



Case Study

North-West University improves user storage mobility with Nextcloud and Collabora Online

May 30, 2018

Mr. Hans Erasmus, Junior Infrastructure Architect
Hans.Erasmus@nwu.ac.za

Potchefstroom, South Africa



October 2017 North-West University deployed Nextcloud and Collabora Online, providing students and employees a way to access, share and collaborate on their data on the go. Aiming for about 6000 users initially, the university has 3 physically distributed locations but maintains the main infrastructure from Potchefstroom. This case study presents the challenges NWU faces, why Nextcloud was chosen as most fitting solution and how it was implemented.



The challenge of providing a modern service

The North-West University is the result of several mergers over the last decades, with the last happening in 2004. The resulting three physical locations are not close – the furthest is over 200 kilometers away from Potchefstroom where the main infrastructure is kept. There are 2 data centers on the Potchefstroom campus, and one each on the Mafikeng and Vanderbijlpark campuses. Support comes in the form of a combination of local IT support offices on each campus and remote VNC sessions.

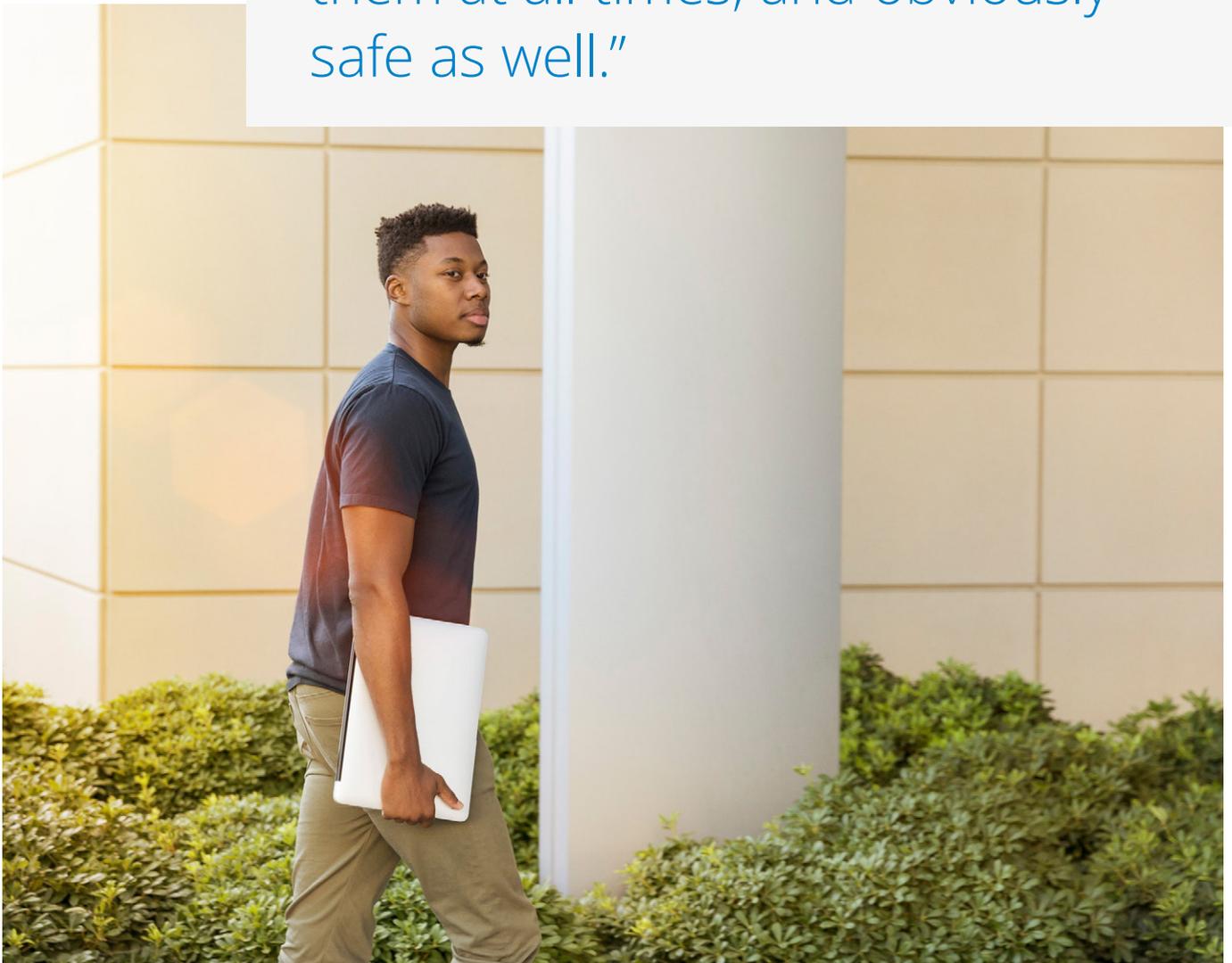
The university still uses Novell for identity management, offering a shared directory and of course managing the user home directories. A brilliant solution when introduced in the early nineties! Times have changed and the mapped network drive solutions from various vendors became a somewhat cumbersome and rather outdated concept that was “missing the futuristic bus in the last decades and failing to innovate”, according to Hans.

