

# Merge-tree: Visualizing the Integration of Commits into Linux

*Evan Wilde, Daniel German.*

*2016 IEEE Working Conference on Software Visualization (VISSOFT),  
Raleigh, NC, pp. 1-10, 2016.*

*Presented by: Nick Bradley*

*March 16, 2017*

# Git Version Control

- Is a distributed version control system
- Supports non-linear workflows
- Uses directed acyclic graph (DAG)
- Commits
- Branches

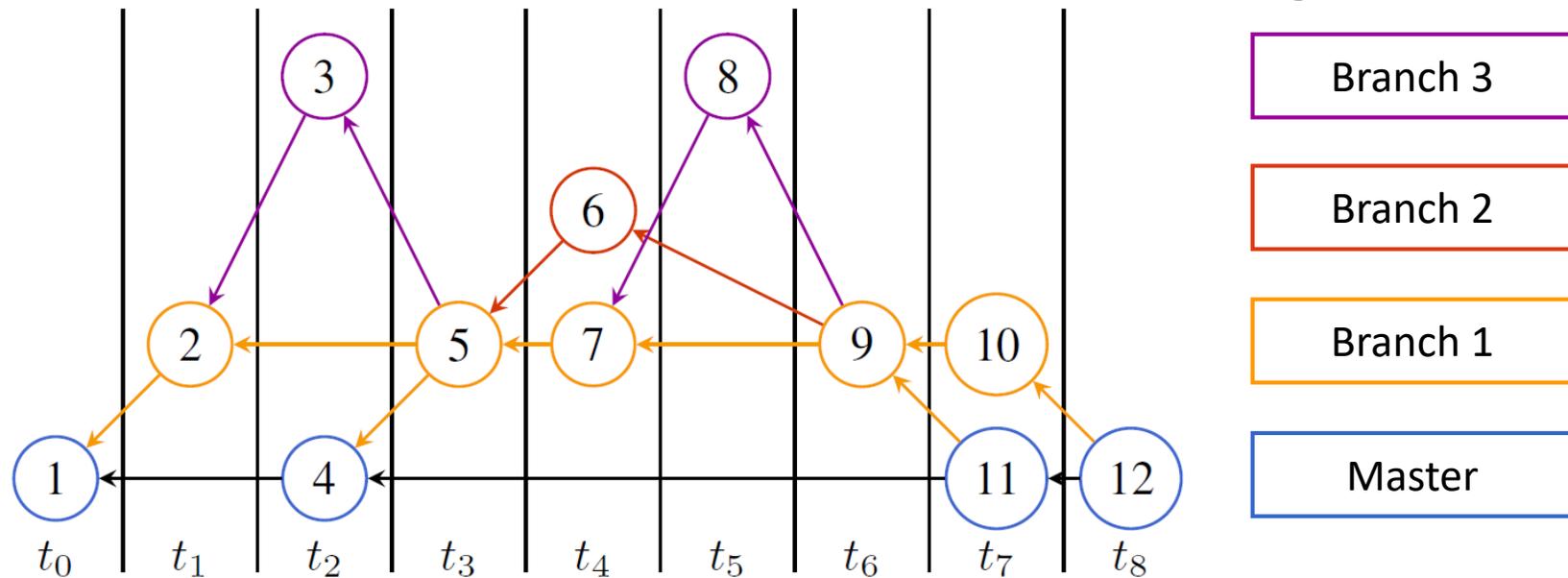
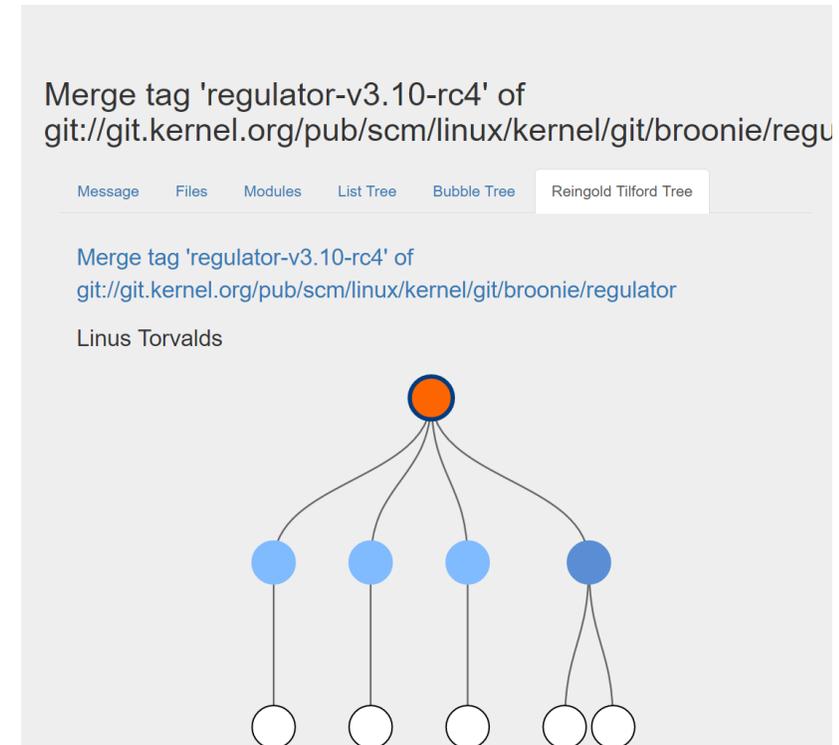


