riverbed

White Paper

Riverbed Hyper-converged Edge

A Look at the History and Trade-Offs of Traditional Hyper-converged Infrastructure (HCI Architecture), and What You Need to Know Before Making a Decision That Will Affect Your Business Front-Lines

Hyper-converged Infrastructure is a revolutionary technology area that has recently seen massive adoption. It is a great infrastructure approach for Data Centers around the globe as the market has shown. However, being “stateful”, makes it an unfit choice for remote branch offices. Hyper-converged Edge takes the value that Hyper-converged Infrastructure (HCI) provides in the Data Center and projects it out to a “stateless” remote office.

Back in the 1990’s, Virtualization changed the way data centers around the world were designed and architected. In the mid-to-late 2000’s, we saw another revolution with the advent of Hyper-converged Infrastructures (HCI). According to IDC, years 2015/2016 will see HCI eclipse the $1 billion dollar revenue mark worldwide. This is a spectacular adoption statistic, since in 2014, the total worldwide revenue number for Hyper-convergence was just over $100 million. The technology is sound, scales well and is in line with the vision of the software defined data center (SDDC).

All major HCI vendors, the likes of Nutanix, Simplivity and even the big storage shops like Dell, EMC and NetApp have had their hat in the ring since the beginning. Although the technology can be very good, it only makes sense for the software defined data center - keyword data center. These are great service delivery apparatus, but not one of the HCI vendors has focused on the service consumption side. The consumers of these services primarily reside outside of the data center at remote offices—the important business front lines. HCI devices, inherently, are stateful beings, meaning they house real application data—company’s intellectual property. For ages, we have been trying to get the data out of branch and remote offices. Putting these in remote offices renders the branch stateful. Stateful branches contain real data which renders them highly susceptible to data loss and security breaches. We need a new way of architecting remote office IT. We want the benefits of the software defined Hyper-converged data center projected out to the remote offices. The solution would result in a stateless branch, meaning no data at these locations. Moving data services away from remote locations would cause network dependencies leading to bad user experience. We want stateless branches without the compromise of bad user experience.

We propose a new revolutionary architecture that takes the stateful-ness of your SDDC, projects it out to the edge with only the stateless bits at these remote offices. No need to provision or manage data in remote offices, no need to protect data in branches, and no need to suffer data loss during outages. We propose a Hyper-converged Infrastructure built and architected specifically for the remote offices. We at Riverbed, propose a Hyper-converged Edge solution - the world’s only stateless Hyper-converged Infrastructure built specifically for remote offices—the front lines of your business.

Why Hyper-converged?

Big data center technology vendors have had their own proprietary flavors of Converged Infrastructure for as long as they have been around. Converged Infrastructure came into being out of the need for a simple data center. Combining multiple aspects of the application stack (network, storage and compute) meant easy design, implementation and support. Simplicity and supportability were huge factors in the early adoption of such solutions since it avoided finger pointing between various groups in IT, sped up the service delivery process and gave data centers the proverbial, one throat to choke for all things IT. The only thing they lacked was scalability. Hence came the next generation of Converged Infrastructure - the Hyper-converged Infrastructures, short for Hyper-scale Converged Infrastructure. The first adopters of the Hyper-converged mantra were not the typical enterprise corporations but large software and computing shops, like Google, Facebook and Amazon etc. These DIYers came up with their own flavor of storage, compute and network that scaled without limits. They changed, not only the way services were delivered, but architected. From the inception of an idea to delivery of that service, the scale-out approach made things expeditious. The rest of us in the enterprise world wanted to follow, but did not have anything to match. The emergence of Hyper-converged Infrastructure gave us the following benefits.

Scalability

Enterprises are seeing unprecedented growth in their data and are looking for more agile ways to deliver their services. HCI benefits both IT organizations providing the services as well as the developers creating the applications. Scale-Out provides fault tolerance and self-healing properties without a single point of failure. Scale is about more than capacity - it’s also about performance. Since your business growth is unlimited, your IT should also scale without limits.

