

Gfarmワークショップ2021

Gfarm-S3-MinIO

GfarmのS3互換ゲートウェイ

株式会社創夢

石橋拓也

2021/03/05

本日の内容

- Gfarm-S3-MinIO概要
- Gfarm-S3-MinIO使い方
- S3クライアント利用例
- Gfarm-S3-MinIOインストール・設定方法
- Gfarm-S3-MinIO実装概要
- Gfarm-S3-MinIOお試し利用方法
- Gfarm-S3-MinIO注意事項

自己紹介: 石橋拓也

- 2002年～: Gfarm 関連お手伝い開始
- gfarmfs-fuse (Gfarm v1)
- gfprep, gfpcopy
- 複製数維持 (replica_check)
- 高負荷時安定化, gfsdスプールチェック
- クォータ, ACL
- HPCI共用ストレージマニュアル作成
- gfsd読み込み専用化
- クライアントライブラリスレッドセーフ化
- gfarm_gridftp_dsi, gfarm_samba
- gfarm-s3-minio

会社紹介: 株式会社 創夢

- <https://www.soum.co.jp/technology/>
- UNIX/Linux, ネットワーク, OSS
- 組み込み
- 研究支援
- 環境構築・運用

Gfarm-S3-MinIO

概要

Gfarm-S3-MinIOとは

- S3クライアントを使用して、Gfarm上のファイルにアクセス可能
- 下記をインストールして利用
- gfarm-s3-minio
 - MinIOリポジトリ(GitHub)をfork
 - Gfarmアクセス用のGateway実装をMinIOに追加
 - Gfarm上の特定ディレクトリのみをS3 APIで公開
- gfarm-s3-minio-web
 - ユーザーごとにMinIOを起動するためのWebインターフェース

MinIOとは

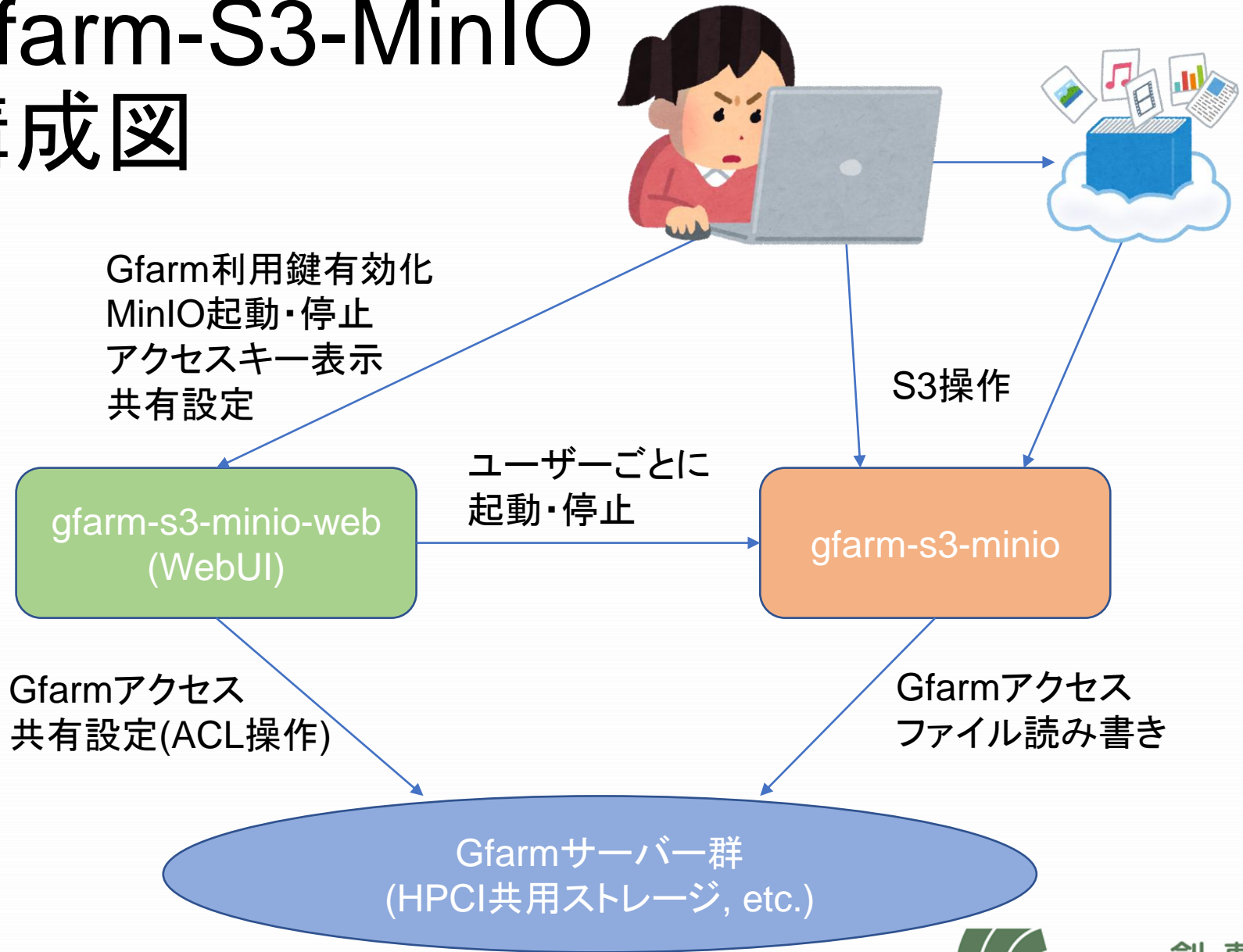
- <https://min.io/>
- Amazon S3互換(オブジェクト)ストレージのひとつ
 - ローカルファイルなどをS3 APIで公開可能
 - Gateway機能: Azureなど他ストレージを中継も可能
 - Go言語
 - 最上位ディレクトリがバケット
- S3互換ストレージ
 - アクセスキーで利用制御
 - 各種S3クライアントを利用してアクセス

