L12

What's new for Linux on System z?

Martin Schwidefsky (schwidefsky@de.ibm.com)

IBM
SYSTEM z9 AND zSERIES EXPO
October 9 - 13, 2006

Orlando, FL

© IBM Corporation 2006

RETURN TO INDEX

Agenda

- Linux on System z development
  - Linux on System z overview
  - Open source development process
- New features
  - Linux kernel
  - Compiler and Toolchain
  - Outlook

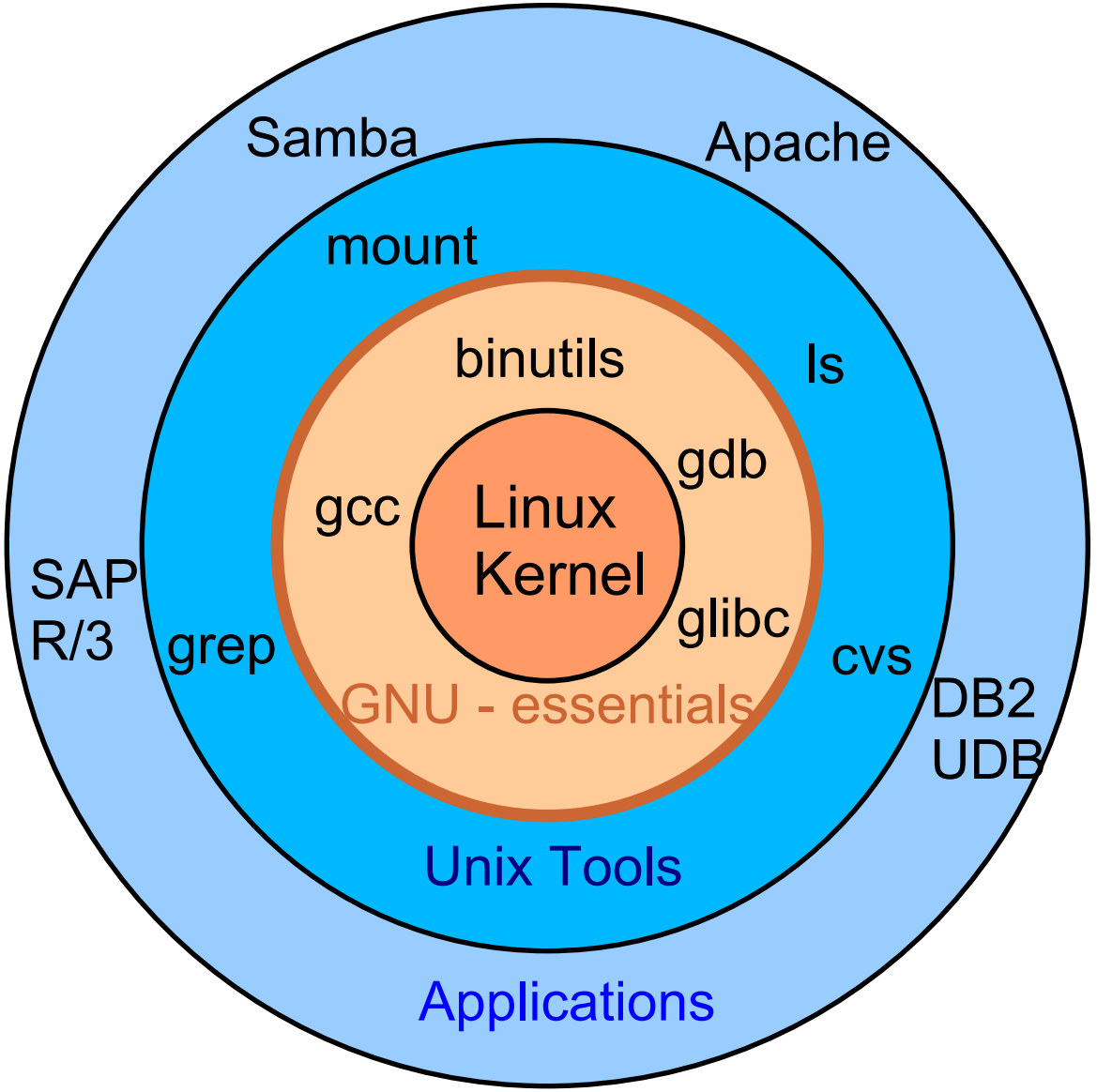
© 2006 IBM Corporation

Linux on System z distributions (Kernel 2.6 based)

