



Linux: The Operating System of the Cloud

By Amanda McPherson

MAY 2009

Linux: The Operating System of the Cloud

By Amanda McPherson



After struggling to reach mainstream status in the technology world for years, utility computing may yet arrive in the coming quarters. This time around, however, it will be known as cloud computing. And when cloud computing does cross the chasm into mainstream adoption, it will be powered by Linux.

The term cloud computing borrows much from the utility and grid technologies that preceded it. Cloud computing describes the availability of off-premise services - traditionally pictured as a cloud in architectural diagrams, hence the name - which themselves may be composed of a variety of hardware and software services. These cloud services may be subsequently re-combined as either a complementary resource, for backup or disaster recovery scenarios as an example, or used as a standalone platform. Cloud computing is characterized principally by its enablement of on demand computing. Even hardware resources such as servers or storage may be purchased in incremental fashion, and billed according to usage. Typical usage may be metered in minutes or hours then, rather than the more typical monthly or yearly terms for traditional contracts.

In many respects, the technology behind cloud computing is anything but revolutionary; the off-site availability of hardware is not a new development. What makes cloud computing more compelling than similar previous offerings is part demand and part technology. On the demand side, more and more enterprises are considering the cloud as a deployment option because the costs of running their own data center are escalating rapidly. But improvements in the technology, particularly in virtualization, distributed computing, and IT management, are what marry the opportunity to the demand.

The progression of cloud computing adoption has been rapid; the initial wave of innovators - primarily consisting of individual developers and startups - has given way to early adopters in the form of cutting edge enterprises and, in some cases, governments. The reasons for selecting the cloud as a deployment target vary from constituency to constituency, but generally center around cost, speed and flexibility. In that respect, adoption drivers for cloud computing often mirror those of open source technologies.

Given that, perhaps the ubiquity of Linux in cloud environments should have been anticipated. To explore why Linux is so compelling for the cloud, it's necessary to examine the characteristics that make Linux a natural foundation for the cloud. The widespread use of Linux in the cloud benefits both those who run and operate clouds, as well as those who build upon them.

Linux: Designed for the Cloud

Linux is the natural technology for enabling cloud computing: it's modular, it's performant, it's power efficient, it scales, it's open source, and it's ubiquitous. And, as the platform upon which the largest cloud infrastructures in the world have been built, Linux - unlike other available operating systems - has little left to prove as a component of cloud infrastructures, be they public or private. "Every time you use Google, you're using a machine running the Linux kernel," as Google's Chris DiBona has said. [1]

Architecture

The Linux kernel supports a degree of componentization that is unmatched amongst general purpose operating systems. Configurable such that it may power everything from a handset to a supercomputer, the Linux kernel is remarkably adaptable to computing environments of all shapes and sizes. "Linux today supports more hardware devices than any other operating system in the history of the world." [2] This is of particular value in highly customized, scale-out cloud platforms, which are required to run on a heterogeneous collection of commodity hardware, networking and storage gear. Beyond the basic compatibility with the mixed nature of the environment, cloud providers will often take advantage of the ability to modify the Linux source code in order to tune and customize the kernel to their specific needs and hardware.

Compatibility

Linux has an extensive application and ISV ecosystem. With thousands and thousands of Linux compatible and certified applications available, users have many options for their specific workload's needs. Customers leveraging Linux for their local and data center needs, then, will be able to extend this advantage to their cloud-based deployments.

For platform providers, Linux is the logical choice. Like the web architectures it spawned from, cloud computing platforms are often composed from many other open source projects, from databases to file systems to application and web servers to language runtimes. By virtue of its quality, ubiquity, and open source nature, Linux is a first choice deployment target for developers of all of the above. As a result, cloud vendors benefit from the wide application catalog available to the Linux platform. Whatever your role, choosing Linux means guaranteeing your application choice.

