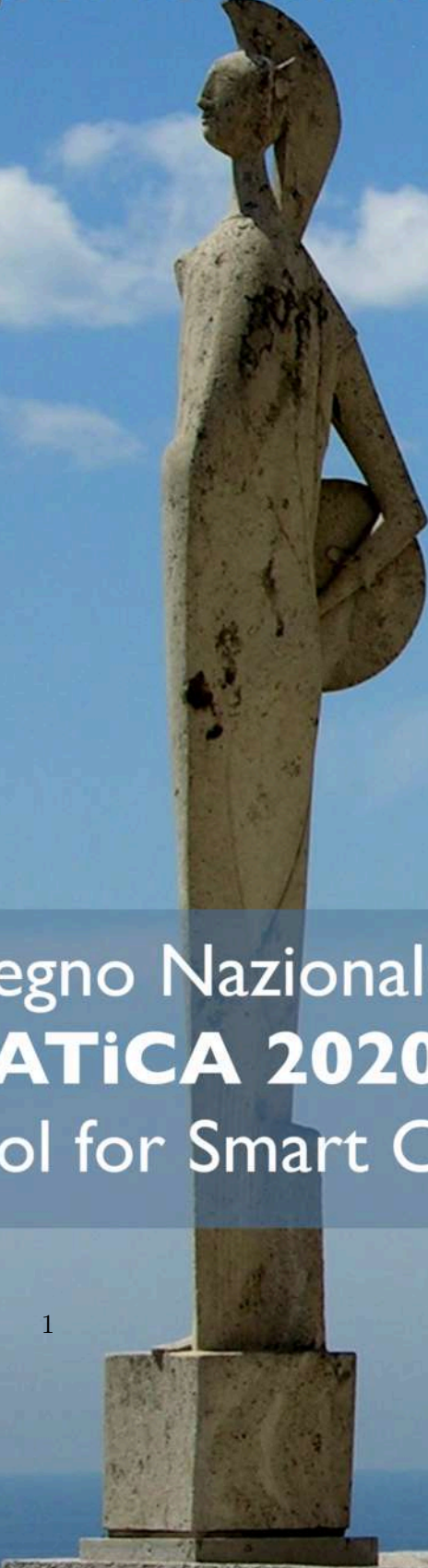


# DIDAMATICA

*informatica per la didattica*



Atti Convegno Nazionale  
**DIDAMATICA 2020**  
"Smarter School for Smart Cities"

# Atti Convegno Nazionale DIDAMATiCA 2020



UNIVERSITÀ  
DEGLI STUDI DI TRIESTE



A cura di: Giovanni Adorni, Andrea De Lorenzo, Luca Manzoni, Eric Medvet

Risorse e aggiornamenti relativi a questi Atti sono disponibili all'indirizzo  
[www.aicanet.it/didamatica2020](http://www.aicanet.it/didamatica2020)

Copyright©2020 AICA-Associazione Italiana per l'Informatica ed il Calcolo Automatico  
Piazzale Rodolfo Morandi, 2 - 20121 Milano  
Tel. +39-02-7645501 - Fax +39-02-76015717  
[www.aicanet.it](http://www.aicanet.it)

The cover of the DIDAMATiCA 2020 proceedings is distributed under the Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0) license (see <https://creativecommons.org/licenses/by-sa/3.0/>). Original photo by user Picchio4ever ([https://it.wikipedia.org/wiki/File:Minerva\\_units.jpg](https://it.wikipedia.org/wiki/File:Minerva_units.jpg)).

Edizione del 17 Novembre 2020

ISBN: 978-8-89-809161-4

# Remotely accessing files in a distributed LDAP+Samba-based infrastructure

M. Marinello

FUSS Team, Autonomous Province of Bolzano, 39100 Italy  
e-mail: [me@marcomarinello.it](mailto:me@marcomarinello.it)

October 31, 2020

## ABSTRACT

**Context.** An in-production infrastructure of 64 schools running Debian-based networks with OpenLDAP and Kerberos. Samba is even provided for Windows compatibility. This O.S. is called "FUSS" \* and is developed by the Autonomous Province of Bolzano.

