



みかん – ミラー選択機能付き 代理FTPサーバ

尾藤 正人(HDE)
舟阪淳一(広島市立大学)



はじめに

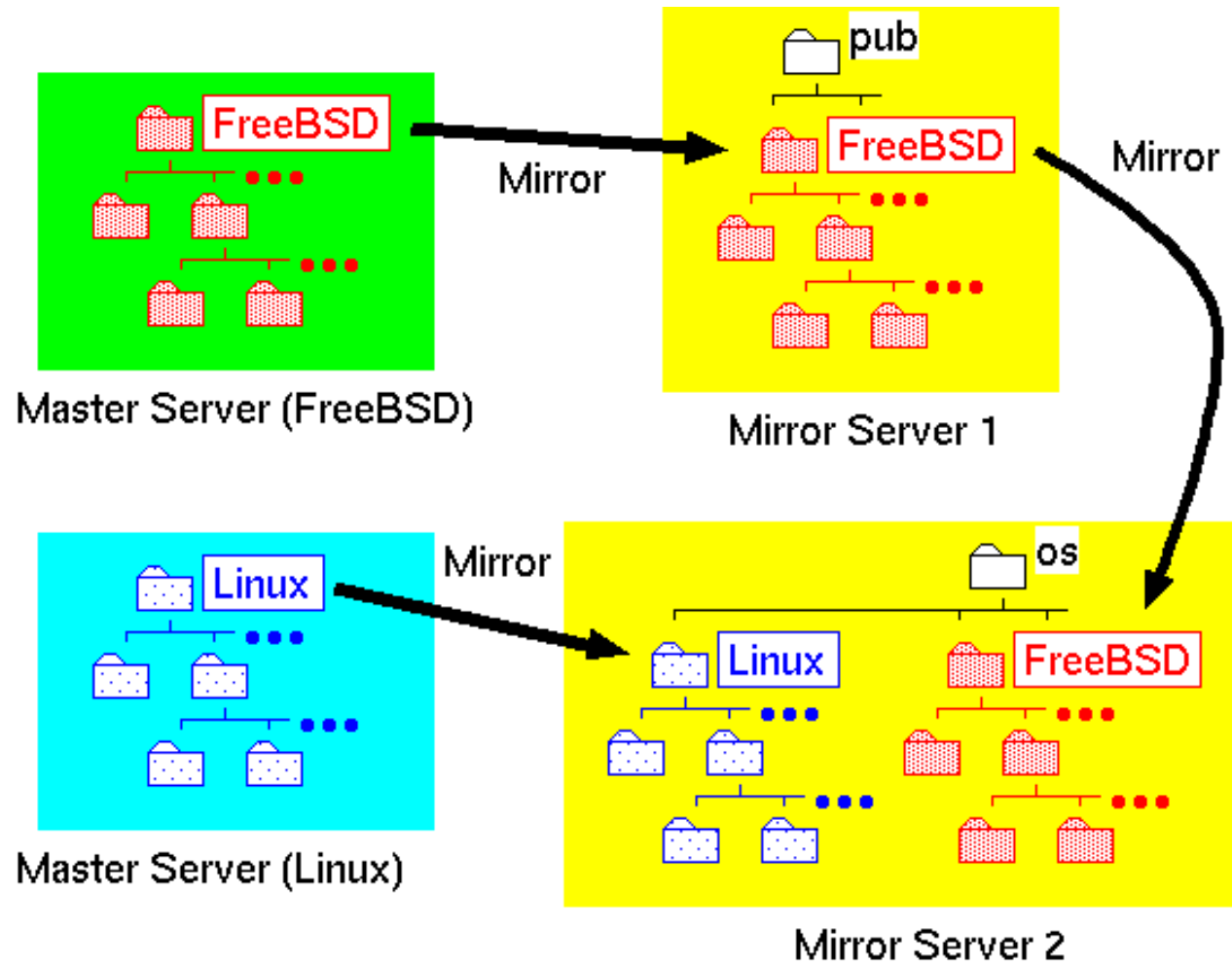
1. 研究背景
2. 関連研究
3. みかんの提案
4. みかんの実装
5. 評価および考察
6. まとめ