Fig. 3

# Domain Task

- **Show when and how a commit was merged into master**
- Challenge: commits cannot be changed
  - can't link to later commits
  - can't track merge dates
- Solution: Linvis
  - Shows topological view of merges
  - Supports aggregation and filtering
  - Supports two use-cases
    1. top-to-bottom: aggregate
    2. bottom-to-top: see flow into master



Source: <http://li.turingmachine.org>

# Merge-tree

- Transforms DAG into trees
  - each rooted on master
  - such that all commits are assigned to exactly one tree
- Algorithm
  - Invert DAG
  - For every commit
    - compute distance to each subsequent commit
    - only keep link to closest (in time) in merge-tree
    - stop at master commit
- Relies on specific Git workflow

# Convert DAG to Trees

## DAG Model

- Newer commits link to older
- All links present

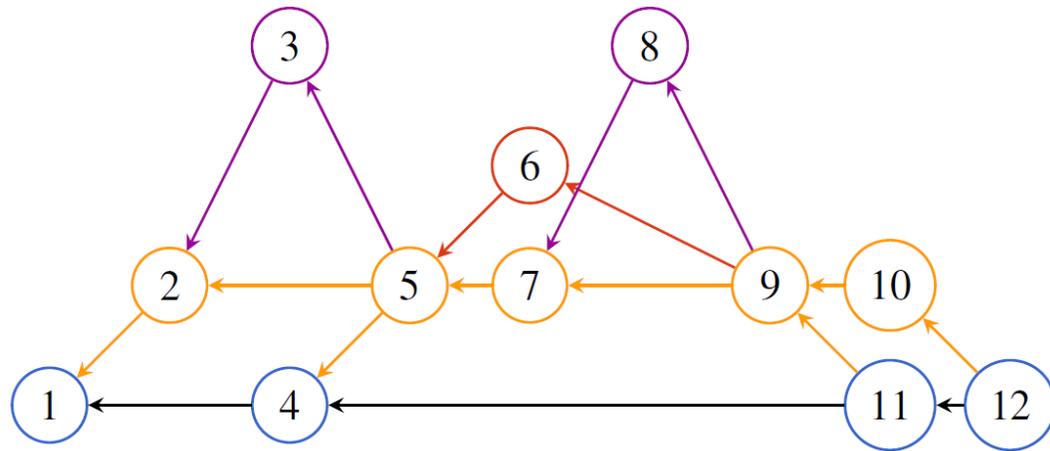


Fig. 4

## Merge-tree Model

- Older commits link to newer
- Removed links
  - Only keep links on shortest path

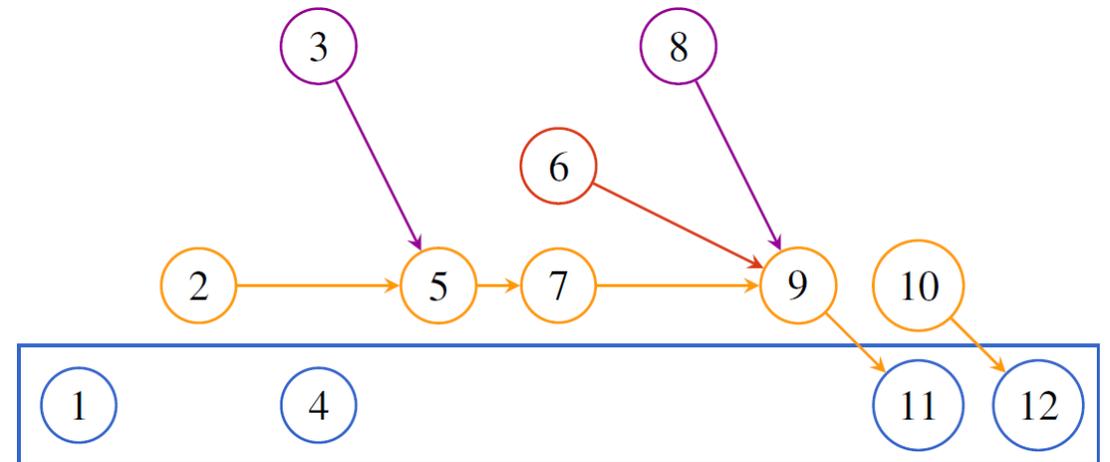


Fig. 5

# Linvis Live Demo

<http://li.turingmachine.org>

# Linvis analysis

- What: data
    - DAG
  - What: derived
    - Merge-tree
  - Why: tasks
    - Search for commits
    - Summarize changes
  - How: reduce
    - Filtering
  - How: Manipulate
    - Navigate with pan/zoom
    - Select
- How: encode
    - indented outline (*list tree*)
    - tree map using nested circles and radial containment (*bubble tree*)
    - vertical node-link (*Reingold-Tilford tree*)

# Limitations + Next Steps

- **No evaluation of Linvis**
  - quantitative user-testing: improvements to user workflow
  - qualitative user-evaluation: do users think tool is helpful
- Merge-tree cannot be constructed for most repositories
- Cannot search by filename
  