- SUSE Linux Enterprise Server 9 (GA 08/2004)
  - Kernel 2.6.5, GCC 3.3.3
  - Service Pack 3 (GA 12/2005)
- SUSE Linux Enterprise Server 10 (GA 07/2006)
  - Kernel 2.6.16, GCC 4.1.0
- Red Hat Enterprise Linux AS 4 (GA 02/2005)
  - Kernel 2.6.9, GCC 3.4.3
  - Update 4 (GA 07/2006)
- Red Hat Enterprise Linux AS 5 (upcoming)
  - Kernel 2.6.x, GCC 4.1.x (t.b.d.)
- Others
  - Debian, Slackware, ...
  - Support may be available by some third party

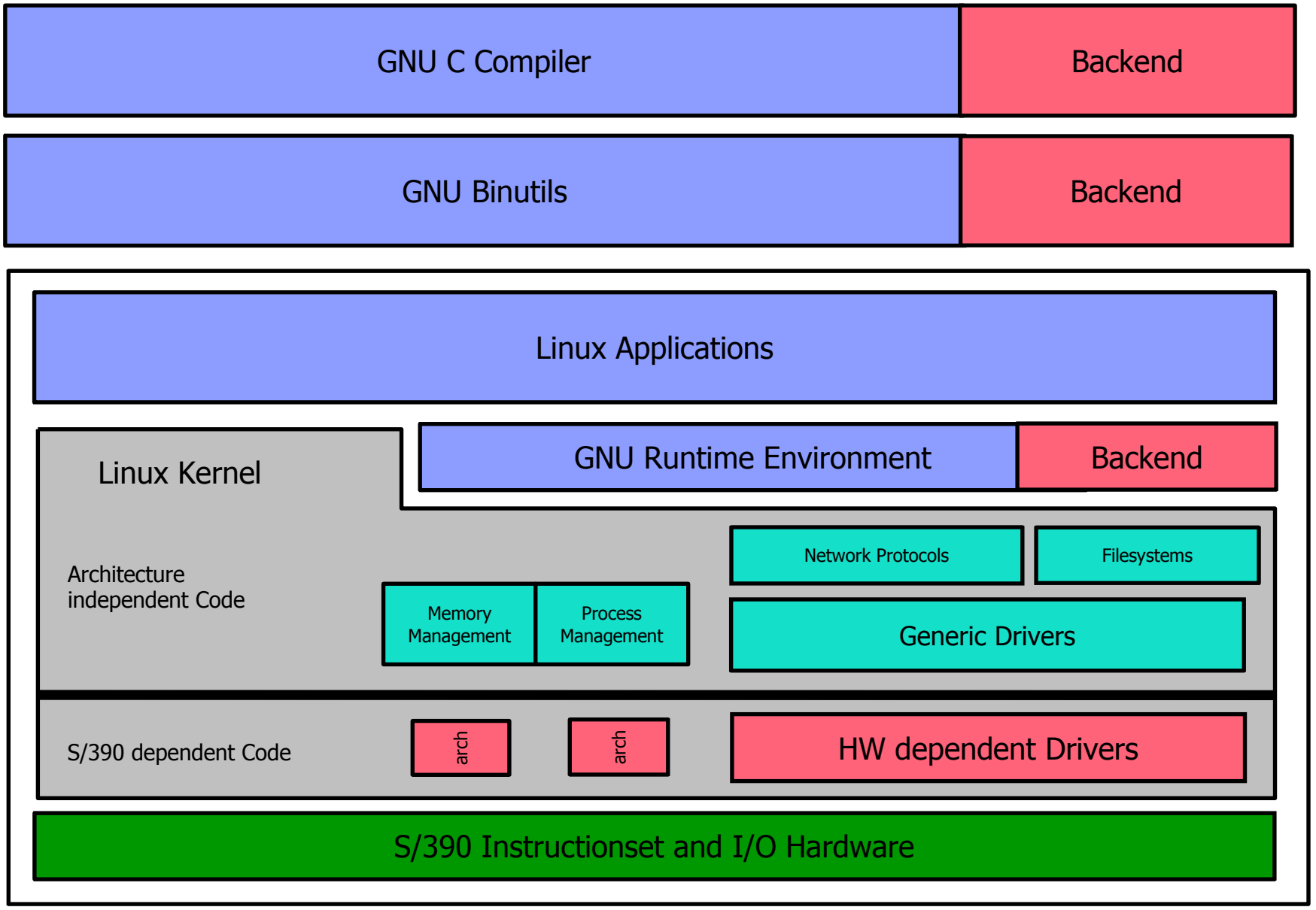
© 2006 IBM Corporation

Linux system components



