

a453004

A story about a commit

Daniel Gustafsson

Pivotal

daniel@yesql.se @d_gustafsson

a453004

This is why Daddy drinks..

Daniel Gustafsson

Pivotal

daniel@yesql.se @d_gustafsson

Hello, I'm Daniel Gustafsson

Pivotal.

I work at with and



**GREENPLUM
DATABASE**



Greenplum: Massively Parallelised Processing

Shared Nothing Cluster Data Warehouse based

on PostgreSQL



BIG DATA!

BIG DATA!

(No really, quite big actually)

**Greenplum was
forked off
PostgreSQL 8.2**

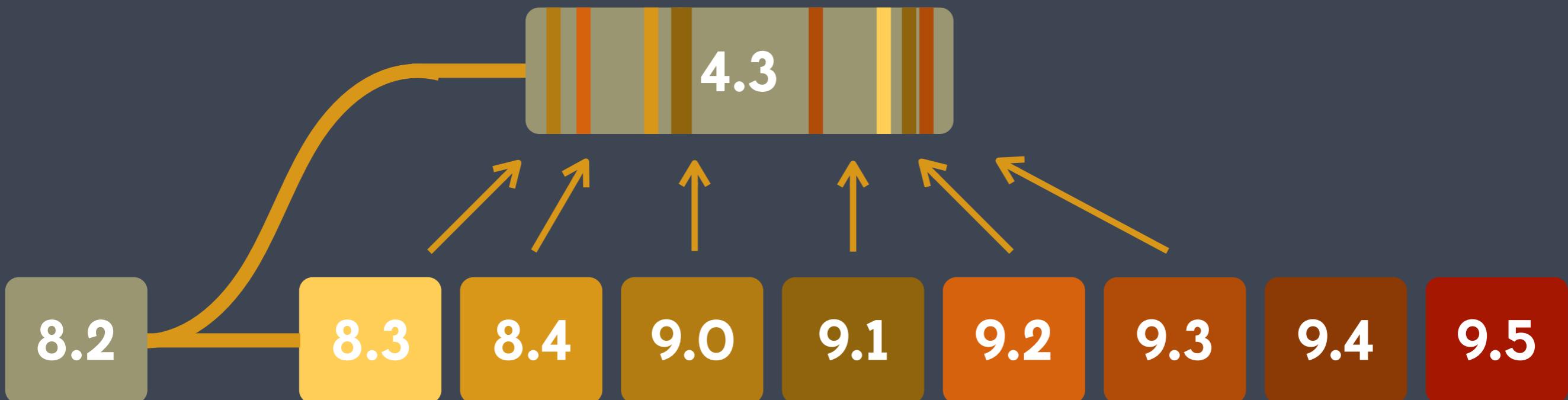
**Greenplum was
forked off**

PostgreSQL 8.2

(.. Kind of Sort of ..)

