

Collabora Online Usability Optimization

Smooth & beautiful,
collaborative editing for all

Michael Meeks

CEO

michael.meeks@collabora.com

and Caolán McNamara

Principle Engineer

caolan.mcnamara@collabora.com

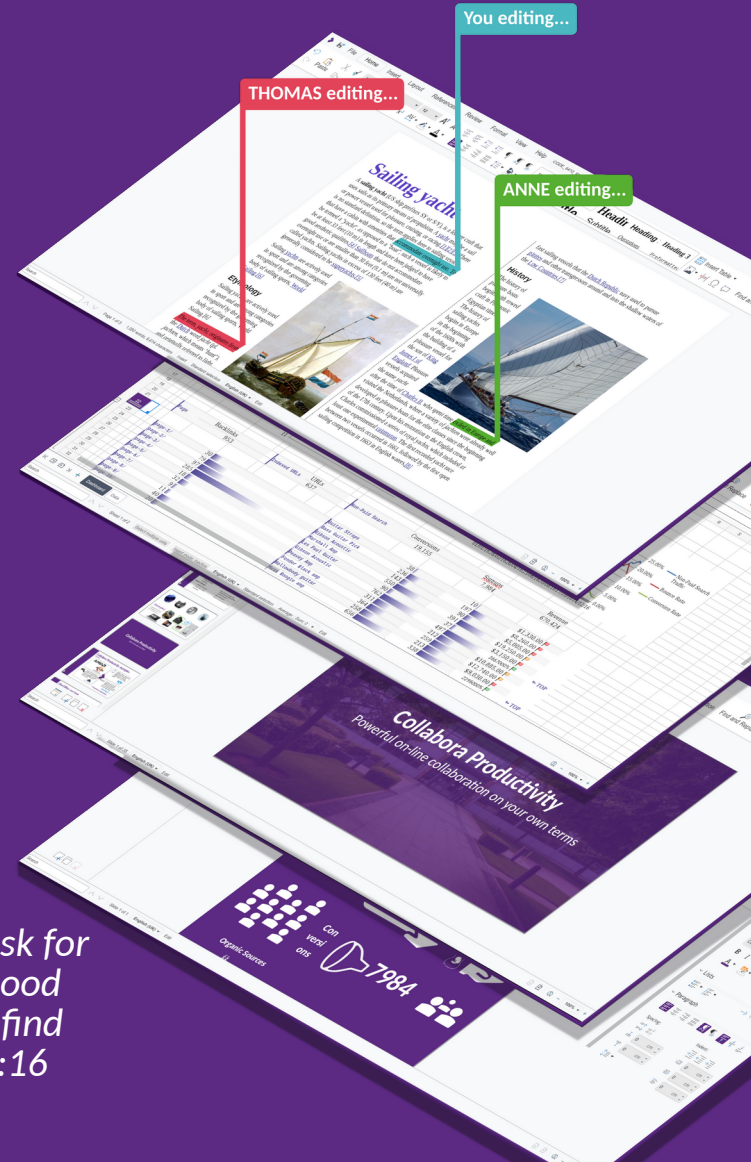


Collabora
Online



FOSDEM

“Stand at the crossroads and look; ask for the ancient paths, ask where the good way is, and walk in it, and you will find rest for your souls...” - Jeremiah 6:16





Overview:

How Collabora Online works

- Getting a feel for latency

How we measure performance

Typical wins

- Deltas
- Tile rendering
- Memory
- Miscellaneous sillies

Usability features

- Accessibility improvements
- The latest bits for everyone.

Future work

How to get involved

Conclusions



LibreOffice Technology

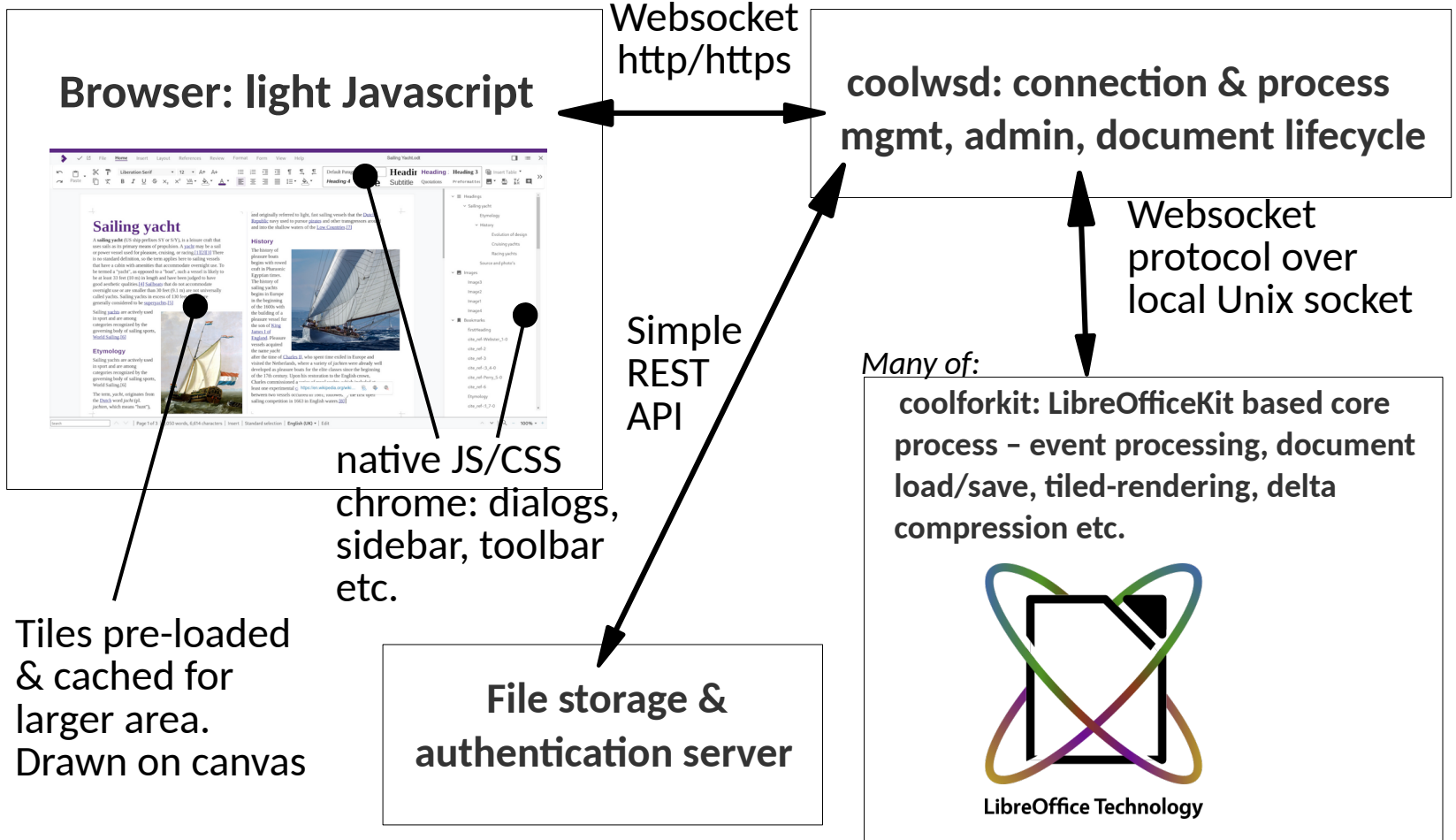


How Collabora Online works:

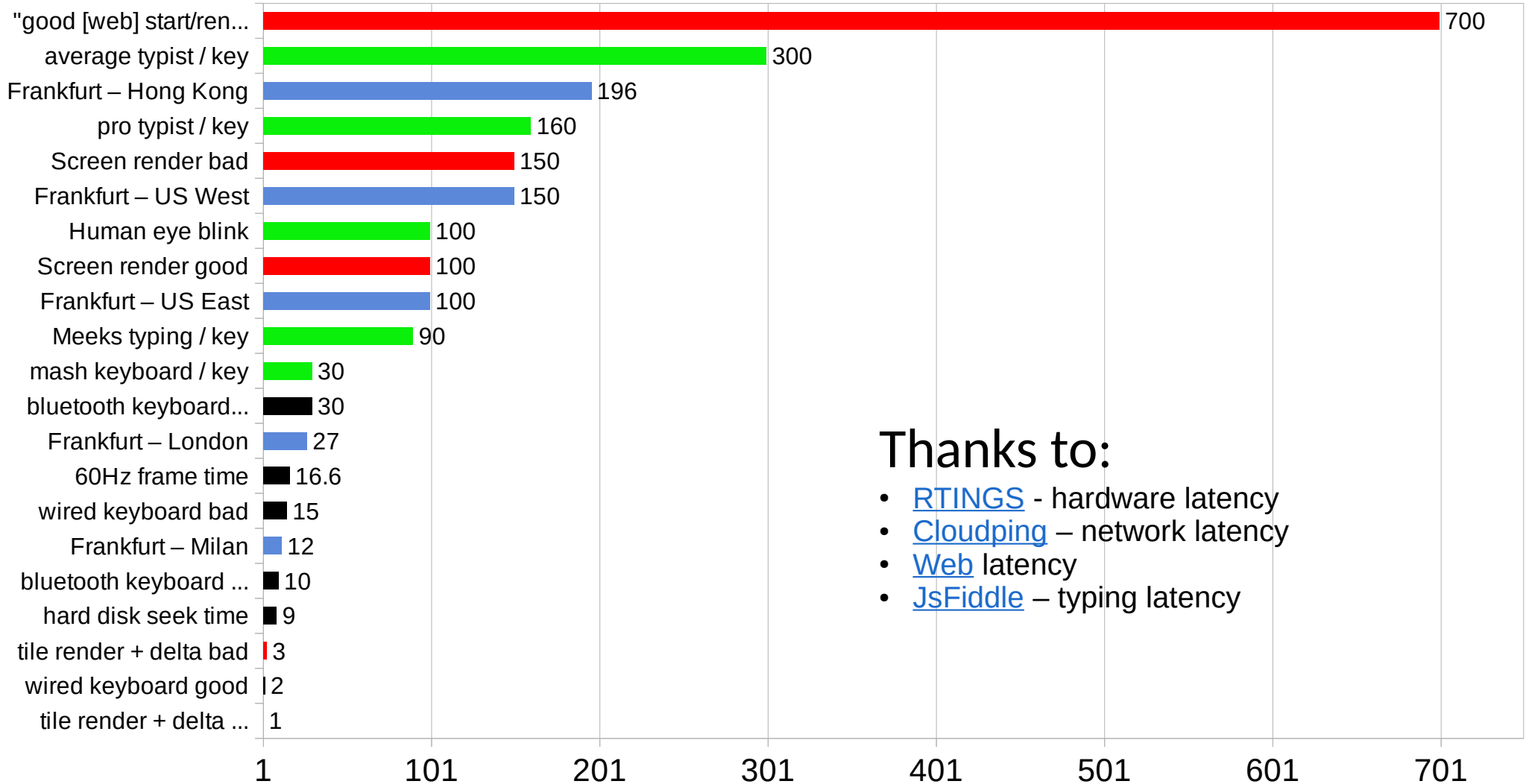




The various pieces:



Sample latencies - Milliseconds - linear plot



Thanks to:

- [RTINGS](#) - hardware latency
- [Cloudping](#) – network latency
- [Web](#) latency
- [JsFiddle](#) – typing latency



**Start counting blinks when
you see green ...**

Stop when you see red.