Simplicity

Pre HCI, a typical application delivery workflow required managing the network, storage and compute as separate entities each with their own management portals. The success and timeliness for the delivery of a new application depended upon the IT admin's hope of all these segmented silos playing harmoniously with each other. Combining these made manageability of these components much simpler. A single pane of glass to manage all components of a service delivery stack meant less reliance on multiple groups and less chance of human error.

Predictability

From a business perspective, this model is a lot more predictable than any of its predecessors. Instead of over-provisioning scale-up architectures to account for 3-5 year growth, now one can design for what is required and grow as needed i.e. pay for what you need today. The "just in time sizing" makes budgeting decisions easier as admins can predict when further scale might be required based on organizational growth. From a technical perspective, it also provides linearity in both capacity and performance. Hyper-converged infrastructure solutions are built as equal lego pieces hence stacking more of these provides a linear and predictable amount of all resources.

Consolidation with Convergence

HCI's align really well with not only the software defined next generation data centers but also with the CIO initiatives of consolidation. Consolidation has always been a primary point of focus since the beginning of the first branch office. It gives you greater control and governance of your intellectual property. It gives you better bang for your buck for a particular resource. Take storage as an example, almost 50% of all storage is unused. Combining multiple buckets of storage means less number of half empty buckets. Consolidation with convergence of functions gives you the ultimate bang for your buck. Instead of just consolidating your storage tier, convergence gives you the same business benefits for your entire IT infrastructure stack.

Remote Office Dilemma

We have seen that as IT administrators, Hyper-converged Infrastructures can provide a great mechanism for architecting and delivering services because they are inherently simple, scalable, and predictable. But now let's shift focus to the consumers of these services - our customers.

Majority of the revenue-generating workforce resides outside the data center at remote office branches. If we design a remote office's IT from a user's point of view, we find ourselves at a paradox as it conflicts with all the benefits mentioned above. Let's take a look at some of these.

Performance Dilemma

Performance equals productivity. Nowadays, performance and experience are two of the biggest factors in retaining a customer—internal or external. If we are not giving LAN speed performance to our users, they cannot function and hence productivity suffers. If we put a conventional HCI at these locations to solve the performance challenge, we forgo the simplicity and ease of centralized management, data protection, governance and support

WAN Dependency Dilemma

Many IT decision makers that end up consolidating everything, try to balance the scales by provisioning additional bandwidth and sometimes using traditional WAN optimization solutions. These work for numerous use cases and applications, however, as a user I am still WAN dependent. My experience becomes unpredictable depending upon the time of day and consumption of services, not to mention WAN outages that can completely inhibit my ability to get my work done.

Stateful Branch Inflation Dilemma

As IT admins, usually stuck between a rock and a hard place, we have to balance the consolidation-distributed service delivery mechanism based on individual applications. Some services do not work well over the WAN, also certain applications are mission critical and require near 100% up-time with no dependency on WAN or data center infrastructure. These normally end up at the branch. Once there is infrastructure to support stateful applications in the branch, they tend to grow and inflate beyond control. You start with one service, and with every user complaint, you end up placing more stateful apps in the branch, resulting in more and more unmanaged and unprotected data at these locations. Hence completely losing the control and consolidation benefits of traditional HCI.

Looking side by side from both the deliverer (admin) and the consumers (user's) perspective, we can see this catch-22

Table 1 Pros, cons and tradeoffs of traditional HCI at Data Center & Remote Branch Office

Solution	Service Deliverer (Admin in DC)	Service Consumer (User in Branch)
• Traditional HCI ONLY at the Data Center, nothing at the Branch	• Good Admin experience • Consolidation. Simplicity • Control • Zero risk of data loss	• Poor User Experience, Poor Performance, Poor Productivity
• Traditional HCI at Data Center and at the Branch	• Poor Admin experience • Distributed management • Data at risk • Stateful Branch take longer to provision and recover	• Good user experience, performance and productivity • Data Protection issues. High risk of loss of user data.

