

Unleashing Cloud Performance

Making the promise of the cloud a reality

UNLEASHING CLOUD PERFORMANCE

Introduction: The reality of cloud services

Thirty-five percent. By 2014, analysts believe that thirty-five percent, or over one-third, of global enterprise IT budgets will be spent on cloud services. While today the percentage of organizations investing in these types of offerings is small, the implication is clear: “the cloud” is not another industry buzz word, but a broad category which will drive the next phase of IT projects. For IT and business managers already inundated with information about the promise of a cloud centric infrastructure the question is not whether or not to use the cloud, but how. Fortunately, the virtualization and consolidation projects that ignited this renewed interest in centralized computing also provide a guide for how to best realize the full potential of the cloud by establishing the criteria that any project should try to address: How does IT more effectively understand, optimize, and consolidate resources to deliver the highest use and highest ROI for the business?

Defining Cloud Services

When the term ‘cloud’ comes up most people envision Amazon Web Services, Google, or Salesforce.com. While these vendors are certainly the icons of the cloud vision, and are attempting to become a one-stop shop for all of an enterprise’s IT needs they are all public cloud solutions and represent one model for cloud services that an organization may adopt. In addition to the public cloud, private clouds where organizations manage applications and infrastructure independently from a central location, and hybrid clouds that combine both public and private clouds round out the cloud services continuum.

Public cloud services like Amazon, Google, Salesforce, et al drive cost down and scale up by trying to support thousands, perhaps millions of businesses on their platforms. They centralize all of their services in a few data centers around the world, and try to deliver global services – at acceptable performance - from these locations. This is the purest vision of the cloud. It completely abstracts away all the complexity of dealing with physical IT infrastructure. Due to thin-provisioning and virtualization, these vendors can present a seemingly limitless data center infrastructure at a low monthly cost.

Private cloud services take the fundamental business and delivery model for public vendors and scale it down to delivering the computing capacity for an individual enterprise. For enterprises that have tens of thousands or hundreds of thousands of employees, they may also reach the tipping point where they can cost effectively provide the type of instant, seemingly endless computing and storage capacity that public vendors have.

By consolidating storage and applications, virtualizing infrastructure, and then providing acceleration to branch offices and mobile workers, businesses are beginning to create private cloud services. In essence, businesses are taking their physical data centers and changing the way they manage the services that run out of that data center. In addition to overcoming issues of availability, security, and lock-in, organizations see one other benefit to the private cloud model: dealing with sunk data center costs.

Many organizations have invested millions over the past two to four years to build the private data center capacity that they need to support their business for the next 5 to 10 years. With such a large investment (and the case that IT departments had to make to their CFO’s and boards), it’s unlikely that they would simply abandon those investments to use the cloud. Rather, they may shift their operational models within their own data centers to mimic what public cloud services are doing. This will give IT organizations to more cost effectively deliver internal services, and if they have a chargeback model they can drive these lower monthly costs back into the business units that consume IT resources and services.

Over time, businesses that adopt a private cloud model may more easily transition to a hybrid model that uses both private and public cloud models. In fact, in a recent discussion with a top-10 engineering company, its CTO looked at the public cloud as “flex-capacity” to support their private cloud infrastructure. As large projects come online, shift locations, or undergo other transitions, public cloud services may supplement their internal capacity to ensure that IT services are not a bottleneck to completing a revenue-generating project on schedule. Other companies may simply look to recoup their investment in their existing data centers and eventually transition completely to a public services model, making the assumption that by the time they begin that transition costs will be even lower and the other barriers to adoption will no longer be an issue.

Regardless of whether technologists choose the public cloud, the private cloud, or a mix of these services to deliver IT for their business, it’s important that they don’t lose sight of the end goal. Their mission is to provide a competitive advantage to their end users so that they can generate more revenue, more easily for the business. Above all, that means IT must provide them with the functionality they need at the fastest possible level of performance. Whether a business builds out its own cloud or buys cloud

services from someone else, they need to be sure that the end product is optimized for the speed that its own users need.

Understanding the limitations of cloud performance

Since the transition to cloud services looks much like the transition that occurred within enterprises over the last decade to centralize business applications to private clouds it is useful to consider the role that the network played in these projects to better understand how it may impact different types of cloud initiatives. Then too the impetus was driven by opportunities on both the business and technical side as business managers saw the chance to reduce costs and improve the security of distributed branch offices, and IT managers leveraged new technologies and more WAN bandwidth to make consolidation projects a high ROI priority for companies. In practice businesses often ran into a significant problem: distributed end users experienced slow wait times to access the business applications they needed to get their jobs done, and a greatly-increased set of demands that were placed on the wide area network (WAN). WAN optimization was used in many of these cases to overcome this performance issue. By including WAN optimization in consolidation projects, performance demands were met, bandwidth growth was controlled or eliminated, and consolidation projects were able to avoid trading off IT cost and management benefits with business performance requirements.

