RED HAT
ON THE MAINFRAME
THE REASONS ARE COMPELLING

IBM.

RHEL 5.3 Update (for System z)

Agenda & Introduction

- RHEL 5.3 released 20-JAN 2009
- What's new?
- What's new specifically for System z?

Agenda & Introduction

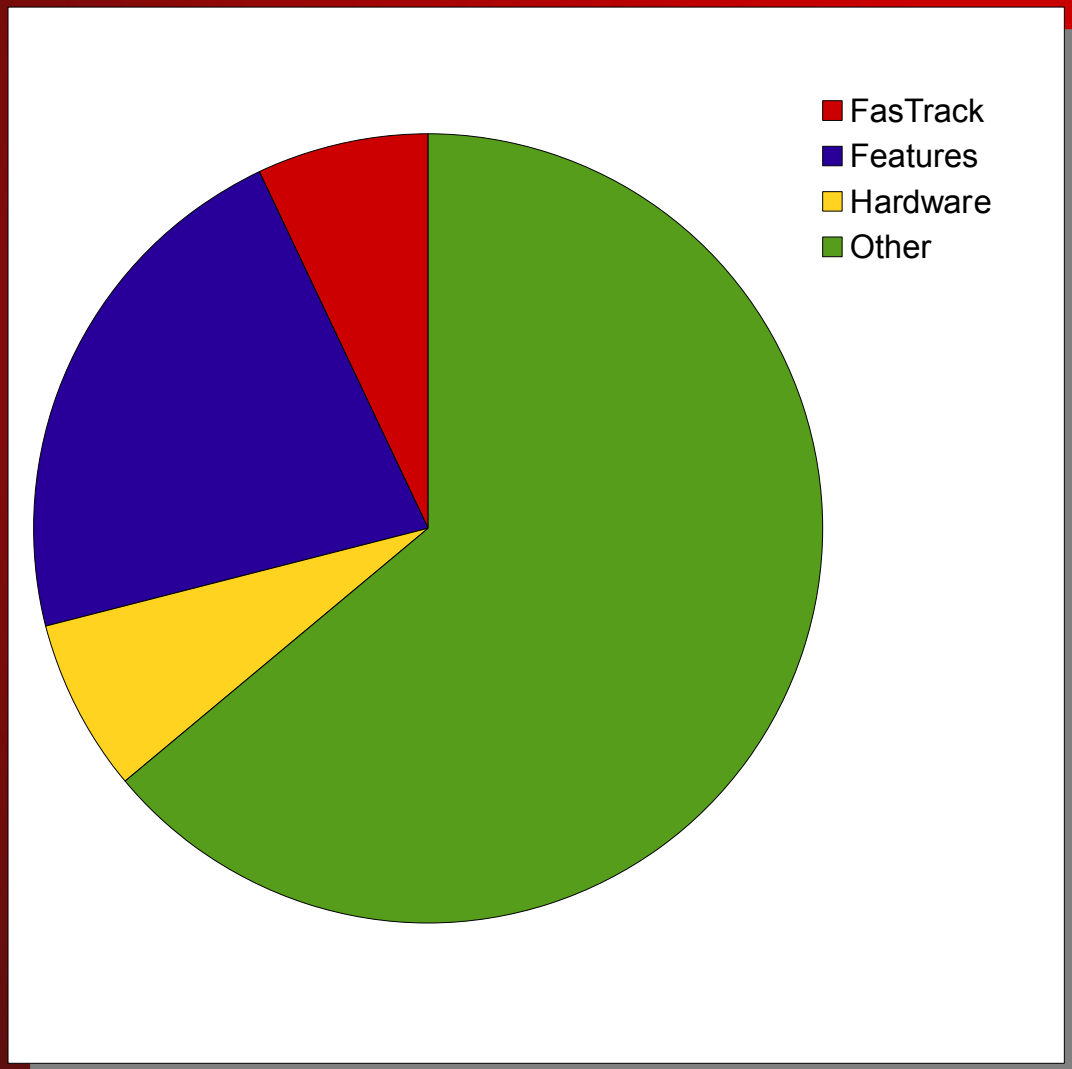
Shawn Wells <swells@redhat.com>

W/W Lead, Linux on System z

(+1) 443 534 0130

- Based in Washington, D.C.
- Global responsibility for Red Hat's System z activities

