National College of Ireland

Data storage and Management Project A

on

Software-Defined-Storage (SDS)

Sindhujan Dhayalan
X17170265

MSc/PGDip Data Analytics – 2018/9

Submitted to: Vikas Tomer

Software-Defined-Storage (SDS)

Sindhujan Dhayalan
X17170265

December 20, 2018

Abstract

SDS is defined by SNIA as Virtualized storage with a service management interface. The model is built based on virtualization of storage such that users can access and control it through a software interface independent of the nature, configuration and location of the storage device. The software layer manages the entire storage infrastructure as an entity. This paper discusses SDS and its advantages over the traditional network storage systems.

1 Introduction

In the era of Big Data, the content and network providers require to manage the heavy amount of data collection in a more flexible way than what is being offered by traditional network storage. The Software Defined Storage (SDS) systems meet such flexibility needs by minimizing the complexity of data management in cloud environments.Mauro et al. (2017).

Software defined storage (SDS) is the emerging storage technology. Inspired by Software Defined Network (SDN), The complexity of the storage device is divided into two parts by SDS - the pure raw storage device(data plane) and the controller software resides in a server(control plane) and manages the set of raw storage device.

The controller creates replicas of each block on multiple devices and fragments of the file are stored on multiple devices. Resulting in very high reliability and increases the throughput as the file is read from multiple places at the same time. The reliability and scalability of SDS is same compared to that of SAN and NAS traditional network storage systems. But the cost of SDS is very low compared to the expenditures made for SAN and NAS storage by data centres. Gupta & Goel (2015)

2 SDS Architecture

SDS is one of the main components of SDI architecture. the SDS storage function is divided into two layers mainly, Data plane and Control plane.

SDS Control plane: the control plane is software layer of SDS that manages the virtualized storage resources. It enables optimized, flexible, scalable and rapid provisioning storage infrastructure capacity. These capabilities span functions such as policy automaton, analytics and optimization, backup and copy management, security and integration

with API and cloud services. SDS data plane: the data plane is the interface to infrastructure where data is processed and stored. it consists of storage management functions such as virtualization, RAID, protection, tiering, copy services, encryption, compression, and data deduplication as per requested by the control plane.

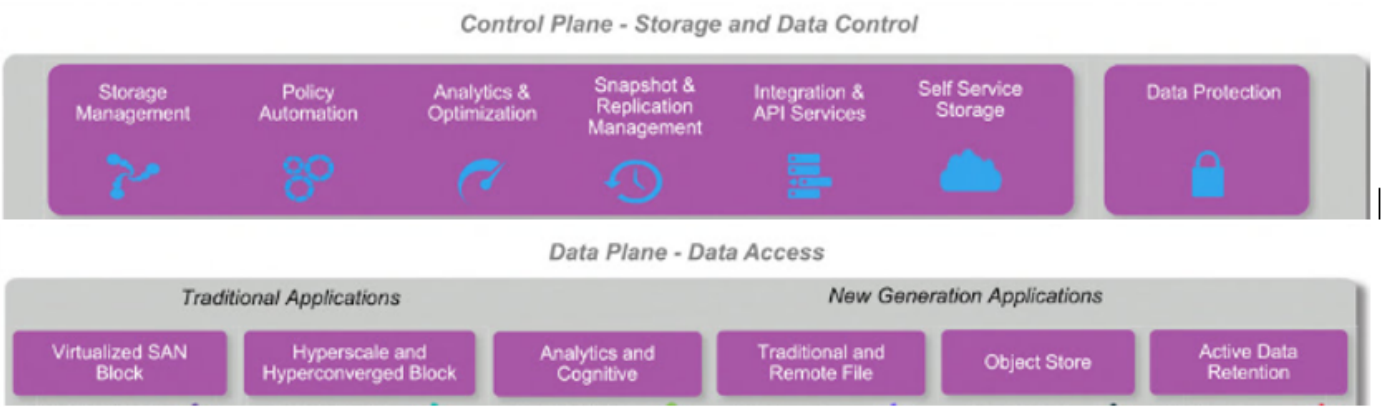

Figure 1: SDS functions through control and data plane, source: Coyne et al. (2018)

SDS provides agility, control and efficiency required to satisfy the rapidly changing business requirements by dynamically optimizing infrastructure capabilities to application service requirements.Coyne et al. (2018)

3 SDS Capabilities and Functions

SDS must have the following capabilities to support SDI with storage services.
Optimal workload allocation:
SDS must manage the dataflow and workloads across storage service level tiers to achieve agreed upon performance, availability, retention and security.
Data Retention and Archive:
SDS must enable archival capability for long term data retention as per the business requirements to maintain regulatory requirement and effective data rationalization. Agility and Scalability: SDS must be agile to respond within the allocated time and SDS must provide storage services that is scalable as per business requirement.
Flexible Data Access:
SDS should be capable of enabling data sharing and multi tenancy across data centers as per business requirement. It must provide access methods and protocols as per business needs.
To support the above discussed capabilities, SDS integrated components should contain API-controlled functions in order to support business application service levels:
Storage virtualization: Facilitates the creation of unique pools of storage capacity like heterogeneous disk storage arrays and devices.
Policy automation: Facilitates the automation of storage contributing as per the storage policies which are correlated to automation workflows.
Analytics and optimization: the function provides metrics to measure the perfomance of storage and capability usage. It also provides the tools to analyze them by making reports and comparing them to the required service level.

