Red Hat Enterprise Linux on System z

Brad Hinson
Worldwide Lead, Linux on System z
bhinson@redhat.com

Agenda

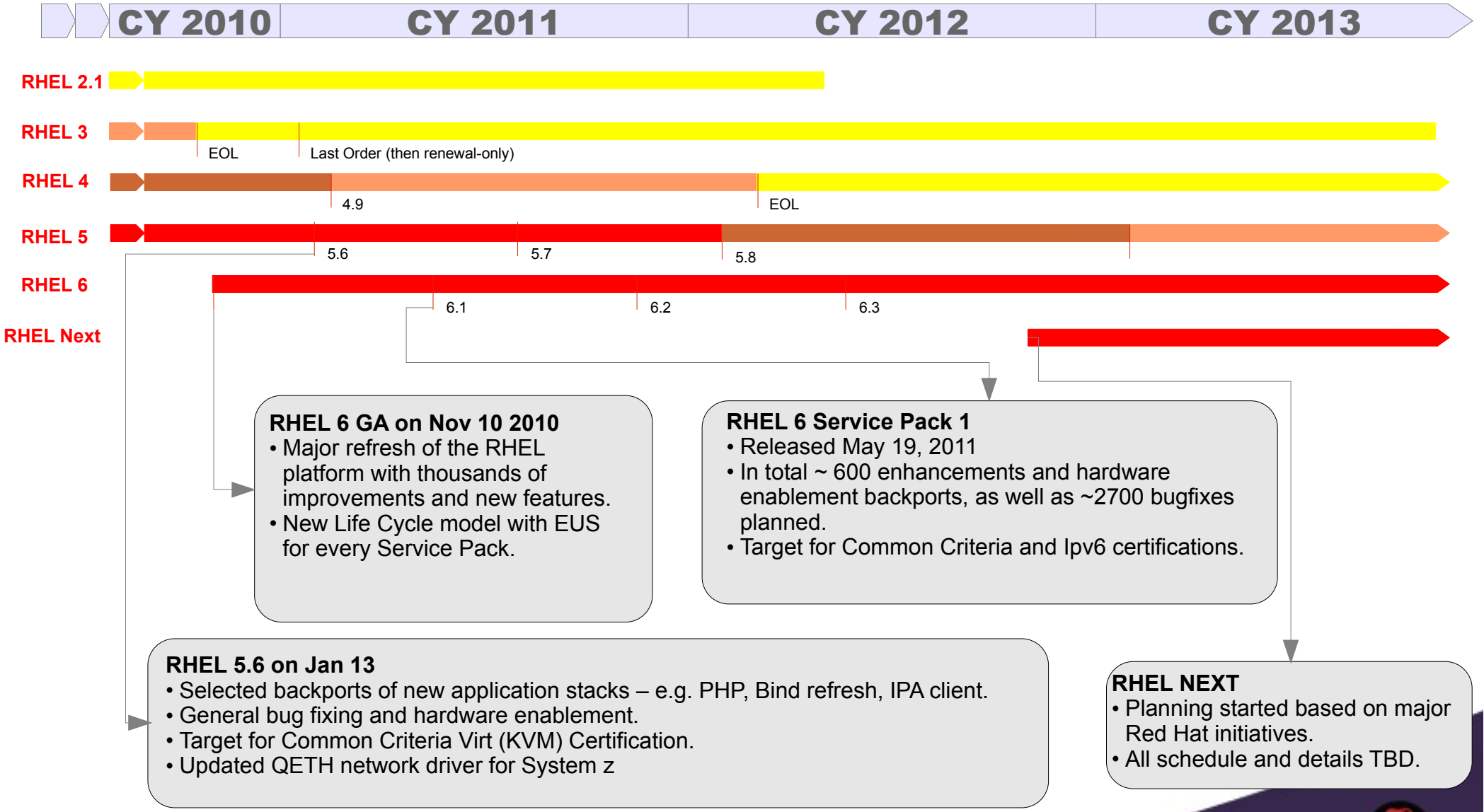
- RHEL Lifecycle
- Latest Updates
  - RHEL 6.1
  - RHEL 6.2
  - RHEL 5.7: Feature Highlights
- Appendix: Systems Management with RHN Satellite