Gfarm-S3-MinIO

構成概要

- Gfarm サーバー群
- MinIO サーバー (Gfarmに中継)
- WebUI (MinIO制御、共有操作)
- (Webブラウザー)
- (S3クライアント)

Gfarm-S3-MinIO 構成図



Gfarm-S3-MinIO 使い方

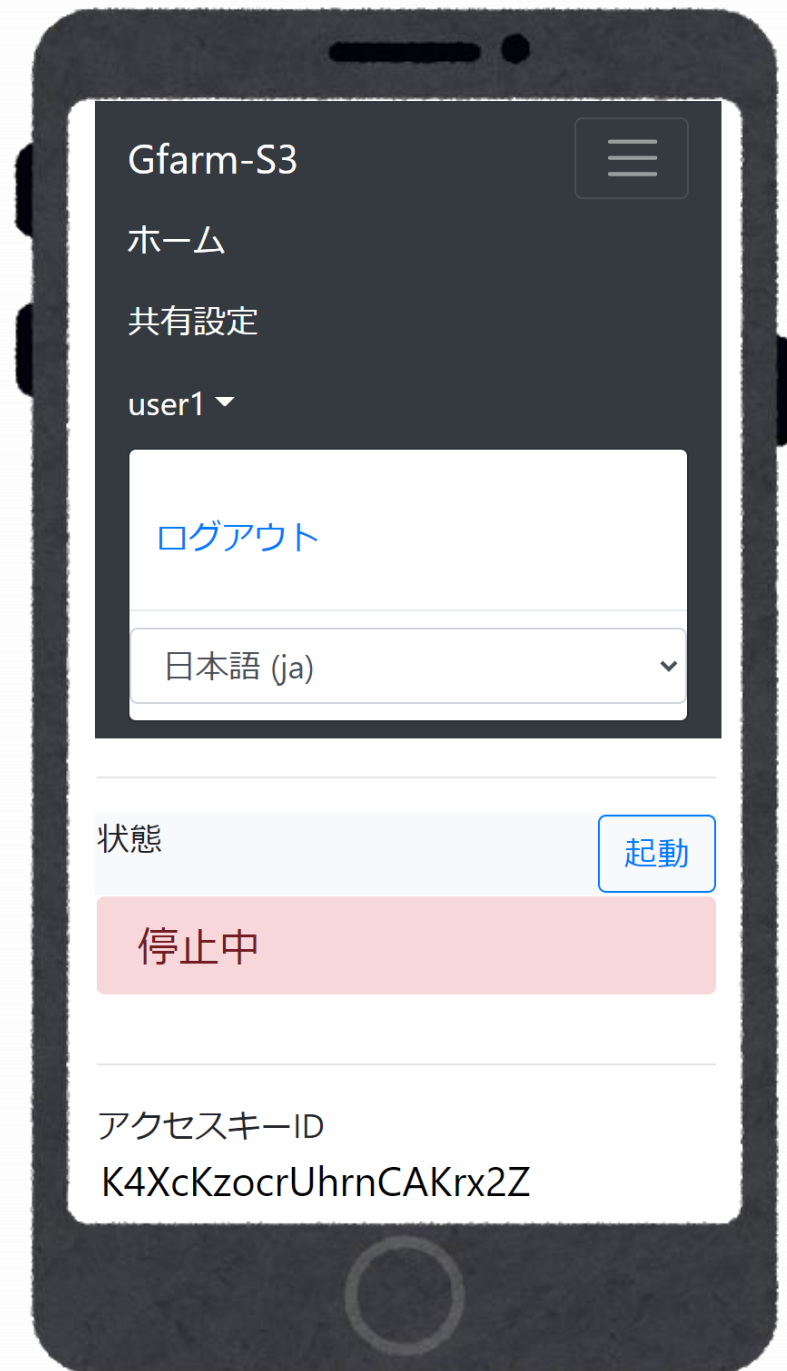
ログイン

- ユーザーごとに操作・管理
- myproxy-logon
 - パスフレーズ
- grid-proxy-init
 - パスフレーズ
- Gfarm共有鍵
 - ~/.gfarm_shared_key由来のハッシュ値がパスワード
 - 事前にCLIで値を表示して把握しておく



メニュー

- ホーム画面
 - 起動制御、状態表示
- 共有設定
- ログアウト
- 表示言語切り替え



起動・停止

- ボタンでMinIOを起動する
 - 自動起動しない
 - ユーザーごとにMinIOプロセス起動
- アクセスキーIDはユーザーごとに管理者が事前に設定
- シークレットアクセスキーは自動生成される
 - 変更可能



