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

Jacopo Gasparetto
INFN - CNAF
on behalf of StoRM developers team

ICSC
Centro Nazionale di Ricerca in HPC,
Big Data and Quantum Computing

International Symposium on Grids & Clouds (ISGC)
24-29 March 2024

INFN
CNAF

Outline

- StoRM overview
  - StoRM components
  - Deployment architectures
- Enhanced StoRM WebDAV deployment
  - NGINX for authN and HTTP requests
  - Deployment model and performance
- Future outlooks and conclusions

StoRM overview

- StoRM components
- Deployment architectures

StoRM: STOrage Resource Manager

StoRM is a storage resource manager for disk-based storage systems that provides a "thin" management layer (SRM, WebDAV) **over a POSIX filesystem**

- typically a distributed file-system such as IBM GPFS or Lustre

StoRM is a suite of components

StoRM provides flexible authentication and authorization mechanisms:

- **VOMS proxies** & **OAuth/OIDC JWT tokens**
- file access control is enforced via **POSIX ACLs**

StoRM supports a tape system through integration with **GEMSS**, a full Hierarchical Storage Management (HSM) system that integrates:

- IBM General Parallel File System (GPFS)
- IBM Tivoli Storage Manager (TSM)
- StoRM Backend

StoRM components

StoRM main components are:

- StoRM Backend + StoRM Frontend
  - SRM disk and tape file management
- StoRM GridFTP
  - GridFTP file transfer
- **StoRM WebDAV**
  - **HTTP/HTTPS file management and transfer**
- StoRM Tape REST API
  - HTTP/HTTPS tape file management

Latest release:
StoRM v1.11.22
(2023-06-21)
release-notes

available on
StoRM stable repo
or UMD repositories

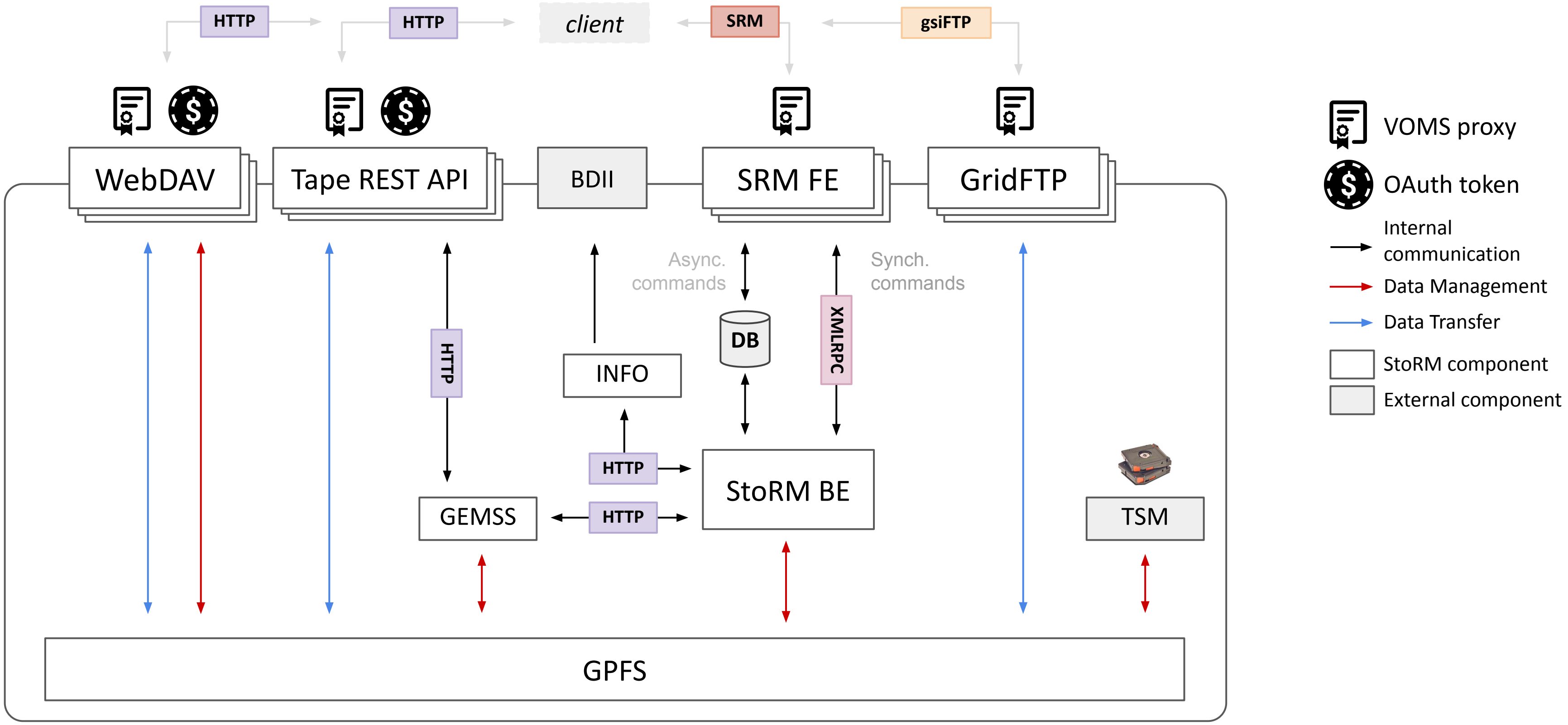
Name	Latest version
StoRM Backend	1.11.22
StoRM Frontend	1.8.15
StoRM WebDAV	1.4.2
StoRM Native Libs	1.0.7
StoRM Info Provider	1.8.3
StoRM SRM Client	1.6.1
StoRM GridFTP	1.2.4
StoRM XMLRPC-C	1.39.12
StoRM Utils	1.0.0
StoRM Puppet module	4.1.0
StoRM Tape REST API	0.7.0