RHEL Lifecycle

Red Hat Enterprise Linux Life Cycle Overview

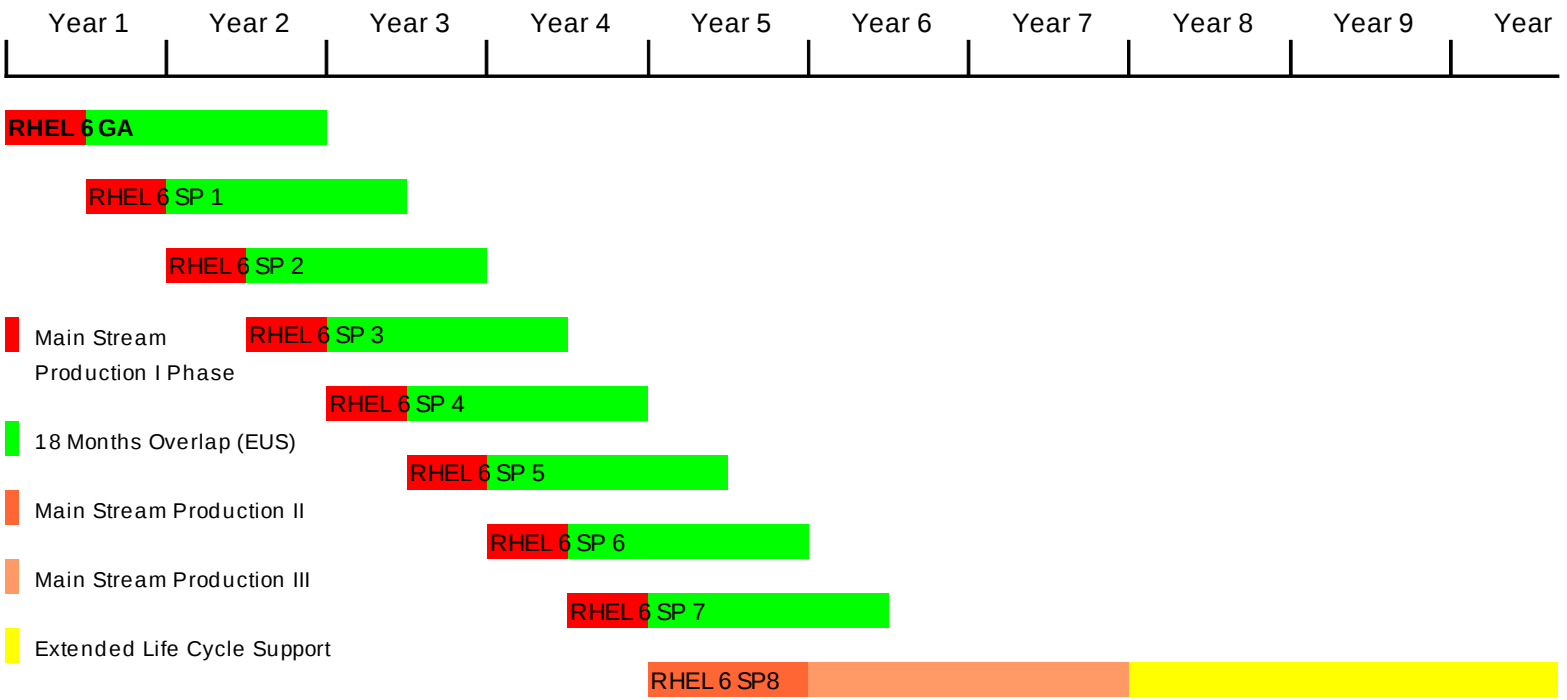
- **Fully supported through standard life cycle of 7 years from GA plus 3-year optional extension.**
- **Asynchronous bug fix, enhancement, and security Errata Advisories for critical issues as needed.**
- **All Errata Advisories and Service Packs are incremental on a single stream.**
- **ABI stability is guaranteed for the full life cycle.**
- **Production 1 Phase**
  - Non-critical Errata Advisories aggregated in Service Packs (aka Minor Releases), approx. 2 per year, may stretch out later in the life cycle; includes Feature enhancements and new Hardware support. Typically 4 years.
- **Production 2 Phase**
  - Transition from Production 1 to Production 3, Bug fixes and minor hardware enablement
- **Production 3 Phase**
  - Time between the final Service Pack and the end of the regular 7-year life cycle.

Red Hat Enterprise Linux High Level Roadmap



Note: This is a current planning snapshot. Dates and details subject to change.

Extended Update Support (EUS)



- Optional add-on to regular RHEL Subscription, that provides independent life cycles for the individual Service Packs during Production I phase.
- The overlap is 18 months leading to a ~ 2 year life cycle for each SP.
- Selective backports of fixes into the respective Service Pack
  - Provides Critical Impact security errata independent of customer requests.
  - Selected urgent priority defect fixes to address production blocking problems reported by EUS customers.
  - Does not provide incremental features or hardware enablement
- Inherits support SLA from underlying RHEL subscription (requires Standard or Premium support).

RHEL 6.1: Feature Highlights
Released May 19, 2011

Performance and Application Scheduling

- CPU scheduling algorithms optimized
  - Results in 3 to 5% performance gain.
- Better concurrent processing by making use of RCU locking in the scheduler.
  - Read-Copy Update: Access to shared data without traditional locks. Designed for today's faster CPUs.
  - http://lse.sourceforge.net/locking/rcupdate.html
- Improvements to tickless timer algorithm.
- Performance improvements since 6.0
  - Java workload improved from 1% to 3%.
  - Transaction Processing workload improved from 6% to 8%.

File Systems and I/O

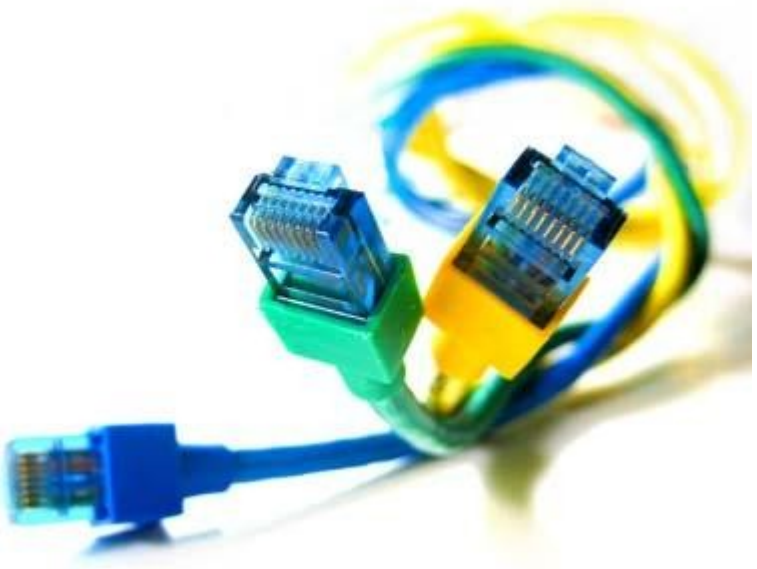
- CIFS improvements for access to Windows Shares
  - Multi-user mounts for more secure access
  - Support for Unix-style symbolic links on CIFS (mfsymlink)
- Quota management
  - Consolidation of quota management tools for file systems

File Systems and I/O

- I/O Barriers
  - Implementation reworked for improved disk performance
- LVM
  - Improved recovery times by skipping scans on failed devices.
  - Snapshot of mirrors
  - Support for mirror devices whose constituent devices are striped (RAID 0+1)

Networking

- Significant optimization in the way network traffic is processed in single and multi-CPU environments.
- Reduced latency for re-transmission of lost packets in time sensitive applications.
- Transparent proxy (TProxy) support.
- DHCP support for IPv6.
- Active-Active bonding for load sharing.

