



# データ レゼボワール

## Data Reservoir

---

**(株)富士通コンピュータテクノロジーズ**

# 背景



理学研究の最先端では巨大な実験観測データが発生し、これらを効率よくデータ解析処理装置(スーパーコンピュータ等)に結びつけることが必要となっております。

## 某研究所様の例

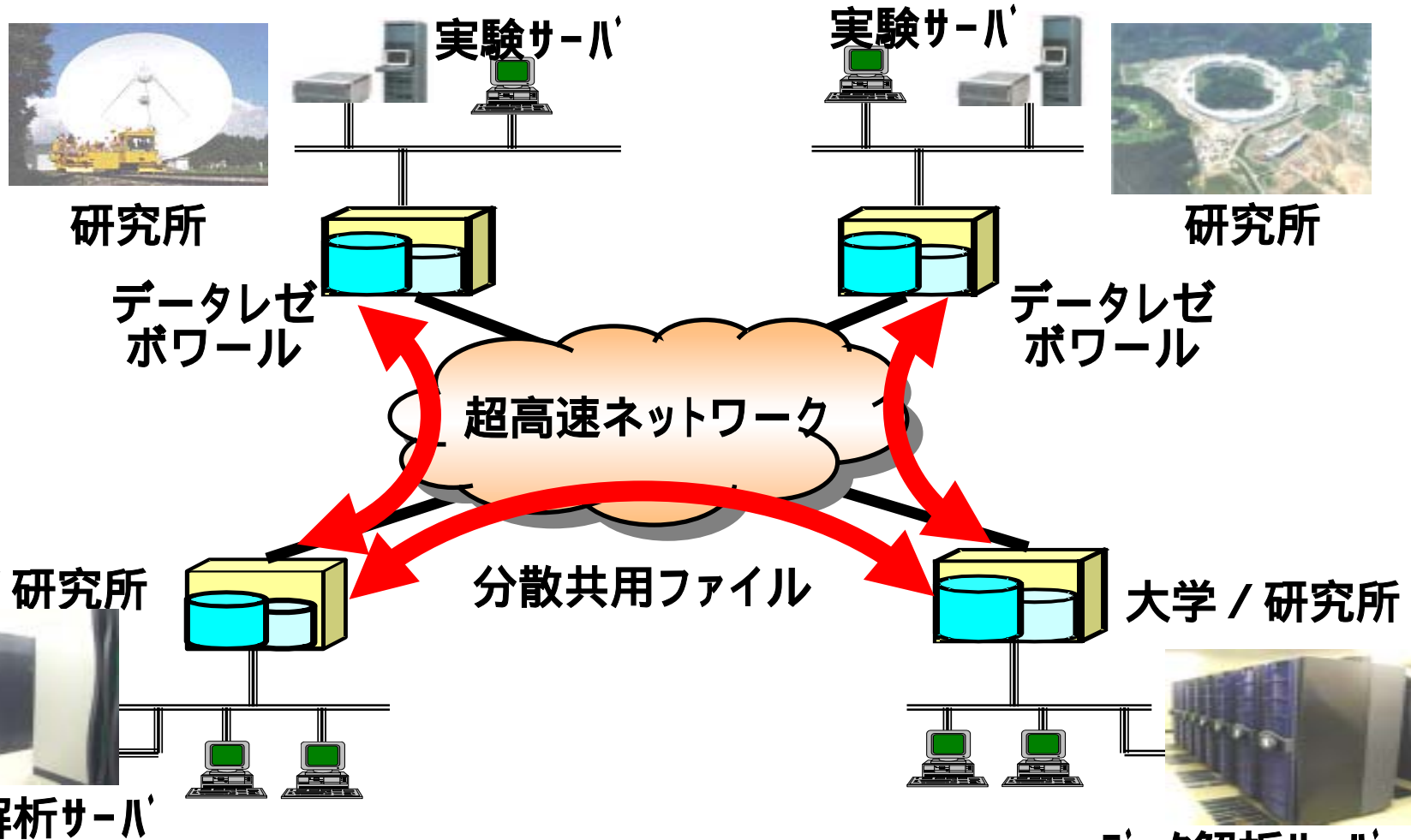
ファイル容量 = 数10TB ~ 100TB  
1日に数10GB ~ 1TB位のデータが生成