Hyper-converged Architecture

By now, we see the advantages and disadvantages of conventional HCI solutions. It clearly favors the admin and the data center, but is not best suited for the users in the remote offices. Before we re-architect the solution specifically for the branch, let's take a look at what a typical HCI architecture looks like and what architectural benefits it provides.

If we look at a typical application stack, it looks something like the following:

Figure 1 Single Application Stack

And architecturally, multiple applications, would look something like this.

Figure 2 Multiple Application Stack

Virtualization logically abstracted individual units of compute and gave us a horizontal compute layer. Making the stack looking like this.

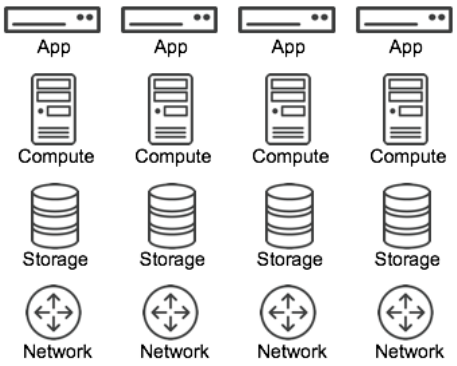


Figure 3 Virtualization In Application Stack

Storage and software defined networking vendors did the same for the other two layers.

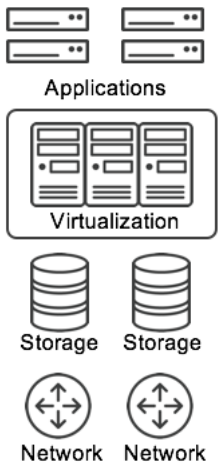
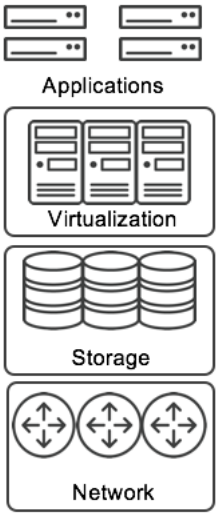


Figure 4 Virtualized Application Stack

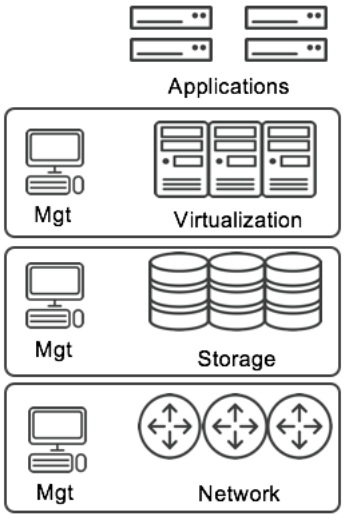
There were two main challenges with this picture.



Management Silos

Each of these layers, now abstracted from each other, required its own separate management tools causing silos of management to delay services. From an application delivery standpoint, all three layers had to be configured and provisioned individually before it can be deployed. Starting at the network layer, IPs, VLANs and firewalls had to be configured to support the need for new services. At the storage layer, LUNs and/or storage volumes had to be provisioned to be presented to the compute layer. At the virtualization layer, data stores had to be created, and Operating Systems deployed before the application owner can deploy that particular service. The picture looked something like this.

Figure 5 Separate Management tools for each layer of functionality



Scale Challenges

All of these layers individually only scaled to a certain limit and then an entirely new infrastructure had to be resurrected, each with its own management silos.

Figure 6 Multiple management silos

This caused a design challenge of over provisioning resources resulting in high initial costs and as the business grew, manageability and support became quite cumbersome.

A new architectural approach was needed.

Since all applications require the same three components (compute, storage and network), the Hyper-scale visionaries came up with an approach of combining all three into a single entity, called a Node, that can be managed through a single pane of glass.

