LCA 2021

Collabora Online & Mobile

By Michael Meeks

General Manager, Collabora Productivity

@mmeeks @CollaboraOffice

"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

Collabora Online - what is it? Scaling & Performance

Wrestling a legacy code base Business & FOSS Engineering

The web: compound suffering Tasters from the last year

Mobile: easy?

How to get involved

Linux secure containment

Collabora Online What's that?



Collabora Online

Built with awesome LibreOffice Technology

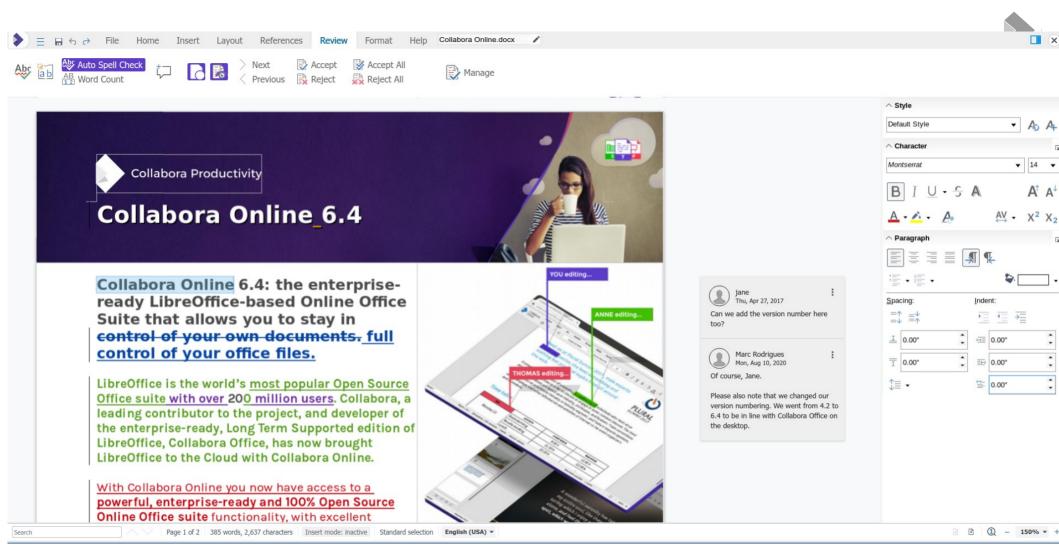


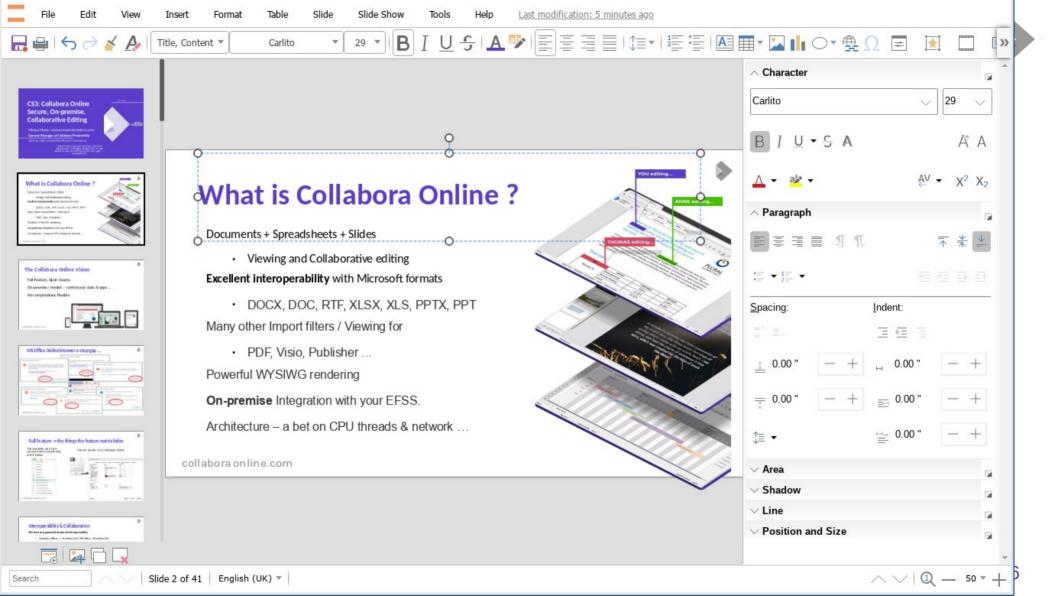


Rich, FOSS, secure, interoperable, collaborative documents, on-premise ...

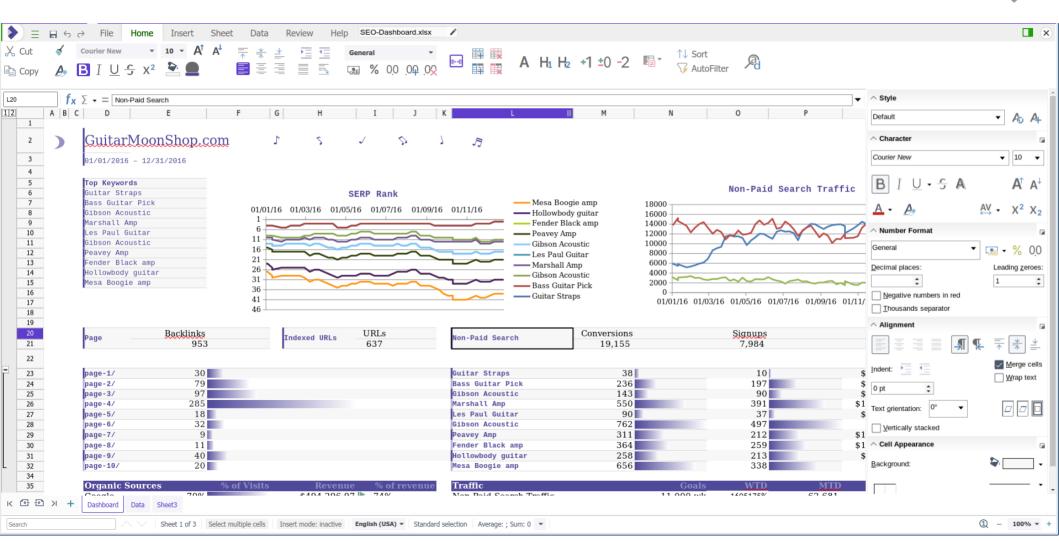


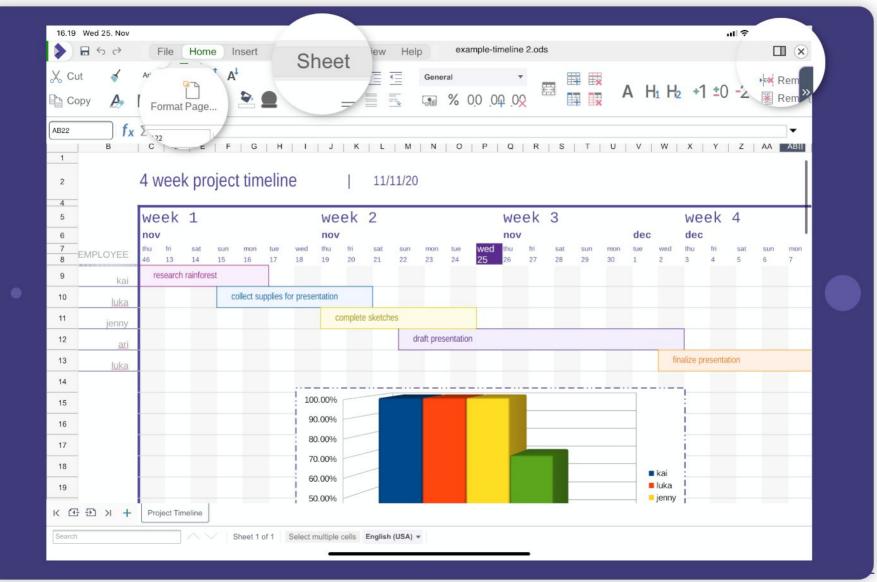
Let me show you some pixels:





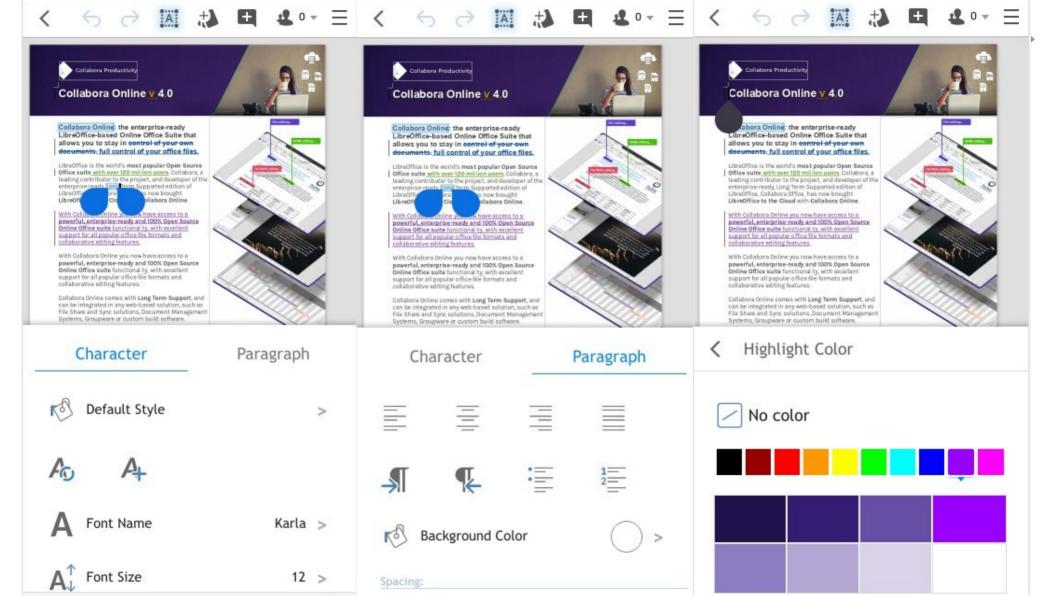


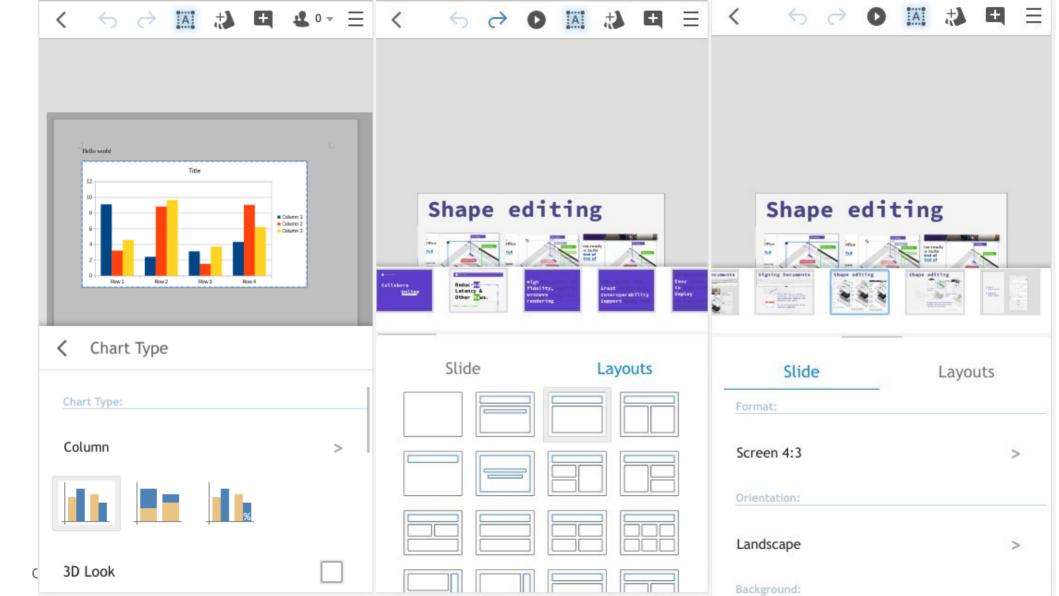


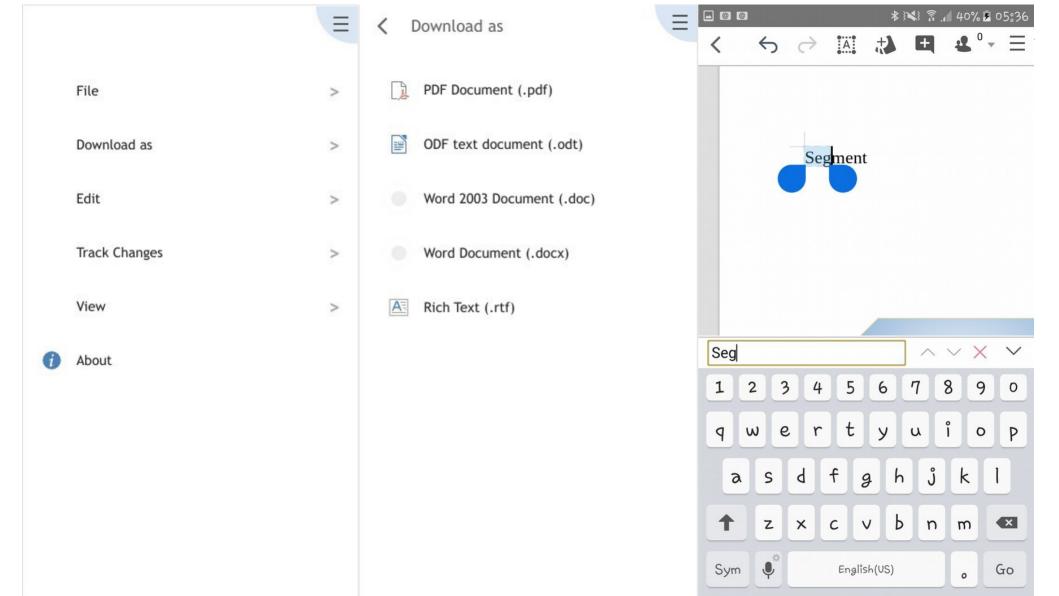




collab









What is Collabora Online?

Built on top of LibreOffice Technology:

- → Excellent interop with Microsoft
- → cf. Office 365 / Google Docs ...

On-premise Integration with your EFSS: ownCloud, Nextcloud, Seafile, EGroupware or LMS: Moodle, or chat: Mattermost

Over 50m pulls from CODE Collabora Online

Development Edition

docker image

Freedom applied → FLOSS giving Digital Sovereignty

Pragmatism: "The best line of code, is the one we didn't have to write" x 10^7 lines

Architecture – a bet on CPU threads & network ... getting cheaper.



Open Document Freedom - sharing the goodness ...

Can share your ODF (and OOXML, DOC) formats with everyone

Without caring about file-format

Web great for Deployment

 Getting ODF everywhere without a PC / Mobile install.



An(other) end to: "I couldn't open your document"

→ follow this link in your inbox for something beautiful – cf. Hybrid PDF, OOXML export etc.

Wrestling a legacy code-base

how did we do that thing?



LibreOfficeKit - wrap and re-use the core

A very simple C & C++ API for LibreOffice

Exposes the core value of LibreOffice:

- File format filters
- Tiled rendering (converting documents to pixels)
- Exposes editing, selections etc.

A very simple ~header-only API - no linking

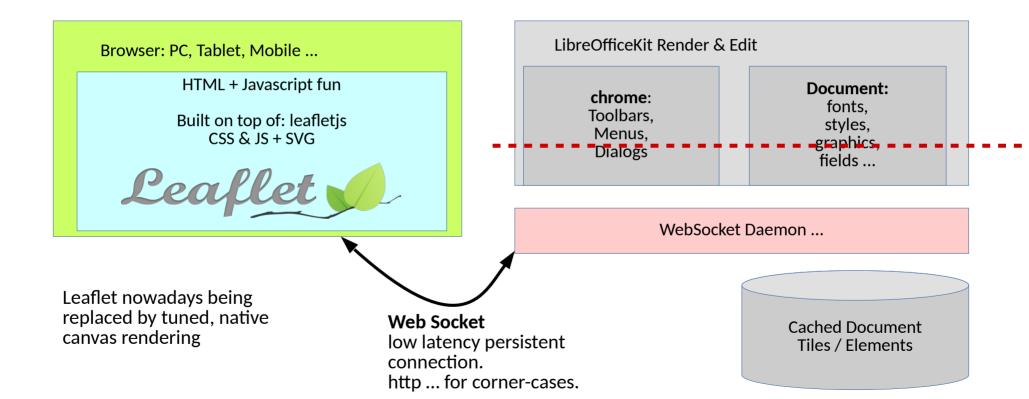
- fully abstract: fn pointers, opaque structs etc.
- No sockets opened, no plugins / simple init.
- Global error messages.

Exposes the Multi-Window (MVCish) core of LibreOffice

Ever wondered about:
 Window → "New Window"
 give it a try ...



Online design: a first cut ...

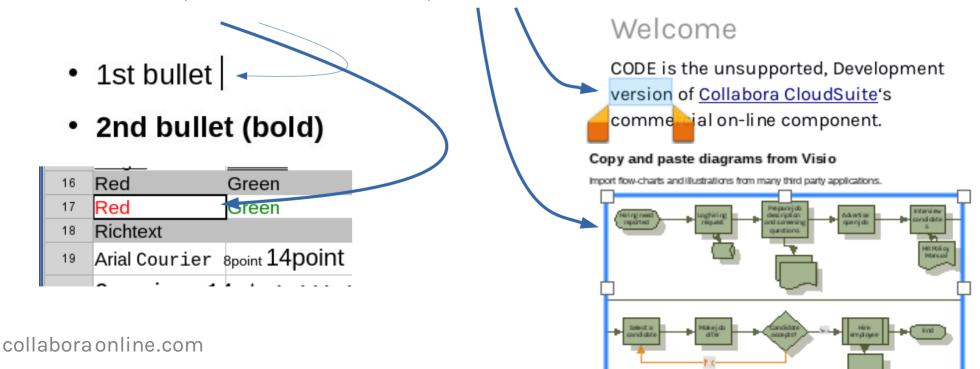




Fast local interactions: local overlays ... (SVG)

Layers over the document:

Cursor, Selection handles, etc.



The web: compound suffering

the bits you don't want to know.



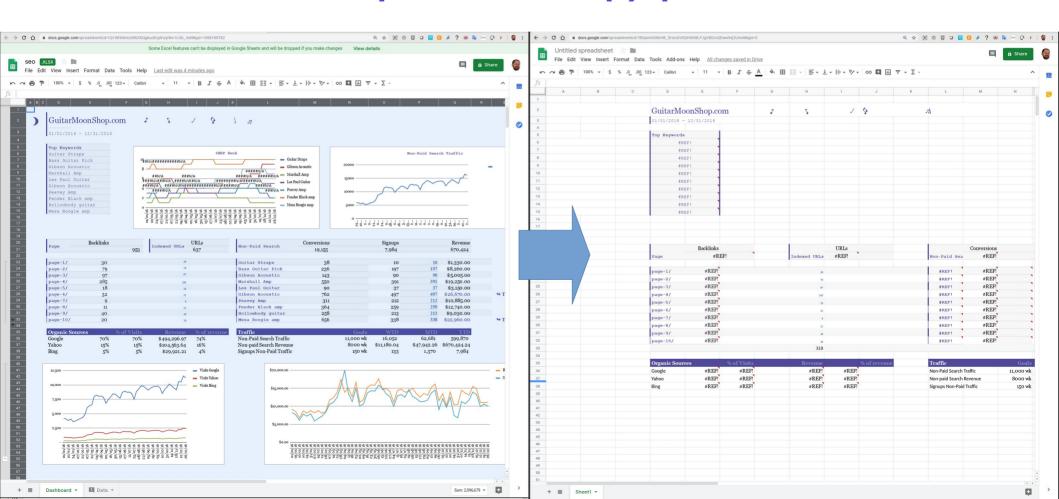
Horrible parts of the web

Keyboard input

- Totally unpredictable
- man months finding out that all the waffle about keyboard & text entry events:
 - Nonsense; no reliable re-construct of text entry from the events possible
 - Mind-blowing cross-browser brokenness [!]
- A hidden entry ... positioned with CSS at the cursor (for ATs)
- Diff it from A→B, see what changed, emit events:
 - Requires a "<space>|<different-space>" to distinguish delete/backspace
- Then you fight the asynchronous keyboard to keep your text-entry clean.

