Maxta Storage Platform

Enterprise Storage Re-defined

WHITE PAPER

Software-Defined Data Center

The Software-Defined Data Center (SDDC) is a unified data center platform that delivers converged computing, storage and networking layers, provides unprecedented automation, flexibility and efficiency and transforms the way IT is delivered. Compute, storage, networking and availability services are pooled, aggregated, and delivered as software and managed by intelligent, policy-driven software. SDDC redefines efficiency in two ways. First, it leverages standard, off-the-shelf commodity hardware achieving CAPEX reduction. Second, by converging compute, storage and networking, the over provisioning of CPU, memory, storage, and networking resources that exists in today's world of separate, distinct, physical entities is eliminated. Instead, these resources are managed in a global manner to significantly improve utilization and efficiency. In SDDC, all resources are specified and provisioned by the application team in an integrated manner. This in turn, greatly simplifies IT and significantly reduces OPEX.

Data Storage Challenges of Virtual Datacenters

On the compute side, VMware and others have delivered for the most part on the promise of the Software-Defined Server paradigm. Server virtualization has enabled aggregation of compute resources (CPU, memory) to create an efficient infrastructure, leveraging commodity hardware and providing elastic scale-out. Thus, server virtualization has redefined the compute paradigm by dramatically improving compute simplicity, flexibility, agility, scalability and availability thereby lowering both capital and operational expenses. Unfortunately, storage technologies have only made incremental innovations in the past decade and have not kept up with the pace of improvements in compute technology. Shared storage is a pre-requisite for utilizing many of the benefits of server virtualization such as migration of virtual machines, dynamic load balancing, high availability and disaster recovery. Currently, the most widely deployed form of shared storage is networked storage, either SAN or NAS. Networked storage solutions are significantly more expensive than other forms of storage and require specialized storage skills for provisioning storage and managing storage constructs such as volumes, LUNs, file systems, and RAID. Moreover, server virtualization increases the complexity of networked storage management due to the mismatch between storage constructs (e.g., LUNs, volumes, file systems) and virtualization constructs (Virtual Machines or VMs) and the need to manage servers and storage systems separately. While the provisioning of VMs can be done in minutes if not seconds, the provisioning of storage and associated storage constructs to match the VM can take much longer and requires much more expertise.

The advent of flash storage has been a boon in terms of significantly improving the performance of storage and in some cases even reducing the cost per I/O ($ per IOPS). Many vendors are positioning all flash arrays or server side flash as the solution to all storage problems. Unfortunately, they do not address the gap between storage constructs and server virtualization constructs or the relatively high storage capacity cost ($ per GB) of flash.

In recent years, alternative approaches to networked storage such as virtual storage appliance (VSA) and distributed file system have emerged. However, these approaches are limited in terms of scalability and performance while not improving storage management. Thus, they are not an alternative to networked storage in most cases. Other alternatives to networked storage such as Hadoop have emerged as well for Big Data environments and specific storage workloads. However, these alternatives are not well suited for general purpose virtualized environments and associated storage workloads.

Software-Defined Storage

Software-Defined Storage (SDS) meets the above-mentioned challenges of storage in virtualized environments by implementing storage functionality entirely in software leveraging commodity components. The software is running on the same physical servers that are running the server virtualization hypervisor. It is abstracting storage resources to enable aggregation and pooling of disperse server-side physical storage resources, and supporting on-demand allocation of VM-centric virtual storage resources. The mapping of virtual storage to physical storage provides high levels of performance, scalability, and availability by optimizing data placement and leveraging data replication. The VM-centric allocation of storage fully addresses the gap between storage constructs and server virtualization constructs. The result is a storage layer with agility similar to that of virtualized compute: aggregated, flexible, efficient and capable of elastic scale-out. In the last couple of years, it has become fashionable to leverage the term "Software-Defined Storage" or SDS, and many vendors have co-opted it to define their offerings. So how can one identify a true SDS that is an essential component of the Software-Defined Data Center?