**Aims.** Allow school's users to remotely access and collaborate on their files.

**Methods.** Using Free and Open Source software only.

**Results.** A Docker+Nextcloud based solution that can be automatically deployed in the single school and an internal PKI to secure communication between the delegate server and the external proxy.

**Key words.** LDAP — Samba — Nextcloud — Open Source — Apache — Internal PKI

## 1. Introduction

FUSS, acronym for *Free Upgrade for a (digitally) sustainable school*, is a project originally launched by the Autonomous Province of Bolzano, Italian school Department, in 2005 and consists of a Debian-based operating system that aims to give South Tyrol's schools an open-sourced alternative to closed OSs.

The schools are connected within a Virtual Private Network that allows technicians to remotely support and maintain the schools. The VPN has a few public endpoints which is the thing that made possible to develop this solution.

A feature that teachers often asked was the possibility to access and edit their files outside the school LAN which was impossible before the beginning of this study.

## 2. Pre-existing infrastructure

South Tyrol's schools have a quite uncommon network infrastructure which makes the goal of this study harder to reach.

First of all, almost every school is wired via optical fiber to the Province's VPN and, through that, is able:

- to reach the internet;
- to reach other hosts and services in the private infrastructure.

Secondly, every network is equipped with a virtualization environment based on Proxmox VE<sup>1</sup> that emulates the FUSS Server, nowadays based on Debian 8. The virtualization environment, the FUSS Server and the gateway have an IP address of the VPN which is reachable from other hosts of the private infrastructure.

\* <https://fuss.bz.it>

<sup>1</sup> <https://www.proxmox.com/en/proxmox-ve>

### 2.1. Normal behaviour of a FUSS network

The FUSS Server is designed to act as server of the network of the single school and is supposed not to be connected with other servers or with the rest of the infrastructure.

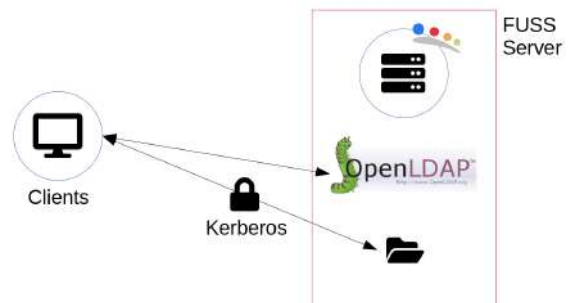


Fig. 1. Diagram of a FUSS network

It provides, for each school:

- a DHCP and DNS server;
- an OpenLDAP server that stores usernames and passwords of the users of the network;
- a Kerberos-secured NFSv4 share to allow clients to securely mount the homes;
- a Samba share to keep the compatibility with the (few) Windows clients.

It doesn't, indeed, replicate the LDAP archive on a central server and the firewall configuration denies queries (and even mounting NFS or Samba) from any host outside the LAN.

In fact, the only way to access both LDAP and Samba is to pretend to be a client and thus deceive the server.

### 3. State of the art

This study can be divided into three parts:

- the cloud part, i.e. the user interface that will allow users to browse, upload and download their files;
- the online collaboration part, i.e. the suite that will allow users to edit collaboratively the same document at the same time;
- the internal PKI part, i.e. the software that will (automatically) issue the SSL certificates signed by a private CA to protect the connection between the public endpoint and the delegate server.

#### 3.1. Cloud

There is a plethora — both open and closed source — of cloud software to achieve this goal. Google Drive<sup>®</sup> and Microsoft Office 365<sup>®</sup> are the proprietary solutions on the market.

For the on-premise solutions we have, in the open source side, Pydio, Seafile, OwnCloud and Nextcloud.

##### 3.1.1. Why Nextcloud?

Nextcloud is widely deployed by public administrations, enterprises, small companies and private users. It has a very large community, it is shipped with Docker and natively supports external storages and LDAP authentication which was, for this project, an essential feature. In version 18 Nextcloud launched *Nextcloud Hub*, a fully on-premise solution that provides the benefits of online collaboration without the compliance and security risks<sup>2</sup>. It integrates in one single portal file sharing, contacts, e-mails, calendar and meetings.