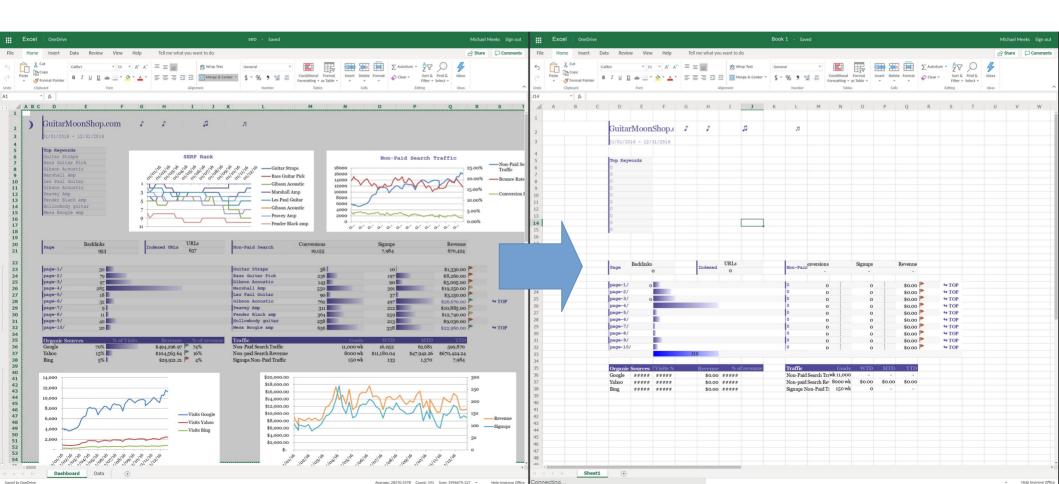


Chrome / GDocs spreadsheet copy/paste





Office.com / Excel copy / paste ...

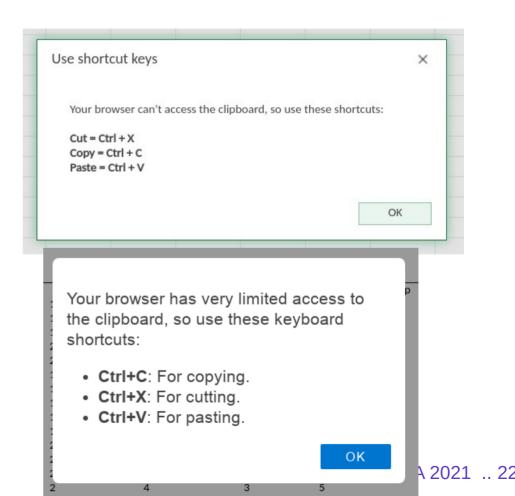




Web copy/paste is baroque & paranoid beyond belief.

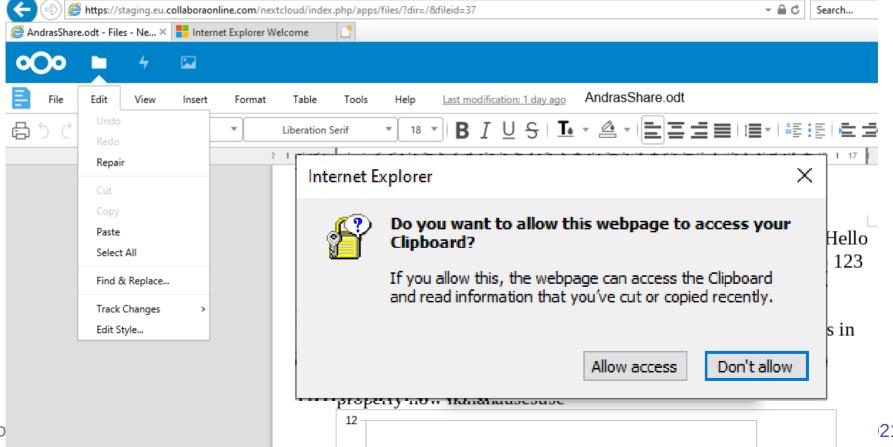
APIs for copy/paste are very mixed

- In many browsers copy/cut can put ~anything (HTML/txt/RTF/IMG) on the clipboard ~whenever they like.
 - But they have to work-around layers of obsolete 'security' foo.
- 1. Find/focus/select content-editable
 - 2. document.execCommand('copy')
 - 3. catch events if you get them
 - 4. tweak your contenteditable,
 - 5. retry from 1.





IE11 - had a great initial approach:



collab



Google Chrome / Docs - joined up thinking?

Google Chrome's

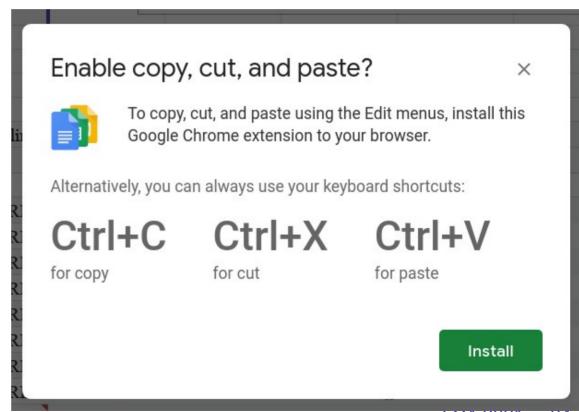
 Security team hardened by many issues → so rightly paranoid.

Google Doc's

 Users cannot copy/paste using context menus → make a plugin for that.

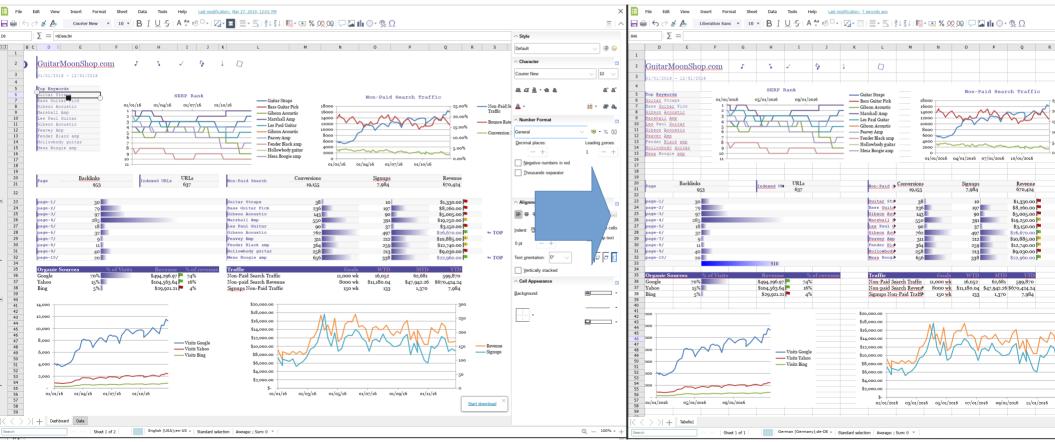
The Security situation?

- Snooping is indeed bad:
 - IE11 solution was a good one.
 - Signed sites you trust ...?





HTML → plus back-channel to fetch the real data



collabora online.com

LCA 2021 .. 25

Mobile: easy?

a decade of work & fun ...



New platform / form-factor fun

A decade getting here

- Cross compliation ...
- Android C++ system library dynamic patching.
- Android: fixed 96 DSOs
 - 45+ system libs
 - crash at the limit ...
- Faster linking, profiling

Native mobile UX

- Fennec seemed a good base
- Wrote in Java

Switch to Online

- Embedded Webkit →
 Online / LibreOfficeKit
- Easy! Except:

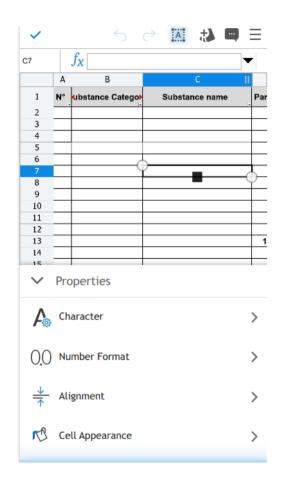
collabora online.com

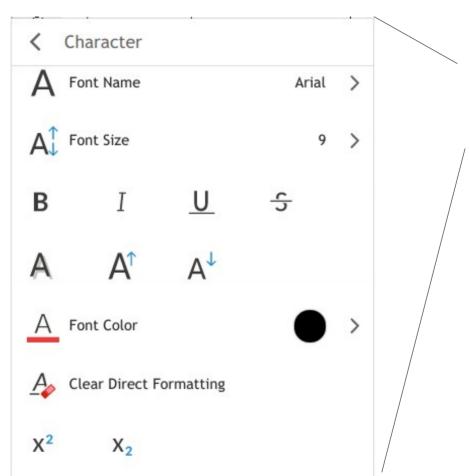
New User Experience

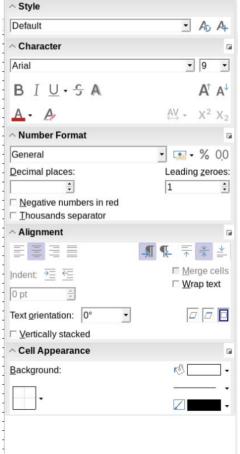
Re-using much of the LibreOffice sidebar



• glade XML \rightarrow VCL \rightarrow JSON \rightarrow Re-layout







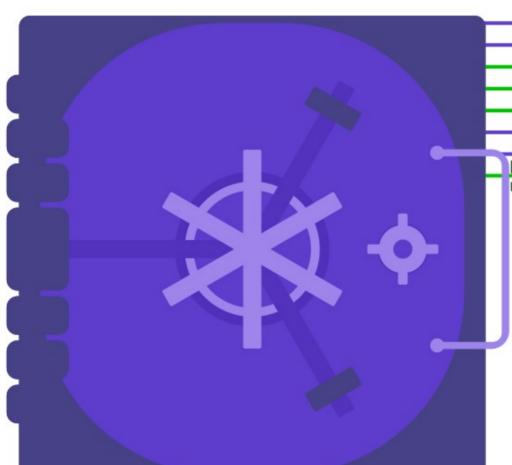
Linux containment

making this secure on servers



Fork & drop capabilties ...

