

# **NEC Scalable Technology File System for AI (ScaTeFS for AI) システムガイド**

#### 輸出する際の注意事項

本製品（ソフトウェアを含む）は、外国為替および外国貿易法で規定される規制貨物（または役務）に該当することがあります。

その場合、日本国外へ輸出する場合には日本国政府の輸出許可が必要です。

なお、輸出許可申請手続きにあたり資料等が必要な場合には、お買い上げの販売店またはお近くの当社営業拠点にご相談ください。

---

# は し が き

このマニュアルは、NEC Scalable Technology File System for AI 用のシステムガイドです。ここではマニュアルの目的、対象読者について説明します。

## マニュアルの目的

このマニュアルでは、NEC Scalable Technology File System for AI の概要やクライアント、IO サーバの SW/HW 構成、ライセンスに関する説明を目的としています。

## 対象読者

このマニュアルは、次の方を対象読者として記述しています。

- システム管理者
- 一般利用者

---

## 備考

- (1) Linux は Linus Torvalds氏 の日本およびその他の国における登録商標あるいは商標です。
- (2) Red Hatは米国およびその他の国でのRed Hat, Inc. の登録商標もしくは商標です。
- (3) CLUSTERPROは日本電気株式会社の登録商標です。
- (4) Windowsは、米国Microsoft Corporationの米国およびその他の国における登録商標または商標です。
- (5) NVIDIAは、米国およびその他の国におけるNVIDIA Corporationの登録商標または商標です。
- (6) その他記載の会社名、製品名は、それぞれの会社の商標、もしくは登録商標です。

---

## 本書の読み進め方

本書は、次の構成となっています。章ごとの対象読者の範囲は、表の一番右の列にその範囲を示しています。

章	タイトル	内容	対象読者
1	NEC Scalable Technology File System for AIの概要	NEC Scalable Technology File System for AIの概要について、記載しています。	システム管理者 一般利用者
2	IOサーバのHW/SW構成	IOサーバのHW/SW構成について、記載しています。	システム管理者 一般利用者
3	クライアントのHW/SW構成	ScaTeFS for AI の ク ラ イ ア ン ト の HW/SW構成について、記載しています。	システム管理者 一般利用者
4	ScaTeFS for AIのライセンスについて	ScaTeFS for AIのライセンスについて記載しています。	システム管理者 一般利用者

## 関連説明書

- 『NEC Scalable Technology File System for AI (ScaTeFS for AI) インストールガイド』
- 『NEC Scalable Technology File System for AI (ScaTeFS for AI) ユーザーズガイド』
- 『NEC Scalable Technology File System for AI (ScaTeFS for AI) ソフトウェアライセンス管理説明書』

各種説明書は以下のNECサポートポータルのWebページから参照できます。

<https://www.support.nec.co.jp/View.aspx?id=3170102881>

## 用語定義・略語

用語	意味
GbE	「Gigabit Ethernet」の略称。
GPU	「Graphics Processing Unit」の略称。
IOサーバ	ScaTeFSを構成するサーバ。最低 2 台必要。
IOサーバデーモン	IOサーバ上で動作するデーモン。
NIC	「Network Interface Card」の略称。 他ノードと通信するためのハードウェア。
ScaTeFS for AI	「NEC Scalable Technology File System for AI」の略称。
フェアシェアI/Oスケジューリング	ユーザ毎、またはノード毎にサーバ資源を公平に割り当てる機能のこと。
ルートIOサーバ	IOサーバの一種。mkfsを実行するサーバであり、クライアントがマウントする際のマウント先のサーバ。システム運用中においては、他のIOサーバと特に相違はなく同様な処理を行う。
仮想ファイル	仮想ファイルシステム上に作成されたファイル。ScaTeFS上のレギュラーファイル。
仮想ファイルシステム	複数のIOターゲットにより構成されるクライアント見えのファイルシステム。ScaTeFSそのもの。
実ファイル	複数のサーバに跨った仮想ファイルの断片。実際には、実ファイルシステム上のファイルのこと。
実ファイルシステム、IOターゲット	仮想ファイルシステムを構成する基本単位。各IOサーバ配下に作成される。実態は、Linuxで使用可能な通常のファイルシステム。