現在は磁気テープ等によりデータの送受を行っておりますが、超高速ネットワークを介して上記データを共有することが求められております。

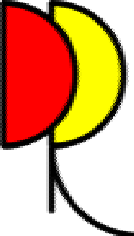
本システムは(東京大学)平木敬教授他のご指導のもとで開発したものです。

# システムの狙い

データレゼポワールは巨大なデータを、超高速ネットワークを用いて、大域でデータ共用可能とするシステムです



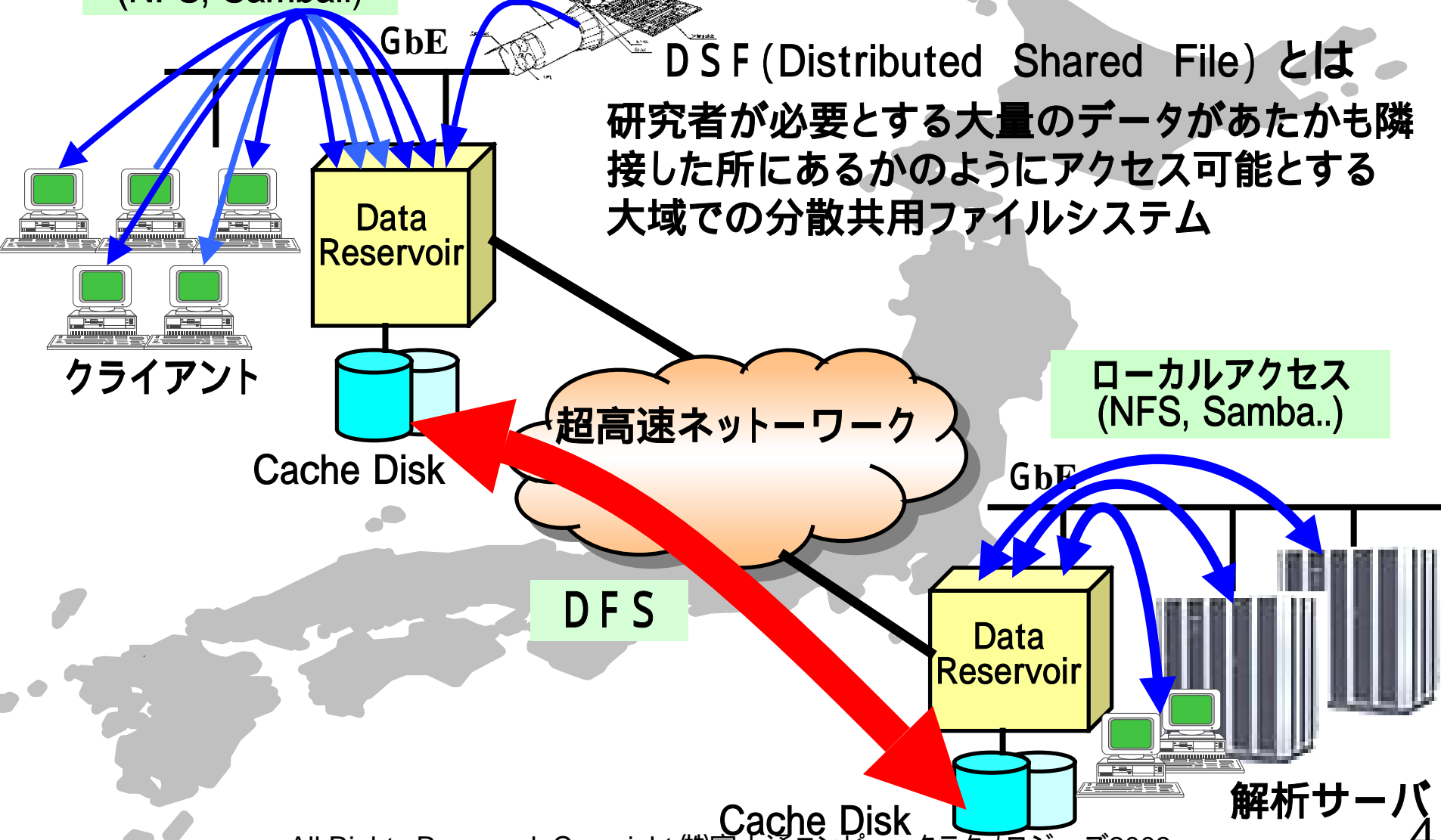
# 分散共用ファイルシステム(DFS)



