

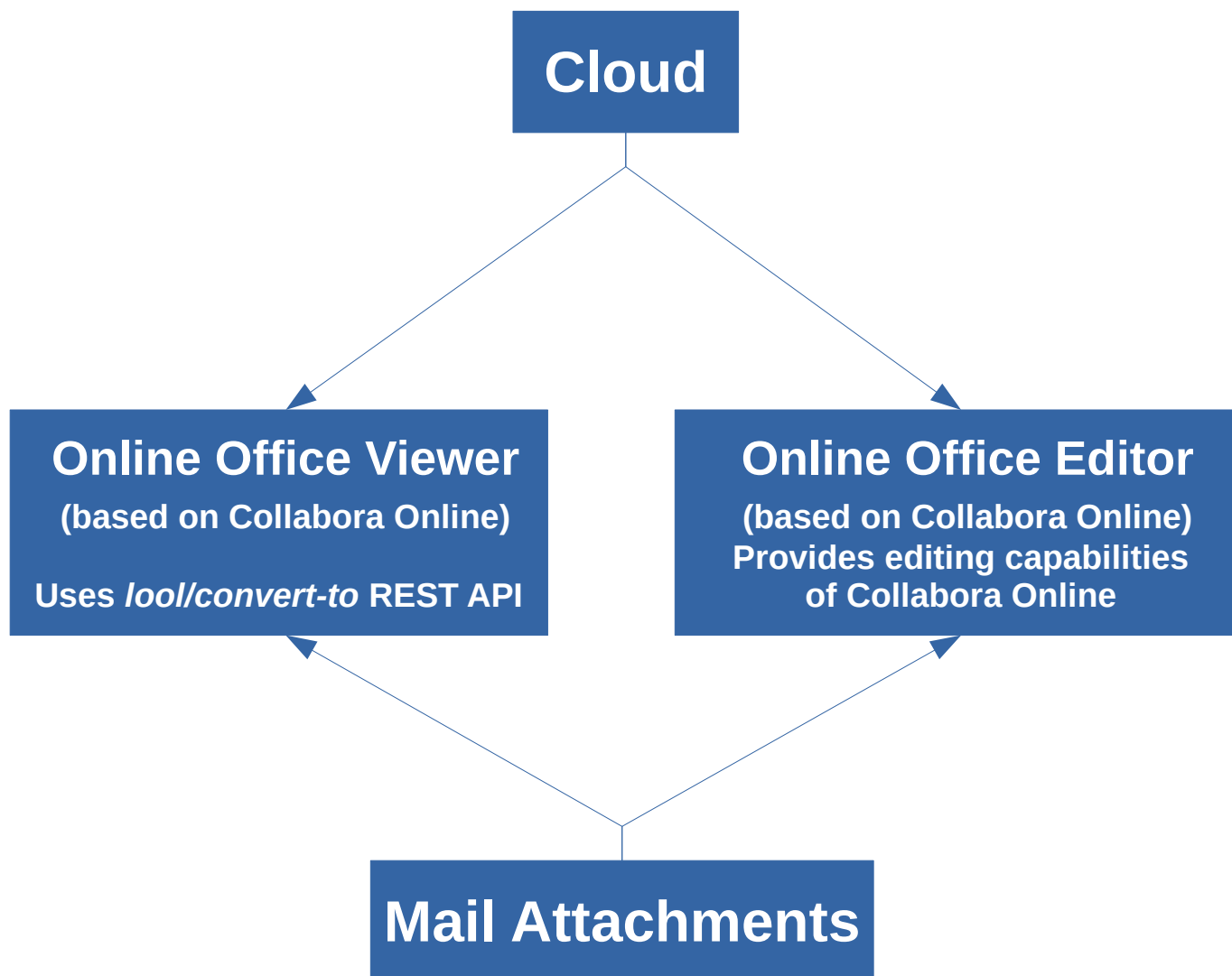
1&1 Mail&Media

STABILITY & CLEANUP IMPROVEMENTS IN ONLINE

Gabriel Masei
Senior C/C++ Developer



Where and how we use Collabora Online



Scale and type of deployment

Online Office Viewer

- > 100 instances of Collabora Online
- > 600,000 conversions / 24h

Online Office Editor

- > 100 instances of Collabora Online
- > 35,000 edited documents / 24h
- > 600 parallel edited documents

Kubernetes & custom components

- Deployment, scaling and management
- One instance / pod

Stability issues

- Crashes
- Abnormal CPU and/or memory resources consumption on multiple pods.
- Most of issues were inside kit processes where the core is loaded.

