PROCEEDINGS OF SCIENCE

Enhancing StoRM WebDAV data transfer performance with a new deployment architecture behind NGINX reverse proxy

Federica Agostini,[a] **Luca Bassi,**[a,b] **Jacopo Gasparetto,**[a,*] **Francesco Giacomini,**[a] **Roberta Miccoli**[a] **and Enrico Vianello**[a]

[a] *INFN-CNAF, Viale Berti Pichat 6/2, Bologna, Italy*

[b] *GARR, Via dei Tizii 6, Roma, Italy*

E-mail: federica.agostini@cnaf.infn.it, luca.bassi@garr.it, jacopo.gasparetto@cnaf.infn.it, francesco.giacomini@cnaf.infn.it, roberta.miccoli@cnaf.infn.it, enrico.vianello@cnaf.infn.it

StoRM WebDAV is a component of StoRM which is designed to provide a scalable and efficient solution for managing data storage and access in Grid computing environments. It specifically focuses on enabling access to stored data through the WebDAV (Web Distributed Authoring and Versioning) protocol, an extension of the HTTP protocol that allows users to create, change and move resources on a web server.

This contribution highlights how data transfer performance can be enhanced by delegating the execution of some HTTP methods to an external, dedicated service: NGINX. This strategic decision is driven by the proven reliability, scalability, and performance capabilities of NGINX in handling such critical operations.

International Symposium on Grids and Clouds (ISGC2024)
24 -29 March, 2024
Academia Sinica Computing Centre (ASGC), Institute of Physics, Academia Sinica Taipei, Taiwan

[*] Speaker

1. Introduction

The StoRM WebDAV service [1] is a component of StoRM (STOrage Resource Manager) [2] and relies on the WebDAV protocol [3] to access resources shared on a file system.

StoRM WebDAV is designed to follow the requirements set forth by the WLCG (Worldwide LHC Computing Grid) [4] community, in particular it supports: Third Party Copies (TPC) [5], authorization based on access tokens or X.509 certificates/proxies, and fine-grained access policies. TPC operations have been one of the main GridFTP features used by LHC experiments data management frameworks to implement scalable data transfers and management. In 2017 the Globus Alliance announced that the open-source Globus Toolkit [6], including, among other services, GridFTP, would no longer be supported. This seriously impacted the WLCG community because of the central role of the Globus Security Infrastructure and GridFTP in the context of data transfer frameworks. As a natural consequence, WLCG is moving towards HTTP-based data transfers, meaning that there is the necessity to make StoRM WebDAV as reliable, performant and efficient as possible.

This article describes the work done in the past months to delegate some data transfer operation to NGINX [7] with the aim of improving performance, and underlies the forthcoming work necessary to fully reach the goal.

2. StoRM Overview

StoRM is a lightweight, scalable, flexible, high-performance, filesystem-independent, storage manager service for generic disk-based storage system, compliant with the SRM specification [8]. It is the SRM solution developed at INFN-CNAF; it serves the Italian Tier-1 data center at CNAF, as well as more than 30 other sites. StoRM provides a *thin* management layer over a POSIX file system, typically a distributed one such as IBM GPFS [9].

2.1 StoRM components

StoRM is a suite of components, which provide a solution for storage management and data transfers on top of flexible authentication and authorization mechanisms, based on X.509 certificates, VOMS (Virtual Organization Membership Service) [10] proxies and JSON Web Tokens (JWT) [11]. The file access control can be further enforced via POSIX Access Control Lists.

StoRM supports a tape system through integration with GEMSS [12], a full Hierarchical Storage Management (HSM) system, integrating the IBM General Parallel File System (GPFS), the IBM Tivoli Storage Manager (TSM) and the StoRM Backend.

The StoRM main components are:

- StoRM Backend and StoRM Frontend: the StoRM frontend service implements the SRM interface exposed to client applications and frameworks, while the StoRM backend service implements the actual storage management logic by interacting directly with the underlying file system;
- StoRM GridFTP: it provides file transfer based on gsiFTP (an extension of the FTP protocol), used to access directly the file system underlying the StoRM deployment;

- StoRM WebDAV: it provides file transfer and management functionality, based on the WebDAV protocol and is the main topic of this paper;
- StoRM Tape: a new component, implementing the WLCG Tape REST API [13] for StoRM and providing HTTP-based file management for a tape storage system.

