

Gfarmファイルシステムの概要と 最新機能

建部修見
筑波大学

Gfarmシンポジウム2023

- ハイブリッド開催（秋葉原）
- プログラム
 - Gfarm最新機能
 - 第三世代HPCI共用ストレージ
 - スマートシティにおけるデータ基盤
 - 次期ストレージシステム調査研究
 - LustreからGfarmへの透過的アーカイブ
 - 次世代HPCデータストア技術

Gfarmファイルシステム



- オープンソース広域分散ファイルシステム
 - <http://oss-tsukuba.org/software/gfarm/>
- サポート
 - NPO法人つくばOSS技術支援センター(日本他)
 - Libre Solutions Pty Ltd(オーストラリア)
- 特徴
 - インターネット経由で安全にアクセス可
 - ストレージを広域分散化
 - 性能・容量がスケールアウト
 - データアクセス局所性、ファイル複製
 - 無停止で拡張、縮小可能
 - 単一障害点なし
 - 複製数維持機能、ホットスタンバイMDSサーバ
 - ローリングアップデート
 - データ完全性を保証しサイレントデータ損傷も対応可

ossTsukuba
oss-tsukuba.org

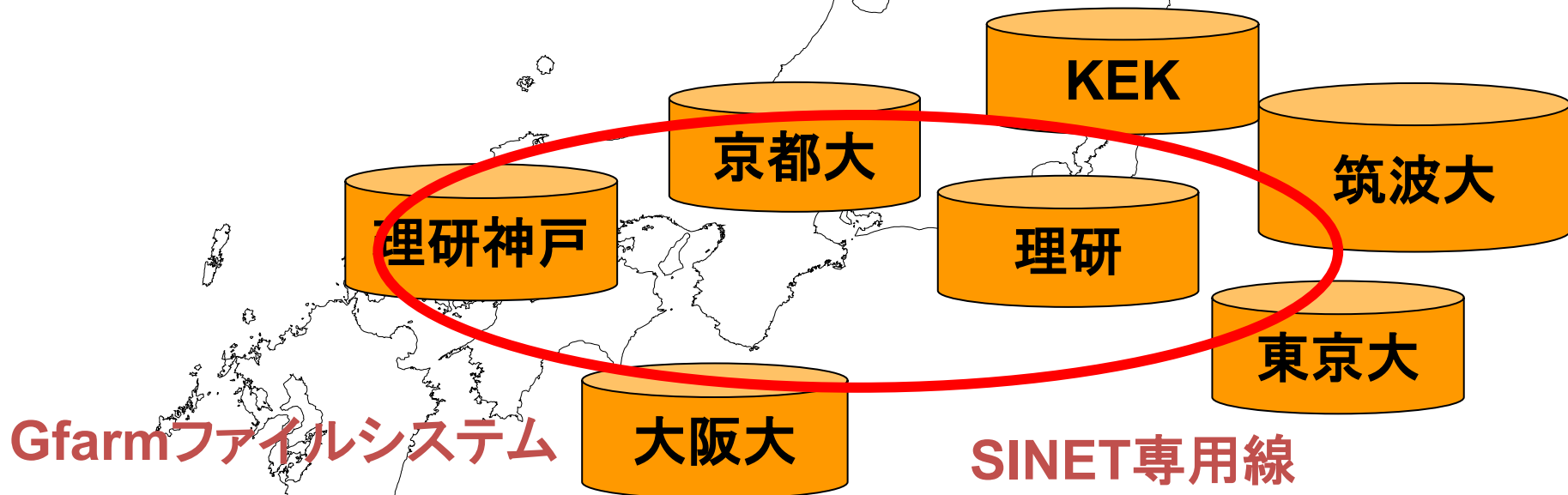


Gfarmファイルシステム(2)