Students, lecturers and researchers are all used to the concept of “data on the go”. Google provided that with Google Drive and the method is so “indoctrinated” in people’s mind that forcing a 90’s style mapped-drive solution on users is not appreciated: “People laugh at you”.

The only way to access data from the network drive remotely was to use a VPN connection to the university. This was hugely limiting to, for example, researchers working overseas, in no small part because the VPN technology the university used had limited client support. It was time for another solution. Keeping up with the constant change in how different operating systems and especially mobile clients operate, it was time for a new approach, a new era so to speak. It was time for a new solution. Any file, anywhere, any time, any device.

Hans: "The IT world has changed in such a way that if the institution, albeit a university, a company or any other organization, does not cater for the needs of their employees, those employees will just go and find these solutions on their own. This can easily be a cause of chaos, as you will end up with 100's or even 1000's of different individual solutions to try and supply end-user support for. So the only logical, rational path to follow is to try and keep up with the technology without losing your sanity. We needed a sync client and all the pertaining features. People want to be mobile, and they want their data to be as well. Above everything else, it seems that end users want their data with them at all times, and obviously safe as well. Hence, a locally hosted cloud solution."

"Above everything else, it seems that end users want their data with them at all times, and obviously safe as well."



Searching solutions

The new technology should offer at least:

- A replacement for the Network Drive
- Be able to sync data for remote, offline work
- Feature good mobile clients
- Allow storage on SWIFT / OpenStack / Object storage
- Offer online, real-time document editing
- Be an easy fit in the existing infrastructure at the university

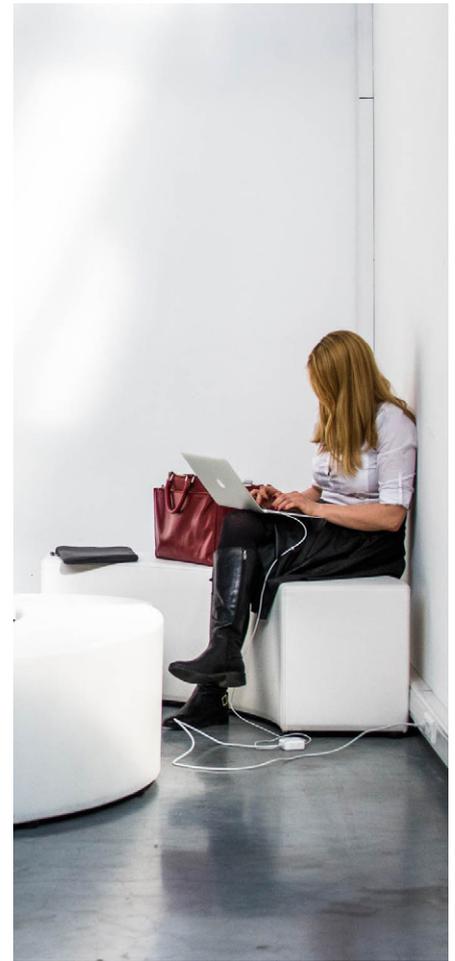
A few EFSS products were tried. A PoC was done with another open source solution but it was deemed too immature to fit the use case of the university. A proprietary solution from another vendor was also in limited use at the university. Desktop clients for all users and platforms as well as good mobile clients were very important to the team so the proprietary nature of the solution and its limited platform support meant it was dropped from the list of options.

The team settled at ownCloud and started a bigger PoC, test driving it within the IT department. Several problems cropped up. The university used SWIFT storage on OpenStack using v3 authentication. Using this as primary storage was initially not supported and when it came with ownCloud 9 it had big technical problems including inability to deal with files bigger than 5 gigabytes.