Supported platforms: **CentOS 7**, *RHEL 9* (next)

StoRM is the Grid/Cloud storage solution adopted by the INFN-CNAF data center and some other TIER 2. It will be maintained and evolved by INFN for the foreseeable future

- including support (through GGUS tickets or mailing-list) to StoRM-based sites

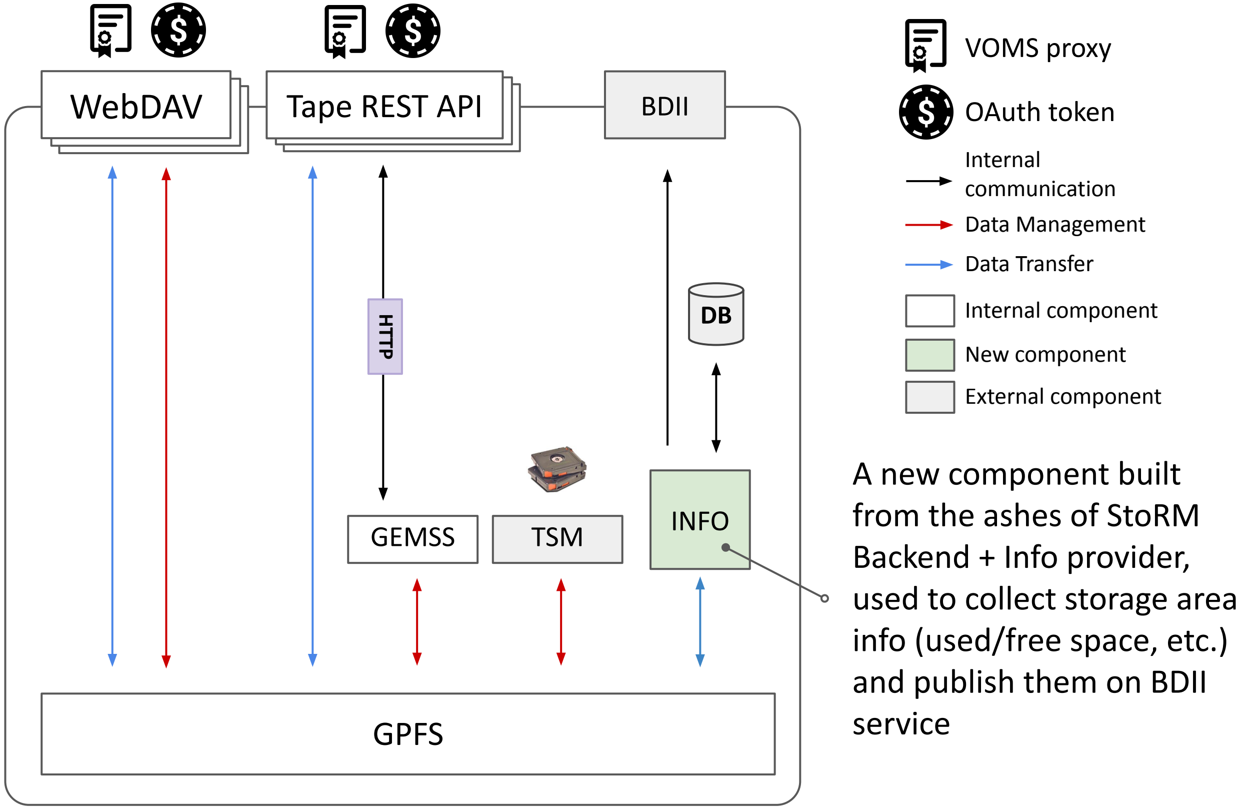
StoRM current typical deployment



Future architecture: no-SRM deployment

All Globus **GridFTP** will be turned off soon.

StoRM WebDAV + Tape REST API components provide the necessary data transfer/management functionalities on a tape-enabled storage system, without the need of legacy SRM components.



