

# 0-day Kernel Performance Test

Yuanhan Liu

<yuanhan.liu@linux.intel.com>

October 19, 2013



# 0-day kernel performance testing

- test latest kernel
- analysis results
- bisect

# example report

Date: Fri, 27 Sep 2013 00:33:00 +0800

From: Yuanhan Liu <yuanhan.liu@linux.intel.com>

To: Yuanhan Liu <yuanhan.liu@linux.intel.com>, Fengguang Wu <fengguang.wu@intel.com>

Subject: [confirmed] perf bisect for ... 7210acd61435c5f3f1387031b150384dc975978c pigz.throughput

f27dde8deef33c9e58027df11ceab2198601d6a6 is the first bad commit

commit f27dde8deef33c9e58027df11ceab2198601d6a6

Author: Peter Zijlstra <peterz@infradead.org>

CommitDate: Wed Sep 25 14:07:49 2013 +0200

sched: Add NEED\_RESCHED to the preempt\_count

In order to combine the preemption and need\_resched test we need to fold the need\_resched information into the preempt\_count value.

....

pigz.throughput



# linux kernel development - git tree

- **mainline(Linus's tree)**

`git://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git`



- **maintainer's tree**

`git://git.kernel.org/pub/scm/linux/kernel/git/viro/vfs.git`

`git://git.kernel.org/pub/scm/linux/kernel/git/tytso/ext4.git`

...



- **developer's tree**

`git://git.kernel.org/pub/scm/linux/kernel/git/kees/linux.git`

`git://github.com/congwang/linux.git`

...

# linux kernel development - release cycle

v3.11-rc1 (one **rc release** per week)

v3.11-rc2

v3.11-rc3

v3.11-rc4

v3.11-rc5

v3.11-rc6

v3.11-rc7

v3.11

|

|

**merge window** (2 weeks)

V

v3.12-rc1

# merge gap

merge date may be **months after** commit date

## commit date

```
commit 4e7ea81db53465ddd753678bc4cebf95369d0984
```

```
Author: Jan Kara <jack@suse.cz>
```

```
AuthorDate: Tue Jun 4 13:17:40 2013 -0400
```

```
Commit: Theodore Ts'o <tytso@mit.edu>
```

```
CommitDate: Tue Jun 4 13:17:40 2013 -0400
```

```
ext4: restructure writeback path
```

There are two issues with current writeback path in ext4. For one we don't necessarily map complete pages when blocksize < pagesize and ...

## merge date

```
tag v3.11-rc1
```

```
Tagger: Linus Torvalds <torvalds@linux-foundation.org>
```

```
Date: Sun Jul 14 15:18:41 2013 -0700
```

# the old way

test each rc release

## issues

- not in time, especially for rc1
- test machines are mostly idle

# the way we want

hunt down and report change **in time** => **0day kernel perf testing**

**Date:** Fri, **27 Sep** 2013 00:33:00 +0800

**From:** Yuanhan Liu <yuanhan.liu@linux.intel.com>

**To:** Yuanhan Liu <yuanhan.liu@linux.intel.com>, Fengguang Wu <fengguang.wu@intel.com>

**Subject:** [confirmed] perf bisect for ... 7210acd61435c5f3f1387031b150384dc975978c pigz.throughput

f27dde8deef33c9e58027df11ceab2198601d6a6 is the first bad commit

commit f27dde8deef33c9e58027df11ceab2198601d6a6

Author: Peter Zijlstra <peterz@infradead.org>

**CommitDate:** Wed **Sep 25** 14:07:49 2013 +0200

sched: Add NEED\_RESCHED to the preempt\_count

....

## 2 key challenges

- test latest code
- bisect - find the first bad commit(for report)



# test latest code: challenge

- **can't test mainline only**

- too late
- can't rebase

• **300+ git tree**      =>      **can't test one by one**

# test latest code: solution

## merge hourly

v3.12-rc5

|----- devel-hourly-2013101501

|

|----- devel-hourly-2013101502

|

|----- devel-hourly-2013101503

|

|----- devel-hourly-2013101504

.