ローカルアクセス  
(NFS, Samba..)

GbE

DSF (Distributed Shared File) とは  
研究者が必要とする大量のデータがあたかも隣  
接した所にあるかのようにアクセス可能とする  
大域での分散共用ファイルシステム



# 技術課題と解決策



- **巨大データの高速アクセス**

  - 多数ディスクに均等分散

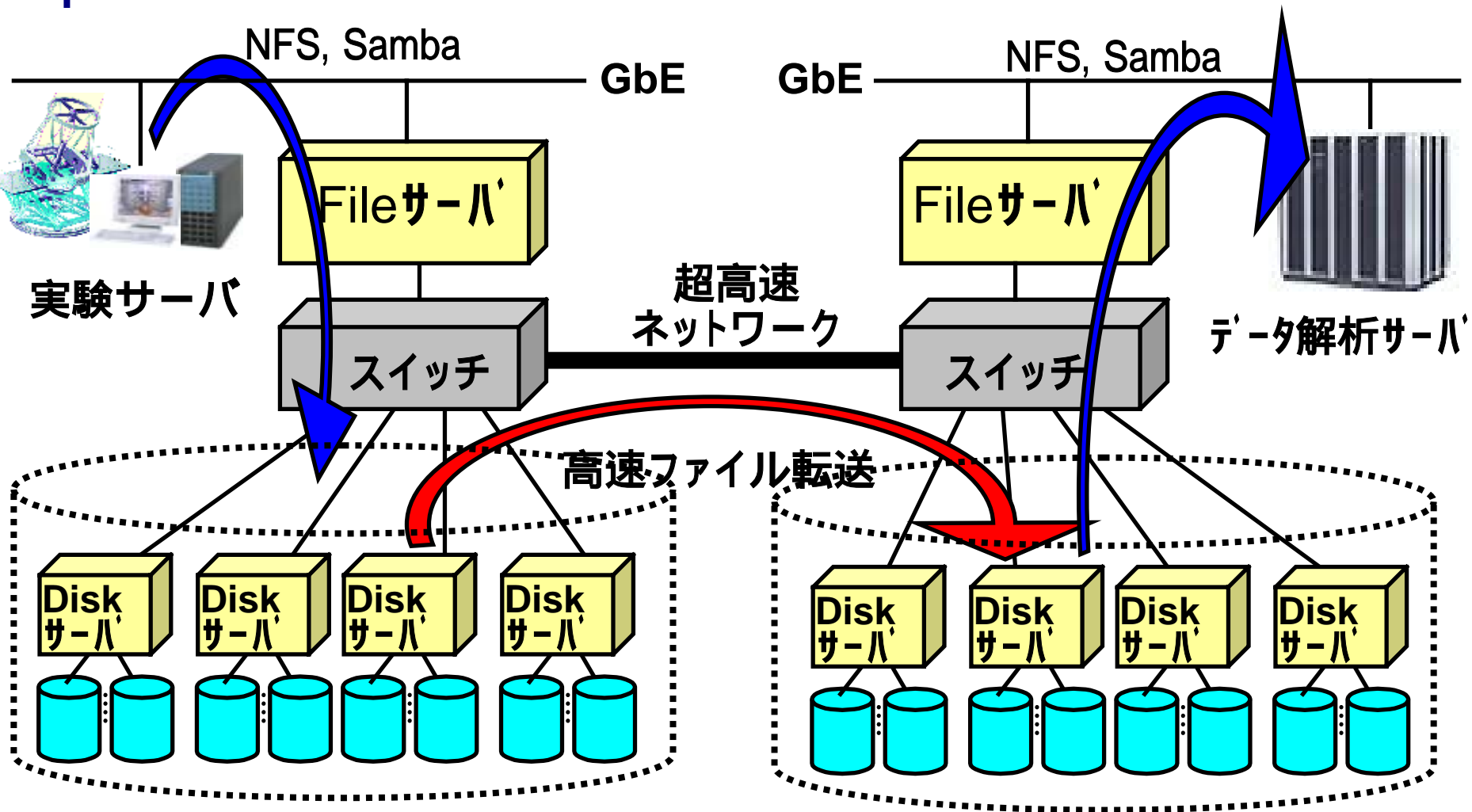
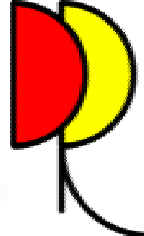
  - ディスクのRawCopy転送

