

White Paper

Hyper! Hyper? Worthwhile or Nonsense?

Hyper-Convergence – The End of Classical IT?

Hyper-converged infrastructures are the rising stars in data centers. What is behind this hype, what are its key benefits and what does it mean in practice? Is hyper-converged always superior to traditional IT architectures? What is the impact on day-to-day operations, expandability and lifecycles? For which use cases is hyper-converged strongly recommendable or even a must? Let's have a deeper look to find the answers.



Content	
The silver bullet for the data center?	2
What is the promise of hyper-converged?	2
How do hyper-converged infrastructures look like?	2
The working principle	3
Potential planned downtime per year	4
What happens, if a server fails?	4
How long will data mirroring take?	5
Bottlenecks and how to counteract them	5
Architectures in comparison and what it means for you	6
Aspects to consider when taking your choice	6
What to consider when going for hyper-converged?	7
Stress-free deployment using FUJITSU Integrated System PRIMEFLEX	7
Summary	8

The silver bullet for the data center?

There is much noise in the market about hyper-converged infrastructures. Analysts predict an ever increasing market for the next couple of years, with a compound annual growth rate of over 60% by 2020, depending on the analyst report you are considering; although the starting point is already a solid base of deployments.

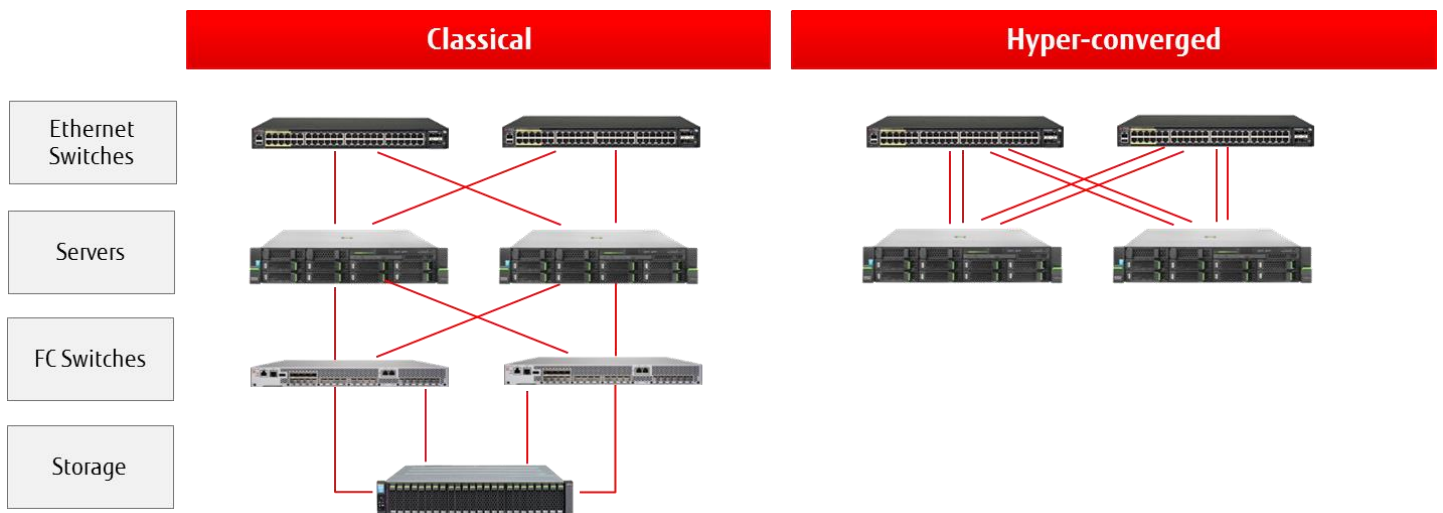
Constantly, articles are published which extraordinarily praise hyper-convergence and demonstrate a high number of benefits. This suggests that hyper-converged infrastructures are the silver bullet for the data center, and you could easily come to the conclusion that hyper-convergence will lead to the end of classical IT infrastructures. But is this really true? Is hyper-converged the right answer to every situation? Is it really jack of all trades? Is it worthwhile, or is there also some overstatement behind?

What is the promise of hyper-converged?