RHEL 5.3 Overview



~150 additions, ~3,400 BugZillas

- **7% FasTrack**
  Early release of low impact fixes
- **7% Hardware Enablement**
  New chipsets & processor feature support
- **21% New Features**
  Feature requests from customers & partners
- **65% "Other"**
  Feature enhancements, Bug fixes, Documentation

RHEL 5.3 Tech Review
(All Architectures)

RHEL 5.3: Technical Review

Kernel Updates

- Added private futexes support
- Added preemt-notifiers implimentation
- Included tracepoint/markers infrastructure
- Added response oriented wake up behavior to scheduler
- Fixed gettimeofday for HPET, PMTimer, TSC
- Updated NMI infrastructure to latest
- Added ACPI tstate support (processor throttling control)
- Enhanced partition statistics
- Enabled CIFS' DFS support and updated CIFS to latest
- Updated Autofs4 to latest
- Added kbrobe-booster and return probe-booster support
- Added PCI domain support
- Added RAID 4/5/10 in dm-raid

RHEL 5.3: Technical Review

System Services

- **Rebased:**
  - Cups (print server), now with full Kerberos support
  - ksh, lm-sensors, lftp, net-snmp, openIPMI, openLDAP
  - Openmotif, python-urlgrabber, rpm, tog-pegasus, vnc, yum,
  - Yum-utils
  - Samba from 3.0.28 to 3.0.32
    - Supports Windows Vista and 2008. Various fixes for Domain Controller functionality (Interoperability with Citrix and Domain trusts)
  - RPM to Fedora 9 version, which includes numerous bugfixes
  - DHCPv6 support

RHEL 5.3: Technical Review

Developer Enhancements

- **SystemTap**
  - New documentation and guides
  - SystemTap script compile server support – no need to replicate debuginfo RPMs
- **GDB Debugger**
  - Rebase, now based on version 6.8 (previously based on version 6.5)
  - Multiple location breakpoints for C++ templates, constructors, inclined functions, etc
- **OpenJDK**
  - Full open-source JDK for Java-1.6 support
  - Tested with Java SE 1.6 Technical Compatibility Kit (TCK) ==> 100%
  - X86 and x86_64 architectures only

RHEL 5.3: Technical Review

Developer Enhancements, cont

- **Tech Preview**
  - SystemTap userspace support
  - GCC 4.3 – experimental support for C++0x, integration of MPFR library
  - OpenMP v3.0 – shared memory parallelism in C, C++ and Fortran

RHEL 5.3: Technical Review

Hardware Enablement (distributed)

- Production 1 lifecycle phase allows the enablement of new hardware capabilities
  - Red Hat Enterprise Linux 5.3 extends the enterprise customers hardware choices.

- New processor and chipset support
  - Intel i7 (Nehalem) support – new Intel architecture with NUMA support

- Better power management
  - Deep C-State support for Intel Tylersburg chipset. Allows ~16w power reduction on Nehalem/Tylersburg platforms

- New graphics support

- PCI domain support
  - Large systems support for multiple PCI segments

RHEL 5.3: Technical Review

Encryption & Security

- Root and swap encryption support in the installer
  - Hibernate / resume support with encrypted disks

- Authentication and Identity
  - Pkinit interoperability fixed & clients can now be configured to use keys for client certificates which may not contain Kerberos-specific extensions
  - Krb5 now applies the correct file context to database lock files
  - Krb5 servers no longer log everything twice by default
  - nss_ldap now configured with support for paged results extension
  - SELinux enablement of new NetworkManager and audit functionality
  - Improved Audit and Logging
    - TTY input audit support
    - Remote audit logging via unencrypted connections

RHEL 5.3: Technical Review

Technology Previews

- AIGLX: X Server & updated Mesa pkg
- Compiz composition manager
- Dm-multipath install
- Dogtail GUI
- Limited eCryptFS support
  - Add ecryptfs support to kernel
  - Add authentication to crypto library in kernel
- Ext4 file system / e4fsprogs
- Firewire
- GCC 4.3
  - Compiler based on gcc 4.3
  - OpenMP3 conformance
- Generic fcoe (potential TP)
- Indic languages: Assamese, Kannada, Sinhalese, Telugu

RHEL 5.3: Technical Review

Technology Previews, cont

- iSCSI boot
- CIFS Kerberos & kernel DFS support
- Ktune: a service that sets several kernel tuning parameters to values suitable for specific system profiles
- SystemTap utrace (user space tracing)
- Trusted Computing Group (TCG) / Trusted Platform Module (TPM Support
  - Include the TCG stack
  - Include the Trousers TSS stack
  - Add trust computing/trust platform module in kernel and tpm-tools boot-loader support will be considered for inclusion in a future release