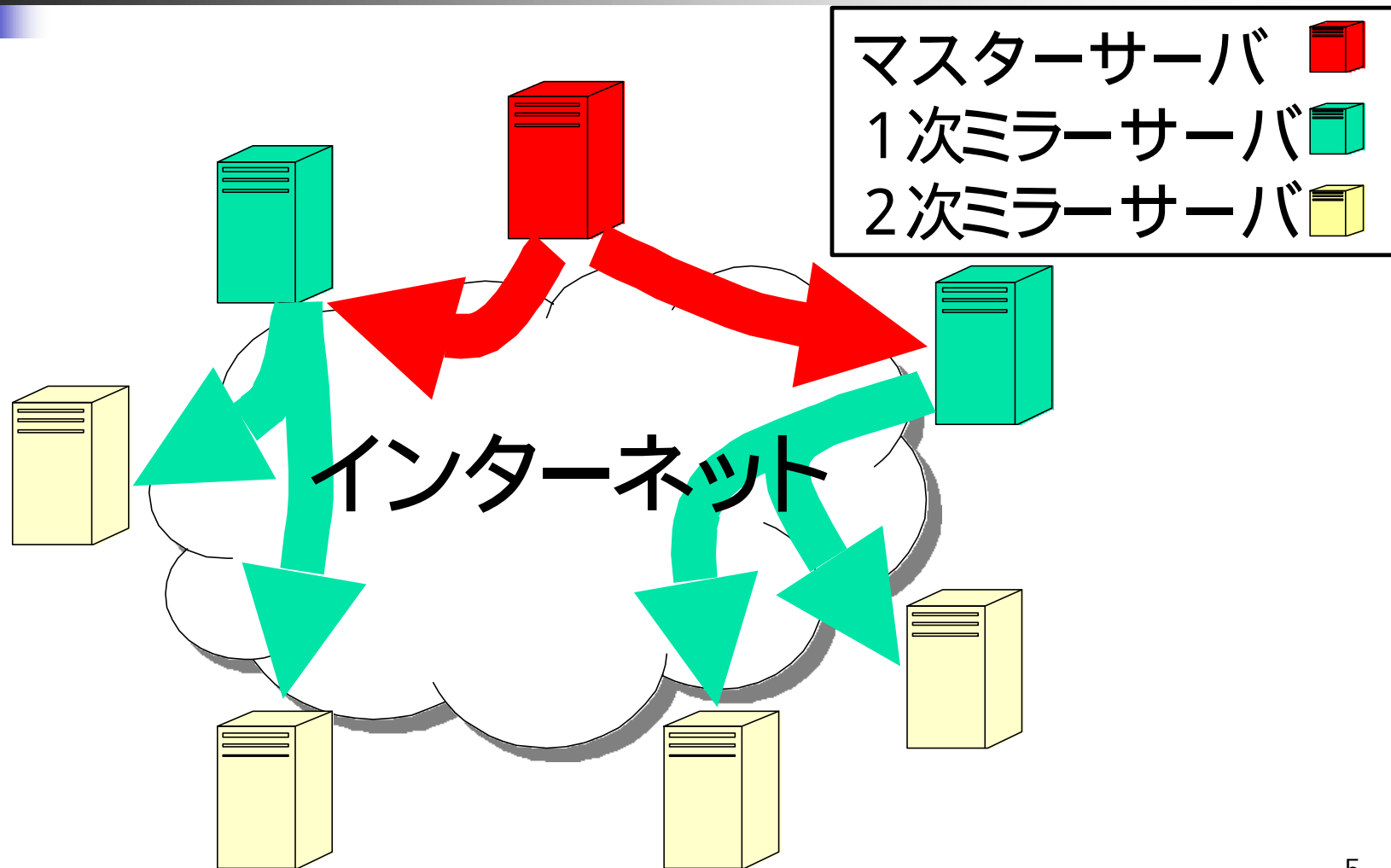
FTPサーバの利用形態

- FTP
 - インターネット上でファイルのやりとりを行うためのプロトコル(RFC 959)
- 匿名FTPサーバによるファイルの配布が一般的
- ミラーサーバの利用
 - マスターサーバと同一ファイルを提供