  - 有効データの選択的転送

- **ネットワーク性能にスケーラブルな転送**

  - i-SCSI プロトコルを用いたディスクの並列転送

# システム構成概要

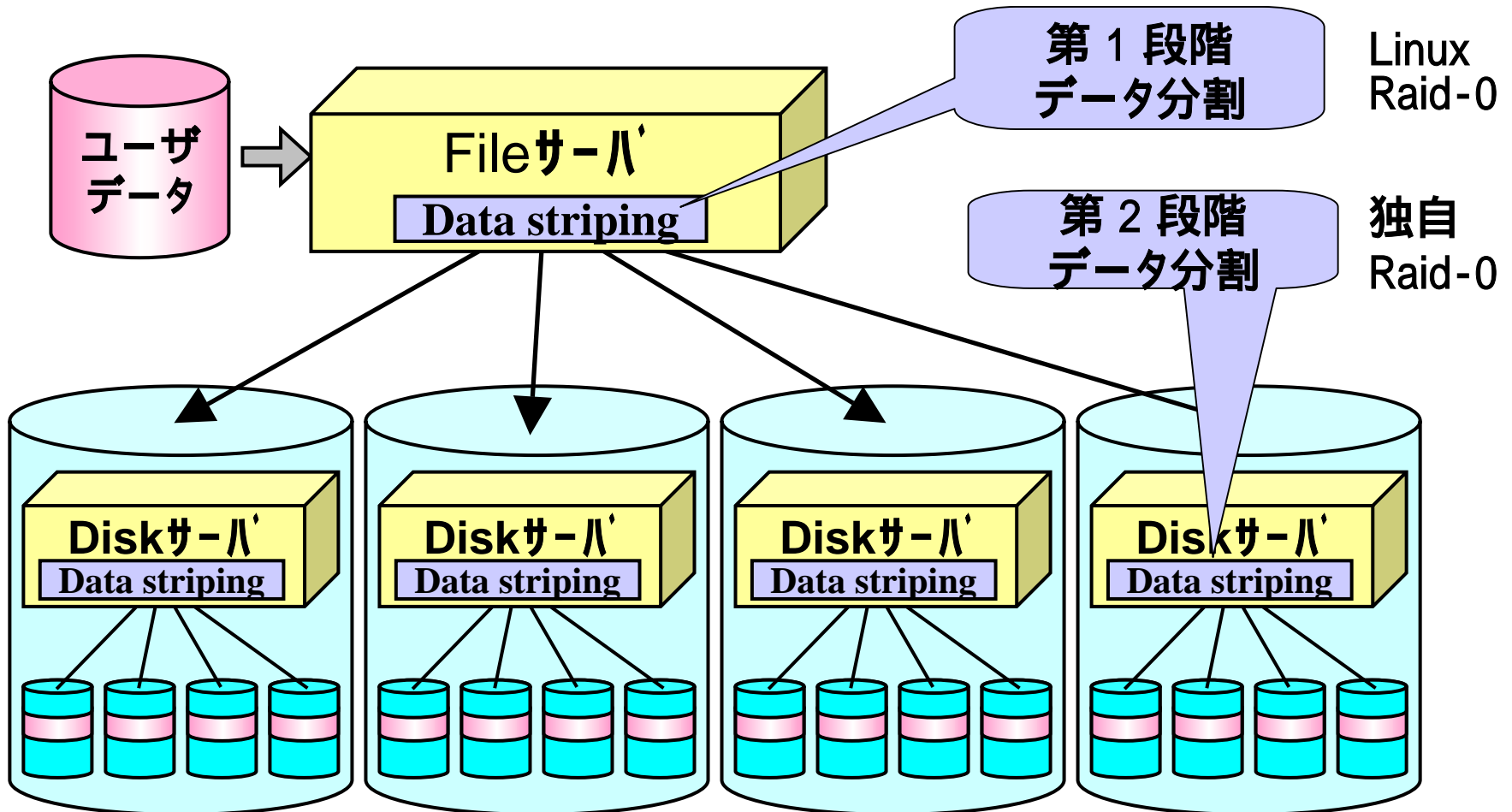


