OPEN SOURCE & CLOUD COMPUTING: ON-DEMAND, INNOVATIVE IT ON A MASSIVE SCALE

White Paper
June 2009

Abstract

While more and more start-ups are taking advantage of the cost-savings and agility that cloud computing provides, enterprises are now exploring how they could leverage an external cloud, build an internal cloud or even become a cloud provider. Because open source is fundamental to cloud computing, enterprises are also seeing that they can avoid the old problem of getting locked in to a single vendor. Sun, a globally recognized leader in open source, is committed to bringing the benefits of open standards and open source to cloud computing.

Table of Contents

Executive Summary

Cloud computing offers enormous opportunities for enterprises to cut computing costs and speed time to market for new Web services. Whether they leverage an external cloud, build their own cloud internally, become a cloud services provider, or choose a blend of these options, enterprises can gain advantages that enable them to become much more agile.

The success of massive-scale systems like Google, eBay, and Amazon led to the rise of cloud computing — and each of these companies built their highly customized systems on a huge scale with freely available, freely modifiable open-source software. Today, the dominant software stacks used in cloud computing environments are also open source, and the integrated, optimized, open-source Apache, MySQL, PHP/Perl/Python (AMP) stack is the preferred platform for building and deploying new Web applications and services.

As a worldwide leader in open source, Sun is committed to bringing the benefits of open standards and open source to cloud computing. Powered by Java™, MySQL™, OpenSolaris™, Open Storage, and other open-source technologies, the Sun Open Cloud Platform is designed to enable enterprises to build both public and private clouds.

Making the Case for Cloud Computing

In many ways, cloud computing is an example of the increasing movement of compute and data resources onto the Web. But there's a difference: cloud computing represents a new tipping point for the value of network computing. It delivers higher efficiency, massive scalability, and faster, easier software development. Cloud computing also brings new levels of efficiency and economy to the delivery of IT resources on demand — and in the process, it opens up new business models and market opportunities for developers and enterprises.

What makes cloud computing so interesting today is that businesses are looking at Amazon, Google, and Yahoo! and wondering why these Web companies are able to produce such phenomenal cost savings over what an individual enterprise can achieve, given the market pricing for basic compute and storage. The fact is that cloud computing costs less — and it allows people to become much more agile in how they're using IT resources.

- **Cost savings.** Cloud computing cost savings are a result of increased efficiencies, which have always been a feature of Web-scale computing. By leveraging technologies like virtualization, companies are able to achieve a much more efficient use of computing resources, improving infrastructure utilization rates and streamlining resource management while providing on-demand scalability. When combined with utility pricing, this removes the need to overprovision in anticipation of future demand peaks. And instead of funding very large capital investments, businesses are able to pay only for the computing resources they actually need and consume, which allows them to convert capital expenses to operating expenses.
- **Business agility.** The self-provisioning features of cloud computing mean much faster time to market. There's no need to negotiate long-term contracts. Built-in services means someone else has already figured out how to build a scalable storage system, so companies don't have to do that work themselves. Cloud computing also enables greater elasticity; if an application grows to be very large, cloud customers only need to pay for increased resources when they're being used. This allows for a more flexible business model behind applications being deployed in the cloud today.

Cloud computing offers enterprises a way to cut costs and increase agility without having to rework or grow internal infrastructures that weren't designed to support Web-based services. It enables IT organizations to increase hardware utilization rates dramatically and scale up to massive capacities in an instant — without having to invest in new infrastructure, train new personnel, or license new software. It also creates opportunities to build a better breed of network services in less time, for less money.

As enterprises see how much higher their IT infrastructure costs are than those of startups using clouds — and how much more quickly startups are able to deliver new services — they will look for more ways to leverage cloud computing. Because cloud computing does not involve long-term contracts (and many cloud vendors charge customers in increments of as little as an hour), clouds are an excellent way for enterprises to quickly prototype new services, conduct testing and development, or run limited-time campaigns.

For example, in a traditional IT environment, developers create applications individually or with a team, then hand them over to an operations team or hosting provider to stage and test — which means less control and more friction as more people become involved. With a cloud, developers can commission servers at a low price via the Internet and have their application up and running quickly. Clouds can enable developers to do a lot more experimentation quickly and scale it up. And faster development and testing cycles mean businesses can accomplish in hours what used to take days, weeks, or even months.

Taking Advantage of Cloud Computing

So how does an enterprise take advantage of the cloud computing trend? It's not just about loading machine images of the business' entire software stack onto a public cloud; there are several different ways to exploit this infrastructure and explore the ecosystem of new business models.

