

Nuvem de Dados Corporativa - Um Caso de Sucesso

Luciano Gonda, Brivaldo A. S. Junior¹

¹Faculdade de Computação (FACOM)
Universidade Federal de Mato Grosso do Sul (UFMS)
CEP – 79070-900 – Campo Grande – MS – Brasil

{gonda, brivaldo}@facom.ufms.br

Abstract. *Cloud storage is a reality for a lot of users that want the peace and agility to be capable of access their important data at any time or device. This reality brings challenges to IT and should be carefully evaluated.*

Resumo. *O armazenamento em nuvem é uma realidade e cada vez mais os usuários desejam a tranquilidade de ter seus dados mais importantes disponíveis a qualquer momento ou dispositivo. Essa realidade traz desafios para a TI e deve ser avaliada cuidadosamente.*

Introdução

A crescente demanda por armazenamento nas organizações públicas tem guiado as decisões dos administradores para o uso de mecanismos de compartilhamento e armazenamento de dados em servidores locais com uso adicional de *storages* e centros de dados com disponibilidades próximas a 98%. Para reduzir custos, optou-se por usar tecnologias gratuitas ou com custo reduzido como Google Drive, Amazon, DropBox, para armazenamento de dados tanto pessoais quanto corporativos. O termo mercadológico para armazenamento na Internet recebeu o nome de nuvem [Peterson and Davie 2012]. Contudo, armazenar dados corporativos em nuvens fora do Brasil deixou de ser aceitável a partir de 04 de novembro de 2013 com o Decreto 8.135 [Presidência da República 2013]. Este decreto define as regras relacionadas a comunicação de dados da administração pública federal direta, autárquica e fundacional.

Dessa forma, e com este decreto, dados institucionais como email, documentos e outras informações do serviço público passaram a ser consideradas importantes e com possibilidade de causar danos a segurança nacional em caso de vazamentos. Isso significa que a partir deste momento, nenhum email corporativo, documento ou dados institucionais podem ser armazenados de forma insegura (ou fora do país) dentro dos moldes estipulados no decreto. Para resolver este problema, a Faculdade de Computação (FACOM) da UFMS replicou diversos serviços como: email, Youtube, páginas corporativas, blogs e, principalmente, ferramentas equivalentes ao DropBox, Google Drive e OneDrive.

Este artigo está assim organizado: na Sessão 2 serão abordadas as necessidades do órgão e as legislações cabíveis. Na Sessão 3 aspectos importantes a serem considerados na implantação do serviço de nuvem de dados. Na Sessão 4 os resultados obtidos após a adesão dos clientes ao novo serviço e os principais problemas detectados. Finalmente, na Sessão 5 as considerações finais e trabalhos futuros.

Tabela 1. Comparação de tecnologias de armazenamento de dados na nuvem.

Ferramenta	Hospedagem	Espaço Gratuito	Plataformas Suportadas
OneDrive	Remota	10GB	Windows, Mac e Mobile
Google Drive	Remota	15GB	Windows, Mac e Mobile
DropBox	Remota	5GB	Windows, Mac, Linux e Mobile
Owncloud	Privada/Remota	Limite do seu espaço de armazenamento	Windows, Mac, Linux e Mobile
Nextcloud	Privada/Remota	Limite do seu espaço de armazenamento	Windows, Mac, Linux e Mobile

Necessidades e Legislação

A necessidade da nuvem de dados para a Faculdade de Computação era possibilitar aos docentes e técnicos administrativos um local privado na instituição para armazenamento de documentos administrativos e relacionados a pesquisas científicas. A UFMS, como fundação, não possui legislação específica para atender ao Decreto 8.135. As definições e estratégias de Tecnologia da Informação são descritas no PDTI (Plano Diretor de Tecnologia da Informação) revisado a cada dois anos pela AGETIC (Agência de Tecnologia e Comunicação). Desta forma, como não existe legislação específica, seguiu-se a legislação federal.

Os professores e técnicos utilizavam tecnologias como Google Drive, DropBox ou OneDrive, mas sempre limitados pelo espaço fornecido gratuitamente por estes serviços ou pela plataforma operacional. Enquanto os técnicos utilizam a plataforma operacional fornecida pela AGETIC, que é o Windows, vários docentes tem maior liberdade e usam outros sistemas como MacOS e Linux (em suas mais variadas versões).