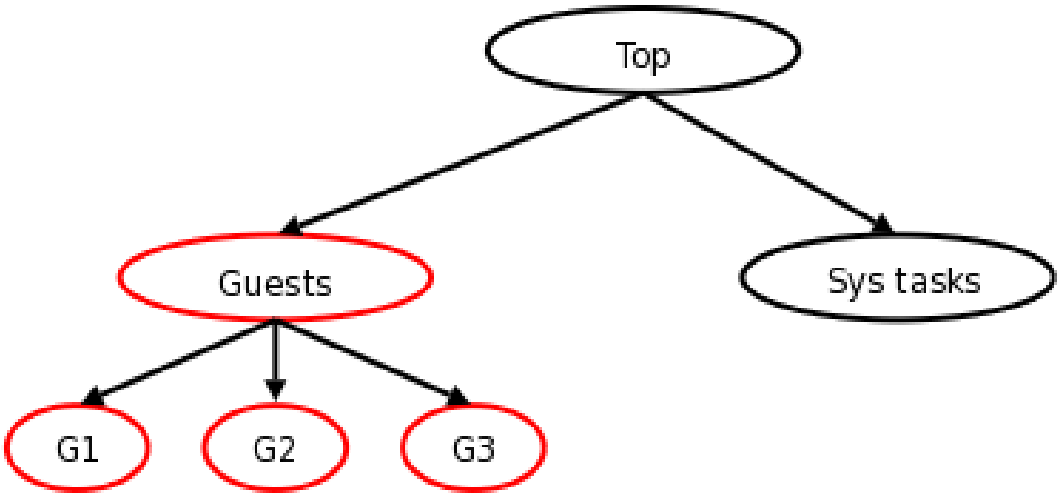


In Depth: Cloud Enablement with Control Groups

Problem: “I want to implement a chargeback model.”

Solution: Control Groups (cgroups)

- Cgroups are “process containers”. Lets you transform groups of applications into workloads



In Depth: Cloud Enablement with Control Groups

- **Resource Limiting**

  Specify limits on CPU, memory, and even file system usage

- **Prioritization**

  Give mission critical workloads higher priority than others

- **Accounting**

  Run report on resource utilization, i.e. for billing purposes

- **Isolation**

  Separate namespaces for groups, so they don't see each other's processes, network connections or files

- **Control**

  Freeze groups for checkpointing or restarting workloads

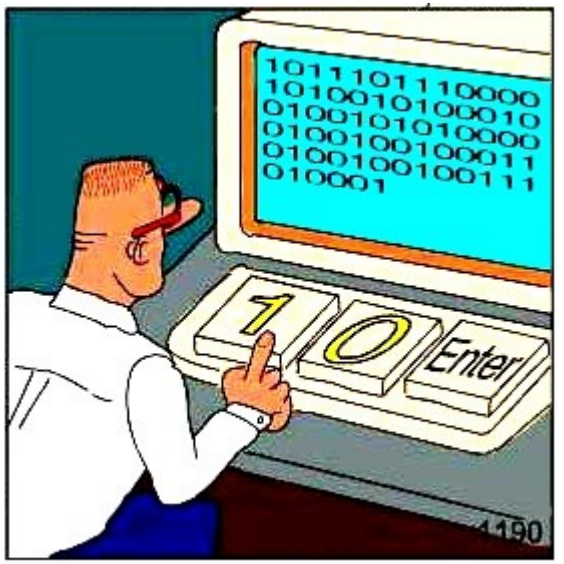
redhat

Resource Management with Control Groups

- Block I/O throttling
  - Limit I/O rate for device based on cgroup membership.
- Balance of throughput and fairness between groups
  - via new tunable “group_idle”.
- Reduce latency for interactive tasks running under CPU intensive workloads via autogroup.
  - Prevents single process from monopolizing the system.

Software Development

- SystemTap
  - Remote scripting capabilities, numerous performance optimizations.
- GDB
  - C++ debugging enhancements and Python support.
- Valgrind
  - Handle CPUs with three levels of cache
- GCC Compiler
  - Bug fixes and optimizations.
- Eclipse
  - Update to the platform (Helios) and plugins



REAL Programmers code in BINARY.

Security and Audit

- Multiple updates to System Security Service Daemon (SSSD)
  - SSSD Integration with identity management services
  - Better DNS-based discovery
  - Auto renewal of Kerberos tickets, plus support for Kerberos FAST protocol
  - Password obfuscation (LDAP)
- Centralized management of SSH keys using LDAP
- Identity Management
  - Password policy management for users and groups

System z Specific Updates

- 49 z-specific features, 56 z-specific bug fixes for RHEL 6.1

- Here are some highlights...

System z Specific Updates

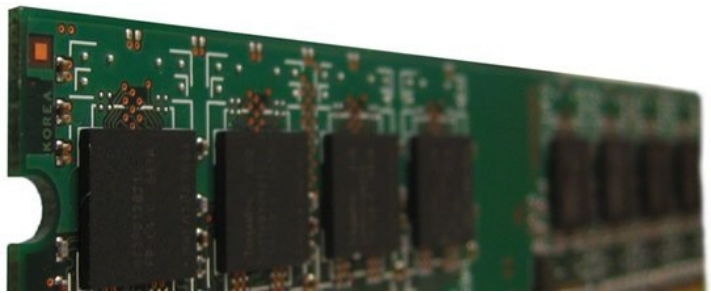
- Fix and recompile openSSH to enable HW Crypto
  - Performance improvement. Enable openSSH to offload secure processing to Crypto card.
- zEnterprise support for 4096-bit RSA FastPath
    - This feature extends the support for current hardware acceleration of RSA encryption and decryption to handle the zEnterprise Crypto Express3 card.
- Installer: /boot partition on LVM
  - zipl bootloader supports device-mapper (LVM & multipath) devices. Installer now allows /boot these devices
- Installer: /boot on ext4 partition
  - zipl bootloader supports ext4 partition. Installer now allows this.