Eating cherries with a fork

Greenplum

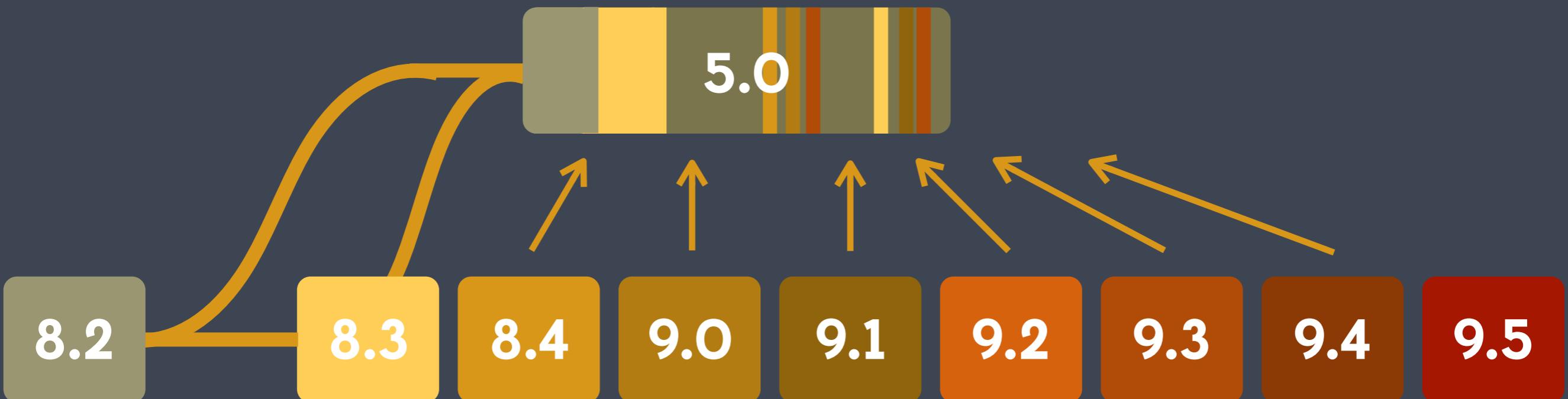


PostgreSQL

Project:
Merge upstream
with entire commit
history intact

Complete upstream merges + GPDB code

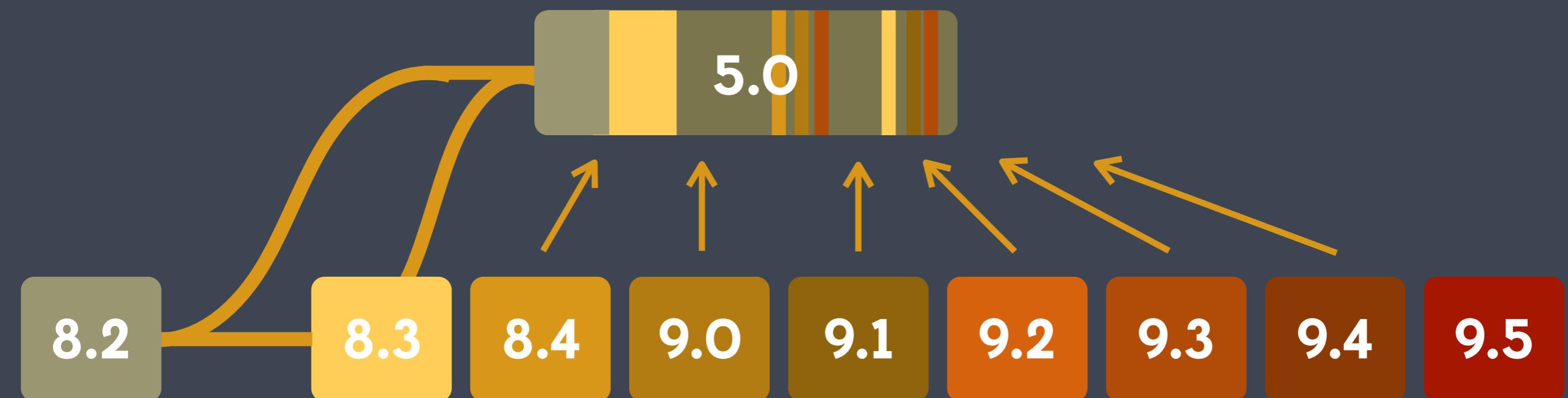
Greenplum



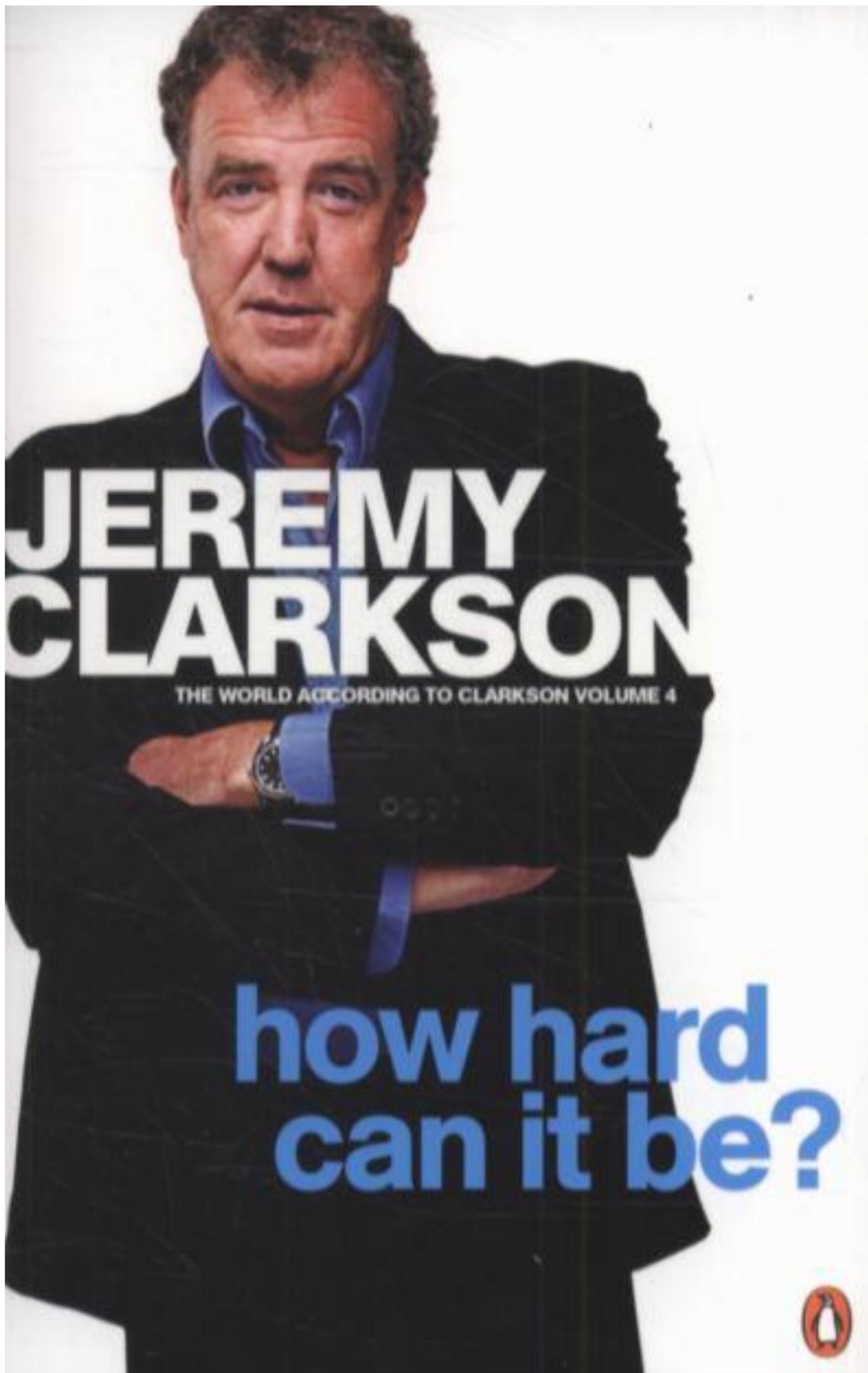
PostgreSQL

Time to catch up!

Greenplum



PostgreSQL



JEREMY CLARKSON

THE WORLD ACCORDING TO CLARKSON VOLUME 4

how hard
can it be?



```
$ git diff --stat REL8_2_23..REL8_3_23
2423 files changed,
461334 insertions(+),
208900 deletions(-)
```



```
$ git merge upstream/REL8_3_STABLE  
$ grep -r "<<<<< HEAD" . | wc -l  
6653
```

BILLY OCEAN



**WHEN THE GOING GETS TOUGH,
THE TOUGH GET GOING**

from the Twentieth Century Fox Motion Picture Soundtrack



6600+ conflicts
2 developers

We.. just might need
some strategy here

1



10 make 

20 merge n commits

30 hack hack hack

40 push to master

50 if more, goto 10

- Hard to parallelise
Probably too slow?*



Contained changes

*The jury is still out

2



10 merge the rest (~75%)

20 push all the conflicts

