White Paper
Enterprise Software Defined Block and Object Storage
- Built for Autonomous Operations -

Table of Contents

Preface 3
1. Goal: Autonomous Operations and Self-Optimizing Systems 4
2. Outcomes, Autonomy and Optimization 4
3. Raising the Level of Interaction – Intent Based Management 5
   3.1. Intent Based Application Example 5
4. System Architecture 6
5. Policy Driven, Managing by Intent 7
6. Utilizing Automated Tuning versus Manual API Tuning 8
7. Continuous Autonomous Optimization 9
8. Autonomous, Self-Service Operation Example 10
9. Conclusions 11

Preface

This document describes the basic architecture of the ETERNUS DSP system and provides the features of the enterprise storage system by explaining the product specifications, benefits, and tradeoffs.

The product lineup and product information stated in this document are current as of January 2020.

■Intended Audience

This document targets the following audience:

- Those who are considering installation or replacement of storage systems
- Those who are proposing installation or replacement of storage systems

■Applicable Model

This document covers the following model.

- FUJITSU Storage ETERNUS Data Services Platform Software

■Naming Conventions

The following abbreviations are used in this document.

- ETERNUS DSP ....FUJITSU Storage ETERNUS Data Services Platform Software
- Web GUI ....Web GUI for ETERNUS DSP
- SSD ....Solid State Drive

1. Goal: Autonomous Operations and Self-Optimizing Systems

In the enterprise storage business, there has long been a measure of efficiency in terms of how many giga-tera- peta bytes a storage administrator could manage. While increasing this number is a worthy endeavor, in the modern data center, infrastructure supports a world that is no longer measured in petabytes (storage), gigabits (networking) or gigahertz (compute). The modern data center is measured by how it enables business agility in response to opportunity or threat – business depends on outcomes; all the other stuff is a means to an end.

ETERNUS DSP storage software was built to be autonomous, and self-governing from initial launch, with the unique characteristic to be able to seamlessly manage an infrastructure that may include multiple classes of storage media, current or future, that through policies setting deliver the optimal business outcome for any application. This is a revolution in storage management, moving from the treadmill of manual process optimization to intent-based management driven by artificial intelligence. Unlike other storage systems, this is not a veneer of management built on legacy storage architectures, but an architecture built expressly for the purpose of autonomous operations.

In this paper we will demonstrate how ETERNUS DSP helps businesses deliver outcomes through advancing storage system administration and management from process driven optimization to intent driven autonomous operations.

2. Outcomes, Autonomy and Optimization

To enable outcomes, the modern data center must continuously exploit increasingly rich data and adapt to rapidly changing needs of the applications that power the business. Furthermore, the modern data center must enable the exploitation of data, fully decoupled from the evolution of the infrastructure.

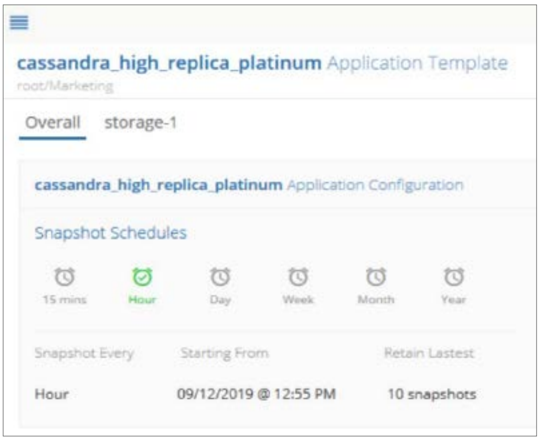
Delivering outcomes cannot depend on frequent human interaction, there are simply too many variables to be considered, too many inter-dependencies all happening at too fast a pace. The enterprise data center is complicated, built on a diverse infrastructure and always a combination of new and legacy applications – all of which are critical to the business. Business not only competes with others externally but often there are competing needs internally. There is an inherent heterogeneity and scale creating complexity as well as the business mandate for velocity and agility.

ETERNUS DSP was built on the premise of intent-based management, an environment where the administrator expresses the desired outcome and the ETERNUS DSP system continuously and autonomously monitors and adjusts to meet the intent.