System z Specific Updates

- Dynamic memory resize tools: lsmem/chmem

```
# lsmem
Range                                  Size (MB)   State   Removable   Device
==========================================================================
0x0000000000000000-0x000000000fffffff  256         online  no          0
0x0000000010000000-0x000000002fffffff  512         online  yes         1-2
0x0000000030000000-0x000000003fffffff  256         online  no          3
0x0000000040000000-0x000000006fffffff  768         online  yes         4-6
0x0000000070000000-0x00000000ffffffff  2304        offline -           7-15

Memory device size  : 256 MB
Memory block size   : 256 MB
Total online memory : 1792 MB
Total offline memory: 2304 MB
```



System z Specific Updates

- CMSFS write support
  - Support for writing to CMS file system. You can now your PROFILE EXEC with vi!
- Exploitation of z10 prefetching instructions
  - This is a toolchain enhancement from IBM. Prefetching instructions have been introduced to enhance memory access.
- Exploitation of z196 out-of-order instruction scheduling
  - Generate faster code sequences, and use CPU facilities to allow better instruction scheduling.
  - Recompile programs with --march=z196 and/or –mtune=z196
- Apply System z optimized sysctl settings by default
  - Apply kernel tuning settings to /etc/sysctl.conf optimized specifically for System z

System z Specific Updates

- hyptop: Hypervisor “top” - Show IFL usage across LPARs

```
# hyptop
14:08:41 | H05LP30 | CPU-T: IFL(18) CP(3) UN(3)                  ? = help
cpuid   type    cpu   mgm visual.
(#)    (str)    (%)   (%) (vis)
0        IFL  96.91  1.96 |############################################ |
1        IFL  81.82  1.46 |#####################################        |
2        IFL  88.00  2.43 |########################################     |
3        IFL  92.27  1.29 |##########################################   |
4        IFL  83.32  1.05 |#####################################        |
5        IFL  92.46  2.59 |##########################################   |
6        IFL   0.00  0.00 |                                             |
7        IFL   0.00  0.00 |                                             |
8        IFL   0.00  0.00 |                                             |
9        IFL   0.00  0.00 |                                             |
             534.79 10.78
```

RHEL 6.2
GA: December 2011

Focus areas

- Resource management
- Networking
- File system and storage
- Identity management
- Installation and advanced storage
- Security and Standards
- Subscription Management
- Feature & Bug Fix Summary

Resource Management - Control Groups

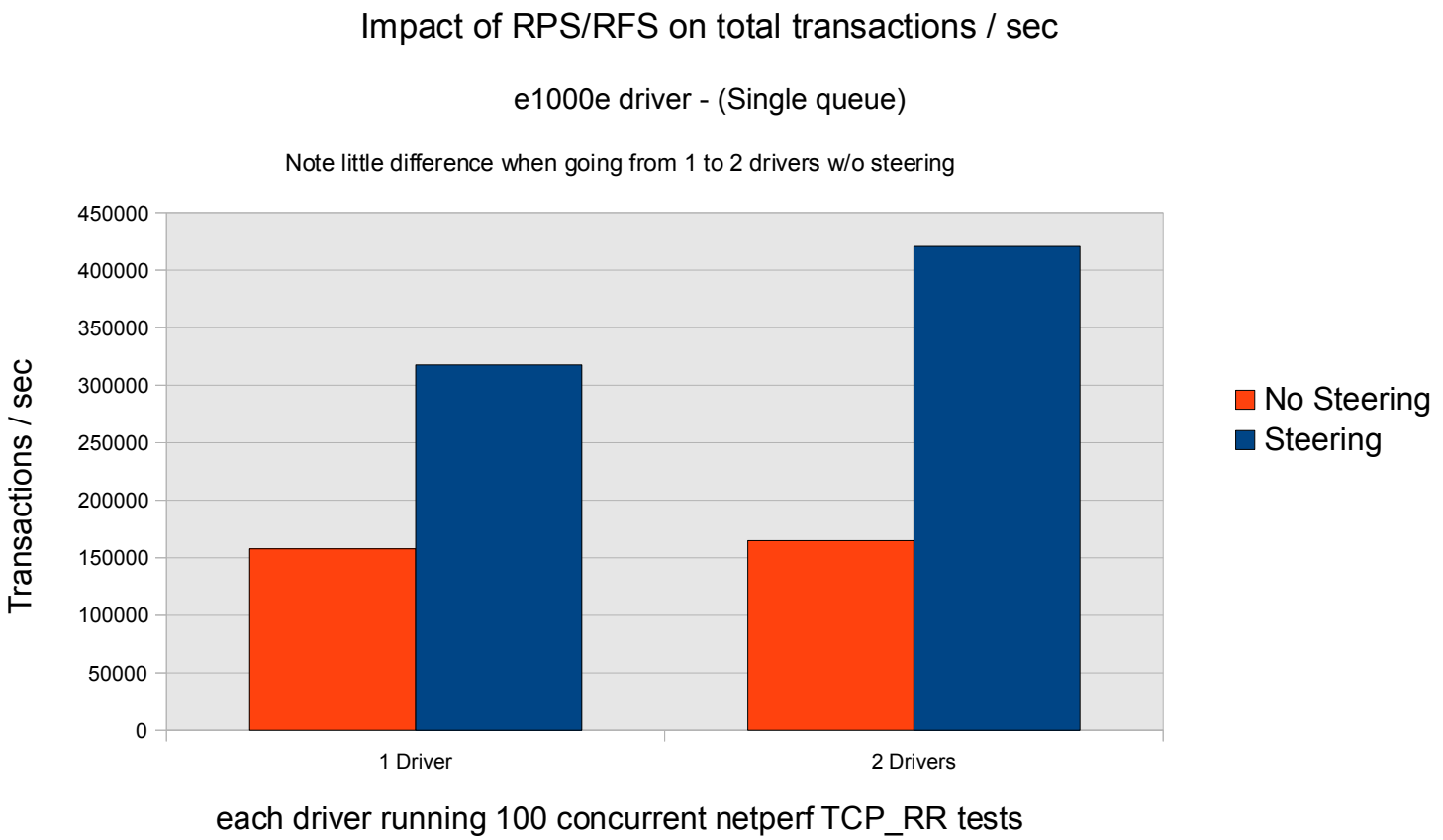
- Ceiling enforcement via CPU cgroup
  - Guarantee precise processor time. Relevant in multi-tenancy environments that are sensitive to over/under allocation.
- CPU controller scalability improvements
  - Relevant for large SMP systems.
- I/O controller performance improvements
  - More efficient use of locks.
- Reduced memory usage by more efficiently allocating pages within cgroups

Networking

- NetworkManager improvements
  - Backend support for NIC bonding.
    - Note: Useful for LPAR-only environments
- Stream Control Transmission Protocol (SCTP)
  - Robust alternative to TCP allowing for “multistreaming” and “multihoming”.
- Transmit Packet Steering (XPS)
  - Selection of transmit queue based on configuration.
  - Throughput improvement by 20% to 30%.