RHEL 5.3 Tech Review

(System z Specific)

RHEL 5.3: System z Specific

BugZilla ID	Summary
46327	stage1: sshd error loading shared lib: libfipscheck.so.1
184770	LTC18425-62140: (big) xDR system Initialization for LPAR Clients
472788	rhel 5.3 snapshot3 scsi mpath install failed on z9bc lpar
439479	LTC:5.3:201474:Include gcc 4.3 as Add-On for latest z10 instruction set support
439440	LTC:5.3:201160:Long Random Numbers Generation
439441	LTC:5.3:201158:Selective Logging of ECKD DASD devices
439482	LTC:5.3:201542:FCP - Enhanced Trace Facility
447379	LTC:5.3:200994:Linux CPU Node Affinity
463917	unable to find DASD drives to install
439484	LTC:5.3:201490:Libica Library: Integration of Icainfo
43946	LTC:5.3:201360:OSA 2 Ports per CHPID Support - Installer Enhancements
466474	[RHEL5.3] *** glibc detected *** /usr/bin/python: double free or corruption (!prev): 0x000 0000080d55e90 ***
466305	cosmetic error message: failure in nl_set_device_mtu
466291	anaconda silently omits uninitialized disk

RHEL 5.3: System z Update

xDR System Initialization for LPAR Clients

[BugZilla 184770, LTC 18425-62140]

RHEL 5.3: System z Update

Include GCC 4.3 as Add-On for latest z10 instruction set support

[BugZilla 184770, LTC 18425-62140, Red Hat Errata 2009:0077-7]

Includes the following z10 specific patches to GCC:

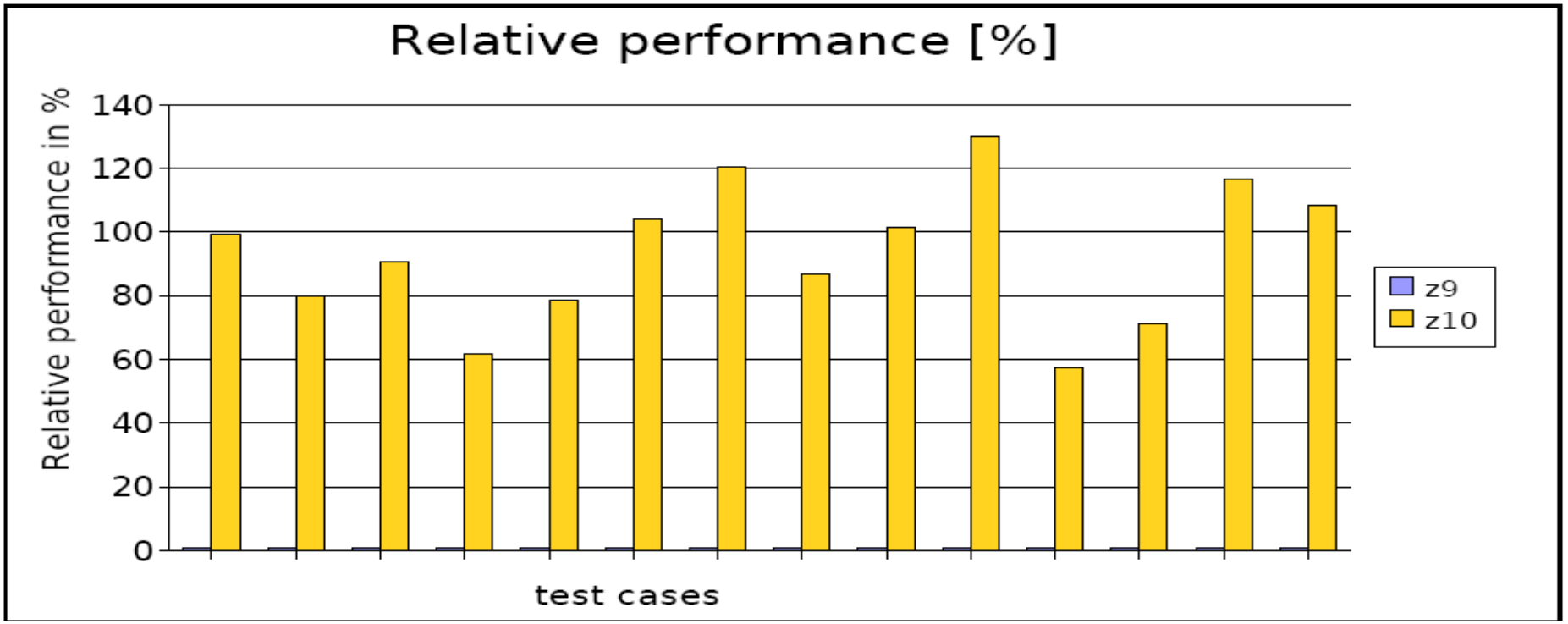
- Introduce TARGET_MEM_CONSTRAINT macro
- Introduce 'enabled' insn attribute
- S/390: Exploit the 'enabled' insn attribute
- S/390: Replace 'm' with 'RT' constraints
- S/390: Add the -march=z10/-mtune=z10 options for z10
- S/390: Support the new instructions introduced with z10
- S/390: z10 pipeline description
- PR36822 recog: Reorder extra memory constraint checks for inlineassemblies
- S/390: Fix -march=z9-ec -msoft-float

