Bare Metal Cloud: A Non-Virtualized Cloud Option for Performance-Sensitive Workloads

Stratecast | FROST & SULLIVAN

Stratecast Perspectives & Insight for Executives (SPIE)
Volume 14, Number 2

Lynda Stadtmueller
Program Director, Cloud Computing

January 17, 2014

INTRODUCTION[1]

Does a cloud configuration require virtualization? It turns out, the answer is "no."

In fact, the National Institute of Standards and Technology (NIST), whose cloud definition is widely accepted in the industry, omits virtualization as a criteria for cloud.[2] NIST's "essential characteristics" include on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service—but not virtualization.

This may surprise many in the IT community who have always assumed that a virtualized server infrastructure was necessary to provide the flexibility and scalability associated with cloud.[3] However, the emergence of "bare metal" clouds—that is, clouds that do not utilize virtualization—is forcing a re-examination of what it takes to offer a cloud service. The bare metal options provide the flexibility and scalability associated with virtualized offers, while promising higher levels of performance and consistency.

> The emergence of "bare metal clouds"—that is, clouds that do not utilize virtualization—is forcing a reexamination of what it takes to offer a cloud service.

Currently, two cloud leaders—SoftLayer (an IBM company) and Internap—have developed a bare metal option as part of their cloud portfolios. Both tout their bare metal services as a way to differentiate themselves from the crowded cloud service market. Both have also had success in attracting cloud-skeptical businesses and performance-sensitive workloads that previously may not have been considered ideal for cloud deployment.

In this report, Stratecast examines the concept of the bare metal cloud from the provider and the customer perspective. We compare benefits and challenges of bare metal cloud configurations with the more common virtualized cloud configurations. Finally, we look at the bare metal cloud offer from SoftLayer, an IBM company.

VIRTUALIZATION – THE VALUE AND THE COST

Server virtualization is well established in enterprise data centers and in hosting and cloud centers. More than half of businesses utilize server virtualization, according to the 2013 Stratecast | Frost & Sullivan Cloud User Survey.

Virtualization separates the logical from the physical components of the workload. Application code and associated operating system are packaged neatly into a virtual machine (VM). Multiple VMs, regardless of operating system, can share a physical server; a hypervisor installed on the server allocates resources and acts as a translator, making each VM believe it has full access to the server resources.

[1] In preparing this report, Stratecast conducted interviews with representatives of the following companies:

- IBM – Jim Comfort, VP, Cloud Computing; Tim Kounadis, Director, Cloud Services Marketing

Please note that the insights and opinions expressed in this assessment are those of Stratecast and have been developed through the Stratecast research and analysis process. These expressed insights and opinions do not necessarily reflect the views of the company executives interviewed.

[2] National Institute of Standards and Technology, *The NIST Definition of Cloud Computing*, Special Publication 800-145. September 2011. http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[3] Stratecast admits to including "virtualization" in its cloud definition for many years. The definition has since been amended.

© 2014 Stratecast. All Rights Reserved.

The virtualized workload is self-contained and highly portable. Like a turtle or a motor home, it carries all it needs on its back—operating system and application code—and isn't fussy about where it sets up housekeeping. Thus, IT technicians do not have to custom-configure a server exoskeleton for a virtualized workload.

Like a turtle or a motor home, a virtualized workload carries all it needs on its back — operating system and application code—and isn't fussy about where it sets up housekeeping.

As such, virtualization is associated with infrastructure conservation and flexibility. Top benefits of virtualization include:

- **Deferral of capital expenses**: By accommodating multiple virtualized workloads per physical server, virtualization optimizes server utilization, and reduces the need for additional servers or expanded floorspace.
- **Faster time to deploy workloads**: In a virtualized environment, VMs can be tested, deployed, spun down, and moved via a management console, without requiring on-site technicians to perform labor-intensive tasks to configure the servers. This rapid deployment reduces operating costs and decreases time to provision servers.
- **Support for high availability environments**: In a virtualized server environment, routine hardware maintenance or unexpected interruptions do not need to shut down applications. Because VMs are portable, they can be moved to another server, in house or outside, that has spare capacity.

The resulting conclusion from these generalized benefits is that virtualization technologies offer the greatest benefit to the infrastructure owner. By optimizing hardware utilization, deferring costs, and allowing for flexibility, virtualization allows infrastructure to be managed more efficiently, easily and cost-effectively.