Hans: "The OpenStack software automatically chunks files larger than 5GB into smaller chunks, and upon successful upload, will present these chunks as a single object to the user again. The ownCloud interface couldn't handle these requests, and it was not clear what the limiting factor was."

This was a serious problem because OpenStack Swift object storage has proven to be the cheapest method of storing large quantities of data (Petabytes) as needed by the university.

The team then found out about Nextcloud, which supported chunking. There were widespread rumours about universities dropping ownCloud and moving to Nextcloud. After some deeper research into development strategies and business models it became clear to the team that Nextcloud was the future and the team moved over.



“There were widespread rumours about universities dropping ownCloud and moving to Nextcloud.”

Implementing Nextcloud and Collabora Online

Hans noted that “the architecture for Nextcloud is laid out well”, making the technical side of implementing it easy. Some concepts were unfamiliar – the team had no prior experience deploying PHPbased solutions and the RHEL servers posed a bit of a challenge. Hans settled on the remi package for PHP¹ and “all was a breeze from there”.



“After some deeper research [...], it became clear to the team that Nextcloud was the future and the team moved over.”

One of the first things users asked after the university implemented Nextcloud was: “Does this mean we can now edit files online? Or work on them simultaneously?”

Hans: “Unfortunately initially the answer was no. But as the service became more popular and more users started working on it, the need grew and it became apparent this is something we would need to implement. Hence, Collabora Online.”

One of the issues any institution runs into when implementing any form of self-hosted open source solution is complexity. Another is scalability. The university did not want to add unnecessary complexity, so the logical choice was to stick to what is known. The seamless integration between Nextcloud and Collabora Online meant the choice was easy to

make. Part of the “available anywhere” approach is to be able to read, update and even delete files from anywhere, any time on any device. Hans notes that “people have become so used to Google docs, that anything similar was the only way of meeting this expectation.”

Control over the data was a primary concern and deciding factor in opting for Collabora Online: “Why not use Google docs you might ask? Well, some laws prohibit us from keeping certain types of data in “the cloud”. It is not allowed outside the boundaries of our country. Data security is therefor key. We needed a solution which could be hosted on premises, where we have total control of the data.”

The team is still coming to terms with maintenance, but Hans does not foresee any issues in the long run.

¹ <https://rpms.remirepo.net/>

“Our systems department is working closely with your technical department to resolve any issues that come up. We try to keep the major updates / upgrades to a minimum to reduce the downtime on this service. Our pilot has expanded to include researchers who needs a safe haven for their large data sets. This means that already we’re held accountable for the uptime of the service.”

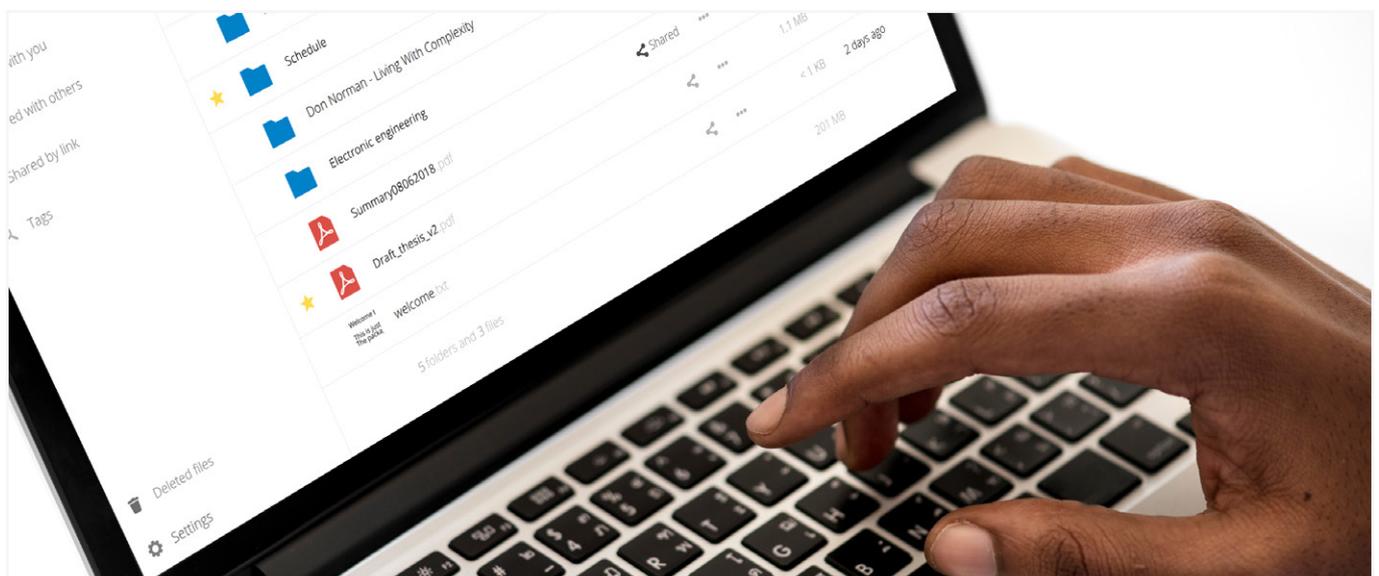
The users are very happy. Hans: “Because they are familiar with products like Google Docs etc., the concept of having stuff in the “cloud” and accessing it through a web browser was not unfamiliar to them. Initially they had some questions about a few functional things, but we sorted those out quickly enough. The product itself is very intuitive (IMHO) so spending a little time on it is all that is necessary to understand how it works.