ミラーサーバとは



マスターサーバとミラーサーバ の関係





ミラーサーバ利用に関する問題点1

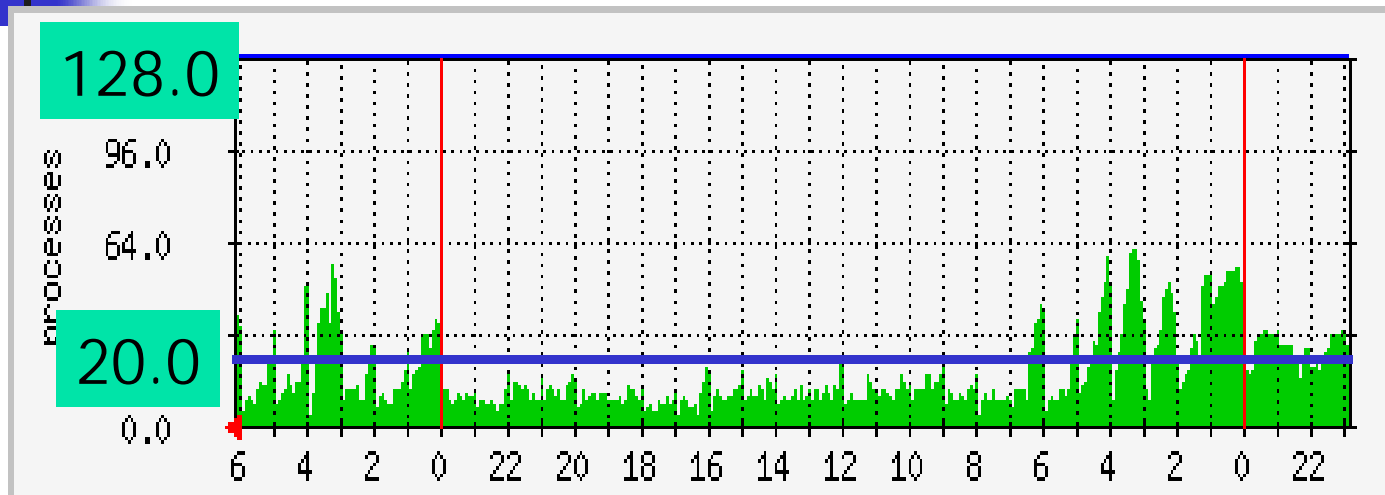
- ミラーサーバの選択時に
 - どのようなミラーサーバがあるのか
 - 何をミラーしているのか
 - どこにミラーされているのか
 - ネットワーク的に近いのはどれか
 - ミラーサーバの更新が遅れていないかが不明