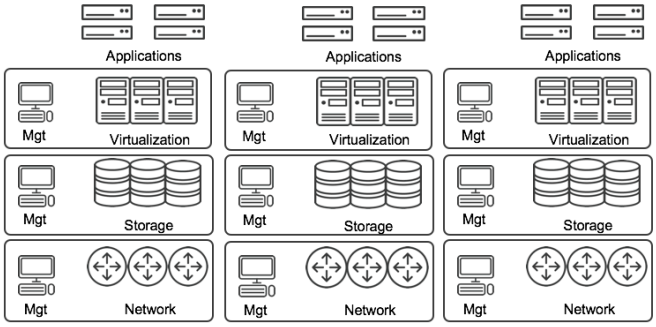


Figure 7 Single HCI Node

This solved the first challenge of manageability, ease and agility of application delivery. The second challenge was to have this infrastructure scale without limits. This gave way to a scale-out way of thinking instead of scaling up. Scale-up requires a single (or dual) brain atop a resource, and then is scaled by add individual resources underneath the master “brain” entity. It’s a master-slave, top down approach of resource allocation. Since each brain (or a set) can only handle a finite amount of resources, this approach could not be used.

A scale-out approach is radically different in the way that it does away with a single brain, master-slave mantra. Each component of the HCI acts as a peer to each other with each node being active. It provides the same services and resources as its peers. It provides fault tolerance to one another and allows peers to speak on a common messaging plane to maintain consistency and resiliency. The nodes scale horizontally and infinitely hence the term “scale-out” is appropriate for the architecture with applications sitting atop, none the wiser.

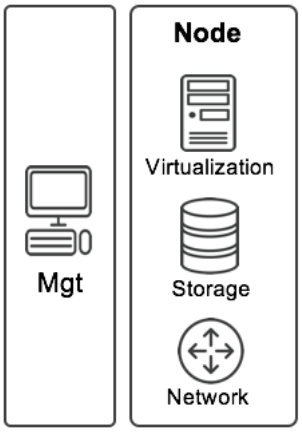
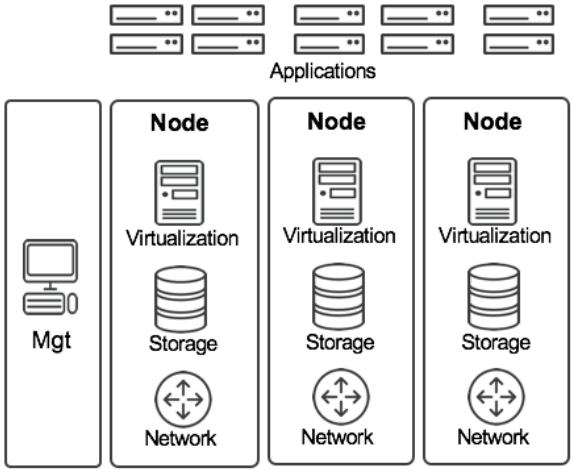


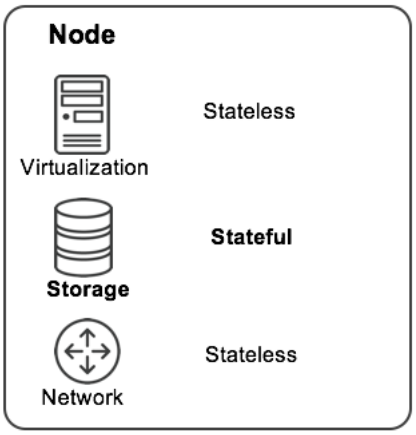
Figure 8 Multiple Nodes In HCI running Applications



Architectural Shift

The benefits of the HCI architecture are resoundingly clear, also proven by its massive adoption rates in the data centers. If we analyze the three components of an HCI node, we see two different characteristics. While compute and network components are stateless, storage is stateful.

Figure 9 Persistence Properties of an HCI Node



Stateful vs Stateless

We can determine the state of these resources by looking at the persistence properties of each. A question to determine the persistence of these resources would be 'if I lose one of these resources, do I lose just the availability or the viability of the service?' Losing network or compute, will just effect the availability of the service. An admin can move the application to a different network or another compute node and would get the same service without issues. Hence these resources are stateless. If, however, I lose the storage piece, I will require some level of effort to bring these back, if at all. I may have to resort to an older copy from a backup, I will likely lose data and effect the viability of the service. Hence storage or data is the stateful entity.