RHEL 5.3: System z Update

Include GCC 4.3 as Add-On for latest z10 instruction set support

[BugZilla 184770, LTC 18425-62140, Red Hat Errata 2009:0077-7]

Overall improvement with z10 versus z9: 1.9x
Work in progress with gcc-4.3 compiler using -march=z10 option



Graph taken from Mustafa Mešanović's T3 Boeblingen presentation, 1-JULY 2008, “Linux on System z Performance Update”

RHEL 5.3: System z Update

Long Numbers Generation

[BugZilla 439440, IBM LTC 201160, Red Hat Errata 2009-0225]

Provides access to the random number generator on the crypto card in order to meet high volume random number requirements.

Specific performance numbers not available at this time.

RHEL 5.3: System z Update

Selective Logging of ECKD DASD Devices

[BugZilla 439441, IBM LTC 201158, Red Hat Errata 2009-0225]

Improves RAS characteristics by performing selective logging of DASD Sense data. Adds more comprehensive messages.

View the patch @ https://bugzilla.redhat.com/attachment.cgi?id=313271

RHEL 5.3: System z Update

FCP – Enhanced Trace Facility

[BugZilla XX, IBM LTC 201158, Red Hat Errata 2009-0225]

Improves RAS characteristics by performing selective logging of DASD Sense data. Adds more comprehensive messages.

View the patch @ https://bugzilla.redhat.com/attachment.cgi?id=313271

RHEL 5.3: System z Update

CPU Node Affinity

[BugZilla XX, IBM LTC 201158, Red Hat Errata 2009-0225]

Improves RAS characteristics by performing selective logging of DASD Sense data. Adds more comprehensive messages.

View the patch @ https://bugzilla.redhat.com/attachment.cgi?id=313271

RHEL 5.3: System z Update

Libica Library: Integration of Icainfo

[BugZilla Bug 439484, IBM LTC 201490, Red Hat Errata http://rhn.redhat.com/errata/RHEA-2009-0064.html]

- icainfo is a part of the SHA & AES enhancements. It shows the customer which CPACF instructions are available in their system.

- libica allows customer applications to speed up cryptographic operations by using the CP Assist for Cryptographic Function (CPACF) facility.

- A new tool called 'icainfo' allows the customer to display a list of all CPACF operations supported by libica. This is helpful to verify that CPACF is correctly enabled on a particular system.

- It makes customer's life easier to gather system informations in a way that there is no need to run through loads of manuals and release notes. The customer only needs to run the tool to see what functions their system supports.

RHEL 5.3: System z Update

OSA 2 Ports per CHPID Support - Installer Enhancements

[Red Hat BugZilla 439461, IBM LTC 201360, IBM BugZilla 43371]

Anaconda now supports both ports on CHPID for OSA Express3 cards. The installer will prompt for the port number in the initial stage of the installation. The value provided for the port also affects installed network interface startup script. When port 1 is selected, the value "portno=1" is added to OPTIONS parameter of ifcfg-eth* file.

Note: When installing under z/VM, you can add either PORTNO=0 (to use port 0) or PORTNO=1 (to use port 1) to the CMS configuration file to avoid being prompted for the mode.