Below are some examples where the SDS concept has been "stretched":

- Storage software that works with one and only one vendor's hardware or with a very limited set of hardware configurations.
- Storage software that works inside a networked storage "box" or a storage array rather than inside the same server hosting the virtualization hypervisor
- Software that accelerates the performance of a storage array and still requires some form of shared storage to deliver storage functionality. To qualify as true SDS, compute and storage have to be delivered on the same pool of industry-standard, off-the-shelf hardware. In SDS, the storage layer is expressed as software, running in virtual machines, leveraging pools of industry standard, cost effective servers and storage devices such as SSD, PCI-attached Flash, and magnetic disk drives. Resource efficiency results from the ability to pool and share all the resources (CPU, memory, storage) in the data center. Operational efficiency results from managing integrated (converged compute and storage) software entities versus isolated hardware resources. Agility is delivered through the ability to quickly react to new requirements. Required storage resources are dynamically specified by the application team as an integral part of their work flow. There is no need to coordinate or negotiate storage resources with a separate storage team. There is no need to forecast and manage different physical storage resource pools.

Over the last 10 years, software-defined servers have completely redefined how we think about computing. Server virtualization is completely mainstream these days (58% of all installed x86 workloads are running in a VM). These same concepts are being applied in the storage domain to deliver a key missing piece of the Software-Defined Data Center.

Introduction

Maxta has developed a ground breaking, highly resilient, scalable, distributed Software-Defined VM Storage Platform (MxSP) that enables IT to fully realize the vision of the virtual data center. Maxta enables the convergence of applications, server virtualization, and storage on commodity servers thereby eliminating the need for storage arrays and realizing the Software-Defined Data Center vision. MxSP is a hypervisor agnostic software-only implementation providing VM level storage abstraction and full integration into server virtualization user interface. It provides the ability to scale server virtualization and storage independently on-demand, one standard server at a time without having to over-provision resources. MxSP delivers VM level storage abstraction rather than block or file level abstraction. It eliminates the need for provisioning and managing low level storage constructs such as LUNs and volumes dramatically simplifying IT. It also delivers VM centric enterprise-class data services to support advanced capabilities such as live migration of virtual machines, dynamic load balancing, high availability, data protection and disaster recovery. MxSP leverages any combination of magnetic disk drives and flash technology to deliver competitive performance and high capacity at an attractive price for all storage workloads in a virtualized environment. MxSP intelligently maps VMs to storage resources, optimizing data layout for virtual workloads and leverages any form of flash technology for read/write caching. It is optimized for flash performance and hard disk capacity, enabling deployment of data centers on standard servers and eliminates the need for IT administrators to make difficult tradeoffs between performance and cost.

In summary, MxSP significantly simplifies IT, increases IT efficiency, and dramatically reduces capital and operational expenditures.

Maxta Storage Platform Architecture

MxSP's innovative, peer-to-peer architecture aggregates storage resources from multiple servers, assimilating a global namespace, creating a Maxta storage pool. An instance of MxSP software is installed on each of the servers that are part of the virtualization cluster. The storage resources can be any combination of magnetic disk drives and SSD. All the servers running MxSP software have access to the aggregated Maxta storage pool. However, it is not a requirement that all servers have storage resources or contribute storage resources to the Maxta storage pool. It is also not a requirement that the servers that contribute storage resources to the Maxta storage pool contribute the same type (e.g., SSD, PCI-attached Flash, SATA magnetic drive, SAS magnetic drive) or capacity of storage resources. The communication between the MxSP instances is over a private Ethernet network or shared Ethernet network. It is highly recommended that the network be a private network of at least 1GbE to avoid VM/application performance degradation due to Maxta data traffic.

The servers, running MxSP software either contribute storage to Maxta storage pool ("Converged Compute/Storage Servers") or do not contribute storage to the Maxta storage pool ("Compute Only Servers"). The hardware requirements of a particular server depend on whether the server is a Converged Compute/Storage Server or a Compute Only Server. A Compute Only Server can be upgraded to a Converged Compute/Storage Servers by:

a) Changing the role definition (if the Compute Only Server meets the hardware requirements of a Converged Compute/Storage Server) or
b) Changing the role definition and upgrading the server (if the Compute Only Server does not meet the hardware requirements of a Converged Compute/Storage Server)

Maxta architecture provides VM level storage abstraction rather than block or file level abstraction. It eliminates the need for provisioning and managing low level storage constructs such as LUNs, volumes, files, RAID, ACLs, and storage networking (zoning). It fully integrates storage management into the server virtualization UI. Additionally, all data services are seamlessly integrated into the virtualization UI, eliminating the inconsistencies and overhead of using multiple management/UI interfaces to further streamline and simplify IT.
Moreover, all data service are performed at the VM level to fully align the virtualization constructs with the storage constructs. All data stored on Maxta storage is protected with strong checksums and is mirrored across at least 2 physical servers to provide high levels of data integrity and data availability. Additional level of mirroring can be utilized within each physical server as well. With extremely fast and efficient VM level snapshots, zero-copy clones, local and remote replication, MxSP delivers VM agility, data resiliency, availability and protection. Combining this with capacity optimization techniques such as thin provisioning, in-line compression and in-line de-duplication, the platform delivers significant storage efficiency.

