

Open Source Cloud Computing

Thomas Cameron, RHCA, RHCSS, RHCVA

Managing Solutions Architect

Red Hat

<http://people.redhat.com/tcameron>

thomas@redhat.com

thomasdcameron on Twitter

Agenda

- What is cloud computing?
 - By whose definition?
- Why cloud?
- What is Open Source Software (OSS)?
 - A brief history
 - Why OSS?
- OSS cloud stack(s)

What is cloud computing?

- By whose definition?

What is cloud computing?

- Wikipedia calls it “the delivery of computing as a service rather than a product.” ¹
- NIST says: “Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.” ²

What is cloud computing?

- How Stuff Works describes the cloud as web-based applications for e-mail and thin clients. ³
- Infoworld says cloud is “a way to increase capacity or add capabilities on the fly without investing in new infrastructure, training new personnel, or licensing new software.” ⁴

What is cloud computing?

- Microsoft says the cloud is a way to watch “Celebrity Probation” while stranded at an airport. ⁵



What is cloud computing?

- Who's right?

What is cloud computing?

- Everyone.

What is cloud computing?

- Everyone. Except Microsoft.

What is cloud computing?

- Everyone. Except Microsoft. Seriously. Remote desktop is **not** cloud computing.

What is cloud computing?

- On a more serious note... Even though it's fun to bust my competitors' chops, it won't happen again. There are a lot of good cloud technologies from outside of Red Hat. We'll look at several OSS examples today.

What is cloud computing?

- A deeper definition – NIST says:
 - Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.

What is cloud computing?

- Five essential characteristics:
 - On-demand self-service
 - Broad network access
 - Resource pooling
 - Rapid elasticity
 - Measured service

What is cloud computing?

- Three service models:
 - Cloud software as a Service (SaaS)
 - Cloud platform as a service (PaaS)
 - Cloud infrastructure as a service (IaaS)

What is cloud computing?

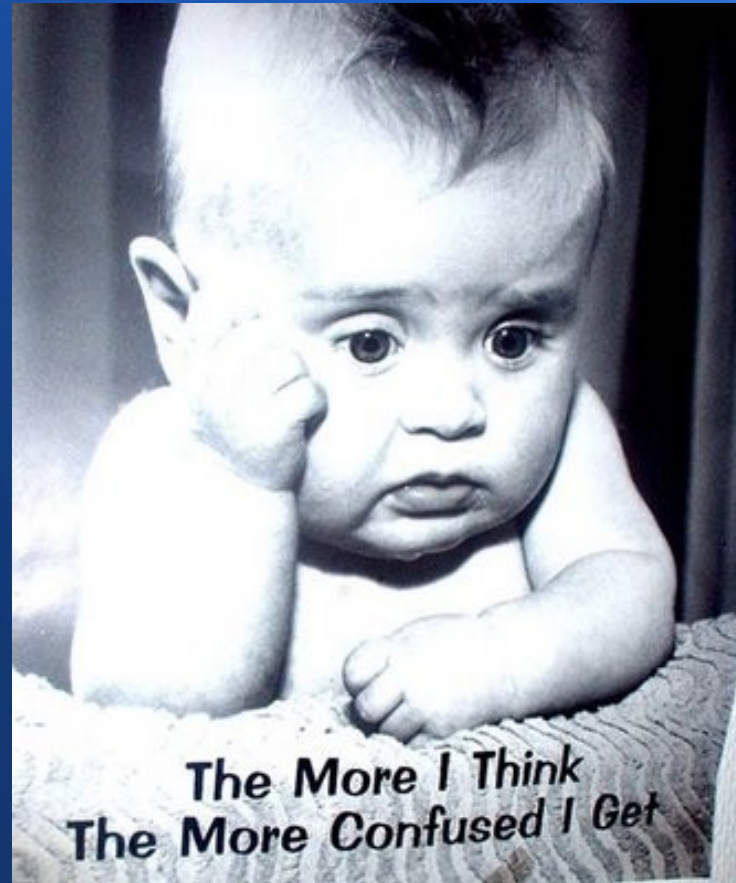
- Four deployment models
 - Private cloud
 - Community cloud
 - Public cloud
 - Hybrid cloud

What is cloud computing?

- Five cloud actors
 - Cloud consumer
 - Cloud provider
 - Cloud auditor
 - Cloud broker
 - Cloud carrier

What is cloud computing?

- Confused yet?



What is cloud computing?

- It's OK... So is the rest of the industry.

Why cloud computing?

Why cloud computing?

What are CIOs interested in for 2011?

- 85% computing capacity is idle
- 70% IT budget spent maintaining legacy infrastructure, not delivering new capabilities
- 50% CIOs say they're understaffed
- Excess capacity is in IT silos
- Organizational resources are stretched
- IT is struggling to keep up with the status quo

Why cloud computing?