A new component built from the ashes of StoRM Backend + Info provider, used to collect storage area info (used/free space, etc.) and publish them on BDII service

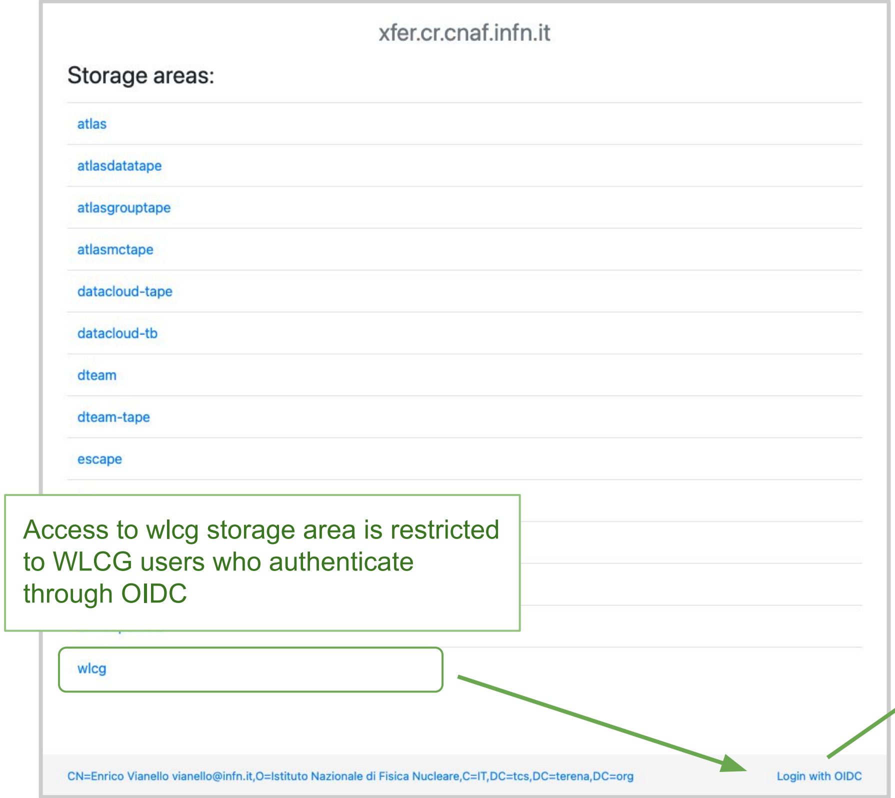
StoRM WebDAV

StoRM WebDAV is a StoRM component which provides a data transfer functionality through the WebDAV (Web Distributed Authoring and Versioning) protocol

- WebDAV is an extension of the HTTP protocol that allows users to create, change and move resources on a web server
- StoRM WebDAV provides a browser-based data management interface
- supports authorization based on JWT tokens, X.509 certificates and VOMS proxies
- supports Third Party Copies (TPC)
  - relies on an extension of the WebDAV COPY verb, which consists in bulk transfer requests between two remote storage endpoints