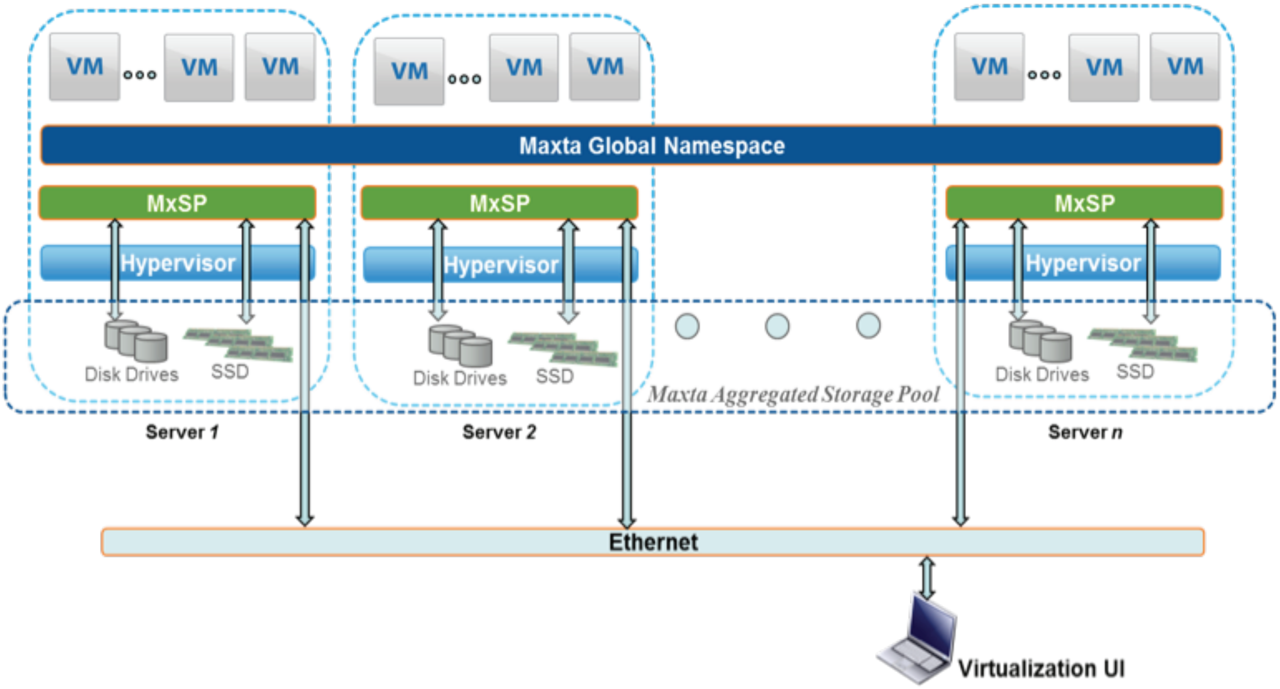
MxSP provides the flexibility to scale compute and storage independently on-demand, one standard server at a time without having to over-provision resources. Additionally MxSP seamlessly co-exists with other storage solutions providing investment protection for customers.
The main components of MxSP are:

– Maxta Distributed File System (MDFS)
– Maxta Management Server
– NFSv3 Server

Maxta Distributed File System (MDFS): An instance of MDFS is running in each of the servers that are part of virtualization cluster. MDFS enables the creation of the global namespace delivering high availability and scalability. MDFS leverages highly available metadata and logic to synchronize the operations between the MDFS instances and eliminates the split-brain condition (e.g., data inconsistency due to a network failure while the servers are still running, each one believing they are the only one running).

Maxta Management Server: The management server is deployed as a service on the servers running an instance of MxSP. The management server will be active on one of the servers and will be in standby mode on the others. In the event of a server failure where the management server was active, the management server automatically fails over to another server running an instance of MxSP. The management server leverage RESTful APIs to assist in administration and configuration of MxSP. These APIs also provide the basis of delivering the integration of user interface into the virtualization UI.