アクセスキー

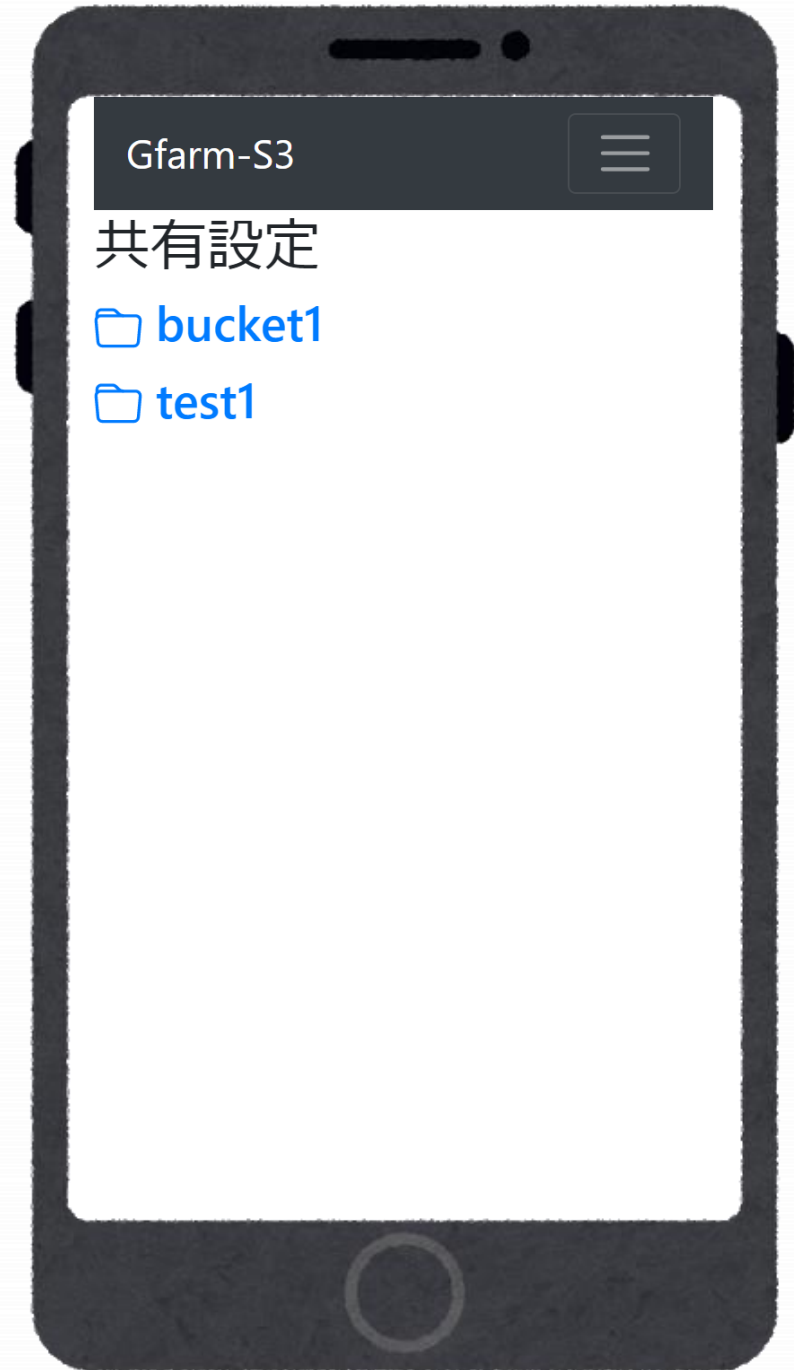
- S3クライアント側に必要な設定を表示
 - アクセスキーID: ユーザー名相当
 - シークレットアクセスキー: パスワード相当
- 各ユーザーのホームディレクトリにシークレット値が保存される
- リバースプロキシ(Apache)がアクセスキーID (HTTPヘッダ内)を見てユーザーごとに起動したminioのポート番号に転送

共有設定 概要

- (例) 名前bucket1のバケットを作成しておく
 - S3クライアントを使用して作成する
 - WebUI ではバケット作成不可
- Gfarmには
/share/user1/bucket1/
として作成される
- WebUI で bucket1 に対して共有設定
 - ユーザーやグループに対してRead or Write を許可
- 他ユーザーのS3クライアントからは
/sss/user1/bucket1 に見える

共有設定 バケット選択

- バケットを選択して、共有設定画面へ



共有設定 設定画面

- 「エントリ追加」ボタン
- ユーザー名orグループ名を検索して追加
 - Gfarmユーザーのrealnameも検索可能
- read,writeスイッチ
- 「変更を適用」ボタン



共有設定 設定画面

- 変更を適用後の画面
- GfarmのACLが設定される
- 特別なバケット名 SSS
- 他ユーザーから共有されたバケット
 - /sss/ユーザ名/バケット名
- アクセスできないディレクトリとファイルは見えない
 - 隠れているだけ



S3クライアント 利用例

S3クライアントの例

- コマンドラインインターフェース
 - AWS CLI (awsコマンド), s3cmd
- ライブラリ
 - Boto3 (Python)
- Windows用クライアント
 - WinSCP, Cyberduck, firedrive
- Mac用クライアント
 - Cyberduck
- スマホアプリ
 - BucketAnywhere for S3
- Webアプリ ... 多数あるはず
 - Nextcloud を試してみた
- FUSEマウント (gfarm2fsを使ったほうが良いが)
 - s3fs, goofys

AWS CLI利用例

- aws configure
 - アクセスキー設定
- aws s3 --endpoint-url <http://ホスト名:18080/>
mb bucket1
- aws s3 --endpoint-url ...
cp file1 s3://bucket1/dir1/file1
- aws s3 --endpoint-url ...
cp s3://bucket1/dir1/file1 file1-copied
- aws s3 --endpoint-url ...
ls s3://bucket1

Nextcloudと連携 外部ストレージ設定 (External storage supportアプリ)



The screenshot shows the Nextcloud External Storage configuration page. The left sidebar contains navigation options: 個人 (Personal), 個人情報 (Personal information), セキュリティ (Security), アクティビティ (Activity), 外部ストレージ (External storage), モバイル & デスクトップ (Mobile & Desktop), アクセシビリティ (Accessibility), 共有中 (Shared), Flow, プライバシー (Privacy), 管理 (Administration), 概要 (Overview), サポート (Support), 基本設定 (Basic settings), and 共有 (Sharing).