Everyone in their own chroot + seccomp-bpf + ...



on your premise / Private network

Virtual machine / Docker container

Seccomp-bpf ~no sys-calls

Extremely sparse filesystem

Chroot per document

Systematic load crash testing

Industry beating coverity score

LibreOfficeKit

rendering instance

chroot

- no shell
- handful of NSS shared libraries, etc.
- fonts
- one document.

Pixels come out ...

Scaling & Performance

making this efficient on servers



Different optimizations

Online

- dlopens ~world (RTLD_BIND_NOW)
- Loads ~all config, dictionaries & hyphenation
- then forks children
 - Copy-on-write-ness (COW) everywhere.
- Burn memory in pre-init to save in children
- Burn CPU time on startup to save it later.

How to catch anti-social behaviour:

• Eg. touching / un-sharing pages:



Compare COW / shared memory pages: loolmap

Shows parent data / diff: here a ref-count in a Unicode string from 1->3

Also dumps all strings it can find with --strings

Totals for heap

shared 23040 (92160kB)
 unshared 982 (3928kB)
 same but unshared 191 (764kB)
 dirtied bytes touched 30718 per page 31.28

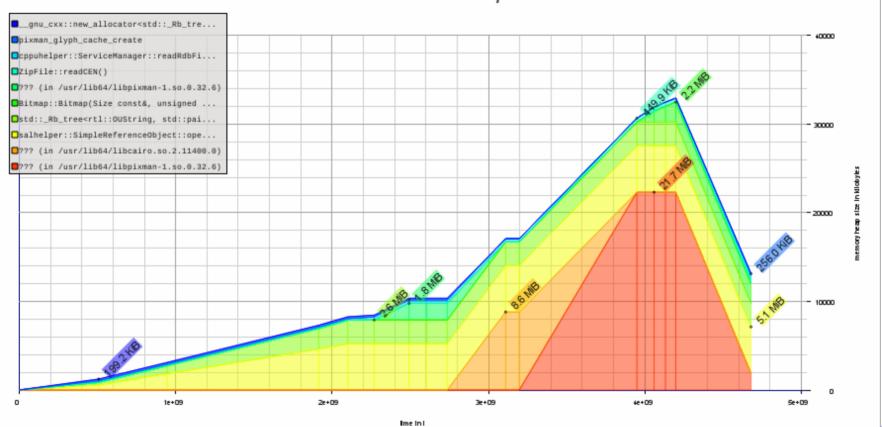
→ staticize our preinit strings

Would love help expanding the tool to track that to the respective allocations ...



Tilebench - simple test tool: 38 Mb peak.

Memory consumption of /data/opt/libreoffice/master/workdir/LinkTarget/Executable/tilebench
Peak of 38.1 MiB at snapshot #48



colla

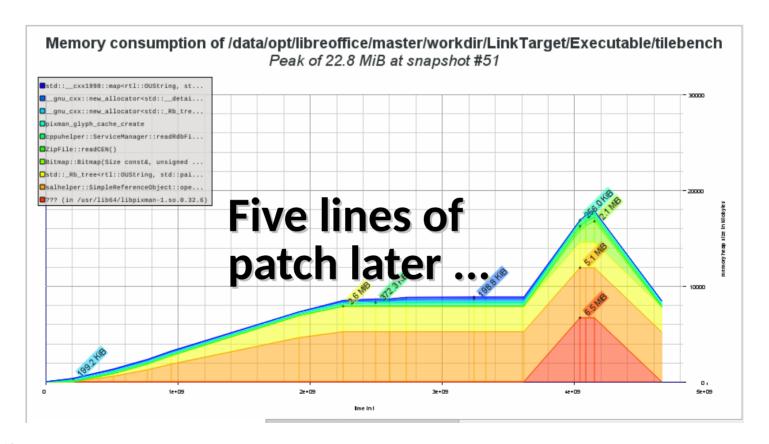


Amazing what you find: running on your machine now:

```
pixman/pixman-glyph.c (pixman glyph insert) - 40Mb
                                                            since May 2012
/* XXX: These numbers are arbitrary---we've never done any
measurements.
#define N GLYPHS HIGH WATER
                               (16384)
#define N GLYPHS LOW WATER
                               (8192)
How many glyph bitmaps to keep around.
cairo/src/cairo-scaled-font.c - since March 2009
    /* XXX: This number is arbitrary---we've never done any measurement
    of this. */
#define MAX GLYPH PAGES CACHED 512
Patches for internal versions to better manage the size of that ...
```



TileBench: with some optimization:





COW sharing across processes is efficient:

Heap is only one way to burn RAM but still a popular way:

Red waste removed: 9Mb

Remaining: ~14Mb

of which

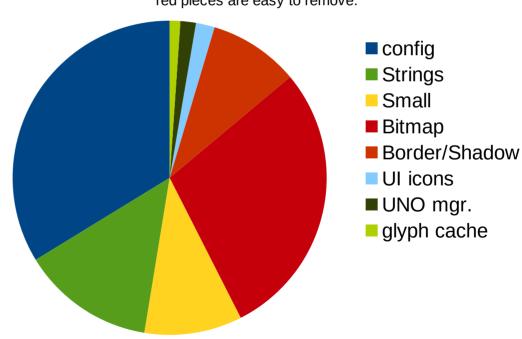
~12Mb is shareable by forkit

~<5Mb ... base use

Real-world guess: ~50Mb / doc.

Breakdown of tilebench heap size by type (24Mb total)

red pieces are easy to remove.



collabora online.com LCA 2021 .. 37



Large deployments / HA pieces

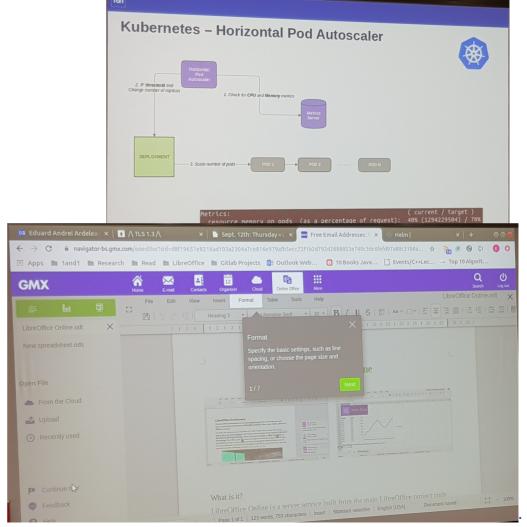
Scales with the network.

Simple architecture:

- No new HA Database
- No new message-bus
- No shared data storage

eg. 1&1 / GMX

~millions of users with access

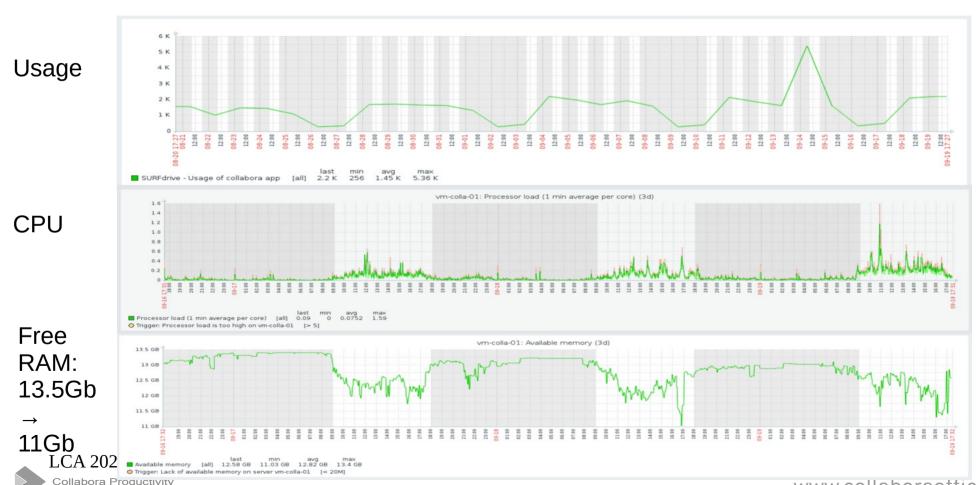


Collaboration Protocol

- Not composable operations: CompSci → get all excited about this.
 - Horrors of async model sync. → consistency, save, perf. etc.
- Real-world use-cases:
 - 90% of the time → "One user working alone on the authoritative document"
 - ~10% of the time → "Lots of users in a call, co-editing / scoring / minuting"
 - \sim 0% of the time \rightarrow conflicting edits with undo/redo on the same document
- Conflicts fundamentally not computer-resolvable always
 - Better to back-off & show the user
- Bring all the keystrokes & mouse movements
 - Into the same process same memory space same CPU

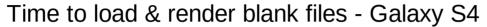


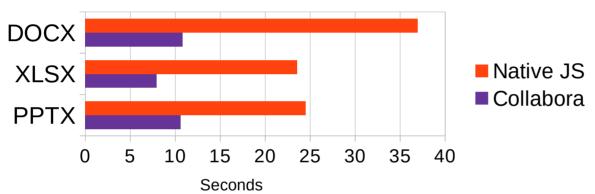
Real world: ~45k users - Collabora Online - SURFdrive





Mobile: do the work on battery, or in the data-center?

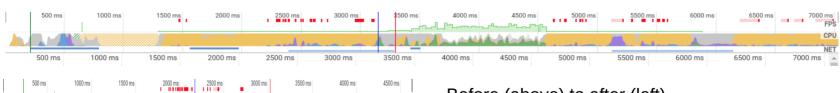




First cut pre-optimized numbers

With data: larger spreadsheets 10x faster → 1000X faster.

Threaded, GPU accelerated spreadsheet core



Before (above) to after (left) – re-ordering & optimizing "Hello World" DOCX load - ~3.5s from start.

1000 ms

Progress this year

a taste of where we placed our investment bets



Quick try-out: built-in demo servers ...

Please make sure you understand that the following will happen if you set up the Collabora Online demo.

- The service will send users documents to Collabora and/or third party demo servers.
- This service is not intended for production use, hence the documents will show tile watermarks.
- The demo service may be under heavy load, and its performance is not representative in any way of the performance of an on-premise installation.
- These servers are used for testing and development, and may run test versions
 of the software. As such they may crash, burn, and re-start without warning.
- The users documents will not be retained by a third party after their session completes except in exceptional circumstances. By using the service, the user gives permission for Collabora engineers to exceptionally use such document data, solely for the purpose of providing, optimizing and improving Collabora Online. Such document data will remain confidential to Collabora and/or any third party providing a demo server.