## 目 次

第 1 章	NEC Scalable Technology File System for AI の概要 .....	1
1.1	NEC Scalable Technology File System for AI とは .....	1
1.2	基本コンポーネント .....	2
1.2.1	クライアント .....	3
1.2.2	ネットワーク .....	3
1.2.3	IO サーバ .....	3
1.2.4	ストレージ .....	3
1.3	ScaTeFS の便利な機能 .....	3
第 2 章	IO サーバの HW/SW 構成 .....	5
2.1	IO サーバの HW 構成 .....	5
2.2	IO サーバの SW 構成 .....	5
第 3 章	クライアントの HW/SW 構成 .....	6
3.1	クライアントの HW 構成 .....	6
3.2	クライアントの SW 構成 .....	6
第 4 章	ScaTeFS for AI のライセンスについて .....	7
付録 A	発行履歴 .....	8
A.1	発行履歴一覧表 .....	8
A.2	追加・変更点詳細 .....	8

## 表目次

表 2-1	IO サーバの動作確認済み Linux ディストリビューションとソフトウェア ..	5
表 3-1	使用できる Linux ディストリビューションとカーネルバージョン .....	6
表 4-1	対象製品一覧 .....	7



## 図目次

図 1-1 ScaTeFS の分散処理のイメージ.....	1
図 1-2 ScaTeFS の構成.....	2
図 2-1 IO サーバ構成例 .....	5



## 第1章 NEC Scalable Technology File System for AI の概要

### 1.1 NEC Scalable Technology File System for AIとは

NEC Scalable Technology File System (ScaTeFS: スケートエフエス) は、複数のサーバに分散してファイルシステムの処理を行う分散・並列ファイルシステムです。データ、メタデータともに複数のIOサーバへ一様に分散配置し、read/writeやファイル、ディレクトリの操作のリクエストを各IOサーバへ分散して処理します。これにより、高いスループットを発揮できるので、多数の計算ノードで並列に大量のIOを行いながら同時に解析処理を行うAIシステムに活用できます。

