



Transformation from legacy storage to software defined storage—a review

Swaraj Paul Chinnaraju^{1*}, G. Gunasekaran², N. Kumar³, R. Anandan⁴

¹Department of Computer Science & Engineering, Vels Institute of Science, Technology & Advanced Studies(VISTAS), Chennai, India. St. Peter's Institute of Higher Education and Research, Avadi, Chennai, India.

²Meenakshi College of Engineering, Chennai, India.

³Department of Computer Science & Engineering, Vels Institute of Science, Technology & Advanced Studies(VISTAS), Chennai, India.

⁴Department of Computer Science & Engineering, Vels Institute of Science, Technology & Advanced Studies(VISTAS), Chennai, India.

*Corresponding author E-mail:mail2swarajpaul@gmail.com

Abstract

In IT Industry the new trend is everything to be in Software defined. And so the storage industry has started its transformation from Legacy storage like storage area network and network attached storage. This is enable through the software defined networks. Software defined networks helps organizations to accelerate its application deployment, thus reducing IT costs by applying some policy enabled workflows. In this paper, we will have a review over the need for it from the storage vendor perspective.

Keywords: Software defined network, virtualization, multitenancy, agility.

1. Software defined network—a brief Introduction

A software-defined networking (SDN) architecture defines how a networking and computing system can be built using a combination of open, software-based technologies and commodity networking hardware that separate the SDN control plane and the SDN data plane of the networking stack.

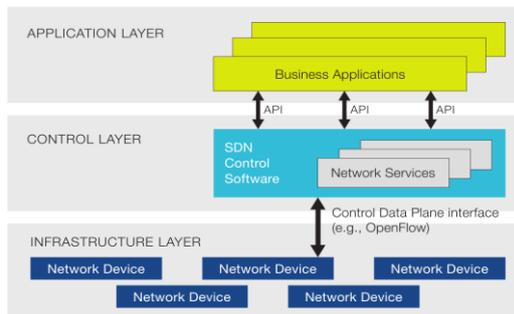


Fig. 1: Software defined network architecture

The three components or functionality groups of SDN architecture are given below :

SDN Applications: The communication of behaviors and the required resources by the programs, called Software defined Application is through the Application programming interface (APIs). In addition, the applications can build an abstracted view of the network by collecting information from the controller for decision-making purposes. These applications could include networking management, analytics, or business applications used to run large data centers. For example, an analytics application might be built to recognize suspicious network activity for security purposes.

SDN Controller: The SDN Controller is a logical entity that receives instructions or requirements from the SDN Application layer and relays them to the networking components.

2. Software defined data center

Applications in an SDDC run entirely on logically defined resources for which the underlying hardware is abstracted away. In this sense, the SDDC architecture is hardware agnostic, with the requirements as basic as x86 servers, with disks, and a simple packet-forwarding network backplane. A completely virtualized infrastructure enables the SDDC cloud management platform to analyze trends and apply machine learning to help self-regulate the IT environment.

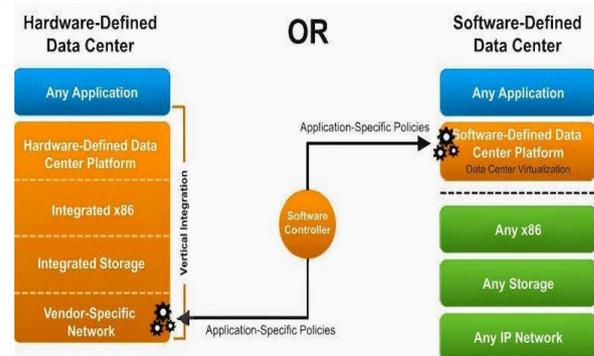


Fig. 2: Hardware vs software defined data center

3. Software defined storage vs traditional Storage from the view of storage vendor

Based on the difference between software and hardware
When the whole IT sector moves towards software defined storage, then there will be a dramatic change in the infrastructure

leading to most of the hardware functionalities to be replaced with some appropriate software, with only a lesser requirement for the hardware system. On the basis of this, some important criteria were discussed.

Cost

On taking cost into account, the vendors of traditional SAN and NAS storage array have spent a huge amount for installing user defined custom ASICs, its circuit board and the real time operating system required for it. For setting a user defined custom ASIC it had costs with around 10 million USD to 20 million USD, with additional charges for dynamically altering the number of storage arrays based on the requirement. It clearly shows that out of the total expenses spent for the Research and Development, only a small fraction of the amount was used for maintaining the storage functionality. The major portion of the R & D expenses were spent on developing and installing the required user defined custom hardware and software infrastructure. Apart from this, as all the sub components are proprietary, the traditional storage vendors has to spend some part of the fund in testing the interoperability of all sub components. For an example, it may cost some millions of dollars and may take many months for setting a simple drive qualification cycle. But on the other side, Software defined storage concerns are using a standard x86 hardware, and an appropriate standard operating system. It supports inter operability, that means the semiconductors here are from large scale industries of the world like Intel, Broadcom etc. The hardware is manufactured by companies like HP, Dell, IBM etc. and the operating systems are developed by companies like Linux, Microsoft, VMware etc. In such a way, company like Zadara storage has forcefully implemented the proven technologies to make its R & D to focus mainly on how the storage works, and thus providing a better result with only a small part of the amount spent for the traditional storage methods. When this is the situation, the question raises why do these traditional storage system still preferring custom hardware? If any storage vendor is asked with why they need such expensive custom hardware, their answer is the reliability of their hardware is much better than the software defined storage system. This is true, but on the other hand the traditional software depends at greater degree on their underlying hardware. If there is a single point of failure, then it can manage, but not to that extent in case of multiple and frequent failures. It clearly shows that Software defined storage companies like Zadara storage, are not relying on the hardware in its infrastructure. It makes its software to function continuously even at the time of multiple and frequent hardware failures.