As an administrator, the majority of management is done on the storage layer. An admin has to make sure it's performant, protected via mulitple copies, and it even has an offsite copy for disaster scenarios. The stateless components are not as crucial to manage. One of the primary reasons of the consolidation movement is to get the storage or data closer to the admin.

But as a user, my performance and productivity depends upon how close I am to the compute component. A user interacts primarily with the compute and the compute interacts with the storage. If we design a 100% consolidated solution, the admin is satisfied, as we have a stateless branch, but the user suffers.

Figure 10 Admin and User experience without Infrastructure in Branch

If we place HCI also at the branch, the user gets better quality of service and experience, but results in a stateful Data Center and a stateful branch and hence the admin suffers.

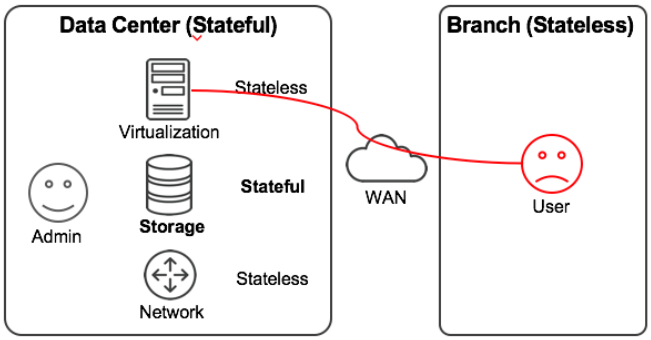


Figure 11 Admin and User experience with Infrastructure in the Branch

In order to achieve happiness at both sides, we need an architecture that keeps the stateful storage at the data center and moves the stateless compute and network close to the user. Theoretically speaking, even if we place just the compute and network at the branch, we are faced with a new challenge of wan dependency between compute and storage, which does not work well in reality.

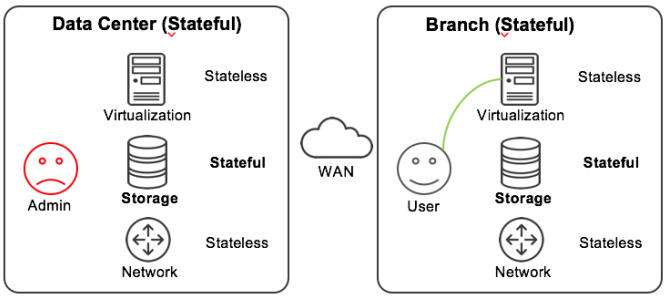


Figure 12 Compute and Network Infrastructure in the Branch

In order to satisfy the IO requirements between compute and storage, we need some level of caching at the edge that is not "real data" but an intelligently cached copy that only portrays what the users are working on. In other words, we need an imitation of the working subset of the data for compute purposes.

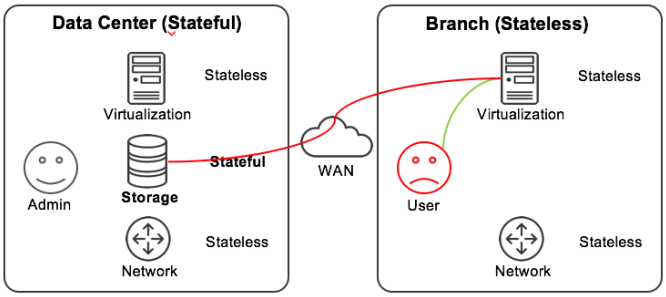


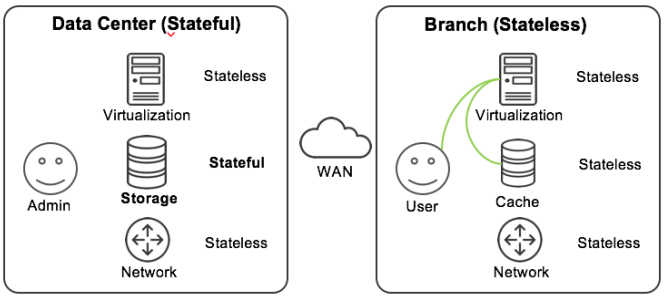
Figure 13 Admin and User experience with cache in Branch

Now we have achieved 100% consolidation, maintained statelessness at the remote office and have finally made both the admin and user equally productive.