30 hack ☕ push hack ☕

40 maintain dirty branch

50 rewrite the branch

50 push clean to master

60 🍺🎉🍸🍰😴zzZ

-

Harder to debug A bit overwhelming



Easy to parallelise Got the job done!

- * a243b84 (**HEAD**) Re-enable psql queries that depend on 8.3 catalogs.
 - | Heikki Linnakangas, 19 hours ago
 - * 40d6ea4 Backport the "proper" fix for CVE-2013-0255.
 - | Heikki Linnakangas, 20 hours ago
 - * a453004 Merge with PostgreSQL 8.3.23.
 - | Heikki Linnakangas, 20 hours ago
 - | \
 - * d13f41d (**upstream/REL8_3_STABLE**) Disable -faggressive-loop-
 - optimizations in gcc 4.8+ for pre-9.2 branches. Alvaro Herrera, 1 year, 5 months ago
 - * 1b7f516 Return error if allocation of new element was not possible. Michael Meskes, 2 years, 9 months ago
 - * 0caf562 (**tag: REL8_3_23**) Stamp 8.3.23.
 - Tom Lane, 3 years, 4 months ago
 - * aa27464 Prevent execution of enum_recv() from SQL.
 - Tom Lane, 3 years, 4 months ago
 - * fea6323 Update release notes for 9.2.3, 9.1.8, 9.0.12, 8.4.16, 8.3.23. Tom Lane, 3 years, 4 months ago
- .
- .
- .