StoRM WebDAV web interface

https://xfer.cr.cnaf.infn.it:8443/



xfer.cr.cnaf.infn.it

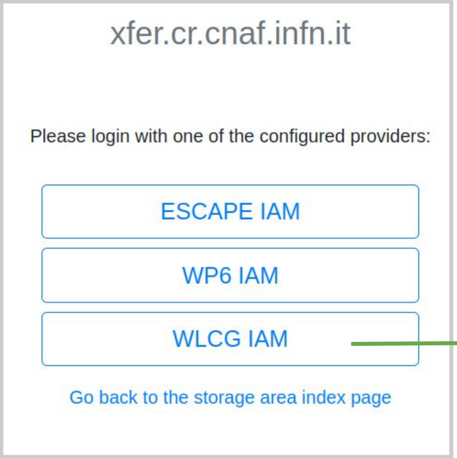
Storage areas:

- atlas
- atlasdatatape
- atlasgrouptape
- atlasmctape
- datacloud-tape
- datacloud-tb
- dteam
- dteam-tape
- escape
- wlcg

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

Login with OIDC

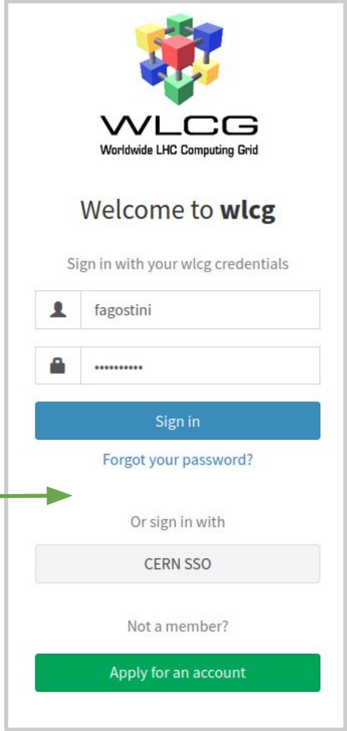
Access to wlcg storage area is restricted to WLCG users who authenticate through OIDC



xfer.cr.cnaf.infn.it

Please login with one of the configured providers:

Go back to the storage area index page



Welcome to **wlcg**

Sign in with your wlcg credentials

fagostini

••••••••••

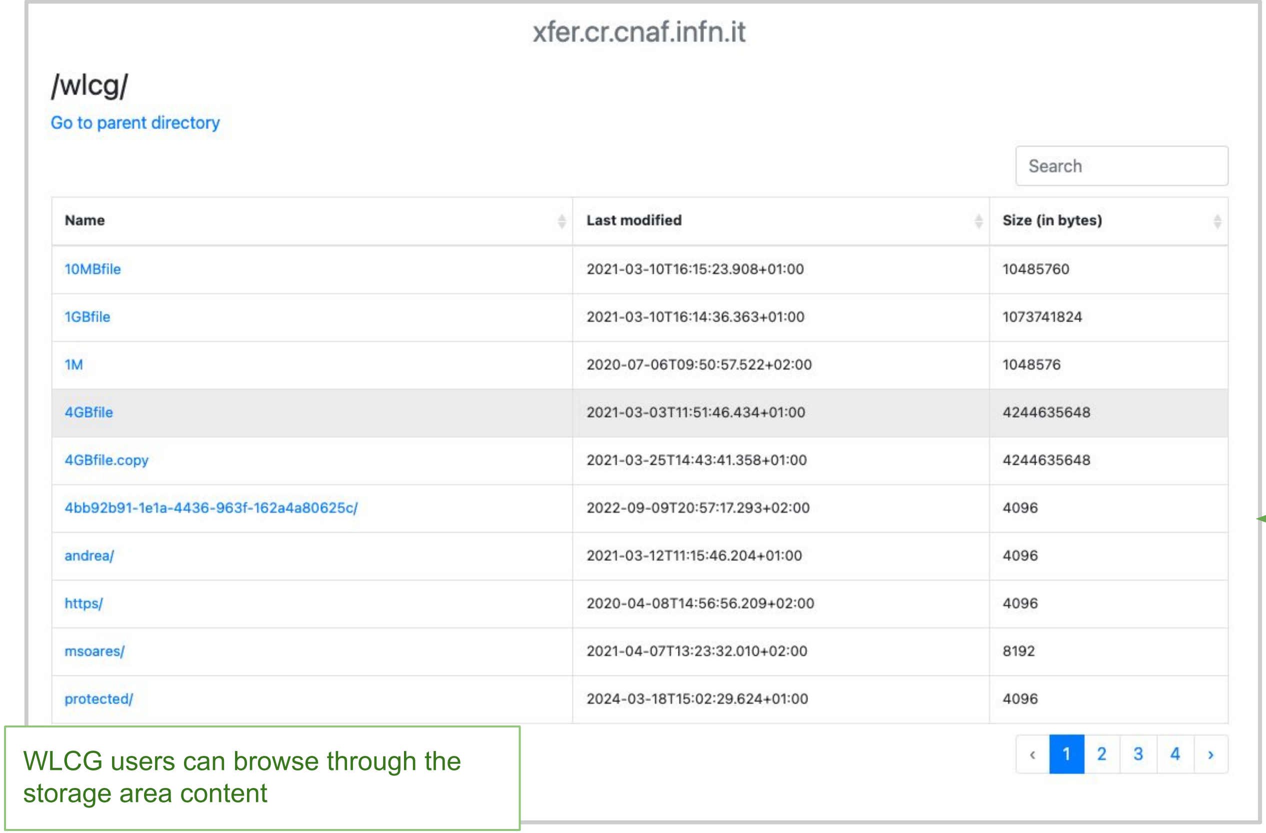
Forgot your password?