Transmit Packet Steering (XPS)

Greatly improved message throughput as a result of network packet flow/steering.



File System

- Parallel NFS (pNFS) *(Technology Preview)*
  - Allow for larger data transfer rates.
    - Typically processing metadata is the bottleneck
    - Process metadata separately from actual data
  - Allows client to access storage devices in parallel
  - Limited to client-side functionality for file layout only.

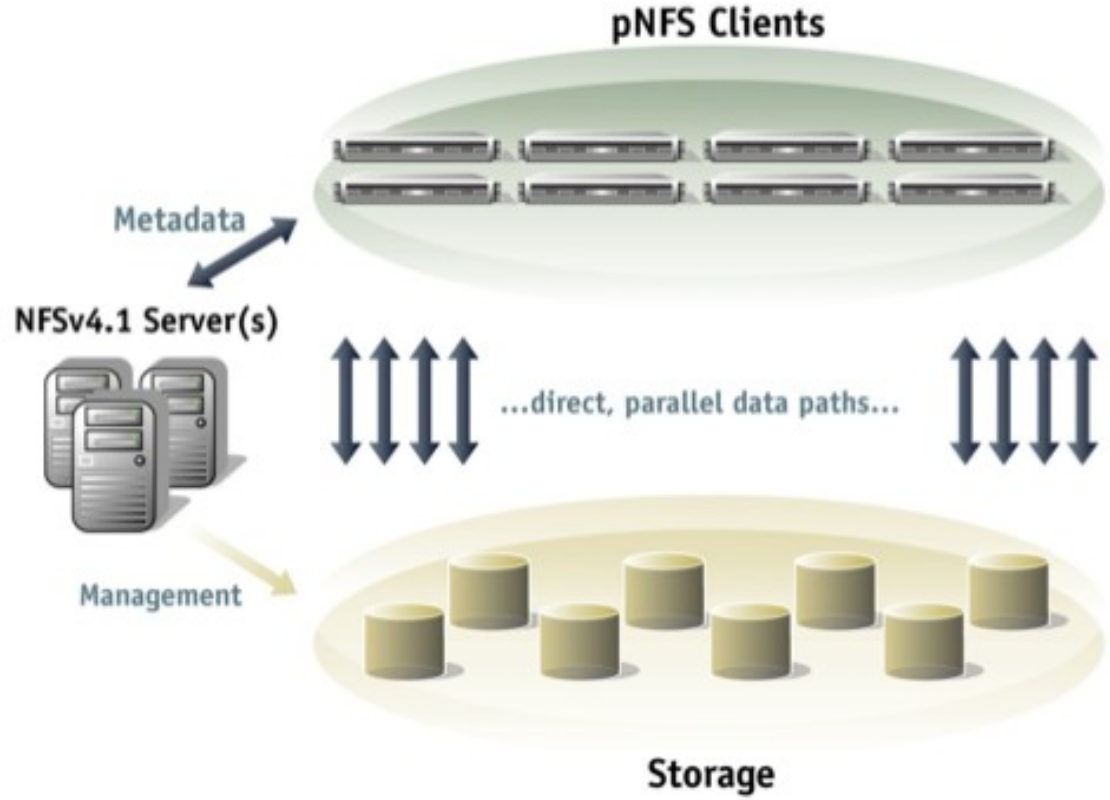
File System

- Parallel NFS (pNFS) *(Technology Preview)*
  - www.pnfs.com/docs/DMG_Parallel_NFS.pdf

Today's NFS Architecture

File System

- Parallel NFS (pNFS) *(Technology Preview)*
  - www.pnfs.com/docs/DMG_Parallel_NFS.pdf



Parallel NFS architecture

File System (cont.)

- EXT4
  - Faster creation times by delaying the initialization of the inode tables.
  - New option “-E lazy_itable_init”
    - Delays initialization until first mounted
    - Once mounted, initialization happens in background, so file system is immediately usable
    - Creation time on large file systems reduced as much as 100%
- CIFS
  - Asynchronous write support resulting in significant performance improvement.
  - Performance benefits up to 200%

Storage

- Reduced activation times for LVM devices
  - Relevant in high-density environments (E.g. large number of virtual guests).
- Improved LVM efficiency for solid-state backend devices

Identity Management

- Centralized administration
  - Allow administrators to manage identities of users and servers centrally.
- Interoperability with Active Directory
- Single-Sign On (SSO)
- Based on open standards
  - Integrates Kerberos, LDAP, DNS and x.509 certificates into a unified identity management solution.

Installation

- Support for partitions greater than 2.2 TB
  - Data devices only
- Support for device identification using WWIDs
  - Fiber Channel devices can be now specified by a World Wide Name (WWN) or World Wide Identifier (WWID) for unattended installations.
  - Easier to identify storage devices

Error detection and reporting

- Improved ABRT framework
  - Easier configuration of plugins and settings.
  - More consistent way to store error reports.
  - Switch to using a non-privileged account for most of the processing that results in a more secure environment.
  - Greater plugin stability.

Desktop and Graphics

- X Server rebase
  - Easier to maintain X.org and Mesa drivers in future RHEL 6 releases. Increased system stability in the long-run
- Monitoring of larger CPU count in gnome-system-monitor.

Security, Standards and Certification

- Kerberos FTP client improvements
  - Command line limit increase to better handle large message processing environments.
- Common Criteria Certification
  - Target for Evaluation Assurance Level (EAL) 4+
- FIPS-140
- FIPS 201 PIV Smartcard

Subscription Management

- Certificate based RHN for new installations
  - Certificate based access that is highly available with the help of the globally distributed content delivery network (CDN).
- Entitlement certificates for disconnected systems
  - Allow for offline registration of up to 25 systems.
- Auto regeneration of renewal certificates
  - It is now possible to automatically regenerate new entitlement certificates after the renewal of a subscription.