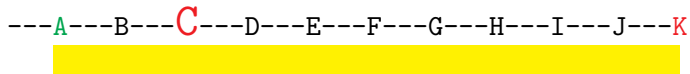
.

.

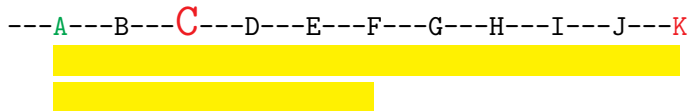
# bisect concept

---A---B---C---D---E---F---G---H---I---J---K

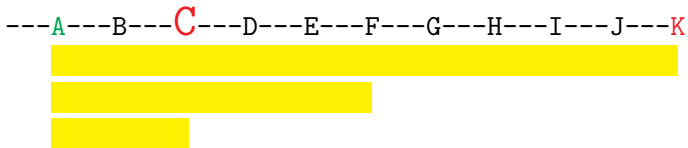
# bisect concept



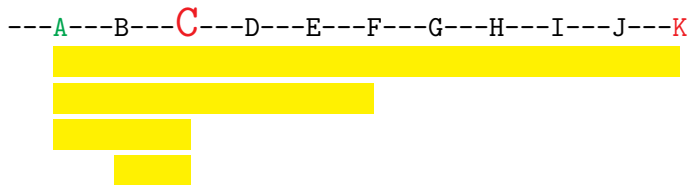
# bisect concept



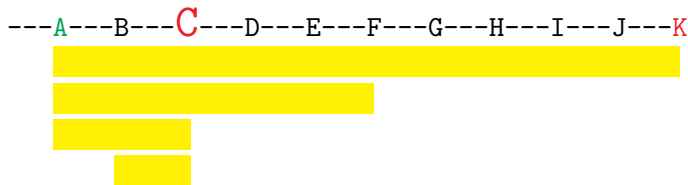
# bisect concept



# bisect concept



# bisect concept



**C** is the first bad commit

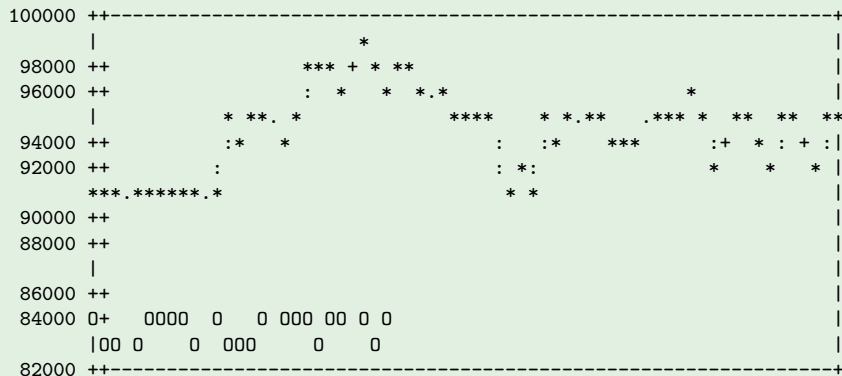


# what we track - I

## benchmark's result

0 -- hourly merged branch HEAD  
\* -- hourly merged branch BASE

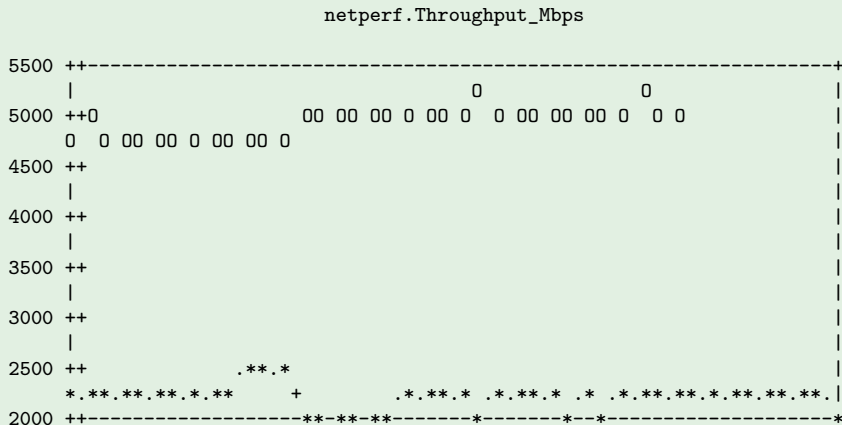
aim7.2000.jobs-per-min



## what we track - I

```
0 -- houlry merged branch HEAD
```

```
* -- hourly merged branch BASE
```



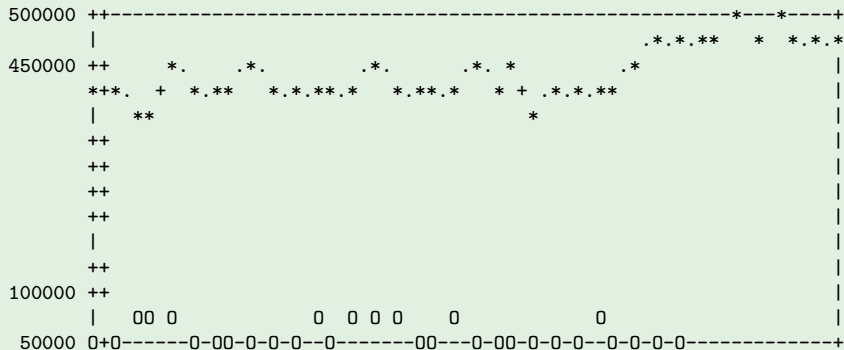