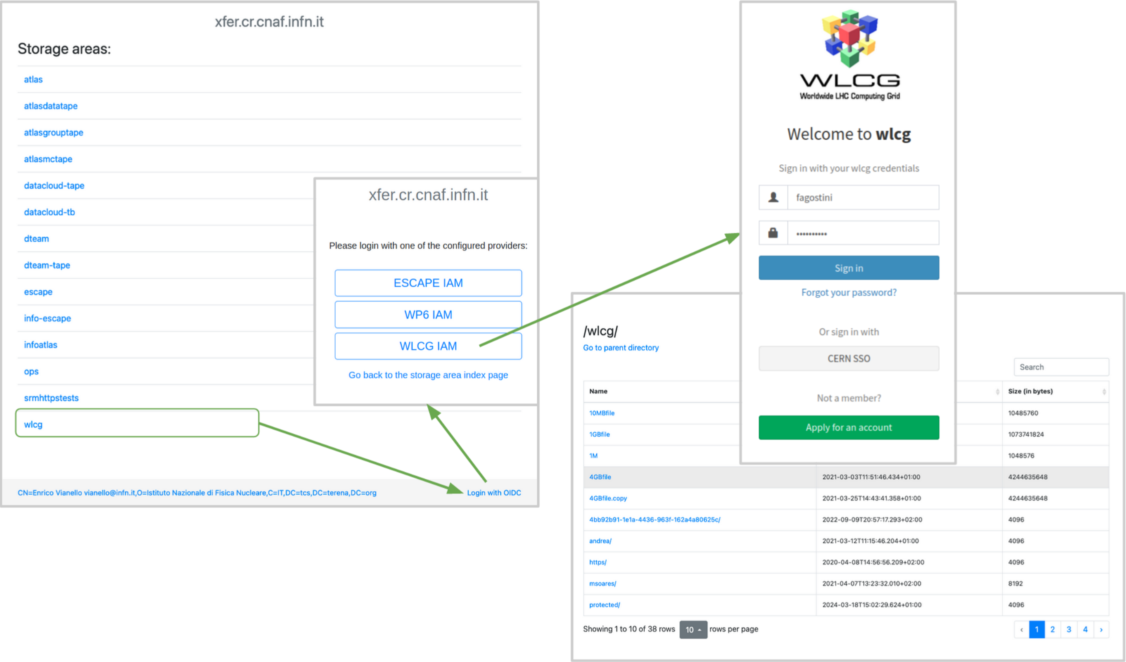
The latest StoRM release is the v1.11.22, supporting the CentOS 7 platform, but packages for RHEL 9 will be available soon. StoRM will be maintained and evolved for the foreseeable future, including support (through GGUS tickets or mailing-list) to StoRM-based sites.

2.1.1 StoRM WebDAV

StoRM WebDAV, besides HTTP data transfer functionality, supports also TPC operations. TPC is an extension of the WebDAV COPY verb, defined by the WLCG community and consists of bulk transfer requests between two remote storage endpoints. The authorization to StoRM WebDAV resources can be enabled for both JWT tokens and X.509 certificates/VOMS proxies. StoRM WebDAV also provides a browser-based data management interface, where access is granted according with some requirements present in the JWT or X.509 credentials. Figure 1 shows an example where access to a certain storage area is granted to users registered within a specific OpenID Connect (OIDC) identity provider.

2.2 Deployment architecture

An example of the current StoRM deployment architecture, where all StoRM components are in place, is shown in Figure 2. Blue arrows indicate data transfers, red arrows highlight data



xfer.cr.cnaf.infn.it

Storage areas:

atlas
atlasdatatape
atlasgrouptape
atlasmctape
datacloud-tape
datacloud-tb
dteam
dteam-tape
escape
info-escape
infoatlas
ops
srmhttpstests
wlcg

CN=Enrico Vianello vianello@infn.it,O=Istituto Nazionale di Fisica Nucleare,C=IT,DC=tcs,DC=terena,DC=org

Login with OIDC

xfer.cr.cnaf.infn.it

Please login with one of the configured providers:

ESCAPE IAM
WP6 IAM
WLCG IAM

Go back to the storage area index page

Welcome to **wlcg**

Sign in with your wlcg credentials

fagostini

••••••••••

Sign in

Forgot your password?

