Real Time BoF ELC 2012

This session provides a forum to discuss Real Time Linux, share how you are using it, and learn from the experiences of others.

Please come prepared to discuss your experiences with and hear about:

- fields of use (for example: stock market, audio, robotics, industrial automation, consumer products)
- what works well
- what needs to improve
- what is missing and needs to be implemented
- experiences using real time, both successes and failures
- where the future Preempt RT patchset is heading
- status of mainlining, what is left to mainline
- possible future areas of development such as scheduler improvements, a deadline scheduler, resource allocation
- benchmarking tools and techniques
- performance measurement tools and techniques
- any other topic of interest to the attendees

Frank Rowand, Sony Network Entertainment

February 16, 2012

Birds of a Feather (computing)

In computing, BoF (Birds of a Feather) can refer to:

> A BoF session, an informal meet-up at conferences, where the attendees group together based on a shared interest and carry out discussions without any pre-planned agenda.

http://en.wikipedia.org/wiki/Birds_of_a_Feather_%28computing%29

Session Format

I would like for this to be an interactive session.

I would prefer to listen to other people talk.

But if no one else wants to talk, I can.....

What to talk about

PREEMPT_RT patch size

Current PREEMPT_RT Versions

Stable Tree

What events are coming up?

Lightning talks

Who are we? How are we using RT Linux?

Various RT Features

How can you contribute?