LibreOffice Technology



How did you score ?

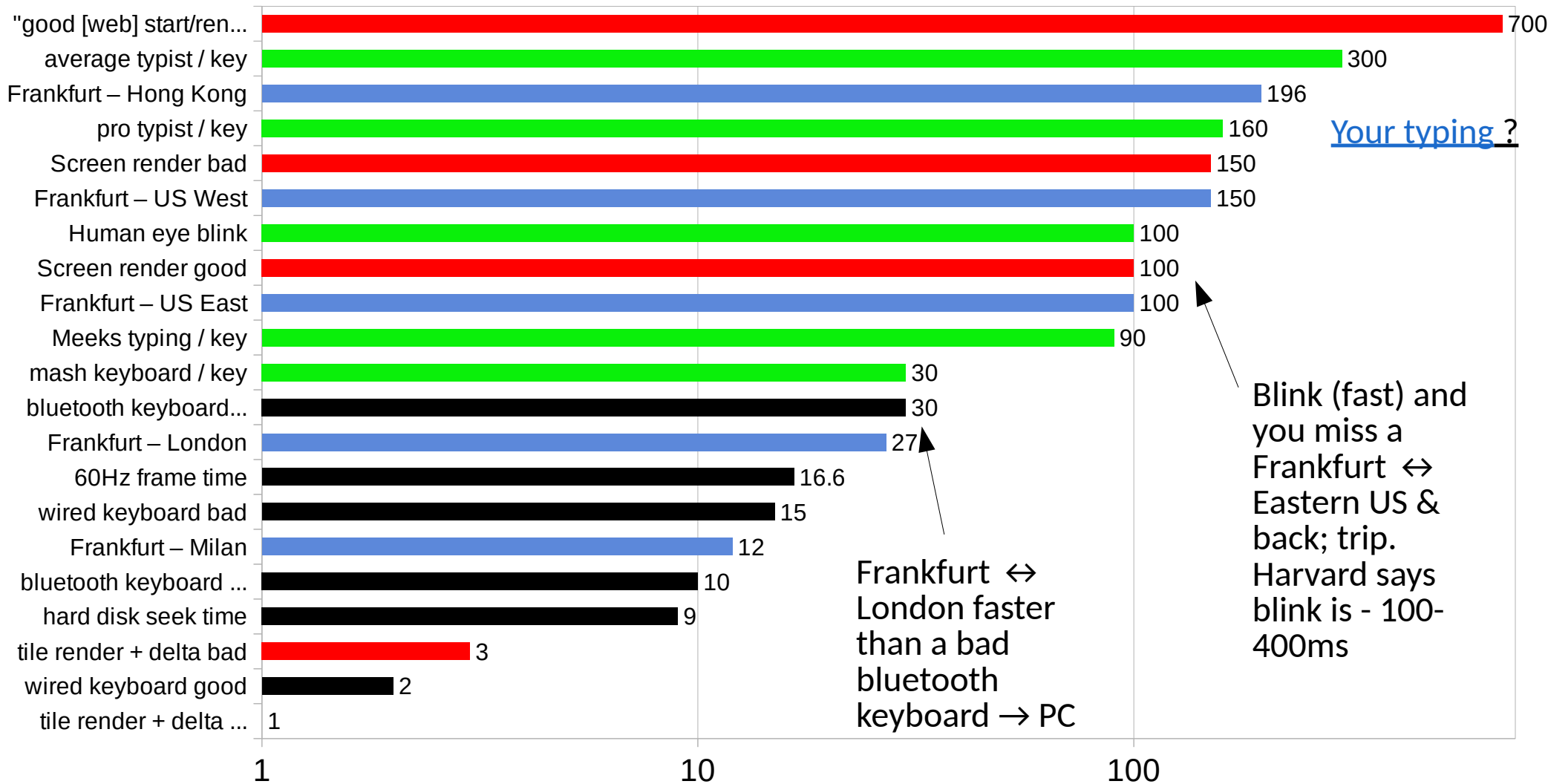
Reciprocation for beginners:

- 1 → 1000ms per blink
- 2 → 500ms per blink
- 5 → 200ms per blink
- 6 → 167ms per blink
- 7 → 143ms per blink
- 10 → 100ms per blink
- ~130 → ~7.7ms (peregrine falcon)



LibreOffice Technology

Sample latencies - Milliseconds - log plot





Measuring Performance

Don't optimize before profiling





Profiling & testing

Demo servers

- Optimizing what people do when they try a demo
- Sample once per second for a week & flamegraph

Internal Collabora users

- daily real-world use
- similar slow burn profiles

Multi-user testing

- Do we see lots of flashing red invalidation rectangles ?
- How does it “feel” in our community call with ~20 people.
- Profiling interactive stress testing.

End to end tracing tooling

- Tracking & aligning times from the three involved processes ...



LibreOffice Technology



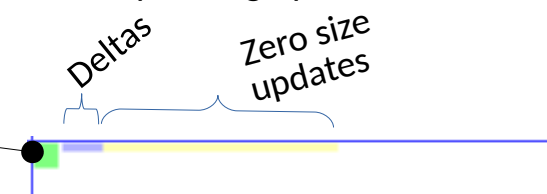
Interactive debugging

Help → About → Triple click.

- Lots of useful data
- Invalidation areas linger as red surrounds after a block red flash.
- If the screen flashes red:
 - Something is wrong

Status
 Green: valid
 Yellow: pending
 Red: invalid

In-tile transparent graph of recent history





Demo profiling: example flame-graph

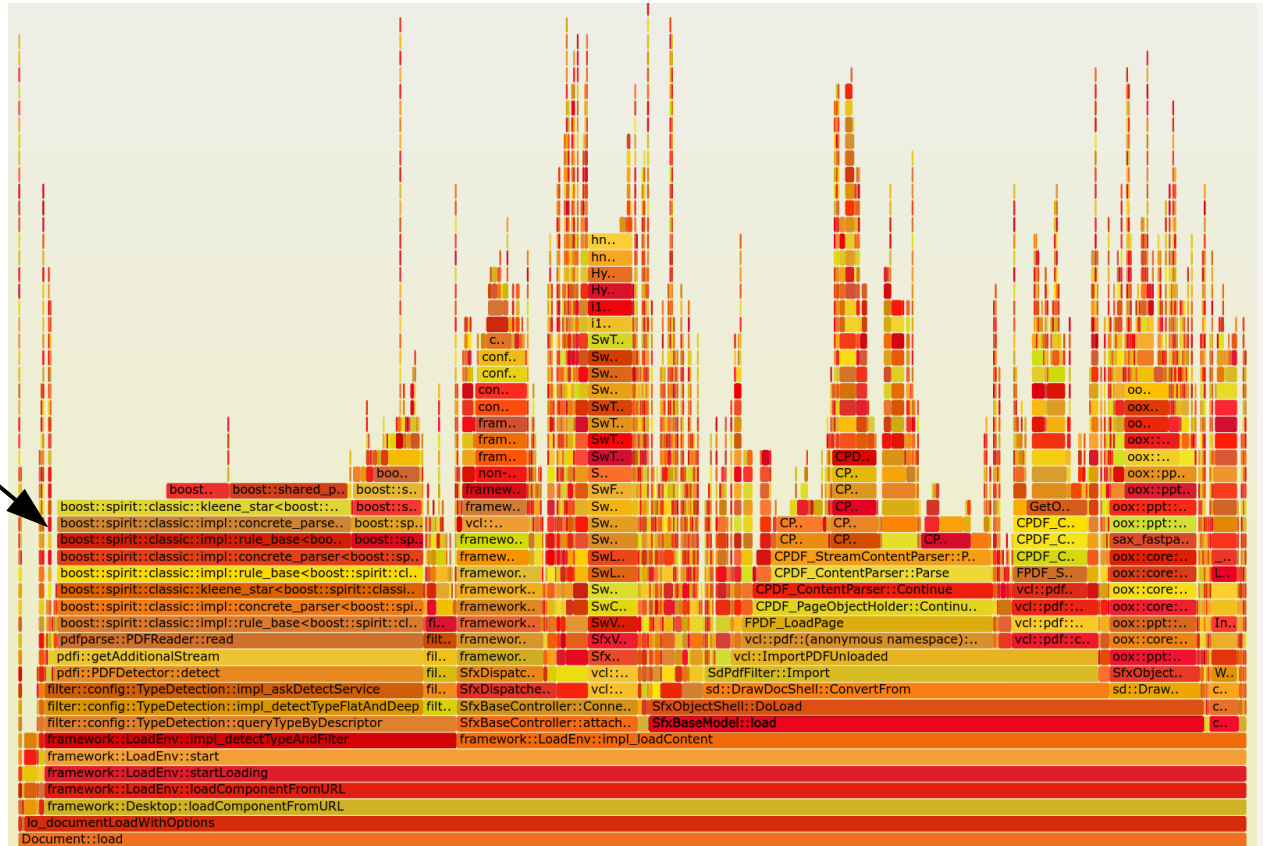
What does it mean ?