NFSv3 server: Maxta Storage Platform implements an NFSv3 server. The assimilated storage across all the servers in the virtualization cluster running the MxSP instances will be presented as a single Maxta storage pool to all the hypervisor instances in the virtualization cluster via an NFS protocol. The Maxta storage pool can co-exist with storage from existing storage arrays eliminating the need to rip and replace existing storage infrastructure.



MxSP Data Layout

MxSP has optimized the data layout for virtual workloads eliminating the tradeoff between cost and performance thereby simplifying administrators job. The pattern of read/write operations in a virtualized environment is predominately random in nature since the storage workloads of multiple virtual machines hosted on a virtualized server are blended together to generate a single blended storage workload. Thus, from a storage point of view, the storage workload always appear random independent of the individual storage workloads generated by the individual virtual machines (phenomena sometimes referred to as "I/O blender"). MxSP accelerates random write operations with a log based layout approach. The layout is very similar to log structured file system layout. A highly available metadata is maintained for mapping data blocks to their storage locations. MxSP leverages SSDs for both read and write-back caching. The writes are first written to the SSD in a sequential manner and replicated synchronously to an SSD on a second server for data availability. The writes are later de-staged on a periodic basis to the magnetic drives in a sequential manner. I/O performance and SSD endurance are maximized by always performing writes to both SSD and magnetic drives in a sequential manner. The metadata and frequently accessed data ("hot data") are cached on SSD to enable read acceleration. Leveraging SSD for metadata caching and read caching improves the performance of random read operations. While the hypervisor is leveraging RAM for frequently accessed data and metadata as a higher caching tier in front of SSD, SSD caching enables significant performance improvement since it is substantially larger than RAM supporting larger working sets and it is capable of providing write-back caching.

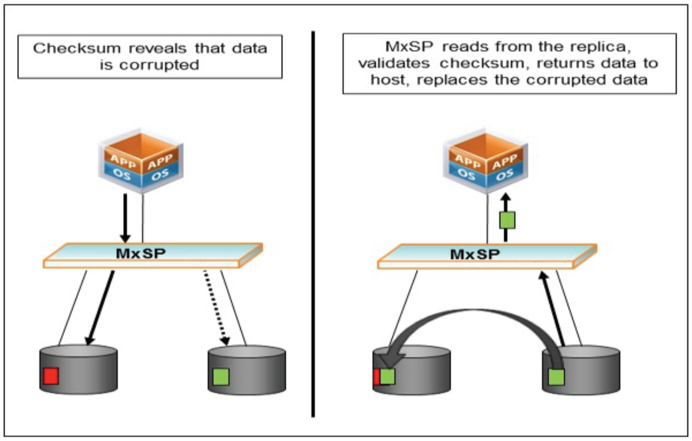
Note: While the preferred configuration of Maxta leverages high density, low $/GD magnetic drives and high performance SSD (either SAS/SATA attached or PCI attached) to optimize for both cost and performance, it is possible to leverage an all SSD configuration for extreme performance configurations.

MxSP intelligently optimizes the mapping of VMs to storage resources. The storage for a given virtual machine in most scenarios is local to the server where the VM is deployed. Under normal operation, both reads and writes are served by the server hosting the VM. Intelligence is built into the software to minimize the probability that a VM and its associated storage are on separate servers. This eliminates the network round-trip latency associated with distributed storage systems that don't optimize the placement of VM data.

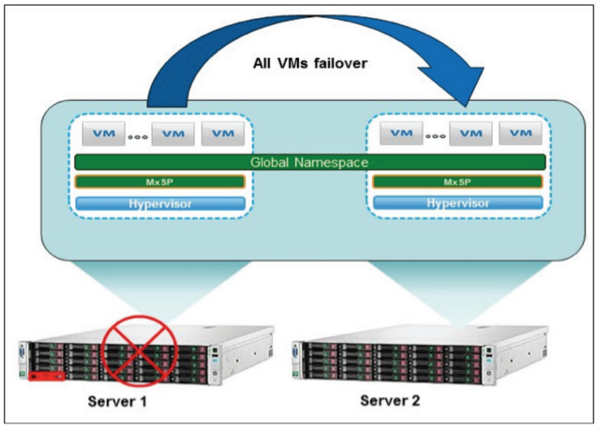
Data Integrity, Resiliency, Availability and Protection

Data Integrity and Resiliency

MxSP delivers best-in-class end-to-end data integrity. MxSP automatically detects and fixes data integrity issues due to hardware failures preventing silent data corruption. For each data block, MxSP maintains an independent checksum. The data for each data block is verified against its checksum after the data has been read. MxSP doesn't only verify the consistency of the data within the block but also validates that the block read by the server is the right block. In addition, MxSP executes self-healing whenever possible in case it detects inconsistencies in the data block.