Or sign in with

Not a member?

StoRM WebDAV web interface

https://xfer.cr.cnaf.infn.it:8443/wlcg



xfer.cr.cnaf.infn.it

/wlcg/

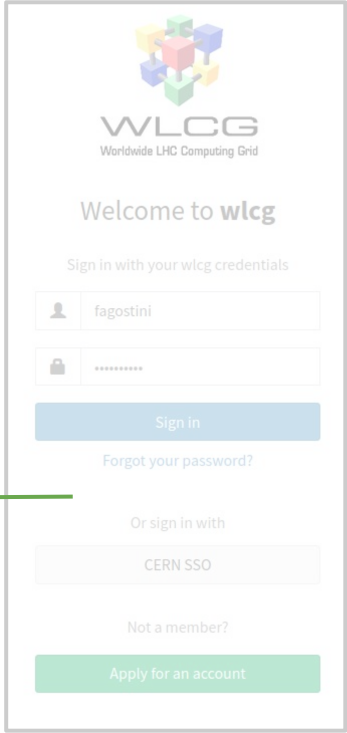
Go to parent directory

Search

Name	Last modified	Size (in bytes)
10MBfile	2021-03-10T16:15:23.908+01:00	10485760
1GBfile	2021-03-10T16:14:36.363+01:00	1073741824
1M	2020-07-06T09:50:57.522+02:00	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

‹ 1 2 3 4 ›

WLCG users can browse through the storage area content



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

Enhanced StoRM WebDAV architecture

- NGINX for authN and HTTP requests
- Deployment model and performance

Motivations

- Simplify StoRM WebDAV codebase by externalizing common functions better implemented by established third-party components
  - including AuthN and TLS termination (and possibly AuthZ)
- Re-use this deployment model for other products we develop
  - StoRM Tape, INDIGO IAM
- Improve transfer performance by relying on an external component that has been designed to efficiently handling HTTP GET and PUT requests
- Improve service scalability

