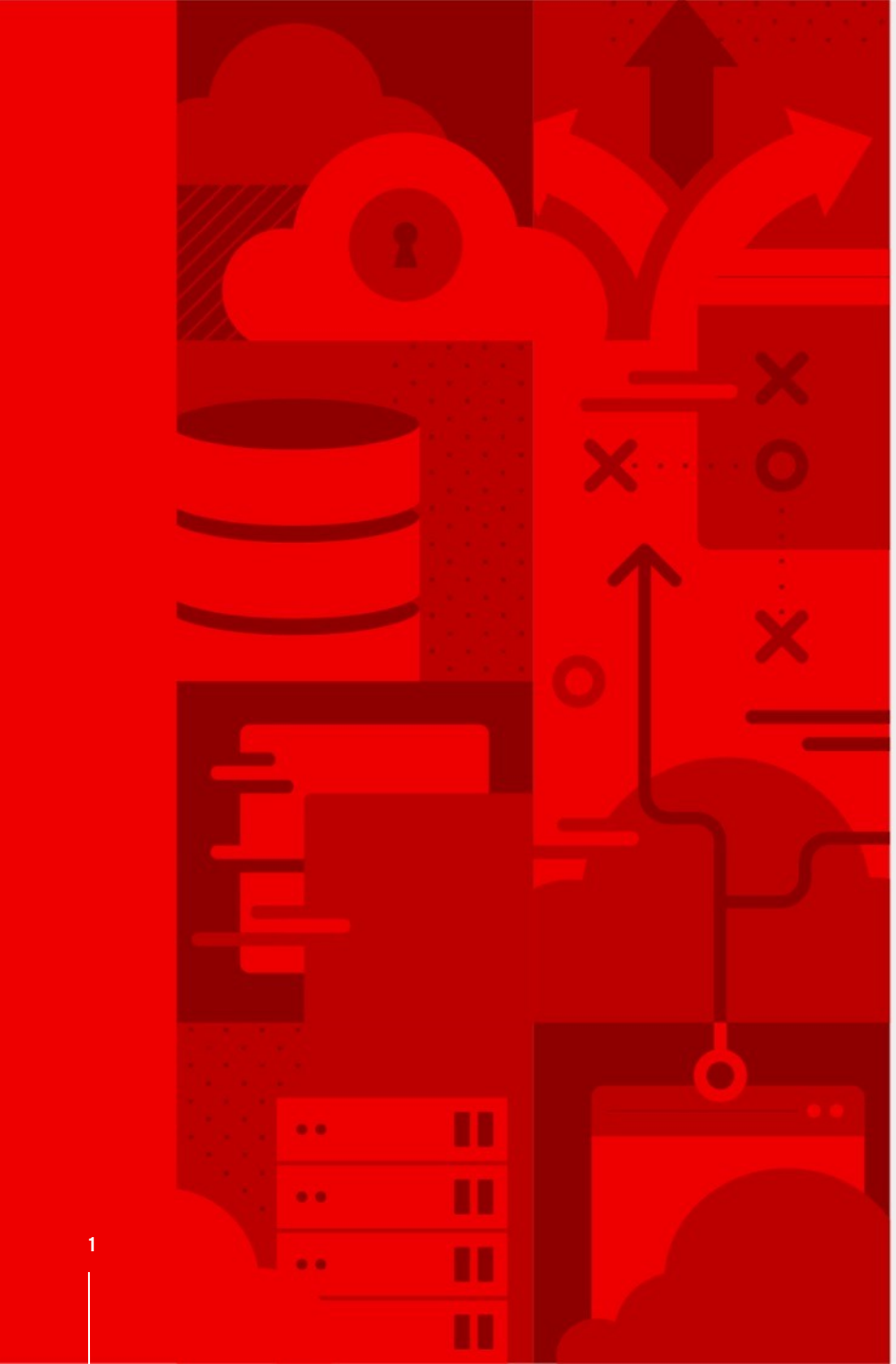


Introducing Git

Eddie Kovsky
ekovsky@redhat.com
IRC: ewk

CC BY NC SA

Red Hat

An Engineer's Toolbox

* Operating system
* Programming language
* Version control
* Your $EDITOR

Red Hat

https://xkcd.com/1597

Red Hat

What changed?

```
$ cat file1.txt
The Royal Tenenbaums is my favorite movie.
Luke Wilson is the best Wilson brother.
He should have won an Oscar for that movie.

$ cat file2.txt
The Royal Tenenbaums is my favorite movie.
Owen Wilson is the best Wilson brother.
He should have won an Oscar for that movie.
```