- JLDG(22PB、7拠点)
 - 筑波大スパコン、富岳でマウント可。他拠点ではGridFTPでアクセス
- HPCI共用ストレージ(~100PB、2拠点)
 - HPCI構成機関(9大学スパコン等)でマウント可
 - 任意場所でクライアント設置可
- 他、NICTサイエンスクラウド等で実運用
- 計算ノードのローカルディスクによるデータ解析
 - すばる望遠鏡データ解析、メタゲノム解析
- Pwrakeワークフローシステム、MapReduce、MPI-IO、バッチキューイングシステム
 - データ局所性を高めるプロセススケジューリング
 - ディスクキャッシュを有効利用するプロセススケジューリング
 - データ局所性を高めるファイル複製作成

Japan Lattice Data Grid (JLDG)

- 国内素粒子物理学研究者のための22PB規模の広域共有ファイルシステム
 - スパコンで数ヶ月～数年計算したシミュレーションデータの共有
 - 各拠点のファイルサーバを束ね、ファイルは必要な数の複製を作成
 - 各拠点では格納場所を意識せずアクセス
 - 複製を持っている拠点はアクセスが高速に



HPCI共用ストレージ

- 大学情報基盤センターをはじめ全国からマウント可能な共有ファイルシステム(～100PB)
- スパコン間のデータ共有、共有データ格納



西拠点 (R-CCS)



東拠点 (東京大)

最新機能

主なリリース

日付	version	新機能、更新機能
2023/9/3	2.8.0	<ul style="list-style-type: none">• TLS, OAuth2認証• マルチテナント、gfcg
2023/3/14	2.7.21	<ul style="list-style-type: none">• gfptar
2022/3/24	2.7.20	<ul style="list-style-type: none">• ARMサポート
2021/4/23	2.7.18	<ul style="list-style-type: none">• network_send_timeout
2020/9/17	2.7.17	<ul style="list-style-type: none">• MTセーフ、read-only FS
2019/9/10	2.7.14	<ul style="list-style-type: none">• Gfarm/BB
2016/12/8	2.7.0	<ul style="list-style-type: none">• InfiniBand RDMAサポート• ディレクトリクォータ

TLS (Transport Layer Security)

- TLS 1.3以降が必要 (OpenSSL 1.1.1以降)
- CA公開鍵
 - /etc/pki/tls/certs/gfarm/
- gfmd証明書
 - /etc/pki/tls/certs/gfmd.crt
 - /etc/pki/tls/private/gfmd.key
- gfsd証明書
 - /etc/pki/tls/certs/gfsd.crt
 - /etc/pki/tls/private/gfsd.key

TLS共有鍵認証 (2.8.0)

- Gfarm共有鍵を用いた認証
- 通信はTLSで暗号化
- 設定
 - gfarm2.conf
 - auth enable tls_shared_secret *