Having compute and storage resources integrated in a single box means less infrastructure components, less data center footprint, less energy consumption and less cooling requirements. The single pane of glass management for both, compute and storage resources, simplifies administration and reduces skills demands, which in turn means that lower paid staff can do what higher paid staff did before. Due to their flexible scalability, most hyper-converged infrastructures can be easily aligned to growing business demands, while built-in high availability ensures business continuity. Finally, all benefits aforementioned have often a positive impact on cost, operational expenditure in particular.

How do hyper-converged infrastructures look like?

Before having a closer look at hyper-converged infrastructures, let us remember the well-known classical virtualized infrastructure approach with several servers interconnected in a Local Area Network (LAN) by redundant Ethernet switches. The data resides on an external storage system which can be accessed through a Storage Area Network (SAN) whose core elements are usually redundant Fiber Channel (FC) switches.



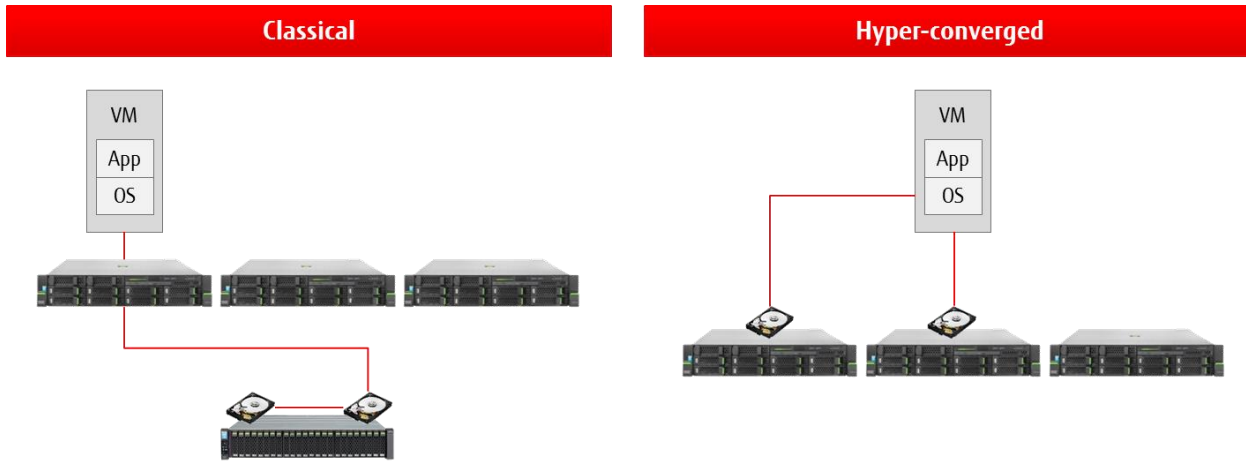
Hyper-converged infrastructures are also built from multiple servers interconnected in a LAN by redundant Ethernet switches. But in contrast to classical architectures, storage spreads across the local disks of the server nodes, making a central storage system and its dedicated physical SAN with its management, or a central Network-Attached Storage (NAS) superfluous.

The built-in storage features, such as data replication, deduplication and compression, turn hyper-converged infrastructures into a software-defined storage platform. Solid State Disks (SSD) can be used as a cache contributing to an enormous performance boost. High availability is built-in, because whenever one of the server nodes fails, another server can take over the workload running on the failed one, and data is not lost due to data replication. Compute performance and storage capacity can be scaled easily, just by adding or removing servers.

To be fair, we should not conceal that also the components of classical virtualized infrastructures can be reduced by using the same Ethernet switches for the LAN and the SAN, or by using a NAS in the same physical LAN. But in all these cases, we are still faced with two logically separated networks, and therefore a separate storage management is still required.