Contents of RT patches

******** What Else?   ********

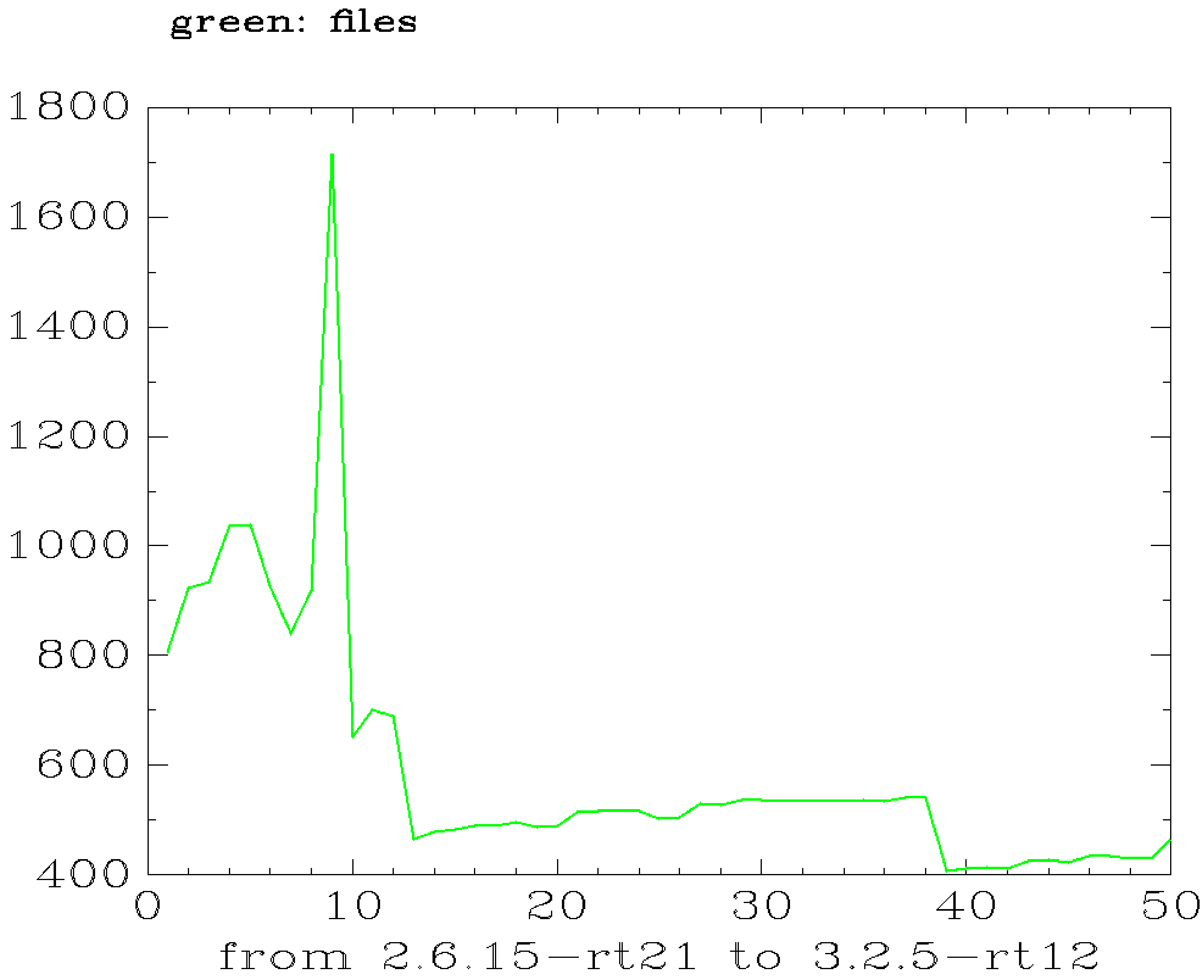
PREEMPT_RT Patch Size

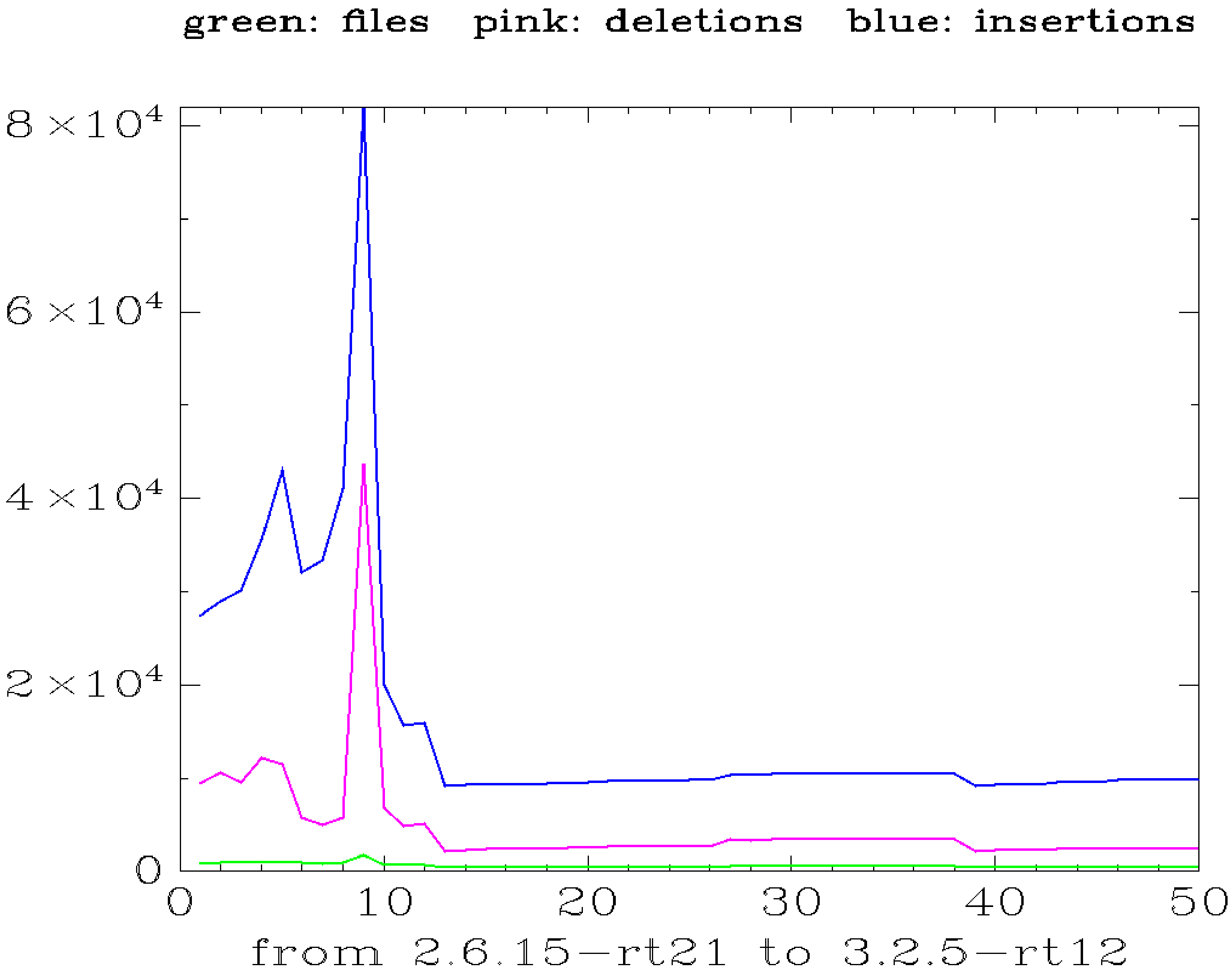