In this paper we will cover the core capabilities that enable customers the ease of management and critical insight that radically simplifies how they can deliver business outcomes effortlessly, while adopting any storage technologies for best price/performance for any application:

1. Rich and open API
2. Powerful and tunable application templates
3. Policies that are easy to construct, and intelligently maintained
4. Integrated, cloud-based analytics engine that provides aggregated telemetry

3. Raising the Level of Interaction – Intent Based Management

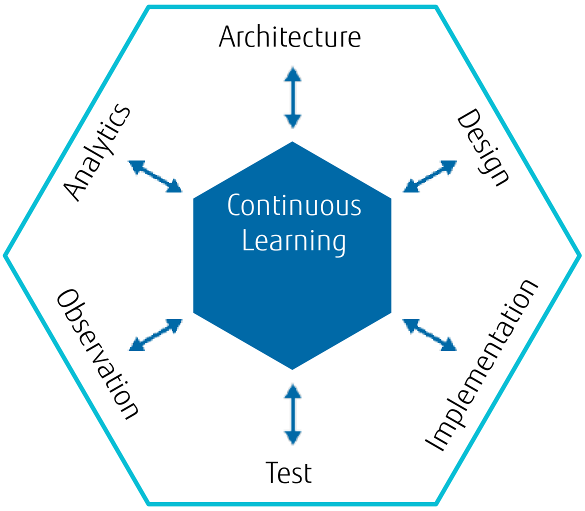


Key to autonomy is raising the level of interaction with the system and making the interaction programmatic, i.e. as-code. ETERNUS DSP raises the level of interaction via codifying applications storage requirements into templates. Templates contain all of the capacity, performance and data management (protection, distribution, security…) needs for an application instance.
When orchestration software instantiates the compute, networking and storage infrastructure for an application instance the template reference is all that is needed. Gone are the days of managing an individual volume's capacity, consistency groups, performance, failures domains, snapshot schedules, media layout, RAID… The "what" is all captured in the template and the who, i.e. who can access the storage, is all that is required.

You can have your cake and eat it too…By way of analogy, it is much like the emerging power of self-driving cars. In 1960, your car told you two things, how fast you were going, and when the engine overheated (if you were lucky!), all while you needed to watch the road, steer, accelerate, brake, etc. Today, a Tesla uses continuous optimization and telemetry gathering in real time, to offer you the luxury to read, as it automates steering, velocity, adjusting for driving conditions, time of day, and even the weather! Fully, self-service user model.

Raising the level of interaction to intent based management is but the first step in the journey. Specifying by intent in-turn lets the system operate with a high degree of autonomy in matching the intent to the continuously inventoried, available capabilities of the storage infrastructure. It is important to point out that matching the intent of a single application instance is done simultaneously with other application instances and tenants competing for the same resources.

3.1. Intent Based Application Example



As described in other white papers, one of ETERNUS DSP's key characteristics is the ability to seamlessly manage disparate types of media, whether that be current or future storage media, and though policies allocate “on-the-fly" the right performance/cost technology to meet the application's requirements. And if later requirements change, then by a simple policy change, the data will be redirected "live" to a new media class. In this paradigm, it radically simplifies customers' planning and adjusting to new requirements.

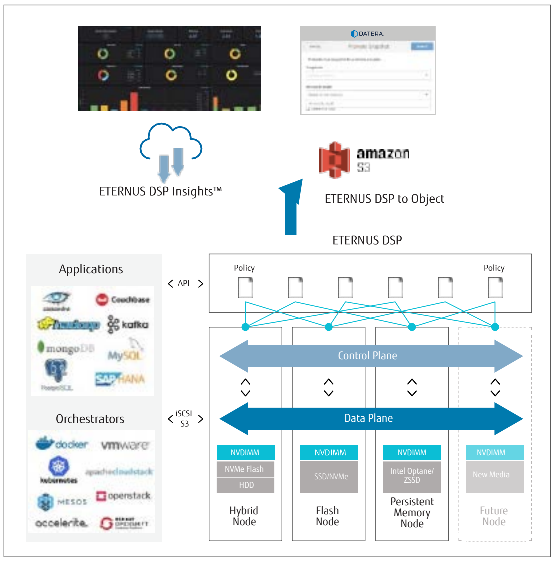
Now for an example on how this gets implemented for a mission critical, Cassandra DB appinstance. You would assign the highest levels of performance, using a previously created media policy, perhaps a "Platinum tier" comprised of only Optane or NVMe nodes. You would also assign a high replica count within the ETERNUS DSP cluster, perhaps 3 or more, as well as hourly snapshots to Amazon S3, via the ETERNUS DSP to Object capability. Encryption would be turned on, while dedupe might be set to off. This is done at the granular, appinstance level, and the ETERNUS DSP intelligence is fully aware of and able to exploit new performance media nodes as they are added. We will cover the mechanism ETERNUS DSP uses in the System Architecture section below.