外部ストレージ i

外部ストレージを使用すると、外部ストレージサービスおよびデバイスをセカンダリNextcloudストレージデバイスとしてマウントできます。また、ユーザーが独自の外部ストレージサービスをマウントできるようにすることもできます。

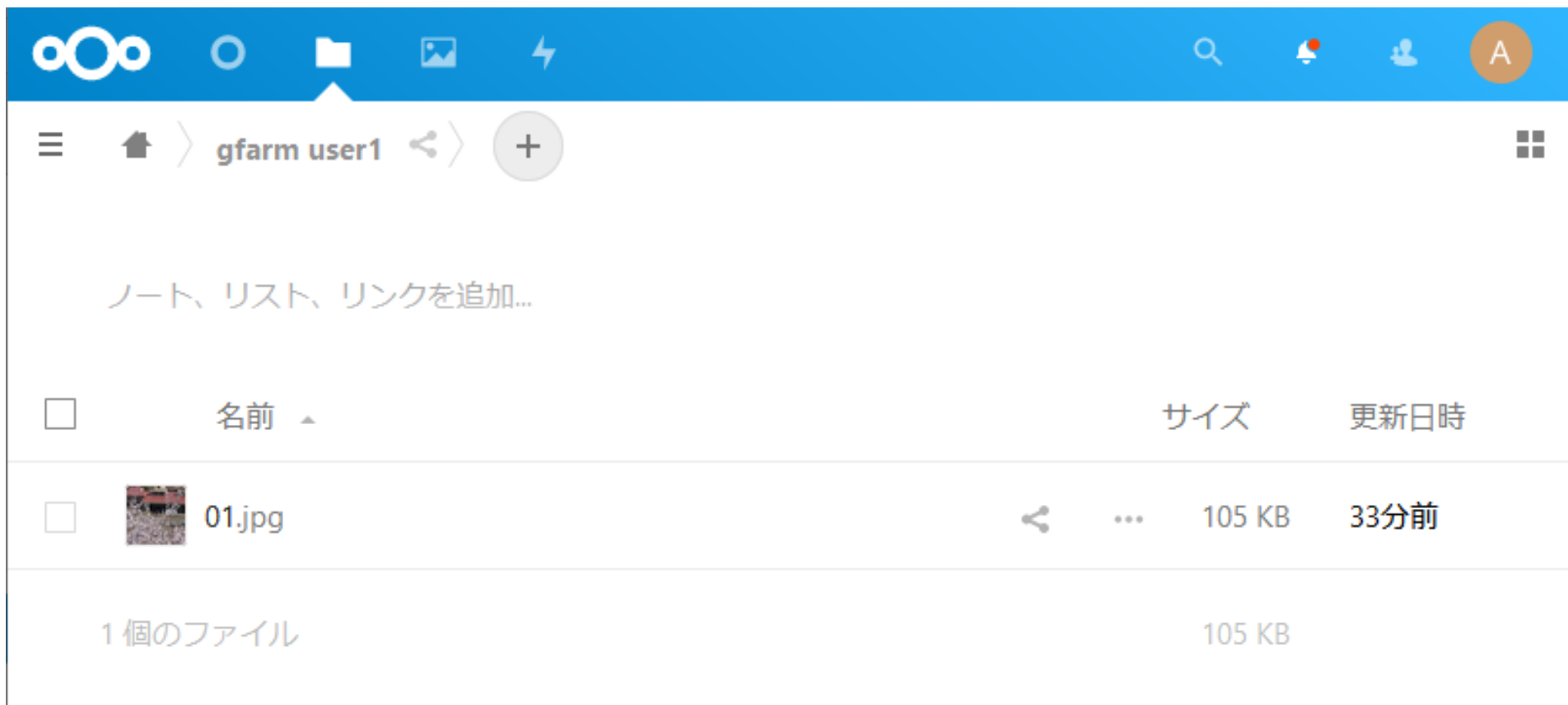
"smbclient" はインストールされていません。"SMB / CIFS", "Nextcloud ログインを利用したSMB / CIFS" のマウントは不可能です。システム管理者にインストールを依頼してください。