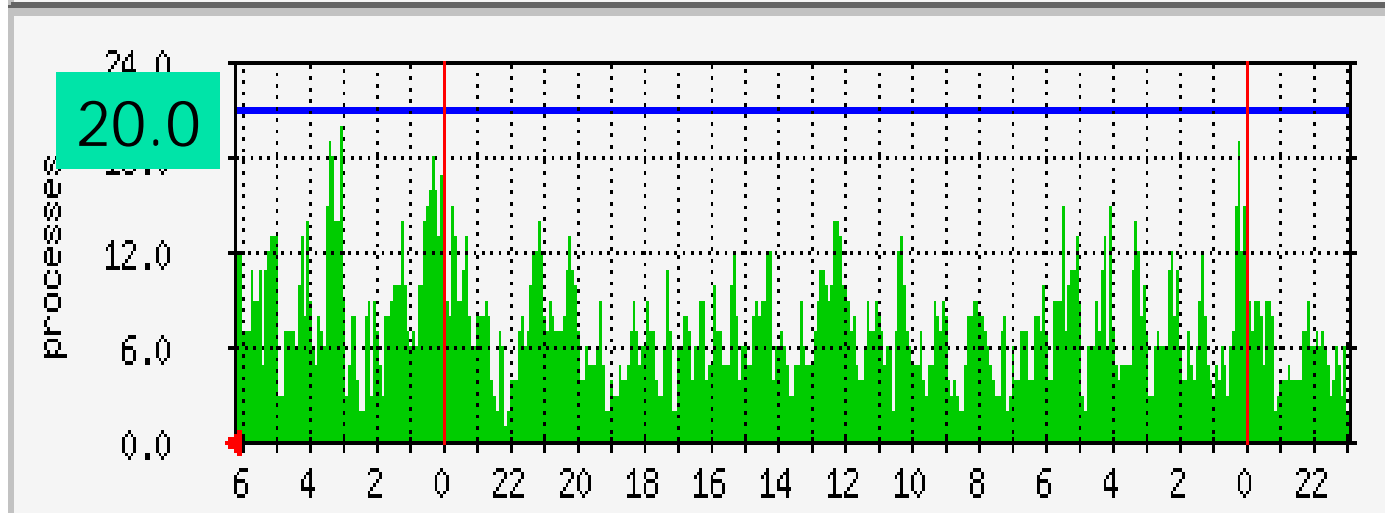
ミラーサーバ利用に関する問題点2

- ミラーサーバ選択のために多くの情報が必要
- ミラーサーバの選択が困難
- マスターサーバに負荷が集中

CVSupサーバの利用統計 (2002/2/8)



マスター
サーバ

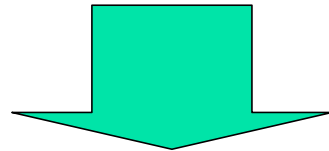


ミラー
サーバ

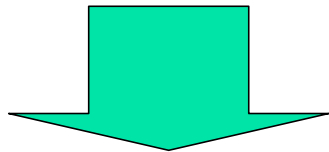


研究背景のまとめ

ミラーサーバを利用するためには、
多くの情報と複雑な手順が必要



ミラーサーバが有効活用されていない



ミラーサーバを自動選択し、最新ファイルを取得する技術が必要



関連研究

- TENBIN
 - DNSを用いた負荷分散
 - 各サーバが同一内容を保持
- FTP Mirror Tracker
 - ディレクトリ内容のMD5・・・識別子
 - 同一内容のディレクトリ候補を提示
 - ディレクトリ内に不整合があると効果小



更新状況調査の方法

- Vine Linuxのディレクトリツリーを調査
- 無作為に抽出した11台のサーバからファイル一覧を取得し、比較
- 比較した結果を以下の3つに分類
 - 同一
 - 差異あり
 - ミラーなし



ミラーリングポリシー

- 次のようなファイルはミラーしない場合がある
 - 先頭に“.”のつくファイル
 - ファイルの最後に“~”のつくファイル
- ミラーに関するポリシーを本研究では**ミラーリングポリシー**と呼ぶ



更新状況の調査結果

11のサーバをマスターとの同期状況により分類
(ftp.jaist.ac.jp)

ディレクトリ	同一	差異あり	ミラーなし
Vine-2.5	7	1	3
VinePlus	4	6	3
VineSeed	2	8	1

(Vine Linuxについて2002年6月3日の調査結果)



