

White Paper

Flash Forward – a guide to new ways to do storage

Contents

Digitization increases data traffic	2
Starting point: Choosing the scalability approach	3
Architectures for independent storage scalability	4
All-flash or hybrid – which way to go?	5
When is Software-defined Storage (SDS) the right way?	6
When to move into hyper-converged IT	7
Fujitsu puts you on the right track	8



Digitization increases data traffic

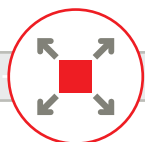
Digitization has a big impact on storage. More data sources create more data, which needs to be gathered, transmitted, processed and stored. Experts predict a growth in data volumes by a factor of 10 in the next three years. This will lead to higher storage demands in terms of response time as well as faster and greater scalability. In order to keep the costs of this explosive growth under control, the costs per terabyte of data need to be reduced significantly. Since storage systems have to host larger data volumes, customers require longer life cycles because the migration of hundreds of terabytes or even petabytes of data is a complex task.

Not all of these new requirements can be solved with one single storage architecture. This is the reason for the emergence of new and innovative storage designs – such as all-flash systems, software-defined and hyper-scale architectures, plus hyper-converged solutions. They add to the classical controller-based hybrid storage designs which host disks and SSDs. What's more, many storage categories partially overlap, making it difficult for enterprises to find just the right technology for their requirements. Expert guidance can make life a little easier. The following considerations are intended to provide you with guidance in choosing the storage approach that is best for you.



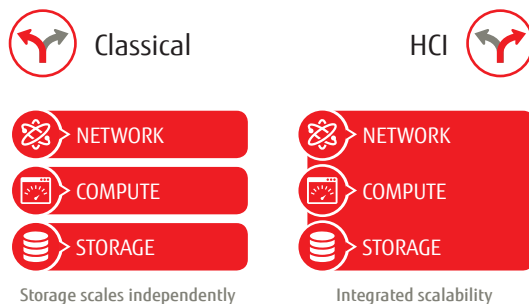
Starting point: Choosing the scalability approach

When selecting the right storage for a particular use case, the starting point involves thinking about which overall scalability approach should be deployed with the new storage solution. A key question is whether the demand for additional storage capacity also requires a linear growth of compute power, or whether these growth paths are only loosely coupled.



In a **classical architecture** there are discrete server, storage and network entities – this is why they are also called “3-tier architectures.” This is an ideal architecture for general purpose scenarios intended to consolidate as many workloads as possible in one central data center infrastructure. It also provides high flexibility when selecting storage, server and networking equipment. The purchase cycles of all three can be decoupled to a certain extent. Currently most mission-critical business applications run on this type of architecture.

Another approach is to merge storage, computing and some networking functionalities in one **hyper-converged infrastructure (HCI)**. In this hyper-converged system, computing and storage functions are not separated in different hardware elements, but are instead provided from one “building block” with one embedded management system. Scalability is achieved by adding more blocks. The growth rates of HCI are amazing. Various analysts predict growth rates of up to several hundred percent over the next few years. Despite the steep growth rate of HCI, it is a safe assumption that both scalability models will coexist for at least the next ten years.



CLASSICAL ARCHITECTURE

If you are thinking in terms of classical architectures, the next question is **what the core application requirements will be in relation to the storage system.**

Applications like server or desktop virtualization, databases or transactional applications, classical CRM, ERP and SCM, plus several types of analytical applications, require short response times from storage systems. Even more important are **reliable response times** in order to prevent transactional application time-outs. Of course, there are also huge amounts of unstructured data which are seldom accessed and do not require high storage performance. Due to the sheer amount of data, the key requirement is to achieve the **lowest cost per capacity** possible. Scenarios to consider are online archive/content depots, 2nd tier cloud storage, gathering historical data for analytical tasks or media streaming.



Reliable performance



- Server virtualization, VDI
- Database/OLTP
- Business applications
- Analytics (e.g. SAP HANA)



Lowest cost of capacity



- Unstructured data
- Content depots/online archives
- Historical data for analytics
- Media streaming

Architectures for independent storage scalability

Whenever high and reliable performance is needed, **all-flash storage** arrays are probably the ideal storage solution. All storage media are non-mechanical and deliver the lowest latency. Scalability is achieved by adding additional shelves (scale up), and fast internal high-speed connections avoid bottlenecks. The system software is optimized for the use of solid state disks (SSDs). The adoption of all-flash storage is rising at a very fast pace because the price – in comparison with fast-spinning hard disks – is eroding dramatically, and data reduction technologies are further reducing capacity costs. This is also supported by many operational benefits like heavily reduced power consumption, less space and less administration effort (e.g. no performance tuning, no replacement of disks).