What are CIOs interested in for 2011?

- 46% IT execs say meeting SLAs “more or much more” difficult/costly
- 14% annual decline in IT execs who say they are becoming “more strategic” to their organization
- IT organizations can't react quickly
- New projects are slow to implement

Why cloud computing?

- The old way:
 - Overbuying computing capacity, based on infrequently used peak demand
 - Purchasing, installing, and provisioning a server for each new request in the organization. Maintaining it for years to come.

Why cloud computing?

- The old way:
 - Responding slowly to business opportunities or threats because it takes a long time to add or change apps and infrastructure.
 - Having to manage and support many different infrastructure platforms.

Why cloud computing?

- The new way:
 - Buying computing capacity as you need it, when you need it
 - Responding to demand for new services in minutes, not months.
 - Meeting SLAs by using all available computing resources to meet users' requirements.

Why cloud computing?

- The new way:
 - Efficiently managing the diverse systems and platforms you already have, without having to replace them.
 - Using IT to drive innovation and respond to changes in the business.

Why cloud computing?

- IDC #1 Top 10 Prediction 2011:
 - “Private Cloud Plans Mature, [and] Dominate the Enterprise Infrastructure Agenda in 2011.”

Future Priorities for CIOs

2011



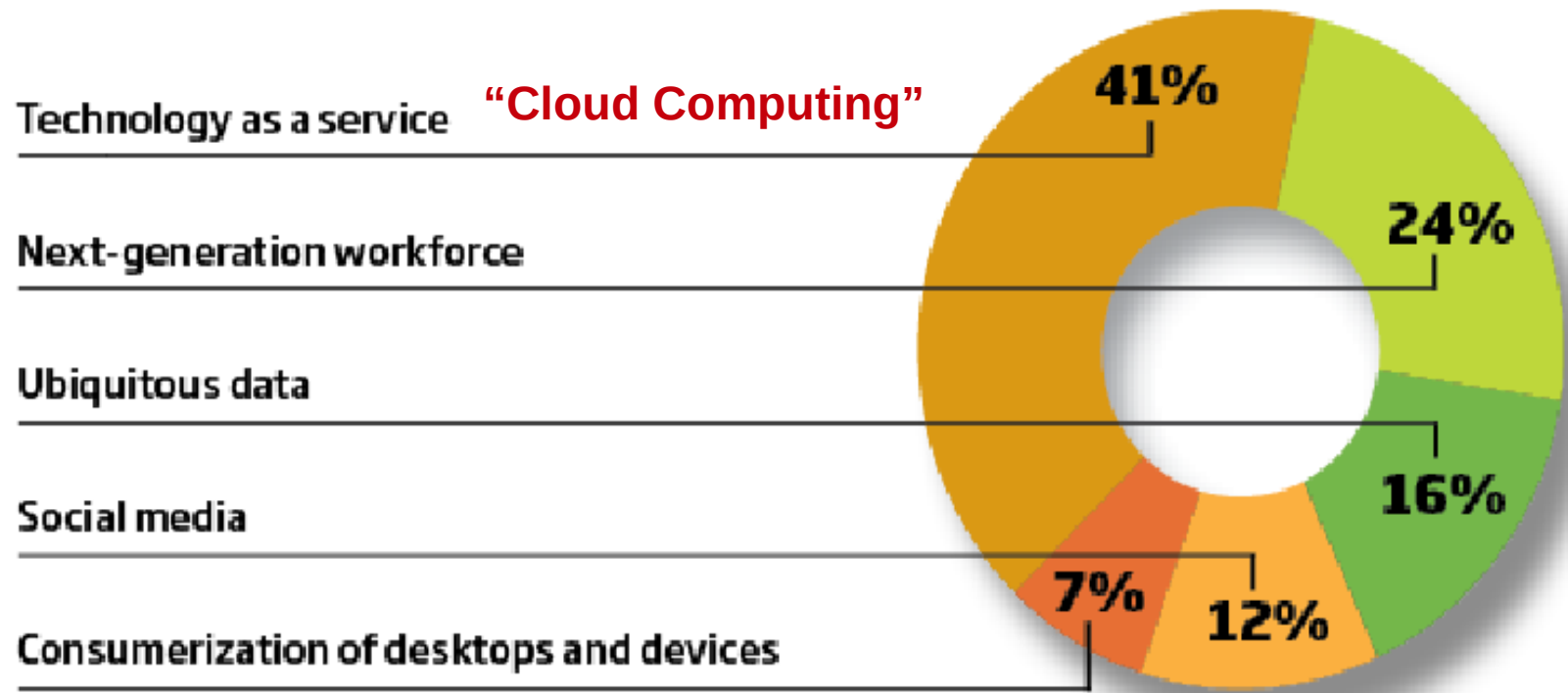
Q. Which of the following activities best characterize your focus and how you spend your time in your current role?