At the first use and after an update, each user will get the warning, explaining all the above.

I agree, and use the demo server

I will setup my own server

Use a demo server

You can use a demo server provided by Collabora and other service

Select a demo server

Friprogramvarusyndikatet – Sweden

Collabora Productivity Ltd. - Ireland

Collabora Productivity Ltd. - North California, USA

Collabora Productivity Ltd. - Tokyo, Japan

Collabora Productivity Ltd. - São Paulo, Brazil

Easy to test your new integration against



Built-in demo servers: iff safe & route-able ...

Use a demo server

You can use a demo server provided by Collabora and other service providers for giving Collabora Online a try.

Your NextCloud setup is not capable of connecting to the demo servers because:

- it is a local setup (localhost)
- it uses an insecure protocol (http)

For use cases like this, we offer instructions for a **Quick tryout with Nextcloud docker**.



Better: Nextcloud Hub / Applmage install ...

A lovely, simple one-click install ...

- A small / fast PHP forwarding proxy & an AppImage
- Please upgrade to a faster, native server quickly ...
- Use your own server

Collabora Online requires a seperate server acting as a WOPI-like Client to provide editing ca

Use the built-in CODE - Collabora Online Development Edition

Easy to install, for personal use or for small teams. A bit slower than a standalone server and



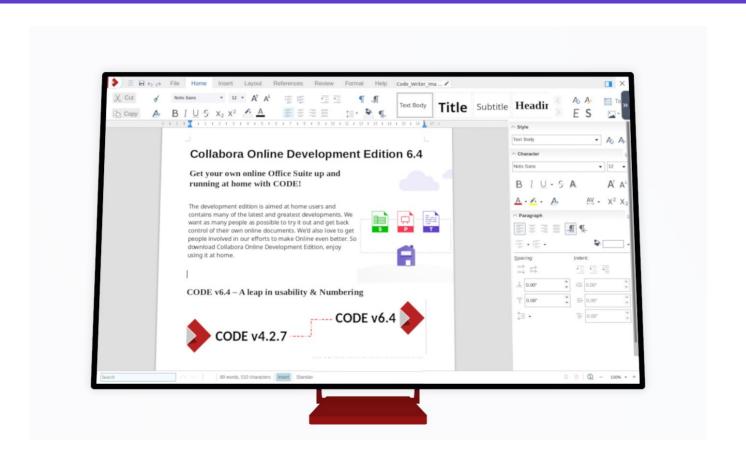
Polling websocket-alikes over keep-alive'd https://

Time / milliseconds

Notify to reply:

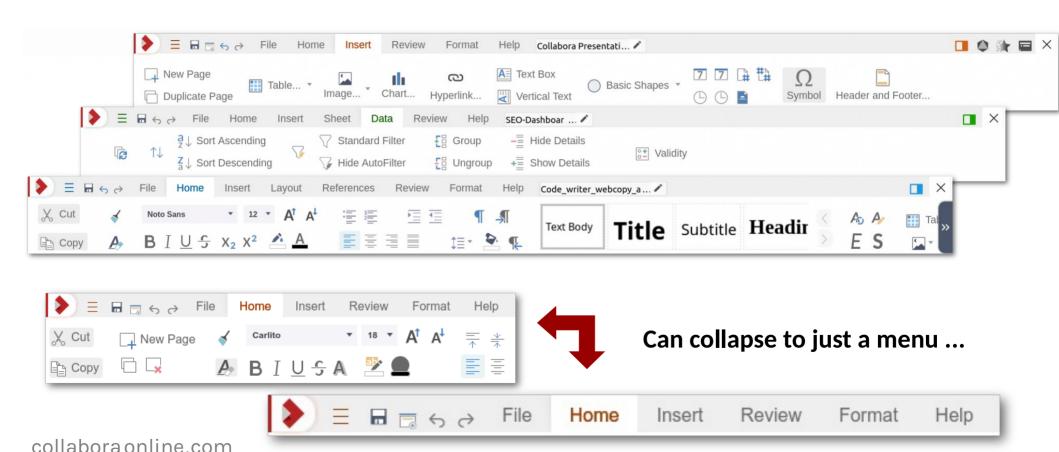
RW Back A SW Out A Ev Out B RW Back B RW Out C Back C Out D RW Back D RW Back E Out E Out F RW Back F Out G ... worst case notify to reply latency ... 180ms (from 50ms ping) 8000 ms 10000 ms 12000 ms 14000 ms 10000 ms 20000 ms 30000 ms 40000 ms 50000 ms 60000 ms 70000 ms 80000 ms

Collabora Online 6.4.0 (Nov 20)



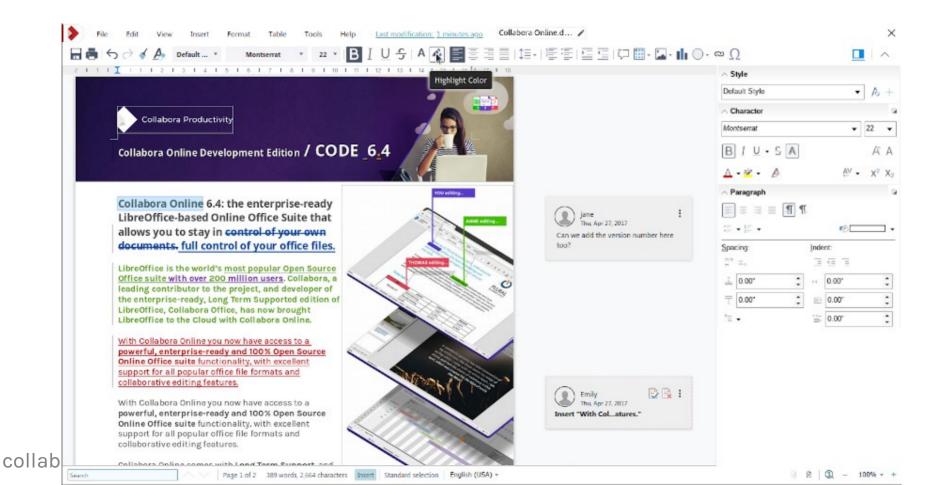


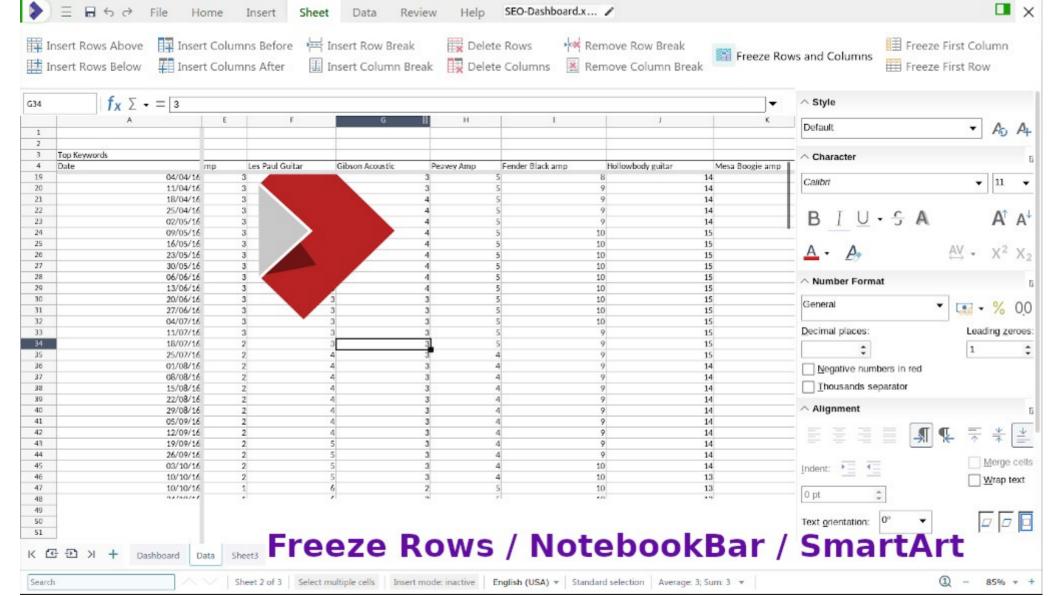
Eclectic Notebookbar UX choices: to taste