Agility

When software defined storage is implemented its benefits does not stops with the cost alone. At what speed we are completing the task? This is also an important characteristic to be considered. So the next big advantage this software defined storage provides to the IT sector is the Agility. In general Agility means doing some work faster, and much faster. It is a significant value add for allowing rapid and effective response to the dynamically modifying requirements of the business. At some situation, Agility plays an important role that the cost spent for that task. We can realize it in terms of weeks vs minute! Yes, for an instance, it will take some weeks of time starting from plan to buy, next to install the required infrastructure, then connecting all with the application servers. On the other hand the same work can be completed within a minute of time with companies like Zadara storage. One more important factor to be considered in terms of Agility is the Automation of work till its completion. Traditional storages uses Command Line Interfaces (CLI) or Storage management initiative specification API to automate all its operation. But software defined storage storage uses a much simpler REST API to automate all its work.

Scalability

Apart from this major hardware and software, many other differences are there between Traditional storage arrays and the software defined storage. Scalability is also an important criteria

among these. Dual controller is the configuration specified for Traditional storage arrays. In this if one scales out, again it can be scaled to a maximum of 4 to 8 controllers. It clearly depicts there is a real time limit to the number of drives that can be attached to a controller. In case of SSDs (Solid state drives), this limit for scalability is too lower because of its higher input output operations per second (IOPS) rate of the drives when compared to its controllers. But software defined storage is equipped as like cloud. In this storage type hundred of nodes, or thousands of nodes are connected to the standard network. This provides the feature that the number of drives that can be connected with the controller is unlimited. It has the flexibility to connect with any number of drives at the dynamic time. In the view of software architecture, a huge difference can be noticed in terms of designing a distributed network infrastructure that can extend dynamically upto 8 nodes and the other with the flexibility of extending over thousands of nodes. But in the case of software defined storage there are some limitations that there should not be any blocking activities, as this may lead to different components that are be in up or down at any specified time that needs the software to change its module up and down. One more thing is that the entire storage system should not get shut down or rebooted not even at the time of upgrading of the software. An Automated software upgrading has to be there on a round robin fashion, whenever there is a need and it has to be applied for all the nodes irrespective of the nature of the nodes.

Reliability

We had already seen that the reliability of stored data in the traditional storage methods depends mainly on the underlying hardware system. Whereas, the software defined storage method depends mainly on its software for the reliability of stored data, with the assumption that there are more chances for the hardware to get failure in case of highly scalable infrastructure. The intelligence, self re-constructing process of the software assures for the continuous working of everything in that environment. These features of the software defined storage makes it more reliable than the traditional storage system. This approach over the network had made internet and its various protocols more reliable. Much in the same way, TCP/IP will ensure that if there is a failure while transferring packets between two nodes, with its intelligence it has to reroute packets via an alternate route to ensure reliability of data.

Multi-tenancy

The traditional storage systems are used for a single purpose only because of its limited scalability, and it may have only one administrator or a limited number. Moreover in the traditional storage system, there is no mutual separation or scheduling at the time of multitasking with many applications. Ultimately it leads to the lower performance and inability to self provisioning. Generally, the customers will buy many arrays for storing from the vendors, later they will make a physical partition among the available workloads. But this is not cost effective because as the unused array was not unable to reuse again by another array. On the other hand, it is very difficult for the storage administrators to efficiently manage the multiple storage arrays, as the arrays are purchased at different times and from different vendors with different management techniques. Moreover, with the main idea of multi tenancy, the software defined storage system infrastructure was built. If when all the applications has to come from a single system, then it will be very much critical to separate workloads and to provide self provision storage by every tenant. This ultimately leads to the separation of performance among all the tenants, so that all the tenants has the flexibility to quickly reallocate the unoccupied storage for some other applications running at the same time or may transfer the storage to some other tenants. Chargeback is the another important aspect of multi tenancy. Tracking the usage and cost of any particular applications is a major problem for the traditional SAN and NAS storage system, whereas in case of multi tenancy, understanding the cost incurred by every tenant is the main thing. Software Defined

Storage designed for the cloud, like Zadara Storage, has built in metering and billing capabilities.

4. Outcomes of SDDC

Software defined storage centre provides business outcomes over two main areas. First, because of its abundance need it enables all the IT companies to transfer its resources towards innovative business growth with a free travel of no more worries on rate-limiting for launching new projects in their business. Secondly, software defined storage centre ensures all its companies to deliver a secure and resilient services with the two main key outcomes–Security control for the underlying infrastructure and high availability with resilient infrastructure.

5. Conclusion

In future the whole IT Industry will trend towards Software defined architecture. When compared to the traditional legacy storage like SAN, NAN etc, efforts required to install and to manage Software defined network is very much lesser. In this paper, a detailed study is given on the software defined data center and its supporting networks along with the criteria like cost, agility, scalability, reliability and multi tenancy which makes SDDC as a better one than the legacy storage.

References

- [1] <https://www.sdxcentral.com/sdn/definitions/what-the-definition-of-software-defined-networking-sdn/>
- [2] <https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/technical-whitepaper-sddc-capabilities-itoutcomes-white-paper.pdf>
- [3] <https://www.networkcomputing.com/storage/6-ways-transform-legacy-data-storage-infrastructure/213214384>
- [4] http://tdhpe.techdata.eu/Documents/gap_analysis_uk/HPQ3_July_Software_Defined_Storage_eBook.pdf?epslanguage=en
- [5] https://www.suse.com/docrep/documents/kknpw8ets2/intel_suse_sb_hi.pdf
- [6] Gupta SC & Anita Goel, “Software Defined Storage Technology”, *Asia-Pacific Software Engineering Conference*, (2015).