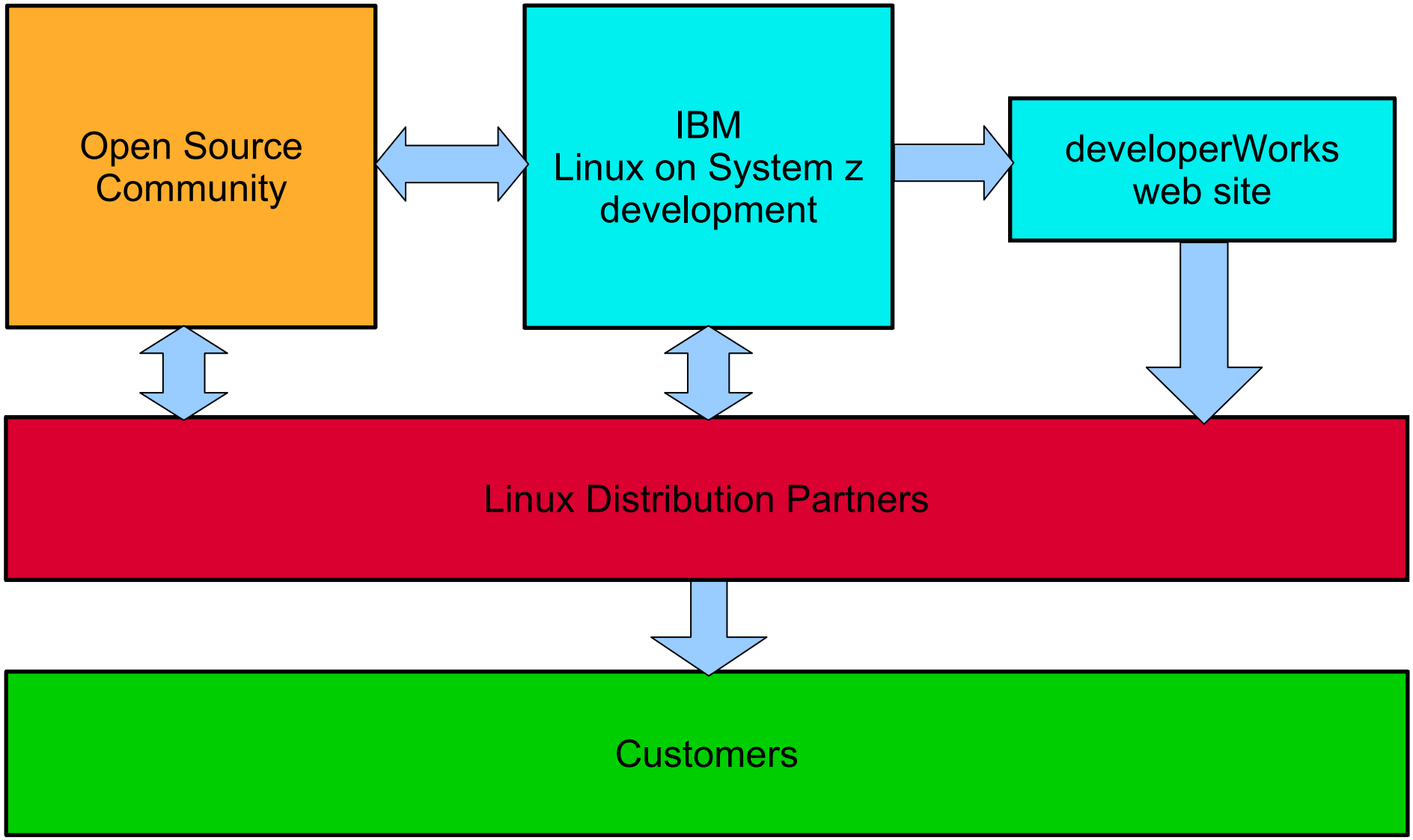
© 2006 IBM Corporation

Linux on System z system structure



© 2006 IBM Corporation

Linux on System z development process



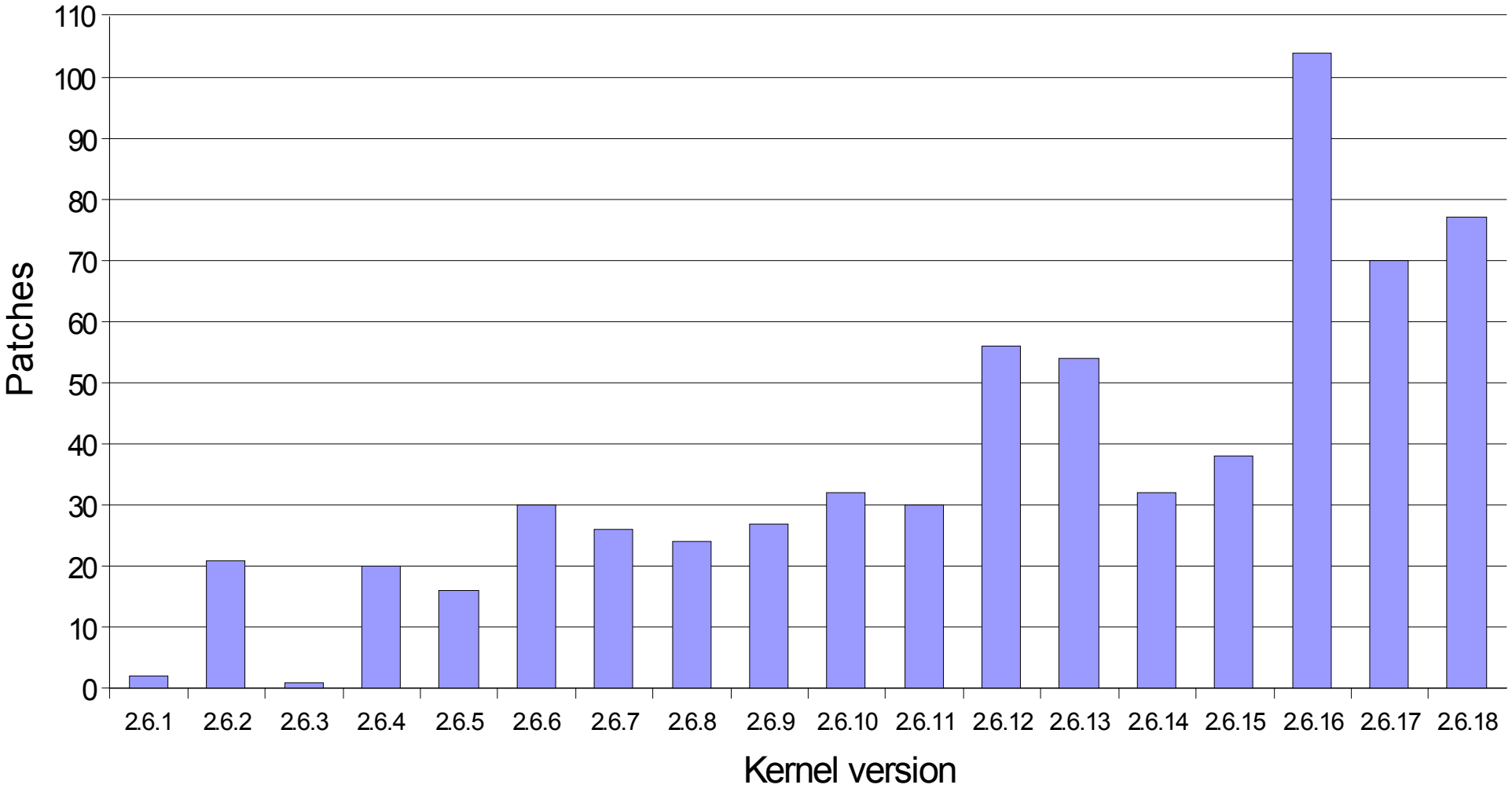
© 2006 IBM Corporation

Open Source development process: Linux Kernel

- Distributed development model
  - Source code control tool: git
  - 'Master' repository maintained by Linus Torvalds
  - 'Experimental' repository maintained by Andrew Morton
  - Secondary repositories maintained by subsystem maintainers and others
  - Flow of code tracked via “Signed-Off” and “Acked-By” statements
- Release process
  - New 2.6.x version released every 2-3 months by Linus
  - First two weeks to merge new features, leading to first release candidate
  - Sequence of multiple release candidates to stabilize
- System z integration
  - Platform subsystem maintainer: Martin Schwidefsky
  - **New:** git repository for System z features hosted on non-IBM site
    - Staging area for IBM and third-party System z patches
    - Experimental System z features