# what we track - II

## monitor's output

0 -- hourly merged branch HEAD

\* -- hourly merged branch BASE

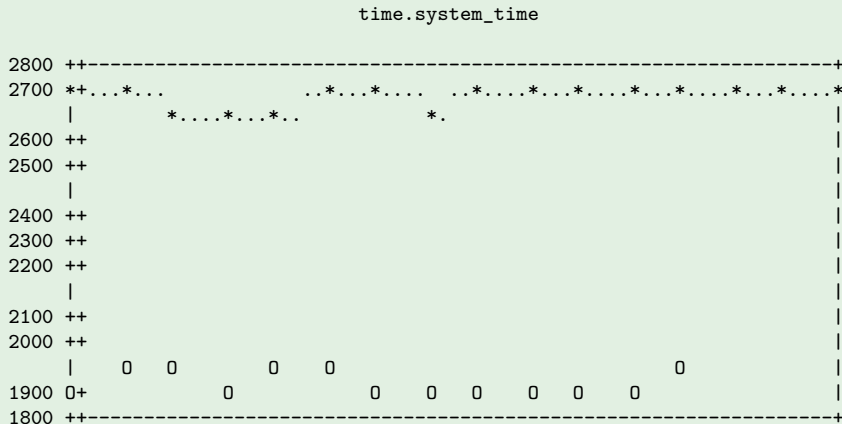
lock\_stat.slock-AF\_INET/1.contentions.tcp\_v4\_rcv



# what we track - II

0 -- hourly merged branch HEAD

\* -- hourly merged branch BASE



# bisect challenge

In a typical day:

- **5,000+** bisectable changes
- we can only bisect about **10** of them

# bisect challenge

In a typical day:

- **5,000+** bisectable changes
- we can only bisect about **10** of them

Question: How to choose **10** among **5,000** changes?

# bisect challenge

In a typical day:

- **5,000+** bisectable changes
- we can only bisect about **10** of them

Question: How to choose **10** among **5,000** changes?