Cost-Licensing

There exists in some quarters the misconception that Linux is always free in the financial sense of the word. In reality, the overwhelming majority of enterprise and governmental production deployments are commercially licensed and supported. For cloud platform providers, however, the option to run non-commercial distributions does exist, and may be compelling. Platform providers will choose this path because creating a cloud infrastructure composed of thousands or tens of thousands of licensed nodes

1 Q&A: Google's open-source balancing act, http://news.cnet.com/8301-13580_3-9952719-39.html

2 The Linux Driver Model: A Better Way to Support Devices, https://www.linuxfoundation.org/en/Linux_Device_Driver_Model#Executive_Summary

would be uneconomical with traditional per-server or socket models. By leveraging this lower cost approach, cloud providers are able to pass on the savings to customers.

Cost-Power

Besides its advantages in licensing, Linux is a more cost effective platform for providers to deploy and customers to target. Partially because of its usage in small, power sensitive devices, Linux has been the beneficiary of a great deal of research in lowering total power consumption. Heavy attention has been paid, for example, to making Linux more power efficient relative to competitors, via projects like the tickless kernel. Combined with the power saving efforts within cloud data centers, Linux is helping to lower the total solution cost for cloud customers.

Manageability & Staffing

For enterprises and governments alike, questions of resourcing and personnel are an important factor in technology deployment and purchase. In addition to evaluating the merits of a given product or project, organizations must consider how their existing skillsets map to the technologies in question, and further, the ability to hire those skills from the general market in the future. Fortunately, because managing and developing for Linux are common skills, the ubiquity of Linux within cloud platforms means that customers deploying to the cloud can avoid costly re-training for system administrators and developers. In addition to re-purposing existing personnel, deployed IT management systems that already target Linux can be better leveraged with regard to Linux-based cloud nodes.

Standards

One of the most common concerns that analysts and other advisers have for potential cloud customers is the lack of standards, and the resulting potential for lock-in. For all of the advantages in deployment speed and flexibility, the nascent stage of many cloud offerings and the absence of common, agreed upon formats for packaging, runtimes, and virtual images introduces risk. Fortunately, customers can leverage Linux as a hedge against this possibility. The differences between Linux instances hosted in cloud environments and those hosted locally or at a data center, after all, are generally less technical than geographical. By standardizing on Linux workloads, customers will have the flexibility to deploy locally or remotely as the economics and circumstances dictate.

Virtualization

Virtualization, a mainstream technology in most data centers and enterprises, is an important enabler of most cloud platforms. In simple terms, virtualization involves the ability to abstract operating system or application instances from the underlying platform. Windows images or applications, for instance, may be hosted and run on top of a Linux platform using these technologies.

Available to Linux users are a diverse array of virtualization technologies, from the hypervisors that make virtualization possible to the management tooling that allows the virtualized resources to be efficiently marshaled and applied. Equally capable of serving as a host for virtualized instances or as a guest itself, Linux is a stable, secure virtualization option.

Coupled with so-called “live migration” functionality, virtualization can and will also be an important bridge from local environments to cloud based hardware. Linux is therefore an optimal cloud platform, as it is equally adept at playing the role of the host operating system, via technologies like KVM or Xen, or the guest.

Linux is the Cloud's Past

The dot-com era, besides being famous for its irrational exuberance, witnessed the first tentative steps towards cloud computing. Driven by the need to contain costs, many web startups eschewed more traditional scale-up architectures of fewer, more powerful servers and mainframes in favor of massively scaled-out architectures composed of commodity hardware running the Linux kernel. Their success has meant the popularity of the scale-out architectural style more broadly.

In a very real sense, Linux was the catalyst for a new architectural approach and a new generation of web oriented businesses. It is difficult, in fact, to imagine the cloud arriving without a readily available open source kernel like Linux. Few of the original online startups could have risen to their current prominence without it, as both the technology and the economics of alternatives such as Windows would be prohibitive. Linux enabled the generation of an entirely new class of businesses, serving as the technical foundation for successful startups and those that have followed in their footsteps.