So, which is best?

So, which is best?

None and/or Both

commit f41803bb39bc2949db200116a609fd242d0ec221

Author: Tom Lane <tgl@sss.pgh.pa.us>

Date: Sat Jan 20 20:45:41 2007 +0000

Refactor planner's pathkeys data structure to create a separate, explicit representation of equivalence classes of variables. This is an extensive rewrite, but it brings a number of benefits:

- * planner no longer fails in the presence of "incomplete" operator families that don't offer operators for every possible combination of datatypes.
- * avoid generating and then discarding redundant equality clauses.
- * remove bogus assumption that derived equalities always use operators named "=".
- * mergejoins can work with a variety of sort orders (e.g., descending) now, instead of tying each mergejoinable operator to exactly one sort order.
- * better recognition of redundant sort columns.
- * can make use of equalities appearing underneath an outer join.



ONE DOES NOT SIMPLY
MERGE..

..AN EXTENSIVE REWRITE
BY TOM LANE



Merge complicated +
supporting patches

Merge big chunks in
dirty branches

Lots of coffee

What happens
when you look at
code no-one else is
looking at?

Not much..

Not much..

(This is PostgreSQL after all)

Small error in contrib docs since 9.1

commit 768d6f90f9e44af96c22135a8eb1e83ed73c422b

Author: Tom Lane <tgl@sss.pgh.pa.us>

Date: Sat May 21 15:43:57 2016 -0400

Improve docs about contrib/intarray's benchmark suite.

Correct obsolete install instructions, as noted by Daniel Gustafsson.

Clarify the test code's prerequisites.

Discussion: <88E617F2-7721-4C4E-84F4-886A2041C1D0@yesql.se>

..referring to a script it seems that not
too many use

Fix bug in contrib/intarray bench script for showing query
plans with EXPLAIN

Ready for Committer

Ye olde = vs == equality

```
commit 7d7b129277eb545286aecf29ec22b5bb7fdf46bd
Author: Tom Lane <tgl@sss.pgh.pa.us>
Date: Mon May 2 11:18:10 2016 -0400
```

Fix configure's incorrect version tests for flex and perl.

awk's equality-comparison operator is "==" not "=". We got this right in many places, but not in configure's checks for supported version numbers of flex and perl. It hadn't been noticed because **unsupported versions are so old as to be basically extinct in the wild**, and because the only consequence is whether or not a WARNING flies by during configure.

Daniel Gustafsson noted the problem with respect to the test for flex, I found the other by reviewing other awk calls.

Some dead code

commit d22b85fb4ca3b4c2f41d8d4372205d56b296eb3

Author: Heikki Linnakangas <heikki.linnakangas@iki.fi>

Date: Mon May 2 10:07:49 2016 +0300

Remove unused macros.

`CHECK_PAGE_OFFSET_RANGE()` has been unused forever.

`CHECK_RELATION_BLOCK_RANGE()` has been unused in `pgstatindex.c` ever since `bt_page_stats()` and `bt_page_items()` functions were moved from `pgstattuple` to `pageinspect` module. It still exists in `pageinspect/btreefuncs.c`.

Daniel Gustafsson

A typo originating in Postgres95(!)

```
commit b7351ced425f3937f0a61adb4ade1d4b93bf751d
Author: Magnus Hagander <magnus@hagander.net>
Date:   Tue Apr 26 10:38:32 2016 +0200
```

Fix typo in comment

Author: Daniel Gustafsson

A really old pg_dump bug

commit 0461b66e361f56560aaa50e1bd006e36e8bedc6e

Author: Tom Lane <tgl@sss.pgh.pa.us>

Date: Fri May 26 12:51:05 2017 -0400

Fix pg_dump to not emit invalid SQL for an empty operator class.

If an operator class has no operators or functions, and doesn't need a STORAGE clause, we emitted "CREATE OPERATOR CLASS ... AS ;" which is syntactically invalid. Fix by forcing a STORAGE clause to be emitted anyway in this case.

(At some point we might consider changing the grammar to allow CREATE OPERATOR CLASS without an opclass_item_list. But probably we'd want to omit the AS in that case, so that wouldn't fix this pg_dump issue anyway.)

It's been like this all along, so back-patch to all supported branches.

Daniel Gustafsson, tweaked by me to avoid a dangling-pointer bug

**PostgreSQL code
quality is unrivalled**

**PostgreSQL code
quality is unrivalled**

(But hey, this is PostgreSQL after all)

Thanks!

daniel@yesql.se @d_gustafsson