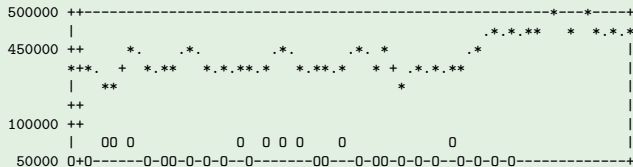
## 5 factors

- 1 change percent
- 2 variation
- 3 run time
- 4 stats field weight
- 5 timeliness

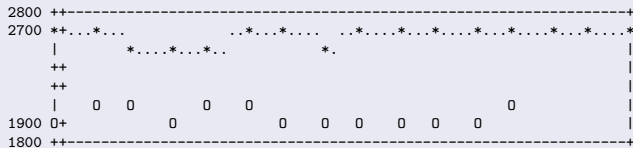
# factor 1: change percent

```
0 -- hourly merged branch HEAD
* -- hourly merged branch BASE
```

lock\_stat.slock-AF\_INET/1.contentions.tcp\_v4\_rcv (900%)



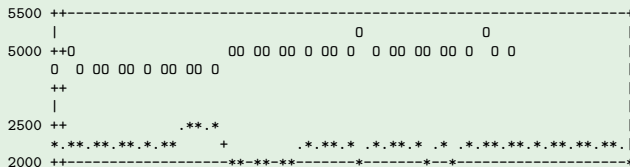
time.system\_time (42%)



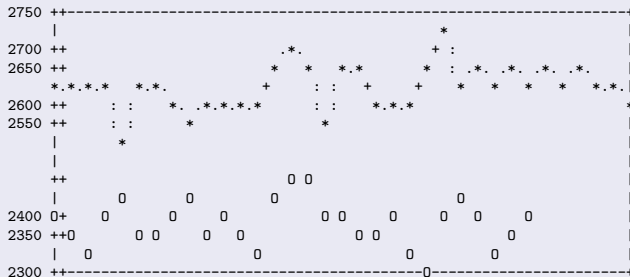


## factor 2: variation

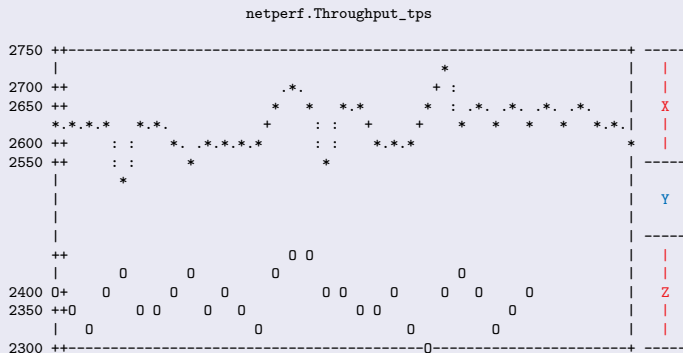
netperf.Throughput\_Mbps



netperf.Throughput\_tps



## factor 2: variation



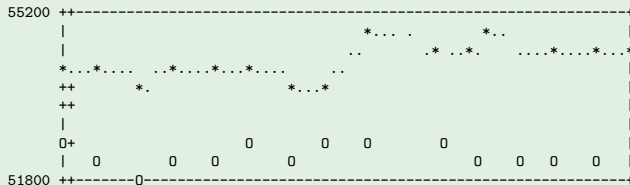
## factor 3: run time

**less run time, faster for bisection**

## factor 4: stats field weight

some stats field(e.g aim7) are more *important*

aim7.2000.jobs-per-min (6%)



lock\_stat.slock-AF\_INET/1.contentions.tcp\_v4\_rcv (900%)



## factor 5: timeliness

**today's changed stats > few days ago's changed stats**

# score based bisect priority calculation

change percent	0 - 20
variation	0 - 10
run time	0 - 10
stats field weight	0 - 30, or more
timeliness	0 - 10

## add them up

```
bisect_priority = change_perecnet_score +  
                  variation_score      +  
                  run_time_score       +  
                  stats_field_score    +  
                  timeliness_score
```

# Authors

Fengguang Wu -- Leader & designer

Yuannhan Liu

Alex Shi

Zhao Lei

Fan Chaoting

# Thank you!