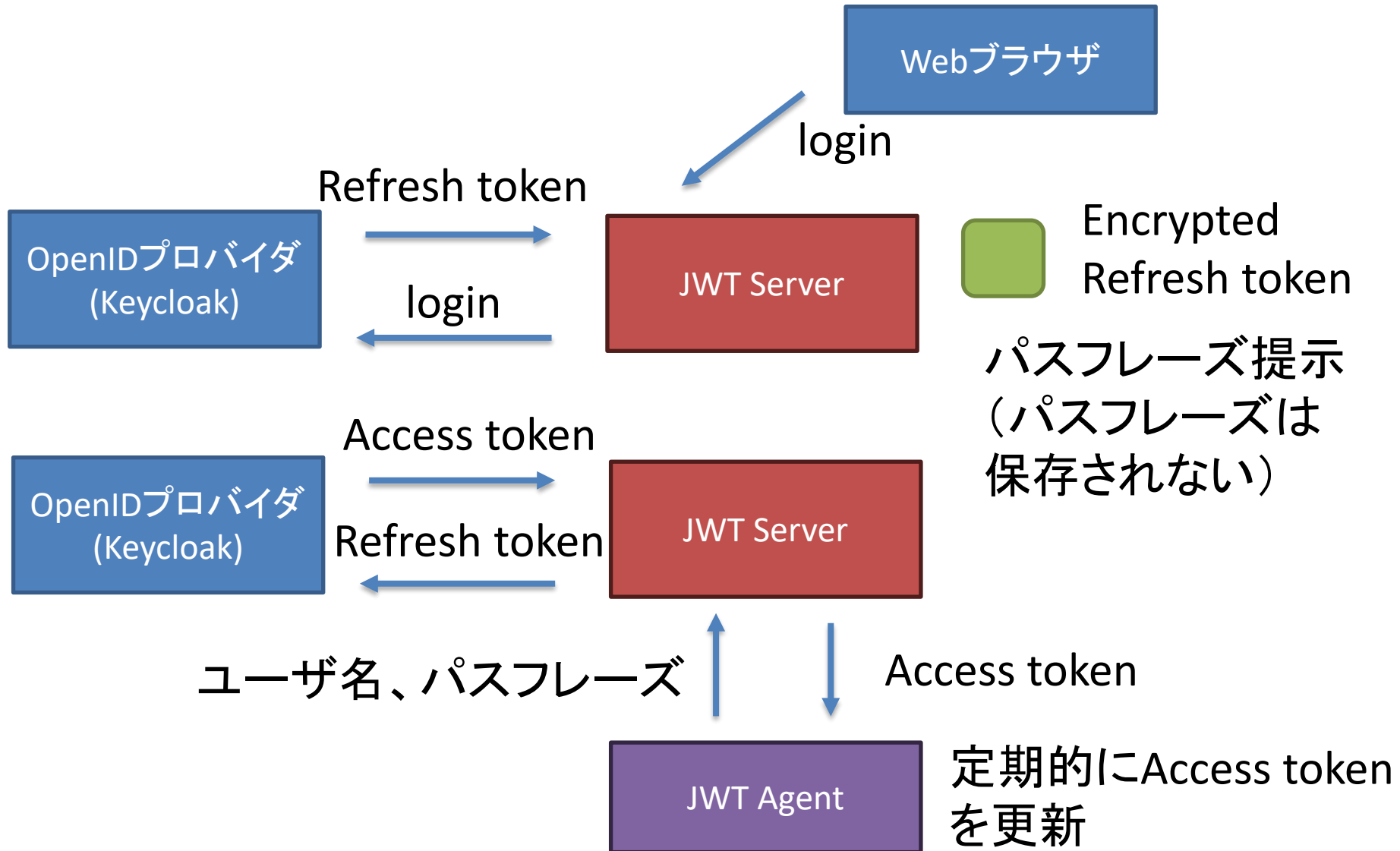
TLS証明書認証 (2.8.0)

- ホスト証明書、代理証明書を利用して相互認証
- GSIの代替
- 設定
 - gfarm2.conf
 - auth enable tls_client_certificate *

SASL (Simple Authentication and Security Layer)のサポート

- メールクライアント等でよく利用される認証セキュリティフレームワーク
- これまでの共有鍵認証、GSSAPI (Generic Security Services Application Program Interface)に加え、SASLをサポート
- SASL XOAUTH2メカニズムのプラグイン開発
 - <https://github.com/oss-tsukuba/cyrus-sasl-xoauth2-idp>

OAuth/OIDC認証における アクセストークンの取得



OAuth認証 (2.8.0)

- 設定
 - gfarm2.conf
 - auth enable sasl *
 - sasl_mechanisms XOAUTH2
 - /usr/lib/sasl2/gfarm{-client}.conf
 - xoauth2_scope: xxx
 - xoauth2_aud: xxx
 - xoauth2_user_claim: xxx
- gfuser -A username SASL sasl_global_name
- jwt-agent/jwt-logonでアクセストークンを取得

