

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2024

- ハイブリッド開催(秋葉原)
- プログラム
 - Gfarm最新機能
 - Japan Lattice Data Grid (JLDG)
 - データコンピューティング基盤
 - GfarmのクラウドデプロイとLustre HSM
 - DDN Infinia
 - 第三世代HPCI共用ストレージ東拠点

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

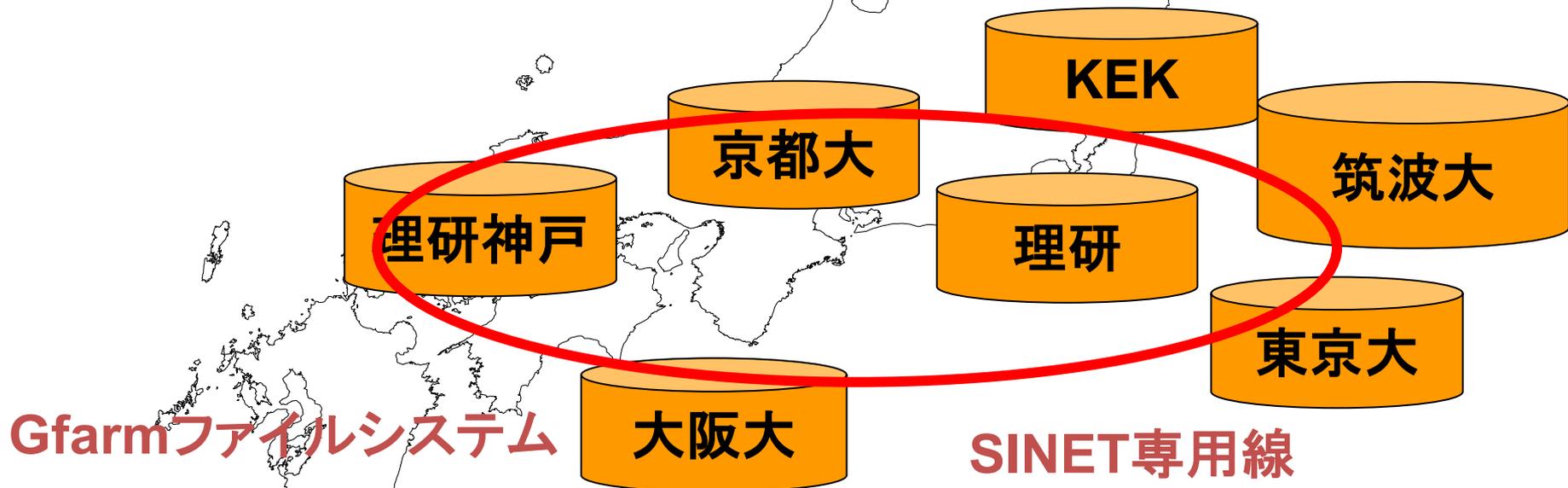


Gfarmファイルシステム(2)

- 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(Locality-aware 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
2024/9/4	2.8.5	• Gfmd起動高速化、バグ修正
2024/6/28	2.0.0	• Nextcloud-Gfarm OAuth2対応、 ラージファイル対応
2024/2/20	1.2.21	• Gfarm2fs OAuth2対応
2024/1/25	1.1.0	• Gfarm GridFTP DSI UTIME対応
2023/9/3	2.8.0	• TLS, OAuth2認証 • マルチテナント、gfcp
2023/3/14	2.7.21	• gfptar