Technical Review: Roadmap

RHEL 5.3 [BugZilla Search Link]

Verified Features

BugZilla	Severity	Priority	Summary
184770	high	high	LTC18425-62140: (big) xDR system Initialization for LPAR Clients
439479	high	high	LTC:5.3:201474:Include gcc 4.3 as Add-On for latest z10 instruction set support
439440	high	high	LTC:5.3:201160:Long Random Numbers Generation
439482	high	high	LTC:5.3:201542:FCP - Enhanced Trace Facility
447379	high	high	LTC:5.3:200994:Linux CPU Node Affinity
439461	high	high	LTC:5.3:201360:OSA 2 Ports per CHPID Support - Installer Enhancements
439484	high	high	LTC:5.3:201490:Libica Library: Integration of Icainfo

Technical Review: Roadmap

- **CPU Affinity** (Red Hat BugZilla 463537)
  - The z10 supports an interface which can be used to get information about the cpu topology of an LPAR. This can be used to optimize the Linux scheduler which bases its decisions on which process gets scheduled to which cpu on the cpu topology. This feature should increase cache hits and therefore overall performance as well. This code has been accepted upstream.
  - **Target: RHEL6**
- **ETR Support** (Red Hat BugZilla 463518)
  - This feature enables Linux images to synchronize with a parallel Sysplex or GDPS. In particular it supports maintaining data consistency groups for the XRC data mover. This code has been accepted upstream.
  - **Target: RHEL6**
- **Link: RHEL5.x Features In Progress**
- **Link: RHEL6 Features In Progress**

Technical Review: System z

- **CPU Affinity** (Red Hat BugZilla 463537)
  - The z10 supports an interface which can be used to get information about the cpu topology of an LPAR. This can be used to optimize the Linux scheduler which bases its decisions on which process gets scheduled to which cpu on the cpu topology. This feature should increase cache hits and therefore overall performance as well. This code has been accepted upstream.
  - **Target: RHEL6**
- **ETR Support** (Red Hat BugZilla 463518)
  - This feature enables Linux images to synchronize with a parallel Sysplex or GDPS. In particular it supports maintaining data consistency groups for the XRC data mover. This code has been accepted upstream.
  - **Target: RHEL6**
- **Link: RHEL5.x Features In Progress**
- **Link: RHEL6 Features In Progress**

Upstream Kernel Development

(stuff we're working on for the future)

Upstream Kernel Development

Generic Kernel 1/4

- **Virtual Memory**
  - Scalability – 1TB ram, 1G page table support (AMD)
  - Scatter list IO support for large page sizes
  - Queued spinlocks – protects large non-numa configs from contention starvation (database stalls)
  - Replicated readonly page cache for NUMA (ie tetx for filesystem backend pages).... very experimental
  - IO throttling – scaling IO device speed to RAM sizes & speed
  - SLUB allocator to scale for large CPU counts
  - Transactional memory – charger member in Velox
- **CFS (completely fair scheduler)**
  - Realtime priority
  - Beneficial for high computer bound, large # of thread
  - Improved network latency
  - Group scheduling – process groups, constrained to cpu sets

Upstream Kernel Development

Generic Kernel 2/4

- **Scalability**
  - Private futexes – avoiding data structure contention (glibc & kernel)
  - Syslets – async syscalls

- **Realtime** – goal of consistency, low-latency determinism (incl in Red Hat MRG product)

- **Storage Enhancements**
  - Seamless SAN/NAS – ease of use / config – make as easy to use as local disks. Enhanced iSCSI config in installer/boot

  - LVM Layering – combinations
    - Striping (raid0) + mirroring (raid1) = raid10
    - Snapshot & mirroring
    - Remote replication – remote copy asynchronous, journaled resync (experimental, feedback welcome)

Upstream Kernel Development

Generic Kernel 3/4

- **Virtualization (distributed)**
  - KVM
  - Paravirt Ops
- **Power Management Work Areas**
  - Tickless kernel – avoid clock tick 1000/sec – allowing true idle
  - Kernel & user space APIs to align timers
  - PowerTOP – useful in identifying “hot” applications
    - Iterative process of cleaning up apps
  - Reworking system startup
    - Only start services / devices as needed
    - Stop idle services

Upstream Kernel Development

Generic Kernel 4/4

- **Ongoing Work Areas**
  - *Security*
    - Hardware drivers, fingerprint readers
    - Runtime tamper checks
    - SHA256 standardized encryption algorithm usage throughout all core services
    - SELinux usability enhancements
    - NFS v4 extended attribute support, allowing SELinux operation

Open Discussion / Q&A