On the other end of the scale, the best architecture for hosting huge volumes of rapidly growing data is scale-out, **software-defined storage (SDS)**. Essentially the storage software transforms the disks of connected servers into a storage pool. The software ensures high availability by automatically placing multiple copies of the data on different nodes. This also supports adding or replacing hardware nodes on the fly for easy scalability and easy

maintenance. Another advantage is the fact that the life cycle can be extended by many years, thus reducing migration efforts. As such architectures typically require a couple of nodes as a minimum configuration, efficiencies of SDS start in the range of 200 to 300 terabytes. Some can scale in extreme petabyte ranges – so there are practically no limits.

In larger environments the combination of all-flash as a performance tier and scale-out, disk-based SDS as a capacity tier may be an ideal combination, providing that the data volume is big enough.

There are a lot of scenarios where it may be necessary to host data for performance-hungry applications like virtualization, databases or OLTP alongside unstructured data, online archives and other “lukewarm” data in a single system. This is where **hybrid storage** systems, using SSDs and large capacity disks, are still the ideal solution. They are true all-rounders in IT environments and allow you to balance speed, capacity and costs in one single storage system.



Reliable performance



ALL-FLASH STORAGE

- Fastest response time
- Lowest operational costs (power, space, maintenance, admin)



HYBRID DISK AND FLASH STORAGE

- All-in-one
- Balancing capacity, speed and costs in one system

Lowest cost of capacity



SOFTWARE-DEFINED, HYPERSCALE STORAGE

- Extreme scalability
- Lowest costs (open source)
- Object and OpenStack storage

All-flash or hybrid – which way to go?



GO HYBRID if you plan to consolidate structured and unstructured data in one system.

HYBRID

If your objective is to consolidate all your data tiers in one or very few systems, then hybrid systems are the best all-in-one solution. By combining fast SSDs with cost-effective low-spinning disks, they allow you to balance speed, capacity and costs in one single storage system. In contrast to all-flash arrays, however, you may encounter higher operational efforts. But if you use modern automation tools, you can compensate for a fair share of the administration workload. Look for storage systems that let you predefine response times for data volumes. In some systems all internal adjustments like bandwidth allocation or storage tiering are made automatically. Fujitsu offers such technologies.

If you are thinking about whether your next storage replacement should be an all-flash system, or whether it is better to again use a hybrid system for the next couple of years, here are some simple recommendations – assuming that your usage scenario is similar.



GO ALL-FLASH when you mainly use fast disks.

ALL-FLASH

If your current storage systems are mainly equipped with fast-spinning disks, you should definitely consider an all-flash array as your next storage platform. The price per capacity for SSDs is still four times higher than fast-spinning disks – but this price gap will probably disappear in the next two years. Thus any additional SSD capacity which is added during the second half of your investment cycle will be on the same cost level. Short-term data reduction technologies like deduplication and compression will also significantly close the price gap. Since the effect of these two technologies depends on the data type, make sure that you invest in systems that can use these data reduction technologies in very flexible

ways – and definitely avoid architectures in which you cannot switch them off.

All-flash arrays are unbeatable from an operational cost perspective. SSDs seldom fail, whereas several disks need to be replaced every year. This alone can reduce maintenance costs up to 80%. Furthermore, all-flash systems need 10 times less power and 10 times less space. The operational workload is further reduced because there are no performance optimization or tuning tasks. Overall you can expect that the administrator workload will decrease by 30% to 40% so that the same IT team can manage more data or more processes.

When is Software-defined Storage (SDS) the right way?

In terms of performance, the use of high capacity disks, the synchronization of multiple data copies and the networking of storage nodes add a lot of latency, making SDS suitable for applications which do not require low latency. Thus SDS today is mainly used as a 2nd tier storage environment for unstructured data. This may change over time with the emergence of all-flash, scale-out SDS systems, so eventually SDS will find more application areas.



The adoption of SDS is gaining momentum because the architecture promises a robust and secure system with easy and efficient scalability of capacity and performance. Cloud native applications and analytics which require efficient, scalable solutions that help you get more out of your data are the main driver. There is considerable saving potential for overall costs because industry-standard server hardware is used, and various software-defined storage platforms are available as open source products. Probably the most mature and scalable SDS open source platform is Ceph, an open source, software-defined storage platform designed for object, block and file storage from a distributed x86 server cluster. Ceph supports scalable clusters up to the Exabyte level.

- Extreme and flexible scalability
- Extended life cycle
- No or fewer migrations needed
- Lower costs (with open source SDS)

- Difficult to build yourself
- Hidden costs – TCO risks!
- Higher latency
- Possible vendor lock-in!