Fileサーバ、DISKサーバはIAサーバ(Linux)を採用

# 巨大データの高速アクセス(1)

## - 多段階層データストライピング -

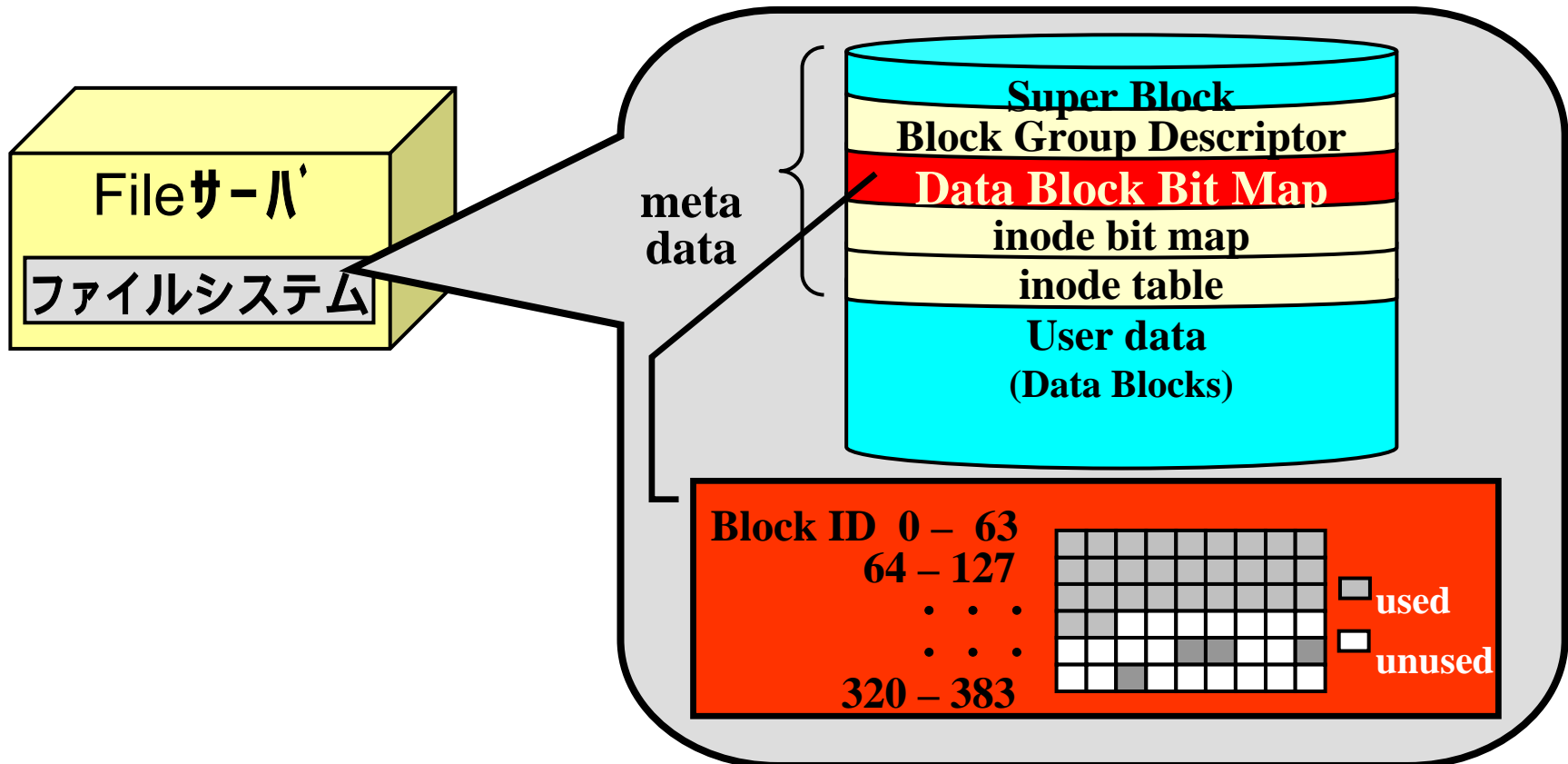
- 大容量データを多数ディスクに均等分割



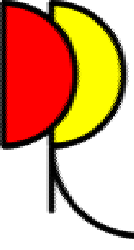