Leverage the cloud

Typically, enterprises are using public clouds for specific functions or workloads. The public cloud is an attractive alternative for:

- **Development and testing**. This is perhaps the easiest cloud use case for enterprises (not just startup developers). There's no need to purchase servers when it's not yet clear if a project will pass the proof of concept.
- **Functional offloading**. Enterprises can use the cloud for specific workloads. For example, SmugMug does its image thumbnailing as a batch job in the cloud.
- **Augmentation**. Clouds give businesses a new option for handling peak load or anticipated spikes in demand for services. This is a very attractive option for enterprises, but also potentially one of the most difficult use cases. Success is dependent on the statefulness of the application and interdependence with other data sets that may need to be replicated and load-balanced across multiple sites.
- **Experimentation**. Why download demos of new software and then install, license, and test them? In the future, software evaluation can be performed in the cloud before licenses or support are purchased.

Build the cloud

Many large enterprises understand the economic benefits of cloud computing but want to ensure strict enforcement of security policies. Some might experiment first with "private" clouds, with a longer-term option of migrating mature enterprise applications to a cloud that's able to deliver the right service levels. Others may simply want to build private clouds to take advantage of the economics of resource pools and to standardize their development and deployment processes.

Some enterprises will transition to cloud computing by working with cloud providers to develop an architecture for a private cloud housed inside the corporate firewall. However, moving data from protected areas inside the firewall to public, multitenant datacenters can be problematic for some enterprises due to regulatory requirements. This may be mitigated when an enterprise utilizes an external cloud that runs an environment similar to its internal datacenter, enabling it to leverage the external cloud when demand spikes while also protecting its data and staying in compliance.

Be the cloud

As enterprises and service providers gain experience with the cloud architecture model and become more confident in the security and access-control technologies that are available, many will decide to deploy externally facing cloud services. The phenomenal growth rates of some of the public cloud offerings available today will no doubt accelerate this momentum.

Cloud service providers can:

- Provide new routes to market for startups and Web 2.0 application developers
- Offer new value-added capabilities such as analytics
- Derive a competitive edge through enterprise-level SLAs
- Help enterprise customers develop their own clouds

An enterprise may choose to use a service provider's cloud or build its own cloud, which is a good option for companies dealing with data protection and service-level issues. A third possibility is to develop a hybrid model where the enterprise owns parts of a cloud and shares other parts, though in a controlled way. Hybrid clouds offer the promise of on-demand, externally provisioned scale but add the complexity of determining how to distribute applications across these different environments. While enterprises may be attracted to the promise of a hybrid cloud, this option will likely see earliest adoption for stateless applications that require no complex databases or data synchronization.

Any enterprise that is building large datacenters today should be thinking about whether they will offer cloud services internally (private cloud) and to external organizations (public cloud).

Open Source in the Cloud

Open source is the great enabler of cloud computing. From Google and Yahoo! to Amazon and eBay, the precursors of cloud computing utilized the freely available, freely modifiable nature of open source to build highly customized systems on a never-before-seen scale to power their Web-based applications. It is the success of these massive-scale systems that has led to the rise of cloud computing — which is a generalization of the same techniques and technologies used by Google and others to enable developers the world over to tap into a model of computing that would otherwise not be affordable (or even available).

To the early purveyors of the massive-scale systems that inspired cloud computing, it was the free availability and modifiability of open source that made it appealing. But with enterprises, the attraction of open source is that it is a way to avoid the continuing problem of single-vendor lock-in.

Open-source technologies tend to attract large and vibrant communities and ecosystems around them, with one result being a variety of products and services tailored for enterprise use. So if an enterprise is not happy with the service or support it is receiving from one vendor, it can turn to a different vendor for that service and support — and if all else fails, it has ready access to the source code and the communities that created and maintain it.

Because open source is fundamental to cloud computing, it is not surprising that the dominant software stacks used in cloud environments are also open source. Today, the integrated, optimized, open-source Apache, MySQL, PHP/Perl/Python (AMP) stack is the preferred platform for building and deploying new Web applications and services. What's more, cloud computing is proving to be the catalyst for the adoption of an even newer stack of more lightweight, agile tools such as lighttpd, an open-source Web server; Hadoop, the free Java software framework that supports data-intensive distributed applications; and MogileFS, a file system that enables horizontal scaling of storage.

Despite open source's central place in cloud computing, its benefits have yet to be fully passed on to early adopters. As Tim O'Reilly, CEO of O'Reilly Media, and others have pointed out, open source is predicated on software licenses, which in turn are predicated on software distribution — and in cloud computing, software is not distributed; it's delivered as a service over the Web. So cloud computing infrastructures — and the modifications to the open-source technologies that enable them — tend not to be available outside the cloud vendors' datacenters, potentially locking their users in to a specific infrastructure.