本研究の目的

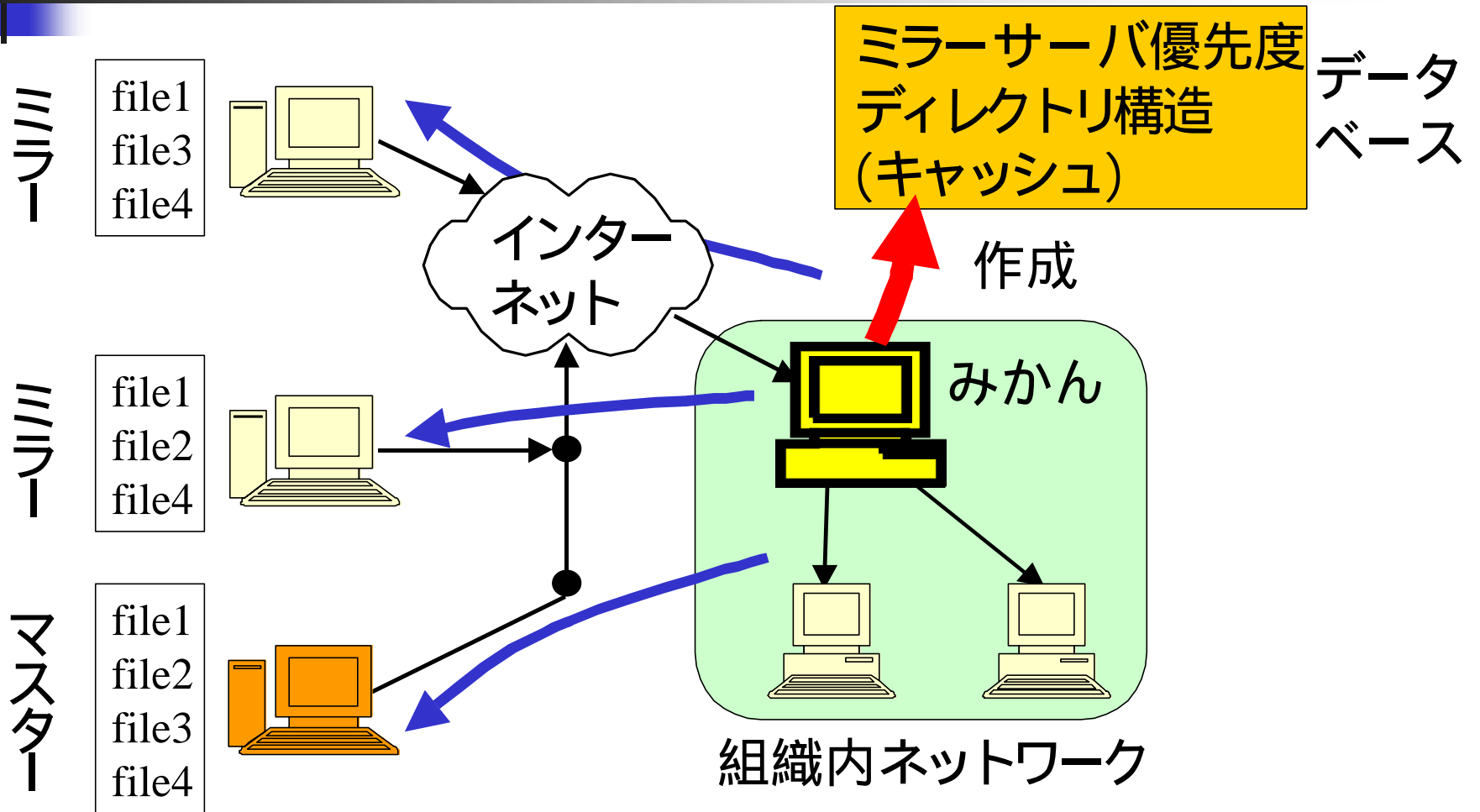
- ミラーサーバの選択を自動的に行う
- サービスの不均一性を吸収する
- 取得ファイルの最新性を保証する
- アクセスを分散させる
- 従来クライアントとの互換性を維持する