# 巨大データの高速アクセス(2)

## - 有効データのみを選択転送 -

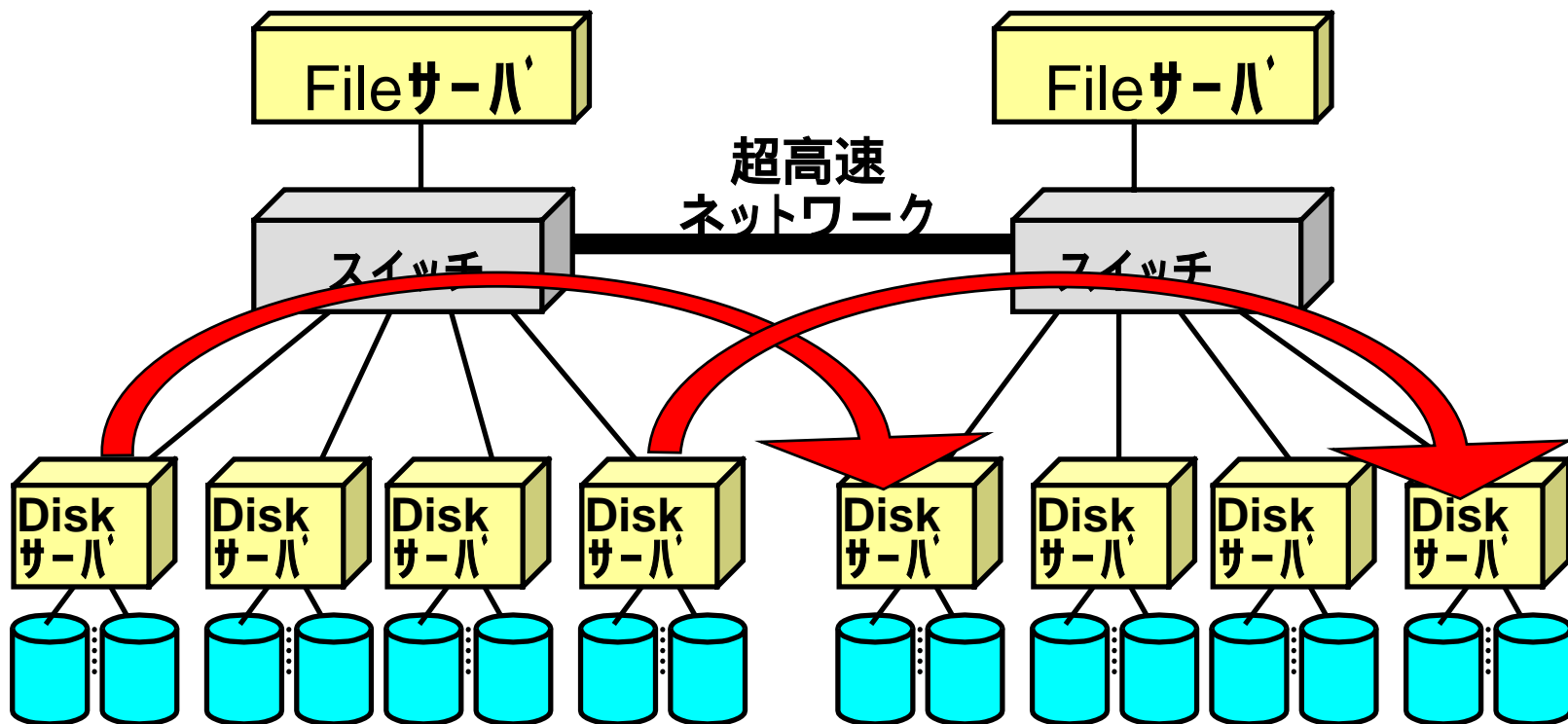
- Linuxファイルシステムのメタ情報からLBA単位にディスクの有効データを検索