Or sign in with

CERN SSO

Not a member?

Apply for an account

/wlcg/

Go to parent directory

Search

Name		Size (in bytes)
10MBfile		10485760
1GBfile		1073741824
1M		1048576
4GBfile	2021-03-03T11:51:46.434+01:00	4244635648
4GBfile.copy	2021-03-25T14:43:41.358+01:00	4244635648
4bb92b91-1e1a-4436-963f-162a4a80625c/	2022-09-09T20:57:17.293+02:00	4096
andrea/	2021-03-12T11:15:46.204+01:00	4096
https/	2020-04-08T14:56:56.209+02:00	4096
msoares/	2021-04-07T13:23:32.010+02:00	8192
protected/	2024-03-18T15:02:29.624+01:00	4096

Showing 1 to 10 of 38 rows 10 rows per page

‹ 1 2 3 4 ›

Figure 1: Access to `wlcg` storage area is restricted to WLCG users who authenticate through OIDC. After authentication, WLCG users can browse through the storage area content.

management operations and black arrows are used for internal communication between components. The figure also shows how some components support only authentication and authorization with X.509/VOMS proxies, while the HTTP-based components support also JWT tokens.

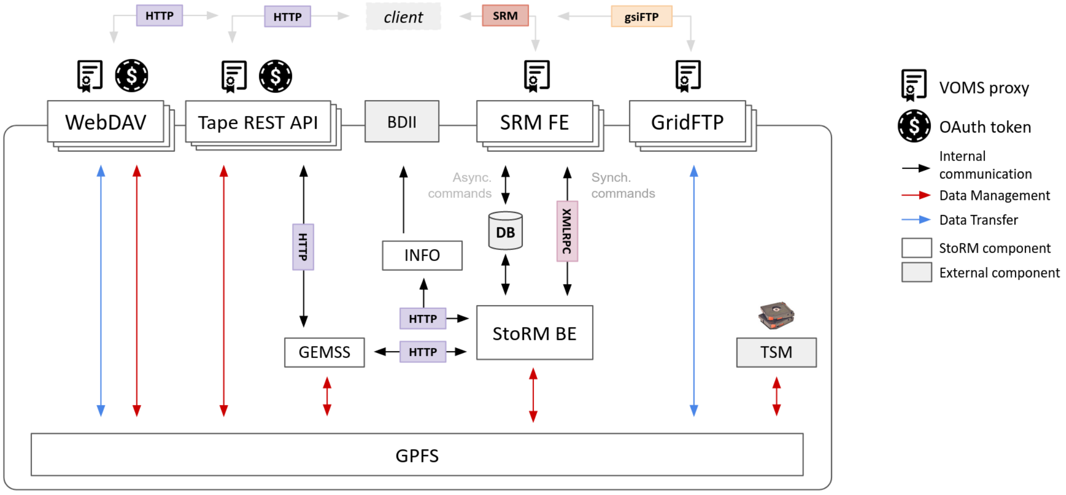


Figure 2: The current StoRM high-level architecture.

The decommissioning of all Globus GridFTP transfer endpoints, replaced by StoRM WebDAV and StoRM Tape instances, is a necessary step towards a deployment scenario free of the SRM legacy (Figure 3). For StoRM, a no-SRM deployment means the removal of the StoRM Frontend and of a substantial part of the current StoRM Backend, updating the component dedicated to collect information about the storage areas (e.g. the used space).