Gfmd起動高速化

- 起動時にジャーナルのPostgreSQLへの反映を非同期に

<info> [1003273] start applying db journal

<info> [1005736] postgresql backend: set **synchronous_commit=0**

<info> [1005736] postgresql backend: set **synchronous_commit=1**

<info> [1003275] end applying db journal

- 大量のファイル作成、削除等の直後のgfmd再起動時間の短縮

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 *

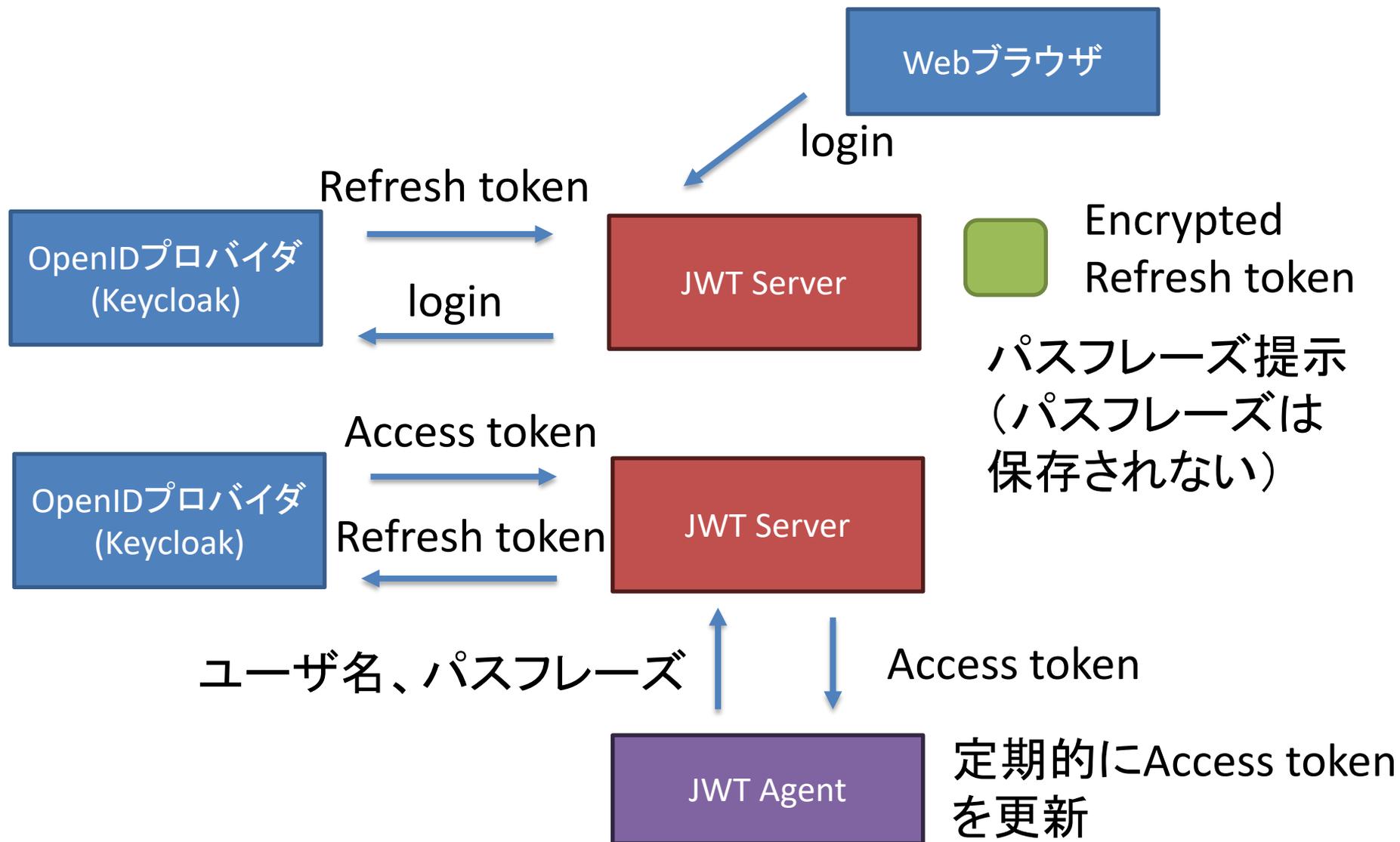
TLS証明書認証 (2.8.0)