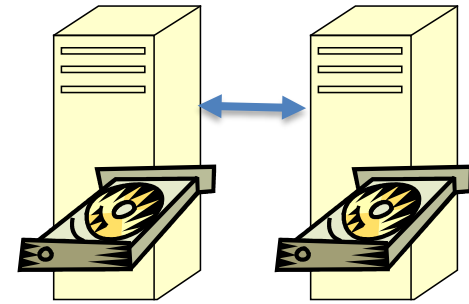
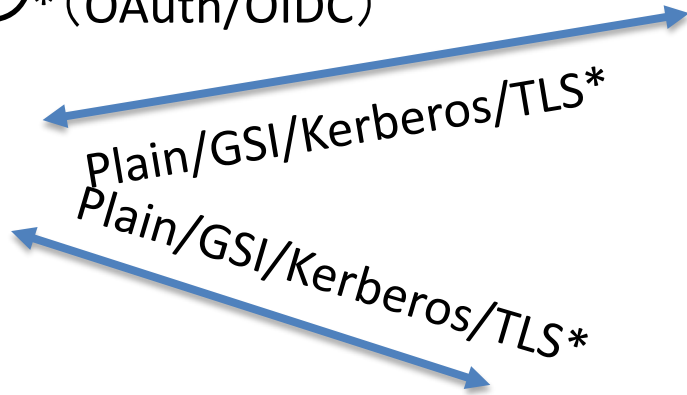
Anonymous認証 (2.8.0)

- 認証なしのanonymousアクセスを提供
 - サーバ側でanonymousアクセスを許している場合
- 設定
 - gfarm2.conf
 - auth enable sasl *
 - sasl_mechanisms ANONYMOUS
 - sasl_user dummy
- gfuser -c anonymous anon / ""

認証・セキュリティについて

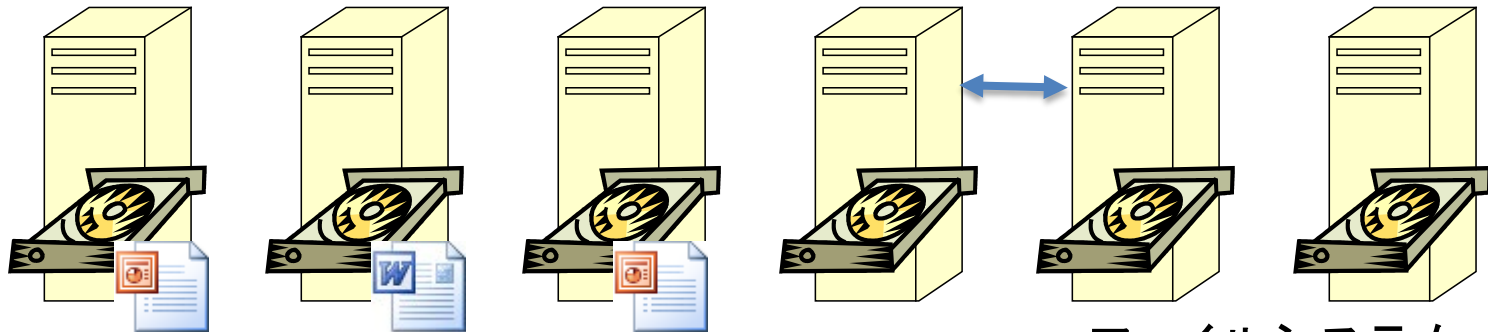
共有鍵 (Gfarm)
(代理)証明書 (GSI/TLS*)
チケット* (Kerberos)

トークン* (OAuth/OIDC)



メタデータサーバ

共有鍵 (Gfarm)
証明書 (GSI/TLS*)
チケット* (Kerberos)