← Width is proportional time →

.17%+ of a week of profiling:

- Detecting: "is it Hybrid PDF"?
- Unbelievably wasteful
- Scan last 4k block 'AdditionalStreams' +17%

Stack of function callers



LibreOffice Technology



Tiles & Deltas





Tile Deltas cache / optimization

Store previous tiles

- So we can delta them
- Previously – generated row CRC while copying & kept all pixels
- Now use RLE bitmask.
- Substantially compressed: 256x256x4 → 256 kB
- RLE compressed: < 26 Kb – 10x size win
- **90** tile cache (per view) vs. **24** (per view)
 - Plus ~2Mb per view size saving.

RLE DeltaBitmapRow:

```
uint64_t  _rleMask[4];  
size_t    _rleSize;  
uint32_t *_rleData;
```

- Split mask bits from Data
- `_rleMask` bit-set – ‘1’
 - copy previous pixel 0 default transparent
- No need for a hash anymore: just compare `_rleSize` & `_rleMask`.



LibreOffice Technology



The magic of AVX2 - branch free loop

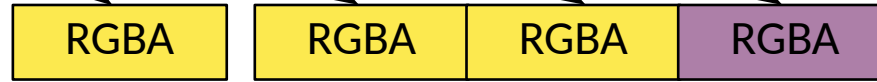
prepend 'last' pixel of the last block

Block of next four pixels to RLE ...

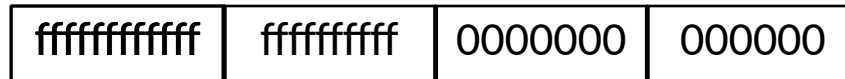


Next time's 'last' pixel:

Build comparison register:

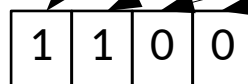


Compare:
cmpeq_epi32



NB. really 8 pixels at a time, not four and more 0's and f's needed ...

Magic: floating point sign mask:
movemask_ps



This is our RLE mask.

How many pixels to copy? popcount(RLE mask)
Which ones? RLE Mask → LUT + AVX2 gather
permutevar8x32_epi32 ... is your friend

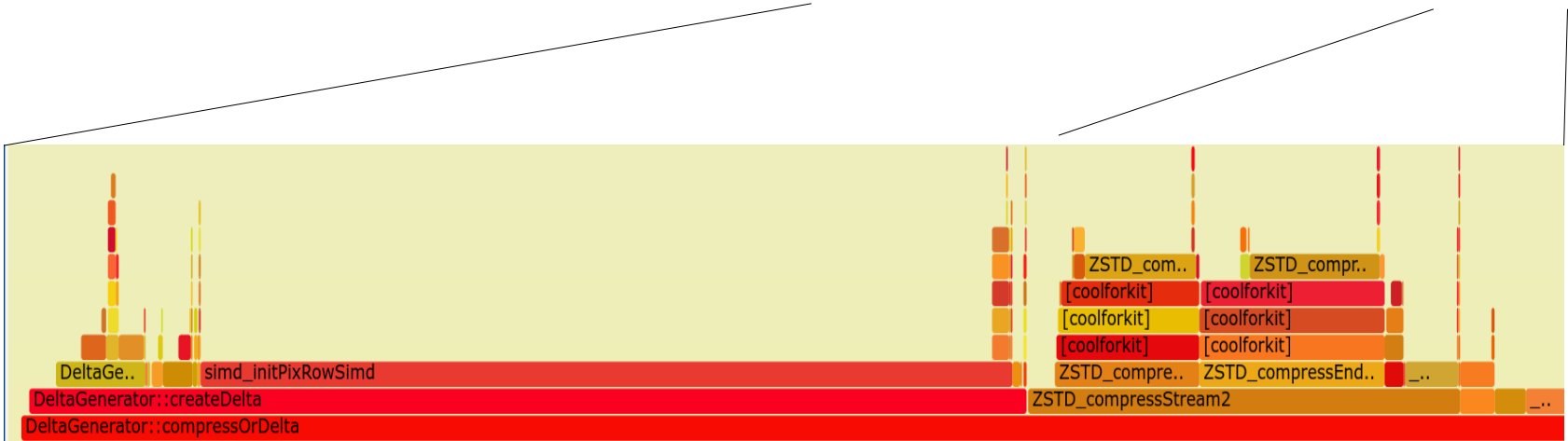
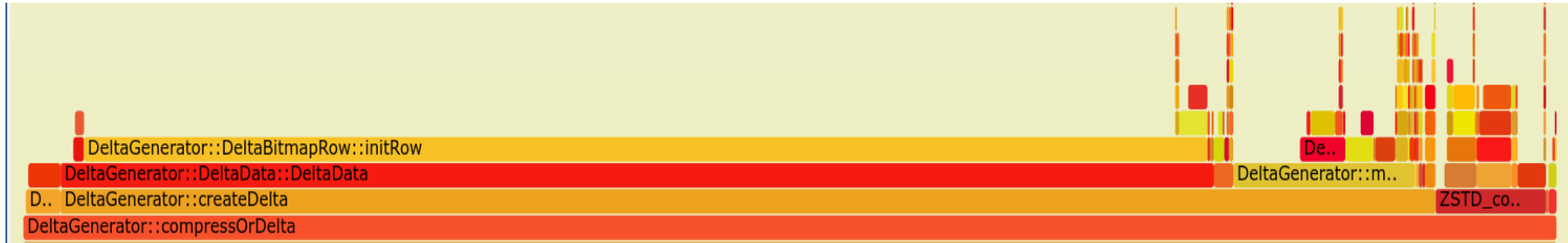


LibreOffice Technology



Performance win - around 2x ...

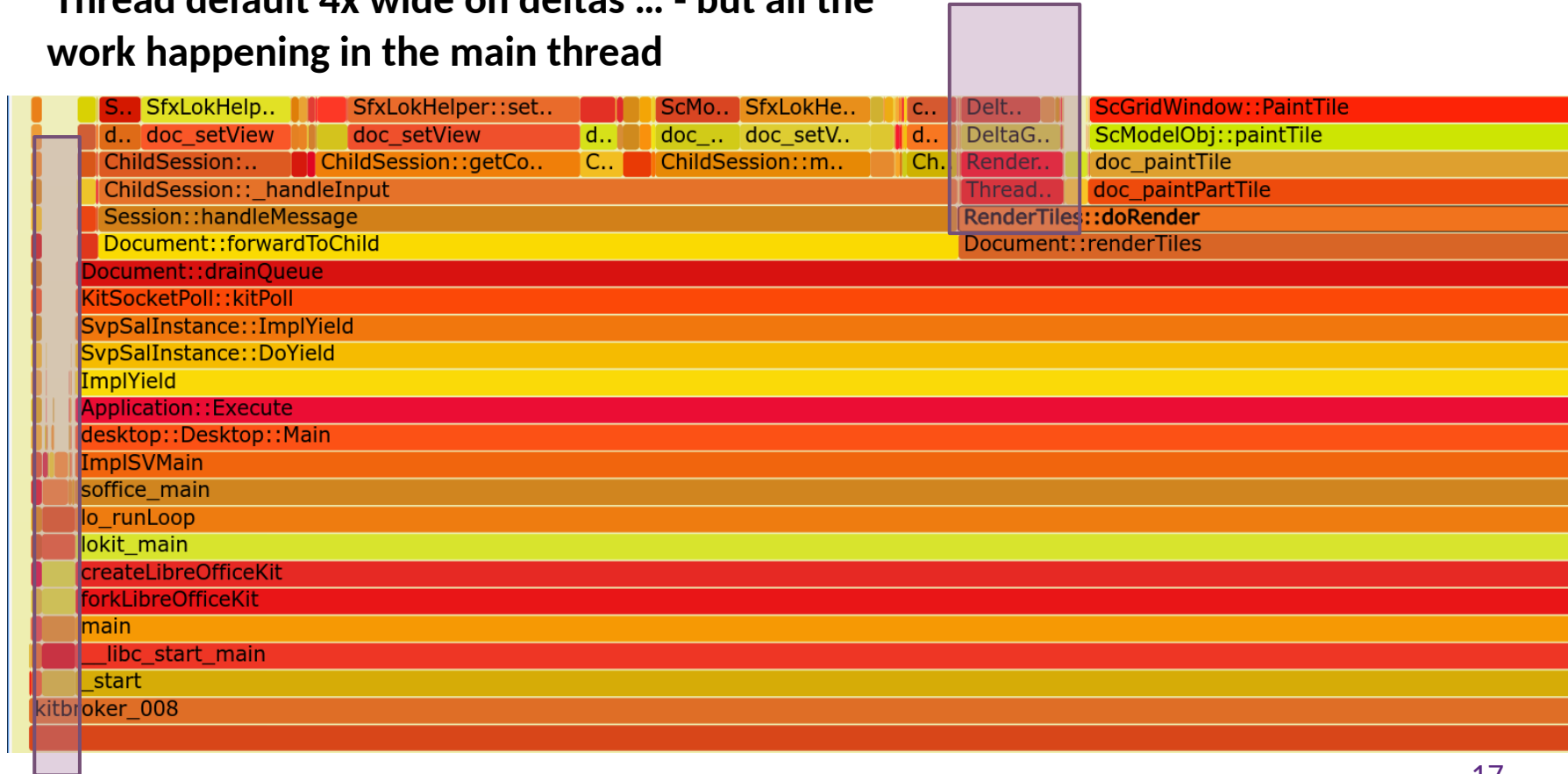
Comparing vs. best hand-optimized CPU RLE code ...