# 巨大データの高速アクセス(3)



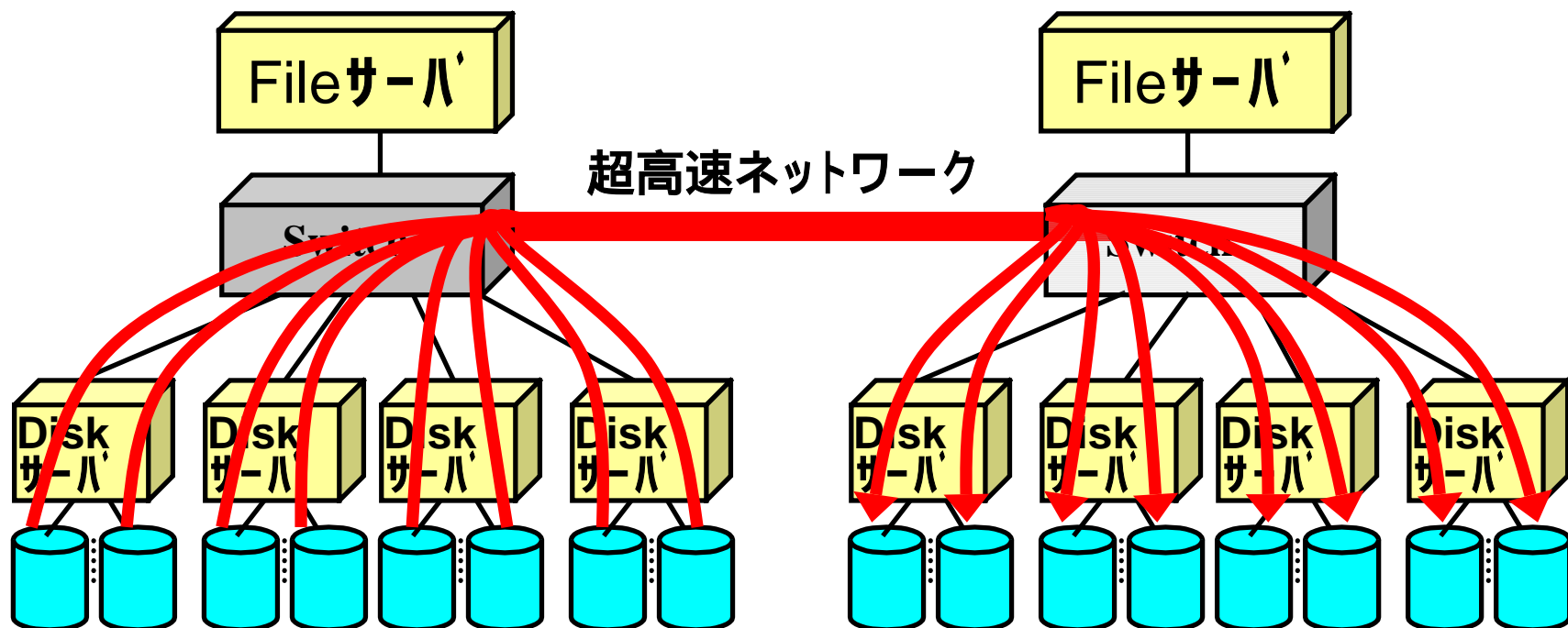
- ファイルシステムを経由しないデータ転送
- Raw Copyによるディスクの最大RD/WR性能の確保
- ディスク内の有効データのみをLBA単位で転送