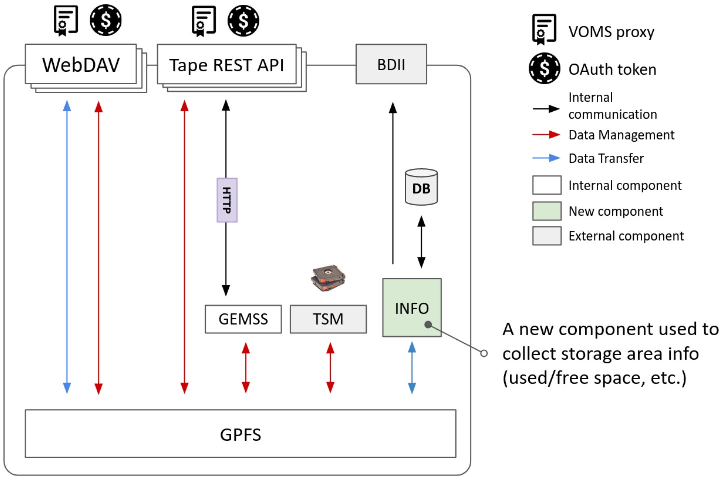


Figure 3: A no-SRM future StoRM deployment, including a tape-enabled storage area.

3. Enhanced StoRM WebDAV deployment

As already mentioned, the decommissioning of all Globus components, including the GridFTP transfer server, raises the need to focus the developments of StoRM WebDAV with the aim of improving the service performance. In this work, we show how the StoRM WebDAV codebase is being simplified by externalizing common functions better implemented by established third-party components. In particular, the transfer performance has already been enhanced thanks to NGINX, designed to efficiently handling HTTP requests. Moreover, the adoption of the new deployment model enhances the scalability of Grid middleware services, ensuring they can efficiently accommodate growing demands and effectively scale in response to increased usage.

3.1 NGINX for authentication and HTTP requests

NGINX is an open-source HTTP server and reverse proxy, known for its high performance and stability, rich set of features and low resource consumption [14]. In addition to the common compelling reasons to use open-source software, e.g. a large community of developers and users who collaborate to inspect the software for security vulnerabilities and other issues, NGINX is highly adopted among the research communities. In the context of StoRM WebDAV deployment, it is used for TLS/VOMS termination and serving HTTP requests.

A dedicated VOMS module for NGINX [15] is loaded into the configuration, in order to enable client-side authentication based on X.509 proxy certificates augmented with VOMS Attribute Certificates (AC), typically obtained from a VOMS server. In particular, the module validates the VOMS proxy, meaning to confirm that the certificate has been issued by a trusted authority, to check its expiration date, to verify the digital signature and to ensure it has not been revoked. It also defines a set of NGINX embedded variables, whose values are extracted from the first AC found in the certificate chain (e.g. the VO name or the FQANs). Thus, after authentication, the VOMS attributes extracted by the VOMS proxy are forwarded to StoRM WebDAV, to be processed for authorization decisions.

Moreover, the goal is for StoRM WebDAV to delegate as much work as possible to NGINX in terms of data transfers, while StoRM keeps the role of authorization decision point. At the moment, we have addressed the GET requests. The implementation is based on an internal redirect from StoRM WebDAV to NGINX, using the `X-Accel` [16] HTTP response header. The way it works is that StoRM WebDAV sends the header `X-Accel-Redirect` with a URI. NGINX will match this URI against its locations as if it were a normal request. It will then serve the location that matches the defined root known by StoRM WebDAV, plus the URI passed in the header.

3.2 Deployment model and performance

Figure 4 shows the interactions between the above mentioned components when requesting a `/sa/file.dat` resource as example.

The flow of a GET HTTP request can be summarized as follows:

1. the user submits a request to obtain a resource. The request is VOMS/TLS-terminated by NGINX;