Q. Which of the following activities would you like to spend more time on in the next 3-5 years?

What's Driving Change

Cloud computing will profoundly influence the CIO's role

SaaS
PaaS
IaaS



Source: CIO Magazine 2011 CIO Priorities Survey

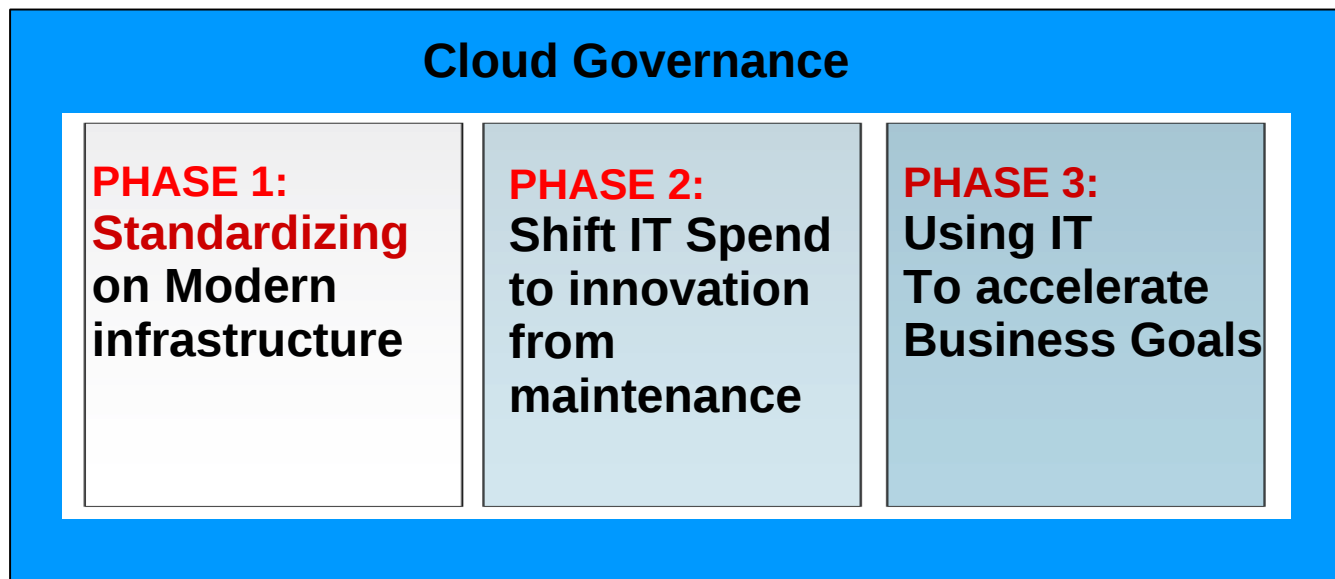
Two Types of Cloud Conversations

Typical Cloud Implementation Cycle Conversation



Focus:
How cloud can be adopted from technical perspective.