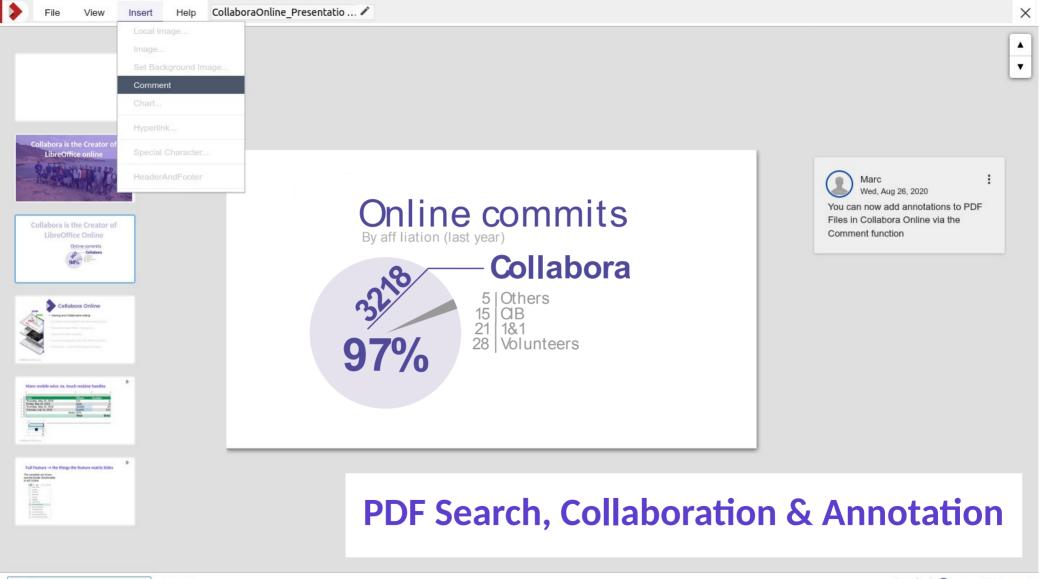


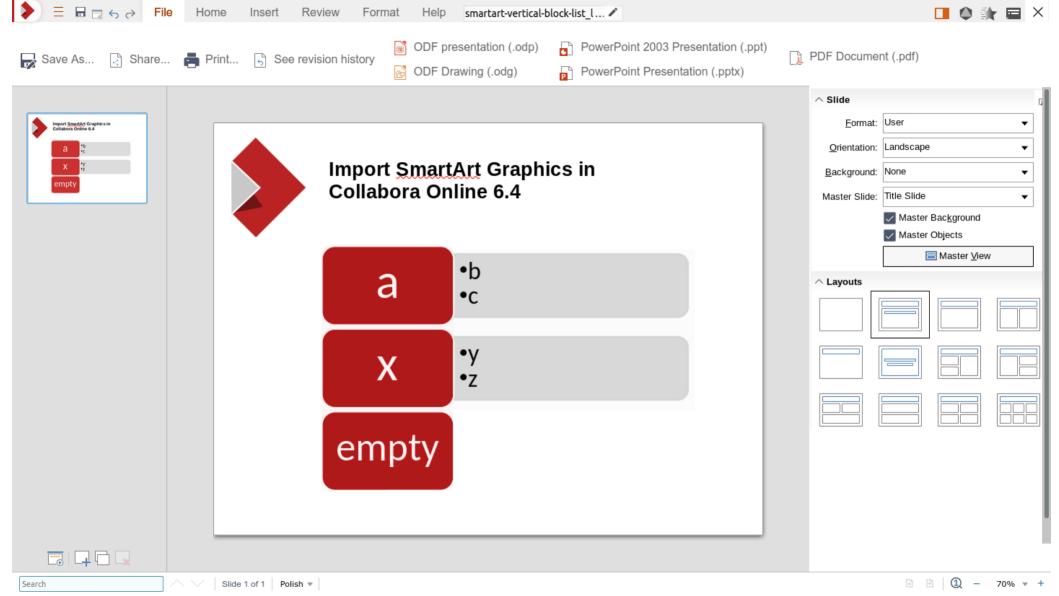


Classic mode retained: configurable & default ...



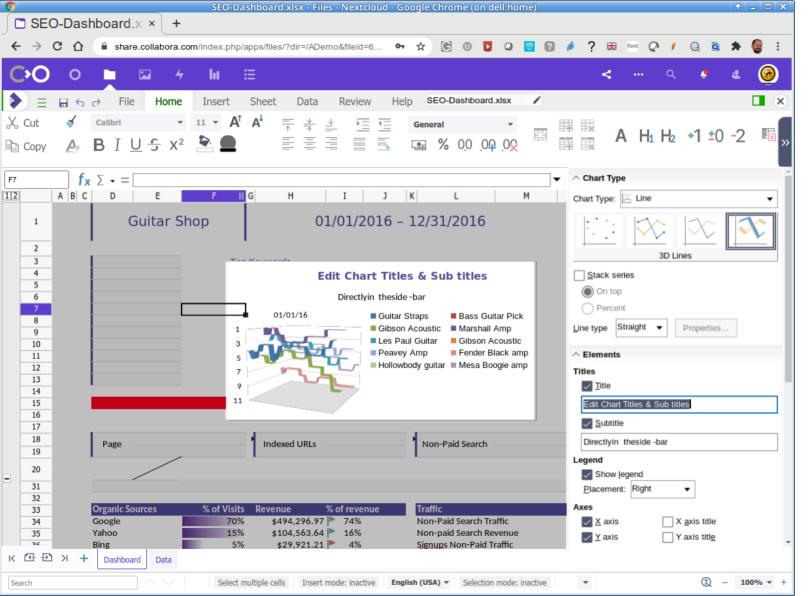






Online 6.4.x ...

rapid fire incremental goodness

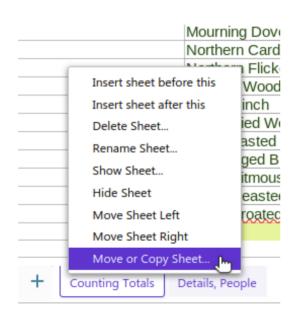


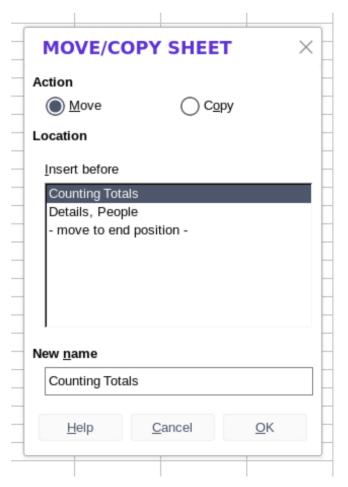


Improved chart editing sidebar functionality



Calc: Improved sheet movement:

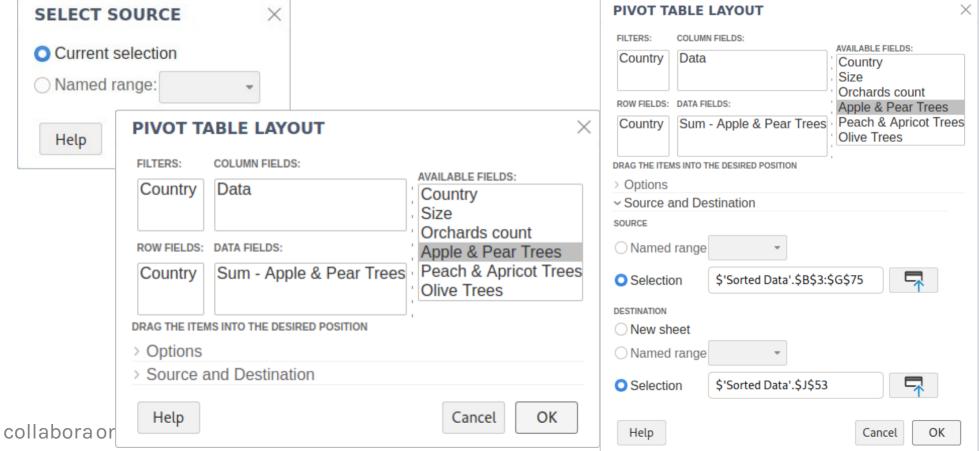




collabora online.com

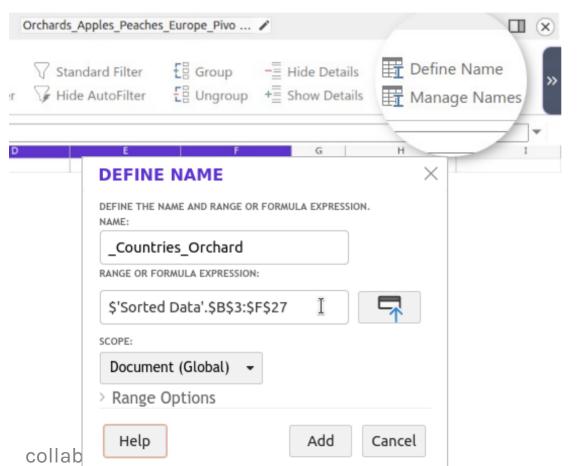


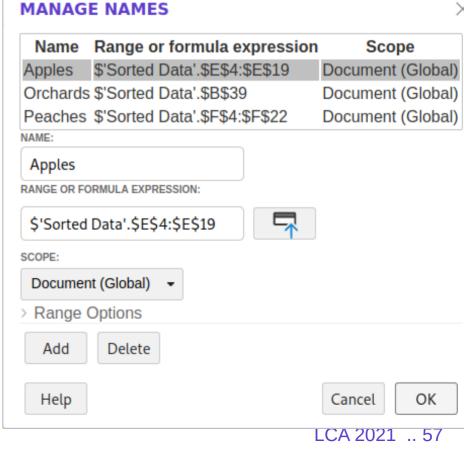
Calc: customize & create Pivot Tables





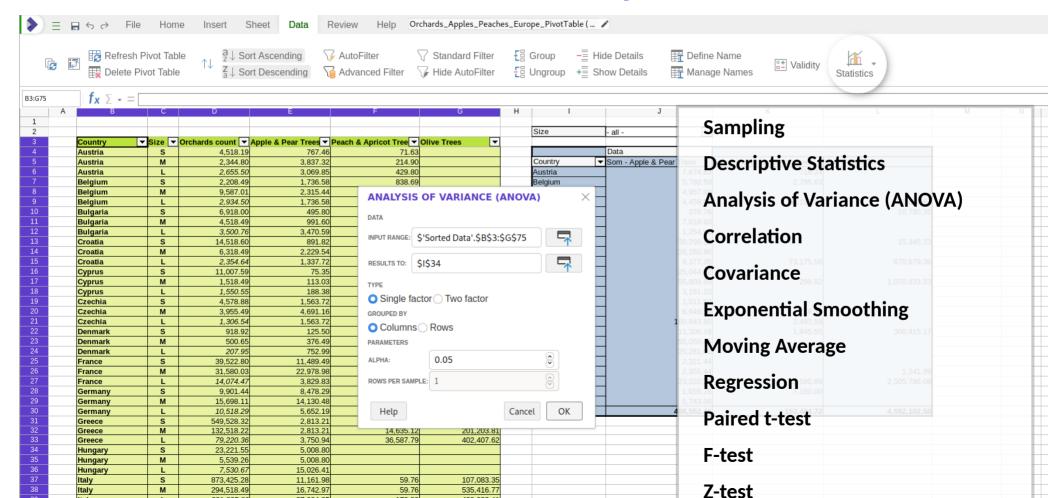
Calc: define & manage range names & print ranges







Calc: Statistical tools for data analysis





The pieces you can't see

Performance

- faster rendering
- Snappier interactive editing of cells
- Bind-mounting of document jails (instead of hard-linking)
- Improved watermark caching
- Slide content pre-fetching
- On-demand slide thumbnail fetching.
- Significant PDF rendering improvements
- Significant thumbnail / download speedups

Misc

- iOS multi-document support
- Substantial cypress automated test coverage.
- Cluster debugging / warnings ... check matching ServerIds

And much much more ...

- 3600+ commits.
- Innumerable bug-fixes & improvements

collabora online.com LCA 2021 .. 59

Mobile 6.4.x ...