Red Hat

diff

```
$ diff file1.txt file2.txt
2c2
< Luke Wilson is the best Wilson brother.
---
> Owen Wilson is the best Wilson brother.
```


Red Hat

unified diff

```
$ diff -u file1.txt file2.txt

--- file1.txt 2019-11-08 11:14:06.927351536 -0700
+++ file2.txt 2019-11-08 11:14:37.986533112 -0700
@@ -1,3 +1,3 @@
The Royal Tenenbaums is my favorite movie.
-Luke Wilson is the best Wilson brother.
+Owen Wilson is the best Wilson brother.
He should have won an Oscar for that movie.
```


Red Hat

unified diff

```
$ diff -u hello1.py hello2.py

--- hello1.py   2020-02-10 18:27:16.748772276 -0700
+++ hello2.py   2020-02-10 18:27:45.261180670 -0700
@@ -1,4 +1,4 @@
 #!/usr/bin/env python3

-greeting = "Hello world!"
+greeting = "Hello gorgeous!"
 print(greeting)
```


Red Hat

Version control is the history of changes to text.

A version control system (or revision control or source control) is software that records those changes over time.

Red Hat

An abbreviated list of prior art

* Concurrent Versions System – CVS (retired)
* Subversion (client-server)
* BitKeeper (proprietary)

Red Hat

How Git is different

```
$ man git

*NAME*
       git - the stupid content tracker

*DESCRIPTION*
       Git is a fast, scalable, distributed revision control system
```


Red Hat

Things we can do with Git

* Everyone has a local copy of project history. (**clone**)
* Develop features concurrently alongside releases. (**branch** and **rebase**)
* Record when a feature was released. (**tag**)
* Identify where a bug was introduced. (**bisect**)
* Identify who introduced the change. (**blame**)
* Quickly search repository history. (**grep** and **log**)

Red Hat

2017 kernel rate of change

* New kernel released every ~70 days
* ~8 changes per hour
* 13,000 to 16,000 patches per release
* Bonus: Red Hat was the third most active corporate contributor

2017 State of Linux Kernel Development – 4.8 through 4.13
https://www.linuxfoundation.org/2017-linux-kernel-report-landing-page

Red Hat

Provide mechanism, not policy.

Red Hat

Commit Identifier

Hashing the content plus some entropy produces a unique ID.

This is the **Commit Identifier**.

The commit ID is a globally unique number that is mapped to the content change.

Red Hat

```
commit a430e67d9a2c62a8c7b315b99e74de02018d0a96
Author: Tadeusz Struk <tadeusz.struk@intel.com>
Date:   Tue Jan 7 14:04:48 2020 -0800

    tpm: Handle negative priv->response_len in tpm_common_read()

    The priv->response_length can hold the size of an response or an negative
    error code, and the tpm_common_read() needs to handle both cases correctly.
    Changed the type of response_length to signed and accounted for negative
    value in tpm_common_read().

    Cc: stable@vger.kernel.org
    Fixes: d23d12484307 ("tpm: fix invalid locking in NONBLOCKING mode")
    Reported-by: Laura Abbott <labbott@redhat.com>
    Signed-off-by: Tadeusz Struk <tadeusz.struk@intel.com>
    Reviewed-by: Jarkko Sakkinen <jarkko.sakkinen@linux.intel.com>
    Signed-off-by: Jarkko Sakkinen <jarkko.sakkinen@linux.intel.com>

--->8
```


Red Hat

```
diff --git a/drivers/char/tpm/tpm-dev-common.c b/drivers/char/tpm/tpm-dev-common.c
index b23b0b999232..87f449340202 100644
--- a/drivers/char/tpm/tpm-dev-common.c
+++ b/drivers/char/tpm/tpm-dev-common.c
@@ -130,7 +130,7 @@ ssize_t tpm_common_read(struct file *file, char __user *buf,
                priv->response_read = true;

                ret_size = min_t(ssize_t, size, priv->response_length);
-               if (!ret_size) {
+               if (ret_size <= 0) {
                        priv->response_length = 0;
                        goto out;
                }
diff --git a/drivers/char/tpm/tpm-dev.h b/drivers/char/tpm/tpm-dev.h
index 1089fc0bb290..f3742bcc73e3 100644
--- a/drivers/char/tpm/tpm-dev.h
+++ b/drivers/char/tpm/tpm-dev.h
@@ -14,7 +14,7 @@ struct file_priv {
        struct work_struct timeout_work;
        struct work_struct async_work;
        wait_queue_head_t async_wait;
```