The application is computed at the remote office with a working subset of data cached locally hence zero dependency on the WAN. This gives user LAN speed performance even in the event of a WAN outage.

All data is centralized, where it can be provisioned, managed and protected all within the confines and tools of the data center.

By combining the three components of app delivery mechanism at the edge (compute, storage or cache, and network), we can now achieve the same benefits of HCI at the remote offices.

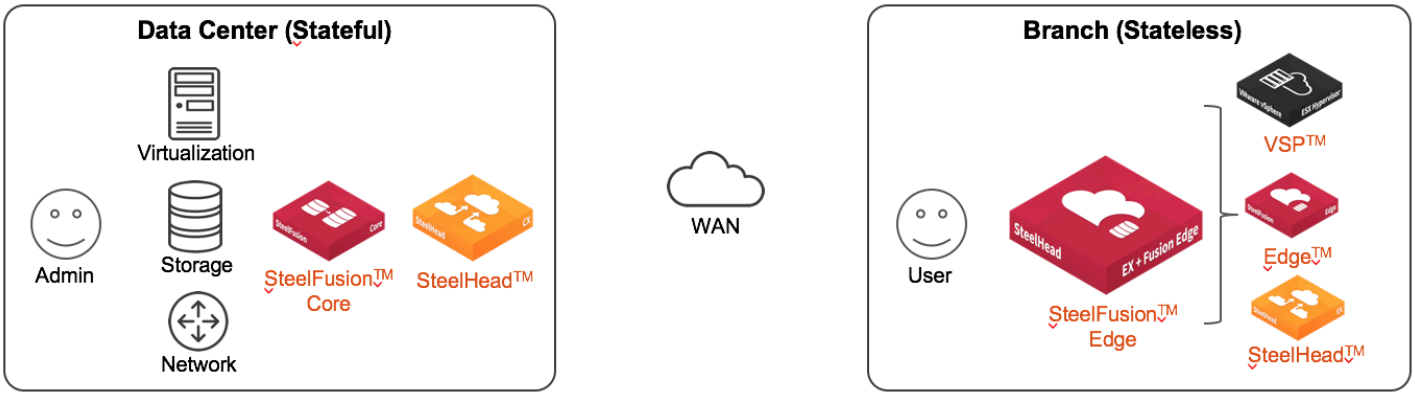


Hyper-converged Edge

Riverbed SteelFusion™ is the world's only Hyper-converged solution designed specifically for the remote office. It facilitates 100% consolidation with LAN speed user performance. It enables data center manageability of all infrastructure components, while giving local performance and WAN disconnected operations to the remote office users. It empowers organizations to extend the capability – services, the resilience, and the security – of their fully consolidated stateful data center to a stateless branch without compromising performance, experience and productivity of the end user.

The overall solution consists of three components.

Figure 14 Riverbed SteelFusion™ Solution



Data Center

At the data center, the two components are the Riverbed SteelFusion™ Core and Riverbed SteelHead™.

Riverbed SteelFusion™ Core

The SteelFusion™ Core acts as a broker and a single management portal to all remote office infrastructure. It is responsible for managing and projecting data to the SteelFusion™ Edge component discussed in the next section. It communicates with the backend storage and does not, itself, house any data.

Data Center SteelHead™

A SteelHead™ exists at the data center to accelerate the projected data in flight.

Both the SteelFusion™ Core and SteelHead™ can be in a highly available active-active configuration for increased performance and fault tolerance.

Remote Office

At the remote office, the Riverbed SteelFusion™ Edge appliance converges all three components of the application stack (compute, storage or cache, and network) in a single appliance. SteelFusion™ Edge consists of the following components.