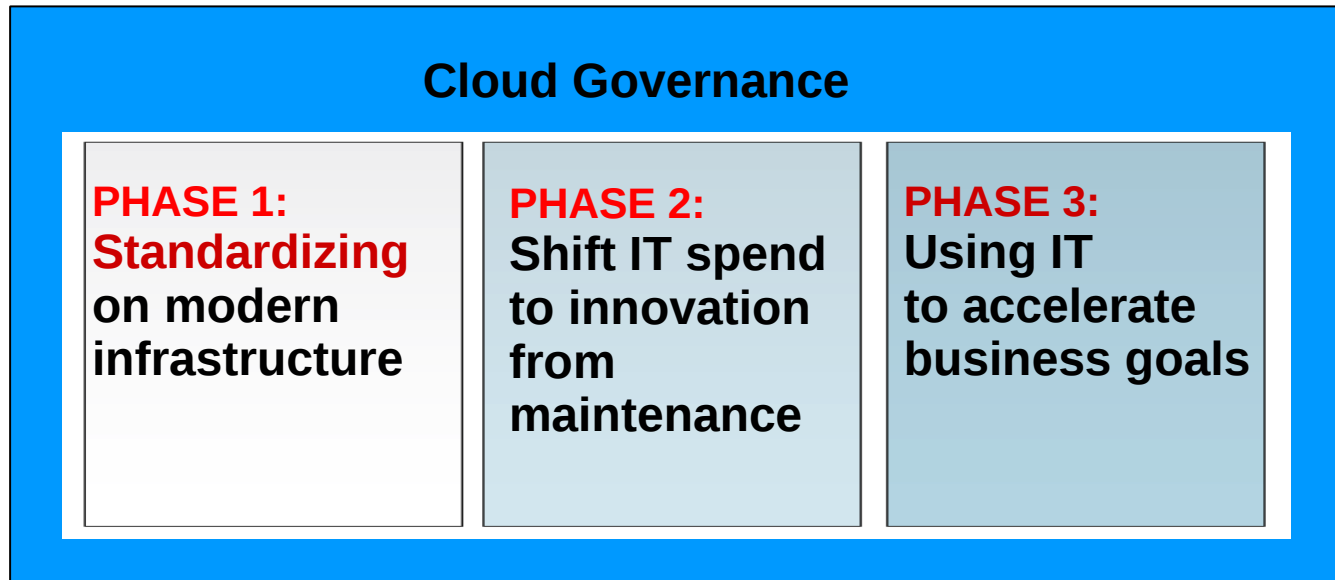
Business Phases of Cloud Conversation



Focus:
CIO view looking at the business benefits.

CIO Cloud Conversation

It's optimizing IT's face to the business, their applications



- The industry has spend the last 3 years reducing costs.
- IT budgets are increasing again...average of 6% in CY'11.
- They are ready to make improvements vs. just cutting costs.
- How can we help them make their business more strategic?

Phase 1: Standardizing on Modern infrastructure

These are the discussions we are having today with CIOs:

Migrating Solaris and legacy Unix systems to Linux

Utilizing commodity x86 hardware to achieve greater choice of supplier

Adding virtualization (perhaps multiple virtualization solutions) to increase system utilization

“IT Organizations must standardize IT infrastructure or their cloud strategy will fail.”

IDC insight Conference, 2011

Benefits of a Modern Infrastructure

More cost efficient

Supports fast pace of innovation

Frees you from single vendor's monolithic stack

Modern Infrastructures can be more easily migrated to the cloud

Modular with no need to Rip & Replace

How IT makes a difference Top CIO Priorities

Expected Accomplishments	2011
Improve end-user workforce productivity	67%
Improve quality of products & processes	52%
Lower company operating costs	51%
Re-engineer core business processes	50%

State of the CIO 2011 Survey

Source: CIO Magazine

Phase 2: Shift Spend from IT Maintenance to Innovation

Automation of manual processes frees up resources for strategic projects

Allows IT to focus on the most visible part of their IT budget...The Applications

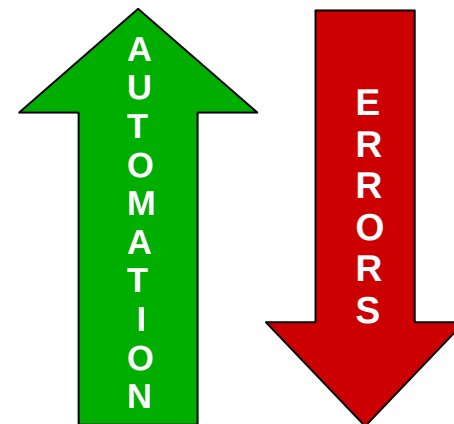
Automate manual processes that are time consuming and error prone

Manage more applications and application images with fewer people

Enable self service

Automate configuration management

Achieve run time management of applications & OS compliance across all deployment models (physical virtual, private and public clouds)