While writing this paper, Nextcloud is being carefully inspected by the German government to acquire it as open-source alternative and re-establish their digital sovereignty<sup>3 4 5</sup>.

#### 3.2. Online collaboration

The choice in the field of online document collaboration in terms of on-premise solutions is quite limited. LibreOffice Online is one of the possible solutions. The main issue of LOOL is that the prebuilt versions made available by TDF or one of their partner is delivered with the limit of 10 concurrent documents and 10 users.

The users which were targeted by this solutions are many more than 10. Therefore, the only way to use LOOL is to recompile it by ourself. The know-how gained in this procedure has been used to write a page on TDF's public wiki.<sup>6</sup>

#### 3.3. Internal PKI/ACME

A critical part of this study is the security of the tunnel from outside to the school.

Since the very beginning, Let's Encrypt has been designated to automatically secure the external proxy. Let's Encrypt is a

project of the Internet Security Research Group (ISRG) that have implemented both an ACME server (called *Boulder*) and the client (*Certbot*).

Unfortunately, Boulder, the server developed by ISRG, can't currently be deployed in a private context, since they use many custom configurations to run it in production<sup>7</sup>. The problem is essentially that Boulder itself relies on other components whose configuration has been released only in testing version, intended for development purpose. This configuration leaves some debug doors opened, exposing the whole solutions to attacks.

The only other open source project that supports ACME right now seems to be Smallstep Certificates<sup>8</sup> and is — therefore — the one we're going to use.

### 4. Goals

The main goal of this project was to allow schools' users to remotely access their files. While developing this concept another requirement was added: allowing users to collaborate on their documents.

At the same time, the Province purchased Microsoft Office 365<sup>®</sup>, which offers built-in collaborative editing. An advantage of our solution is that it keeps the data in the schools' servers on our territory, without sending them anywhere, which is very good for the GDPR-compliance.

The common behaviour of a cloud is the following: you copy a document to a directory on your computer which is synchronized by a daemon that sends the file you just copied to the cloud's server. The respective daemons on your other devices (laptop, tablet, smartphone, and so on) can also sync a copy of the file locally.

At the end of this process you will have:

- the original document, on the device you wrote it;
- a copy on the cloud's servers;
- a copy on every device you synchronize with the cloud.

While just thinking of ourselves, having many copies of the same document would not seem a problem. Scaling this to all users, instead, results in a waste of archivation space.

Accessing the servers' storage won't just allow users to access their files but allows even to take advantage of the (unused) storage present in every school. It achieves also a non-replication strategy: the cloud storage coincides with user's home.

Finally, it will not force users to remember another password since will be authenticated in SSO via LDAP.

### 5. New infrastructure

This project requires a huge (new) infrastructure to be realized in order to complete the challenge.

Once again, we'll divide the different requirements of infrastructure.

On the public side we need:

- a new DNS server to provide a sub-zone of the main domain;
- a VM with a public IP address that will host the LibreOffice Online instance. Since it would be useless to generate extra traffic from the VPN endpoint to the school's server and run there the instance of LOOL, the approach will be to

<sup>7</sup> <https://community.letsencrypt.org/t/boulder-deploy-in-production/100050>

<sup>8</sup> <https://smallstep.com/>

<sup>2</sup> <https://nextcloud.com/hub/>

<sup>3</sup> <https://www.zdnet.com/article/eu-turns-from-american-public-clouds-to-nextcloud-private-clouds/>

<sup>4</sup> <https://datenschutz.hessen.de/pressemitteilungen/stellungnahme-des-hessischen-beauftragten-f%C3%BCr-datenschutz-und>

<sup>5</sup> [https://nextcloud.com/fr\\_FR/blog/eu-governments-choose-independence-from-us-cloud-providers-with-nextcloud/](https://nextcloud.com/fr_FR/blog/eu-governments-choose-independence-from-us-cloud-providers-with-nextcloud/)