- Aggregate commit patches
- Aggregate authorship information

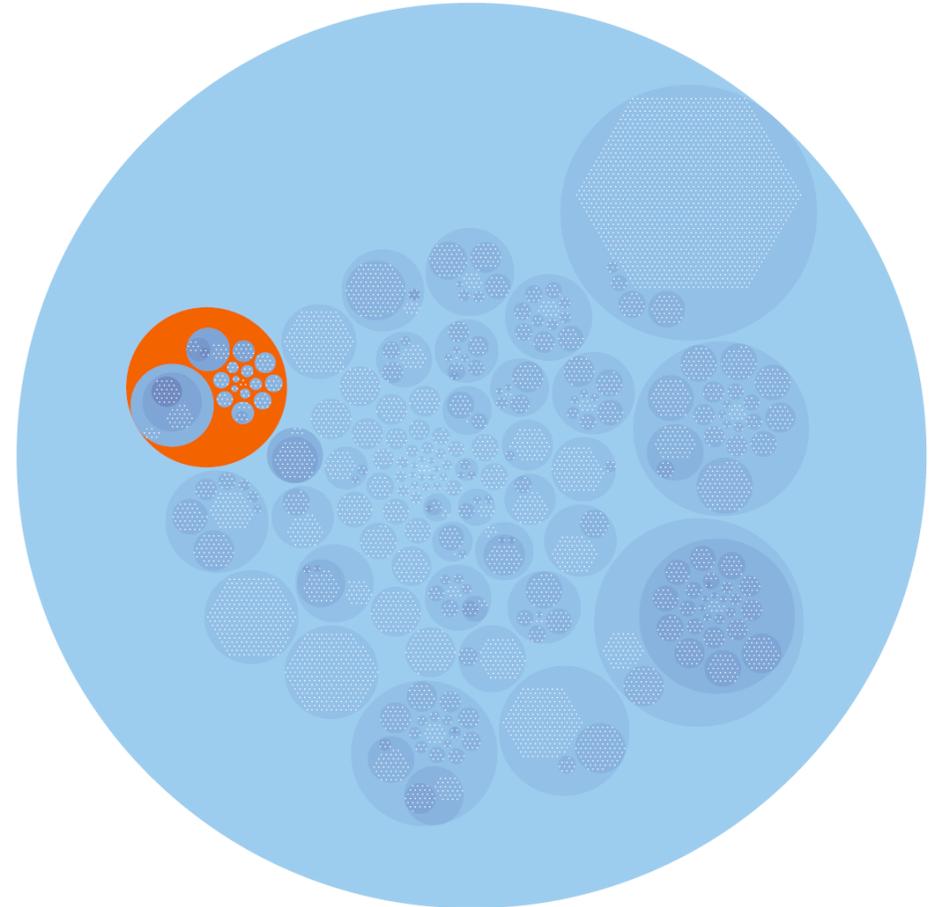
# Compare existing - GitHub



**Couldn't load network graph.**

Too many forks to display.