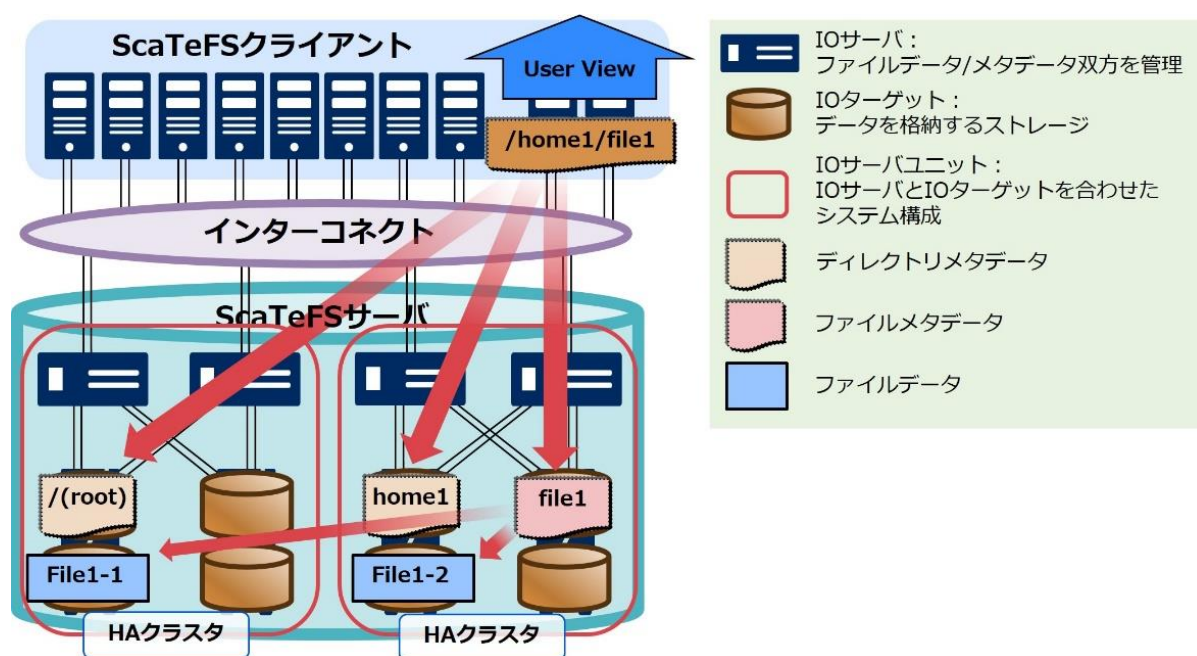


図 1-1 ScaTeFS の分散処理のイメージ

- **高スケーラビリティ**  
分散配置、負荷分散の仕組みによりPBクラスの大容量、高スループットを実現します。
- **高性能**  
同一ファイルでも内部的に分散されたデータに並列にIOすることができ、高いIO性能を実現し、メタデータキャッシュ機能により高いメタデータ性能を実現します。
- **データ/メタデータの高信頼性**  
未コミットデータの再送によるデータの保証や、オンメモリジャーナル機能によるメタデータの保証により高い信頼性を実現します。

- 高可用性

IOサーバの障害時やクライアントとIOサーバ間のパス障害時でも運用継続を可能とする高い可用性を実現します。

- IOサーバ障害時、ジャーナリングによるデータ保全のできるIOサーバのフェイルオーバー機能
- IOサーバとストレージ間のパス障害時のためのパスフェイルオーバー機能
- IOサーバのネットワークインターフェース障害への対処