Availability, backup and copy management: The function should provides back-up/restore capabilities and local/remote copies for recovery purposes.
Integration and API services: The features provided by SDS should be integrated internally as well as with external functions available in SDI and with business applications.
Security: Enables secure data access by authorized persons.Coyne et al. (2018)

4 SDS VS Traditional network storage system

This section discusses on the advantages of SDS over the traditional network storage system by comparing them on Cost, Flexibility, Agility, Resource utilization and storage.

4.1 Cost:

Traditional networked storage solutions require high capital expenditures to procure, house, operate, cool, maintain, support, scale and upgrade an installed base of equipment. Whereas SDS lowers the storage cost by switching over from capital expenditure to operating expenses. It follows pay as your grow approach benefiting upgrowing companies. Instead of purchasing storage hardware and software through up-front capital expenditures, Companies can purchase software services on a monthly fee basis by switching to SDS. This is commonly called the Storage-as-a-Service (STaaS) model. SDS provides significant savings by allowing the use of inexpensive COTS (commercial, off-the-shelf) x86 servers and commodity Hard disk drives in place of dedicated and costly proprietary storage appliances when a customer decides to keep their own on-premises storage infrastructure, or to maintain portions both on-premises and in the cloud.van Duijvenbode (2017*a*)

4.2 Greater Flexibility

The function of SDS implementations resides in software rather than hardware, the functionality of the system is not limited by the unique characteristics of individual storage units. With the policy-based guidance of the SDS software, the most basic to the advanced types of storage arrays and subsystems can be configured with the best advantage of the capabilities of each type of device. This gives storage administrators the flexibility to mix and match devices and media (such as HDDs and SSDs) in ways that maximize their effectiveness in meeting the requirements of specific workloads. This is not possible in traditional network storage model which is dependent on hardware storage units.van Duijvenbode (2017*a*)

4.3 Agility

Agility is the speed at which organization adjust to changes in technology or in its business environment and it is one of major factor required to survive in the IT industry. Making large changes to an organizations traditional network storage infrastructure mostly took weeks or months because of time required to fund, order, install, and configure new compatible hardware. But In case of SDS, Users have the flexibility to quickly and freely scale the amount of storage using the software they use without any purchase of new hardware on-premises or in the public cloud. Organizations can implement additional

capacity through software in minutes rather than days or weeks, providing companies with the agility to respond in near real time to changes in business conditions.van Duijvenbode (2017*a*)

4.4 Resource Utilization

The traditional storage model requires data centres to have unused additional storage units for any unforeseen spikes or need in demand. To prevent such scenarios from occurring, a prudent storage administrator will make sure that spare capacity is instantly available when needed. The result is that in normal operations a significant proportion of the data centers available storage capacity will remain unused. With the SDS model customers pay only for the storage and services used, STaaS provider will have adequate reserve capacity to be allocated as needed. there is normally no need to purchase and maintain unused capacity. van Duijvenbode (2017*a*)

4.5 Storage

Storage capacity: SDS can offer more storage capacity with commodity storage with enterprise-grade hardware. It can offer the same storage capacity that is offered by SAN for much lesser expenditure. Organization can benefit from high performance/high capacity storage nodes with a third to double as much raw data at a lower cost with SDS.van Duijvenbode (2017*b*) **Storage Management:**Traditional storage model requires high end hardware configuration and resources with storage knowledge to maintain, operate the storage units to function in this big data era. Whereas the complexity of the storage management task is reduced by SDS. Using SDS a single unified software interface centrally manages the entire storage infrastructure. Reducing the hassle over the configuration, setup and maintenance of traditional storage model. Enterprise-level storage features such as replication, inline data deduplication and compression, snapshots, encryption, and best-of-breed backup and disaster recovery functionality can be implemented once and applied system wide as the storage system is embedded in software that controls the entire infrastructure at a granular level.van Duijvenbode (2017*a*)

References

Coyne, L., Dain, J., Forestier, E., Guaitani, P., Haas, R., Maestas, C. D., Maille, A., Pearson, T., Sherman, B., Vollmar, C. et al. (2018), *IBM Software-Defined Storage Guide*, IBM Redbooks.

Gupta, S. C. & Goel, A. (2015), Software defined storage technology, *in* '2015 Asia-Pacific Software Engineering Conference (APSEC)', pp. 6–7.

Mauro, M. D., Longo, M., Postiglione, F., Carullo, G. & Tambasco, M. (2017), Software defined storage: Availability modeling and sensitivity analysis, *in* '2017 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)', pp. 1–7.

van Duijvenbode, J. (2017*a*), 'How sds is disrupting the traditional storage industry'.
URL: *https://www.dezyre.com/article/overview-of-hbase-architecture-and-its-components/295*

van Duijvenbode, J. (2017*b*), 'How sds is disrupting the traditional storage industry'.
URL: *https://www.dezyre.com/article/overview-of-hbase-architecture-and-its-components/295*