*Fig. 2*



*Fig. 15*

# Compare existing - Gitk

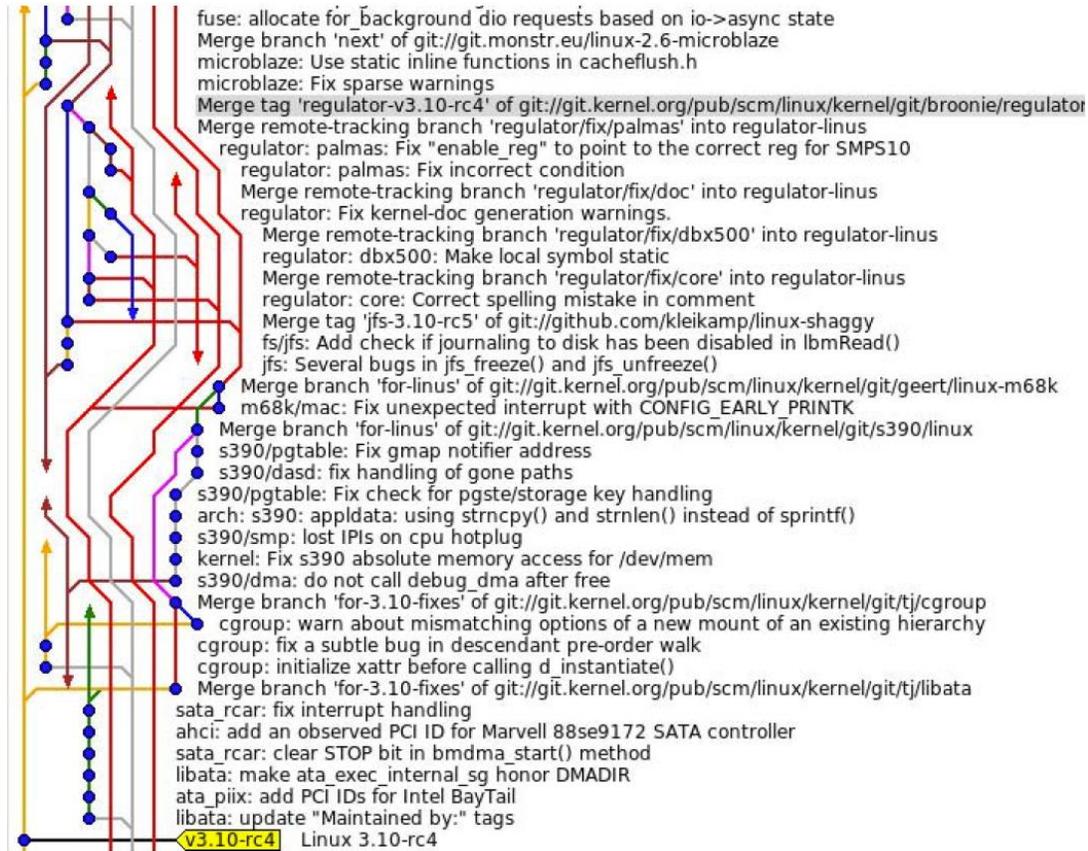


Fig. 16

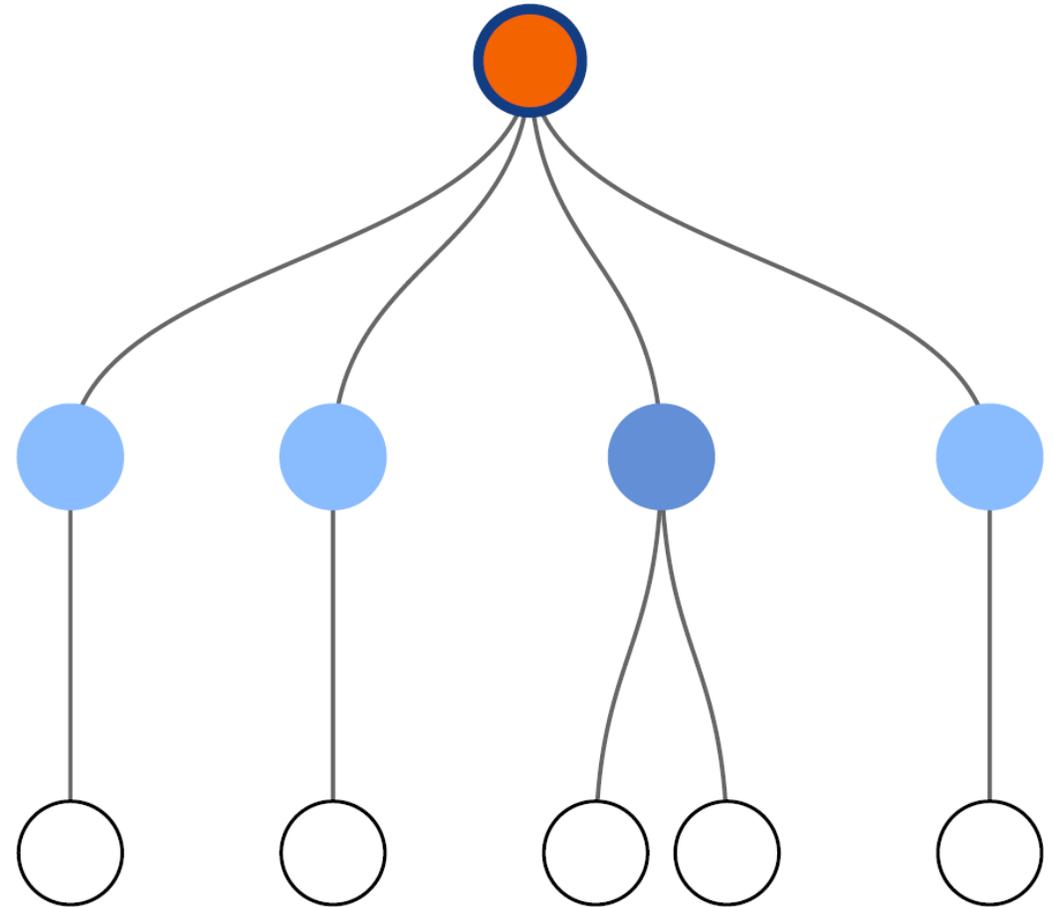


Fig. 17

# Critique

- Strengths
  - Main contribution of merge-tree
  - Demonstrated on most complex DAG
  - Natural interactions (e.g. pan, zoom)
  - Different encodings of tree structure, all intuitive
- Weaknesses
  - Merge-tree algorithm not robust
  - Navigation between views cumbersome
  - Vis felt like afterthought

# Suggestion

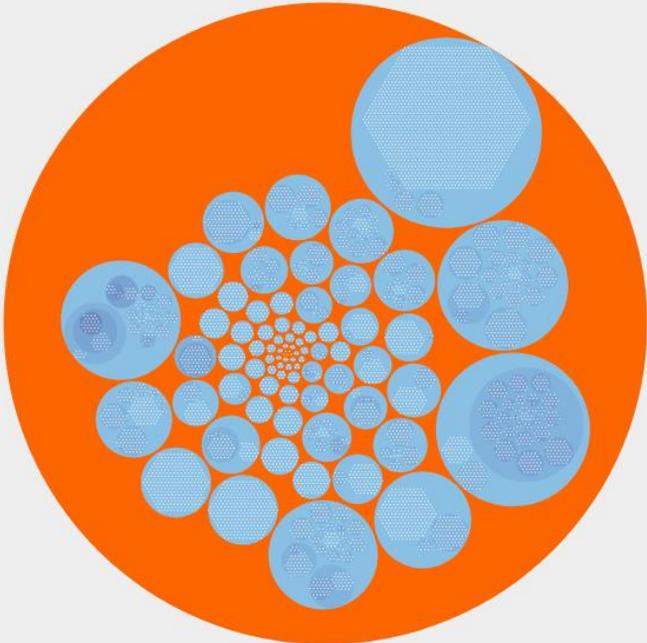
- Juxtapose:
  - Git log
  - Files
  - tree vis