<sup>6</sup> <https://wiki.documentfoundation.org/Development/BuildingOnline>



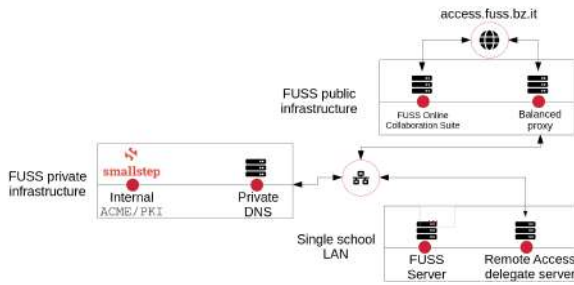


Fig. 2. Diagram of the new architecture

have only one strong instance of LOOL, running on a separate network, saving bandwidth for the school's traffic and providing the service for all the instances of Nextcloud;

- a few VMs with both a public IP address and a VPN's IP address. These will be the proxies from the outside to the schools.

On the private side, instead, we need:

- a VM, with a private IP only, that provides both the private DNS server and the Smallstep's ACME server. The CA will be secured by a LUKS-protected volume and a strong firewall completes the protection of the solution;
- a new VM on every school's server running Debian 10 with Docker. For simplicity, the technicians will just restore a Proxmox template with Debian 10 and the remaining setup<sup>9</sup> will be done by an Ansible<sup>10</sup> playbook.

### 5.1. Deploy strategy

Reconfiguring 64 servers in likewise different institutes, often geographically far, can quickly become a nightmare. Therefore, we needed an efficient and reliable deploy strategy.

We strongly rely on Ansible, the open source IaC software developed by RedHat, to quickly re-orchestrate both public and private infrastructure.

The whole procedure is divided into three playbooks.

#### 5.1.1. Installation of the template in Proxmox

The task of the first playbook is to copy a dump of a Debian 10 template with cloud-init installed and restore it into the Proxmox instance. After playing this playbook, a technician logs manually into the server, checks that there is the required space and RAM for running the new VM, clones it, adds the second NIC and the cloud-init drive, sets IP address and SSH access key and starts it.

At this point the template can be safely<sup>11</sup> deleted to save some space.

#### 5.1.2. DNS and Apache orchestration

While installing the new VM, another playbook runs to prepare the present infrastructure to receive a new delegate.

<sup>9</sup> Installation of Docker, clone of the required Git repository, creation of a .env file with the installation specifications and debootstrap of the docker-compose environment.

<sup>10</sup> <https://www.ansible.com/>

<sup>11</sup> The container has to be cloned with the *full clone* and not *linked*.

This script reads the configuration (school, internal IP, intended external proxy) that has to be applied from a file and then connects to the different hosts to apply it.

It starts from the internal DNS: regenerates the entire zone of bind9 and restarts it.<sup>12</sup>

Next it reconfigures the external DNS: even here the whole zone will be rewritten, setting the external subdomain as a CNAME of the proxy, and the service restarted.

Finally the playbook will connect to the proxy, add a VirtualHost of Apache2 for the new instance, restart the service, ask a certificate to Let's encrypt and restart Apache again.

### 5.1.3. Configuration of the delegate server

The last playbook is supposed to connect to the newly created VM, update and configure that by cloning the repository containing the Dockerfile, generating some random password and filling the local configuration file with the informations stored on the central configuration file.

After this playbook, the instance is ready to be accessed by users.

## 6. Conclusions

### 6.1. Bandwidth utilization

Although schools should have more than enough bandwidth to provide both this solution and the normal internet connection, we wanted to be sure right away. That's why we monitor all servers with Zabbix<sup>13</sup>.

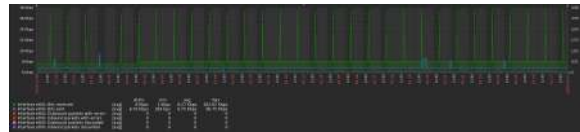


Fig. 3. Graph of bandwidth usage of a single delegate server

Partly because of the very strong safety criteria of Nextcloud, we found out that bandwidth utilization were quite good. Compared to when the traffic originated only from the LAN, the delegate server had a negligible rise in load.