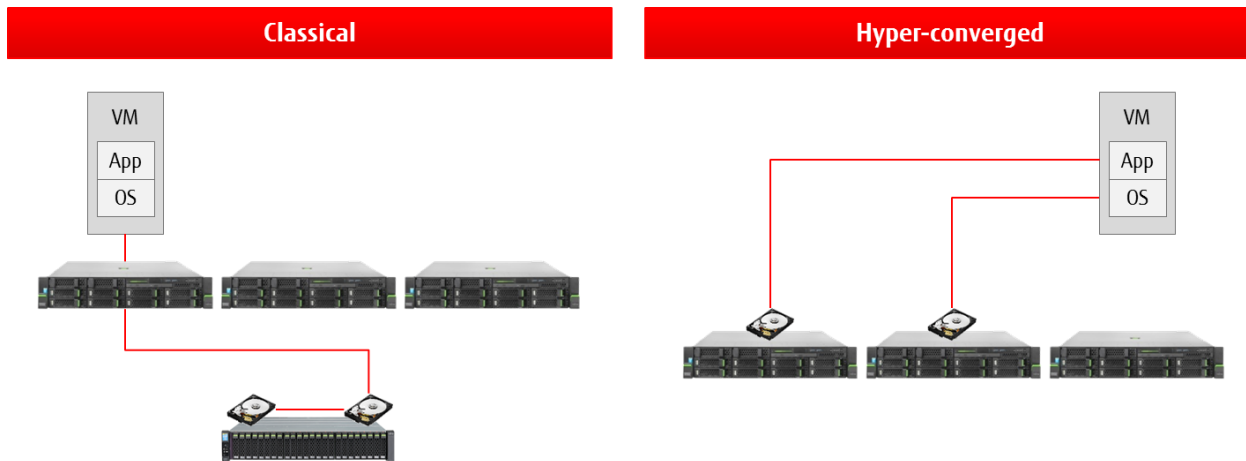
The working principle

Let us start with the classical architecture, where virtual machines (VMs) run on the servers and their data resides on the central (external) storage system. If you have to move VMs from one server to another one, e.g. because of planned downtime, the VM can still access their data on the central storage. In order to protect data against being lost in the event of a disk failure, any type of RAID level should be implemented in the storage system.



In a hyper-converged architecture, the situation is somewhat different. If a server fails, for sure, the VMs running on that server will also be restarted on another server of the cluster. But as there is no external storage, server failure affects all the data which reside on this server, too. In order to prevent data loss in the event of a server failure, the data connected to a VM is redundantly distributed over different server nodes by data replication.

It is worth mentioning that a VM does not necessarily run where its data is. Having the VM and its data on the same server is certainly an advantage for read operations; but when it comes to write operations, you will always have to move data through the network to another server. Otherwise you would not be able to leverage all benefits of virtualization, e.g. providing the flexibility to move VMs to any node in the server cluster. This means, that data locality does not really matter in any case, quite in contrast to what some vendors try to make their customers believe. Remember with write operations, it is always the slowest member which determines the overall speed.



Potential planned downtime per year

Why are we talking about planned downtime? The answer is fairly simple: Statistically, every server is out of order in the range of 4 to 19 times per year, according to plan. 3 to 8 times, the hypervisor requires an update, what may happen during operation. At the end, it is just the restart which causes approximately 15 - 20 minutes of downtime.

Apart from this, some firmware updates are due again and again. With the integrated remote management controller, this occurs maximally once per year at the average during operation with a subsequent restart of the controller itself which usually causes no downtime. The BIOS requires 1 to 6 encapsulated updates with a subsequent restart causing ca. 30 minutes of downtime. Finally, disk controllers as well as I/O adapters require max. 2 encapsulated updates with a subsequent restart causing some 20 minutes of downtime.

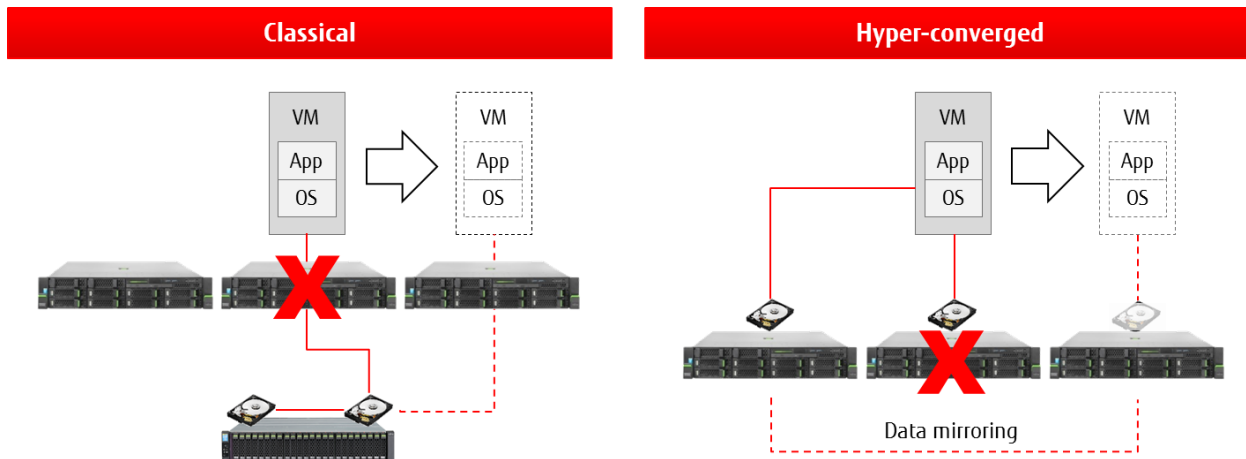
Apart from the servers, storage firmware is updated maximum 3 times per year at the average. Here, a controller restart taking some 15 minutes is required during which the controller itself is not available. But this restart causes no downtime of the overall system, if at least two controllers are built in.