Merge branch 'next' (accumulated 3.16 merge window patches) into master

Message Files Modules List Tree **Bubble Tree** Reingold Tilford Tree

Merge branch 'next' (accumulated 3.16 merge window patches) into master

Linus Torvalds



```
Merge branch 'next' (accumulated 3.16 merge window patches) into master
Now that 3.15 is released, this merges the 'next' branch into 'master',
bringing us to the normal situation where my 'master' branch is the
merge window.

* accumulated work in next: (6809 commits)
ufs: sb mutex merge + mutex_destroy
powerpc: update comments for generic idle conversion
cris: update comments for generic idle conversion
idle: remove cpu_idle() forward declarations
nbd: zero from and len fields in NBD_CMD_DISCONNECT.
mm: convert some level-less printks to pr_*
MAINTAINERS: adi-buildroot-devel is moderated
MAINTAINERS: add linux-api for review of API/ABI changes
mm/kmemleak-test.c: use pr_fmt for logging
fs/dlm/debug_fs.c: replace seq_printf by seq_puts
fs/dlm/lockspace.c: convert simple_str to kstr
fs/dlm/config.c: convert simple_str to kstr
mm: mark remap_file_pages() syscall as deprecated
mm: memcontrol: remove unnecessary memcg argument from soft limit functions
mm: memcontrol: clean up memcg zoneinfo lookup
mm/memblock.c: call kmemleak directly from memblock_alloc/free
mm/mempool.c: update the kmemleak stack trace for mempool allocations
lib/radix-tree.c: update the kmemleak stack trace for radix tree allocations
mm: introduce kmemleak_update_trace()
mm/kmemleak.c: use %u to print ->checksum
...
```

Show 10 entries Search:

Filename	Added	Removed
arch/alpha/include/asm/atomic.h	0	5
arch/alpha/include/asm/bitops.h	0	3
arch/alpha/include/asm/pci.h	0	5
arch/alpha/include/asm/thread_info.h	2	2
arch/arc/include/asm/atomic.h	0	5
arch/arc/include/asm/barrier.h	0	37
arch/arc/include/asm/bitops.h	1	4
arch/arc/include/asm/sections.h	0	1
arch/arc/kernel/devtree.c	1	1
arch/arc/kernel/troubleshoot.c	6	4

Showing 1 to 10 of 6 627 entries Previous 1 2 3 4 5 ... 563 Next

Questions

# Linvis Search View

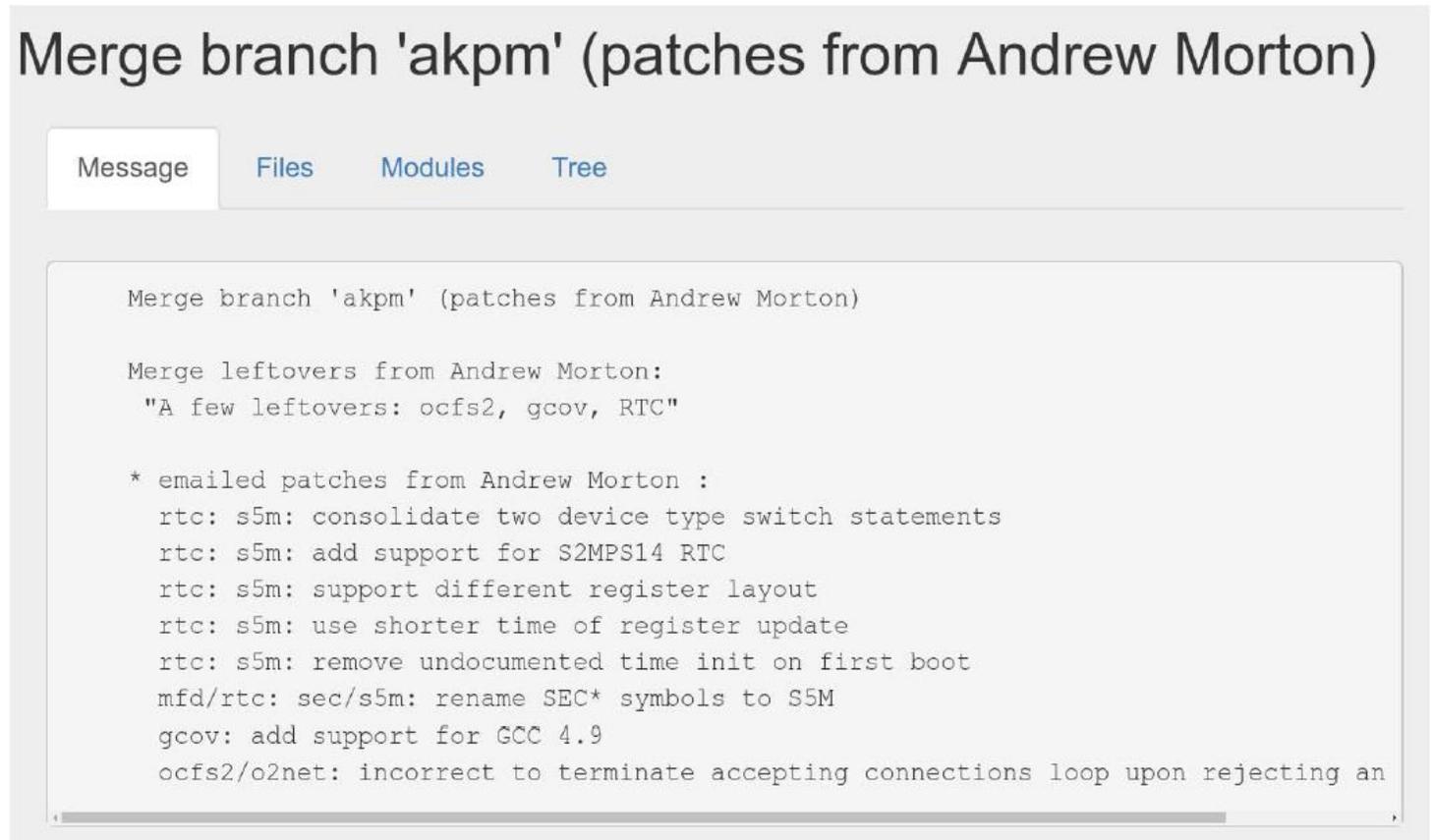
- Allows filtering by
  - Merge date range
  - Commit author
  - Keyword
  - Commit ID