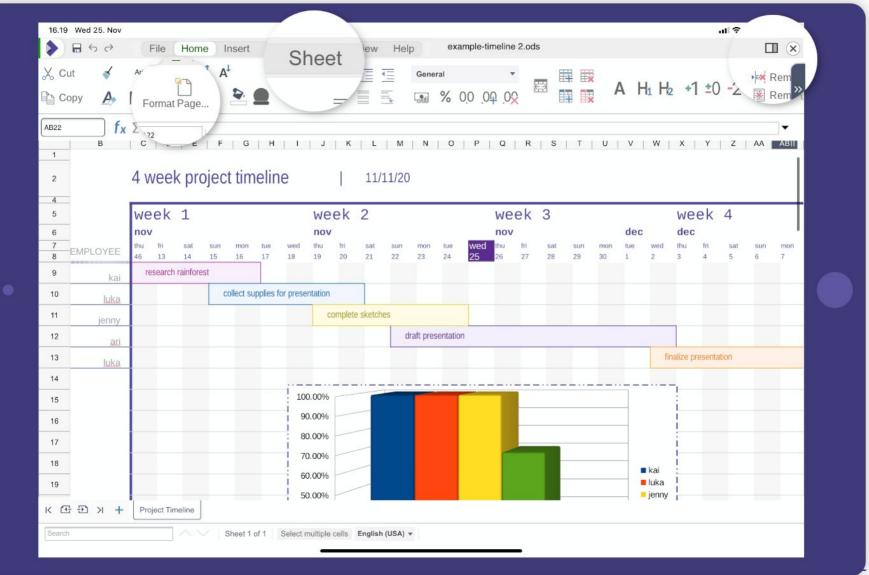
bringing goodness to your phone





Collabora Office

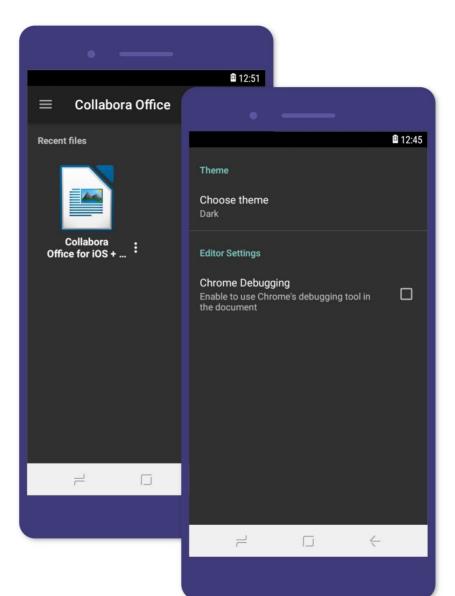
6.4





collab

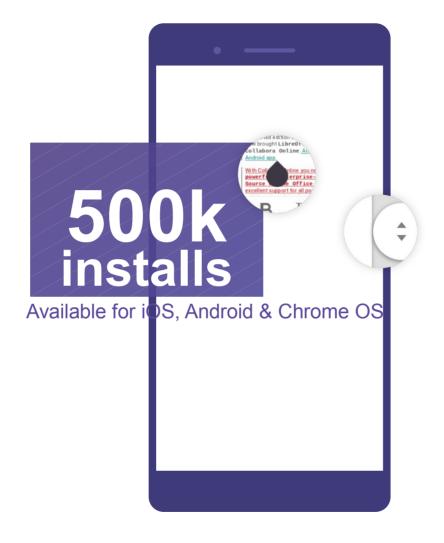
Android: start of a Dark Mode for the Online shell.



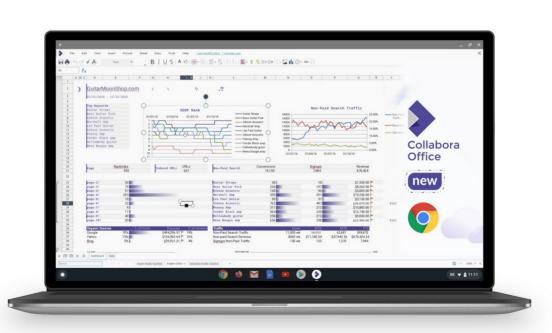


collabora online.com





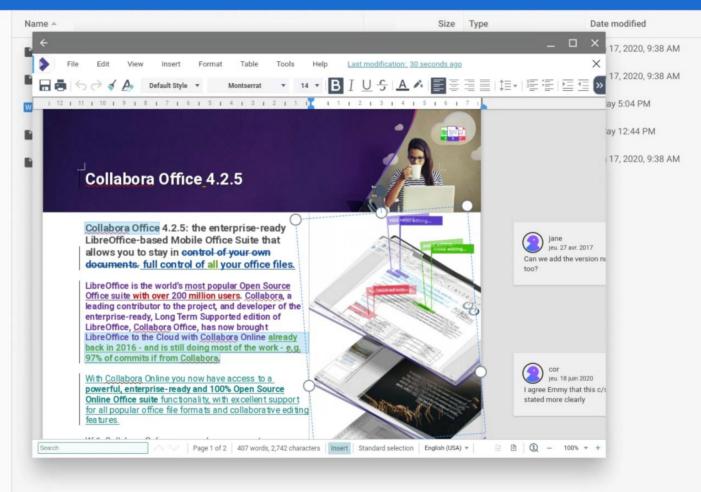
ChromeBooks:







- Collabora Office 6.2-13
- Collabora Office 6.4
- Collabora Online Nextcloud Edition
- Collabora Online v1
- Collabora Online v2
- Collabora Online v2.1
- Collabora Online v3
- Collabora Online v4.0
- Collabora Online v4.2
- Collabora Online v4.2.1
- Collabora Online v4.2.2
- Collabora Online v4.2.3
- Collabora Online v4.2.4
- Collabora Online v4.2.5
- Collabora Online_DemoServer
- Collabora Online_General
- ✓ Demo Test Docs
 - Collabora Office_DemoDocs to be finished





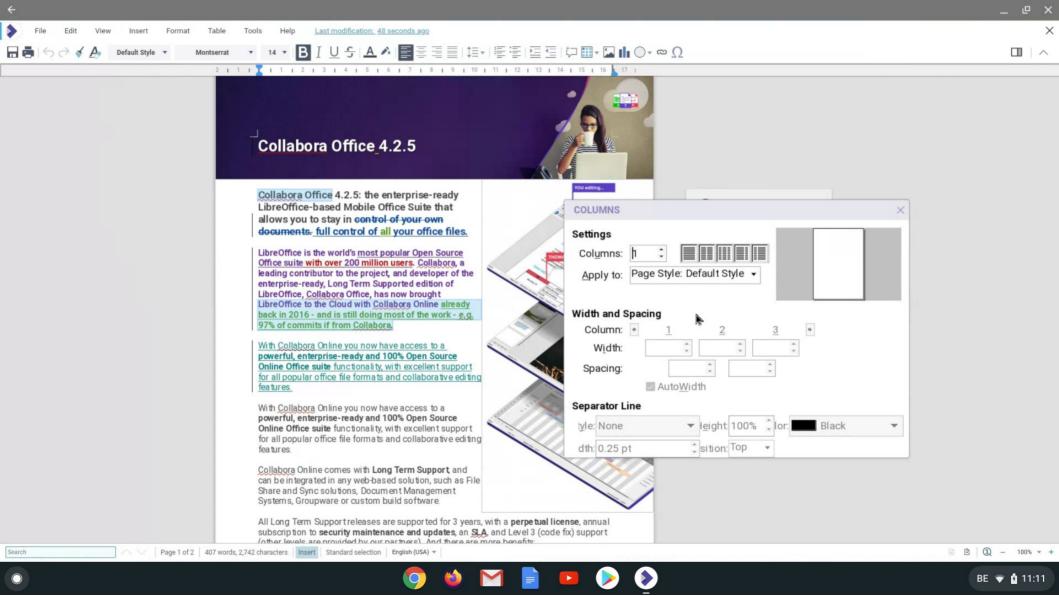










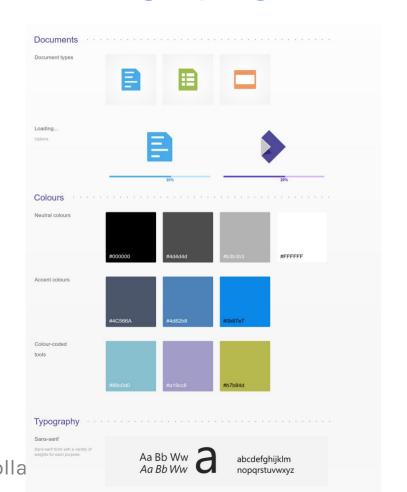


UX integration:

Letting you theme your your integration to match (cf. blog)



Getting style guide: a how-to ...



Iconography																- 22
	-		ā	6	\rightarrow	*	A	A	1	1 (8)	»					
	В	I	\underline{U}	5	A											
	(A)	٥	23	1	_	+		口	ΑΞ		<u> </u>	Ili	\bigcirc	ര	Ω	
	K	×	\equiv	≣	≣		₽	\equiv	=	₫	\blacksquare	+ +				
	Σ	fx														
	=		0.0	0,0	\$	%										
		•	***	₪												
	+ -	+ +	*	*												
				×												
Dialog Windows	Sele F		ND NUME				×						11.74			



ui_defaults: posted with the iframe:

<input name="ui_defaults" value="TextRuler=false;StatusBar=false"
type="hidden"/>

";" delimited options in a post field.

uiMode=classic/notebookbar - select menu/toolbars

StatusBar=false

- hide status bar

TextStatusBar=false

- hide status bar only for Text

other prefixes: Presentation/Spreadsheet

Sidebar=false

- hide sidebar by default

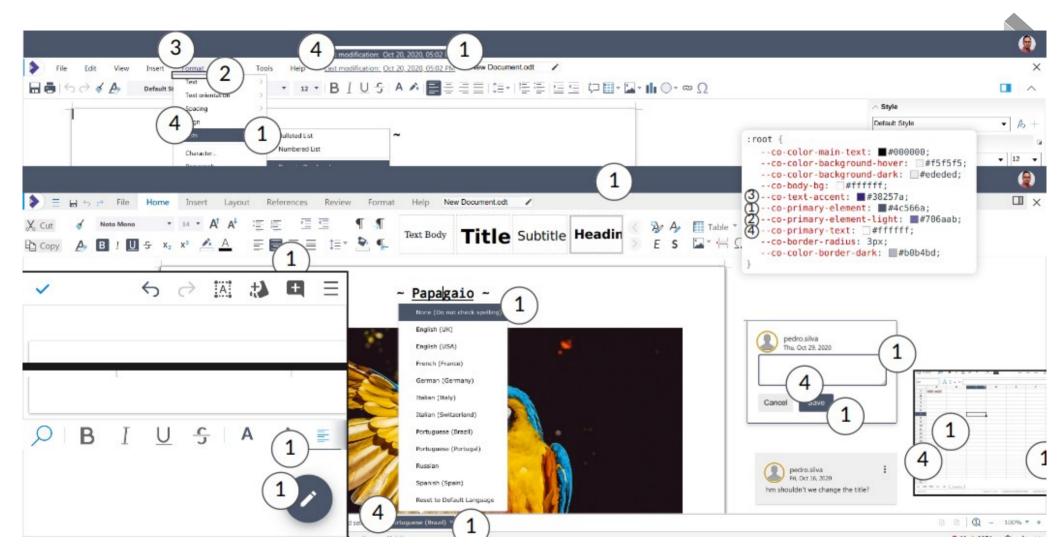
Ruler=false

- hide ruler by default ... etc.