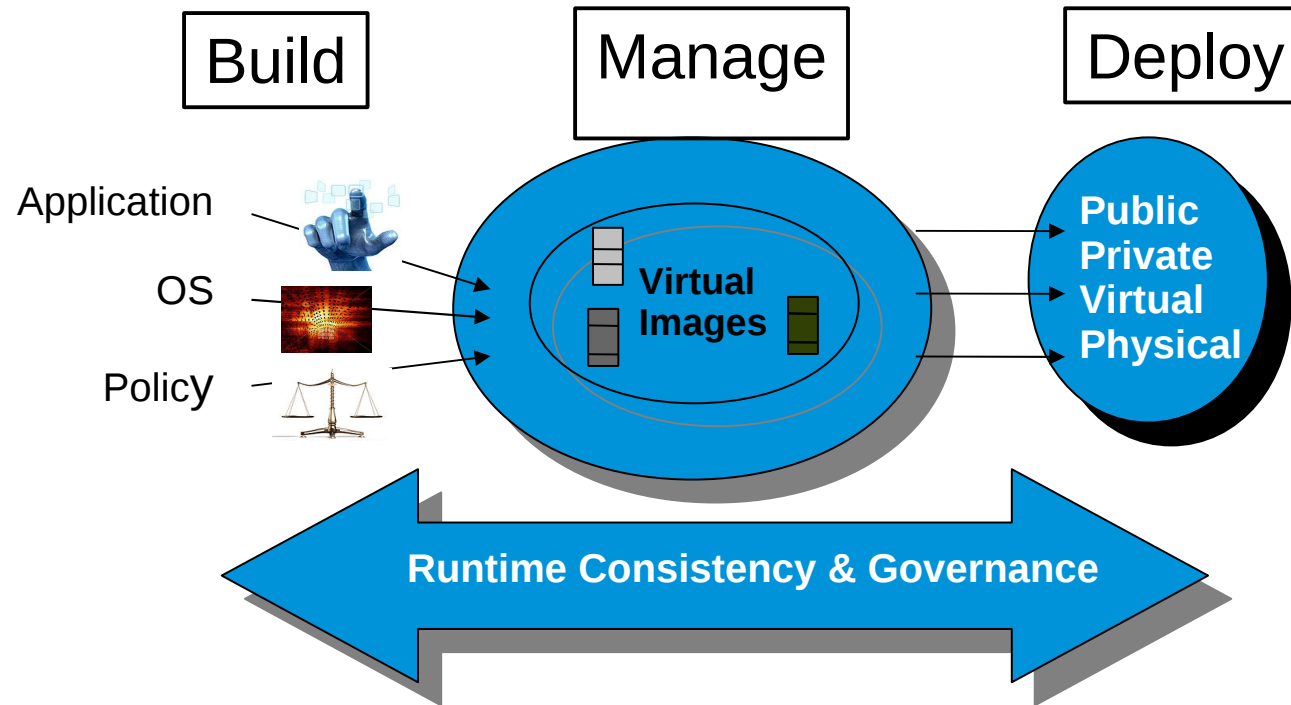


Through 2015, 80% of mission critical outages will be caused by people and process issues.

Source: Gartner

The cloud breaks down infrastructure silos and adds automation to optimize your business application lifecycle

Governance is Important in the Cloud



*Application
Lifecycle
Management
In the Cloud*

Extending on-premise policies to the cloud

Security, access, audit, regulatory compliance & change processes

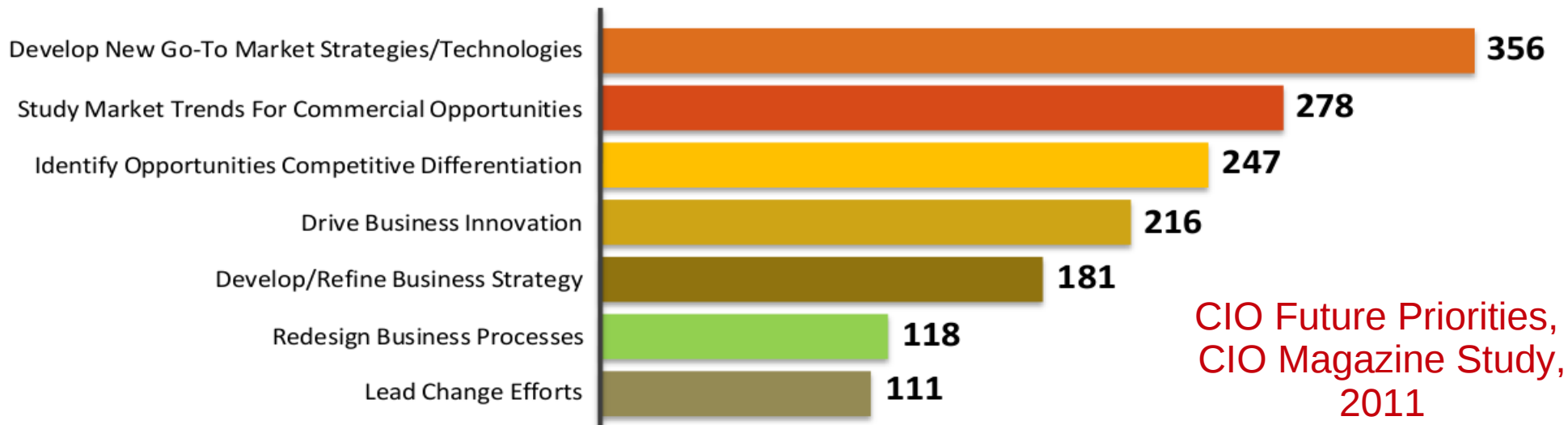
Portability of data & applications

Provides flexibility to move between on-premise & in cloud all with in your stated policies

Risk mitigation

Understanding the risk & balancing benefits with costs

Phase 3: Using IT To Accelerate Business Goals



Hard to separate IT capabilities and business success today...

**IT is redefining how technology can fuel business success ---
run leaner, be more flexible, innovate faster**

Automation frees up resources to engage on new initiatives

Why cloud computing?

- THE CLOUD IS THE SOLUTION
 - Provide new services despite budget constraints
 - Increase enterprise agility
 - Improve service levels
 - Reduce capital requirements
 - Increase operational efficiency

What is Open Source Software?