© 2006 IBM Corporation

Linux kernel – System z contributions



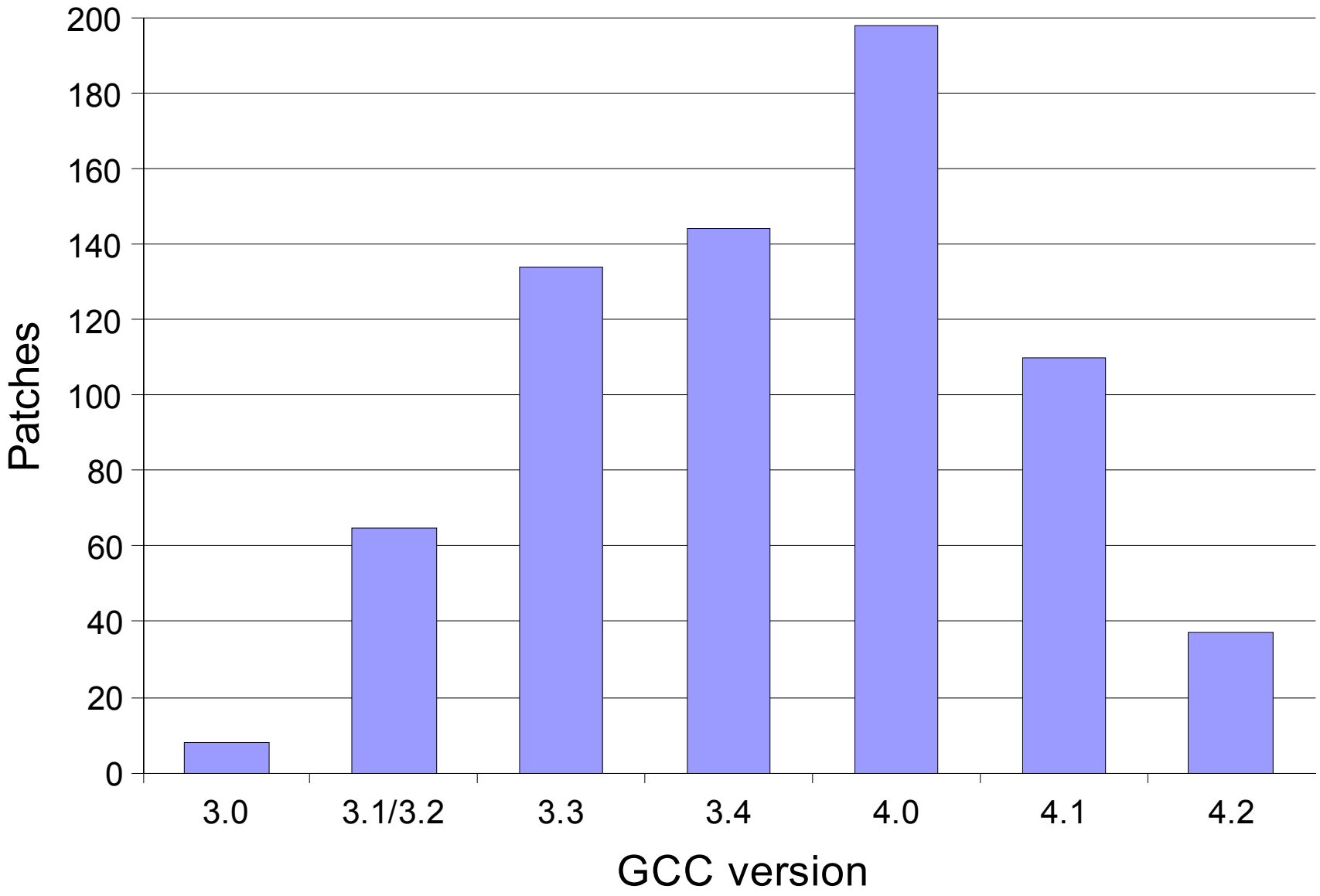
© 2006 IBM Corporation

Open Source development process: GCC

- Centralized development model
  - Source code control tool: subversion
  - Master repository hosted by the Free Software Foundation
    - Read access to the general public, write access to maintainers
    - All copyright owned by / transferred to the FSF
  - GCC Steering Committee oversees the project
  - SC delegates design/development to maintainers
    - Global maintainers (ca. 12), Subsystem maintainers (ca. 130)
- Release process
  - New major release every 8-12 months
  - Development stages: Major changes, minor changes, bugs, regressions
  - “Dot releases” every 2 months containing regression fixes only
- System z integration
  - Platform back-end maintainers: Ulrich Weigand, Hartmut Penner
  - Generally all System z features merged upstream

© 2006 IBM Corporation

GNU Compiler Collection – System z contributions



© 2006 IBM Corporation

How to get new features into distributions ...

- Upstream feature (ideal case)
  - Develop feature against mainline kernel, accepted in kernel version 2.6.x
  - Distribution release based on 2.6.x or later will usually include feature
- Backport of upstream feature (usually acceptable)
  - Code already accepted in some kernel version 2.6.x
  - Develop back-port against previous kernel release, provide on developerWorks and/or to distributor
  - Distribution release/update based on earlier kernel may add the feature as additional patch
- Feature not upstream (difficult)
  - Code provided only on developerWorks and/or to distributor, not yet accepted in any upstream kernel
  - Distributors are generally reluctant to add such features as additional patches due to maintenance concerns

© 2006 IBM Corporation

Object-code only kernel modules

- Issues
  - OCO modules need to be re-built with every kernel change
  - Distributors reluctant to include OCO modules
- Currently, we have no OCO module
  - lcs: open source since 2002-03-04, upstream in 2.4.x
  - z90crypt: open source since 2002-07-31, upstream in 2.4.x
  - qdio: open source since 2002-09-13, upstream in 2.4.x
  - qeth: open source since 2003-06-30, upstream in 2.4.x
  - tape_3590: open source since 2006-03-28, upstream in 2.6.17
- Future strategy: No more OCO modules!