The recent trajectory

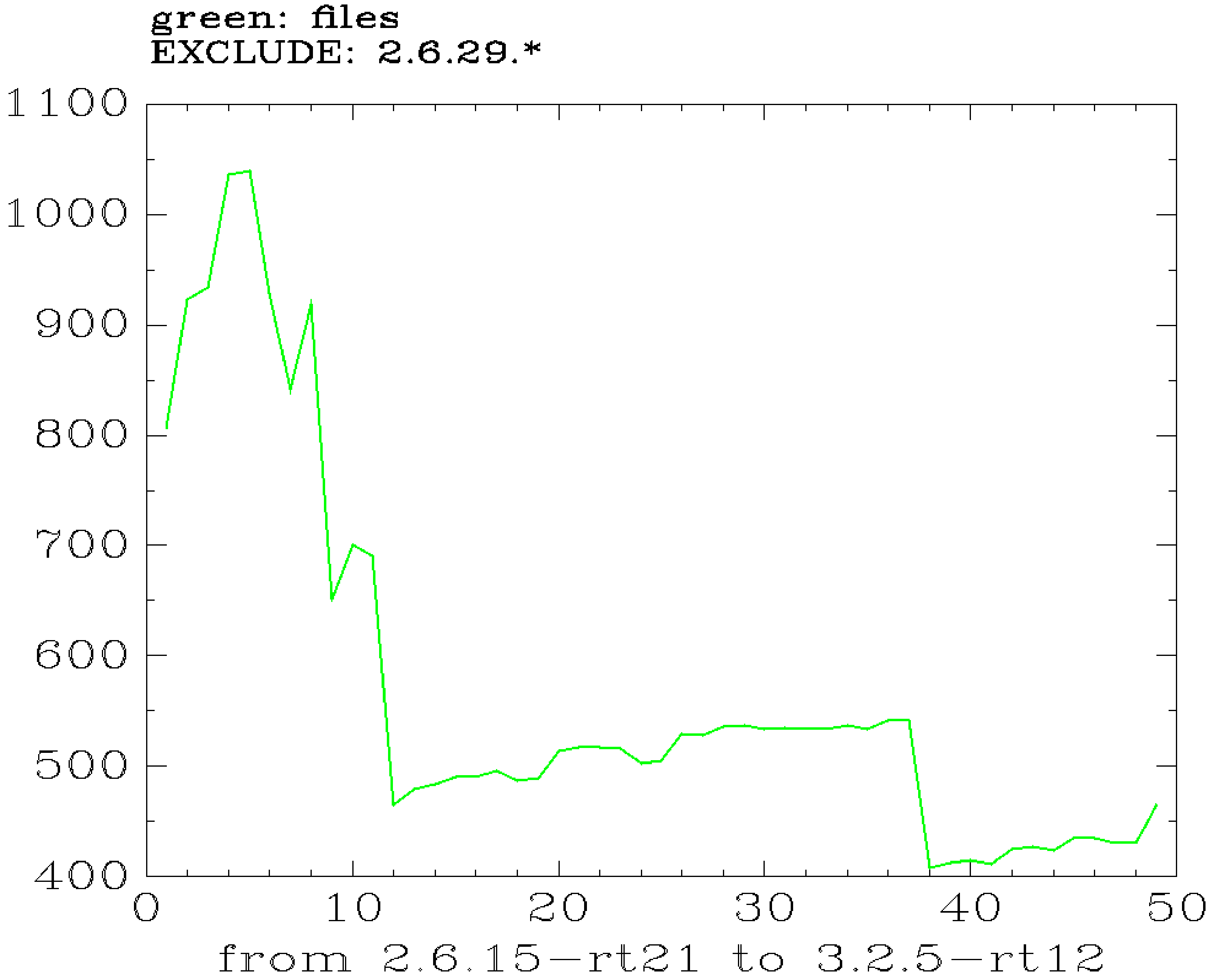
- Number of files modified
- Number of lines added
- Number of lines deleted