What is Open Source Software?

- What is Open Source Software (OSS)?
 - Open source is a development method for software that harnesses the power of distributed peer review and transparency of process. The promise of open source is better quality, higher reliability, more flexibility, lower cost, and an end to predatory vendor lock-in. ⁶

What is Open Source Software?

- Meets the following 10 criteria:
 - Free redistribution
 - Source Code
 - Derived Works
 - Integrity of The Author's Source Code
 - No Discrimination Against Persons or Groups

What is Open Source Software?

- Meets the following 10 criteria:
 - No Discrimination Against Fields of Endeavor
 - Distribution of License
 - License Must Not Be Specific to a Product
 - License Must Not Restrict Other Software
 - License Must Be Technology-Neutral

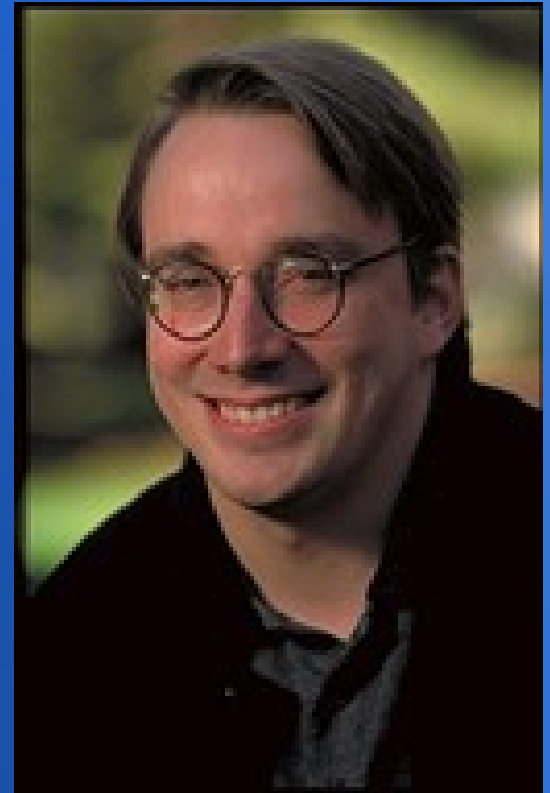
History of Open Source Software

- Richard Stallman at MIT
 - Wanted access to source code so that it could be understood and shared. Father of the Free Software concept.
 - Founded GNU project to build a Free replacement for Unix.
 - Perceived as inflexible in his and GNU's beliefs.



History of Open Source Software

- Linus Torvalds at the University of Helsinki
 - Grad student, came up with a Unix-like kernel for x86 hardware
 - Married his kernel to the GNU utilities to make Linux as we know it today



History of Open Source Software

- Eric Raymond, developer and author
 - Wrote “The Cathedral and the Bazaar,” which Netscape credited with its decision to release the code to the browser
 - Coined phrases like "given enough eyeballs, all bugs are shallow," and “RE,RO.”



History of Open Source Software

- “Open Source” chosen to represent a more pragmatic view of free computing.
- Free as in speech, not just as in beer.

Why Open Source Software?

Why Open Source Software?

- Quality
 - Carnegie Mellon University's CyLab Sustainable Computing Consortium has found that typical commercial code contains 20-30 bugs per 1000 lines of code (KLOC). This has not changed since the 60s.
 - The 2.6 kernel contained approximately 5.7 million lines of code at release.

Why Open Source Software?

- Quality
 - At typical commercial quality, the Linux kernel should have contained 114,000 to 171,000 bugs

Why Open Source Software?

- Quality
 - A Coverity study of the 2.6 Linux kernel found 985 bugs... Most of which had already been discovered and fixed by the community before Coverity released its findings.⁷

Why Open Source Software?

- Quality
 - In this case, Open Source Software yielded 0.434 bugs per 1000 lines of code
 - Typical commercial software has 20 to 30 bugs per KLOC.

Why Open Source Software?

- Who relies on Open Source?
 - NYSE/Euronext is the world's largest exchange, with a \$27.3 trillion dollar market cap. They perform \$141 billion in trades per 6.5 hour day. That's \$6 million in trades per second.
 - Migrated from proprietary Unix to Linux on commodity X86_64 hardware.

Why Open Source Software?

- Who relies on Open Source?
 - By show of hands, how many of you fly on a regular basis?

Why Open Source Software?

- Who relies on Open Source?
 - The Federal Aviation Administration tracks the 7000-8000 airplanes over US airspace at any one time using Linux.

Why Open Source Software?

- Who relies on Open Source?
 - OSS keeps our sailors and soldiers safe in live combat situations via real time weapons control systems and the Land Warrior project.

Why Open Source Software?