Red Hat

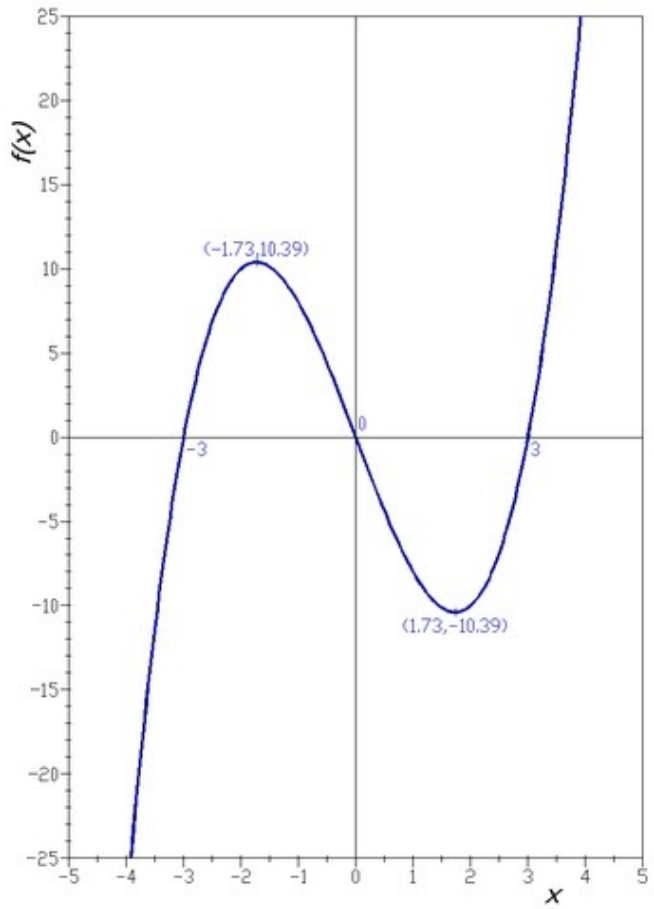
Directed Acyclic Graph - The DAG

Commit identifiers are not very useful if we cannot see how they relate to each other.

Git uses a data structure called a graph to build a timeline of commits.

Red Hat

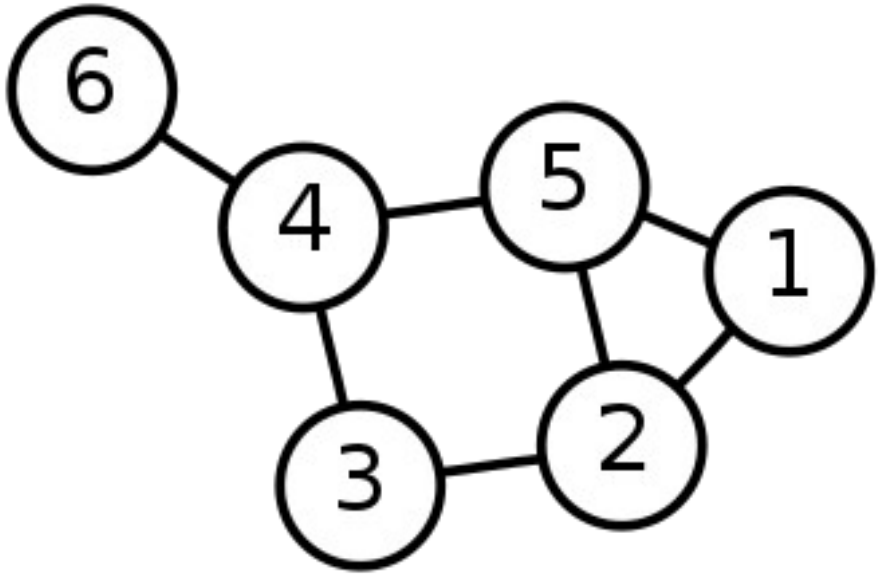
Not this kind of graph



CC BY-SA 4.0 https://en.wikipedia.org/wiki/Graph_of_a_function#/media/File:F(x)_=_x^3_%E2%88%92_9.PNG

A graph is a structure with nodes and edges.