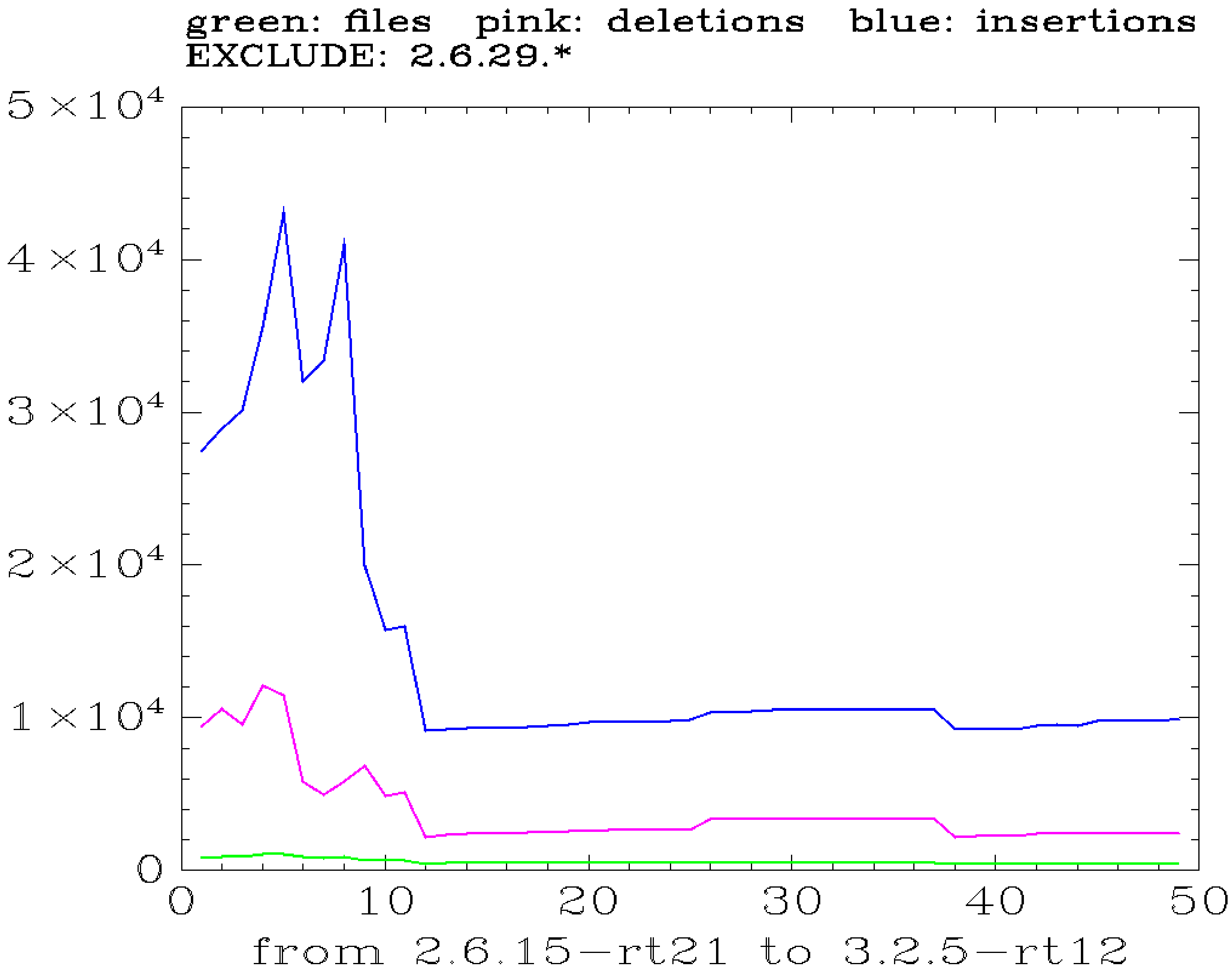
For example:

```
    2.6.15-rt21                    806       27449    9372
```