ETERNUS DSP implements a closed loop system, where intent is captured as a service-level objective, converted to policy, which in turn drives data placement and management. Placement drives performance and distribution, which drives telemetry and analytics. Telemetry is compared to the intent and adjustments made as needed. Intent for application data can be changed at any time and the system will autonomously adjust. Equally important, the infrastructure can independently grow or shrink, fail and recover, speed up or slow down, … And for the grand finale, many applications and tenants can share the system simultaneously.

4. System Architecture

Scalability, low latency, policies, autonomous operations and analytics come together to deliver an agile and efficient data infrastructure paradigm

ETERNUS DSP's enterprise software defined storage (SDS) architecture provides a continuously available, service level objective-based data service for bare metal, VM or cloud-based applications. A broad range of orchestration layers are tightly integrated and are fully optimized by the ETERNUS DSP OS. ETERNUS DSP's low latency platform can seamlessly scale to PBs capacity and millions of IOPs.



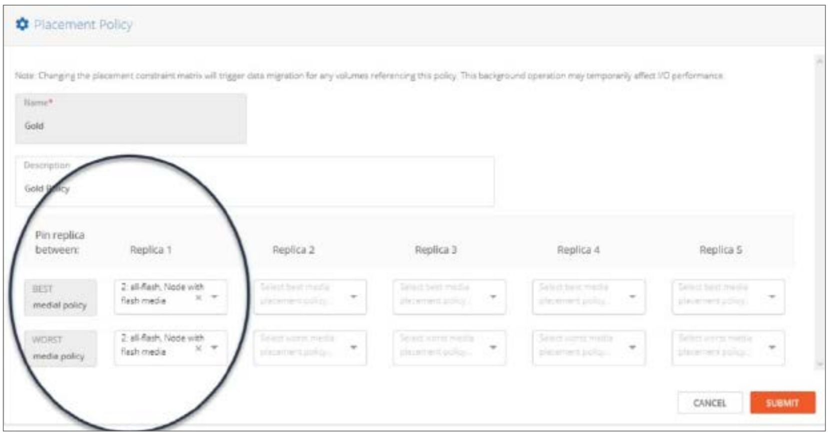
ETERNUS DSP is deployed in true scale-out fashion, using a heterogeneous mix of commodity storage media, including storage class memory, such as Optane™, NVMe, SATA flash, and conventional, spinning HDD.

The control plane handles policy management, as the data plane provides optimal data placement as set by policy.

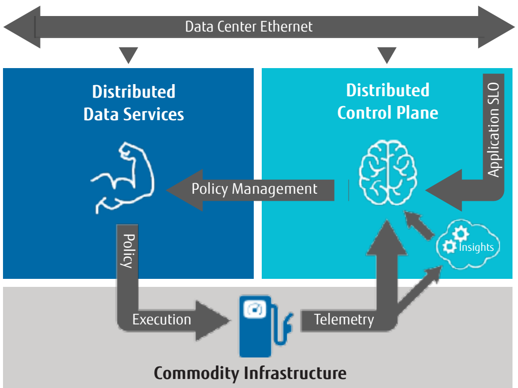
5. Policy Driven, Managing by Intent

ETERNUS DSP provides a volume service akin to cloud based Elastic Block Storage, but with a key distinction – the preferred method of provisioning storage for an application is via the use of application templates that specify the storage needs and policies for the application.