© 2006 IBM Corporation

System z kernel features – Virtualization

- CPU virtualization enhancements
  - CPU hotplug support *(in 2.6.8, DW 1Q05)*
  - Adjust CPU accounting for virtual servers *(in 2.6.11, DW 4Q05)*
  - APPLDATA enhancements (steal time, cpu hotplug) *(in 2.6.18, no DW)*
- DCSS exploitation
  - DCSS mixed memory layout support *(in 2.6.10, DW 1Q05)*
  - DCSS block-device driver enh. (swap to DCSS) *(in 2.6.10, DW 1Q05)*
  - Merge DCSS xip2 file system into ext2 *(in 2.6.13, DW 4Q05)*

© 2006 IBM Corporation

System z kernel features – Virtualization (2)

- Reduction of virtualization overhead
  - QDIO pass-through stage 2 *(in 2.6.16, DW 1Q06)*
  - Collaborative memory management stage 2 *(under discussion, no DW)*
  - z/VM DIAG250 I/O support for 64-bit *(in 2.6.14, DW 1Q06)*
- Usability enhancements
  - z/VM watchdog support *(in 2.6.10, DW 1Q05)*
  - FCP: N-Port-ID Virtualization *(in 2.6.14, DW 4Q05)*
  - Guest LAN sniffer support *(in 2.6.15, DW 1Q06)*

© 2006 IBM Corporation

System z kernel features – Performance

- Scalability enhancements
  - TCP segmentation offload (both HW and SW) *(in 2.6.12, DW 1Q05)*
  - Large number of OSA Express virtual devices *(in 2.6.12, DW 1Q05)*
  - Multiple Subchannel Set support *(in 2.6.16, DW 1Q06)*
  - Linux PAV support for LPAR *(in 2.6.18, no DW)*
- Hardware/kernel performance data collection
  - FCP performance statistics *(under discussion, DW 4Q05)*
  - Channel path measurement data *(in 2.6.17, no DW)*
  - Access to LPAR performance data *(in 2.6.18, no DW)*
- User and kernel space code profiling
  - Oprofile support *(in 2.6.12, DW 1Q05)*
  - Oprofile in-kernel call graph support *(in 2.6.16, DW 1Q06)*

© 2006 IBM Corporation

System z kernel features – Operational Simplification

- Communication Controller support
  - Linux NCP CDLC support via OSA *(in 2.6.15, DW 4Q05)*
  - OSA Layer 2 sequence numbers *(in 2.6.14, DW 4Q05)*
- FCP enhancements
  - Point-to-point support *(in 2.6.12, DW 1Q05)*
  - Re-IPL from SCSI *(in 2.6.14, DW 4Q05)*
  - Export SCSI IPL parameter list *(in 2.6.15, DW 4Q05)*
  - SAN discovery tool *(DW 4Q05)*
- z/VM integration
  - User space access to CP commands *(in 2.6.13, DW 4Q05)*
  - Support 64-bit VMDUMP format *(DW 4Q05)*

© 2006 IBM Corporation

System z kernel features – RAS

- Kernel
  - Enhanced kernel machine-check handling *(in 2.6.13, DW 4Q05)*
- DASD
  - Write barrier support *(in 2.6.12, DW 4Q05)*
  - Fast fail support *(in 2.6.16, DW 1Q06)*
  - Enhanced error reporting *(in 2.6.17, no DW)*
- FCP
  - Best effort SAN notifications *(in 2.6.16, DW 1Q06)*