フォルダー名	外部ストレージ	認証	設定	利用可能
			bucket1	
			192.168.56.2	
			48080	
			リージョン	
<input checked="" type="checkbox"/>	gfarm user1	Amazon S3	アクセスキー	<input type="checkbox"/> SSLを有効
				<input checked="" type="checkbox"/> パス形式を有効
				<input type="checkbox"/> レガシー認証(v2)
			:XcKzocrUhrnCAKrx2Z	
			

フォルダー名 ストレージを追加

ユーザーに外部ストレージの接続を許可する

Nextcloudと連携 MinIO経由でGfarmアクセス



The screenshot displays the Nextcloud web interface. At the top, there is a blue navigation bar with various icons including a home icon, a folder icon, a document icon, a lightning bolt icon, a search icon, a notification bell, a user profile icon, and a circular profile picture with the letter 'A'. Below the navigation bar, the breadcrumb path shows 'gfarm user1' with a share icon and a plus icon. The main content area contains the text 'ノート、リスト、リンクを追加...' (Add notes, lists, links...). Below this, there is a table listing files with columns for '名前' (Name), 'サイズ' (Size), and '更新日時' (Update Time). The table shows one file named '01.jpg' with a size of 105 KB and an update time of 33 minutes ago. At the bottom, a summary row indicates '1 個のファイル' (1 file) with a total size of 105 KB.

<input type="checkbox"/>	名前 ▾	サイズ	更新日時
<input type="checkbox"/>	 01.jpg	105 KB	33分前
1 個のファイル			105 KB

Gfarm-S3-MinIO インストール・設定方法

インストール・設定 環境の条件

- Gfarmクライアントを利用可能なホスト
- CentOS 7で動作を確認
- SSHログインノード(複数ユーザー利用想定)
または個人利用専用(自分で管理)ホスト
- Webサーバーを設置可能なホスト
 - 既にWebサーバーがある場合は、
その設定を変更可能であること
- Gfarm上でファイル共有する場合は、
Gfarm管理者に専用ディレクトリ作成を依頼

インストール・設定 手順概要

- ソースコード用意
 - gfarm-s3-minio : developブランチ
 - gfarm-s3-minio-web : gfarmブランチ
- 依存パッケージインストール
- ./configure --with-いろいろ...
 - インストール先など指定
- make install-go, make, sudo make install
- ユーザー登録コマンド実行
- Apacheの設定に対して apache-gfarm-s3.conf 内容を追記
- gunicornの自動起動設定
- (Gfarm管理者に依頼)共有用ディレクトリ作成

インストール・設定 ホスト管理者による操作

- gfarm-s3-minio-webインストール
- gfarm-s3-useradd
 - ユーザーを登録する
 - グローバルユーザー名 (Gfarm user)
ローカルユーザー名
S3アクセスキーID
を関連づける
- アップロード時キャッシュ用一時ファイル領域作成

インストール・設定 Gfarm管理者に依頼

- 共有用ディレクトリ作成
- gfarm:/share/ユーザー名

インストール・設定 自動で設定される項目

- /etc/sudoers.d/gfarm-s3
 - 必要なコマンドのみ代理実行を許可
 - wsgiユーザーが、各ユーザーの代わりにminio起動、Gfarm操作
 - wsgiユーザーが、Apacheのリバースプロキシ設定を変更

インストール・設定

/etc/sudoers.d/gfarm-s3

- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfarm-s3-server
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/bin/grid-proxy-info
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/bin/grid-proxy-init
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/bin/myproxy-logon
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfkey
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfuser
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfgroup
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfls
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfgetfac
- wsgi ALL=(GFARMS3) NOPASSWD: /usr/local/bin/gfsetfac
- wsgi ALL=(root) NOPASSWD: /usr/sbin/apachectl