- Who relies on Open Source?
 - Federal departments running OSS:
 - NOAA FAA GSA
 - DOE DOC PTO
 - NASA DHS/FEMA DISA
 - U.S. Courts DOI USDA
 - US Air Force US Army US Navy
 - US Marines US Marshal Service
 - Intelligence Community

Why Open Source Software?

- Community = Strength
 - Fedora Linux has had over 33 million unique connections to its content repository mirror system.⁸
 - Ubuntu Linux has had over 12 million unique connections.⁹
 - The Java application server alone from JBoss.org has been downloaded over 7 million times.¹⁰

Why Open Source Software?

- The code these community members work with and on winds up in commercial offerings from SUSE, Red Hat, Canonical and many others
- You will *never* get that sort of QA on any commercial product.

OSS Cloud Stack(s)

OSS Cloud Stack(s)

- Eucalyptus (<http://www.eucalyptus.com/>)
 - Private & hybrid cloud IaaS suite
 - Works across multiple virtualization stacks
 - Compatibility with Amazon Web Services API
 - Installation and deployment from source or DEB and RPM packages

OSS Cloud Stack(s)

- Eucalyptus
 - Secure communication between internal processes via SOAP and WS-Security
 - Support for Linux and Windows virtual machines
 - Support for multiple clusters as a single cloud

OSS Cloud Stack(s)

- Eucalyptus
 - Elastic IPs and Security Groups
 - Users and Groups Management
 - Accounting reports
 - Configurable scheduling policies and SLAs

OSS Cloud Stack(s)

- Eucalyptus
 - For full features, requires “Open Core” enterprise edition not fully OSS
 - Small development community
 - Small corporate backing
 - No application service management

OSS Cloud Stack(s)

- OpenStack (<http://openstack.org>)
 - Very impressive, public/private IaaS joint project between Rackspace and NASA
 - Hardware/hypervisor agnostic
 - Fully OSS (Apache license)
 - Compute, Storage and Image management as separate projects

OSS Cloud Stack(s)

- OpenStack Compute
 - Manage virtualized commodity server resources
 - Manage Local Area Networks (LAN)
 - API with rate limiting and authentication
 - Distributed and asynchronous architecture
 - Virtual Machine (VM) image management
 - Live VM management (Instance)

OSS Cloud Stack(s)

- OpenStack Compute
 - Floating IP addresses
 - Security Groups
 - Role Based Access Control (RBAC)
 - Projects & Quotas
 - VNC Proxy through web browser
 - Advanced Scheduler (pending)
 - Federated Auth with Zones (pending)

OSS Cloud Stack(s)

- OpenStack Storage
 - Store and Manage files programmatically via API
 - Create Public or Private containers
 - Leverages Commodity hardware
 - HDD/node failure agnostic
 - Unlimited Storage
 - Multi-dimensional scalability (scale out architecture)

OSS Cloud Stack(s)

- OpenStack Storage
 - Account/Container/Object structure
 - Built-in Replication
 - Easily add capacity unlike RAID resize
 - No central database
 - RAID not required
 - Built-in Mgmt. Utilities
 - Drive auditing
 - VNC Proxy through web browser

OSS Cloud Stack(s)

- OpenStack Image Service
 - Image-as-a-service
 - Multi-format/container support
 - Image status
 - Scalable API
 - Metadata
 - Image Checksum
 - Extensive Logging

OSS Cloud Stack(s)

- OpenStack Image Service
 - Integrated testing
 - Back-end store options
 - Version control
 - CLI access
 - Built-in Mgmt. Utilities
 - Drive auditing
 - VNC Proxy through web browser

OSS Cloud Stack(s)

- Red Hat Cloud Engine
 - IaaS offering
 - Hypervisor agnostic but RHEL/RHEV Designed Specifically for Cloud Environments
 - Quality of service (QoS) protection
 - Granular, policy-based security in the kernel (SELinux)

OSS Cloud Stack(s)

- Red Hat Cloud Engine
 - Reliability, availability, scalability (RAS)
 - Exceptional performance
 - Virtualization of servers, networks, storage, clients and apps.
 - Full software lifecycle management for hypervisors, guests and apps
 - Private, public and hybrid clouds

OSS Cloud Stack(s)

- Red Hat Cloud Engine
 - Open standard, high throughput messaging infrastructure (MRG)
 - Leverages solid components from the open source community (libvirt, KVM, Puppet, ApplianceTools, Qpid, CloudFiles, Condor, GlusterFS, Rails, Deltacloud, etc.)