perf: surprisingly little delta threading

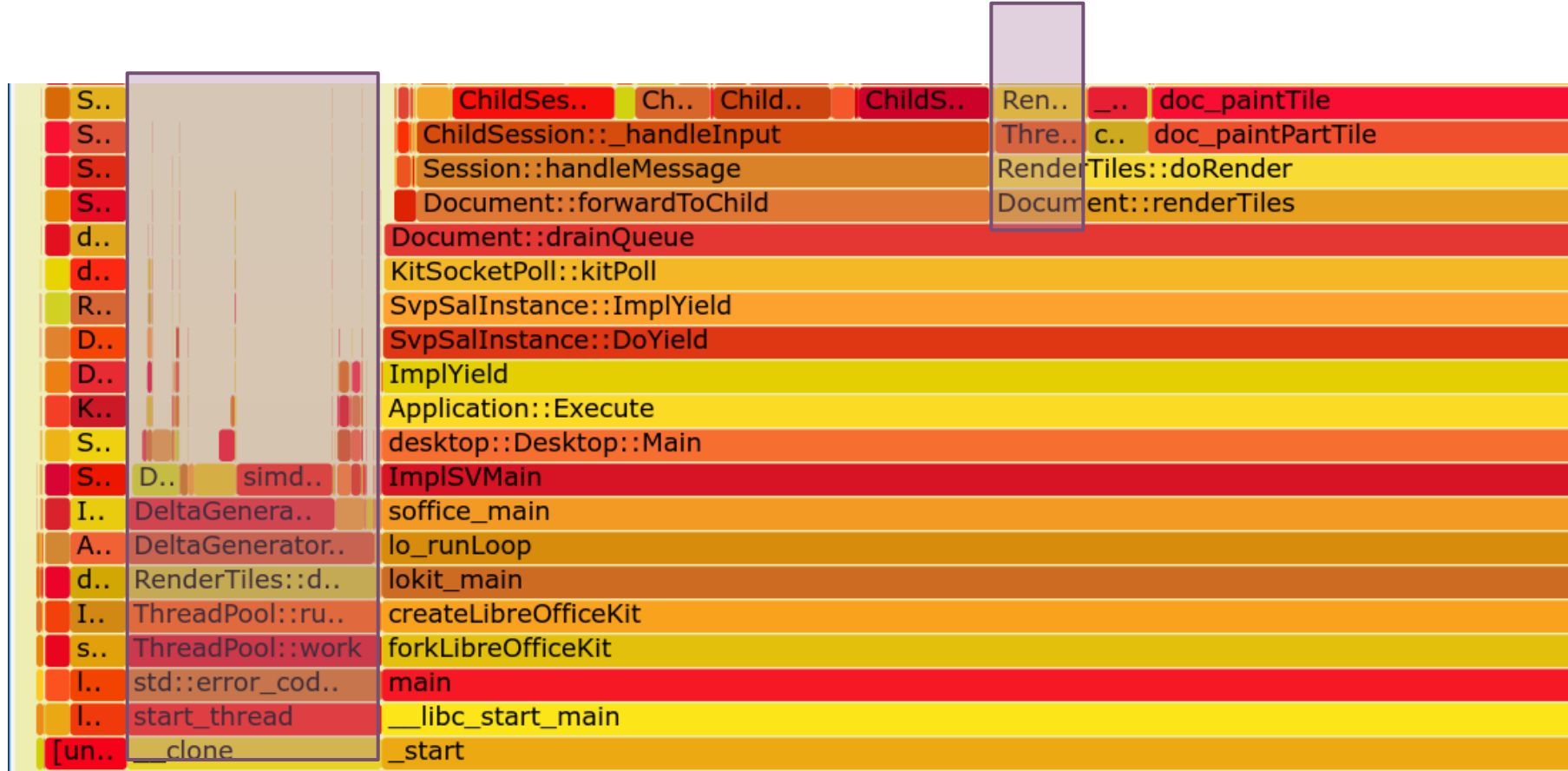
Thread default 4x wide on deltas ... - but all the work happening in the main thread





perf: surprisingly little delta threading

An 'if (work)' instead of 'while (work)' 1 line fix 4x latency redux in delta'ing.



LibreOffice Technology



RGBA & pre-multiplied alpha

Documents rendered to an alpha surface

- Pre-multiplied the sensible way to go so of course:
 - HTML5 canvas API – not pre-multiplied
 - HTML5 canvas implementation – pre-multiplied [!]
 - cf. [complaints](#) about not getting back RGBA you put into it ...

Change the approach and win

- COOL → un-premultiply → **net** → canvas API → re-pre-multiply → graphics
- COOL → **net** → un-premultiply → canvas API → re-pre-multiply → graphics
 - Leave the web's problems to the [browser](#) JIT.
- Also RGBA support to Cairo from libpixmap, to avoid BGRA conversion.



LibreOffice Technology



Tile Rendering





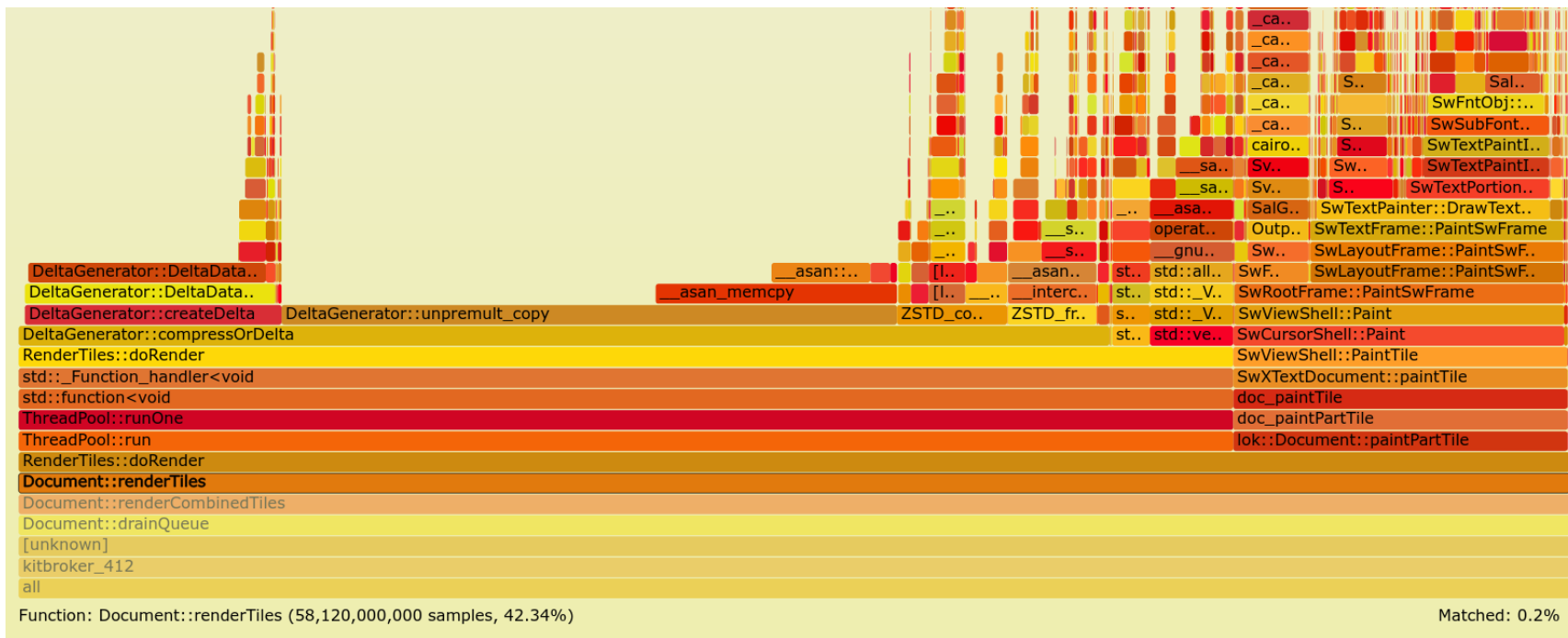
Rendering tiles 40% of your profile ?

Plausibly could be good or is it bad ?

- How does it feel ?

With lots of bogus invalidations

- you see lots of re-rendering ...





Lets do a lot less invalidation / render

Desktop app: “Rendering is free”

- COOL: “that’s not cool!”

Avoiding pointless invalidations:

- Whole doc when user joins document
- Whole doc on switching slides
- Whole doc on enlarging call sheet
- Whole doc on click in header/footer
- Entire row when editing calc cell

Saving huge amounts of rendering ...

- Really impactful on latency
- Dropping unnecessary 100ms waits is good ...
- Wasted rendering is cheap bandwidth wise: empty deltas ...



LibreOffice Technology



Better Latency Hiding

More aggressive Javascript tile caching

- An old tile is better than nothing
 - 150 - 250 tiles as canvases (30-60Mb)
 - Shepherd canvas memory better
 - JS 'GC' is not your friend; need to explicitly memory manage these.

More aggressive pre-fetching

- Next Previous / Next Slide in direction of movement - 100ms after switch
- Fetching and caching around the view area
- Tracking global invalidations to manage larger cache properly.