green: files
1800
1600
1400
1200
1000
800
600
400
0
10
20
30
40
50
from 2.6.15-rt21 to 3.2.5-rt12


green: files pink: deletions blue: insertions
8×10^4
6×10^4
4×10^4
2×10^4
0
0
10
20
30
40
50
from 2.6.15−rt21 to 3.2.5−rt12


green: files
EXCLUDE: 2.6.29.*
1100
1000
900
800
700
600
500
400
0
10
20
30
40
50
from 2.6.15-rt21 to 3.2.5-rt12


green: files pink: deletions blue: insertions
EXCLUDE: 2.6.29.*
5 ×10⁴
4 ×10⁴
3 ×10⁴
2 ×10⁴
1 ×10⁴
0
0
10
20
30
40
50
from 2.6.15−rt21 to 3.2.5−rt12

PREEMPT_RT Patch Size

2.6.15-rt21	806	27449	9372	
2.6.22.1-rt9	923	28964	10597	
2.6.23-rt1	934	30141	9591	
2.6.23.11-rt14	1036	35565	12107	
2.6.24.7-rt27	1039	43057	11481	peak
2.6.25.8-rt7	928	32060	5848	
2.6.26-rt1	842	33384	4985	
2.6.26.8-rt16	918	41005	5852	
2.6.29.6-rt24	1715	81837	43653	peak
2.6.31.6-rt19	652	19944	6893	
2.6.33.7.2-rt30	701	15723	4870	
2.6.33.9-rt31	690	15994	5123	

PREEMPT_RT Patch Size

2.6.33.7.2-rt30	701	15723	4870	
3.0-rt1	465	9183	2211	drop
3.0-rt2	479	9255	2325	
3.0-rt3	484	9312	2414	
3.0-rt4	490	9356	2432	
3.0-rt5	490	9357	2433	
3.0-rt6	496	9392	2467	
3.0-rt7	487	9468	2508	
3.0.1-rt8	489	9596	2542	
3.0.1-rt9	514	9712	2609	
3.0.1-rt10	517	9746	2649	
3.0.1-rt11	517	9751	2649	
3.0.3-rt12	516	9739	2640	
3.0.4-rt13	503	9831	2674	
3.0.4-rt14	505	9875	2690	

PREEMPT_RT Patch Size

3.0-rt1	465	9183	2211	drop
3.0.4-rt14	505	9875	2690	
3.0.4-rt15	529	10393	3395	rise
3.0.6-rt16	528	10363	3375	
3.0.6-rt17	536	10471	3390	
3.0.6-rt18	537	10510	3405	
3.0.7-rt19	534	10478	3394	
3.0.7-rt20	535	10509	3406	
3.0.8-rt21	534	10505	3403	
3.0.8-rt22	534	10499	3388	
3.0.8-rt23	537	10524	3393	
3.0.9-rt24	534	10506	3389	
3.0.9-rt25	542	10519	3401	

PREEMPT_RT Patch Size

3.0-rt1	465	9183	2211
3.0.9-rt25	542	10519	3401
3.2-rc1-52e4c2a05-rt1	408	9225	2212
3.2-rc1-52e4c2a05-rt2	413	9244	2225
3.2-rc2-rt3	415	9280	2264
3.2-rc4-rt4	412	9277	2253
3.2-rc4-rt5	425	9523	2413
3.2-rc4-rt5	425	9523	2413
3.2-rc4-rt6	427	9535	2425
3.2-rc5-rt7	424	9520	2423
3.2-rc5-rt8	435	9766	2441
3.2-rc6-rt9	435	9766	2441
3.2-rt10	431	9764	2441
3.2.5-rt11	431	9764	2441
3.2.5-rt12	466	9955	2469

Contents of 3.2.5 RT Patches

# UPSTREAM changes queued for 3.2 or in 3.1
2 patches

# Stuff broken upstream, patches submitted
1 patch

# Stuff which needs addressing upstream, but requires
# more information
1 patch

# Submitted on LKML, mips ML, ARM ML, ppc-devel
8 patches

Contents of 3.2.5 RT Patches

# Pending in peterz's scheduler queue
2 patches

# Stuff which should go upstream ASAP
49 patches

# Stuff which should go mainline, but wants some care
4 patches

# REAL RT STUFF starts here
182 patches

Current PREEMPT_RT Versions

OSADL latest stable: 2.6.33.7.2-rt30

- https://www.osadl.org/Latest-Stable-Realtime.latest-stable-realtime-linux.0.html
- Paul Gortmaker created a broken out version
  - http://marc.info/?l=linux-rt-users&m=129588844818236&w=2

