Nextcloud

Case Study

tubCloud migration to Nextcloud 11

June 2017

Dr.-Ing. Thomas Hildmann
thomas.hildmann@tu-berlin.de

tubIT IT-Service-Center TU Berlin

Readying cloud storage for a future with new goals and challenges

May 2017 the Technische Universität Berlin migrated from ownCloud to Nextcloud, achieving over 38 % lower database load at peak times. This brought faster service for users as well as user interface improvements and better security. The migration was planned and executed in under 4 weeks without any problems and did not need the stand-by support from Nextcloud. This case study details the architecture and infrastructure at the TU Berlin, migration planning and execution as well as the results.

History

tubIT, TU Berlin's IT-Service-Center, began providing a cloud storage/sync-n-share service sized for up to 30.000 students, employees and guests of the University in May 2012. This made the TU Berlin one of the first big universities providing a self-hosted Dropbox/Google Drive alternative for all members of the university.

After evaluating the available sync-n-share software solutions in 2011, the TU Berlin started with ownCloud 5 community edition and changed to the Enterprise Edition when the service became critical for thousands of users at the University.

TU Berlin was one of the first providers of cloud storage within the DFN cloud research program.

The infrastructure is now used by 16 other DFN member institutions, consisting of many research and higher education organizations in Germany.

While initially, only a few thousand users relied on the system, usage has grown steadily. Late 2014, about 7.000 users had been active and in 2017 the service was used by about 22.000 users at TU Berlin.

More than 80.000 out of 100.000.000 files are changed daily in the tubCloud. About 70 TB disk space is used at the moment for TUB itself with more in use by the other DFN members.

Architecture

tubIT employs a setup with F5 Big-IP load balancers ahead of multiple web application servers. Data is stored on a GPFS-based storage system. For running the database, tubIT decided to use a Galera MySQL cluster together with an SQL Proxy for read/write splitting. The cluster nodes are running in in LXD containers managed with OpenStack. File locking and session management are provided using Redis, also on LXD managed with OpenStack. Authentication is handled through LDAP and Kerberos.

Currently, there are 4 application servers in use, running an LNMP (Linux/NGINX/MySQL/PHP) stack with PHP 5.6 on CentOS 7.3. Additional application servers are running the DFN cloud instances. A locally replicated LDAP instance runs on each of the application servers. Each server has 16 virtual CPU cores with 18 GB of memory.

The Galera cluster employs 3 containers with 32-core CPUs and 128 GB memory each. Its performance is very critical to the performance of the overall cloud. To serve more incoming queries the number of instances can be increased, however, every new instance increases the resources required to synchronize within the cluster, limiting scalability.

The GPFS cluster is load balanced between four Spectrum Scale nodes to improve response times and provide failover.

Students can use a maximum of 20 GB of data, while staff members can go up to 100 GB and units up to 250 GB.

Decision to migrate

The first support contract ended in May 2017. tubIT issued a tender to the most popular players in the sync-n-share market. Nextcloud won the tender. The contract covered the migration of the TU Berlin and all DFN partners without their own service contract. It also included support for the next two years including all updates and improvements we would like to implement.

A partnership with Nextcloud gives us the opportunity to re-adjust our service to a more collaborative vision. This helps us fulfill our main goal of providing the university's members with the best infrastructure for learning and researching.

Nextcloud 11, the version we migrated to, promises significant improvements to scalability and security, on top of a number of refinements to the web interface.

"A partnership with Nextcloud gives us the opportunity to re-adjust our service to a more collaborative vision."

New security capabilities like brute force protection and browser security improvements like CSP 3.0 and same-site cookies will improve our resilience to hacking attempts. Thanks to the native SAML app it was possible to migrate from Apache to NGINX, from which we expected to achieve lower memory usage.

tubIT users and management want to explore further capabilities available or under development in Nextcloud. The text, audio, and video chat features will enable closer and more efficient collaboration and offering real-time document editing is another area of interest. The possibility of providing access to data through other protocols like NFS/Samba/RSYNC/SFTP while respecting access rights and other restrictions are under investigation.

We are also excited about the upcoming Global Scale architecture which gives us further options for scaling the DFN cloud services and decreasing total cost of ownership significantly.

"Overall, the migration was like a typical upgrade."

Migration

tubIT regularly announces a calendar with update windows (down time) for the whole year. Our challenge was to make a migration plan from ownCloud 9 to Nextcloud 11 including the branded clients (Windows, Mac OS, Linux) and mobile apps (iOS, Android) in about three weeks and execute in one week, fitting the update window.