What happens, if a server fails?

Let us now consider what happens, if a server fails. In the classical architecture, the VMs usually running on the failed server will be restarted on another server node. The data of the VMs on the central storage is not affected, and is still accessible by the VMs.

If your servers are not configured for more performance than needed, every server failure will degrade performance, because a smaller number of servers have to take over all the load of the failed servers, in addition to their usual load. And this will be critical especially, if it comes to the rare event that several servers fail at about the same time. However, the good thing is that the operation can still go on. Just if once the central storage system fails, all VMs will be stopped, and the entire operation will be out of order.

If a server fails in a hyper-converged infrastructure, the VMs running on that server will also be restarted on another server. But in addition, we have to bear in mind, that data is also affected. To ensure that the next planned or unplanned server downtime will cause no data loss, the original data redundancy has to be rebuilt; e.g. by mirroring the data to other intact servers. While data is being mirrored, operation is going on. Similar to classical environments, spare compute capacity is needed in hyper-converged environments in order to avoid performance degradation. But the servers need to provide the extra storage capacity, in addition.



How long will data mirroring take?

It sounds trivial that in the event of a server failure, data copies need to be mirrored between servers. But do you have an idea how long data mirroring will take?

Data mirroring will not just happen while clicking one's fingers. Assuming you have a 10 Gigabit Ethernet with a maximum throughput of 80%, transferring 1 TB of data through the network will take 17 minutes. Transferring 20 TB will take 5 hours and 33 minutes. Conducting the same exercise in a 1 Gigabit Ethernet will take 2 hours 47 minutes with 1 TB, and 55 hours 33 minutes with 20 TB. The conclusion might be that you are bound to 10 Gigabit Ethernet when going for hyper-converged.

It is true, that built-in data services, such as deduplication and compression, may reduce the data volumes by a factor of 7; nonetheless data mirroring will always take a considerable amount of time. And this exactly will be the critical thing, if the server, on which the data to be mirrored resides, fails before data mirroring is completed.

In this case, at least a part of the operation will be out of order. It is only a part of the operation, because VMs not depending on the data mirroring can still continue running. But usually it is not quite clear what is still running and what is stopped; and this is an uncertainty which is usually not acceptable.