MxSP delivers top-notch data resiliency with no single point of failure. MxSP supports the ability to synchronously replicate the data locally within the servers and between the servers. This provides the ability to sustain any 3 disk drive failures without losing data or access to data.



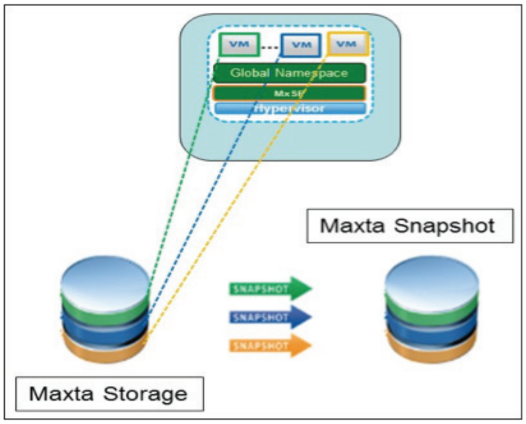
Data Availability

High availability features built into MxSP architecture addresses multiple failure scenarios providing business continuity. All writes are synchronously replicated to a second server within the virtualization cluster. This protects the application against disk and server failures. MxSP optionally provides the ability to replicate data locally within the server synchronously. This avoids rebuild over the network when drives fail, improving rebuild duration and minimizing network traffic. High availability features of MxSP seamlessly integrates with high availability features of the hypervisor to enable automated failover of virtual machines and associated storage minimizing downtime and administrative errors.

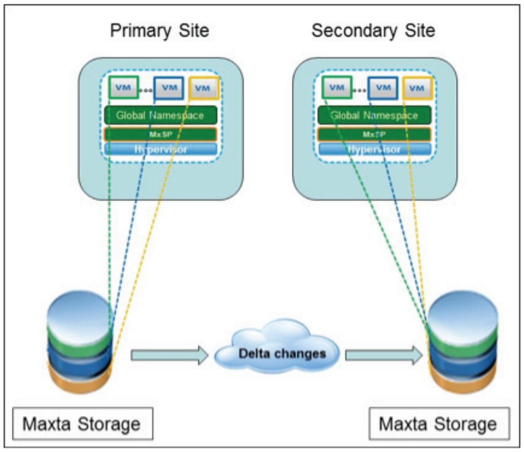
Data Protection

MxSP provides multiple features to protect your applications simply and efficiently. MxSP delivers unlimited number of time, performance, and capacity efficient VM level snapshots and zero-copy clones. Additionally, snapshots and zero-copy clones are configured and managed from the virtualization UI at the VM level rather than from a storage specific UI at the storage level. This enables the VM administrator to leverage advanced storage features without the need for deep storage and vendor specific expertise. The source for a snapshot or clone can be a VM, another snapshot, or another clone. MxSP leverages a log based layout approach for data placement and metadata for mapping data blocks to their storage locations. Therefore, once updated, the new image of a data block is not stored "in-place" to the same storage location that contains the previous image of the data block. Instead, the new image of the data block is written to a new storage location and the metadata is updated to reflect the change in the storage location of the data block. With this approach, the creation of a snapshot/clone is merely the creation of metadata for the snapshot/clone that points to the same data blocks as the source. Once a data block is updated by a VM or a clone, the metadata for the respective VM/clone is updated to reflect the storage location of the new image while the metadata of all snapshots and clones of that VM/clone is kept unchanged.

It takes less than a second to create a snapshot or a clone independent of the capacity of the source. There is no upfront space reservation when creating snapshots or clones. Snapshots and clones do not consume any capacity unless data is modified. Creating snapshots and clones does not impact the performance of virtual machines. All snapshots and clones are crash consistent. MxSP also enables application consistent snapshots in these environments that provide the appropriate tools for application consistent snapshots (e.g., VSS quiescing). Consistency groups provide the ability to take a consistent snapshot or clone of an application spanning multiple virtual machines. Snapshots provide the first layer of data protection without sacrificing performance or capacity. In case a recovery is required, a snapshot can be cloned and powered up as a recovered VM (of a previous point in time) instantaneously.