Our users mostly ask question relating to Collabora Online, as this feature is used a lot. The initial implementation with the load balancer required up to date documentation, which is now provided by the Nextcloud Customer Portal.”

“The most used feature”, Hans noted, “is the ‘available anywhere’ aspect: Working on a document at work, going home, and keeping on going where you left off. The feedback we get is that Nextcloud helps a lot with sharing stuff effortlessly with other people in the organization.”

Asked about what could be improved, Hans shared that he’d like to see the Collabora Online documentation improve further, and for Nextcloud itself: “Things like how much additional disk space is needed for file versions etc. was at one point a concern for us. Not knowing if we have catered enough storage space for versioning, trash, etc., to work properly.”

“The feedback we get is that Nextcloud helps a lot with sharing stuff effortlessly with other people in the organization.”



Hardware setup

The setup used by the North-West University relies on 2 application servers running Apache, one database server with MySQL, one NFS provisioned with CEPH storage and one Collabora server. Authentication happens with Shibboleth connected to a LDAP. We use CAS as SSO provider. There is also a Redis server in the mix. The App servers are load balanced (least con method) with HA-Proxy load balancer. All users connect to the load balancer. This also helps with security, as the only server reachable from the outside is the load balancer.

All of the servers are virtual machines (Vmware) and run CentOS-7. The university has the following infrastructure assigned:

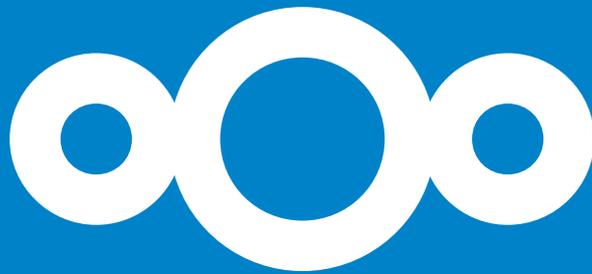
Load balancer	Application server 2	NFS server
CPU – 2x	CPU – 8x	CPU – 2x
RAM – 4Gb	RAM – 16Gb	RAM – 2Gb
Disk – 50 Gb	Disk – 50Gb + 200GB	Disk – 50Gb + 64TB
Application server 1	Database server	Collabora server
CPU – 8x	CPU – 8x	CPU – 8x
RAM – 16Gb	RAM – 16Gb	RAM – 32Gb
Disk – 50Gb + 200GB	Disk – 50Gb + 60GB	Disk – 50Gb + 500GB

In the future, the university will again investigate the possibility of migrating to using Object Storage directly on OpenStack - Nextcloud 14 is scheduled to provide the capabilities needed for this. There are still some technical requirements that have to be determined and “ironed out” before this implementation will be commissioned.

Conclusion

North-West University moved to a Nextcloud solution backed by a Nextcloud Enterprise Subscription. Starting off with a self installed implementation, it was only natural progression that led to the purchase of the subscription. The system is now running well and is providing the university with a future-ready, always adapting and market leading technology to improve productivity for employees.





Nextcloud GmbH
Hauptmannsreute 44A
70192 Stuttgart
Germany

Email sales@nextcloud.com
Phone +49 711 252428-90
Fax +49 711 252428-20

nextcloud.com