Store & manage zstd compressed tiles

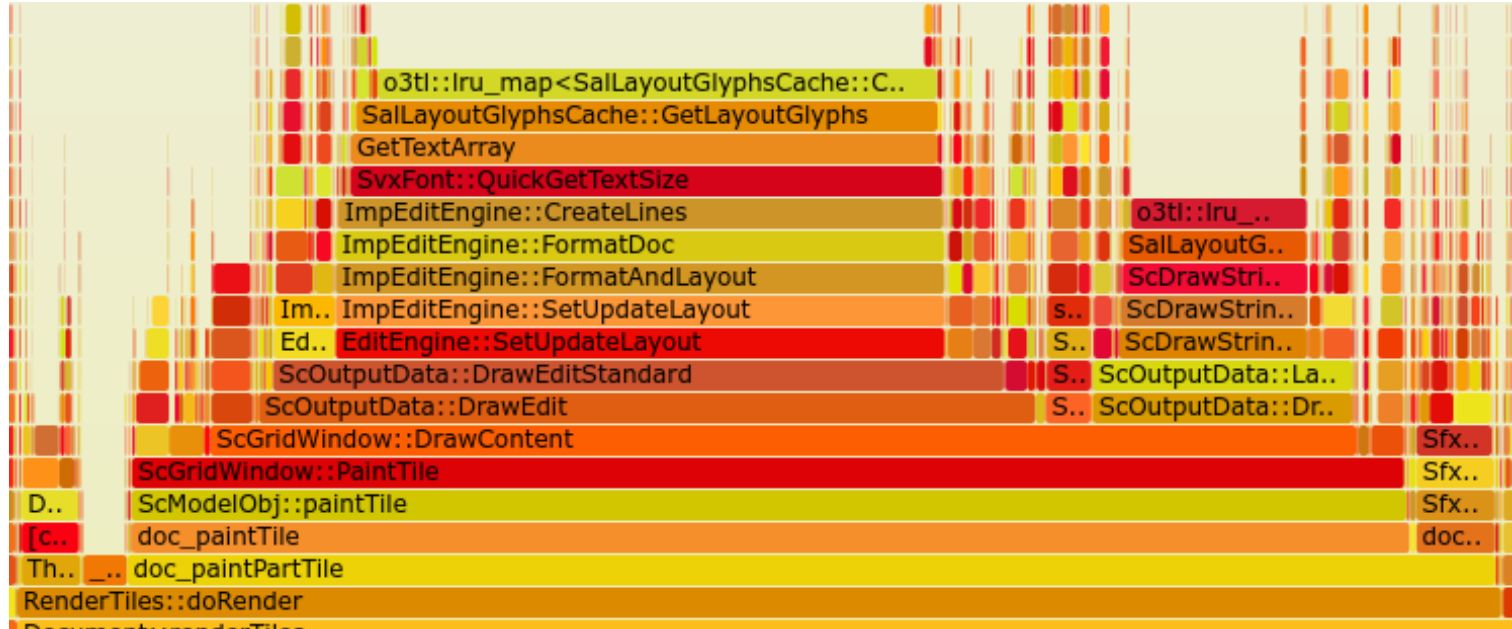
- 1000 - 2000 tiles zstd compressed
 - Keyframes + Deltas



LibreOffice Technology



LRU: std::list::size as std::distance()



Symbol	cpu-clock:ppp (incl.)	cpu-clock
SvxFont::QuickGetTextSize(OutputDevice const*, rtl::OUString const&, int, int, KernArray*) const	24.6%	0.00776%
GetTextArray	24.4%	0%
SalLayoutGlyphsCache::GetLayoutGlyphs(VclPtr<...>, rtl::OUString const&, int, int, long, vcl::te	24.2%	0%
o3tl::lru_map<...>::insert(std::pair<...>&&)	23.4%	0%
o3tl::lru_map<...>::checkLRUItemInsert()	23.4%	0%
std::list<...>::size() const	23.4%	0%
std::list<...>::_M_node_count() const	23.4%	0%
std::iterator traits<...>::difference_type std::distance<...>(std::list const iter	23.4%	0%





Kill paint to a giant virtual device

Older Writer rendering path:

```
//Refresh with virtual device  
to avoid flickering.
```

```
VclPtrInstance<VirtualDevice>  
pVout( *mpOut );  
pVout->SetMapMode( mpOut->  
GetMapMode() );  
Size aSize( VisArea().SSize() );  
aSize.AdjustWidth(20);  
aSize.AdjustHeight(20);  
if( pVout->SetOutputSize( aSize ) )
```

Un-necessary PC 'flicker reduction' optimization

- We push tiles to JS for a flicker-free scroll/zoom anyway.

Giant / whole document area

- Plus a bit.

← Back that with lots of memory & do lots of rendering into it



LibreOffice Technology



Memory use





Lots of space (& time) saving:

Discovered a lurking **benchmark**

- Allocate 64Mb of RAM, and performing a CPU rendering benchmark before loading each document ...
- Good to get initial dirty-page count down to ~20Mb in one line.

Image caching

- Compressed images are small!
 - Not so TIFFs → swap them.
- Cache & Images & glibc allocator trim on idle → mobile-phone style.

Sparse documents:

- Calc – file save used to allocate all 16k columns – making many things slower.
- Calc – discourage users to leap to limits of document

\$ make run-inproc

- Run under massif / valgrind as a single process in the build-tree ...
- Avoiding real-CPU timing jitter:
 - flat profiles for no change ... vital.



LibreOffice Technology

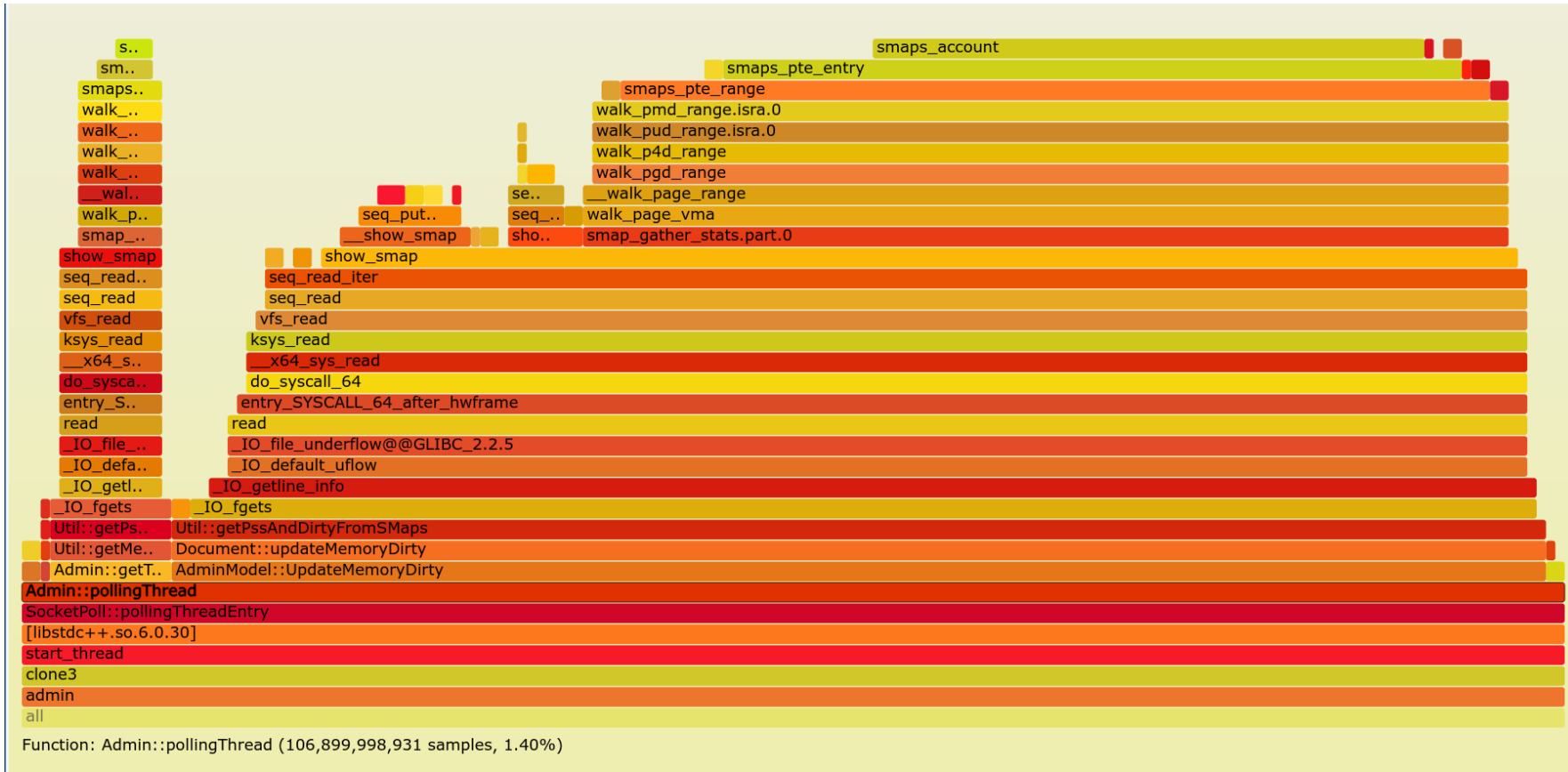


Misc. sillies





Week long demo profile: 1.4% time:



LibreOffice Technology



Kernel craziness: /proc/smmaps*

Need to gather PSS for each kit process

- We have 100+ memory maps
- /proc/self/smmaps fd passed by UDS between privilege domains to supervisor process.
- Have to rewind, can't close & re-open

/proc/smmaps_rollup joy !

- Does just what we want: better!
- But: **Rewind, re-read: constantly increases memory reported ...**
- Dynamic check for Linus' [horror bug fix](#)



LibreOffice Technology



Hunspell inner loop ...

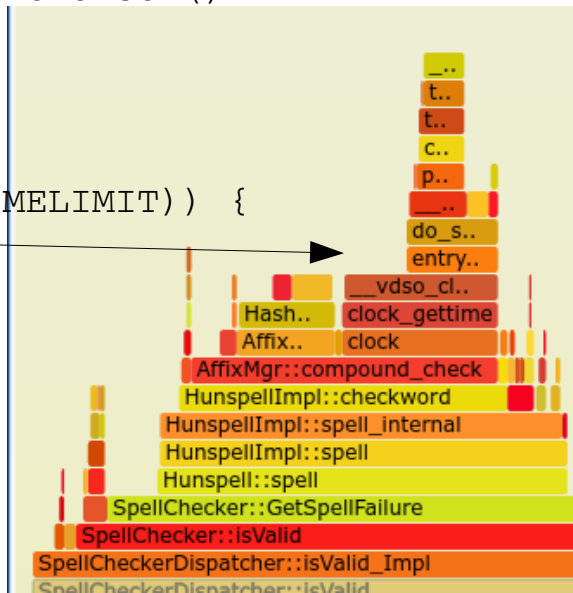
```
int AffixMgr::compound_check_morph(const std::string& word ...
// add a time limit to handle possible
// combinatorical explosion of the overlapping words
HUNSPELL_THREAD_LOCAL clock_t timelimit;
if (wordnum == 0) {
    // get the start time, seeing as we're reusing this set to 0
    // to flag timeout, use clock() + 1 to avoid start clock()
    // of 0 as being a timeout
    timelimit = clock() + 1;
}
else if (timelimit != 0 && (clock() > timelimit + TIMELIMIT)) {
    timelimit = 0;
```

...