Riverbed Virtual Services Platform

A compute layer called Virtual Services Platform (VSP™) runs VMWare ESX hypervisor (up to version 6.0). This is where all branch specific applications run. These applications are typically local to the specific remote office location. Examples of these applications are branch utility servers i.e. file, print, DHCP/DNS, AD etc.

Riverbed BlockStream™

VSP™ gets its storage from the BlockStream cache. All SteelFusion™ Edges communicate with the SteelFusion™ Core and are pushed the configurations, and the working set data that reside within this cache. Utilizing patented prefetch algorithms, only the working set data traverses the WAN and is cached in this layer. This ensures that the users are getting LAN speed performance no matter the state of the WAN, even in the case of a total black out. The new data created at the branch is immediately and continuously streamed back over the WAN through the SteelFusion™ Core to its final stateful resting place in the data center.

SteelHead™

At the network layer is the SteelHead™. It serves a dual purpose. Firstly, it is responsible for accelerating all the prefetched and delta change traffic between the SteelFusion™ Edge and SteelFusion™ Core for a near real time RPO (Recovery Point Objective). Secondly, it is responsible for optimizing all data center-to-user traffic to give the user a LAN-like experience even for the apps that are not locally executed on the VSP™.

Benefits of Hyper-converged Edge

Taking a Hyper-converged Edge Infrastructure approach provides many unprecedented advantages.

Consolidation and Performance

Typical consolidation initiatives favor the IT organization and are usually CIO run initiatives. The cost comes in the form of user performance, experience, security and productivity. Riverbed SteelFusion™ Edge provides the only Hyper-converged solution in the world that enables 100% consolidation of data without any compromise to the end user.

Centralized Data Management and Protection

All branch data is managed and protected centrally. SteelFusion™ solution automates and orchestrates the storage at the data center and enables consistent backups of all remote applications.

Unprecedented Business Continuity

With Riverbed SteelFusion™ solution, branch infrastructure provisioning and recovery times go from days and weeks, to hours and minutes. From the data center, simply point and shoot the data over to the remote office and boot/bring up applications over the WAN for instantaneous availability, anywhere in the world. In times of remote office failure, the same application can either be brought up in the data center and accelerated via the Riverbed SteelHead™, or be projected to another branch location closer to the user for better performance. In case of a primary data center failure, Riverbed FusionSync™ ensures that all branch data is consistently replicated to a secondary data center location for immediate disaster recovery.

SteelFusion™ empowers organizations to deliver an “always on” promise to their end users regardless of the networking or data center infrastructure.

Table 2 Advantages of Traditional Infrastructure in the Data Center, with a Stateless Hyper-converged Edge

Solution	Service Deliverer (Admin in DC)	Service Consumer (User in Branch)
• Traditional Infrastructure in the Data Center, with a Stateless Hyper-converged Edge	• Great Admin Experience • Everything is Centralized; Consolidation, Simplicity, Control, Agility • Zero risk of data loss • Instant provisioning of new services and remote sites • Instant recovery and centralized Backups	• Great User Experience • Performance high • Productivity high • Instant access to new services

Summary

Traditional Hyper-converged Infrastructure is a revolutionary step forward in support of a software defined scalable data center for the enterprise, but the architecture falls short when it comes to supporting the branch and remote offices – the front lines of business growth.

HCIs are great for application management and delivery. They are stateful and hence better suited in the data center. Looking at the numbers, the market also backs this claim as we see majority of the HCI sales being in the data center. Looking at the service consumption model, a new architectural approach is required that takes advantage of the stateful data center and maintains stateless-ness at the edge. The Riverbed SteelFusion™ Hyper-converged Edge is the world's only HCI solution built specifically for the remote office. It maintains 100% consolidation, complete manageability and protection at the data center, while still providing LAN speed performance and experience to the end user with no network dependency. It enables data center like agility of instant provisioning and recovery for remote offices all managed centrally. It enables the ultimate consolidation initiative that both, the business and IT can advocate.