Feature & Bug Fix Summary

- Automatic detection of read-only DASD
  - Useful in read-only root setups
- Access to raw ECKD data from Linux
  - Linux can now see the entire disk, useful for “dd”
- Crypto CPACF improvements for zEnterprise
  - Support for new algorithms
- Optimal qeth network settings
  - Enable checksum and GRO, increase buffers

Feature & Bug Fix Summary

- New HiperSocket communication infrastructure
  - Now supports TCP/IP congestion management
- cpuplugd service updates
  - Updated rules for CMM
  - Now disabled by default

- 88 total z-related features & bug fixes

RHEL 5.7
GA: Q2 2011

New Feature: SCAP

- Security Content Automation Protocol (SCAP)
  - Open source framework for maintaining security of enterprise systems
  - Verify presence of patches, check system security settings, examine system for signs of compromise
- Includes a library and a set of utilities
- Allows security managers to use OVAL and XCCDF to verify security configuration and vulnerability status
- Backported from RHEL 6

New Feature: Developer Tools

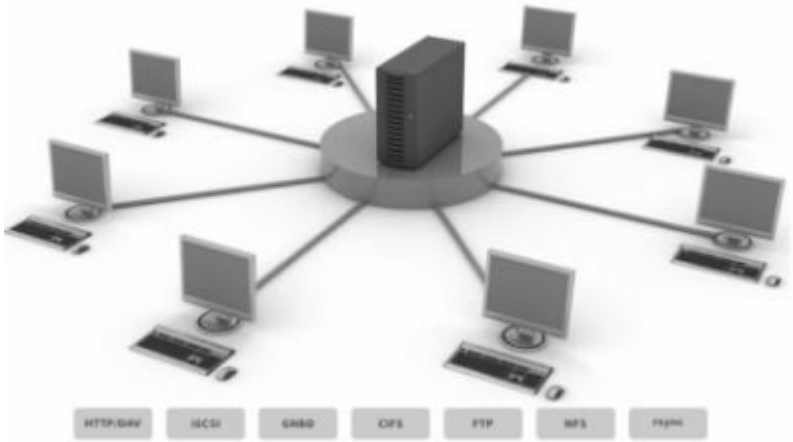
- CMake
  - Cross-platform build system
  - Generates native makefiles and compiler independent configuration files
  - Supports developers with targets across multiple operating systems.
  - Backported from RHEL6
- Buildsys-macros
  - Support for dist tags for developers building RPMs
  - Backported from RHEL6

Feature: Remote Sync (rsync)

- Rsync has been updated to version 3.0.7
  - Improved replication speed
  - Replication starts while file list is still being compiled
- Companies requiring global data set replication should see major benefits

Features: Remote File Systems and Storage

- Updated Automounter (autofs)
  - Support for localityName attribute in LDAP maps
  - Encrypted secret for LDAP authentication
- iSCSI initiator
  - Support for s390x architecture



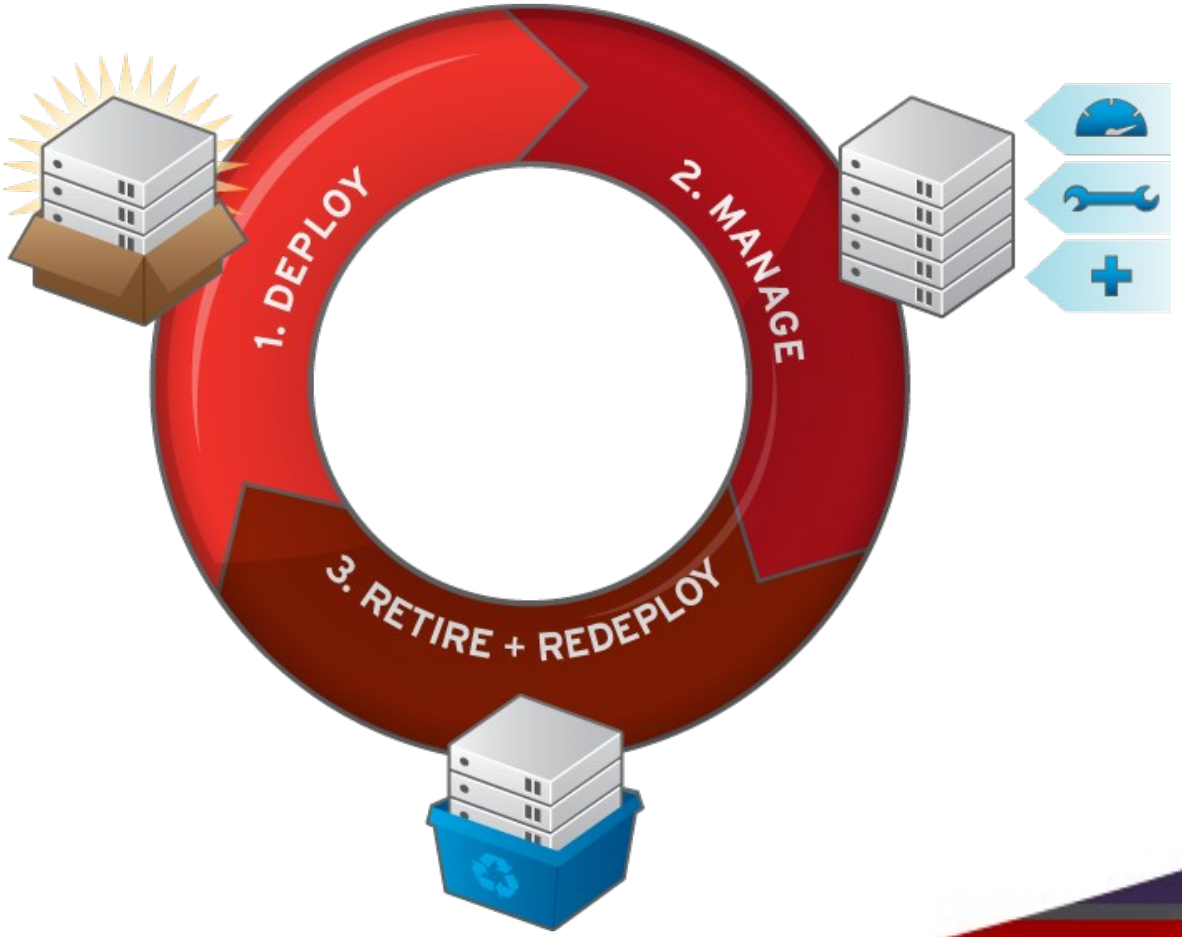
Documentation Links

- Documentation/Getting Started
  - **Redbook, “z/VM and Linux on IBM System z: The Virtualization Cookbook for Red Hat Enterprise Linux 6.0”**
    - http://www.redbooks.ibm.com/abstracts/sg247932.html
    - Covers RHEL 6 and z/VM 6.1
  - **DeveloperWorks**:
    - http://www.ibm.com/developerworks/linux/linux390/documentation_dev.html
  - **Knowledgebase**:
    - http://kbase.redhat.com/
    - Search “s390”
  - **http://www.redhat.com/z**