Mesmo com a possibilidade de uso de ferramentas gratuitas, nem todas tem suporte para todas as plataformas operacionais. A Tabela 1 apresenta de forma sintetizada o suporte das principais tecnologias de armazenamento de dados (incluindo as com tecnologia aberta) e o suporte a plataformas operacionais diferentes. Ou seja, uma tecnologia para ter aceitação pela comunidade acadêmica deveria executar nas principais plataformas e oferecer algum diferencial.

Após análise criteriosa, optou-se pelo uso do Owncloud [Owncloud 2017], que em seguida foi atualizado para o NextCloud [Nextcloud 2017], pois este tem estado mais ativo em termos de desenvolvimento, é mais rápido, mais leve e seguro que o Owncloud. O gerenciamento de aplicações de terceiros, 2FA (*two factor authentication*) e atualização simplificada, tornaram o NextCloud a ferramenta padrão para armazenamento de dados na nuvem. Além disso, tanto o NextCloud quanto o OwnCloud permitem que a ferramenta seja hospedada em servidores da própria instituição ou em alguma empresa de hospedagem e também possibilitam o sincronismo federado e conexão com outras nuvens de dados como o DropBox e Google Drive na mesma plataforma de acesso.

Implantação

O NextCloud como nuvem corporativa executa em um único servidor (embora exista a possibilidade de uso bem descrito na documentação para criar um ambiente com balan-

ceamento de carga) executando Debian 8.7 em um servidor virtual com 16GB de RAM, 8 núcleos de 2.0Ghz Xeon e espaço de 2TB em um *storage* na rede (conexão via iSCSI).

As aplicações e ajustes utilizados no serviço foram:

- Apache 2.4.10:
StartServers 100
MinSpareServers 100
MaxSpareServers 2000
ServerLimit 6000
MaxClients 6000
MaxRequestsPerChild 4000
- PHP 5.6.30:
Limite de Memória: 512MB
Tempo Máximo de Execução: 3600s
Upload Máximo: 511MB
- PostgreSQL 9.1 (ocupando 3.0GB de dados para referenciamento de arquivos).

O uso do Apache ao invés do Nginx (alternativa mais rápida para diversos cenários) foi pela automação e simplicidade do uso do Apache com o Certbot [Foundation 2017] para criação e manutenção de certificado digital para a nuvem sem necessidade de contratação de um serviço de terceiros (precisa ser renovado a cada três meses e a renovação automática é simples com o Apache) e o conhecimento da equipe com a segurança deste servidor web [Ristic 2005].

A Figura 1 mostra o uso intensivo das CPU's em relação a sincronização constante dos usuários e interface de acesso web. Nesse caso, o responsável por esse uso intensivo é o Apache. Em comparação, o PostgreSQL, servidor de banco de dados SQL, no mesmo servidor, não causa tanto impacto. Testes com Nginx mostraram que a carga das CPU's reduz em torno de 25%, mas vários módulos automatizados deixaram de funcionar.



Figura 1. Uso da CPU com média de 35 usuários simultâneos.

Este servidor já está em execução a aproximadamente 4 anos e o consumo de dados não cresceu muito. Cada usuário do sistema (que é integrado a uma base de dados LDAP com Samba4), possui uma cota de espaço de 200GB. O uso atual do *storage* de dados deste servidor está em torno de 75%. O crescimento inicial era em torno de 50GB/mês, mas agora, após a fase inicial de sincronização, estabilizou e varia menos de 1GB/mês. Já a média de uso é de 15GB por usuário. Mesmo que o ambiente não suporte



Figura 2. Uso da memória com média de 35 usuários simultâneos.

todos os 200GB por usuário de fato, a referência de uso estatística mostra que a média de uso é do espaço fornecido é de 7,5% no cenário acadêmico verificado.

Muito embora o servidor possua 16GB de RAM, o sistema não chega a usar no uso diário (com o sistema, ignorando o SO e outros programas adicionais) menos de 1GB. Mesmo com uma quantidade grande de links compartilhados, como aparece na Figura 3. Notem na Figura 2 a quantidade necessária de memória RAM pelo Nextcloud durante o uso diário.