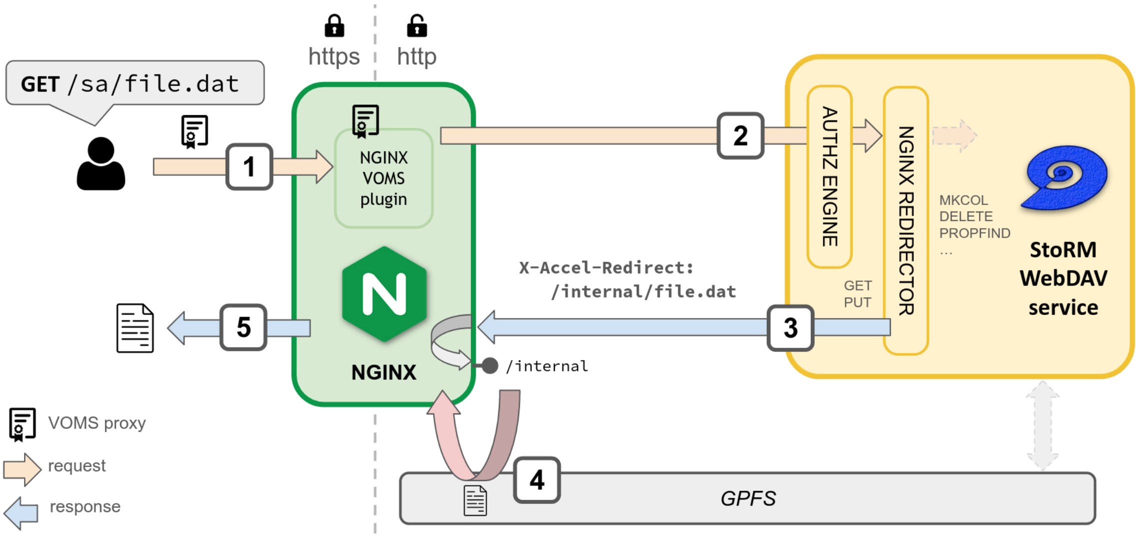


Figure 4: The StoRM WebDAV internal deployment schema. The steps of a GET request to the `/sa/file.dat` file shown as example are highlighted from number 1 to 5. The interaction 4 consists of an internal redirect.

2. the NGINX VOMS module parses and validates the provided proxy and forwards the request to the StoRM WebDAV service for authorization;

3. the request processed by StoRM WebDAV is internally redirected to NGINX, including the `X-Accel` HTTP response header (i.e. `X-Accel-Redirect: /internal/file.dat`);

4. the internal endpoint of NGINX is configured to retrieve the resource from the file system (GPFS, in this deployment);

5. the resource is returned to the user.

In case of all the other HTTP requests, the flow is valid up to step number 2., where then the full request is directly served by StoRM WebDAV. In any case, NGINX will still play the role of authentication engine for X.509 proxy certificates.

Beside the unit tests of the StoRM WebDAV code base, preliminary stress tests have been performed with the aim of checking the improvements in performance. The client used for testing is called Vegeta [17], and is a versatile HTTP load testing tool built out of a need to drill HTTP services with a constant request rate. The preliminary results refer to services installed in a Virtual Machine with an AlmaLinux 9 distribution, 4 CPUs and 8 GB of memory. Figure 5 shows the result of stress tests, which highlights the ~10% of improvement in terms of transfer performance when using NGINX.

Also, the improvements in performance is visible in terms of memory, where NGINX makes almost half usage with respect to StoRM WebDAV. A significant contrast is evident in CPU usage, with StoRM consuming approximately 10%, whereas NGINX only utilizes about 0.1%. This variance arises from the inherent characteristics of the StoRM WebDAV application, which is Java-based and defaults to a higher CPU allocation.

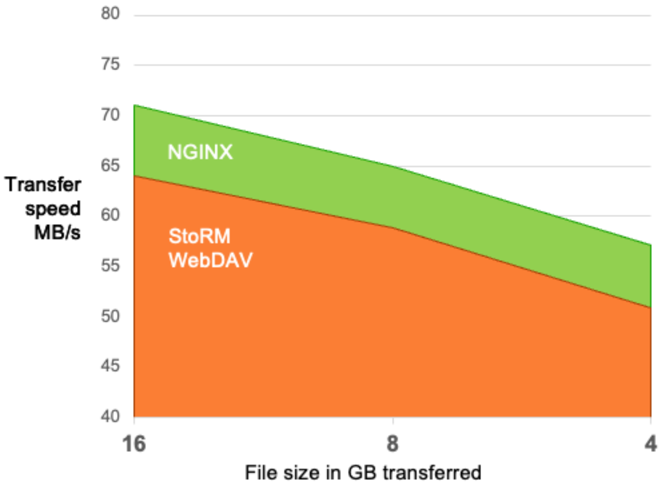


Figure 5: The StoRM WebDAV preliminary performance tests result.