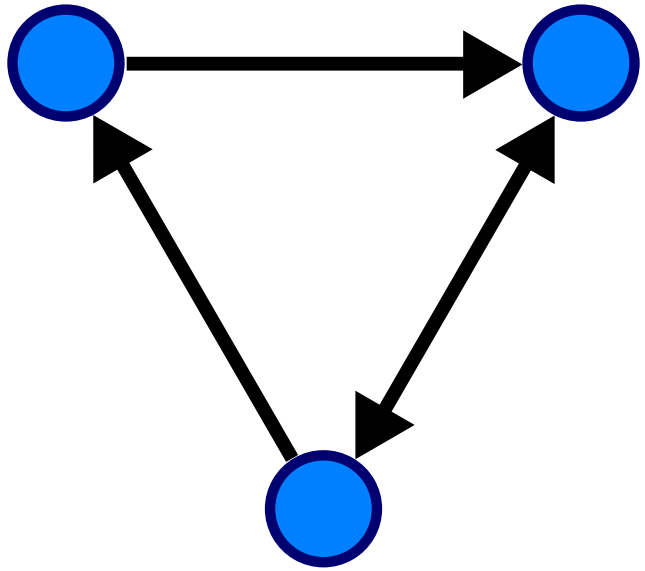
Edges connect each node.



Public domain https://upload.wikimedia.org/wikipedia/commons/thumb/5/5b/6n-graf.svg/333px-6n-graf.svg.png

Red Hat

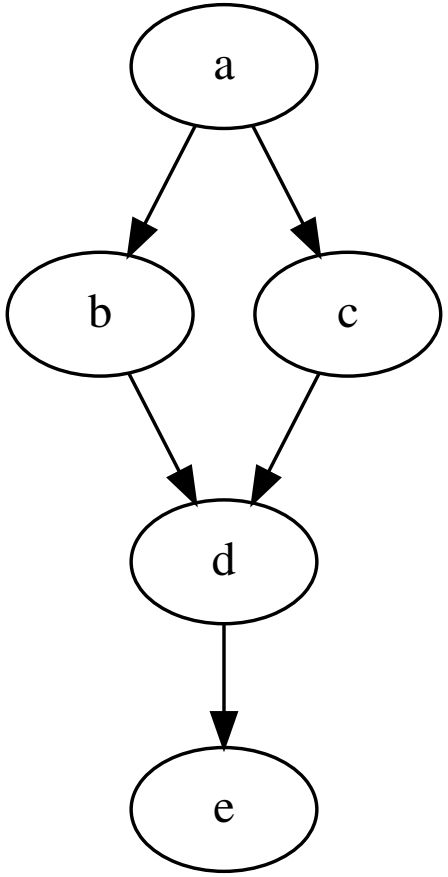
In a directed graph, edges have a direction.