It is good to remember that one of the strength of this solution is that we avoid busying bandwidth while the school is operative (typically in the morning) by running cloud synchronization agents or manually downloading files. The teacher will immediately find the files she/he uploaded in his home directory.

## 7. Future developments

### 7.1. Bugs currently affecting Nextcloud

Up to the version 18, Nextcloud is affected by a bug: a scan of the Samba-shared folder is done just on the first log-on.

As an user creates or deletes a file or directory from the desktop, the database of the cloud application will no longer reflect the filesystem status, running into errors (in the case the user tries to access to deleted files) or into the impossibility to access the newly-created files.

<sup>12</sup> A DNS record will be required later to ask the certificate via certbot.

<sup>13</sup> <https://www.zabbix.com/>

At the time of writing, the problem is still affecting Nextcloud. As a workaround, some records from 2 different tables of the database are periodically dropped to force the system to rescan the filesystem.

Fixing this bug and submitting a pull request for this problem is certainly on the roadmap.

### 7.2. Integration of the present solution with other infrastructures

The result of this study is quite far from an "universal" plug & play solution that can be applied to almost every network. A few steps of the container initialization script strictly depend on other components of the network infrastructure<sup>14</sup>.

Outside of the province, the standard for the schools is to own their own connection with one or more public IP addresses. This solution has to be available even for such kind of infrastructure, certifying directly the public domain name with Let's encrypt.

In order to make this solution suitable for a more common infrastructure, a flag for disabling those critical steps will be made available.

### 7.3. Publication of the container

Since now the container is built on every single delegate server from scratch, getting that built once and published on the docker hub would speed up deploy and upgrade of the infrastructure.

After having the new variables into our setup scripts, this would be helpful even for other user who want to deploy this solution.

## References

1. Davoli, P., Bressan, S. and Lorenzi, P. 2007. Full migration to Open Source systems in education: investigation on "Bolzano case". Piet Kommers. 8. ISBN 978-972-8924-35-5.
2. Butcher, M. 2007. Mastering OpenLDAP: Configuring, Securing, and Integrating Directory Services. Packt Publishing Ltd., Birmingham - Mumbai. 482. ISBN 978-1-847191-02-1.
3. Goldman, R. 2016. Learning Proxmox VE: Unleash the power of Proxmox VE by setting up a dedicated virtual environment to serve both containers and virtual machines. Packt Publishing Ltd., Birmingham - Mumbai. 217. ISBN 978-1-78398-178-6.
4. Merkel, D. 2014. Docker: lightweight linux containers for consistent development and deployment. Linux Journal. 5.
5. Barnes, R., Hoffman-Andrews, J., McCarney, D. and Kasten J. 2019. Automatic Certificate Management Environment (ACME). RFC Editor. 95. ISSN 2070-1721.
6. Paternò, G. 2004. Single Sign-On con Kerberos e LDAP: Una soluzione per ambienti eterogenei. 103. ISBN 88-901141-1-8.
7. Hochstein, L. and Moser, R. 2017. Ansible: Up and Running: Automating Configuration Management and Deployment the easy way. 401. O'Reilly & Associates Inc. ISBN 978-1-491-97980-8.

<sup>14</sup> e.g. the ACME or the private DNS.

## Special thanks

I am grateful to all the people who helped me realizing this study. In particular, I would like to thank

- **Paolo Dongilli**  
Technical inspector - head of the FUSS Project  
Autonomous Province of Bolzano  
*and*  
**Stefania Fiore, Andrea Bonani, Piergiorgio Cemin and Claudio Cavalli**  
for their support and their help in the development of this project
- **Marina Latini**  
Former chairperson of the board – The Document Foundation  
*and*  
**Emiliano Vavassori**  
Director – The Document Foundation  
for their help in recompiling LibreOffice Online
- all the teachers and headmasters who decided to endorse and try this solution and contributed improving it by reporting bugs and requesting new features.