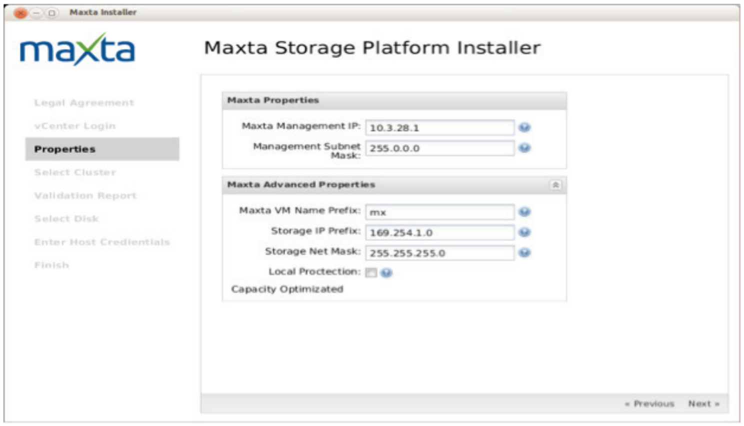


Disaster recovery of applications in greatly simplified with asynchronous replication. Configuring replication is dramatically simplified by providing the ability to set replication policy at the VM level from the virtualization user interface. MxSP's implementation of consistency groups enables comprehensive recoverability for applications that span multiple VMs. Only the delta changes between the replication interval is replicated, minimizing network bandwidth. Additionally, capacity optimization techniques such as in-line compression and de-duplication are applied to the delta changes before replicating to the disaster recovery site providing capacity and bandwidth savings.



Simplify IT Management

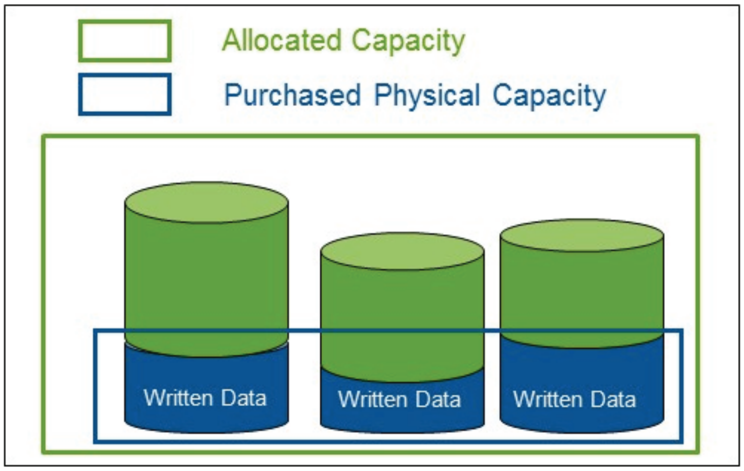
MxSP dramatically simplifies storage management by eliminating the need for storage provisioning and managing volumes, LUNs, file systems, RAID, ACLs, etc. This eliminates the need to develop the mapping between virtual machines and storage constructs such as LUNs and volumes minimizing administrative errors. Converging compute and storage on standard servers eliminates storage networking tasks such as zoning simplifying storage management even further. MxSP provides a single-pane-of-glass for both VM and storage administration by eliminating the need for a separate storage UI and incorporating all the functionality within the virtualization UI, minimizing learning curve and administrative time. Once MxSP is installed, the administrator can just point to Maxta storage pool from the virtualization UI during VM creation and MxSP will take all steps to optimally provision and allocate resources for the new VM.



The installation and configuration of MxSP takes only few minutes enabling all system/VM administrators to install and manage storage. The step-by-step process will guide the administrator through the installation process. During installation, MxSP aggregates storage across all the servers in the cluster and presents it as Maxta storage pool with a global namespace that is used to provision storage for VMs. MxSP will not add to Maxta storage pool any storage that is already allocated to an existing data store. It is also possible to selectively exclude any storage device from Maxta storage pool. The installer automatically installs and configures all the necessary pre-requisite software to enable a single pane of glass for both VM and storage management. Additionally, the installer identifies the minimum hardware and software requirements for installing MxSP. The installer guides the user through the resolution process in case the minimum requirements are not met.

MxSP VM level clone feature enables the provisioning of unlimited number of VMs with a predefined configuration instantaneously. This capability enables rapid provisioning of new VMs further improving IT agility and efficiency.