2.6.34.8 from Paul (_not_ stable, best effort)

- https://lkml.org/lkml/2011/3/4/281

3.0.20-rt36 stable, from Steve Rostedt

3.2.5-rt12 top of tree from Thomas Gleixner

Stable Tree

Maintained by Steve Rostedt

Same type of policy as the Greg KH stable trees

- Fixes only (no new features)
- Patch must first be in one of:
    - more recent tglx RT_PREEMPT patch
    - Linux mainline
    - Linux stable

To nominate a patch for stable:

Cc: stable-rt@vger.kernel.org

Stable Tree

Typical release cycle is a multi-stage process

1 Move existing RT patches to a new
Greg KH stable version [release]

2.a Add additional RT fix patches [-rc]

2.b Add additional RT fix patches [release]

All announced on lkml and linux-rt-users

Stable Tree

Patches can be dowloaded from

ftp://www.kernel.org/pub/linux/kernel/projects/rt/3.0/

Split out patches (quilt model) through 3.0.14-rt31.
Unified patch for all versions.

Maintained in a git repository

- Most recent description:

  https://lkml.org/lkml/2012/1/24/222

Stable Tree git Repository

git://git.kernel.org/pub/scm/linux/kernel/git/rt/linux-stable-rt.git

v3.0-rt branch

- will never rebase
- should be what users develop git repositories with

Stable Tree git Repository

v3.0-rt-rebase branch

- will rebase at every -rt version
- should ***not*** be used by other developmental git repositories.
- will allow pulling out the commits that will apply to the stable tree

At each version, the rebase branch will be tagged with a "-rebase" name after it.

Stable Tree git Repository Tags

v3.0.10-rt27
v3.0.10-rt27-rebase
v3.0.11-rt28
v3.0.11-rt28-rebase
v3.0.12-rt29
v3.0.12-rt29-rebase
v3.0.12-rt30
v3.0.12-rt30-rebase
v3.0.14-rt31
v3.0.14-rt31-rebase
v3.0.14-rt32
v3.0.14-rt32-rebase
v3.0.17-rt33
v3.0.17-rt33-rebase

v3.0.18-rt34

v3.0.20-rt35
v3.0.20-rt35-rebase
v3.0.20-rt36
v3.0.20-rt36-rebase

Stable Tree git Repository

“IMPORTANT NOTE: The rebase branch and tags are very low priority.

If it becomes too time consuming to maintain, I ***will*** stop adding them.

They will be created when I have time to do so.

This rebase branch is a convenience for people that may want it.

It may be discontinued at any time without notice!”

Stable Tree 3.0.x History

DELTA: patches, files, insertions, deletions

3.0.9-rt26	6	6	19	16
3.0.10-rt27	-			
3.0.11-rt28	-			
3.0.12-rt29	-			
3.0.12-rt30	12	21	302	180
3.0.14-rt31	-			
3.0.14-rt32	4	7	227	8
3.0.17-rt33	-			
3.0.18-rt34	-			
3.0.20-rt35	-			
3.0.20-rt36	8	17	191	29

What events are coming up?

14th real time linux workshop

October 18 - 20, 2012
University of North Carolina at Chapel Hill

Linux Kernel Summit, LinuxCon North America, Linux Plumbers Conference

August 27 - 29, August 29 - 31

LinuxCon Europe /
Embedded Linux Conference Europe

November 5 - 7

Lightning Talks

Random current RT topics of interest...

Who are we?

How are we using RT Linux?

How would we like to use RT?

What is missing or needs to improve?

What works well already?

Various RT Features

Deadline Scheduler

Juri Lelli has taken over from Dario:

> Is there any news on the possible integration of the EDF scheduling
> class (SCHED_DEADLINE) in the RT branch?

Cannot speak about "the integration" :-P, but I can tell you that I'm currently taking over Dario's work (he has a new job now, and I started working at his previous lab), even if he remains behind the scenes.