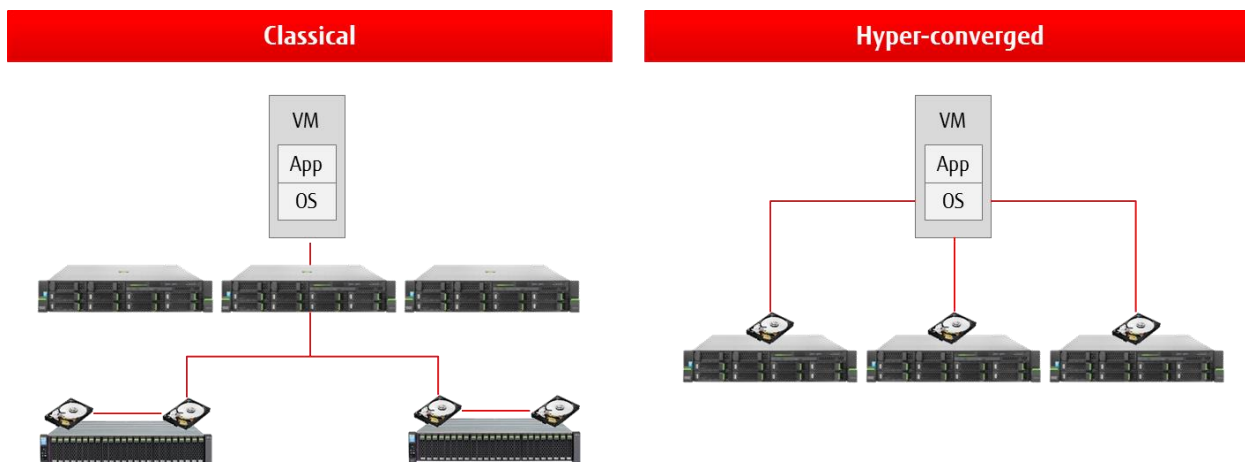
Bottlenecks and how to counteract them

From what we described afore, we recognize that in both cases, with classical and with hyper-converged infrastructures, there are bottlenecks and events which compromise or even stop operation.

Doubtlessly, there are options to counteract these bottlenecks. In the classical approach, you may just add a second external storage system to have an insurance against the failure of one of them, while of course more than doubling your storage budget.

In the hyper-converged approach, you may apply multi-mirroring of all data. Having multiple data copies reduces not only the risk of data loss, but also potential interruption times in the event of server failures. But for sure, multi-mirroring requires many more disks and therefore a higher budget. But hoping that after the failure of a server no further server will fail before data mirroring is completed is perhaps not the safest foundation for business continuity.

At the end, it is just a matter of investment and it is about the question what you are willing to pay for which service level.



Architectures in comparison and what it means for you

Let us now summarize the main characteristics of both architectures. A classical architecture includes servers, storage systems, as well as a LAN and a SAN or NAS. Consequently, you have to look after the lifecycle management for multiple components with different lifecycles; and you have to maintain two separate networks. Storage activities (also denoted as data services) such as snapshots are separated from servers. The classical architecture scales by adding multiple components, primarily separately. And it is about an approach which has been established for many years.

A hyper-converged architecture is based on servers and a LAN, with storage being included in the servers. While lifecycle management is needed for fewer components compared to the classical architecture, a more powerful LAN is required. Servers are involved in every storage activity, which should be considered when sizing the compute resources. According to our design guidelines, up to 10% of the CPU performance should be considered as storage activity overhead. Hyper-converged architectures scale by adding servers, primarily in a linear manner, which means compute and storage resources usually scale in tandem. And we should bear in mind that the hyper-converged approach is less established, because it just entered the market some 3 years ago.

What do these differences mean for you? Classical architectures are used for general purposes. They support virtual environments including multi-hypervisor environments and even physical (bare-metal) environments, if necessary. There is certainly a high level of hardware dependence, primarily related to storage. Typically the operational complexity is higher compared to hyper-converged architectures. Hardware cost is possibly higher, but software cost may be lower instead. The vendors of classical architectures that you find in the market are mainly well-established ones.

Hyper-converged architectures are usually tied to virtualization. There is basically no hardware dependence at all, but software dependence related to the virtualization software instead. Compared to classical architectures, operational complexity is definitely lower. Hardware costs are possibly lower, but software costs may be higher instead. Hyper-converged infrastructure solutions are mainly driven by newcomers, but more and more established vendors are joining.

As you can see, both approaches provide unique advantages. This means that there are good reasons for both architectures to coexist and to be taken into account.

Aspects to consider when taking your choice

In order to give you some guidance, we are subsequently addressing a number of aspects which should be considered when taking your decision for or against hyper-converged.

If your workloads scale horizontally, hyper-converged architectures are a perfect fit, especially if compute and storage resources need to scale in tandem. This applies to workloads which require a fixed amount of CPU performance, main memory, disk space and IOPS. Typical examples are Hosted Virtual Desktops and Hosted Shared Desktops.

This does not necessarily mean that hyper-converged is no option for workloads where compute and storage resources don't scale in tandem. Running different applications with different needs in terms of storage and compute resources may even improve resource utilization and performance compared to uniform workloads. As in every virtual infrastructure, there is a limited pool of resources available that has to be shared among multiple VMs running the workloads. To utilize these resources optimally, an application landscape with balanced resource utilization would be ideal rather than just having compute or storage intensive workloads in parallel.