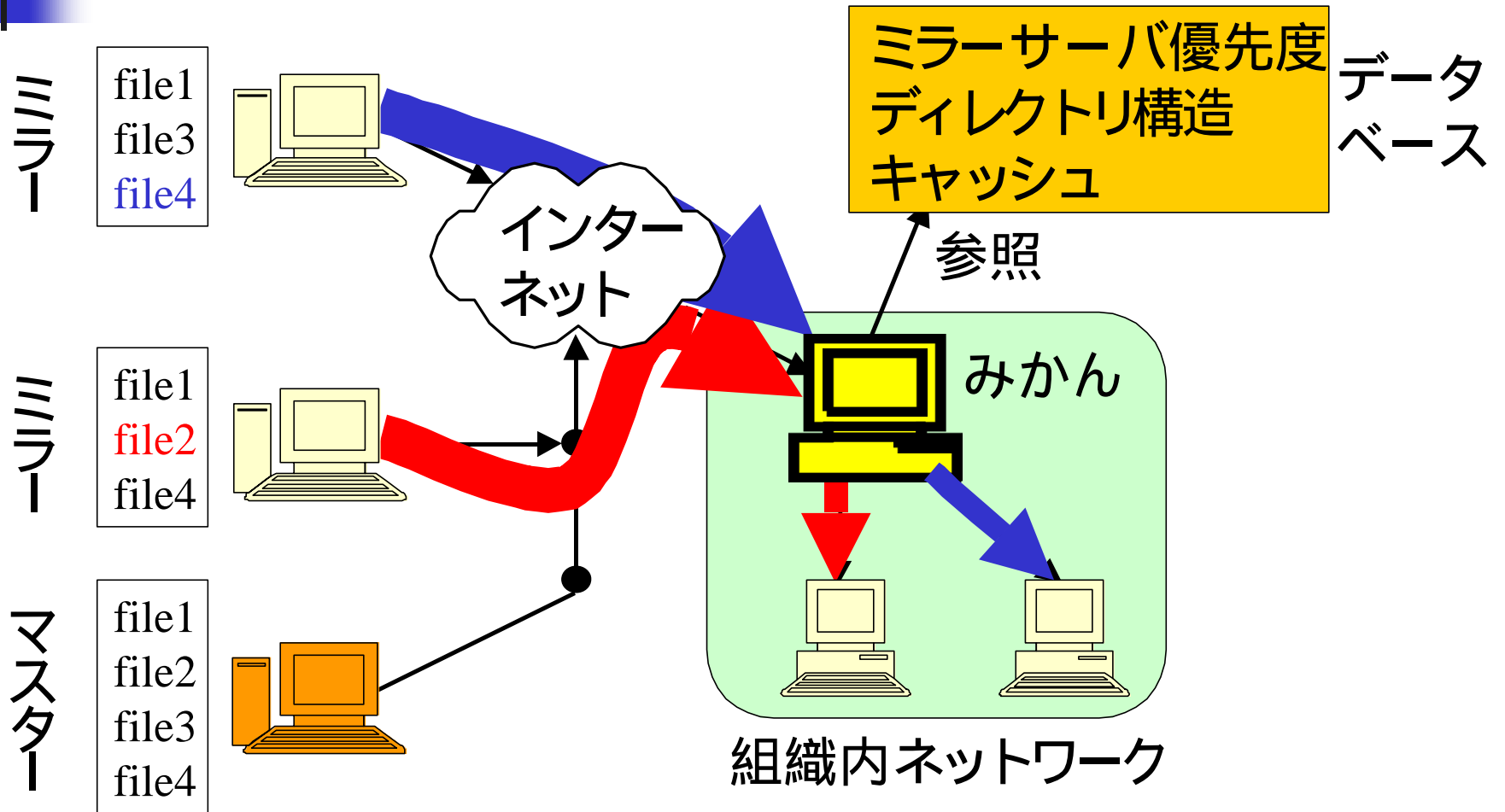
みかん – ミラー選択機能付き 代理FTPサーバの提案

- 中継サーバ
 - 従来クライアントで接続可能
 - 仮想的なミラーサーバ
 - キャッシュ機能
- 更新状況・ネットワーク状況を常時監視
- ファイル単位でサーバ選択