The screenshot shows the 'Linux Kernel' search interface. At the top, there is a search bar with a 'Search' button. Below it is a dropdown menu currently showing 'Linux 3.10'. Underneath is the 'Commit Date Range' section, which includes two input fields: 'Begin Date' with a placeholder 'Begin Date (MM/DD/YYYY)' and 'End Date' with a placeholder 'End Date (MM/DD/YYYY)'. Below these is a text input field containing 'c26'. At the bottom of this section is another dropdown menu showing 'Commit ID'. Below the dropdowns is a checkbox labeled 'Merges by Linus' which is currently unchecked. At the very bottom is another 'Search' button.

*Fig. 7*

# Linvis Message Tab

- Displays Git log
- Included for completeness



Merge branch 'akpm' (patches from Andrew Morton)

Message Files Modules Tree

```
Merge branch 'akpm' (patches from Andrew Morton)

Merge leftovers from Andrew Morton:
"A few leftovers: ocfs2, gcov, RTC"

* emailed patches from Andrew Morton :
rtc: s5m: consolidate two device type switch statements
rtc: s5m: add support for S2MPS14 RTC
rtc: s5m: support different register layout
rtc: s5m: use shorter time of register update
rtc: s5m: remove undocumented time init on first boot
mfd/rtc: sec/s5m: rename SEC* symbols to S5M
gcov: add support for GCC 4.9
ocfs2/o2net: incorrect to terminate accepting connections loop upon rejecting an
```

*Fig. 9*

# Linvis File Tab

- Files changed in leaf commits
- Added/Removed columns show number of changed lines
- Aggregates number of changes to single file across all commits

Merge branch 'akpm' (patches from Andrew Morton)

Message Files Modules Tree

Show 10 entries Search:

Filename	Added	Removed
drivers/rtc/Kconfig	2	2
drivers/rtc/rtc-s5m.c	241	147
fs/ocfs2/cluster/tcp.c	29	2
include/linux/mfd/samsung/rtc.h	48	38
kernel/gcov/base.c	6	0
kernel/gcov/gcc_4_7.c	5	0

Showing 1 to 6 of 6 entries

Previous 1 Next

Fig. 10

# Linvis Module Tab

- Linux specific
- Groups files
- Count column shows number of changed files for all leaf commits

Merge branch 'akpm' (patches from Andrew Morton)

Message Files **Modules** Tree

Show 10 entries Search:

Module	Count
gcov	1
mfd/rtc	1
ocfs2/o2net	1
rtc	5

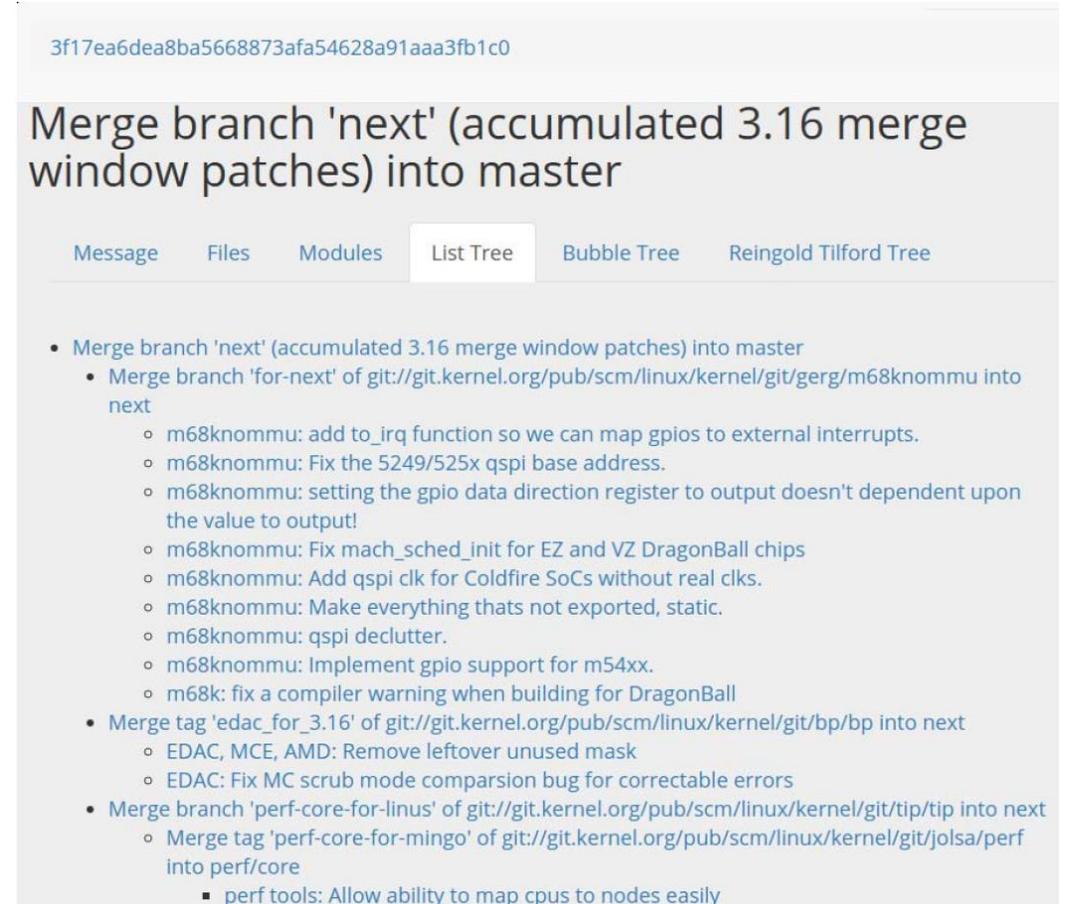
Showing 1 to 4 of 4 entries

Previous **1** Next

Fig. 11

# Linvis List Tree

- Text-based representations of the merge-tree
- Nested lists show the hierarchy
- Designed to model tree-views of file browsers which are familiar to developers
- Easy to search and navigate



3f17ea6dea8ba5668873afa54628a91aaa3fb1c0

Merge branch 'next' (accumulated 3.16 merge window patches) into master

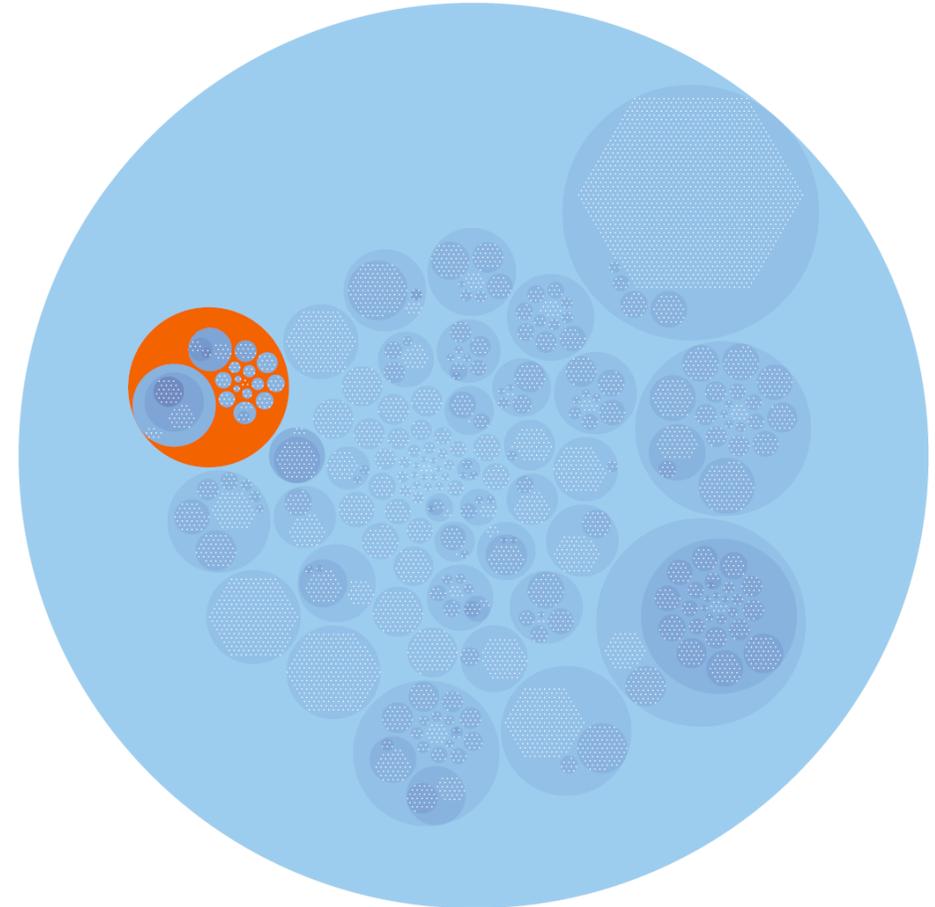
Message Files Modules **List Tree** Bubble Tree Reingold Tilford Tree