Although the software stacks that run on top of these cloud computing infrastructures are predominantly open source, the APIs used to control them (such as those that enable applications to provision new server instances) are not entirely open, further limiting developer choice. And cloud computing platforms that offer developers higher-level abstractions such as identity, databases, and messaging, as well as automatic scaling capabilities (often referred to as "platform as a service"), are the most likely to lock their customers in.

Ideally, users of cloud computing would be able to move their applications among a variety of standardized providers that offer open interfaces to common services. Developers and startups would be able to target public clouds, allowing them to focus their scarce resources on the applications and services they are developing rather than the infrastructures that power them — without limiting their ability to later migrate to their own hosted infrastructure. Enterprises utilizing clouds for experimentation at the departmental or workgroup level would be able to easily migrate their prototypes to a private cloud hosted behind the company firewall or build their private clouds to expected capacity and offload excess capacity to public clouds. Without open interfaces linking the variety of clouds that will exist — public, private, and hybrid — these use cases will be difficult or impossible to deliver.

Cloud Computing with Sun

For more than 25 years, Sun has been driving open standards. Today, Sun is a globally recognized leader in open source, and we're committed to bringing the benefits of open standards and open source to cloud computing.

The cornerstone of Sun's cloud computing offering is called the Sun Open Cloud Platform, an open cloud computing infrastructure powered by open-source technologies including Java, MySQL, OpenSolaris, and Open Storage. The Sun Open Cloud Platform is designed to run both public and private clouds, enabling enterprises to build an internal cloud and even mirror an external cloud's infrastructure with components that fully optimize the features of each system.

Sun also is building a public cloud based on the Sun Open Cloud Platform called the Sun Cloud, targeted at developers, startups, students, and enterprise workgroups. The Sun Cloud will enable these groups to tap into the economic and time-to-market advantages of cloud computing. Because it is built on the Sun Open Cloud Platform, applications and services written for the Sun Cloud will be able to easily be migrated to other compatible public clouds or to private clouds within an enterprise's own datacenters.

Initially, the Sun Cloud will provide two services — storage and compute. Each service will fully leverage existing standards and industry best practices — and in cases where those standards and best practices have yet to emerge, Sun will lead the way toward establishing them. To that end, the Sun Cloud APIs are available under a Creative Commons License, and Sun is working with the community and partners to refine the APIs to best meet developers' needs. And as Sun rolls out new services, the company will be utilizing its deep expertise in creating ubiquitous industry standards such as NFS and the Java platform to bring the benefits of open standards and open source to the cloud computing marketplace.

The Sun Cloud Storage Service will bring Sun's breakthrough Solaris ZFS™ technology to the cloud and will provide both file- and object-based interfaces to data storage. For file-based storage, it will implement the industry-standard Web-based Distributed Authoring and Versioning (WebDAV) protocol, enabling existing operating systems and software libraries to easily connect to the service. For object-based storage, the Sun Cloud Storage Service will provide an interface compatible with Amazon's Simple Storage Service (S3) API, making it easy for developers familiar with Amazon semantics to adopt the service and to migrate existing applications easily. Because the Sun Cloud Storage Service will be based on the open-source Solaris ZFS technology, it also will deliver the unparalleled performance and scalability, volume management, and snapshot and rollback capabilities that ZFS provides.

The Sun Cloud Compute Service is taking a new approach to cloud application provisioning and management. In a departure from the prevalent model of individual virtual machine images, which requires developers and operators to manually set up the often-complex interconnections between virtualized resources, the Sun Cloud Compute Service will implement the concept of a virtual datacenter (VDC). A VDC includes all the components used to build and deploy a multitier cloud application, including virtualized servers (supporting Linux and Windows as well as the OpenSolaris OS), storage, networking, load balancers, firewalls, and so on — in one complete package. A VDC can be provisioned via an easy-to-use, drag-and-drop GUI as well as through a command-line interface or an API that enables programmatic access to the VDC, useful for scaling the application up (or down) as required.

Underneath the cloud services of the Sun Cloud is the solid foundation of the Sun Open Cloud Platform — comprised of both Sun and community open-source technologies:

Open Operating System — Solaris, the most powerful and popular cloud operating system, is available on Sun, Intel, IBM, HP, and Dell systems.

- Delivers the performance, stability, and security enterprise users and customers demand
- Has more available applications than any other open OS

Open Virtualization — From NFS to Dynamic System Domains, chip multithreading (CMT), and Solaris Containers, Sun has the experience and expertise to take virtualization to a new level. Sun is one of the few companies able to address all types of cloud virtualization:

- Hypervisor (xVM Server)
- OS (Solaris Containers)
- Network (Crossbow)
- Storage (Solaris ZFS)
- Applications (GlassFish™, Java CAPS)

Open Storage — Sun is leading the open-storage movement, combining open-source software with industry-standard system components to reduce storage costs by up to 90%.