Setting a media placement policy takes seconds. You define how many replicas, and which type of media each replica should reside on. Many use the standard Platinum, Gold, Silver nomenclature, others application names. If you are modifying existing policies, intelligent data movement is automatically triggered, and the movement begins to build new current and future maps that are unique to the ETERNUS DSP platform. The key is, once set, ETERNUS DSP maintains the intent.



Inside the system, the templates map to policies driving service-level objectives (SLO's). SLO's are part of a closed feedback loop system of reading / writing data, gathering telemetry, comparing to the SLO's and then adjusting the placement and management of the data to meet the application SLO's. The administrator needs to specify the applications intent via the templates and the system will automatically adjust to meet the intent.



The most compelling part of managing by intent is that the policies can be changed at any time. As part of the closed loop system the change to the SLO's will be detected and any needed changes will be initiated and managed transparently and non- disruptively to the application.

Unlike previous enterprise storage systems, the ability to change the policies is not a simple management veneer over a rigid architecture. As you will see below, ETERNUS DSP software was built from the ground up to support the dynamic nature of modern data centers.

6. Utilizing Automated Tuning versus Manual API Tuning

ETERNUS DSP provides a policy driven environment where the system will automatically adjust to deliver the appropriate capacity, performance and functionality to the proper tenants and application instances.

The system allows manual tuning via both API's and UI's for volumes. An open, API first mindset, enables each volume to be perfectly tailored to meet the objectives of the applications. Provision and manage storage immediately on-demand through GUI or API, much as you would with the public Cloud.

The ETERNUS DSP best practice is to utilize the automated capabilities of the system unless the automation is over- ridden. Manual tuning requires the individual doing the tuning to understand both the application as well as the internal constructs of the ETERNUS DSP system. The best practice is to codify the needs of the application into the application templates and associated policies. This will enable the ETERNUS DSP automation to deliver both the application needs as well as balance the needs of other application and tenants as well as enable the system to exploit its internal processes to manage scale, change and failures in an automated fashion.

Creating a Performance Policy

```
POST
/app_instances/:id/storage_instances/:storage_instance_name/volumes/:volume_name/perfo
rmance_policy { "write_bandwidth_max": 20000, "read_bandwidth_max": 40000,
"write_iops_max": 80000, "total_iops_max": 100000, "total_bandwidth_max": 100000,
"read_iops_max": 100000 }
```

Modifying a Performance Policy

Performance policies specify performance guarantees and limits. The performance policies are enforced on the individual volume. Different volumes within the same Storage Instance can have different performance policies. You can input individual values to be modified.

```
PUT
/app_instances/:id/storage_instances/:storage_instance_name/volumes/:volume_name/perfo
rmance_policy {"write_bandwidth_max": 40000}
```

Note: You can input individual values to be modified.

Deleting Storage

Before deleting AppInstances, Storage Instances, or volumes there are two important steps that must be completed:

1. The host or application must cease any I/O to the Storage Instance. There is no way to guarantee this from the Datera side alone. 2. Either the AppInstance or Storage Instance must be taken offline.

Taking a Storage Instance Offline

```
PUT /app_instances/:id/storage_instances/:storage_instance_name[ "admin_state":
"offline" }
```

7. Continuous Autonomous Optimization

Historically enterprise storage systems had a small amount of autonomy, usually limited to having spare disk drives that could automatically replace failed drives. Most other changes to the system were either urgent or delicate, anything but autonomous. For example, dealing with controller failures in dual node systems required high-cost fast-response-time maintenance contracts to urgently replace the failed controller and restore performance and resiliency. Data migrations for better performance, resiliency etc. were delicate operations and normally were done by taking applications down. The solution – massive overprovisioning - $$$$ - and it only made incremental improvements. Urgent and Delicate!

What distinguishes ETERNUS DSP is how it continuously monitors how the system is performing relative to the intent and how it autonomously adjusts in response to internal or external changes. If a node (controller) fails in a 10 ETERNUS DSP node cluster the system will autonomously redistribute data among the remaining 9 nodes to restore performance, resiliency... to meet the intent - replace the failed node at your leisure. Similarly, if the policy is changed to improve performance, resiliency... the system will migrate data to better fit the aggregate resources autonomously.