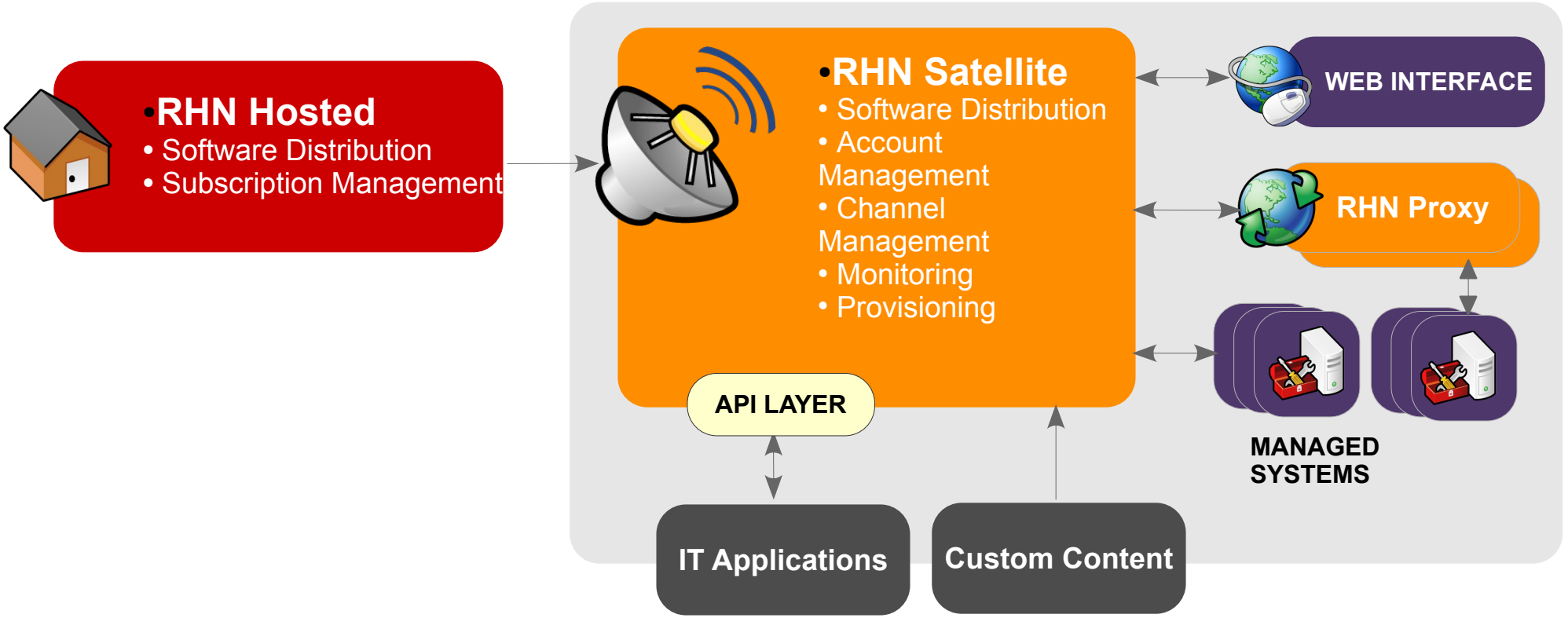
Thank You

Appendix

Systems Management with RHN Satellite



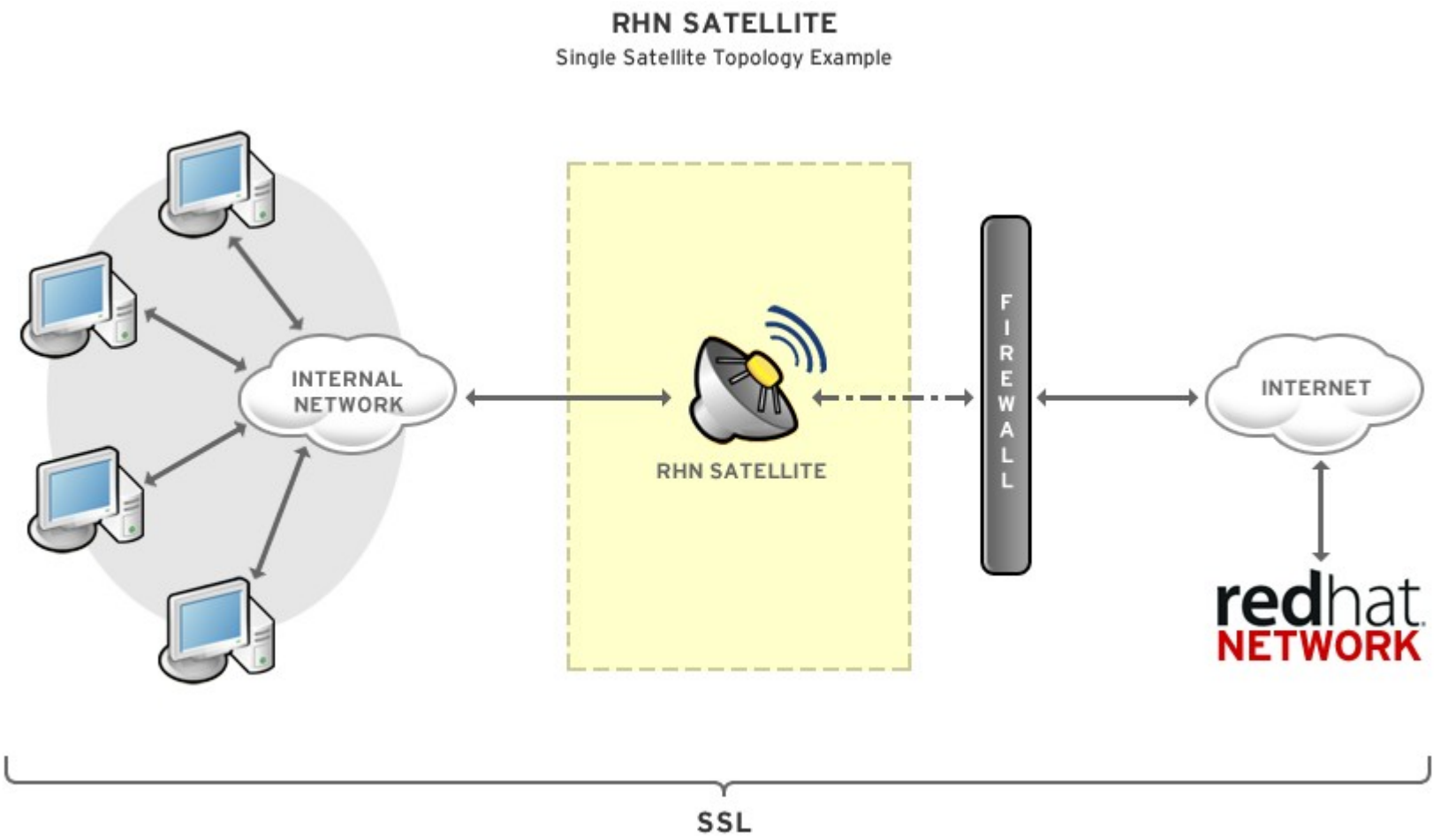
Satellite deployment model



- Enterprise management solution – enhanced control
- Local database stores all packages, profiles, and system information
- Syncs content from RHN Hosted
- Custom content distribution
- Can run disconnected from the Internet

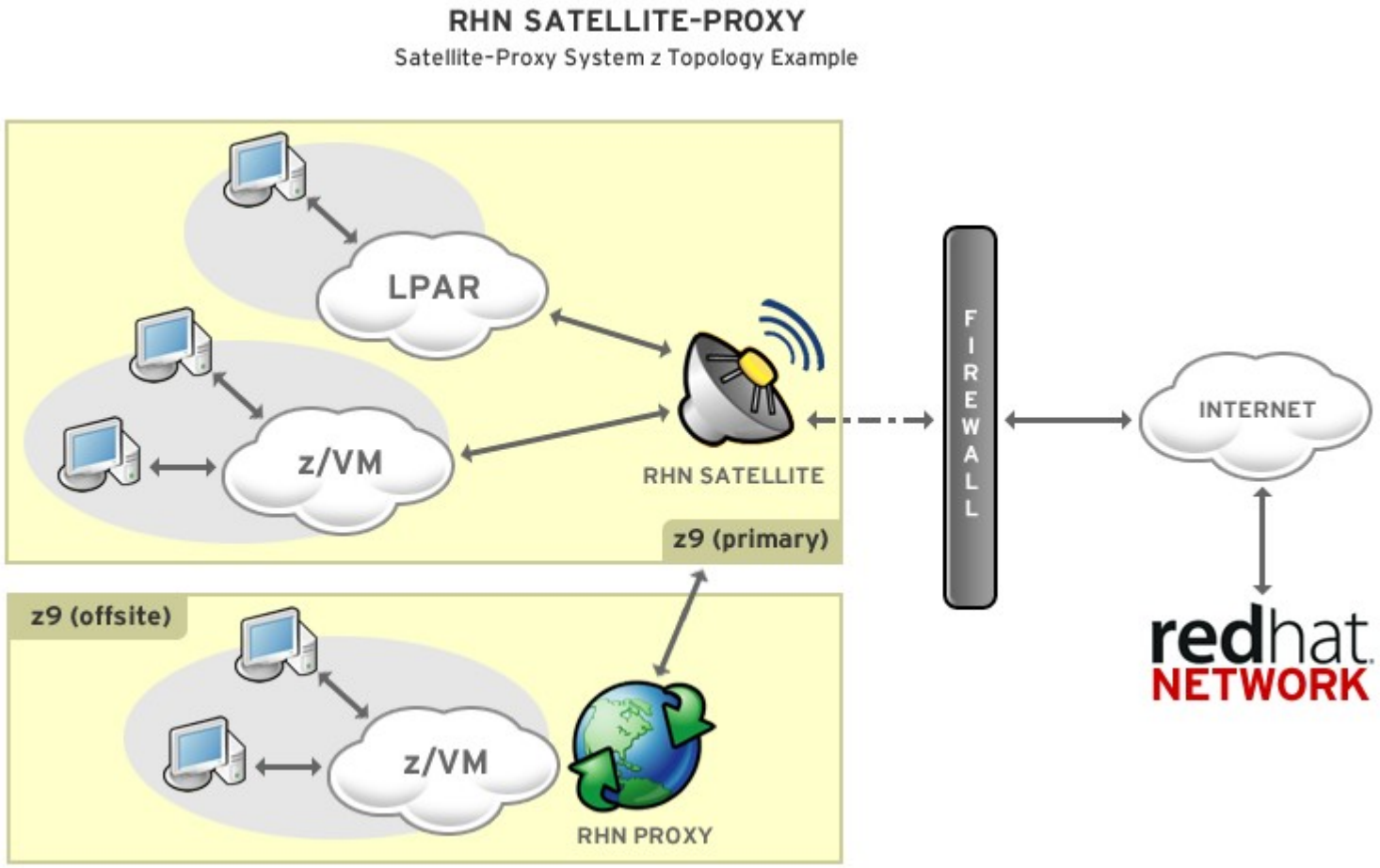
RHN Satellite

RED HAT NETWORK
:: Enterprise Systems Management



RHN Satellite on System z

RED HAT NETWORK
:: Enterprise Systems Management



PXE Deployment on System z

zPXE

- Same configuration/profile on all guests
- Read-write 191 disk not required for each guest
- All changes kept on management server
- Flexibility of kickstart
- Same principles as traditional PXE, adapted to System z
- Fits with configuration management tools (cobbler)
- Easy to update

https://fedorahosted.org/cobbler/wiki/SssThreeNinety