Public domain https://en.wikipedia.org/wiki/File:Directed.svg

Red Hat

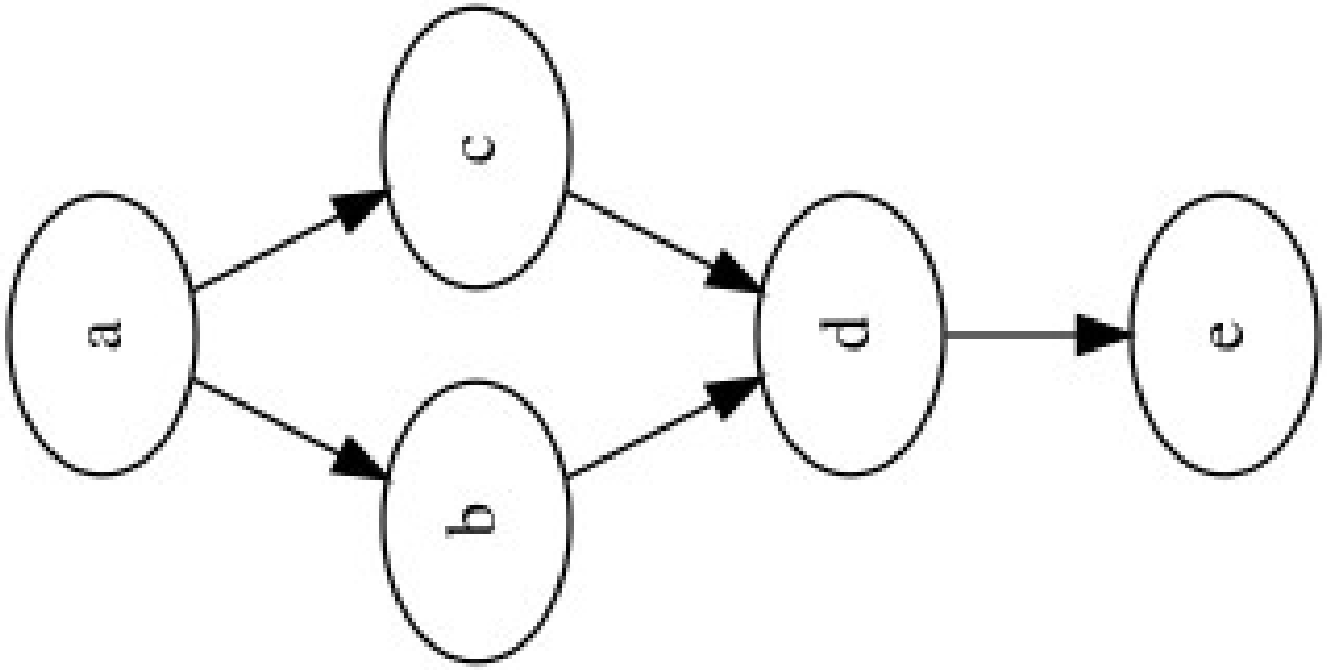
An acyclic graph does not have a closed path.



CC BY-CC0 1.0 https://upload.wikimedia.org/wikipedia/commons/e/ef/Tred-Gprime.svg

Red Hat

For us humans, time is linear.



Red Hat

Build and survey history

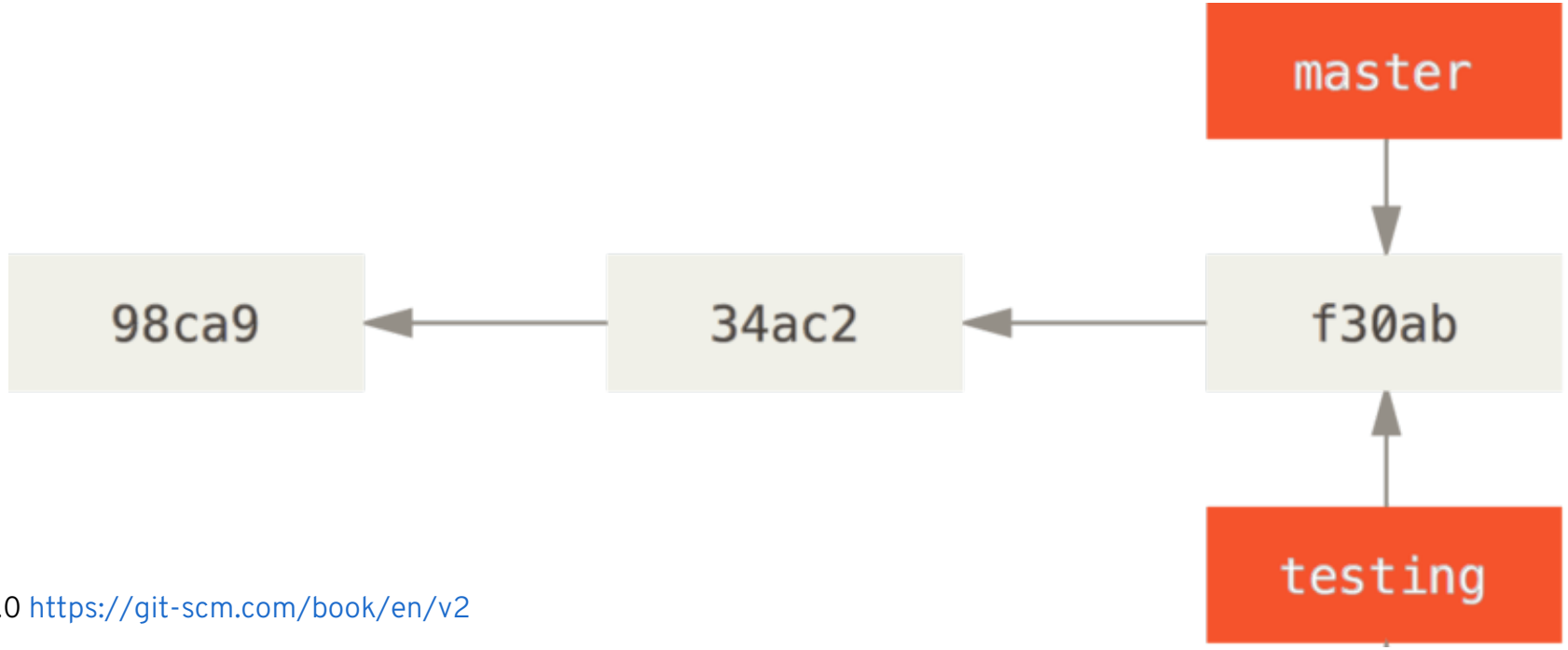
Nodes are the commit IDs.