We focused on the RT branch (3.2-rc1-rtx) for the next release and I'm quite close to post it (hope before the end of November).

http://marc.info/?l=linux-rt-users&m=132135121824549&w=2

CPU Isolation

[git pull] CPU isolation extensions
From: Max Krasnyansky
https://lkml.org/lkml/2008/2/7/1

Linus, please pull CPU isolation extensions from ...

The patchset consist of 4 patches.

- Make cpu isolation configurable and export isolated map
- Do not route IRQs to the CPUs isolated at boot
- Do not schedule workqueues on the isolated CPUs
- Do not halt isolated CPUs with Stop Machine

CPU Isolation - In Progress

[PATCH v7 0/8] Reduce cross CPU IPI interference

From: Gilad Ben-Yossef

https://lkml.org/lkml/2012/1/8/109

https://lkml.org/lkml/2012/1/26/70

CPU Isolation - In Progress

From: Frederic Weisbecker

[RFC PATCH 00/32] Nohz cpusets (was: Nohz Tasks)

https://lkml.org/lkml/2011/8/15/245

Status of Nohz cpusets (adaptive tickless kernel) for January 2012

https://lkml.org/lkml/2012/1/17/486

May include tick avoidance when there is a single userland process running.

https://tglx.de/~fweisbec/TODO-nohz-cpusets

CPU Isolation

One further suggestion for moving solutions forward:

- create a partition using cpusets
- run a while(1); userspace in it, and trace the thing
- Work up patches that remove all useless perbutations

Real-Time virtualisation

???

Real-Time Linux Failure
Frank Rowand
From ELC 2010

Using KVM as a Real-Time Hypervisor
Jan Kiszka
From KVM Forum 2011

Xenomai / ipipe vs. PREEMPT_RT

???

Tools:
What has changed since ELC 2008 BoF?

- ftrace, KernelShark
- perf
- cyclictest enhancements
- lpptest no longer in RT patch set
- RealFeel-ETRI
  - Jung: ELC 2010 Measuring Responsiveness ...
- Woerner test (serial port interrupt latency test)
  - web site gone
- Senoner test (audio)

What has not changed?

LRTB
Hourglass

Other Topics?

How can you contribute?

- OSADL.org
  - ???
- mainline RT_PREEMPT
  - ???
- test new RT releases
  - !!!
- rt.wiki.kernel.org
- elinux.org
  - http://elinux.org/Real_Time
  - http://elinux.org/Realtime_Testing_Best_Practices
- linux-rt-users
  - http://vger.kernel.org/vger-lists.html#linux-rt-users

Trojan Horses

Re: linux-next: add utrace tree
From: Linus Torvalds
Date: Mon Jan 25 2010 - 13:14:03 EST

And yeah, I still think the hard-RT people are mostly crazy.

So I can work with crazy people, that's not the problem. They just need to _sell_ their crazy stuff to me using non-crazy arguments, and in small and well-defined pieces. When I ask for killer features, I want them to lull me into a safe and cozy world where the stuff they are pushing is actually useful to mainline people _first_.

In other words, every new crazy feature should be hidden in a nice solid "Trojan Horse" gift: something that looks _obviously_ good at first sight.

Gleixner ELC 2011 Slides

Resources

Status of Preempt-RT and why there is no roadmap
Thomas Gleixner, ELC 2011
http://elinux.org/images/c/ca/Elc2011_gleixner.pdf
https://lwn.net/Articles/440064/

Trojan Horses
http://lkml.indiana.edu/hypermail/linux/kernel/1001.3/00384.html
http://lkml.indiana.edu/hypermail/linux/kernel/1001.3/00389.html

14th Real Time Linux Workshop 2012
https://www.osadl.org/RTLWS-2012.rtlws-2012.0.html

Resources

Using KSM as a Real-Time Hypervisor
Jan Kiszka, KVM Forum 2011
    http://www.linux-kvm.org/wiki/images/0/03/KVM-Forum-2011-RT-KVM.pdf

Real Time Wiki
    https://rt.wiki.kernel.org

Embedded Linux Wiki
    http://elinux.org

Open Source Automation Development Lab
    http://www.osadl.org/

How to get a copy of the slides

1) leave a business card with me

2) frank.rowand@am.sony.com