outage or failure that halts CRR or lifecycle transitioning processing. As long as the system has sufficient capacity to store events into the workflow queues, normal operations can continue during a network connectivity failure. When network connectivity to the target locations is restored, background processing will automatically commence data traffic to the intended targets.

As applications make API requests to the system, or as policy rules are triggered, a new event is created in the corresponding queue. Queues contain information about the requested API change action and the corresponding object key for the action. In the background, asynchronous worker jobs process the queues in a first-in-first-out (FIFO) manner to preserve time ordering and consistency on the target. For CRR policies, the system creates queue entries for change actions such as PUT requests, since these will need to be replicated to the target location. During queue processing, a PUT event on the source will be directly replicated through another PUT to the target location (using target credentials, secure authentication and adhering to the target's access control policies).

The system depends on network connection to be maintained between the local ARTESCA instance and the target location (or cloud) as well as sufficient bandwidth to maintain throughput for large object transfers. Because network failures and timeouts can and do occur, ARTESCA automatically manages retries of failed replication events.

Retry management is based on a predefined number of attempts with an exponentially increasing back-off time window between retries. Once the maximum number of retries has been exceeded, the system fails the operation and indicates that in its status. The CRR status can be monitored through the ARTESCA UI or queried with a HEAD Bucket API command, which returns a flag indicating in-progress, error or completion status. The system UI provides metrics (key performance indicator) displays to help monitor these data management workflows.

XI. Deployment and Reference Architectures

Secure Operating System

Scality ARTESCA is bundled as a secure software appliance, providing an optimized Linux operating system, and a storage-centric Kubernetes environment with the ARTESCA software providing all necessary components together. ARTESCA OS only installs the minimum required packages for it to run. Everything else is removed, reducing the possibility of bugs, vulnerabilities, and consequently, cyber attacks. The secure operating system is updated with every software release of ARTESCA ensuring any common vulnerabilities and exposures (CVEs) are quickly and easily patched using ARTESCA's simple update procedure.

Scality works to the principles of zero-trust architecture and validates each release against the CIS (Center for Internet Security) Rocky benchmark 2.0.0 to continually audit and maximize security at all levels in ARTESCA.

- The default operating system user account is called artesca-os and is an unprivileged user. This account is the only one able to log in using SSH and who can run sudo commands and any attempts to become a root user, such as running the sudo -s command will fail. Restricting remote access to only one account reduces the attack surface of the system and greatly increases difficulty for remote attackers.
- For greater security, the artesca-os user can only execute a limited subset of commands required for the administration and maintenance of the system. The artesca-os user must use sudo for any procedures requiring root privileges. All other operations are forbidden, further restricting the ability of any malicious actors to cause damage.
- This artesca-os user only has permissions to access files in its home directory and not the data storage area of the system.
- ARTESCA enforces password hardening based on criteria such as length, complexity, and history.

Deployment Options

ARTESCA is available in three on-premise secure deployment options. Functionality remains the same whichever way ARTESCA is deployed, however each method offers unique benefits which could be more suited to your organization. In all cases the software is bundled with all the necessary components including the ARTESCA software, Kubernetes and a secure Linux OS.

Software appliance

The most flexible option. Deploy ARTESCA software on your preferred industry-standard hybrid or all-flash storage servers. Choose between one, three or six server options.

Each server should contain a minimum of:

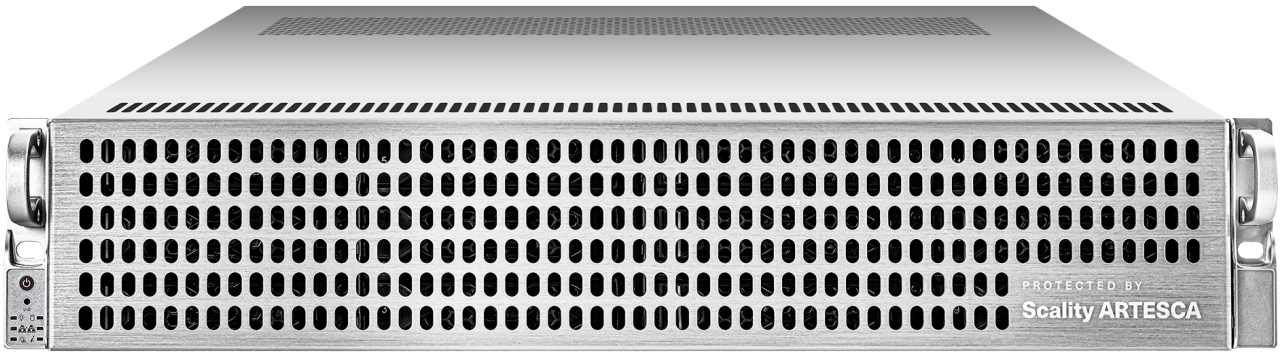
- Twelve (12) identical SAS data drives (HDD or Flash) for secure placement of all data stored in ARTESCA. HDDs data drives must be managed by a Hardware RAID controller, with a minimum of 2GB battery-backed write cache. Data drives will be formatted by the system during software installation. Data drives should be configured with individual RAID0 logical volumes for each disk if using HDD. No specific RAID configuration is required if using flash disks.
- Two (2) service drives (RI/MU Flash) Software RAID1 mirrored for ARTESCA Data Services including metadata, the ARTESCA storage services index, and the ARTESCA kubernetes platform database. Service drive required capacity will vary depending on the application workload and total capacity of the system, ARTESCA product documentation has up to date sizing recommendations.
- Boot drives capacity of at least 480GB from either
    - One (1) Boot Optimized Hardware RAID1 module with two SSD or NVMe drives
    - Two (2) SATA SSD or NVMe drives in software RAID1
- 192GB RAM
- 2x 10Gb/s network ports for the workload plane (LACP bonding for resiliency)
- 2x 1Gb/s network ports for control pane (LACP bonding for resiliency)
- CPU of either
    - Single AMD Epyc, 24 core minimum, equivalent or better performance than
        - Gen 2 (Rome): 7402P
        - Gen 3 (Milan): 7443P
        - Gen 4 (Genoa): 9224

- Dual Intel Xeon, 12 core minimum, equivalent or better performance than
  - Gen 2 (Cascade Lake): 4214R
  - Gen 3 (Ice Lake): 4310
  - Gen 4 (Sapphire Rapids): 4410Y

Hardware appliance for Veeam

Available in a variety of capacity sizes, the ARTESCA hardware appliance for Veeam is designed for ultra-fast deployment, plus the simplest procurement and support experience. Be up and running in less than an hour with an integrated appliance delivered with pre-configured hardware and ARTESCA software tuned for Veeam Backup & Replication v12, Veeam Backup for Microsoft 365 and Kasten.

The appliance includes a QuickStart wizard to simplify integration into your existing network environment. Leverage the Veeam Assistant to assure foolproof (and ransomware-proof) configuration of Veeam and ARTESCA. The appliance is available in limited geographies, see the 'ARTESCA hardware appliance datasheet' for more details and capacity options.



ARTESCA hardware appliance for Veeam

Virtual appliance

ARTESCA software designed and packaged as an OVA for industry-leading VMware virtual environments. The OVA is easy to deploy as a virtual machine for vSphere/ESXi. Deploy on one VM for up to 106TB (OVA) of usable capacity or install the software appliance into a three server VM for larger capacities.

The OVA is ideal to test ARTESCA functionality and features and is available for download as a time limited full-featured trial from scality.com.

XII. Partner Applications

Scality comprehensively tests and validates partner applications for ARTESCA giving customers the utmost confidence their applications interact with the object storage in the anticipated manner. Applications are placed in one of three categories; validated, certified or compatible.