Bringing that recent-history lesson to bear on the cloud, it's quite easy to see some parallels. A public cloud may or may not be 'consolidation,' that is, an enterprise might be splitting its assets across its data center and someone else's cloud, but some of these assets will be living further away in that public cloud. This shift of assets looks much like the consolidation of assets from branch offices to the private cloud. And, as history showed, that shift of assets away from users leads to performance problems. There is an assumption on the part of the cloud provider that, if they manage the performance of their infrastructure within the cloud, performance *between* the cloud and the end user (whether it's a business user or IT user) will not impede the technical and business value proposition. This presumption should not be taken for granted, as cloud services will put more distance between users and their data resulting in greater latency that adversely affects performance.

Recent tests validate this issue. For example, below is an analysis of accessing a document management system in a LAN-like environment versus a cloud-like environment. The difference is not in the actions a user would make, but instead in the latency of the network in accessing an application.

Action/ Latency	-LAN (<10ms)	Cloud (~100ms)	Cloud (~200ms)
Login	2.08	23.08	34.54
"Reports" Tab	1.19	9.80	22.62
Upload 3MB Doc	26.74	88.54	109.03
Download	10.00	38.00	40.00

Figure 1. Document Management System – Performance Analysis

The results of these tests are consistent with concerns raised by in a paper by the UC Berkeley Reliable Adaptive Distributed Systems Laboratory which listed data transfer bottlenecks and performance unpredictability as two of the top five barriers for the growth of cloud computing.¹ They also match the experience that has driven thousands of organizations to rely on WAN optimization from Riverbed Technology to support their distributed environments in a non-cloud world. This is because cloud does not eliminate a fundamental set of challenges associated with distributed computing: bandwidth is limited and costly, while latency still limits the performance of an application over a network.

Private vs. Public: Does it make a difference?

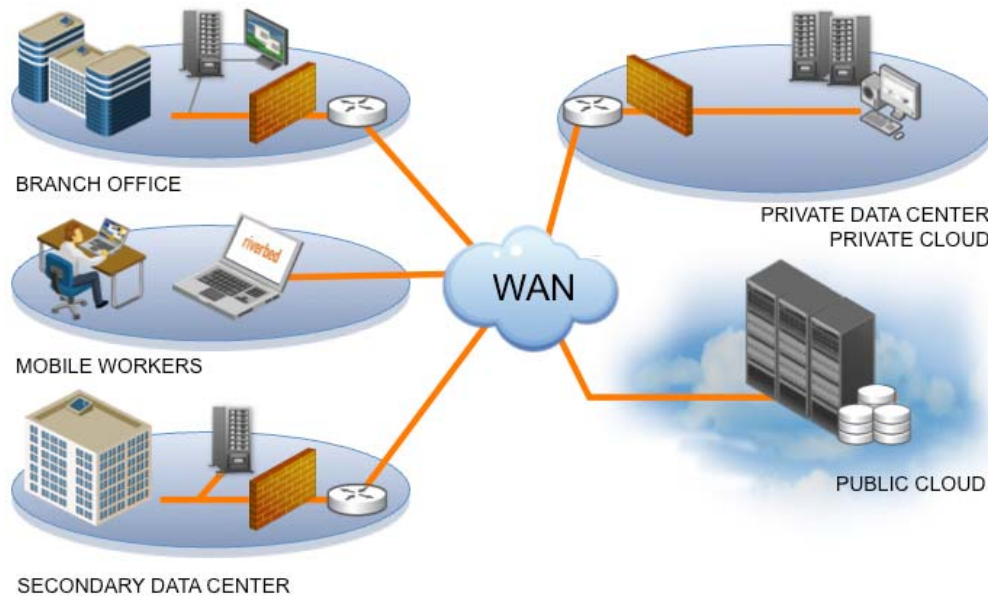
Regardless of whether an organization is considering a private, public, or hybrid cloud deployment the consequences of performance and network limitations remains the same. The central role that the network plays in determining the success or failure of a cloud project has brought the need for WAN optimization to the forefront to address these concerns.

With a private cloud, where consolidation, virtualization, and automation simplify an enterprise's existing data center, there is no leap for many enterprises to see how performance could be a problem. They already know their environment and its limitations;

¹ *Above the Clouds: A Berkeley View of Cloud Computing* UC Berkeley Reliable Adaptive Distributed Laboratory 10 February 2009

they know how much bandwidth they use, for which applications, and how much it costs.