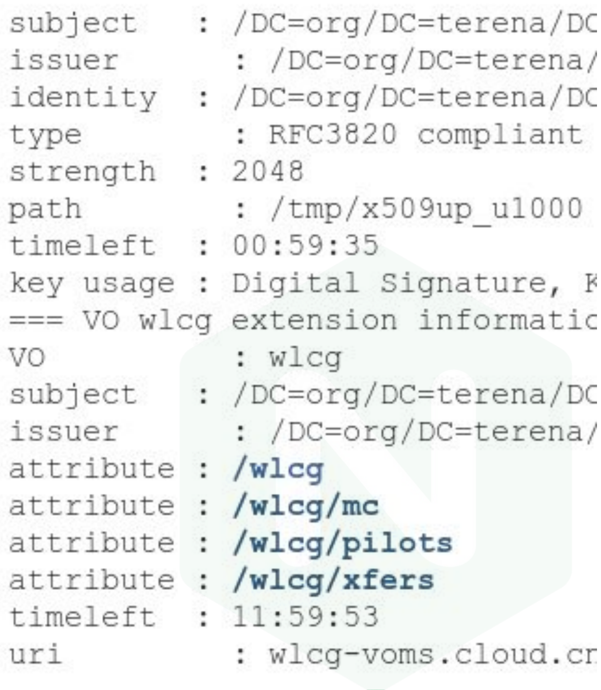
NGINX role in StoRM WebDAV architecture

- NGINX is an open-source HTTP server and reverse proxy, known for
  - High performance
  - High stability
  - Rich feature set
  - Simple configuration
  - Low resource consumption
- NGINX has been chosen as part of this architecture for
  - HTTP request handling
  - TLS termination
  - Flexible authentication via custom modules

NGINX dedicated module for authentication

- ngx_http_voms_module is an NGINX module we developed, that:
  - enables client-side authentication based on X.509/VOMS proxy certificates
  - validates the VOMS proxy
  - defines a set of NGINX **embedded variables** whose values are extracted from the Attribute Certificate
    - e.g. **voms_fqans** contains the FQANs included in the VOMS AC
- VOMS attributes are sent to WebDAV for AuthZ

VOMS proxy content



```
subject   : /DC=org/DC=terena/DC
issuer      : /DC=org/DC=terena/
identity  : /DC=org/DC=terena/DC
type        : RFC3820 compliant
strength  : 2048
path        : /tmp/x509up_u1000
timeleft  : 00:59:35
key usage : Digital Signature, K
=== VO wlcg extension informatic
VO          : wlcg
subject   : /DC=org/DC=terena/DC
issuer      : /DC=org/DC=terena/
attribute : /wlcg
attribute : /wlcg/mc
attribute : /wlcg/pilots
attribute : /wlcg/xfers
timeleft  : 11:59:53
uri         : wlcg-voms.cloud.cr
```