- Merge branch 'next' (accumulated 3.16 merge window patches) into master
  - Merge branch 'for-next' of [git://git.kernel.org/pub/scm/linux/kernel/git/gerg/m68knommu](https://git.kernel.org/pub/scm/linux/kernel/git/gerg/m68knommu) into next
    - m68knommu: add to\_irq function so we can map gpios to external interrupts.
    - m68knommu: Fix the 5249/525x qspi base address.
    - m68knommu: setting the gpio data direction register to output doesn't dependent upon the value to output!
    - m68knommu: Fix mach\_sched\_init for EZ and VZ DragonBall chips
    - m68knommu: Add qspi clk for Coldfire SoCs without real clks.
    - m68knommu: Make everything thats not exported, static.
    - m68knommu: qspi declutter.
    - m68knommu: Implement gpio support for m54xx.
    - m68k: fix a compiler warning when building for DragonBall
  - Merge tag 'edac\_for\_3.16' of [git://git.kernel.org/pub/scm/linux/kernel/git/bp/bp](https://git.kernel.org/pub/scm/linux/kernel/git/bp/bp) into next
    - EDAC, MCE, AMD: Remove leftover unused mask
    - EDAC: Fix MC scrub mode comparsion bug for correctable errors
  - Merge branch 'perf-core-for-linus' of [git://git.kernel.org/pub/scm/linux/kernel/git/tip/tip](https://git.kernel.org/pub/scm/linux/kernel/git/tip/tip) into next
    - Merge tag 'perf-core-for-mingo' of [git://git.kernel.org/pub/scm/linux/kernel/git/jolsa/perf](https://git.kernel.org/pub/scm/linux/kernel/git/jolsa/perf) into perf/core
      - perf tools: Allow ability to map cpus to nodes easily

Fig. 12

# Linvis Bubble Tree

- Organizes the commits hierarchically by having the parent commit contain the child commits
- Similar to tree maps but clearly shows leaf commits
- Good for providing clear visualization of wide, hierarchical data



*Fig. 15*

# Linvis Reingold-Tilford Tree

- Intuitive representation of merge-tree
- Not effective at display large trees

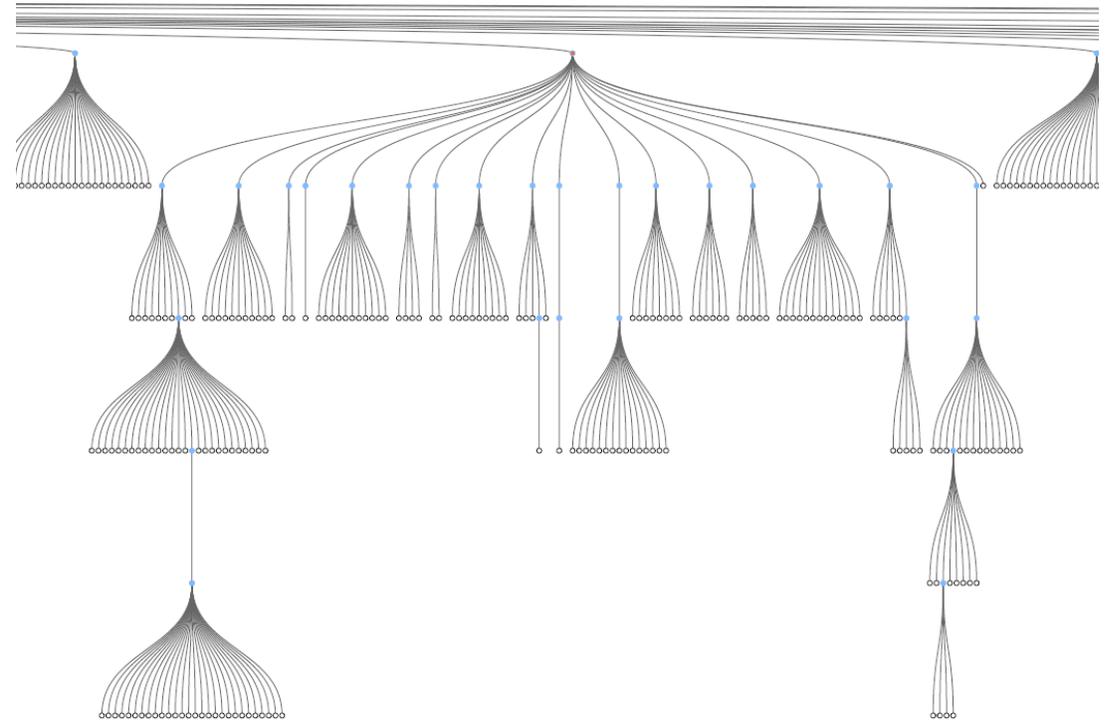


Fig. 13