Possible approaches

- The ideal approach: solve the underlying core issues
 - time/resource consuming
 - never-ending because of constant evolution of application.
- Use Kubernetes and restart the problematic pod.
 - Healthy kit processes will also be affected, increasing number of customers that experience issues (like document reloading and lose of edit context).
- Implement an automatic cleanup mechanism that dismisses only the problematic kit processes.

Our approach

- General approach is to dismiss (SIGABRT/SIGKILL) only the problematic kit processes.
- Split problematic kit processes into 2 categories:
 - Processes that are **still** referenced by loolwsd and consume resources.
 - Processes that are **not** referenced by loolwsd.
- Each needs a different approach:
 - Those still referenced by loolwsd need a more careful approach since they could still be processing user input.
 - Those lost can be handled more directly because they are useless and should not exist.

Cleanup of resource consuming kit processes

- A kit process is a resource consuming one suited for disposal if:
 - The document is in **idle state** for at least a **particular amount of time**.
 - After that the document must **consume resources** for at least a **particular amount of time**.
 - In this time the **CPU or memory usage** must reach at least a **minimum threshold** and must not go below that threshold.

Cleanup of resource consuming kit processes

- These values can be set through `loolwsd.xml` configuration file.
- They can be found under `config.per_document.cleanup` tag
 - **idle_time_secs** (default: 300 seconds)
 - **limit_cpu_per** (default: 85%)
 - **limit_dirty_mem_mb** (default: 3072MB)
 - **bad_behavior_period_secs**: (default: 60 seconds)
 - **cleanup_interval_ms**: (default: 10,000 milliseconds)
- It can be enabled/disabled through „**enable**“ attribute of **cleanup** tag. It is disabled by default.

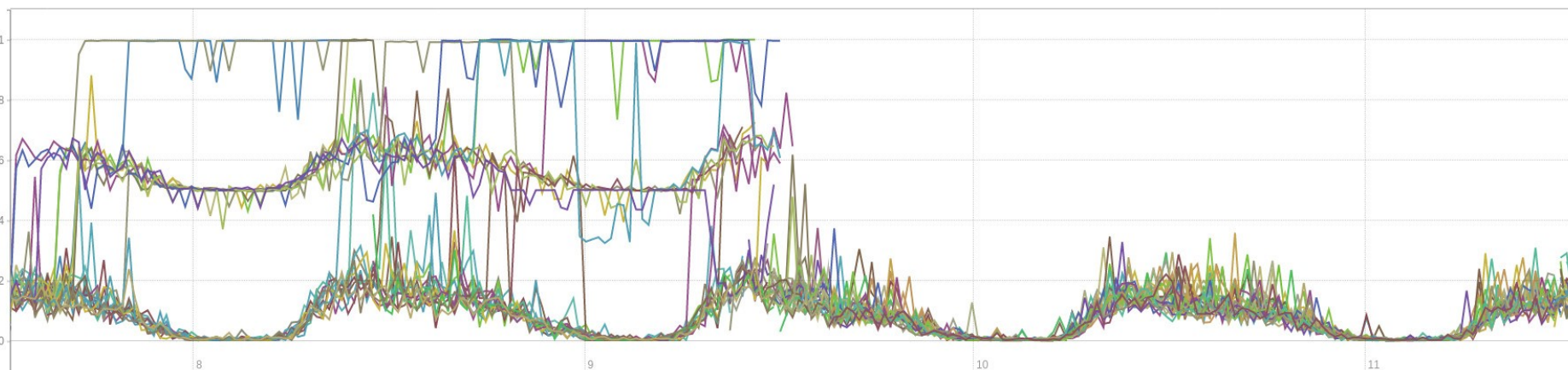
Cleanup of lost kit processes

- A kit process is lost if it's not referenced by loolwsd in either NewChildren or DocBrokers maps for a **particular amount of time** (**lost_kit_grace_period_secs** in loolwsd.xml)
- **lost_kit_grace_period_secs** should not be very low. The default value is 120 seconds. A value of 0 disables this cleanup.
- The pace of search for lost kits is governed by the **cleanup_interval_ms** setting. This is the same setting that applies to resource consuming documents cleanup.
- **kit_lost_terminated_count** metric counts the number of lost kit processes successfully terminated. It can be accessed through **/lool/getMetrics** REST endpoint.

Our experience with cleanup mechanism

- We are using it successfully already for ~2 years.
- 100-200 lost kits dismissed / 24h most on conversions
- 50-100 resource consuming kits dismissed / 24h all on editor
- An example of cleanup effect:

CPU usage / Time (days)



Thank you!



Online Office