Figure 2. Hybrid – Public & Private Cloud Deployment Diagram



Public cloud environments are not as well known a quantity; and naturally it's difficult to infer the impact of moving particular applications to the public cloud without actually testing. One of the biggest differences for public cloud computing is an unintended consequence. When services are moved to a public cloud, even headquarters employees become remote workers. While often HQ houses a large set of employees in addition to applications, moving those applications to a cloud provider will subject HQ employees to the same challenges that branch office employees have faced for years. If the public cloud is located nearby that might not make a noticeable difference; if the organization leverages a public cloud provider that is either far away or does not give any control over where an application is located, then headquarters-based users may have significant issues accessing those applications due to latency.

Strategies for unleashing cloud performance

IT managers are stuck in the middle of a tug-of-war. On one hand, distributed employees need full, fast application access. But on the other, management wants to consolidate IT infrastructure and make full use of any real cost advantages that a cloud provider may offer. How can an organization balance these needs, particularly when the challenge spans almost all applications that organizations use, from common end-user applications such as MS Office down to less visible administrative applications such as backup and replication?

In order to overcome these seemingly divergent needs organizations should consider these three strategies for unleashing cloud performance within their business:

- **Understand your own environment.** It's not enough to know what applications the business would like to move to the cloud. IT managers need a deep understanding of how applications are performance across the WAN today and what users are most dependent on particular applications. IT managers need to proactively aggregate information based on geography, application, and even individual user. They also have to be ready to more quickly assess, discover, and eliminate network-related problems in order to support consolidated cloud environments. Ideally managers should be equipped to aggregate this information without requiring more distributed hardware that goes against the grain of the consolidation initiatives.
- **Optimize what you already have, and expect the same performance from a cloud provider.** Tens of thousands of organizations already use WAN optimization either across their organization or in key locations in order to accelerate end-user computing and collaboration, disaster recovery operations, and cut bandwidth needs. Organizations now need to leverage WAN optimization across the board to prepare all business locations for a more distributed world. At the

same time, as your business considers both cloud services providers and WAN optimization solutions, make sure that the two map to each other. Your WAN optimization provider should have a form-factor (usually a virtual appliance) that will easily slot into a public cloud computing environment or a private cloud implementation, and your cloud service provider should be one who embraces the fact that performance-enhancing products like WAN optimization are necessary to make their cloud worthy of production use in enterprises.

- **Consolidate to the core and at the edge.** Ensure that you have a good plan to discover all of the applications and servers in your environment and which ones can effectively be consolidated today. Such a plan will also allow you to quickly map the applications or services that could potentially be moved to a public cloud, as well as which services must remain distributed. For services that must remain distributed, however, look to combine them using branch-office-in-a-box technologies. These technologies extend many cloud benefits such as simplified management and virtualization all the way to the edge of your network. That enables you to drive cost efficiency from the core to the edge of your IT operations.

Conclusion

While cloud services aim to eliminate cost and complexity from the world of enterprise IT, the unintended consequences of these services may do exactly the opposite if not carefully planned for. Riverbed® is squarely focused on the problem of making cloud services deliver local-like performance to distributed end users and IT managers. Whether enterprises are looking to deliver private cloud services or public cloud services Riverbed provides organizations with the IT performance tools to help ensure that the cost and management promises of cloud services are complemented with the performance needed to unleash the potential of the cloud.

About Riverbed

Riverbed Technology is the IT infrastructure performance company. The Riverbed family of wide area network (WAN) optimization solutions liberates businesses from common IT constraints by increasing application performance, enabling consolidation, and providing enterprise-wide network and application visibility – all while eliminating the need to increase bandwidth, storage or servers. Thousands of companies with distributed operations use Riverbed to make their IT infrastructure faster, less expensive and more responsive. Additional information about Riverbed (NASDAQ: RVBD) is available at www.riverbed.com



Riverbed Technology, Inc.
199 Fremont Street
San Francisco, CA 94105
Tel: (415) 247-8800
www.riverbed.com

Riverbed Technology Ltd.
Farley Hall, London Road, Level 2
Binfield, Bracknell
Berks
RG42 4EU
Tel: +44 1344 401900

Riverbed Technology Pte. Ltd.
391A Orchard Road #22-06/10
Ngee Ann City Tower A
Singapore 238873
Tel: +65 6508-7400

Riverbed Technology K.K.
Shiba-Koen Plaza Building 9F
3-6-9, Shiba, Minato-ku
Tokyo, Japan 105-0014
Tel: +81 3 5419 1990