Switch to shared memory CLOCK_MONOTONIC ?

```
auto clock_now = std::chrono::steady_clock::now();
```

- Thanks to Noel Grandin.





Performance

A faster, slicker experience

- Memory trimming on idle: cleans caches & frees up memory
- Swap out compressed image data as well as uncompressed
- Faster load of large spreadsheets
- Significantly reduced re-rendering (even more pending)
- Continual style re-thumbnailling redux – caching of generated JSON
- Improved tile pre-loading (with more caching work queued)
- Clamp over-sizing of threads
- Accelerated transparent text rendering
- Avoid background whole document renders
- Compress RLE'd tiles on the wire

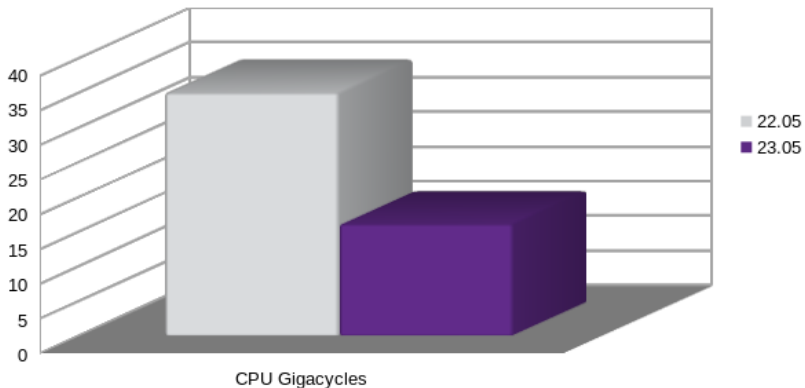


LibreOffice Technology



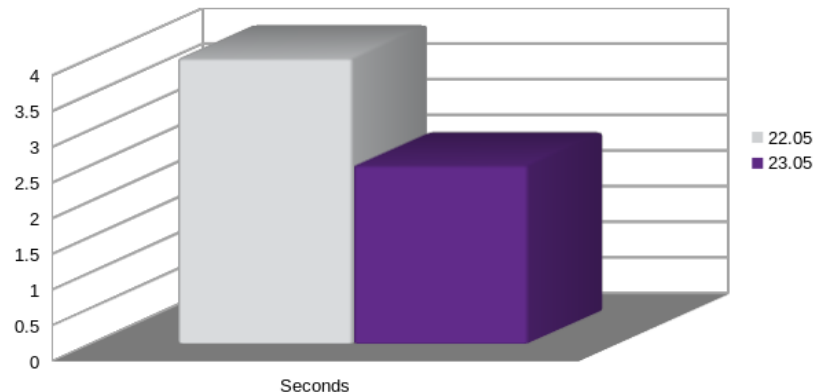
Joining Writer Shared Document

Writer document with 5 pages. 20 users joining simultaneously, until all clients rendered. perf stat -e cycles



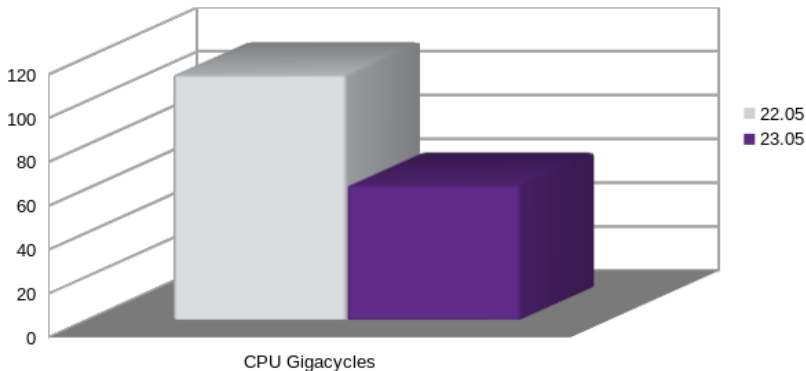
Impress Full Screen Presentation Generation

60 graphics heavy slides. Wall clock time.



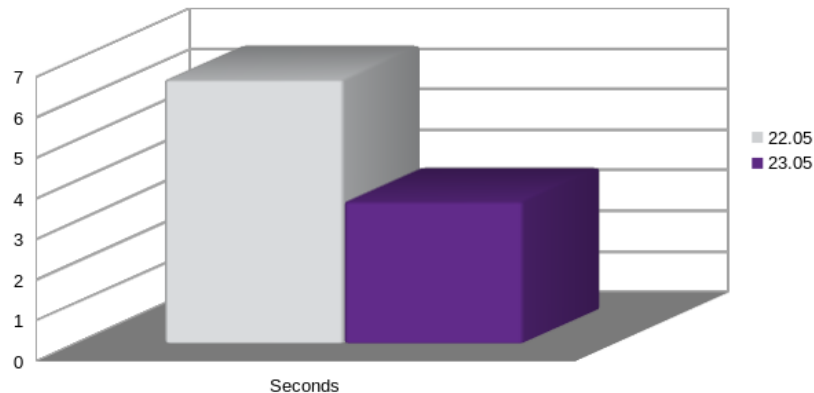
Writer Scrolling

Writer document with 157 pages, 164 shapes, 32 comments, 13 images, 115 tables. Wheel scrolled from start to end. perf stat -e cycles



Calc XLSX Comments Load

1000 Rows, 30 Columns, 2500 Comments. Wall clock time.



LibreOffice Technology

Some older numbers ...

Smaller is better!



Usability & UX

Making things easier to use ...





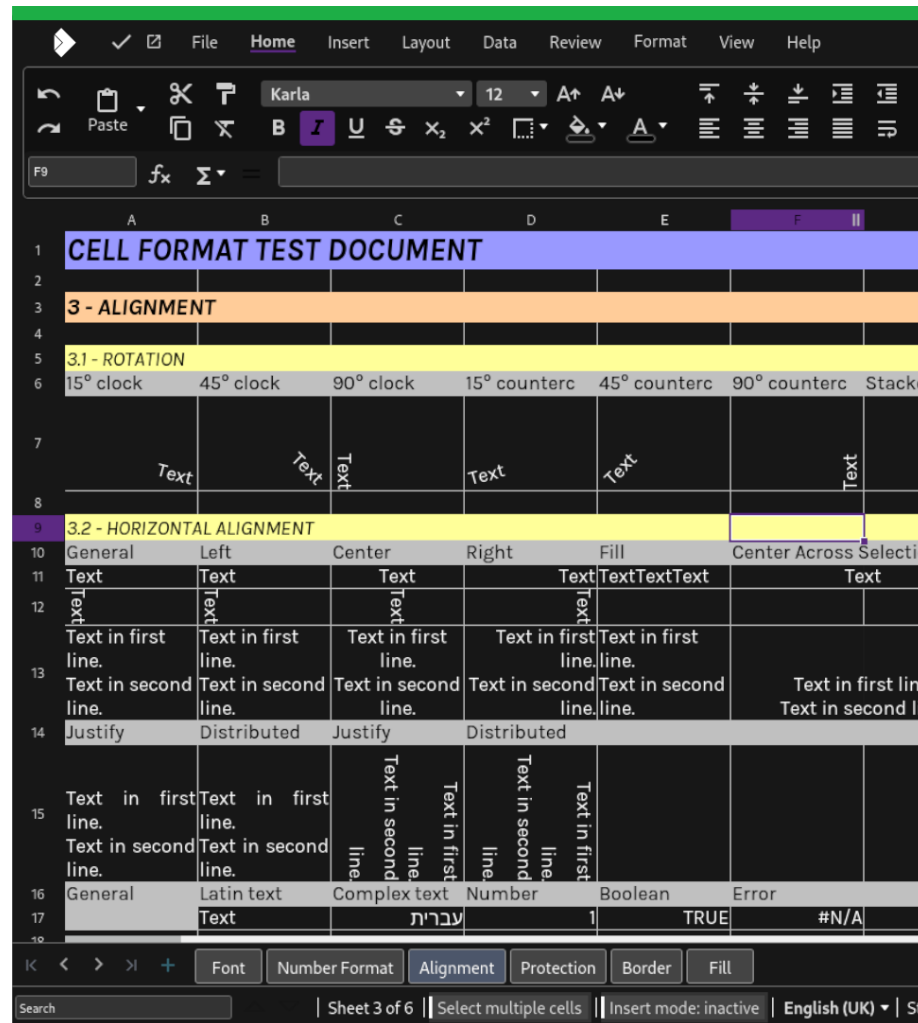
Accessibility

Dark mode + view settings

- Per view settings, allows us to render views differently – including various options such as showing non-printing characters or spell checking underlines in different views, in addition to Dark Mode.