4. Future looks and conclusions

In this article we have described the work done to delegate some data transfer operations to NGINX, currently handled by the StoRM WebDAV service, with the aim of improving performance and reliability. Performance stress tests focused on transfer throughput have already given promising results. We have also explored a new deployment model, which can be adoptable by other products we develop (e.g. StoRM Tape, INDIGO IAM, etc.).

The work ongoing at the time of writing involves delegating also PUT HTTP requests to NGINX. This adds complexity, particularly due to the needs for checksum calculation. In the near future we will focus on enabling JWT authentication also in NGINX, which is already in place and used in production for the StoRM Tape service [18]. Also, since the goal for StoRM WebDAV is to delegate as much work as possible to external, more reliable and specific components, we want to explore the possibility of using a dedicated service for authorization and policy enforcement, such as Open Policy Agent [19]. Moreover, we will apply specific optimization in the NGINX configuration, according to established best practices.

Acknowledgements

The work presented in this paper has been supported by the NextGenerationEU European initiative through the Italian Ministry of University and Research, PNRR Mission 4, Component 2 - ICSC [20]: Investment 1.4, Project code CN00000013 - CUP I53C21000340006; TeRABIT [21]: Investment 3.1, Project code IR0000022 - CUP I53C21000370006.

References

[1] *The StoRM WebDAV service* source code, https://github.com/italiangrid/storm-webdav

[2] *STORage Resource Manager (StoRM)*, https://italiangrid.github.io/storm

[3] *HTTP Extensions for Distributed Authoring – WEBDAV*, https://datatracker.ietf.org/doc/html/rfc2518

[4] *Worldwide LHC Computing Grid*, https://wlcg.web.cern.ch/, last seen April 2024

[5] *HTTP/WebDAV Third-Party-Copy Technical Details*, https://twiki.cern.ch/twiki/bin/view/LCG/HttpTpcTechnical

[6] *Globus, a UChicago no-profit service*, https://www.globus.org/, last seen April 2024

[7] *NGINX*, https://nginx.org/en/docs, last seen April 2024

[8] *The Storage Resource Manager Interface Specification Version 2.2*, https://sdm.lbl.gov/srm-wg/doc/SRM.v2.2.html

[9] *IBM General Parallel File System*, https://www.ibm.com/docs/en/gpfs, last seen April 2024

[10] *Virtual Organisation Membership Service (VOMS)*, https://italiangrid.github.io/voms

[11] *The JSON Web Token RFC*, https://tools.ietf.org/rfc/rfc7519.txt

[12] Pier Paolo Ricci *et al* (2012) *J. Phys.: Conf. Ser.* **396** 042051

[13] *WLCG Tape REST API (v1) reference document*, https://docs.google.com/document/d/1Zx_H5dRkQRfju3xIYZ2WgjKoOvmLtsafP2pKGpHqcfY, last seen April 2024

[14] D. Kunda, S. Chihana, S. Muwanei, (2017) *Web Server Performance of Apache and Nginx: A Systematic Literature Review* **8** 43-52

[15] *VOMS module for NGINX*, https://baltig.infn.it/cnafsd/ngx_http_voms_module

[16] *X-Accel*, https://www.nginx.com/resources/wiki/start/topics/examples/x-accel/, last seen April 2024

[17] *Vegeta*, https://github.com/tsenart/vegeta, last seen April 2024

[18] F. Giacomini, T. Diotalevi, E. Vianello, L. Cappelli, F. Agostini, R. Miccoli, M. Vilaça Pinheiro Soares, A. Galavotti, *A RESTful approach to tape management in StoRM*, https://indico.jlab.org/event/459/contributions/11360/ (2023)

[19] *Open Policy Agent*, https://www.openpolicyagent.org/docs/latest/, last seen April 2024

[20] *ICSC*, https://www.supercomputing-icsc.it/en/icsc-home/, last seen April 2024

[21] *TeRABIT*, https://www.terabit-project.it/it/, last seen April 2024