Edges connect ancestor commits to construct a timeline.

An (arbitrary) ordering of commits produces a DAG.

Red Hat

Putting it all together



CC BY-NC-SA 3.0 https://git-scm.com/book/en/v2

Red Hat

9520b835ffda - mfd: 88pm860x-core: Convert to i2c_new_dummy_device (3 months ago) <Wolfram Sang>
83215897356f - mfd: 88pm800: Convert to i2c_new_dummy_device (3 months ago) <Wolfram Sang>
aee36174b22d - mfd: timberdale: Use dev_get_drvdata (3 months ago) <Chuhong Yuan>
e406b832d89d - mfd: da9063: Remove now unused platform_data (3 months ago) <Wolfram Sang>
a604e5b29ce6 - mfd: tps80031: Convert to devm_i2c_new_dummy_device (3 months ago) <Wolfram Sang>
624e3fceb533 - mfd: cros_ec: Update cros_ec_commands.h (3 months ago) <Yicheng Li>
ed075453d527 - dt-bindings: mfd: rn5t618: Document optional property system-power-controller (3 months ago) <Jonathan Neuschäfer>
cbfe612d471f - mfd: aat2870: No need to check return value of debugfs_create functions (3 months ago) <Greg Kroah-Hartman>
64e8a9bacadb - mfd: ab8500: No need to check return value of debugfs_create functions (3 months ago) <Greg Kroah-Hartman>
efecc3b531a3 - mfd: ab3100: No need to check return value of debugfs_create functions (3 months ago) <Greg Kroah-Hartman>
b5e29aa880be - mfd: davinci_voicecodec: Remove pointless #include (4 months ago) <Arnd Bergmann>
197df18f7038 - mfd: max77620: Add of_node_put() before return (4 months ago) <Nishka Dasgupta>
d0b3cfee333e - Merge tag 'backlight-next-5.4' of git://git.kernel.org/pub/scm/linux/kernel/git/lee/backlight (7 weeks ago) <Linus Torvalds>
c0b64faf0fe6 - backlight: pwm_bl: Set scale type for brightness curves specified in the DT (10 weeks ago) <Matthias Kaehlcke>
511a204638d7 - backlight: pwm_bl: Set scale type for CIE 1931 curves (10 weeks ago) <Matthias Kaehlcke>
d55c028f8b25 - backlight: Expose brightness curve type through sysfs (10 weeks ago) <Matthias Kaehlcke>
6451e123dec3 - MAINTAINERS: Add entry for stable backlight sysfs ABI documentation (10 weeks ago) <Matthias Kaehlcke>
ec665b756e6f - backlight: gpio-backlight: Correct initial power state handling (10 weeks ago) <Peter Ujfalusi>
28a1d72a221e - video: backlight: tosa_lcd: drop check because i2c_unregister_device() is NULL safe (10 weeks ago) <Wolfram Sang>
bcd69da98e36 - video: backlight: Drop default m for {LCD,BACKLIGHT_CLASS_DEVICE} (10 weeks ago) <Geert Uytterhoeven>
30f644fd6ec9 - backlight: lms283gf05: Fix a typo in the description passed to 'devm_gpio_request_one()' (3 months ago) <Christophe JAILLET>
072e2c8192cf - backlight: lm3630a: Switch to use fwnode_property_count_uXX() (3 months ago) <Andy Shevchenko>
865cac14c2da - backlight: rave-sp: Leave initial state and register with correct device (3 months ago) <Lucas Stach>
299d14d4c31a - Merge tag 'pci-v5.4-changes' of git://git.kernel.org/pub/scm/linux/kernel/git/helgaas/pci (7 weeks ago) <Linus Torvalds>
c5048a73b477 - Merge branch 'pci/trivial' (7 weeks ago) <Bjorn Helgaas>
55507aca5882 - PCI: Remove unnecessary returns (2 months ago) <Krzysztof Wilczynski>
b071c1fd7a8a - PCI: OF: Correct of_irq_parse_pci() documentation (2 months ago) <Lubomir Rintel>
d1bbf38aaf88 - PCI: Fix typos and whitespace errors (2 months ago) <Bjorn Helgaas>
cdc7e184abb3 - Merge branch 'remotes/lorenzo/pci/vmd' (7 weeks ago) <Bjorn Helgaas>
a1a30170138c - PCI: vmd: Fix shadow offsets to reflect spec changes (7 weeks ago) <Jon Derrick>
e3dffa4f6c36 - PCI: vmd: Fix config addressing when using bus offsets (7 weeks ago) <Jon Derrick>
3efa7f1febe6 - Merge branch 'lorenzo/pci/tegra' (7 weeks ago) <Bjorn Helgaas>