LibreOffice Technology





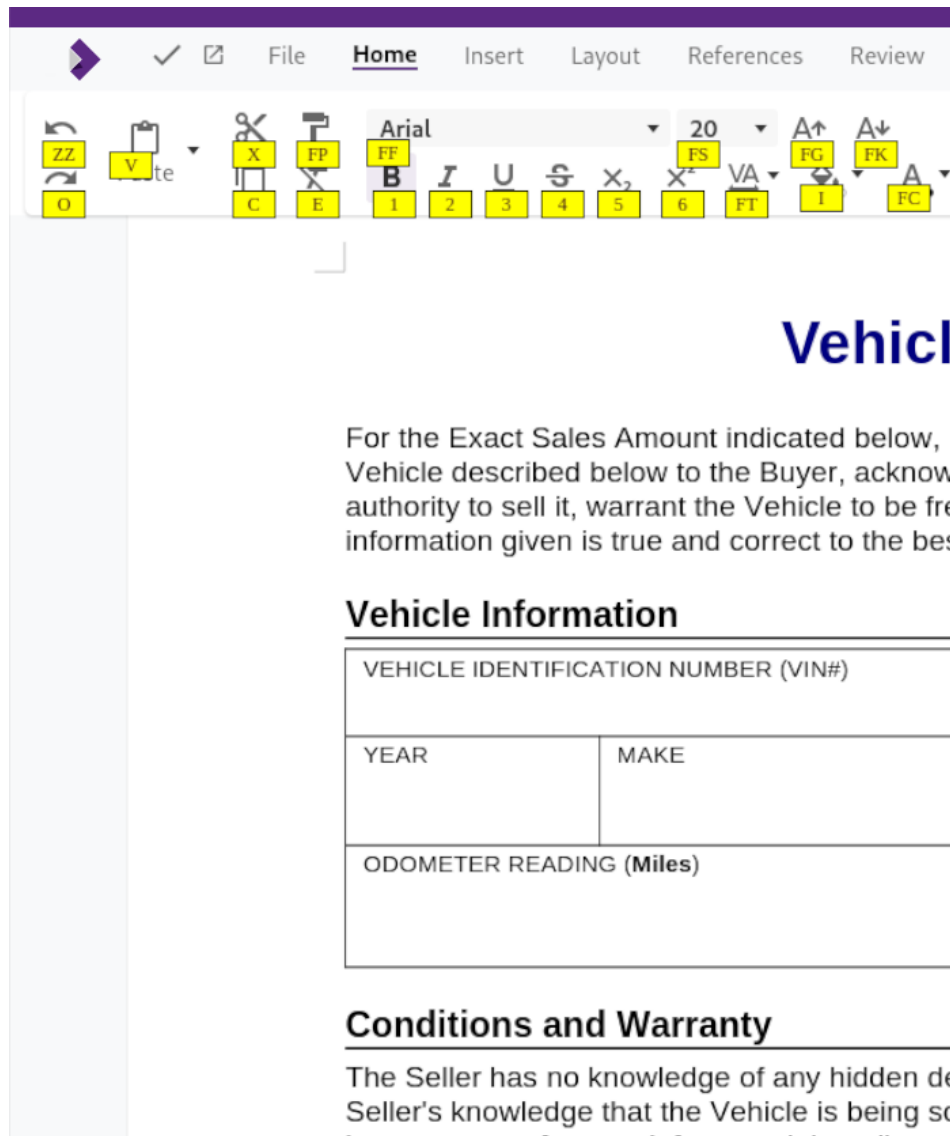
Accessibility

Keyboard accelerators

- Expansion for languages
 - Keyboard shortcuts have been expanded to allow all users to be able to use their own language's keyboard shortcuts when editing a file collaboratively.
- Holding down the 'Alt' key will also highlight options for learnability.



LibreOffice Technology



The screenshot shows the LibreOffice Writer interface with a document titled "Vehicle Information". The ribbon is set to "Home". The document content includes a paragraph of text, a section titled "Vehicle Information" with a table, and a section titled "Conditions and Warranty".

Vehicle Information

VEHICLE IDENTIFICATION NUMBER (VIN#)	
YEAR	MAKE
ODOMETER READING (Miles)	

Conditions and Warranty

The Seller has no knowledge of any hidden defects...

Collabora Online.odt

File Home Insert Layout References Review Format Form View Help

Karla 12 A+ A+ [List Icons] [Text Icons] [Paragraph Icons] Default Paragr Text Body **Heading** **Heading** : **Headin**

ste [Copy] [Paste] [Undo] [Redo] B I U [Link] [Image] [Table] [Text Color] [Background Color] [List Icons] [Text Icons] [Paragraph Icons] **Heading 4** **Title** Subtitle Quotations Preform

title formats and collaborative editing features.

Collabora Online comes with **Long Term Support**, and can be integrated in any web-based solution, such as File Share and Sync solutions, Document Management Systems, Groupware or custom build software.

All Long Term Support releases are supported for 3 years, with a **perpetual license**, annual subscription to **security maintenance and updates**, an **SIA**, and Level 3 (code fix) support (other levels are provided by our partners). And there are more benefits:

Excellent Interoperability

- Collabora Office Mobile allows editing of the following formats:
- Word/Writer documents: .doc, .docx, .dot, .dotx, .odt, .ott,
- Excel/Calc spreadsheets: .xls, .xlsx, .xlt, .xltx, .ods, .ots,
- Powerpoint/Impress presentations: .ppt, .pptx, .pps, .ppsx, .odp, .otp
- /Draw image creation:



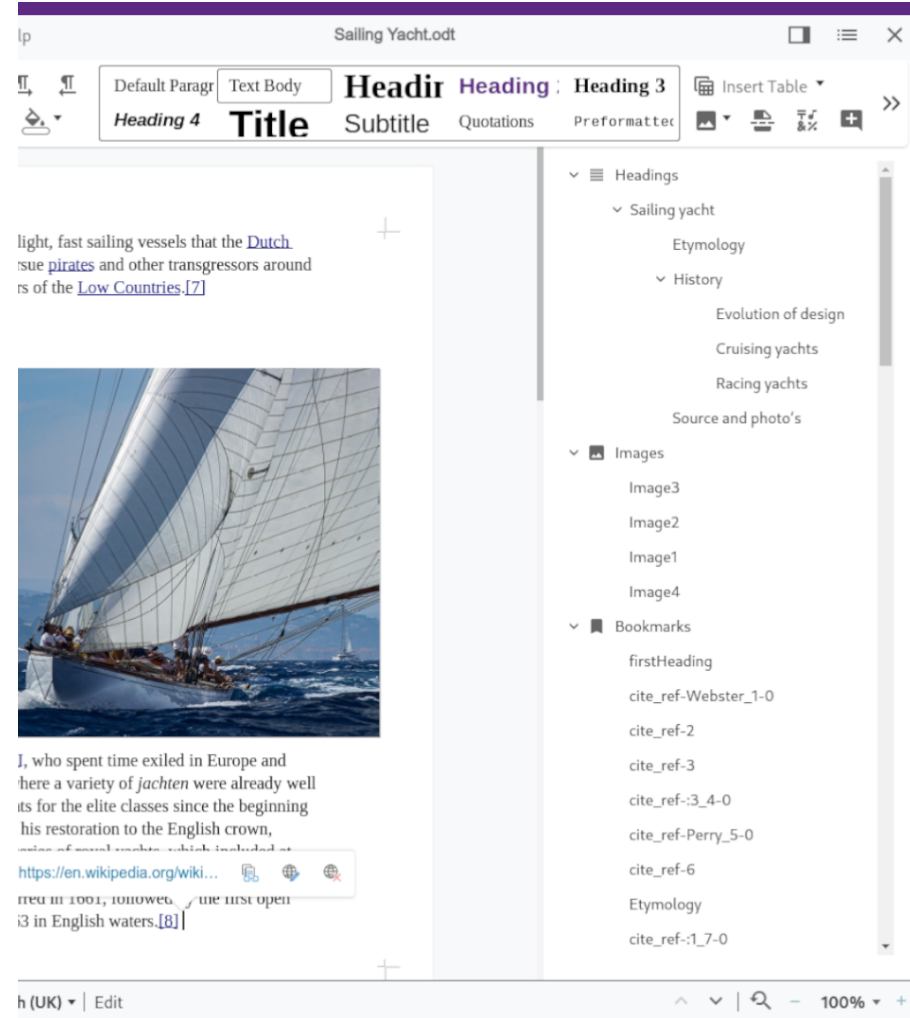
Accessibility

Page Navigation

- Navigator functionality is viewable in the sidebar, allowing you to jump to each section by simply clicking the headings.



LibreOffice Technology



The screenshot shows a LibreOffice Writer window with the title bar 'Sailing Yacht.odt'. The top toolbar includes options for text formatting, such as 'Default Paragraph', 'Text Body', 'Heading 3', 'Heading 4', 'Title', and 'Subtitle'. The main document area contains text: 'light, fast sailing vessels that the Dutch issue pirates and other transgressors around of the Low Countries.[7]'. Below the text is a large image of a sailboat on the water. The right sidebar is open to the Navigator pane, which is organized into three sections: 'Headings' (containing 'Sailing yacht', 'Etymology', and 'History'), 'Images' (containing 'Image3', 'Image2', 'Image1', and 'Image4'), and 'Bookmarks' (containing 'firstHeading', 'cite_ref-Webster_1-0', 'cite_ref-2', 'cite_ref-3', 'cite_ref-3_4-0', 'cite_ref-Perry_5-0', 'cite_ref-6', 'Etymology', and 'cite_ref-1_7-0'). The status bar at the bottom indicates 'h (UK) | Edit' and '100%' zoom.



Recent Usability Polish

Amazingly users when asked care mostly about UX efficiency

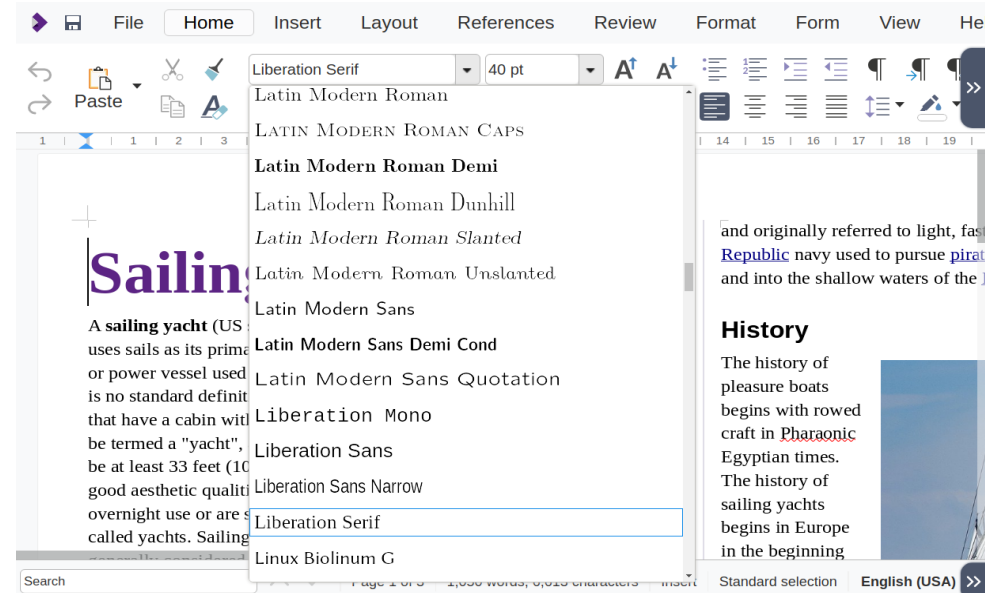




Usability

Font previews

- You can now see a preview of fonts to see what they look like before selecting it in the drop down box.