Feedback from our installed base, and our cloud-based telemetry engine, Insights™ gives us an overview into their use of automation. Here are a few of the top, real world intelligence applications:

1. **Recovery** – An ETERNUS DSP system will autonomously recovery from a variety of failures including simultaneous failures of many different components. Recovery includes storage devices, racks, networking links, CPU, memory, operating system and the ETERNUS DSP software.
2. **Policy Changes** – Policies can be changed at any time and the system will autonomously adjust. Changes to policy include capacity, resiliency, performance, tenancy, fault domains and data management (snapshot, compression, deduplication...).
3. **Scaling** – An ETERNUS DSP system will autonomously incorporate new nodes including configuring networking, upgrading the software and distribution of data. Conversely nodes can be also be retired from the system with data evacuated and cleansed from the retired nodes.
4. **Re-Balance after a Capacity Expansion** - The ETERNUS DSP system is constantly taking an inventory of resources. As new nodes are added, current policies will take advantage of new resources, with workloads being moved to the appropriate, new nodes based upon performance.
5. **Software Upgrades** – When new software is available the cluster will autonomously deploy new software, on a non-disruptive, per node, rolling basis, including managing any networking changes to insure continuous availability.
6. **Incorporate New Technology** – When new technology becomes available, the technology can be autonomously incorporated into a cluster that was created before the technology existed. In some cases, this may require a "live" software upgrade while in others it only requires a description (performance, capacity...) of the capabilities.
7. **Autonomous Redistribution to Meet Policies** – The system allows applications instances to be created even if the resources to meet the intent cannot be guaranteed. As part of the closed loop continuous autonomous optimization, when new resources are added, if they enable the system to meet the intent, data will automatically be redistributed to meet the SLO with no user interaction.
8. **Insights™ Cloud based AI** – ETERNUS DSP systems send telemetry information to a cloud based ETERNUS DSP application that analyzes the information against the ETERNUS DSP installed base and will autonomously make recommendations if improvements are possible.

As discussed in the white paper Built for Change, the engine for ETERNUS DSP's ability to quickly, continuously and autonomously adapt consists of four key components:

- The Control Plane that monitors telemetry against the intent and adjusts as needed
- A cloud based artificial intelligence application that evaluates individual systems against the installed base trends and makes recommendations for improvement.
- A dual data placement map approach that allows data placement to be changed on the fly, any time, non- disruptively.
- An efficient time-based coherency protocol that ensures data correctness while delivering sub 40 microsecond data path (*plus network and media on reads)

ETERNUS DSP's holistic design approach envisioned these capabilities operating in concert to create autonomous operations.

8. Autonomous, Self-Service Operation Example

Let's walk through an example of autonomous operation to deliver business outcomes. This example highlights how simple it is for developers to programmatically deploy persistent volumes, and then for ETERNUS DSP to autonomously enhance and protect Kubernetes performance as the cluster gains additional performance nodes, or nodes are removed (as planned or unplanned events).

In a DevOps environment an application developer needs K8s statefulset to instantiate 10 MySQL instances for a business-critical high-performance application. The application instances will run in containers, so the developer uses Kubernetes (K8s) to provision the needed containers. ETERNUS DSP's application templates integrate K8s's label-based provisioning so in a single operation the compute, network, persistent storage, operating system and application for the 10 instances can be instantiated. The MySQL application template is setup for 10 gigabytes of platinum quality storage for the logs / journals and 200 gigabytes of gold quality storage for the database data. Deduplication and compression are enabled, and snapshots are scheduled daily with incremental data being replicated to an on-prem S3 object store. ETERNUS DSP nodes are placed in each of 10 rows in the data center to denote fault domains. K8s, via the Restful APIs programs the ETERNUS DSP system to create the storage for the MySQL instances and the storage is presented to the operating system of the respective containers. This all takes a few milliseconds and all volumes; data management and data protection policies are configured programmatically autonomously.