## 1.2 基本コンポーネント

ScaTeFSは、図 1-2のように大きく分けて下記の4つより成っています。

- クライアント (計算ノード)
- ネットワーク
- IOサーバ
- ストレージ

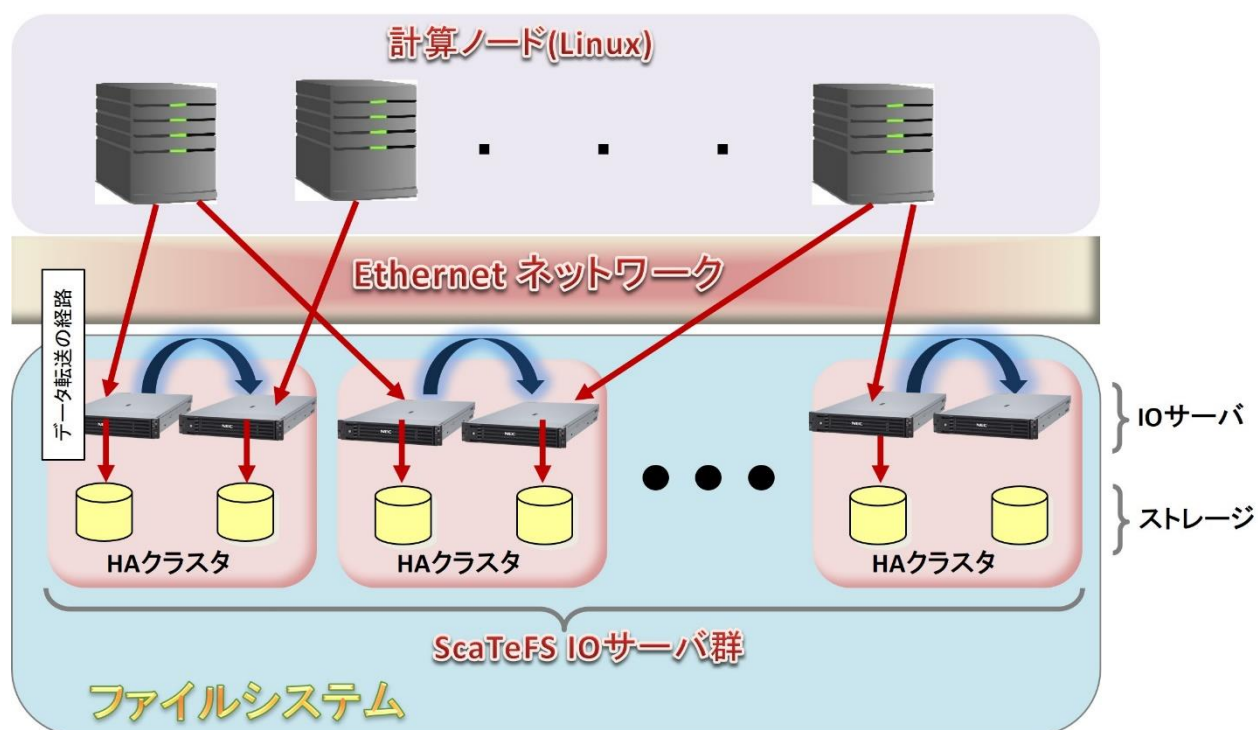


図 1-2 ScaTeFS の構成

### 1.2.1 クライアント

Linuxマシンの計算ノード。GPUを搭載したLinuxマシンを使用することもできます。

### 1.2.2 ネットワーク

管理やメンテナンスのためのネットワーク負荷がファイルシステムの性能に影響しないように、下記の別々のネットワークを用意してください。

- ファイルシステム用ネットワーク  
クライアントとIOサーバ間でデータやメタデータの転送を行うEthernetのネットワーク
- 運用・管理用ネットワーク  
メンテナンスのためクライアントやIOサーバにログインする場合や、ntpによる時刻同期等で使用するEthernetのネットワーク。ファイルシステム用ネットワークより帯域が狭いネットワークでも問題ありません。

### 1.2.3 IO サーバ

クライアントからのリクエストに基づき、配下に接続されたストレージに格納されるファイルデータの断片、メタデータを操作します。なお、IOサーバの障害時でもフェイルオーバーにより運用を継続できるよう2台のIOサーバによりHAクラスタを構成します。

### 1.2.4 ストレージ

各IOサーバ配下に接続され、ファイルデータの断片、メタデータを格納します。

## 1.3 ScaTeFSの便利な機能

#### (1) 運用、システム構築を容易にする機能

- 運用を停めないIOサーバのメンテナンスが可能
- 1つのコマンド実行のみにより、複数のIOサーバに跨るファイルシステムを構築可能
- ファイルシステムの整合性チェック、修復機能
- ログ、統計情報の収集機能
- 性能、容量アップのためのIOサーバ、ストレージの追加機能
- 既存ファイルのデータの配置を変更して、IOサーバ間、IOターゲット間のストレージ使用量の不均衡を解消する機能

(2) 多様な環境、用途に対応するための機能

- フェアシェアI/Oスケジューリングによる多数ユーザの利用環境におけるフェアなI/O処理
- Linuxクライアント上のNFSサーバを使って、ファイルシステムをNFSクライアントへエクスポートすることが可能
- Linuxクライアント上にSambaサーバを構築して外部公開することにより、Windowsからのアクセスが可能