Validated designs carry the highest level of testing ensuring predictable performance and compatibility with relevant application features. Validated designs are supported with the following items all available in the standard product documentation:

- Integration documentation
- Supported architectures
- Performance data
- BOMs (bill of materials)
- Performance tuning procedures and recommendations

Certified solutions undergo thorough testing with ARTESCA. These applications are tested with near real-world use cases, and data to ensure a predictable performance for the solution. Integration documentation, supported architectures and performance data is available.

Compatible partner applications have been analyzed to ensure that they meet the API requirements, and can integrate successfully with ARTESCA. Integration documentation is available.

Product documentation will list all of the partner applications and to which level they have been tested, providing confidence to organizations wishing to use that application with ARTESCA.

Veeam Integration

ARTESCA includes Veeam Assistant, a simple-to-use tool that quickly creates and configures all ARTESCA resources necessary to run a secure Veeam repository on ARTESCA object storage. The tool runs automatically when installed with the Veeam profile and can also be run anytime from the Data Management section. It can be enabled to support Veeam Backup & Replication or Veeam Backup for M365.

You are asked four simple questions and ARTESCA creates all the necessary resources to run an immutable Veeam repository including user, buckets, SOSAPI, object lock, policy and access key pairs. The assistant takes a few seconds to run; once completed, any information you need to enter into the Veeam UI can be copied to clipboard for simple error-free entry.

ARTESCA supports Veeam SOSAPI as used by Veeam Backup & Replication v12 and above, which primarily allows for storage reporting metrics to be displayed in the Veeam UI. The free capacity of an object store is displayed in the repository overview screen of Veeam, equipping the administrator to make informed decisions regarding backup placement. A capacity warning is displayed within the Veeam UI if the capacity of an object storage repository reaches its limits. These features simplify storage operations within Veeam, removing the need to switch between UIs to determine free and total capacities.

Note there is capability in SOSAPI called 'smart entities,' which is not relevant for ARTESCA. The main purpose of this functionality is to allow Veeam to effectively perform a basic 'load balancing' feature that associates jobs to different S3 endpoints, sharding the load over them. It is not needed in ARTESCA because all writes are automatically distributed over the different servers in a multi-server configuration. This built-in ARTESCA functionality has the major benefit of allowing all data to be read in the event of a server outage, whereas endpoints 'load balanced' with SOSAPI will not be able to serve their data in such an event. In addition, ARTESCA will evenly spread data over its servers, automatically manage data placement, eliminate data hotspots, and doesn't require manual tasks to monitor and redistribute data.

ARTESCA supports direct-to-object API calls, allowing it to provide capacity for the Veeam performance tier as well as being the capacity tier as part of a SOBR. The primary benefit is to guarantee the immediate immutability of backups by leveraging built-in object lock functionality. Backups stored on the performance tier of ARTESCA still benefit from key functionality such as VM Instant Recovery and VM SureBackup.

Veeam Ready

Having completed the process to meet Veeam's compatibility and performance standards, ARTESCA has officially achieved "Veeam Ready" validation for:

- Veeam Ready Repository - Primary backup storage solutions meeting or exceeding functional and performance tests for backup and restore operations
- Veeam Ready Object - Verifies compatibility with Veeam Backup & Replication object storage capabilities
- Veeam Ready Object with Immutability - Verifies compatibility with Veeam Backup & Replication use of S3 Object Lock abilities on object repositories

Veeam Ready Object includes Smart Object Storage API and IAM & STS validation providing:

- On-premise object storage specific enhancements including capacity reporting from the Veeam UI
- Secure and direct communication between object storage and Veeam agents

XIII. Summary

Further information, including ARTESCA data sheets, industry papers and demonstration videos, is available on Scality's website: www.scality.com.

TRY ARTESCA FOR FREE

Download a full-featured free trial of ARTESCA at scality.com/try

Visit **www.scality.com**
San Francisco • Paris • Washington, D.C • Tokyo • London