Inside the ETERNUS DSP system, the platinum level of storage for the logs / journals results in these volumes being distributed to nodes containing NVMe SSDs, with three replicas, denoting the fastest available storage in the ETERNUS DSP Cluster. The database itself is placed on SATA SSD's with two replicas denoting a balance between cost and performance. Nodes are distributed across different rows in the data center based on availability zones setup by the data center architects.

In the initial database fill the telemetry information is collected and the control plane indicates the service level objectives are being met. Telemetry sent and analyzed by the cloud-based AI / ML engine indicates that the database is achieving performance levels consistent with the installed base for this type of application.

During a data center expansion project an entire row of racks, including some of the ETERNUS DSP nodes, were taken down by a careless electrician, tripping breakers on both power feeds to the row. The scale-out application was setup for row-level fault domains and with the data autonomously distributed across fault domains there was zero impact to the availability of the application. Once power was restored both the ETERNUS DSP nodes in the affected row were autonomously brought back to normal operations.

The mission critical importance of the application increased, and the performance of the application directly impacted the bottom line of the company creating a competitive advantage. The CIO was asking if the performance of the application could be enhanced and an evaluation of the stack indicated that log / journal read performance was a limiting factor in overall database performance.

A storage architect familiar with ETERNUS DSP's capability suggested adding a new type of media technology, Intel Optane SSD's. A single ETERNUS DSP node containing Intel Optane drives was added to the cluster and the configuration database was provided with a description of the node indicating the enhanced performance.

Here is where the magic happens – where outcomes are delivered. The ETERNUS DSP Control Plane, going through its continuous autonomous optimization process, incorporates the new node into the ETERNUS DSP cluster and evaluates the current applications instances and their intents, against the available resources, including the newly added Optane drives. Recognizing the Platinum level volumes in the MySQL applications instances, it creates a new set of placement maps, called future maps, and provides the maps to the Data Plane to utilize for data placement and distribution. The future maps instruct the Data Plane to move one copy of the data onto the node containing the Optane drives. A few minutes later the migration is complete. Log / journal read latency goes from 500 microseconds to 200 microseconds and overall application response time improves by 50%.
The result is a 20% improvement in the business application transaction time creating an additional competitive advantage for the company.

Let's review what happened in this example:

- Using Kubernetes, 10 containers were created to run a MySQL database. No storage management was required, all storage functions including persistent capacity, resiliency, data management and fault domain management were programmatically and autonomously provisioned by the ETERNUS DSP cluster for all 10 instances using Application Templates integrated with K8s label-based provisioning.
- ETERNUS DSP's Continuous Autonomous Optimization insured data was placed, protected and managed according to the intent specified by the developer. The system was operating as intended.
- An entire row of racks, including racks containing ETERNUS DSP nodes, were inadvertently taken offline and ETERNUS DSP, because of its fault domain awareness, autonomously transitioned all work to nodes in other rows while continuing to meet the intent specified by the developer.
- An improvement in application performance was identified as an enabler to a competitive business advantage. Through the addition of a single node containing a new type of high-performance media the system, without any administration, autonomously incorporated the new media and redistributed data improving application response time by 50%.

If you are familiar with legacy storage, array-based architecture and first-generation software defined storage solutions you will appreciate that previous storage systems simply were not capable of accomplishing the benefits listed above. In most cases the activities listed above would have required significant downtime or administrative and data migration efforts that could have taken months. ETERNUS DSP delivered these business outcomes without the benefit of a storage administrator – **it was all accomplished autonomously**.

9. Conclusions

ETERNUS DSP's autonomous operation for enterprise storage is not an evolutionary step over previous architecture, it is the revolutionary culmination of a comprehensive and holistic approach to a new type of storage system.
Continuous Autonomous Optimization is a design philosophy that raises the level of interaction to the desired outcome, the intent. The system converts this intent into policies to store, protect, distribute and manage the data while always collecting telemetry that is fed back into closed loop process. At any time, in response to external changes in intent or internal factors such as recovery, scaling or upgrades, the system will adjust the way data is stored and managed to deliver the intended outcome.

Contact

FUJITSU LIMITED

Website: https://www.fujitsu.com/eternus/