Open Heart Surgery with 60 Terabytes of Data

Migrating 40.000 users from PowerFolder to Nextcloud

Project summary
CS3, 25.01.2021

sascha.wiswedel@nextcloud.com

Situation

Customer / Industry	German university
Critical Business Issue	Running a state-wide file sync & share solution **since 2014** for **45 universities** and colleges – based on **PowerFolder**, including **ONLYOFFICE**
Problems / Reasons	– Unsatisfied with PowerFolder support – Outdated user interface – Unclear business situation of the PowerFolder company
Specific Capabilities	– **Migration of all data and shares** from PowerFolder to Nextcloud – Strict user provisioning workflow based on invitations
Critical Date / Value Realization Event	Semester spring break 2020, since there is less traffic affected by a migration downtime

The job: Service transformation

from PowerFolder to Nextcloud

60 TB data

> 65 million files and folders

> 40.000 users

> 50.000 shares

Project history wrap-up

- First contact: Oct 2018
- Proof of Concept: Nov 2018
- Kick-off: Jul 2019
- Go-live Nextcloud: 01. April 2020

Challenges

1. Special user de-provisioning workflow
2. Special user provisioning workflow
3. Migration from PowerFolder to Nextcloud

1. Special user de-provisioning workflow

Situation	High user fluctuation (students, teaching and research staff)
Need	Transfer files' ownerships to someone else before leaving
Challenge	So far only possible by admin and for whole accounts, but not selective
Solution	– **New function** with Nextcloud 18 – Users can invoke **selective ownership transfers** on their personal settings pages – Transfer target user has to **accept incoming transfer**

2. Special user provisioning workflow

Situation	SAML authentication against federated proxy (no LDAP)
Need	Share with entitled users who have never logged into Nextcloud before
Challenge	No way to share with unknown users
Solution	– Share by using sharee’s email address – Match either existing user or invite by creating a **guest account** – **Convert guest account** into SAML provided account – transfer files and shares from guest account to SAML account

3. Migration from PowerFolder to Nextcloud

Situation	PowerFolder = LAMP stack web application
Need	Lossless migration of - User accounts - User data - Shares (internal and public)
Challenge	– Match users and metadata from PowerFolder logic on Nextcloud logic – relocate and repair physical data
Solution	A well rehearsed weekend-long migration process ☞

Migration from PowerFolder to Nextcloud

Stats and technologies

- 40.000 users
- 60 TB data, 65 million files and folders
- 57.000 shares
- 72 parallel file scan threads on 4 Nextcloud servers
- 52 hours migration time over all

Used technologies for batch processing: python™ SAXONICA.COM XSLT AND XQUERY PROCESSING

A safe home for all your data

nextcloud.com

github.com/nextcloud

sascha.wiswedel@nextcloud.com

Appendix

Extra slides

Migration from PowerFolder to Nextcloud

The steps

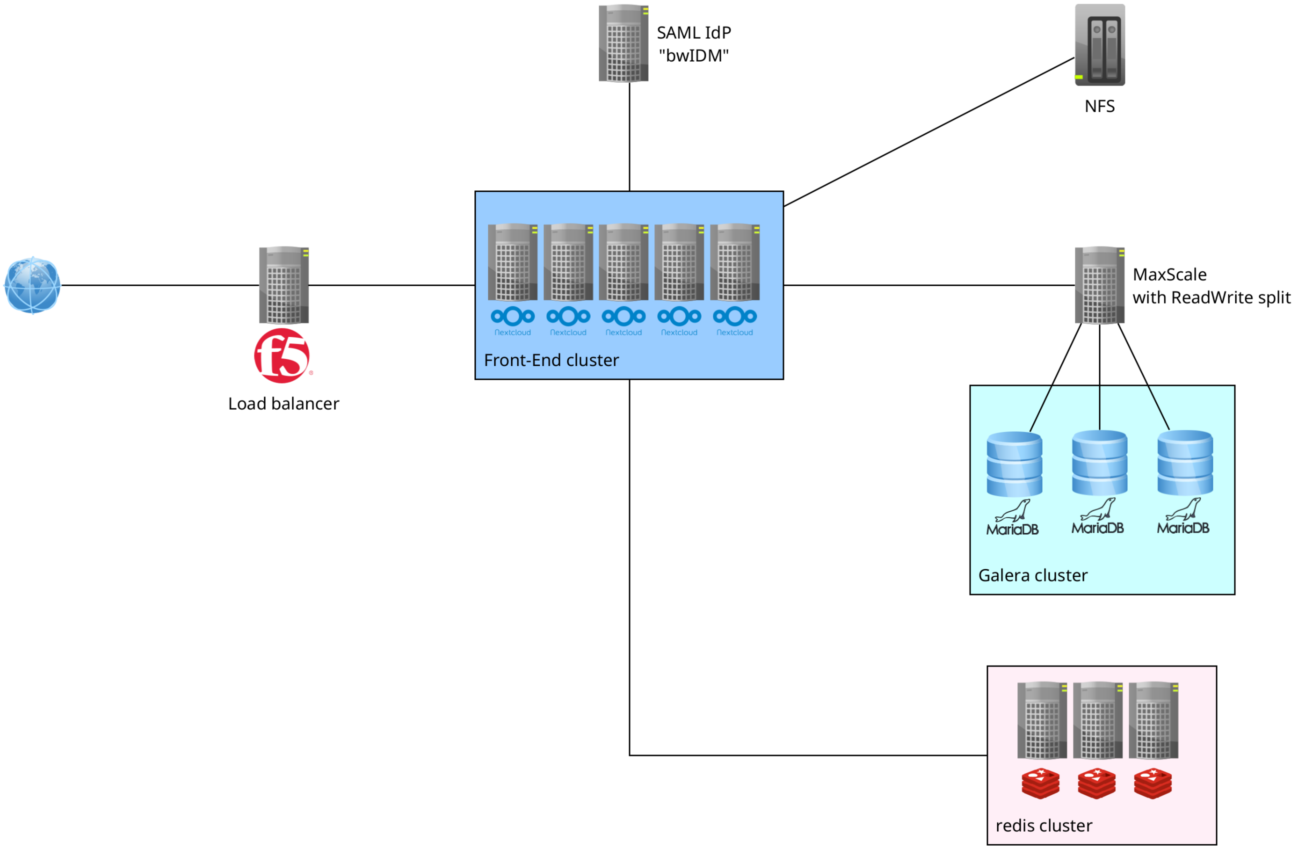
1. Extract users from PowerFolder database and create them in Nextcloud
2. Nextcloud filescan
3. Retrieve Nextcloud File IDs and map them on shared items
4. Create shares in Nextcloud
5. Export a list of password protected shares (passwords have to be manually re-set)

Migration from PowerFolder to Nextcloud

Challenges along the way

1. Map Shibboleth users on bwIDM IdP logic (newly introduced)
2. Distinguish between Shibboleth users and Guest users
3. Fix weird filename encodings
4. Delete PowerFolder meta files from the file system (all over the users' data folders)

Nextcloud architecture


SAML IdP
"bwIDM"
NFS
Nextcloud
Nextcloud
Nextcloud
Nextcloud
Nextcloud
Front-End cluster
MaxScale
with ReadWrite split
f5
Load balancer
MariaDB
MariaDB
MariaDB
Galera cluster
redis cluster