- ホスト証明書、クライアント証明書(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

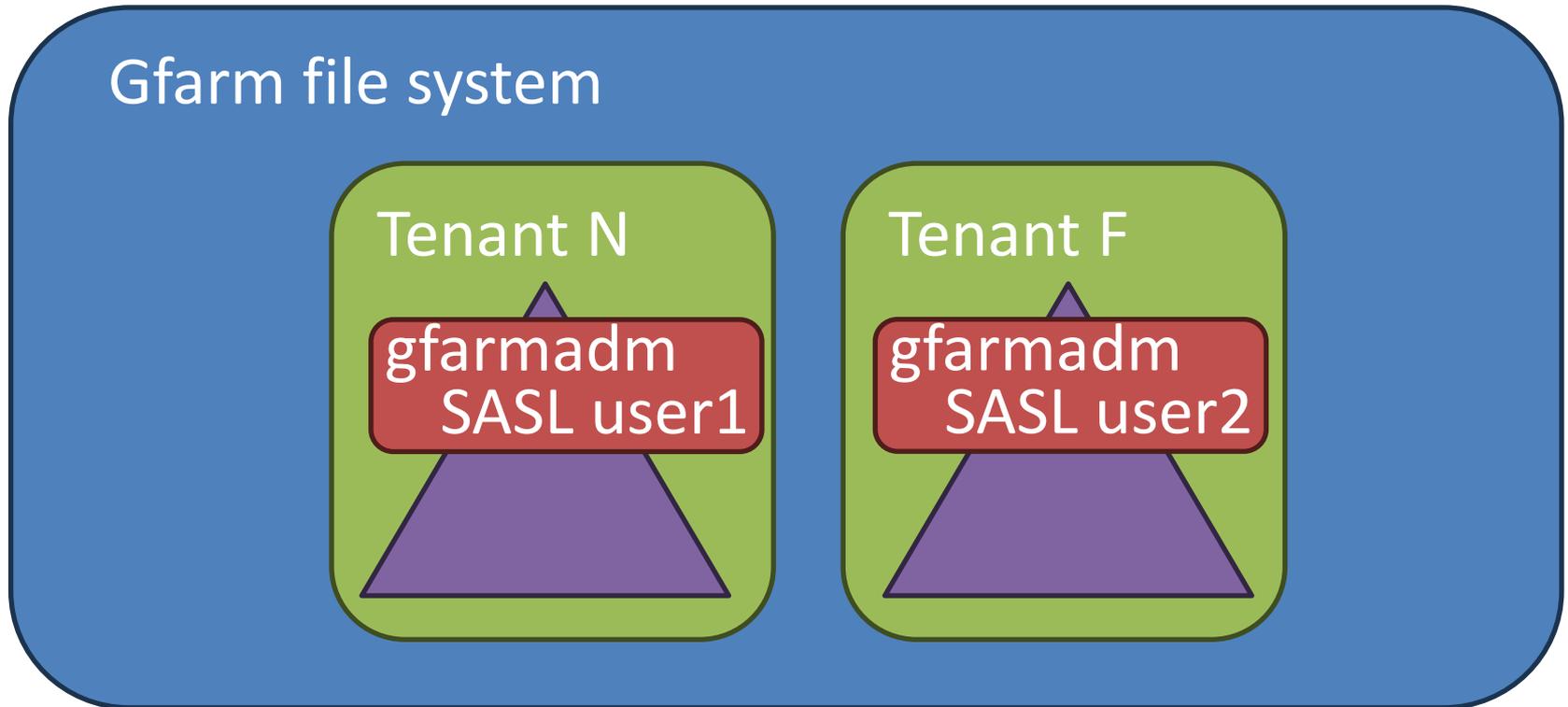


OAuth認証 (2.8.0)

- 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
 - xoauth2_issuers: xxx
- gfuser -A username SASL sasl_global_name
- jwt-agent/jwt-logonでアクセストークンを取得

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)

- gfpconcat – 複数のファイルを結合して並列コピー

```
$ gfpconcat [ -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트리 ...
```

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

```
gfptar -t アーカイブディレクトリ
```

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 2.8.5を9/4にリリース！
 - Gfmd起動高速化、バグ修正
- S3コンテナ、Nextcloudコンテナ