Red Hat

Let’s rewrite history!

CC BY 2.5 https://commons.wikimedia.org/w/index.php?curid=11293395

Red Hat

git init

```
$ diff -u file1.txt file2.txt
--- file1.txt 2019-11-08 11:14:06.927351536 -0700
+++ file2.txt 2019-11-08 11:14:37.986533112 -0700
@@ -1,3 +1,3 @@
The Royal Tenenbaums is my favorite movie.
-Luke Wilson is the best Wilson brother.
+Owen Wilson is the best Wilson brother.
He should have won an Oscar for that movie.
```


Red Hat

Let’s rewrite history!

```
$ git rebase
```

```
$ git reflog
```


Red Hat

The art of finding stuff

Like regular grep, but faster!

```
$ git grep <pattern>
$ git log --grep=<pattern>
$ git blame <file>
$ git bisect
```


Red Hat

Demo 1

Version 1 is working.

Version 2 has a bug.

Can we find the commit that introduced the bug,

even if there are hundreds of commits between V1 and V2?

Red Hat

Let’s play a game!

I am thinking of a number between 1 and 100.

What should your first guess always be?

Red Hat

Binary search

If the problem set can be ordered, each pass reduces the problem space by half.

Red Hat

Demo 1

```
$ git bisect
```


Red Hat

Demo 2

Customer sees a new error after updating kernel:

```
[drm] Unable to create a private tmpfs mount, hugepage support
will be disabled(-22).
```


Red Hat

Demo 2

```
$ git grep '[drm] Unable to create a private tmpfs mount,
hugepage support will be disabled(-22).'
```


Red Hat

Demo 2

```
$ git grep '\[drm\] Unable to create a private tmpfs mount,
hugepage support will be disabled\(-22\)'
```


Red Hat

Demo 2

```
$ git grep 'Unable to create a private tmpfs mount, hugepage
support will be disabled'

drivers/gpu/drm/i915/gem/i915_gem_shmem.c:593:
DRM_NOTE("Unable to create a private tmpfs mount, hugepage
support will be disabled(%d).\n"
```


Red Hat

Demo 2

```
$ git log --grep='Unable to create a private tmpfs mount,
hugepage support will be disabled'
```


Red Hat

Demo 2

```
$ git blame drivers/gpu/drm/i915/gem/i915_gem_shmem.c
```


Red Hat

Demo 3

Find this definition.

```
$ git grep 'struct task_struct'
```


Red Hat

Demo 3

We only want the definition, but we are getting every usage.

```
$ git grep 'struct task_struct'

struct task_struct *task;       /* main task structure */
syscall_get_arch(struct task_struct *task)
...
```


Red Hat

Demo 3

Domain specific knowledge of the C language quickly narrows our search:

```
struct foo {
        field1
        field2
        ...
    }
```