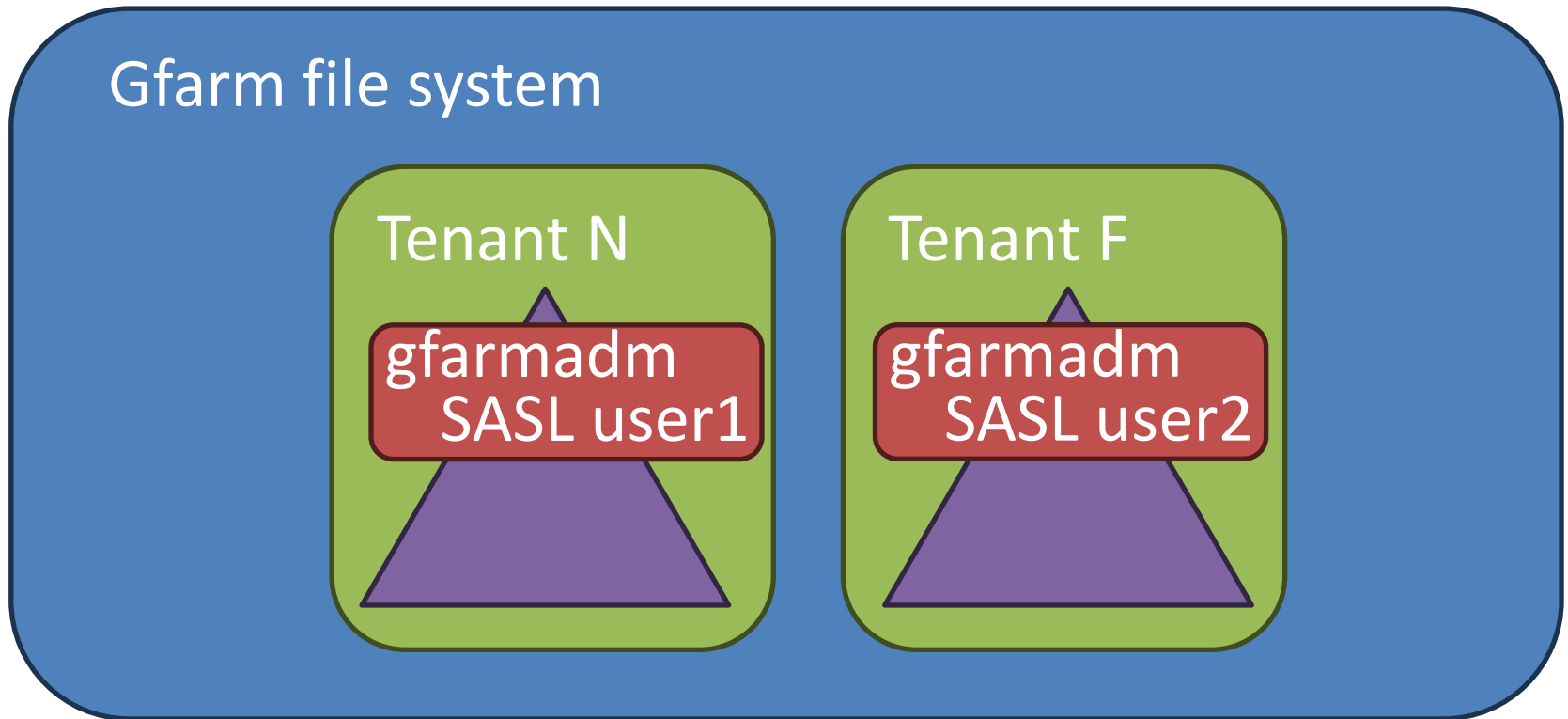


ファイルシステムノード群

共有鍵 (Gfarm) 証明書 (GSI/TLS*) チケット* (Kerberos)

マルチテナント (2.8.0)

- それぞれのテナントが独立してファイルシステムを管理



並列コピー (2.8.0)

- `gfpcat` – 複数のファイルを結合して並列コピー

```
$ gfpcat [ -j 4 ] -o out file1 file2 ...
```

- `gfcop` – 単一ファイルの並列コピー

```
$ gfcop [ -j 4 ] src dest
```

gfptar (2.7.21)

