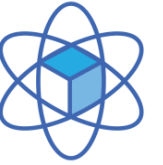


# Samba and CERNBox

Providing HA online access to Windows-based users

*Aritz Brosa Iartza, Giuseppe Lo Presti*

*CERN IT-ST*



# Home folders at CERN

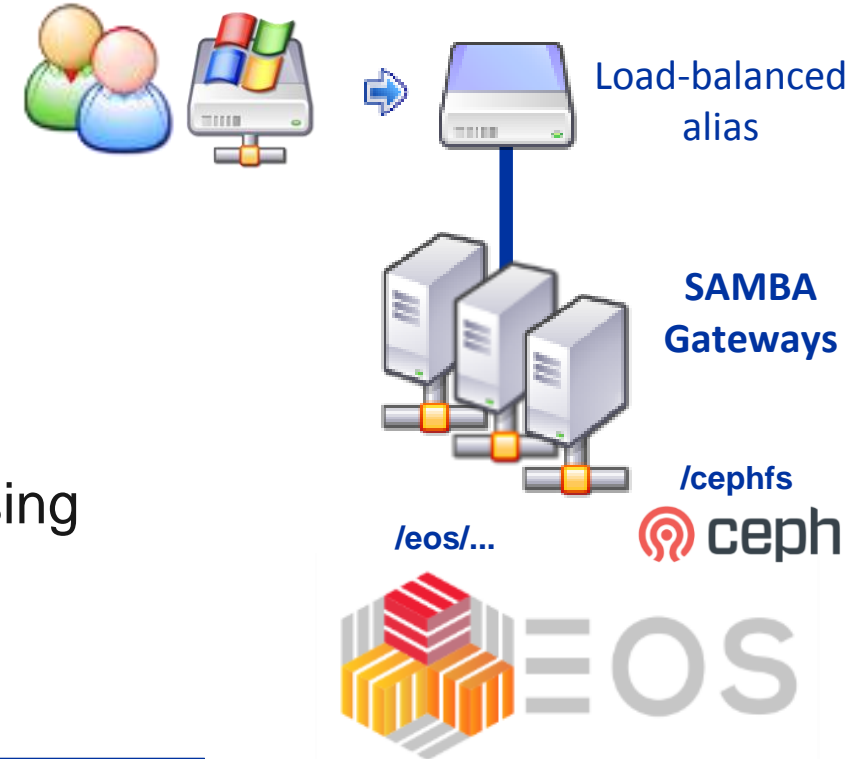
## EOS/CERNBox proposed as “user home space” to all CERN users

- Windows users being migrated out of DFS
- Sync client is seen as a plus, a DFS-like high-performance **online native access from Windows is a must**
- The Samba software suite currently is the natural choice to implement that
  - *Looking forward to the [EOS native client for Windows](#) we've just heard about!*

# HA Samba Architecture



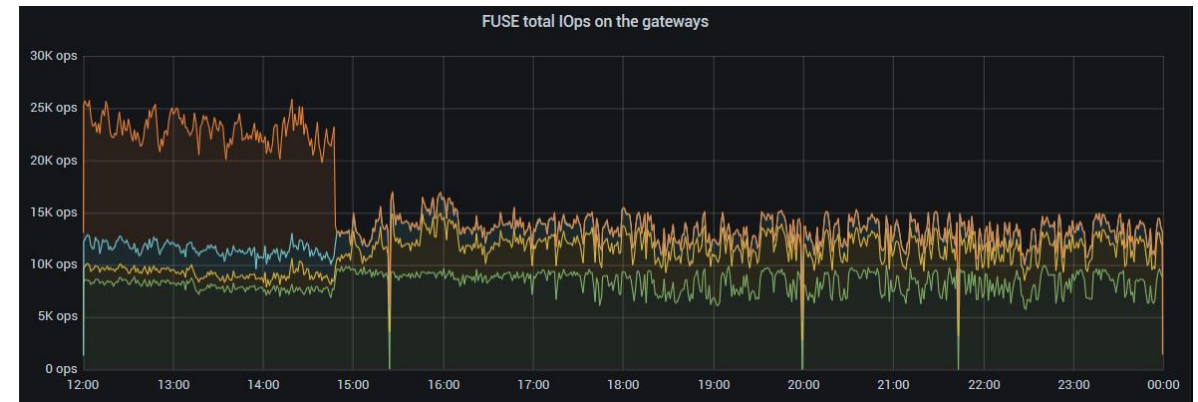
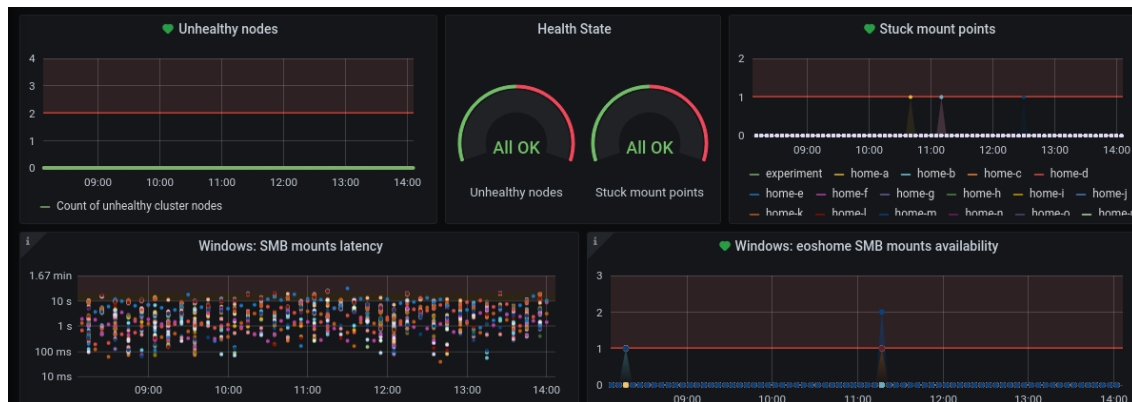
- A high-available, clustered Samba setup requires
  - *Floating IPs*, one per host
  - A small /cephfs mount to share the state
    - Yes, to provide a shared filesystem we need a(nother) shared filesystem...
- EOS FUSE mounted
- Windows Domain (AD) joined in **dedicated keytab** mode, using the **credentials** of the load-balanced alias
- **File locking supported, even across multiple gateways**
  - A must to support Microsoft Office notifications
  - Specific configuration in place with a VFS plugin



```
vfs objects = fileid  
fileid:algorithm = fsname
```