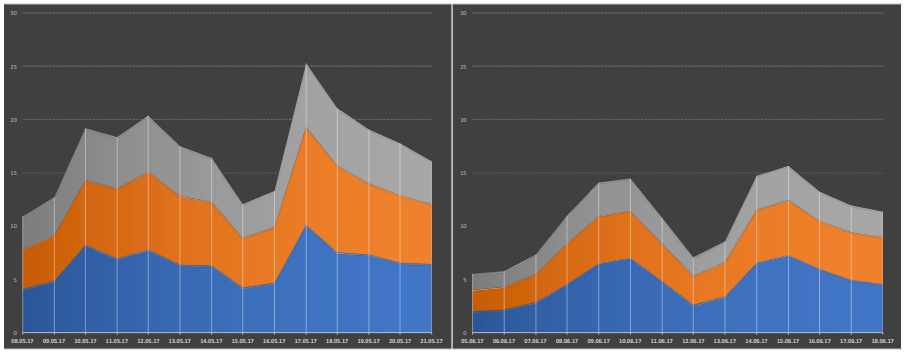
Our test environment is based on a production like VMware/RedHat environment. The database was cloned to a much smaller Galera cluster on OpenStack. The test runs showed some issues that were fixed or could be solved with minor environment/migration process adaptations. We used the opportunity of this migration to migrate from Apache to NGINX, made possible thanks to Nextcloud's native SAML support. We expected NGINX to lower the memory usage of our application servers. We also made some major OS and deployment updates, including an upgrade from RHEL 6 to CentOS 7 with current packages of OpenLDAP, NGINX and a GPFS update.

The migration process itself was ready in time and worked exactly as planned. We did not need our stand-by support from Nextcloud. The first week uncovered two issues that were fixed on-site by Nextcloud. Overall, the migration was like a typical upgrade.

The first problem encountered was related to PHP errors. A second problem, related to the first, were large spikes in database queries caused by empty sessions from the Redis servers at a specific time on the day caused by restarting Redis into a pristine state due to maintenance work. This caused all clients to login again and thus sync metadata per user from LDAP. The problem was solved by emitting the 'user change' event only in those cases when user information was indeed modified.

First impressions

ownCloud 9 used to create heavy peak load on our database cluster during working hours, limiting scaling. We no longer see this behavior after the migration. Seeing our load since the migration we estimate that database peak load has decreased by more than 38 % by moving to Nextcloud. The off-peak times even show an impressive 49 % lower load. This gives us confidence that we can scale up our service without requiring more hardware investments the coming years.



> "Seeing our load since the migration we estimate that database peak load has decreased by more than 38 % [...]"

The numbers in Table 1* are the sum of the daily average load on the database servers during the two weeks before and after the migration. The average load of ownCloud is 17, the average of Nextcloud is 11. Nextcloud thus achieves a 37 % lower database load on the Galera Cluster. The maximum average values are 38 % lower than with ownCloud 9 and the minimum (average) load is nearly 50 % lower.

	ownCloud	Nextcloud	% decrease
average	17,14	10,81	36,94
max	25,26	15,63	38,12
min	10,88	5,51	49,36

Unsurprisingly, the peak load on the front end servers actually increased a few percent. This is due to the faster response times from the database cluster which enables more transactions per second. The result is a better user experience with a snappier user interface, faster syncing and up and downloads for the users.

The migration has prepared TU Berlin for growth by lowering the load on our infrastructure. It also puts us in a position to benefit from the fast feature development in Nextcloud, bringing new capabilities as well as further performance, scalability, security and stability improvements to our deployment.

About tubIT - IT Service-Center TU Berlin

The tubIT IT-service-center provides network- and base-infrastructure for research, teaching, and administration of the Technische Universität Berlin since 2006. All processes of tubIT are students and staff member centered.

We aim at consistently implementing the cooperative supply model: efficient supply with basic services and support of the development and introduction of an IT infrastructure adapted to research requirements. tubIT sees the employees and students are the center of all processes and understands itself as a companion for research and education in an ideal environment.

For more information, visit tubit.tu-berlin.de or follow tubIT_TU_Berlin on Twitter.

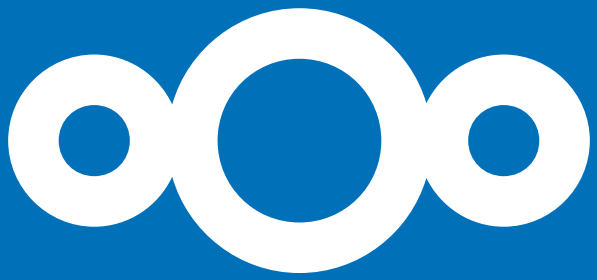
About Nextcloud

Nextcloud offers the industry-leading fully open source solution for on-premise data handling and communication with an uncompromising focus on security and privacy and unprecedented scalability. Nextcloud brings together universal access to data with next-generation secure communication and collaboration capabilities under direct control of IT and integrated with existing compliant infrastructure. Nextcloud's open, modular architecture, emphasis on security and advanced federation capabilities enable modern enterprises to leverage their existing assets within and across the borders of their organization.

For more information, visit nextcloud.com or follow Nextclouders on Twitter.

* Statistics on server load calculated as follows over data gathered by the TU Berlin:

- daily average load from monitoring
- sum of the three cluster nodes (overall load)
- the two weeks before the two migration weeks
- the two weeks after the two migration weeks
- min, max and average from data in the charts for the percentage



Nextcloud GmbH
Kronenstr. 22A
70173 Stuttgart
Germany

Email hello@nextcloud.com
Phone +49 711 896656-0
Fax +49 711 896656-10

nextcloud.com