- Delivers breakthrough economics
- Provides massive capacities and extreme scalability without vendor lock-in
- Offers the flexibility to scale, reconfigure, or repurpose an enterprise's infrastructure

Open Database — Sun's acquisition of MySQL enables Sun, its ISV partners, and cloud service providers to create and deploy new data-fed, data-driven applications and services.

- MySQL: the *M* in the AMP software platform
- A foundation technology for Web applications
- Combines with OpenSolaris, GlassFish, Java, and NetBeans™ technologies to create a powerful cloud platform
- Can be used to build next-generation applications and new revenue streams — without licensing restrictions

Open Tooling — Application developers can develop for the cloud, incorporate cloud services into their applications, and benefit from cloud-based services that make the development process easier, faster, and cheaper.

- The NetBeans IDE offers rich support for developing standards-based applications for the cloud with the Java and JavaFX™ platforms (for front ends to the cloud). NetBeans technology also has excellent support for dynamic languages such as Ruby, PHP, and soon Python — all popular choices for cloud development.
- Project Kenai consists of cloud services for the open-source development process. Today, these include source code management (Mercurial, Git, Subversion), bug tracking (Bugzilla, JIRA), wikis, forums, mailing list support, and more. Soon, project Kenai will offer continuous integration (CI) in the cloud.
- Zembly is a cloud-based application development platform for creating, hosting, and scaling Web applications in the cloud. Zembly includes a Web-based IDE, an integrated hosting environment, and facilities for deploying applications into social networks such as Facebook.
- Project Speedway (previewing at CommunityOne West 2009) consists of a cloud-based solution that enables developers without local SPARC® workstations to build, test, and tune their applications for OpenSolaris software on SPARC technology. Project Speedway is tightly integrated with the NetBeans IDE and Sun Studio (for build/test/tune) as well as project Kenai (for accessing source repositories hosted there).

Looking Ahead

For those developers and enterprises that want to embrace cloud computing, Sun is developing critical technologies to deliver enterprise scale and systemic qualities to this new paradigm:

- **High-density horizontal computing** — Sun is pioneering high-power density compute node architectures and extreme-scale InfiniBand fabrics as part of our top-tier HPC deployments. This high-density technology is being incorporated into our large-scale cloud designs.
- **Data in the cloud** — More than just compute utilities, cloud computing is increasingly about petascale data. Sun's Open Storage products offer open-source software and powerful hybrid data servers with unprecedented efficiency and performance for the emerging data-intensive computing applications that will become a key part of the cloud.

These technologies are focused on driving more efficient large-scale cloud deployments that can provide the infrastructure for next-generation business opportunities: social networks, algorithmic trading, continuous risk analysis, and more.

Sun is working toward a vision of offering enterprises the ability to utilize a public cloud, build a private cloud, or use a hybrid approach by expanding research and development efforts in four key open-source areas:

- **Software.** Providing the open-standards-based tools that developers and architects need to build agile services that can be deployed in the cloud — from Sun's Web stack to software elements from other vendors
- **Systems.** Delivering compute, storage, and networking systems that interoperate with each other and integrate with systems from other vendors, whether they're based on AMD, Intel®, or SPARC architectures
- **Microelectronics.** Pushing the envelope for chip multithreading and multicore computing; moving to ever-higher compute densities within the cloud
- **Services.** Supporting development efforts through a broad range of professional services, network services, and value-added service offerings from partners (ISVs, OEMs, channel partners, and systems integrators)

Sun is uniquely positioned to bring together open-source knowledge, innovative products and technologies, expert services, and real-world experience helping to build some of the world's most efficient and scalable datacenters. As a systems company, Sun can supply hardware, software, storage, and services that are fully integrated and optimized for cloud applications. Sun products are integrated across the cloud computing stack and are designed to be integrated with standards-based technologies from other vendors.

Many of Sun's products and technologies are on-ramps to cloud computing, including virtually all of Sun's server and storage systems, the Solaris Operating System, the ZFS file system (included with Solaris), the Sun virtualization portfolio, and Sun Ray™ desktops. And Sun's experience building and running the Sun Cloud provides a real-world blueprint available for any enterprise that wants to build and run its own cloud.

Resources

"Is Cloud Computing Ready for the Enterprise?"
James Staten, Forrester, March 7, 2008

"Future View: The New Tech Ecosystems of Cloud, Cloud Services, and Cloud Computing,"
Frank E. Gillett, Forrester, August 28, 2008

Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 USA **Phone** 1-650-960-1300 or 1-800-555-9SUN (9786) **Web** sun.com

06/09