## 第2章 IO サーバの HW/SW 構成

## 2.1 IOサーバのHW構成

IOサーバは、2台のExpress5800/R120j-2M でHAクラスタのActive-Activeを構成します。  
以下にIOサーバ構成を記載します。

IOサーバとストレージは2port 16G/2port 32G-FC HBA x 2枚で直接接続します。

IOサーバとScaTeFSクライアントはネットワークスイッチを介して接続します。

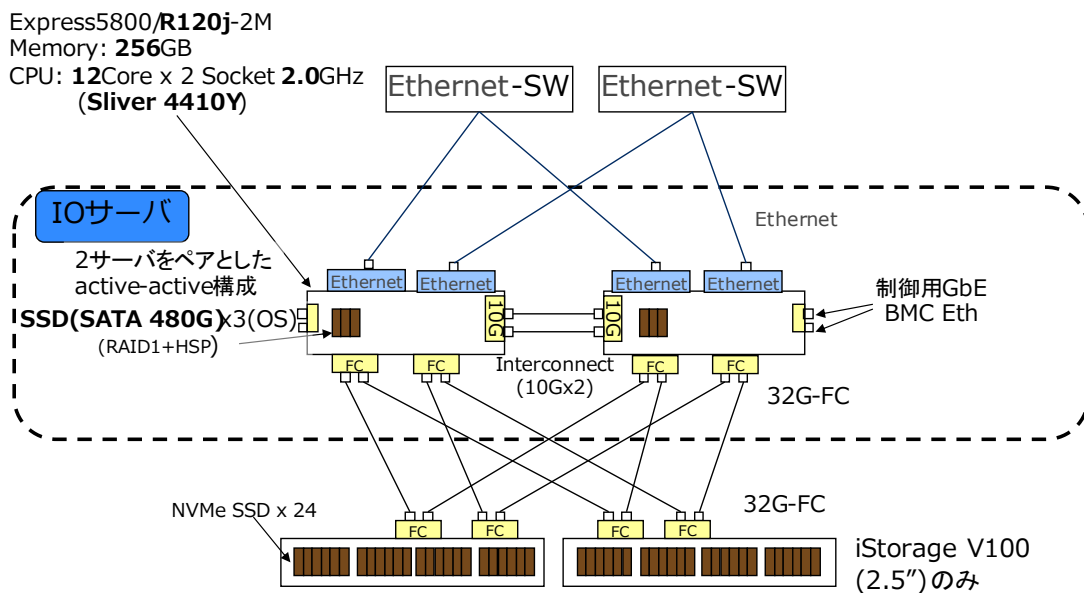


図 2-1 IOサーバ構成例

## 2.2 IOサーバのSW構成

IOサーバはRed Hat Enterprise Linux (x86\_64) 上で動作します。以下にIOサーバの動作確認済みのLinuxディストリビューション、ソフトウェアを記載します。

表 2-1 IO サーバの動作確認済み Linux ディストリビューションとソフトウェア

項目	動作確認済みバージョン	用途
Linuxディストリビューション	Red Hat Enterprise Linux 8.6 (4.18.0-372.32.1.el8_6.x86_64)	
HA Dynamic Link Manager	8.8.6-00	ストレージシステムとIOサーバを結ぶ経路を管理
CLUSTERPRO X for Linux	5.1.1-1	HAクラスタを管理

## 第3章 クライアントの HW/SW 構成

### 3.1 クライアントのHW構成

クライアントとして使用できるマシンは、x86-64アーキテクチャをサポートしているLinuxマシンになります。

使用できるネットワークインターフェースは、イーサネットのNICになります。

### 3.2 クライアントのSW構成

使用できるLinuxディストリビューションとカーネルバージョンは、以下のとおりです。

表 3-1 使用できる Linux ディストリビューションとカーネルバージョン

Linuxディストリビューション	カーネルバージョン
RHEL 8.6	4.18.0-372.32.1.el8_6.x86_64



## 第4章 ScaTeFS for AI のライセンスについて

ScaTeFS for AI のライセンス製品は、以下の通りです。ご利用になるファイルシステムの容量にあわせて製品を組み合わせることでお選びいただけます。

表 4-1 対象製品一覧

製品	型番	説明
ScaTeFS for AI 10TB ライセンス	UFSH41N1XA10-I	10TB の容量ライセンス
ScaTeFS for AI 50TB ライセンス	UFSH41N1XB10-I	50TB の容量ライセンス
ScaTeFS for AI 100TB ライセンス	UFSH41N1XC10-I	100TB の容量ライセンス
ScaTeFS for AI 250TB ライセンス	UFSH41N1XD10-I	250TB の容量ライセンス
ScaTeFS for AI 500TB ライセンス	UFSH41N1XE10-I	500TB の容量ライセンス

ScaTeFS for AI のライセンスファイルの取得方法やトラブルシューティング等の詳細については、別冊のマニュアル『NEC Scalable Technology File System for AI (ScaTeFS for AI) ソフトウェアライセンス管理説明書』をご参照ください。

## 付録 A 発行履歴

### A.1 発行履歴一覧表

2024 年 1 月 初版

### A.2 追加・変更点詳細

- 初版  
新規作成

# NEC Scalable Technology File System for AI

(ScaTeFS for AI)

## システムガイド

2024年 1月 初版

日本電気株式会社

東京都港区芝五丁目 7 番 1 号

TEL(03)3454-1111 (大代表)

© NEC Corporation 2024

日本電気株式会社の許可なく複製・改変などを行うことはできません。

本書の内容に関しては将来予告なしに変更することがあります。