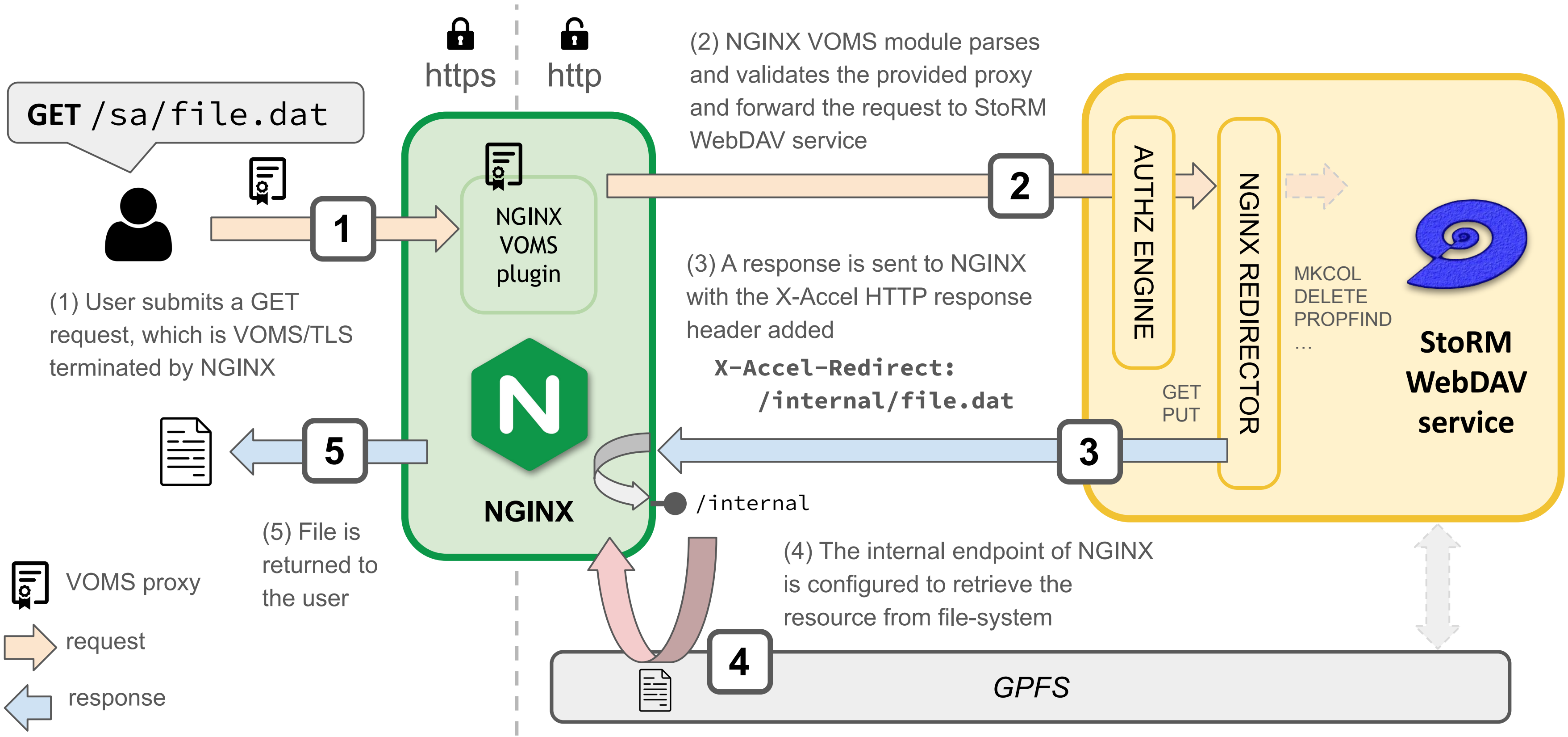
Delegating transfer requests to NGINX

- The goal is for StoRM WebDAV to delegate as much work as possible to NGINX in terms of data transfers
    - StoRM keeps the role of AuthZ decision point
- At the moment, we have addressed the GET request
- The implementation is based on an internal redirect from StoRM WebDAV to NGINX, using the X-Accel HTTP response header
    - `X-Accel-Redirect: /internal/{file-path}`

```
server {
  location /internal {
      internal;
      alias /storage/sa-root;
  }

  location / {
      proxy_pass http://storm-webdav:8086;
      ...
  }
}
```

Delegating transfer requests to NGINX



(1) User submits a GET request, which is VOMS/TLS terminated by NGINX

(2) NGINX VOMS module parses and validates the provided proxy and forward the request to StoRM WebDAV service

(3) A response is sent to NGINX with the X-Accel HTTP response header added

```
X-Accel-Redirect:
   /internal/file.dat
```

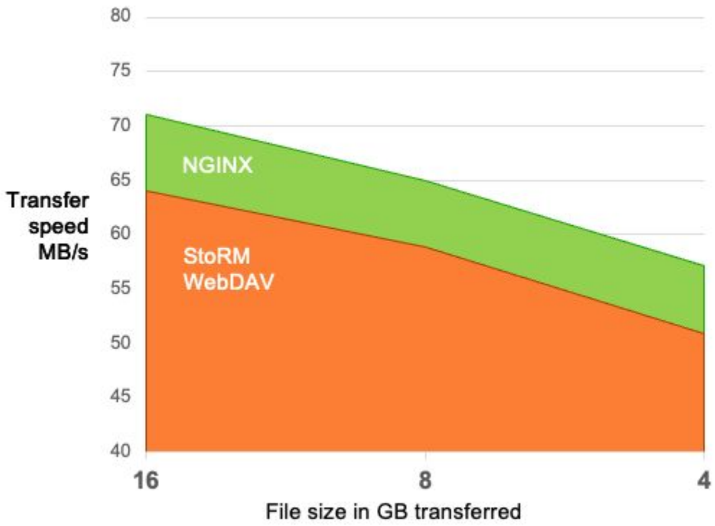
(4) The internal endpoint of NGINX is configured to retrieve the resource from file-system

(5) File is returned to the user

GET performance tests

Stress tests with the Vegeta tool have been performed to check the improvements in performance. On a VM with 4 CPU and 8GB of memory, we observed:

- an average of +10% in terms of transfer performance
- half usage in terms of memory
  - WebDAV ~820 MB vs NGINX ~491 MB
- a huge difference in terms of CPU consuming
  - WebDAV ~10% vs NGINX ~0,1%



Future outlooks and conclusions

Future outlooks

- Delegate PUT requests to NGINX
- Enable JSON Web Token (JWT) authentication in NGINX
  - It is already in place for StoRM Tape REST API (presented at CHEP 2023)
- Rely on an external policy engine for authorization
  - Configure StoRM WebDAV service in order to submit the request and user’s parsed authorities to an external Open Policy Agent engine
- Explore advanced load balancing solutions
  - another reason why we decided to move towards NGINX

Conclusions

- StoRM is a storage resource manager for disk-based storage systems
  - it is the Grid/Cloud storage solution adopted by the INFN-CNAF TIER 1 and some other TIER 2
- StoRM WebDAV is the StoRM component which provides a data transfer functionality through the WebDAV protocol
  - supports Third Party Copies
  - supports authorization based on JWT tokens or X.509/VOMS certificates
- To improve StoRM WebDAV performance and scalability, NGINX has been adopted in a new deployment model
  - handles authentication with VOMS proxies
  - handles HTTP GET requests
  - performance tests already give promising results
- Further improvements are on our roadmap and will be addressed soon !

Contacts and references

StoRM WebDAV source: https://github.com/italiangrid/storm-webdav

StoRM Documentation: http://italiangrid.github.io/storm/

Support mailing lists: storm-support@lists.infn.it storm-users@lists.infn.it

Developers mailing list: storm-devel@lists.infn.it

Other useful links:

- ngx_http_voms_module source code
- WebDAV protocol spec. Web Distributed Authoring and Versioning
- A RESTful approach to tape management in StoRM, CHEP 2023

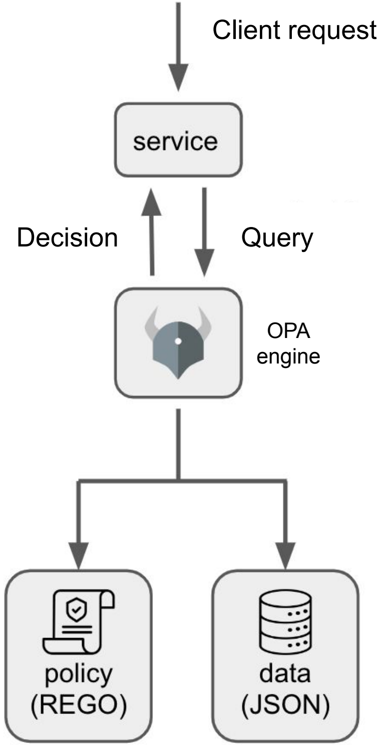
Many thanks to all the contributors

Federica Agostini, Luca Bassi, Roberta Miccoli, Enrico Vianello, Stefano Zotti, Francesco Giacomini

Bkp

Open Policy Agent

- Open Policy Agent (OPA) is an open-source authorization engine that:
  - unifies policy enforcement across the stack
  - is based on an high-level declarative language
  - allows the definition of policies as code
- OPA has is a very good candidate to implement **authorization** policies based on X.509/VOMS proxies or JWT tokens
  - already in use for other middleware softwares, such as StoRM Tape REST API
- It seems flexible enough to replace other authorization engines (*e.g.* Argus)



NGINX optimizations

In case of serving static content, NGINX documentation suggests some specific minor optimizations in order to help reaching optimal performance.

- Enabling sendfile

  → enabling the sendfile directive eliminates the step of copying the data into the buffer and enables direct copying data from one file descriptor to another.

- Enabling tcp_nopush

  → together with the sendfile on; directive, enables NGINX to send HTTP response headers in one packet right after the chunk of data has been obtained by sendfile.

- Enabling tcp_nodelay

  → with the tcp_nodelay directive on, the Nagle's algorithm is disabled. This algorithm is designed to consolidate small packets into a larger one, sending it with 200 ms delay. When serving large static files, the data can be sent immediately regardless of the packet size. This directive is used only for keepalive connections.

```
sendfile on;
tcp_nopush on;
keepalive_timeout 65;
tcp_nodelay on;
```