Recommendation:
Look for complete SDS solutions

A complete SDS setup consists of many components. Servers, a network which connects the servers and a storage software suite which takes care of the storage and data management functions. This involves a lot of integration work and requires substantial testing and evaluation in order to build a flawless and balanced system. Network performance management and incident root cause analysis can be very challenging, not to mention keeping software and hardware updates in sync – a task which is done by the vendor when it comes to traditional storage. The complexity requires expertise going beyond the skills of a storage administrator, and this is keeping IT organizations from adopting SDS.

SW-DEFINED

In order to facilitate the deployment and operation, Fujitsu offers integrated hardware and software solutions with support and maintenance services. The system is based on open standards throughout and optimized Ceph Open Source Software.

When to move into hyper-converged IT

Today HCI has become more than just a niche solution. Analysts expect that 20% of all applications will run on hyper-converged infrastructures by 2020. Various kinds of external service providers are the biggest early adopters. Even if HCIs replace some traditional systems, it is quite probable that they will coexist with classical storage systems.

HYPER-CONVERGED INFRASTRUCTURE



- Servers act also as storage hardware
- Everything is virtualized (server, storage, network)
- High and modular scalability

In most cases HCI is based on x86 servers that function both as server and storage systems. As servers play a dual role by providing storage functionality and serving as an application platform as well, this can dramatically reduce the diversity of system hardware. Storage capacity and computer power scale hand in hand. And since everything is virtualized and distributed, it is easy to expand the IT infrastructure and perform hardware refreshes – without downtime and without any migration. What's more, server administrators usually have or can quickly acquire expertise based on their experience in server virtualization.

HCIs are less suitable for scenarios that need high scalability in data volumes but do not require much compute power to move data. As the applications, the virtualization software layers and the data management rely on the same server resources there can be situations where the quality of service of the storage function is impacted resulting in insufficient response times. As an HCI is typically a distributed environment the network and the ongoing data copy transactions between servers add latency. This may require complex tuning and optimization efforts on the operational level but also in adapting server and network configurations. Finally, HCI could add another storage silo into data centers which could counteract operational efficiency and end up in adding another platform with its own provisioning, management, backup and disaster recovery, and capacity planning tools.



- Less suitable for heterogeneous storage workloads
- Maintaining QoS for data
- Additional storage silo

HYPER-CONVERGED

Microsoft and VMware are the leading commercial vendors of integrated virtualization stacks for HCI. There are also open source developments like OpenStack which are bundled by distributions like Red Hat or SUSE into hyper-converged software stacks. Fujitsu has strong partnerships with all of these vendors. The advantage

of HCI software provided by these vendors is that HCI can be run on any suitable hardware. Thus hardware and software support – as well as management – will be separate. This means that the challenges of implementation, operation and maintenance could increase operational effort, costs and risks.



Recommendation:
Look for complete HCI solutions

That is why Fujitsu offers complete hyper-converged solutions comprised of the right hardware, software stacks from VMware, Microsoft and from an open source like OpenStack. These solutions are fully supported and maintained, allowing you to fully benefit from new agile IT concepts without any pitfalls after purchase.

Fujitsu puts you on the right track

There are many new ways to handle storage. All have their particular advantages for individual application scenarios. As a vendor of storage, servers and hyper-converged solutions, we support you in finding the right storage approach to your requirements.

ALL-FLASH

We offer **ETERNUS AF all-flash** arrays with leading-edge performance and the highest levels of operational efficiency. They support the flexible use of modern data reduction technologies and have the richest options for implementing cost-effective disaster recovery configurations.



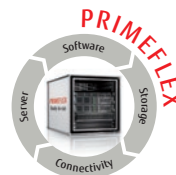
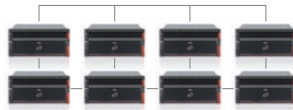
HYBRID



With **ETERNUS DX hybrid storage** we enable you to achieve a maximum of storage consolidation for structured and unstructured data by balancing speed, capacity and cost in one system. The latest automation technologies allow you to manage more data without additional IT personnel.

SW-DEFINED

ETERNUS CD10000 is an extremely scalable **SDS platform** ideal for cost-effective storage of huge volumes of 2nd tier and unstructured data. The combination of the right hardware with the most advanced open source storage software allows you to implement SDS in a fast way.



HYPER-CONVERGED

Our **PRIMFLEX hyper-converged solutions** for VMware, Microsoft and OpenStack deliver all the benefits of this new and agile approach to storage infrastructures.



When planning your next storage project, be sure to contact Fujitsu!

<http://business-datacenter.global.fujitsu.com/#storage>

Contact

Fujitsu Technology Solutions GmbH
Mies-van-der-Rohe-Straße 8
80807 Munich

© 2017 Fujitsu - All rights reserved, including intellectual property rights. Technical data subject to modifications and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.