If compute and storage resources don't scale in tandem, it might look at the first glance as if scaling compute resources brings unused storage resources with it, or scaling storage resources adds compute resources which are not needed. But a cost consideration could turn this perception on its head, considering the fact that you save a high amount of cost for the external storage infrastructure and especially its operation. That's why hyper-converged has become an attractive option for remote offices and branch offices.

If your workloads benefit from the built-in data services, hyper-converged certainly represents an advantage for you. You may just use these services without spending any additional budget. On the other hand, if you don't need them, you might be wondering whether you have indirectly paid for something you don't use.

The unified management of compute and storage resources certainly simplifies administration and reduces administration efforts. But bear in mind that going this new way will change existing staff roles and require other organizational changes. Think about whether this is reasonable for you.

Is there any intention to repurpose your new infrastructure any time in the future? If the uniform building blocks of hyper-converged without external storage are not appropriate, a classical approach might be recommendable.

Hyper-converged infrastructures are less appropriate for workloads that need a granular expansion on the component level. Likewise you should not go hyper-converged with monolithic applications, which can cope with increasing data volumes only using a scale-up approach. Moreover, there are still workloads which run optimally on bare-metal (because a virtual environment would be ineffective and slow down the application); therefore a hyper-converged platform with virtualized compute and storage resources will not be extremely helpful. Moreover, if a mixed operation of multiple hypervisors and bare-metal is needed, hyper-converged will not be an option either. And finally the same is true, if the rare case occurs, that an application generates more IOPS than an individual node of a hyper-converged infrastructure is able to cope with.

And last but not least, there is the important aspect of cost. As mentioned before, operational expenditure always tends to be much lower with hyper-converged infrastructures compared to classical ones. But when it comes to capital expenditure, it is hard to make a general statement. You will find lots of examples with cost advantages on either side. Make a simple cost comparison for your concrete project and you will find out which is the more cost-effective option for you.

Our strong recommendation is to decide per each use case. Go for hyper-converged, if the benefits outweigh the drawbacks.

What to consider when going for hyper-converged?

Coming from classical architectures and going for hyper-convergence requires a certain rethinking. Most important is the question how to deal with potential downtime. Will you reduce the risk of downtime by multi-mirroring or rather aim the gap and hope that data mirroring will never be disrupted by further failures.

But it is also about considering spare capacity in terms of processor performance, memory, disk space and IOPS on every server. It is true that servers may be sized individually, but you are well advised to size them equally. Life will be much easier for the scheduler of your hyper-converged infrastructure, and of course it will consume fewer system resources, then.

A last recommendation is related to software licenses. While the hardware costs decrease, the overall software costs increase in hyper converged environments. Typically, the software licenses are based on the number of CPU cores, what implies that fewer but large servers with less CPU cores in total help save money; similar to what you may have experienced with the licensing of your operating systems and applications.

Stress-free deployment using FUJITSU Integrated System PRIMEFLEX

We have a clear picture now, how hyper-converged infrastructures look like compared to classical architectures, we discussed the working principles of both approaches and their behavior in the event of failures. But the question remains how to get there. There is no doubt that building an infrastructure is complex, the classical one certainly more than the hyper-converged one. You need to select the right components, procure and configure them, before you integrate the individual components onsite. As the compatibility of the components is not guaranteed at all, extensive testing is a must. The fact that these components may originate from multiple vendors does not make things easier. All these activities are time-consuming and expensive, while presenting businesses with multiple risks, that things won't work as desired at the end of the day. A deep knowledge of all components is required, and an understanding of their various dependencies on each other. And as every installation is different, maintenance will be complex, too. Therefore the question suggests itself, if there is a better way of doing it.

Of course there is a better way to go: FUJITSU Integrated System PRIMEFLEX, a pre-defined, pre-integrated and pre-tested combination of data center components, such as servers, storage, network connectivity and software with management software being a mandatory part. Based on best practices and real-life project experience, PRIMEFLEX systems are designed in a way that their components will work optimally together.

The benefits resulting from Integrated Systems are manifold. Complexity is reduced; you need less time for planning; you accelerate deployment; no trouble through trial-and-error testing, and you minimize the risk that anything won't work at the end of the day. Operation will be more efficient and maintenance efforts will be reduced. All these aspects help reduce cost, both CAPEX and OPEX. Finally, we should not ignore the fact that all these benefits enable IT organizations to focus on the really important aspects of the business.