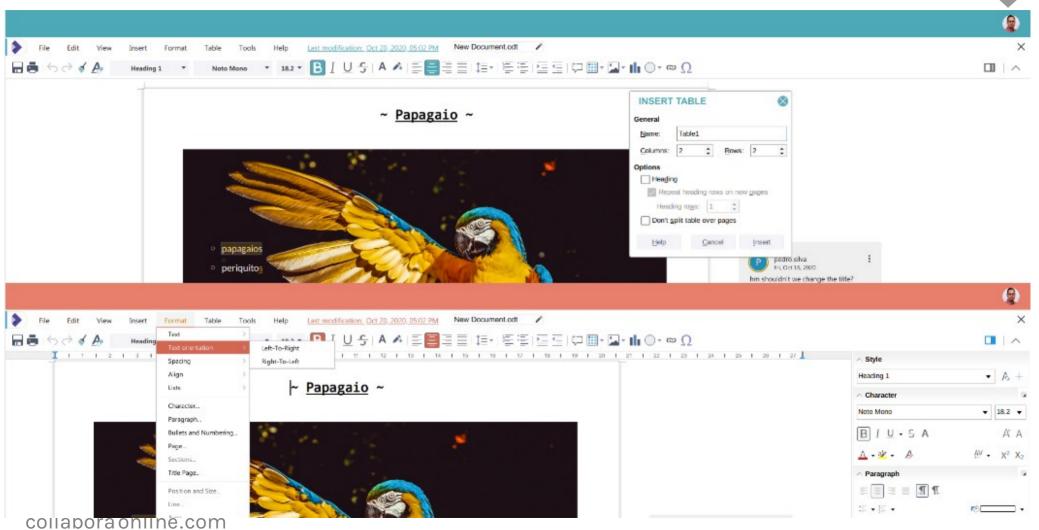


css_variables: posted with the iframe:

```
<input name="css variables" value="co-color-main-text=#000;--co-body-bg=#FFF--co-txt-</pre>
accent=#38257a; "type="hidden"/>
      ":' delimited options in a post field, default (for now) are:
--co-primary-element: #4c566a
--co-primary-element-light: #706aab
--co-txt-accent: #38257a
--co-primary-text: #ffffff
--co-border-radius: 3px
--co-body-bg: #ffffff
--co-color-main-text: #000000
```

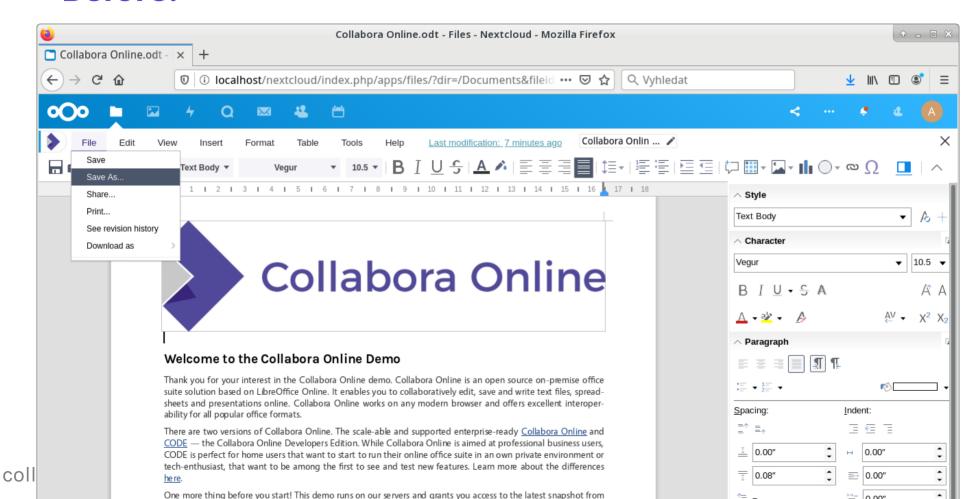






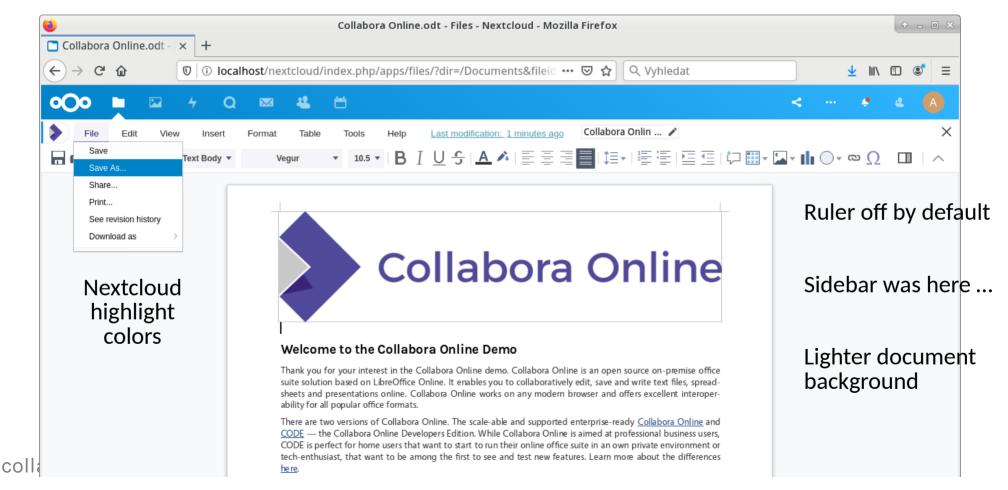


Before:





After: theming / palette & settings ...



Business & FOSS Engineering

time to market, branding & who dares pays ...



Engineering - Business / Community

How to fill this gap

- Paris → GDK-Broadway based prototype
 - 2011
- Collabora Online 1.0
 - 2016
- Why? Investment → return.
 - Eur million(s)

Engineering choices

Getting a product bootstrapped
 "ship it when it's perfect"
 vs.
 "release early & iterate to
 perfection"

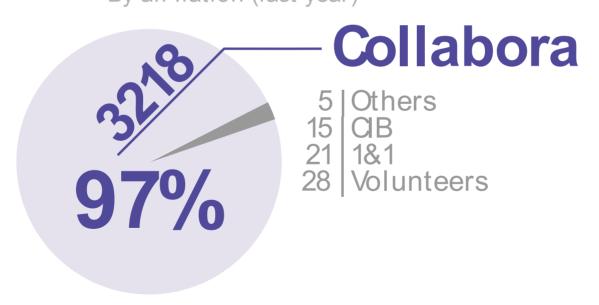
- Technical / release schedule choices.
- Economics / Marketing / Sales ...





Putting in the investment ...

Online commits By aff liation (last year)



collabora online.com LCA 2021 .. 79



Moving our development to github

Clear brand story

- Helps to drive investment and avoid user / customer confusion
- Unwinds LibreOffice vs. Collabora brand conflict

Removed reminder for users in CODE

Developer pool growth ...

Continue to support, work on LibreOffice

 TDF doing well – Eur 1.3m raised last year from Desktop +~30%

Tackles the "Free Compliments" problem

At least marketing-wise ...

Similar to Nextcloud/ownCloud/ pydio/seafile/etc.

- A clear flow of users through their projects
- An obvious place for support
- Clear credit given in the product name
- No complex structures necessary.

FLOSS company assets: cash + their brands

ROI is important → in marketing.

collabora on line.com LCA 2021 .. 80



Example new Online volunteer contributions

Lots of Design / CSS work

Thanks to Andreas Kainz

Improving icon-sets

Thanks to Rizal Muttaqin

Python tooling improvements

Thanks to Alan Verdugo, corleone77,
 Batuhan Görkem Benzer, Kuesji
 Koesnu and Mücahid Aydin

Javascript code cleanups & wins

 Thanks to Dilaver Demirel, Pelin Kuran, Buğra Kurt, Irina Hristova, Rashesh Padia and Ayhan Yalçınsoy

FreeBSD port improvements

Thanks to Gleb Popov

C++ core code cleanups

• Thanks to Jan Dagefoerde, Yusuf Keten

Easy to contribute - with gitpod

collabora on line.com LCA 2021 .. 81

How to get involved ...

Easy wins left & right for many skills



It is easy to get involved:

LibreOffice features:

http://www.libreoffice.org/

Code

https://git.libreoffice.org/core

Get involved:

- https://www.libreoffice.org/community
 Forum: /get-involved/
- C++, python, l10n and more
- https://translations.documentfoundation.org/

Collabora Online

https://collaboraonline.github.io/

Easy Hacks:

https://github.com/CollaboraOnline/online

• C++, Javascript, CSS, Design ...

https://hosted.weblate.org/projects/collabora-online/

https://forum.collaboraonline.com/

Weekly dev. meetings

collabora online.com LCA 2021 .. 83





Conclusions:

Easy to deploy on your existing:

 Nextcloud, ownCloud, Moodle, Seafile, Mattermost, EGroupware ...

Regain control of your data & documents today

Easy to integrate into your web app

Server load-wise, small compared to EFSS

Improving at significant pace

Created by Collabora

- with the generous help of many
- easy to contribute to

based on LibreOffice Technology

Thanks for your support!

collabora on line.com LCA 2021 .. 85