# ネットワーク性能にスケーラブルな転送

- IPネットワークを介しディスク装置間で並列転送 -

- ネットワークの能力に応じスケーラブルなシステムを構築
- コマンドキューイング(i-SCSI Queue:最大256個)によるトワーク遅延の補償

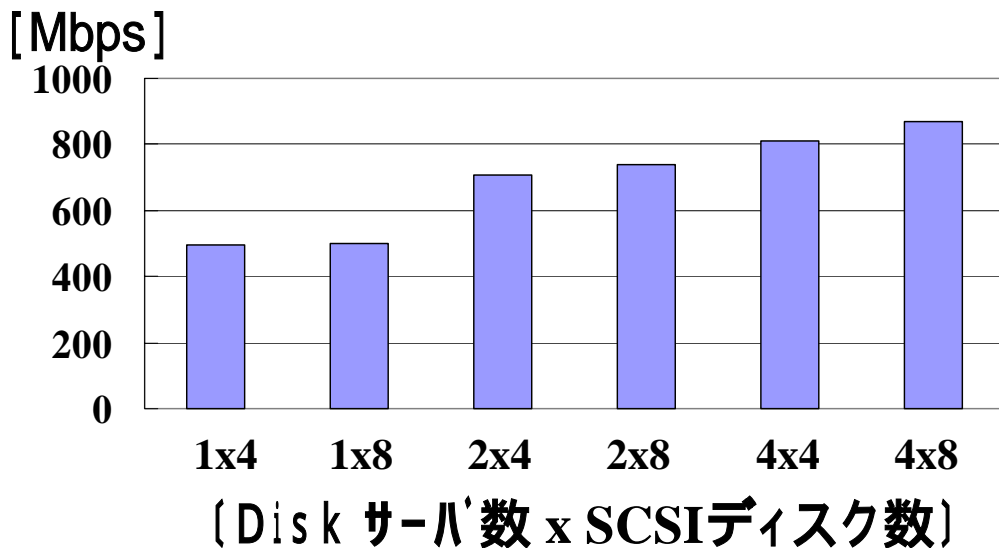


# 1 Gbpsモデル長距離実験

## ● 本州 1 周 1600km 実験

[ Super-SINET 帯域: 1 Gbps ネットワーク遅延(往復): 36ms ]

転送性能

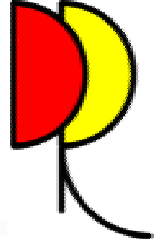


Diskサーバ4台で1 Gbps帯域の  
ネットワークをフルに活用可能



東京大学 京都大学  
大阪大学 東北大学  
東京大学, 延べ 1600 Km.

# まとめ



- データレゼボワールはテラバイト級の巨大データを、1 Gbps帯域のネットワークを用いた長距離 (~ 1600 Km) 伝送で十分な高速転送能力持つことを実証。