Linux is the Cloud's Present

The dominance of Linux within the current crop of cloud computing vendors is eye opening. Virtually every cloud player of any significance features Linux in either primary or supporting capacities, and this adoption is accelerating. Google's recently launched App Engine and Amazon's competitive EC2 product both leverage the Linux kernel, as do cloud offerings from vendors such as 10gen, 3Tera, Media Temple, Mosso, and Zimory. Different providers choose to take different approaches to their products, with some choosing to expose the underlying operating environment (e.g. EC2) and some abstracting it (e.g., App Engine).

One such example of explicit cloud offerings includes commercial Linux distributor Red Hat's partnership with Amazon to offer Red Hat Enterprise Linux, JBoss Enterprise Application Platform, and Red Hat Enterprise MRG Grid and Amazon EC2 Execute Node on EC2. Other Linux distributions that work on the EC2 platform include Oracle Enterprise Linux, SUSE Enterprise Server, and--most recently--Ubuntu Server 9.10.

Linux is still present as the underlying platform in IBM's partnership with Amazon, where IBM's DB2 Express-C 9.5, Informix Dynamic Server Developer Edition 11.5, WebSphere Portal Server and Lotus Web Content Management Standard Edition and WebSphere sMash are the products customers can use in the EC2 cloud. But whether it's an implicit or explicit role, cautious or cutting-edge, Linux is playing a major part in the overwhelming majority of cloud environments.

The fact is that Linux is already the de facto operating system of choice for cloud computing.

Linux is the Cloud's Future

Linux was the core component powering the first generation of web businesses. These businesses could not have been built without a low-cost, flexible software solution as the foundation. Windows, meanwhile, was not a major player for reasons ranging from licensing costs to technology limitations.

The second wave of online business, typified by Google and Amazon, will move farther into consumers' digital lives, running in multiple devices, handling off-line interruptions, improving the browser interface, facilitating mashups between diverse user-chosen services, and a myriad of other issues that are just starting to be glimpsed. This flexibility and utility, based on Linux, is now pushing the cloud into enterprises, governments, and small businesses the world over.

Having proven its worth in high scale, high demand environments, Linux is today being chosen time and again by cloud providers and their customers. Where the likes of Amazon and Google once benefited from the economics that open source and Linux afforded them, their cloud customers will now benefit from the economies of scale that the large providers can leverage. Economies of scale enabled both by Linux and the cloud.

Linux is nothing less than the foundation upon which cloud platforms will be built going forward.

Appendix

The following table lists cloud computing products powered by Linux.

Vendor	Products	URLs
10gen	Mongo	http://www.10gen.com/
3Tera	AppLogic Cloud Computing Platform	http://www.3tera.com/
Amazon EC2	Amazon EC2, Amazon S3, Amazon SimpleDB	http://aws.amazon.com/
Cassatt Corporation	Cassatt Active Response	http://www.cassatt.com/
CohesiveFT	Elastic Server	http://www.cohesiveFT.com/
Dell DCS	Cloud Computing Solutions	http://www.dell.com/cloudcomputing/
Elastra	Enterprise Cloud Server	http://www.elastra.com/
ElasticHosts Ltd.	ElasticHosts	http://www.elastichosts.com/
EMC	Mozy	http://www.mozy.com/
Enomoly	Elastic Computing Platform	http://www.enomoly.com/
Flexiscale	Cloud Computing On-Demand	http://www.flexiscale.com/
Google	Google Apps	http://www.google.com/apps/
IBM	Blue Cloud	http://www.ibm.com/ibm/cloud/
Media Temple	(mt)	http://www.mediatemple.net
Morph Labs	Morph eXchange	http://www.morphexchange.com/
Mosso	Cloud Sites, Cloud Files	http://www.mosso.com/
SalesForce.com	Force.com Cloud Computing	http://www.salesforce.com/
VMware	vCloud	http://www.vmware.com/technology/cloud-computing.html
Zimory	Public Cloud	http://www.zimory.com/