# Experience and issues

- In production since September 2019
- Samba over FUSE stack is **extremely latency-sensitive**
  - Many issues fixed in EOS to address latency
- Monitoring/alerting developed in Grafana from scratch
  - Aim to address issues before users are impacted



# Experience and issues: a few examples

## Folder browsing => fetch of all xattrs of all target files

- In POSIX jargon, `fstatat`, `listxattr`, and `getxattr` called for *each file* and *each extended attribute* of the folder
  - Aggressive caching + disabling of eos-specific system-level xattrs
- "Surprisingly", `smbclient` on Linux does NOT query all xattrs

# Experience and issues: a few examples

## Folder browsing => fetch of all xattrs of all target files

- In POSIX jargon, `fstatat`, `listxattr`, and `getxattr` called for *each file* and *each extended attribute* of the folder
  - Aggressive caching + disabling of eos-specific system-level xattrs
- "Surprisingly", `smbclient` on Linux does NOT query all xattrs

## Most recurrent filenames: **Thumbs.db**, **desktop.ini**

- Representing ~40-50% of the Samba traffic in terms of requests count!
- Considering caching them in a separate SSD-based pool

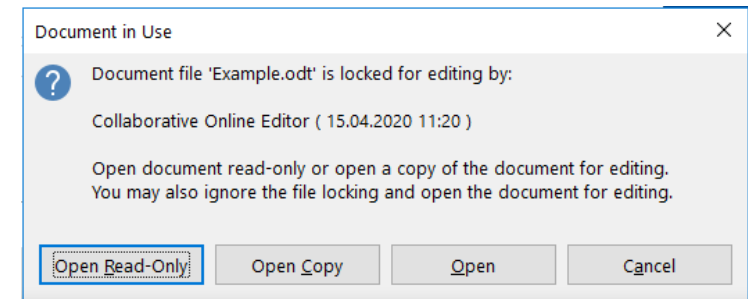
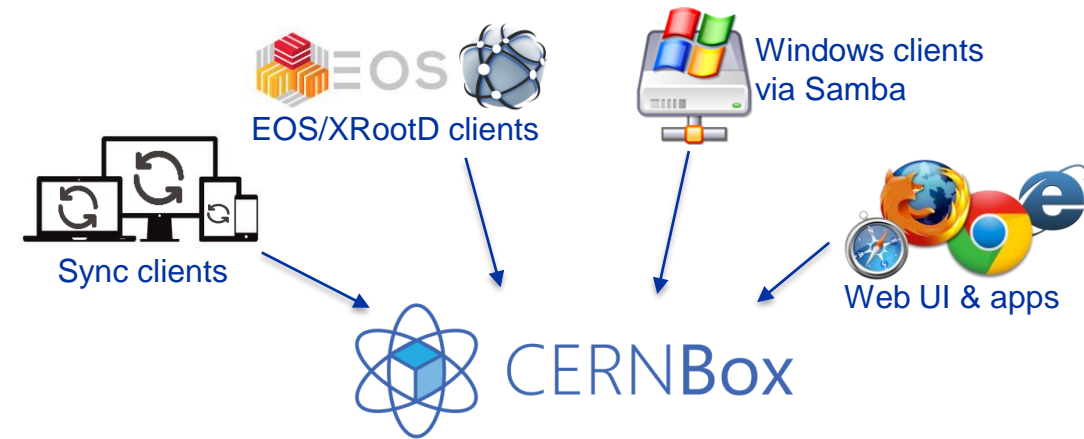
# Coming next

- **Usage growth** expected ahead:  
about to migrate more Windows-based use cases
- Possibly need to isolate the most demanding use cases  
in a separate EOS backend, optimized for low latency operations

# Conclusions: the bigger picture

## Multiple access paths . . .

- Windows Desktop/Documents/... system folders
  - Either synchronized, or mapped to Samba
- Office concurrent editing **across all platforms**
  - Desktop applications vs. web
  - See [OnlyOffice and Collabora at CERN](#)



. . . aiming at a coherent cross-platform UX



Thank you for your attention!  
Time for questions