© 2006 IBM Corporation

System z kernel features – Security

- New hardware support – crypto cards
  - Crypto Express 2 Accelerator *(in 2.6.16, DW 4Q05)*
- New hardware support – z9 processor
  - Support user-space AES+SHA+PRNG crypto CP Assists
  - Support in-kernel AES+SHA crypto CP Assists *(in 2.6.16, DW 1Q06)*
- Functional enhancements
  - Secure Key cryptography *(queued for 2.6.19, no DW)*

© 2006 IBM Corporation

Some common kernel features

- Scalability
  - Per page-table locks
  - RCU enhancements
  - 4-level page tables
- Configuration
  - Enhanced hotplug / udev infrastructure
  - CPU hotplug
  - Memory hotplug *(future)*
- New features
  - OCFS2 cluster file system
  - New POSIX system calls: message queues
  - New Linux-specific system calls: splice/tee/vmsplice
  - Futex enhancements (robustness, priority inheritance)

© 2006 IBM Corporation

Compiler – Common features

- General optimizer improvements
  - SSA-based common optimization infrastructure (GCC 4.0)
  - Inter-procedural optimization infrastructure (GCC 4.1)
- Languages and language features
  - Fortran 95 front end (GCC 4.0)
  - Decimal Floating Point support (GCC 4.2)
- Other improvements
  - Stack Protector feature (GCC 4.1)
  - Builtins for atomic operations (GCC 4.1)

© 2006 IBM Corporation

Compiler – System z features

- System z9 109 processor support (GCC 4.1)
  - Exploit instructions provided by the e*xtended immediate facility*
  - Selected via `-march=z9-109` / `-mtune=z9-109`
- Support for 128-bit IEEE quad “long double” data type (GCC 4.1)
  - Provide extended range of floating point exponent and mantissa
  - Selected via `-mlong-double-128`
- Kernel stack overflow avoidance/detection (GCC 4.0)
  - Compile time detection: `-mwarn-framesize` / `-mwarn-dynamicstack`
  - Run-time detection: `-mstack-size` / `-mstack-guard`
  - Stack frame size reduction: `-mpacked-stack`
- GCC support for the z/TPF OS (GCC 4.0/4.1)
  - z/TPF uses Linux / GCC as cross-build environment
  - New target `s390x-ibm-tpf`

© 2006 IBM Corporation

Compiler – System z performance

- Compiler back-end improvements
  - Improved condition code handling (GCC 4.0)
  - Improved function prologue/epilogue scheduling (GCC 4.0)
  - Improved use of memory-to-memory instructions (GCC 4.0)
  - Added sibling call support (GCC 4.0)
  - Enhanced use of string instructions (SRST, MVST, ...) (GCC 4.1)
  - More precise register tracking (r13, r6, ...) (GCC 4.1)
  - Use LOAD ZERO (GCC 4.1)
  - ICM/STCM, BRCT, vararg enhancements (GCC 4.1)
- Overall performance enhancement 8%
  - Industry-standard integer performance benchmark
  - Comparing GCC 3.4 and GCC 4.1 on System z

© 2006 IBM Corporation

Outlook

- New hardware exploitation
- Enhanced Linux – z/VM synergy
- Enhanced integration with z/OS
- Keep current with open source

© 2006 IBM Corporation

Questions?

© 2006 IBM Corporation

Trademarks

- Linux is a registered trademark of Linus Torvalds
- IBM, zSeries, S/390, z/Architecture, System z9, z/VM, PowerPC are trademarks of International Business Machines in the United-States, other countries or both
- Red Hat™, Inc., Red Hat® Enterprise Linux® and Red Hat® Linux® are trademarks of Red Hat, Inc. in the United States, other countries, or both.
- SUSE® is a trademark of SUSE AG, a Novell business, in the United States, other countries, or both.
- XEN is a trademark of XenSource, Inc.
- VMware® is a registered trademark of VMware, Inc. in the United States, other countries, or both.
- Other company, product or service names may be trademarks or service marks of others