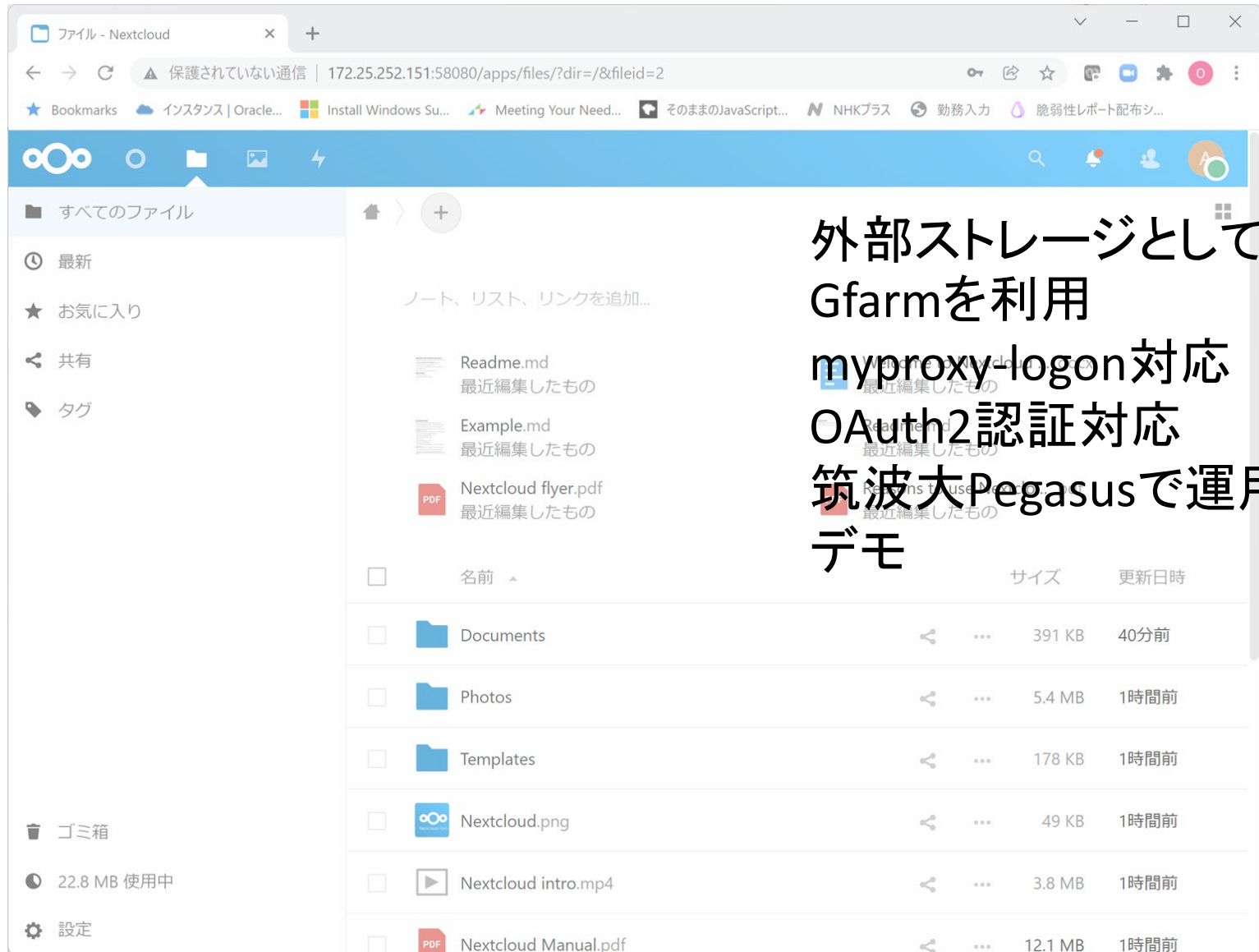
- 多数の小ファイルを固めてGfarmに保存
- サーバ障害時のデータ保全対応(2.7.23)
- Too many open filesエラーの解消(2.7.24)
- tarファイルを分散配置、並列転送性能向上(2.7.25)

```
gfptar -c 出力ディレクトリ 入力エントリ ...
```

```
gfptar -x 出力ディレクトリ アーカイブディレクトリ [エントリ ...]
```

```
gfptar -t アーカイブディレクトリ
```

Nextcloudコンテナ



外部ストレージとして
Gfarmを利用
myproxy-logout対応
OAuth2認証対応
筑波大Pegasusで運用中
デモ

まとめ

- Gfarmファイルシステム
 - NPO法人つくばOSS技術支援センターによるサポート
 - <https://github.com/oss-tsukuba/>
- HPCI共用ストレージ、JLDGなど実運用実績
- 暗号化ファイルシステム、完全性
- Gfarm 2.8.0を9/3にリリース！
 - TLS対応, OAuth2, Kerberos認証
- S3コンテナ、Nextcloudコンテナ