Resultados

A adesão após a implantação do sistema levou em torno de 6 meses entre a ativação do serviço como experimental e o início do uso pelos clientes (docentes e técnicos administrativos da FACOM). O principal motivador do uso da cloud privada via Nextcloud foi a compatibilidade com diversos sistemas operacionais e pelo fato da oferta inicial de espaço por usuário ser de 50GB, o que era 5x mais do que o OneDrive e 3,3x mais do que o Google Drive. Após a primeira fase de adesão, foi realizada a primeira expansão de espaço, seguindo para 100GB.

Com a segunda expansão de espaço, praticamente todos os docentes e técnicos requisitaram o uso do serviço. Após dois anos, realizamos a última expansão para 200GB por usuário. Esse valor é mais do que suficiente para qualquer necessidade acadêmica ou administrativa atual.

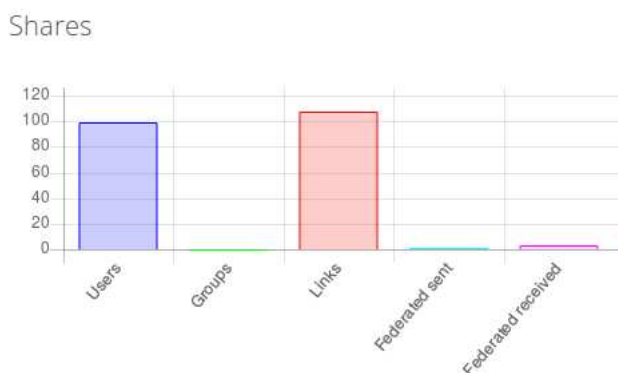


Figura 3. Quantidade de usuários e links compartilhados publicamente.

Com relação aos problemas, os principais foram encontrados nas primeiras versões do Owncloud com relação ao sincronismo de grandes quantidades de arquivos,

já que a tecnologia usada para sincronizar é o WEBDAV, as sessões HTTP acabavam falhando após uma quantidade muito grande de requisições ao servidor Apache. Outro fato foi com relação ao próprio Owncloud que tinha problemas de código ineficiente para mostrar dados dos usuários no navegador de Internet, causando uso excessivo de processamento. Com a migração para o Nextcloud, além da possibilidade de alterar o logotipo e mensagens do sistema, ficou mais fácil manter o ambiente atualizado e com correções de segurança [Spiegel 2017]. Além da redução do uso intensivo de CPU pelo Apache.

Conclusão

A criação de uma nuvem de dados privada dentro da instituição trouxe adesão ao processo iniciado pelo Decreto 8.135, além de facilitar e agregar a possibilidade de aumentar o espaço cedido aos professores e técnicos dependendo apenas da disponibilidade de recursos para compra de armazenamento. Outro fato importante foi a possibilidade de agregar também os acadêmicos (segunda etapa deste projeto).

Para tornar a nuvem privada ainda mais confiável, a equipe está trabalhando na implementação do Nextcloud com balanceamento de carga e integrado ao servidor web mais versátil e rápido Nginx em substituição do Apache. O processo de migração, pelo menos, é transparente para o usuário final.

Referências

- Foundation, E. F. (2017). Automatically enable HTTPS on your website with EFF's Certbot, deploying Let's Encrypt certificates. <https://certbot.eff.org/>. [Online; acessado 07-Abr-2017].
- Nextcloud (2017). A safe home for all your data. <https://nextcloud.com/>. [Online; acessado 07-Abr-2017].
- Owncloud (2017). A safe home for all your data. <https://owncloud.org/>. [Online; acessado 07-Abr-2017].
- Peterson, L. L. and Davie, B. S. (2012). *Computer Networks: A Systems Approach*. Elsevier.
- Presidência da República, C. C. (2013). Decreto Nr. 8.135, de 4 de Novembro de 2013. http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2013/Decreto/D8135.htm. [Online; acessado 07-Abr-2017].
- Ristic, I. (2005). *Apache Security: The Complete Guide to Securing Your Apache Web Server*. Feisty Duck.
- Spiegel, D. (2017). Nextcloud security scanner. <https://scan.nextcloud.com/>. [Online; acessado 07-Abr-2017].