Gfarm-S3-MinIO

実装概要

Gfarm-S3-MinIO ソフトウェア構成

- libgfarm, gf*コマンド
- MinIO (gfarm-s3-minio版)
- リバースプロキシ (Apache)
- WSGIサーバー (gunicorn)
- Webアプリフレームワーク (Django)
- Go, Python, JavaScript 各標準的なライブラリ

MinIOにGatewayの実装を追加

- pkg/gfarm/gfarmClient.go
 - GfarmのGo言語クライアントライブラリを追加
- cmd/gateway/...
 - gfarm を追加
 - (参考)他の実装: azure, gcs, hdfs, nas, s3

マルチパートアップロード (1)

- S3クライアントはサイズが大きいファイルを分割して並列アップロードすることがある
- MinIO gateway 実装側では、パートごとに受信し、最後に結合が必要
- パートごとのファイルをGfarm上に直接書く仕組みで実装するとしたら、結合コストが発生してしまう
 - Gfarmサーバー内で結合できないため
 - Gfarmから読んで結合して書く必要がある
- ローカルFSで一旦キャッシュし、最後にGfarmにアップロードしながら結合する実装とした
- 改善の余地ありそう

マルチパートアップロード (2)

- ローカルキャッシュはGfarmより小さいことが多い
 - キャッシュが溢れた場合はGfarm上に書くようにした
 - その場合、結合に時間がかかる
- S3の仕様では、
事前に各パートごとのサイズが分からない、
どのパートが結合対象となるかわからないため、
最後に結合する必要がある
 - もしパートごとにサイズと位置が事前にわかれば、
pwrite (位置指定書き込み)でパートごとに直接Gfarmに
書く方式を使えた

共有設定の実装

gfarm.effective_perm 拡張属性

- この機能をgfmdに新設した
- この機能をgfarm-s3-minioから利用
- 他ユーザーから共有されていないディレクトリをS3クライアントからは隠す目的
 - 隠すだけ
 - Gfarm直接利用では見える
- ユーザーが読み書きできるかどうかをgfmd側で判断した値を拡張属性で返す
 - GfarmのAPIは、readdir() 相当の処理とともにファイル・ディレクトリごとに拡張属性を返す
 - この機能を使用しても通信回数は増えない

Gfarm-S3-MinIO お試し利用方法

開発用環境で試す手順

- docker, docker-composeをインストール
- git clone -b 2.7 https://github.com/oss-tsukuba/gfarm.git
- cd gfarm
- git clone https://github.com/oss-tsukuba/gfarm2fs.git
- cd docker/dev
 - Gfarmサーバー環境をDockerコンテナで構築
 - 開発用: 特権を利用しているため、他用途環境では使用不可
- touch config.mk
- (必要なら編集) vi config.mk
 - config-default.mk 参考
- cd dist/centos7/src
- make reborn
 - 実行するたびにこの環境のGfarmデータは初期化される
- make s3setup
- ブラウザで <http://このホスト名:18080/>

開発用環境で試す手順 補足

- WebUIログインのためのパスワードは
コンテナ内でコマンドを実行して表示
 - gfarm-s3-sharedsecret-password
 - 開発環境ではGfarmの共有鍵をベースとしたログイン
となっている
 - 開発用環境では環境構築メッセージ最後に自動出力

Gfarm-S3-MinIO

注意事項

制限事項

- Nextcloud, WinSCP ではディレクトリを改名できない
 - オリジナルMinIOでも不可
 - goofys は空ディレクトリを改名できない
 - 他にも改名できないクライアントがあるかもしれない
 - S3には、そもそも改名APIが無いので、各クライアントが、コピー・削除APIを使って改名する
 - (補足: Nextcloudには、改名処理が影響しない方法でS3をプライマリストレージとすることもできる機能があるが、パス名は連携されない)
- (S3の仕様? MinIOの仕様? S3クライアント次第?)

注意事項

- 共有バケットsssを削除操作しないよう注意
 - sss自体は仮想的な名前なので消せないため、この操作は意図と異なる
 - S3クライアントは、sssから参照されているエン트리すべてを消そうとする
 - 自分所有のファイルはすべて消える
 - それだけでなく、他人が共有しているファイルのうち、自分が削除可能なファイルも消える
 - readonlyファイルのみが残る
 - `rm -rf` と同じこと
 - サーバー側では禁止できない

ありがとうございました