But in a cloud environment, infrastructure responsibility falls to the cloud service provider, so those benefits of virtualization do not automatically accrue to the customer. The enterprise customer can benefit indirectly from vitualization if the provider chooses to pass on cost savings in the form of lower rates, for example. Nonetheless, in comparing the end-user experience or application performance, a virtualized workload offers no advantages over a non-virtualized workload.

In a cloud environment, infrastructure responsibility falls to the cloud service provider, so the benefits of virtualization do not automatically accrue to the customer.

© 2014 Stratecast. All Rights Reserved.

In fact, virtualization comes at a cost to the user. For some workloads, virtualization can offer infrastructure efficiency for the cloud service provider, at the cost of diminished performance for the customer. Primary sources of concern are "noisy neighbor syndrome" and the "hypervisor tax."

Noisy Neighbor

As noted, virtualization is an excellent way to optimize use of server capacity. By loading multiple virtualized workloads on a shared physical server, overall resource utilization improves. However, the different applications are all contending for the same processor and memory resources, which inevitably brings the risk that the computing resource will not be available at the capacity level and at the instant it is needed. For many apps, the risk may be minimal—for example, if an internal intranet page loads slowly occasionally, employees will not go elsewhere. Also the performance impact is likely to be sporadic and unpredictable, occurring only on occasions when multiple apps attempt to access the shared resources simultaneously. However, for latency-sensitive applications such as e-commerce, gaming, and streaming media, any delay can be intolerable.

In a private data center, the enterprise can control the risks of resource contention by making decisions regarding assignment of VMs across available physical servers, monitoring and balancing loads as needed. However, that level of control is not possible for customers of a shared cloud, as only the provider has visibility across the entire, multi-tenant environment. In a shared cloud environment, customers have little control over where their VMs are loaded and which other customers' workloads are sharing the processor. Furthermore, like an airline overbooking flights to ensure full planes, the cloud service provider has an incentive to "oversubscribe" each physical server. The greater the resource utilization, more customers are served at a lower cost per customer.

Like an airline overbooking flights to ensure full planes, the cloud service provider has an incentive to "oversubscribe" each physical server.

For customers eager to avoid the "noisy neighbor" risk, many providers offer a hosted private cloud or virtualized private cloud option. In these services, the server hardware and, perhaps, other infrastructure components are dedicated to a single enterprise. Thus, the virtualized workloads that share physical server resources all belong to the same enterprise, giving the enterprise some control over capacity utilization.

Hypervisor Tax

Even if there are no strangers sharing the facility—for example, in a dedicated or private cloud environment—virtualization extracts a toll on available capacity. The "hypervisor tax" is the amount of processing capacity that is consumed by the hypervisor layer. While virtualization providers have enhanced their hypervisor software to be as thin as possible, a hypervisor can still consume as much as 10 percent of the available capacity of a server. For high-performance workloads that require large amounts of capacity, the tax can be significant, even impacting performance of the application.

© 2014 Stratecast. All Rights Reserved.

> A hypervisor can consume as much as 10 percent of the available capacity of a server.

In addition, as with every additional software layer, the hypervisor layer subjects data to delay; minuscule amounts, to be sure, but noticeable for latency-sensitive workloads.

Thus, enterprises are faced with trade-offs in running their high-capacity or high-performance workloads in the cloud; that is, trade optimal performance for the efficiency and low cost structure of the virtualized cloud, or trade efficiency and low cost for high performance in a dedicated hosting environment.

But suppose enterprises had the choice of a low-cost, scalable, easily managed hosting option without virtualization? This is the operating principle behind the bare metal cloud.

BARE METAL CLOUD

The concept of a bare metal cloud is not new; Stratecast first wrote about it back in 2009. But most leading cloud providers do not offer a bare metal option. However, interest in and adoption of bare metal is growing as more enterprises turn to the cloud as a scalable, cost-efficient way to run high-performance workloads.

What is a Bare Metal Cloud?

A bare metal cloud combines characteristics of traditional hosting and virtualized cloud Infrastructure as a Service (IaaS). As in traditional hosting, a bare metal cloud server is dedicated to a specific tenant, not shared among tenants. As such, the basic "unit" is an actual server: customers choose the type of server configuration they want from a menu offered by the provider; specifying, for example, a single, dual, or quad core processor, with associated memory and storage.

> A bare metal cloud combines characteristics of traditional hosting and virtualized cloud Infrastructure as a Service (IaaS).

Unlike traditional hosting, the bare metal cloud offers most of the key characteristics that define any cloud service. With variations based on the provider's terms, a bare metal cloud offers:

- **On-demand capacity** – Servers are deployed in a few hours.
- **Usage-sensitive pricing** – Billing is based on the amount of capacity committed or used; with minimal time increments ranging from one hour to one month.
- **Self-servicing** – Enterprise customers deploy workloads via a portal.