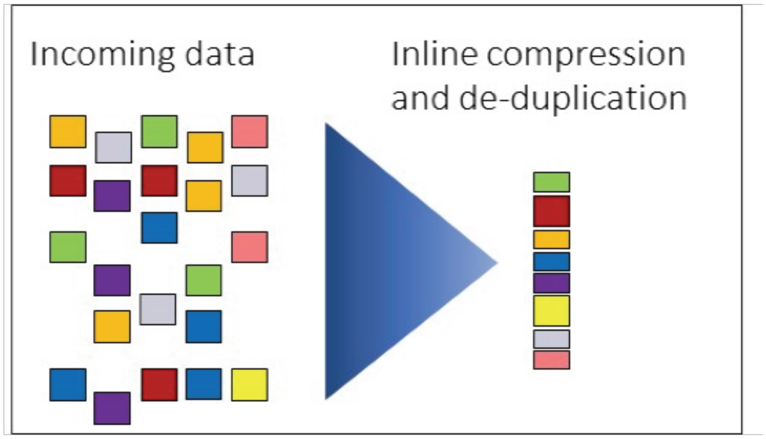
Maximize CAPEX and OPEX savings



MxSP delivers the capabilities of storage arrays in software by leveraging server-side SSD and internal drives or JBODs connected to standard servers. MxSP enables significant CAPEX savings by converging compute and storage resources on standard servers leveraging commodity components, without compromising performance or scalability. MxSP leverages any combination of commodity disk drives and flash technology to deliver competitive performance and high capacity at an attractive cost for all storage workloads in a virtualized environment.

MxSP eliminates all the storage provisioning activities and dramatically simplifies day-to-day data management activities such as recovery of virtual machines from snapshots, and creating new VMs rapidly using clones. Additionally, MxSP eliminates storage management tasks such as changing RAID parameters or storage tier management enabling administrators to focus on strategic tasks. MxSP provides significant reduction in power, cooling, and floor space by converging compute and storage resources on standard commodity servers and eliminating the need for storage arrays.

Capacity Optimization



MxSP delivers capacity optimization capabilities such as thin provisioning, in-line compression and in-line de-duplication to enable significant storage efficiency. Thin provisioning provides the ability to provision capacity several times larger than the amount of physical storage resources. This enables procuring incremental capacity on demand. There is no upfront reservation of capacity when utilizing thin provisioning and it is enabled by default.

MxSP performs in-line compression of data with no performance penalty. Compression is also enabled by default on MxSP. A combination of in-line compression and de-duplication significantly reduces the capacity consumption for all virtualized workloads.

By significantly simplifying IT, eliminating the need for storage arrays, increasing IT efficiency, and delivering capacity optimization, MxSP provides significant OPEX savings.

Conclusion

MxSP dramatically simplifies IT by eliminating the need for storage provisioning and managing volumes, LUNs, file systems, and RAID. The installation and configuration of MxSP takes only few minutes. Additionally, all data services such as local and remote replication, snapshots and zero copy clones are configured and managed from the virtualization UI at the VM level rather than from a storage specific UI at the storage level. This enables the VM administrator to leverage storage without the need for deep storage and vendor specific expertise. With MxSP, administrators can Manage VMs and NOT storage.

MxSP delivers enterprise-class data services with best-in-class resiliency continuous availability, data protection and agility. With unlimited VM level snapshots, backup performance is improved and recovery times are dramatically reduced. By leveraging VM level clones, rapid provisioning of VMs is possible. With asynchronous replication and continuous data protection, disaster recovery of applications is greatly simplified and accelerated for business continuity. MxSP seamlessly integrates with all the advanced capabilities of the server virtualization software.

MxSP enables significant capital savings by converging compute and storage resources on standard commodity servers, without compromising performance or scalability. This provides considerable up-front capital savings vs. storage arrays. In addition, MxSP leverages any combination of magnetic disk drives and flash technology, snapshots, zero-copy clones, thin provisioning, in-line compression and in-line de-duplication to increase storage efficiency and reduce storage expenses. By significantly simplifying IT, increasing IT efficiency, and enabling administrators to focus on managing applications and VMs, MxSP enables dramatic reduction in operating expenses.

100 Mathilda Place, #170, Sunnyvale, CA 94086
408-789-5749
Email: sales@maxta.com
www.maxta.com