Red Hat

Demo 3

```
$ git grep 'struct task_struct {'

include/linux/sched.h:629:struct task_struct {
tools/include/linux/lockdep.h:26:struct task_struct {
tools/testing/selftests/bpf/progs/strobemeta.h:14:struct task_struct {};
tools/testing/selftests/bpf/progs/test_core_reloc_kernel.c:24:struct task_struct {
```


Red Hat

Resources

man pages

```
$ man git
```

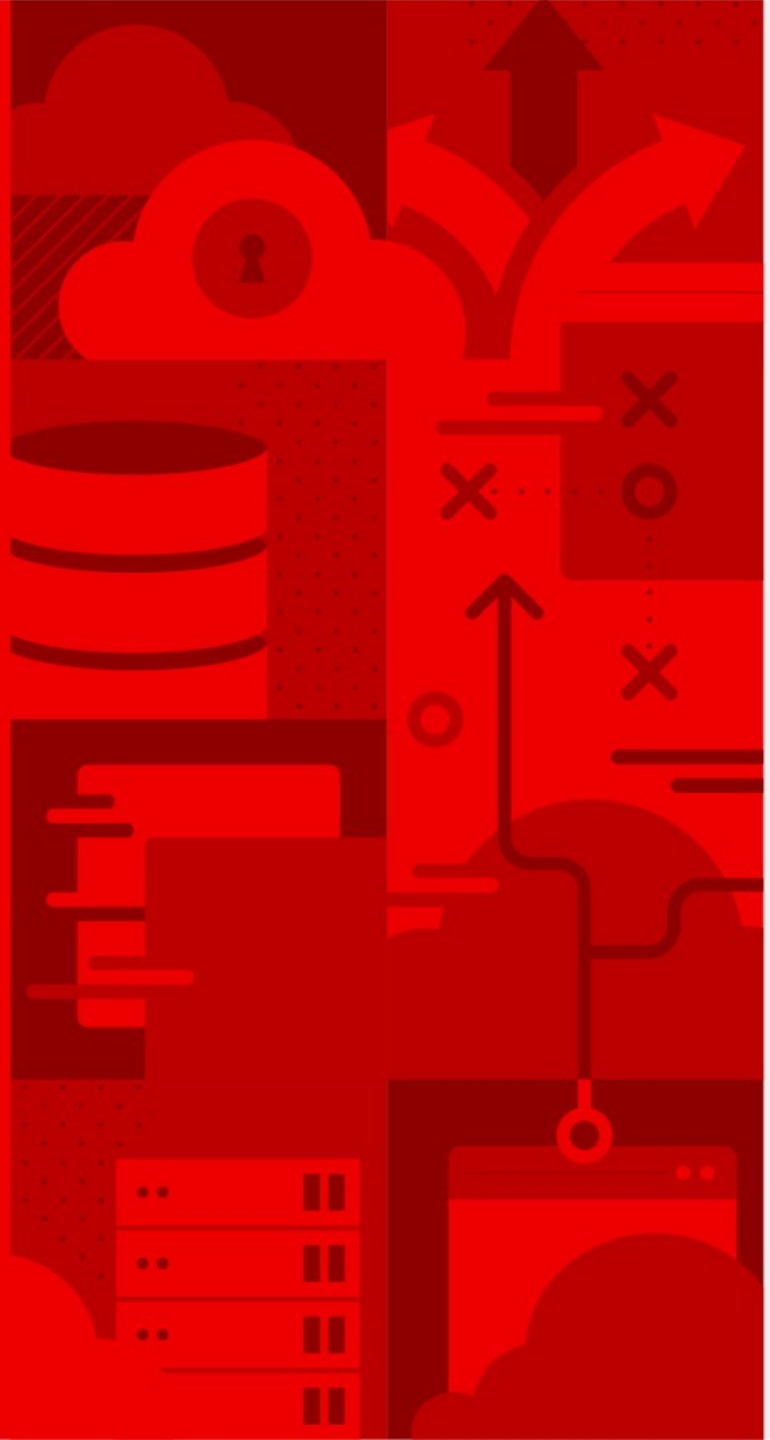
Pro Git

https://git-scm.com/book/en/v2

Git from the inside out

https://maryrosecook.com/blog/post/git-from-the-inside-out

Red Hat



Thank you

Eddie Kovsky
ekovsky@redhat.com
IRC: ewk

linkedin.com/company/red-hat

facebook.com/redhatinc

youtube.com/user/RedHatVideos

twitter.com/RedHat

Red Hat