© 2014 Stratecast. All Rights Reserved.

- **Network access** – Enterprise users access their applications via the public Internet or private networks.

As with a virtualized cloud, a bare metal cloud relies on a sophisticated management and enablement platform. The platform, which is proprietary to each service provider, provides:

- Tools that automate infrastructure provisioning tasks, enabling rapid deployment
- A self-service portal designed for end users to spin workloads up and down
- A usage meter
- Performance monitoring
- Reporting

By utilizing a common management platform for their bare metal option and virtualized cloud offers, providers extend the value and ease-of-use of their services.

How it Works

Absent from the bare metal cloud is the hypervisor layer. In a bare metal cloud, the workloads are deployed onto servers that are pre-configured with the appropriate operating system.

In configuring a bare metal cloud, providers pre-deploy banks of servers with specific operating systems. The bare metal servers are managed via the management platform, which provides automated tasks for deployment, self-servicing portals, capacity and usage monitoring, and reporting. In the virtualized cloud configuration, a smaller percentage of total server resources (processing, memory, storage) are available for individual applications or workloads than in a bare metal configuration. In fact, on a per-server basis, the bare metal cloud looks a lot like a traditional (non-virtualized) server configuration.

Pricing

Bare metal cloud servers generally command a higher per-unit price than virtualized cloud servers. This is because they cost more to provide, as well as offer additional value. Usage-based rates are calculated on either an hourly or a monthly basis. For example, IBM offers a 2-core bare metal server with 2 GB RAM for $0.50 per hour or $159 per month, whereas a 2-core virtualized server with 2 GB RAM is just $0.20 per hour or $119 per month.

> Despite the higher per-unit prices, bare metal servers can be more cost-efficient overall than virtualized servers for high-performance workloads.

While cost is unlikely to be a major factor in the decision to purchase bare metal cloud rather than virtualized cloud, providers report that their customers believe that the value they receive is worth the price premium over a virtualized server. Bare metal servers can be more cost-efficient overall than virtualized servers for high-

© 2014 Stratecast. All Rights Reserved.

performance workloads (e.g., workloads with high Input/Output Processing requirements). Compared with virtualized servers, fewer bare metal cloud units are required to provide equivalent or superior processing power; thus, the overall performance can more than compensate for the higher per-unit price. In some cases, the higher available capacity in a bare metal server provides greater data throughput; which means that fewer cloud "units" may be required to meet the customer's performance parameters.

Workloads

Workloads that benefit most from a bare metal cloud configuration include those that have the following characteristics:

- Input/Output (I/O) intense, such as large NoSQL databases
- High performance computing, such as Big Data analytics
- Latency-sensitive, such as gaming or streaming media.

CHALLENGES TO PROVIDING A BARE METAL CLOUD OPTION

While the number of providers offering virtualized cloud services grows each year, only a handful of providers offer a bare metal version of their cloud services. For providers, the decision to offer bare metal involves some financial risk. In the shaky economics of the public cloud, with market leader Amazon Web Services (AWS) continually driving down prices, most providers have yet to reliably recover the costs of their data center investments. Raising their cost structure by offering a bare metal option could further reduce the return on investment (ROI) for their cloud services.

This lower ROI is because a cloud infrastructure based on bare metal is less flexible than a virtualized infrastructure. Correspondingly, providers must invest in more infrastructure and labor in their bare metal clouds, as follows:

- **Greater capital investment** – To a much greater extent than for virtualized cloud centers, the provider needs to anticipate how much capacity customers will require for each type of operating system, and have appropriate servers ready to go. This puts the provider in the same position that its enterprise customers are looking to avoid: the need to overbuild infrastructure capacity to meet peak demand. In a virtualized environment, the hypervisor makes all servers equal; any workload (regardless of operating system) can go on any server (regardless of operating system).
- **Greater operating costs and effort** – In a virtualized environment, the pool of available capacity can be allocated via software alone. The customer spins up a VM via the user portal; and, as long as any server has spare capacity (meaning that it can reasonably accommodate another VM), servers can be provisioned without physical intervention from the provider's team of on-site infrastructure technicians. In a bare metal cloud, capacity is not pooled. Instead, physical servers are physically configured as demand dictates. This means that technicians will be involved in spinning servers up and down (for example, swapping out the operating system), as customer demand changes.