Usability

Page number insertion – familiar, simplified pop up

- Combining inserting the required headers, footers, and page number fields into a single, easy to use, familiar dialog for users.
- Featuring all the common alignment options, support for a variety of languages, and a preview.



students study for the same course regardless of which college they attend.[2] For postgraduate students, research is conducted centrally in the faculties, departments and other university-affiliated research centres, though the colleges provide a central social and intellectual hub for students.

Facilities

Colleges provide a range of facilities and services including catering, library facilities, extracurricular life at Cambridge is centred around college. Student activity is typically organised through postgraduate students. Another important element is the frequency from weekly to every night of the week.

Colleges also provide funding, accommodation, or both, for some of the academic posts in university, with the majority of Cambridge academics being a fellow of a college in addition to their faculty/departmental role.[3] Fellows may therefore hold college positions in addition to

1 https://en.wikipedia.org/wiki/University_of_Cambridge

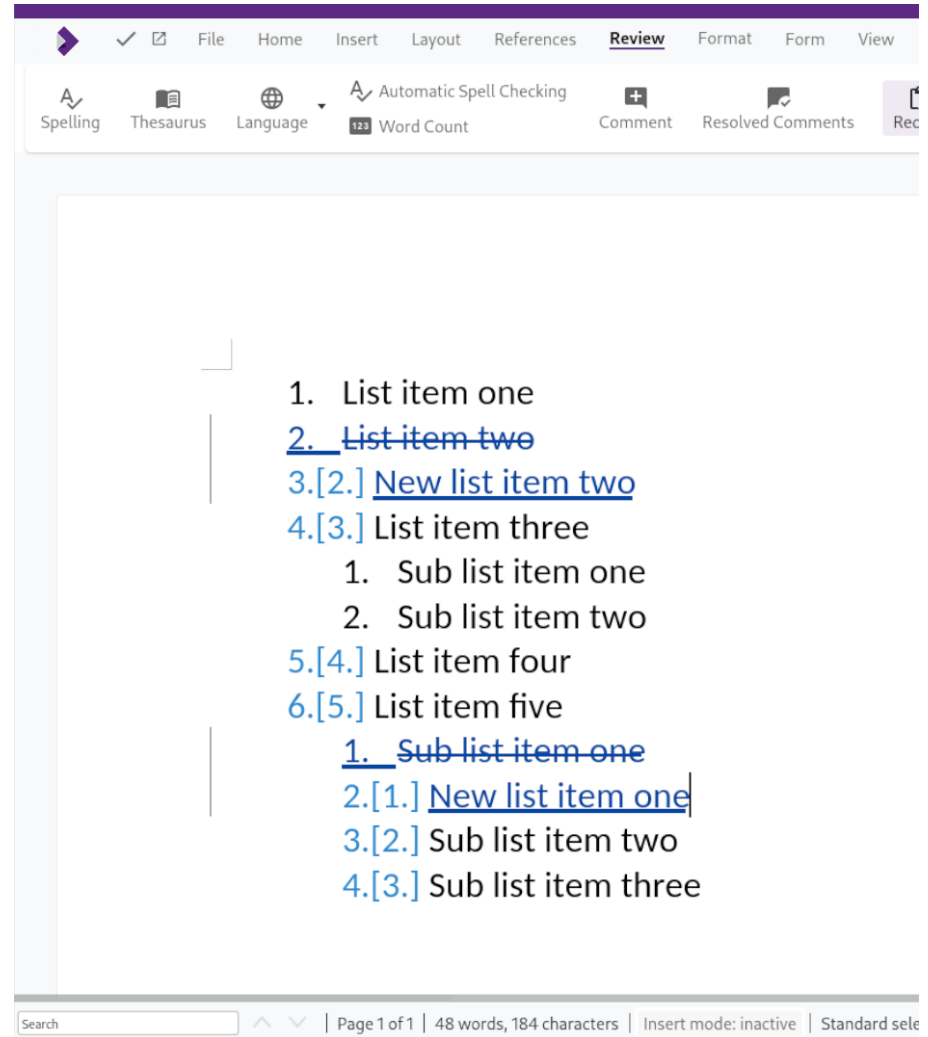
Page 3 of 4 | 9 words, 47 characters | Insert mode: inactive | Standard selection | English (UK) | Edit



Usability

Change tracking in numbered lists

- When making changes within numbered lists, the numbering has now been fixed to show actual and original numbers within the document.



File Home Insert Layout References **Review** Format Form View

Spelling Thesaurus Language Automatic Spell Checking Word Count Comment Resolved Comments

1. List item one
- ~~2. List item two~~
- 3.[2.] New list item two
- 4.[3.] List item three
 1. Sub list item one
 2. Sub list item two
- 5.[4.] List item four
- 6.[5.] List item five
 - ~~1. Sub list item one~~
 - 2.[1.] New list item one
 - 3.[2.] Sub list item two
 - 4.[3.] Sub list item three

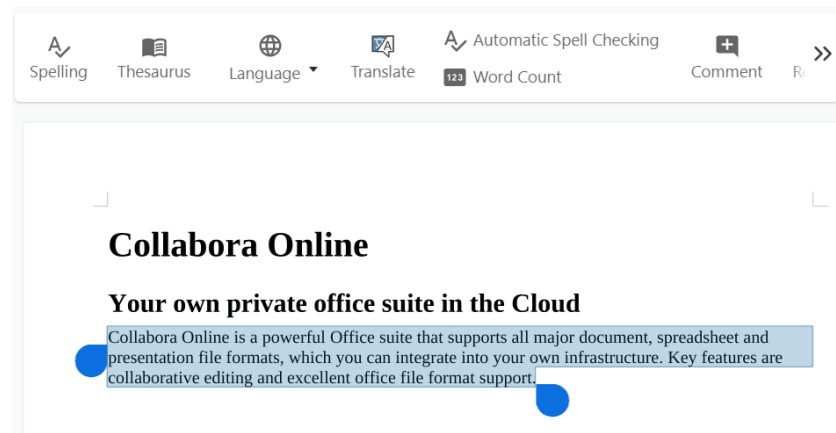
Search Page 1 of 1 | 48 words, 184 characters | Insert mode: inactive | Standard sele



Usability

AI based translations with DeepL

- Translating text with DeepL inside a Writer document is as easy as selecting it, clicking the “Translate” button and choosing the target language from dialog.





Future Performance optimization



LibreOffice Technology





Future tile wins:

Simple examples:

- Scope for improvement
 - eg. 'typing space' →
- Writer: adding page – invalidate only the end ...
- Delta & RLE only rows we know changed: →
- Writer: don't invalidate to the end of the row ... →

Tile update waste: 89%
New Tile ratio: 24%

wireId: 130051
nview
invalid
tile: 1
misse
dita s

Small area

wireId: 247
invalidFrom: 246
nviewId: 1001
invalidates: 84
tile: 2 A: 77 upd: 6

Invalidate before and to end of row

misses: 1 gce: 0 dta size/kB: 0.37
misses: 1 gce: 0 dta size/kB: 7.20
misses: 1 gce: 0 dta size/kB: 6.37
misses: 1 gce: 0 dta size/kB: 0.89
misses: 1 gce: 0 dta size/kB: 0.05



LibreOffice Technology



Other fun areas:

RGBA rendering

- Currently we render RGB
- Then re-render 'A'
- Then merge the two, then ...
 - Great potential 2x win ...
- Already moved to Opacity/Alpha not Transparency channel
 - Thanks to Noel Grandin

Performance regression testing

- Using valgrind to get flat lines ...

Pipelining loading

- Fetch cool.html ...
 - `<async CheckFileInfo>`
 - `<async document load>`
- cool.html's JS connects on websocket
 - gets a pre-loaded document.
 - `<currently we start loading here>`
- Halve the load time ?
- Patch in review from Ashod Nakashian.
- Also: to come: asynchronous locking ...



LibreOffice Technology



**With thanks to our Partners,
Customers & Community !**



LibreOffice Technology





Conclusions

Try out the latest Collabora Online 23.05.8+ and/or 24.04 ...

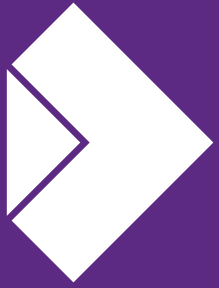
- Already smooth, getting smoother & more beautiful.
- Architecture: a bet on CPUs and networks getting faster & cheaper
 - In race with the hardware folk to get the biggest wins.
- Modern CPUs are -amazingly- quick ...

Still lots of fun to be had: do get involved !

- Lots of easy UX wins and polish to work on: JS, CSS, C++ to taste ...
- Follow our [calc](#) & [writer](#) tracker performance bugs: a profile a week ...
- [Get involved in COOL](#) and [LibreOffice Technology](#)



LibreOffice Technology



Collabora Online



LibreOffice Technology



Oh, that my words were recorded, that they were written on a scroll, that they were inscribed with an iron tool on lead, or engraved in rock for ever! I know that my Redeemer lives, and that in the end he will stand upon the earth. And though this body has been destroyed yet in my flesh I will see God, I myself will see him, with my own eyes - I and not another. How my heart yearns within me. - Job 19: 23-27

Thank you!



@CollaboraOffice
hello@collaboraoffice.com
www.collaboraoffice.com