The PRIMEFLEX line-up includes classical and hyper-converged integrated systems, both built from best-in-class components, either from Fujitsu itself or leading technology partners. PRIMEFLEX systems are either pre-installed in the factory, and arrive ready-to-run at the customer's site; or they are delivered as reference architectures giving you the flexibility to adapt them to your specific requirements. For all PRIMEFLEX reference architectures, installation and configuration guidelines are available as a standard. What is more, customized configurations based on PRIMEFLEX reference architectures are delivered ready-to-run as an option, too. PRIMEFLEX is supplemented by services throughout all lifecycle phases, either delivered by Fujitsu or our local partners.

Fujitsu has the longest track record in terms of integrated systems; the first system being shipped in 2002. Since then we have continuously optimized our processes for end-to-end solutions, be it in product management, quality assurance, manufacturing and support.

The following table shows the PRIMEFLEX integrated systems offerings addressing virtualization use cases including some of their main characteristics.

	PRIMEFLEX for VMware VSAN	PRIMEFLEX for VMware Cloud Foundation	PRIMEFLEX Cluster-in-a-box	PRIMEFLEX vShape
Hypervisor	VMware	VMware	Microsoft	VMware, Microsoft
Hyper-converged	Yes	Yes	Yes	No
Number of VMs	25 – thousands	230 – thousands	4 – 200	25 – thousands
Scalability	Very simple, 2-64 servers	Very simple, 8-192 (8x 24) servers	Very simple, 2-4 servers	Simple, on component level, 2-64 servers
Physical workload	No	No	Alternative to virtual	No
DR option	Sync / Async	Sync / Async	Async	Sync / Async
Delivery model	Reference architecture, Ready-to-run	Ready-to-run	Ready-to-run	Reference architecture, Ready-to-run

PRIMEFLEX for VMware VSAN is our hyper-converged system based on VMware vSphere and VMware Virtual SAN. The hyper-converged multi-rack system **PRIMEFLEX for VMware Cloud Foundation** goes even a step further. It represents a fully software-defined data center with compute, storage and networking resources being virtualized. At the same time it represents an attractive foundation for a cloud infrastructure.

PRIMEFLEX Cluster-in-a-box is our hyper-converged system, based on Microsoft technology, targeted at smaller deployments, such as small and medium businesses, but also remote offices and branch offices. Although it is ready-to-run, many components can be configured individually. As an option, workloads can even run in a bare-metal environment, if useful.

PRIMEFLEX vShape is our virtualization offering based on the classical architecture, giving you free choice in terms of hypervisor (VMware vSphere or Microsoft Hyper-V), storage (Fujitsu's ETERNUS DX or NetApp FAS) and flexible configuration options.

Summary

We started with the question whether considering hyper-converged would be worthwhile or nonsense. We pointed out a number of advantages such as fewer components, less data center space, as well as less energy and cooling requirements. Certainly, these advantages might be helpful for you, but are probably of no strong consequence. Ease of administration, in particular due to the omission of the central storage, might have a high relevance for you, because it directly translates into massive savings of operational costs. Another advantage appreciated by everyone is business continuity due to built-in high availability; but business continuity can also be achieved using the classical approach. Outstanding is the easy, fast and flexible adjustment to changing requirements, especially if scale-out is demanded.

So from our perspective, hyper-convergence is worthwhile, but it always depends on the use case; that's why it does not herald the end of classical IT yet. Fujitsu will consult its customers and elaborate the most reasonable architectural option. No matter, how this will look like, with the right integrated system from our PRIMEFLEX line-up we will help you deploy the new infrastructure – fast, efficiently and without any risk.

Contact

FUJITSU Technology Solutions GmbH
 Address: Mies-van-der-Rohe-Strasse 8,
 D-80807 Munich, Germany
 Website: www.fujitsu.com/primeflex
 2016-10-15 WW EN

© 2016 Fujitsu, the Fujitsu logo, and other Fujitsu trademarks are trademarks or registered trademarks of Fujitsu Limited in Japan and other countries. PRIMEFLEX is a registered trademark in Europe and other countries. Other company, product and service names may be trademarks or registered trademarks of their respective owners. Technical data subject to modification 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.