OSS Cloud Stack(s)

- Red Hat Cloud Engine
 - Cloud Conductor - Cloud management server to build, run and manage a hybrid cloud
 - Image Factory/Warehouse – Define the images for the cloud and manage the operational life cycle

OSS Cloud Stack(s)

- Red Hat Cloud Engine
 - Management Interfaces – Resource, instance and information (REST) APIs for customized integration
 - Cloud Services – To evolve service abstractions for the cloud to ensure portability

CLOUD ENGINE CAPABILITIES

- Simultaneously build private, hybrid clouds and manage/federate public clouds
 - Management of the cloud is by policy/mass actions
 - Utilizing Quota, QoS, limits
 - Users, groups, permissions
 - Admins create pools, add cloud provider accounts to enable self-service
 - Storage Management
 - Operational/Archival
 - Image repository hides complexities of image placement and management
 - Engine model allows configuration of services and resources
- Self Service
 - Admin to user Self-Service via portal
 - Self service API (Deltacloud)
 - Policy
 - Scheduling constraints (location, quota, data cost)
 - Quota / QoS visibility
 - Overflow
 - Security constraints
 - Scale from private cloud to cloud hybrid
 - Resource Mgmt/Scheduling
 - Uses MRG-Condor engine
 - Accounting/ Billing
 - We will provide data, to be integrated with external billing/ accounting

IMAGE FACTORY

- Enable Red Hat content, other content and your private content to be included in system definitions to run in the cloud
- Create content descriptions for the deployments in the cloud and the associated (VM) images
 - Templates – The “recipes” to generate an Image given a suitable (source) content repository (e.g. list of packages, services offered, requires, etc.)
 - Assemblies – An aggregation (e.g. stack) or association (e.g. topolog) of one or more templates
 - Deployables – A set of one or more assemblies and the information (e.g. operational parameters and configuration) in order to boot, initialize and provide the defined services
- Build the appropriate images (bootable or non-bootable) for the provisioned Provider “targets”
 - AMI, QCOW, VMDK, etc.

CLOUD STORAGE SERVICE

• Archival Store (Image Repo)

- large objects
- explicit whole-file get/put operations (via HTTP)
- explicit data placement
- virtual-machine images, media assets, etc
- S3, Google, CloudFiles, Azure replication comparable

• Operational Store

- filesystem-like
 - directories, small files, single-byte read/write, etc.
- higher performance
- caching and/or replication

Daemon-based storage on non-RHEL/RHEV cloud

Meta data daemons and repository manage storage domains

ARCHIVAL STORE DETAILS

- Uses table backend (Tabled or CloudFiles)
- ...**PLUS** a distributed database for tags/metadata
 - tags are key/value, whatever you want them to be
- ...**PLUS** federation/ replication daemon
- Data inserted centrally, accessed globally
- Replication via “policy equations”
 - evaluated in context of object plus site
 - tag comparisons, booleans, pointer following, etc.
 - (creation_date > xxx) && (creation_date < yyy)
 - security_level != zzz
- Remote sites can always “pull” non-replicated data

OPERATIONAL STORE DETAILS

- Follow-on (post v1) offering
- Based on parallel filesystem
- Participates in iwhd/repo naming schemas
- Improved scalability and ease of deployment
- Security and billing/accounting
- Caching
 - hierarchical, synchronous, invalidation based
- Next: WAN replication
 - peer to peer, asynchronous, version-vector based

REPLICATION OVERVIEW

- Any number of sites
- All can write even when partitioned
 - eventually consistent (“eventual” usually means $<1s$)
 - operation transfer (“log shipping” or “patches”)
 - conflict resolution via vector clocks if possible...
 - ...then timestamps if sufficiently different...
 - ...then “pecking order” if all else fails.

ON-PREMISE JBOSS AND CLOUD ENGINE

- Customers today are building internal solutions that leverage JBoss Enterprise Middleware to streamline and simplify application and service provisioning
- Red Hat will continue to help organizations evolve their middleware to incorporate attributes associated with private clouds such as self-service provisioning
- Cloud Engine will provide automation and provide portability across multiple clouds

OSS = Community = Strength

OSS = Community = Strength

- Pssst – that's YOU!
- We all learned it in grade school: sharing is good for everyone!
- Come on in, the water's fine!

Thank you!

References

- 1 http://en.wikipedia.org/wiki/Cloud_computing
- 2 <http://www.nist.gov/itl/cloud/index.cfm>
- 3 <http://computer.howstuffworks.com/cloud-computing/cloud-computing.htm>
- 4 <http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031>
- 5 <http://www.youtube.com/watch?v=Lel3swo4RMc>
- 6 <http://www.opensource.org/>
- 7 <http://www.linuxtoday.com/developer/2006031800826OSCYDV>
- 8 <http://fedoraproject.org/wiki/Statistics>
- 9 <http://ostatic.com/blog/canonical-announces-12-million-ubuntu-users-google-makes-a-comeback>
- 10 <http://www.jboss.org/jbossas/downloads/>