Most cloud providers, including market leader AWS, have not introduced a bare metal cloud offer. In defending its position, AWS has publicly stated its focus is on performance, rather than infrastructure configuration. A valid point: if customers get what they need, should they care what the underlying infrastructure configuration is? But

© 2014 Stratecast. All Rights Reserved.

the counter-argument is that a responsible IT department needs to understand how its suppliers and partners provide services; this enables the enterprise to evaluate marketing claims, predict performance, make smart purchase decisions, and manage vendors.

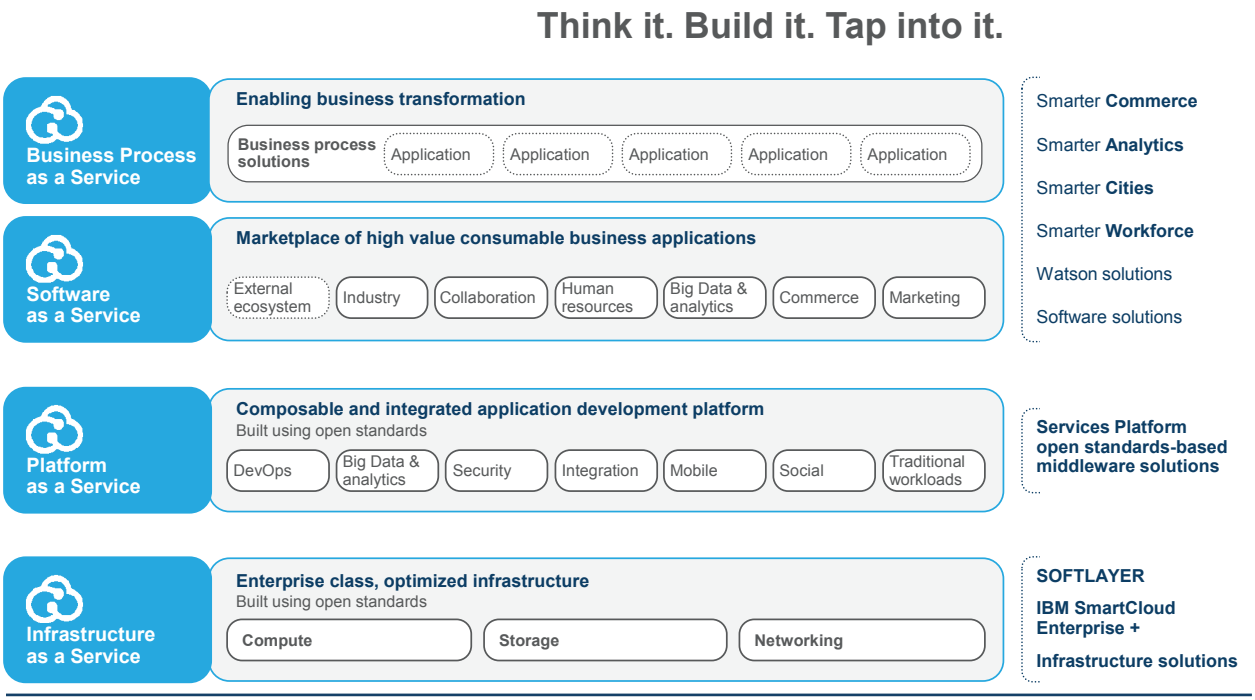
BARE METAL CLOUD SERVICES

SoftLayer, an IBM company, offers a full portfolio of cloud services that include public, private, managed, and bare metal options. SoftLayer allows its customers to select the best option per workload, using a single management portal to deploy, manage, scale, and configure all their cloud workloads.

SoftLayer Bare Metal Servers

IBM boasts the industry's most comprehensive cloud strategy, encompassing a broad range of enterprise-grade hardware, software, and services. Fitting of this strategy, IBM offers customers the option of a bare metal cloud as part of the IBM Cloud portfolio. As shown in Figure 1, IBM cloud services extend up the entire cloud stack, encompassing Infrastructure as a Service, Platform as a Service, Software as a Service, and into a category it calls Business Process as a Service.

Figure 1: IBM Cloud Services



Source: IBM

The foundation for the IBM Cloud family is the newly acquired SoftLayer infrastructure platform. The SoftLayer platform supports a flexible, high-performance enterprise-grade environment in any configuration: traditional IT or cloud, dedicated or shared servers, virtualized or bare metal. The platform flexibility allows businesses to build, configure, change, and tweak their cloud solution mixtures to provide optimal price-performance for every workload. A robust infrastructure management system provides high levels of automation, simplifying complex processes such as auto-scaling. A "triple network" architecture—comprising

© 2014 Stratecast. All Rights Reserved.

public transport networks, a private, inter-data center network, and a separate management network—minimizes capacity issues and contention.

SoftLayer's CloudLayer services include public cloud (multi-tenant, virtualized servers); private cloud (single-tenant, virtualized servers); and bare metal (single tenant, physical server). All are available on-demand, via a Web interface, at hourly or monthly rates. Users have a choice of four basic bare metal server options, ranging from a 2-core server with 2 GB RAM/250 GB storage for $0.50 per hour, to a 16-core server with 16 GB RAM/250 GB storage for $1.25 per hour. Additional RAM and storage are also available.

While all SoftLayer cloud services, including bare metal, can be purchased "a la carte," IBM's go-to-market strategy is based on solution selling. As such, the company often includes SoftLayer bare metal servers as part of a comprehensive cloud solution for customer situations such as:

- Accommodate peak usage for a high performance computing workload that is based in on the customer's premises
- Scale a rapidly growing NoSQL database
- Ensure consistent performance for an I/O intense workload

THE LAST WORD

As Stratecast is fond of pointing out, the cloud is primarily a provider model—an efficient way to deliver IT resources to users. But the infrastructure configuration that works best for the provider may not be ideal for the customer, especially when it comes to high-performance workloads. That's where the bare metal cloud can make a difference.

SoftLayer (now an IBM company) is one of few leading cloud service providers to offer a bare metal option as part of its cloud portfolio. Because bare metal cloud is so poorly understood, the companies is taking on responsibility for educating the market. The company touts its newly-acquired SoftLayer platform for its flexibility, with bare metal as one of the supported options. Key to the market positioning is that bare metal is one more deployment option in an enterprise hybrid cloud strategy; an option that is likely to be more effective for certain high-performance workloads.

For cloud providers who are willing to take on the greater financial responsibility, bare metal can be a valuable addition to the cloud portfolio. Enterprises continue to be squeamish about entrusting their critical and high-performance workloads to a hosted cloud environment. In fact, in the 2013 Stratecast|Frost & Sullivan Cloud User Survey of IT decision-makers, 73 percent cite concerns over "poor or inconsistent application performance" as a key reason for choosing not to place certain workloads in the cloud. For these reluctant decision-makers, a bare metal cloud may offer the right combination of price, performance and control to convince them to begin moving critical workloads out of the on-premises data center. For the provider, the bare metal cloud option can prove valuable in retaining and growing customers: enterprises frequently start with a single bare metal workload, and later consolidate all their virtualized cloud workloads with the provider.

Lynda Stadtmueller

Program Director – Cloud Computing
Stratecast | Frost & Sullivan
lstadtmueller@stratecast.com

© 2014 Stratecast. All Rights Reserved.

Silicon Valley
331 E. Evelyn Ave., Suite 100
Mountain View, CA 94041
Tel 650.475.4500
Fax 650.475.1570

San Antonio
7550 West Interstate 10, Suite 400
San Antonio, Texas 78229-5616
Tel 210.348.1000
Fax 210.348.1003

London
4, Grosvenor Gardens,
London SWIW ODH,UK
Tel 44(0)20 7730 3438
Fax 44(0)20 7730 3343

877.GoFrost • myfrost@frost.com
http://www.frost.com

ABOUT STRATECAST

Stratecast collaborates with our clients to reach smart business decisions in the rapidly evolving and hyper-competitive Information and Communications Technology markets. Leveraging a mix of action-oriented subscription research and customized consulting engagements, Stratecast delivers knowledge and perspective that is only attainable through years of real-world experience in an industry where customers are collaborators; today's partners are tomorrow's competitors; and agility and innovation are essential elements for success. Contact your Stratecast Account Executive to engage our experience to assist you in attaining your growth objectives.

ABOUT FROST & SULLIVAN

Frost & Sullivan, the Growth Partnership Company, works in collaboration with clients to leverage visionary innovation that addresses the global challenges and related growth opportunities that will make or break today's market participants. For more than 50 years, we have been developing growth strategies for the Global 1000, emerging businesses, the public sector and the investment community. Is your organization prepared for the next profound wave of industry convergence, disruptive technologies, increasing competitive intensity, Mega Trends, breakthrough best practices, changing customer dynamics and emerging economies? Contact Us: Start the Discussion

For information regarding permission, write:
Frost & Sullivan
331 E. Evelyn Ave. Suite 100
Mountain View, CA 94041

GTW03042-USEN-00