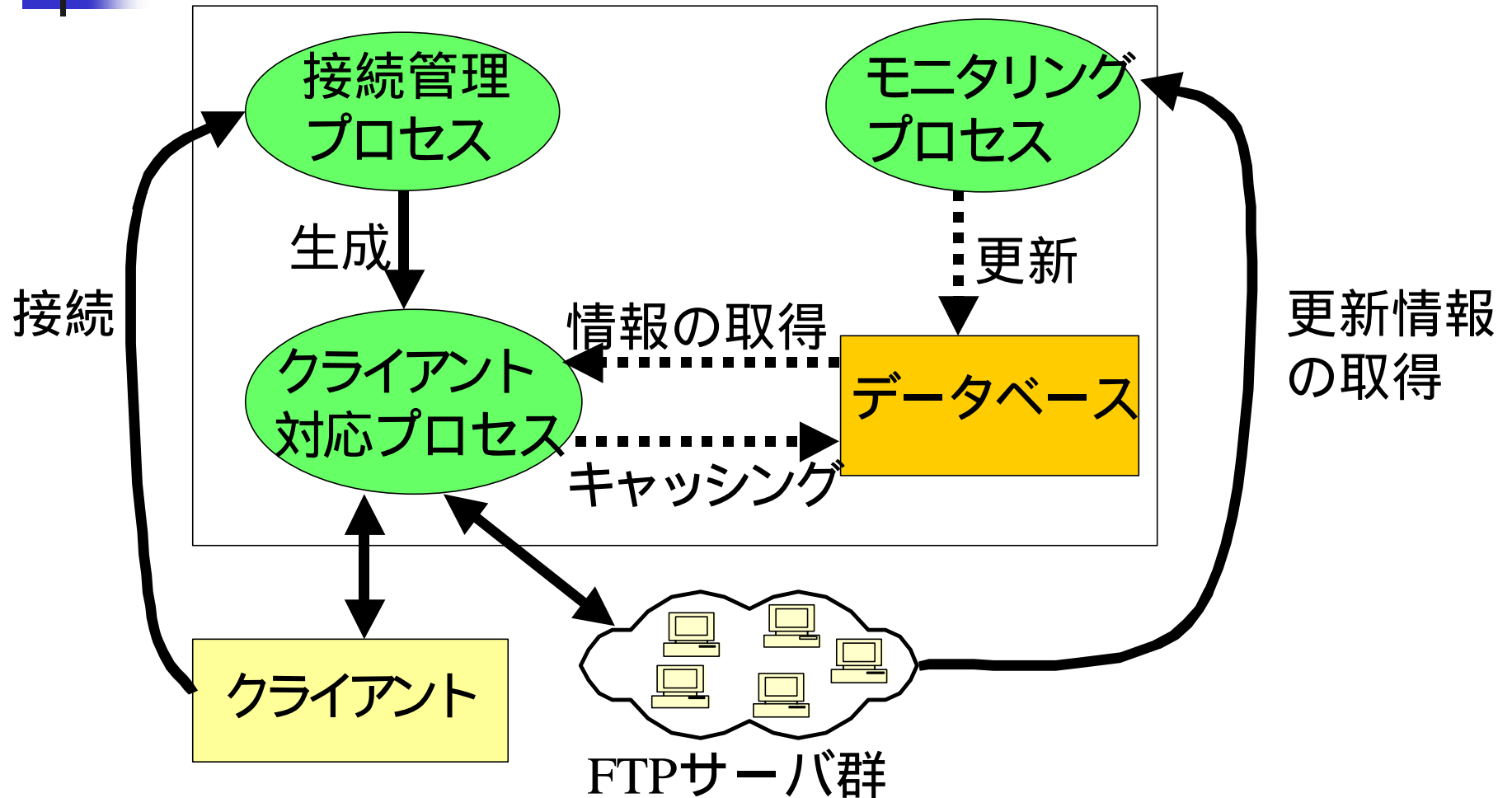
みかんの動作(1)



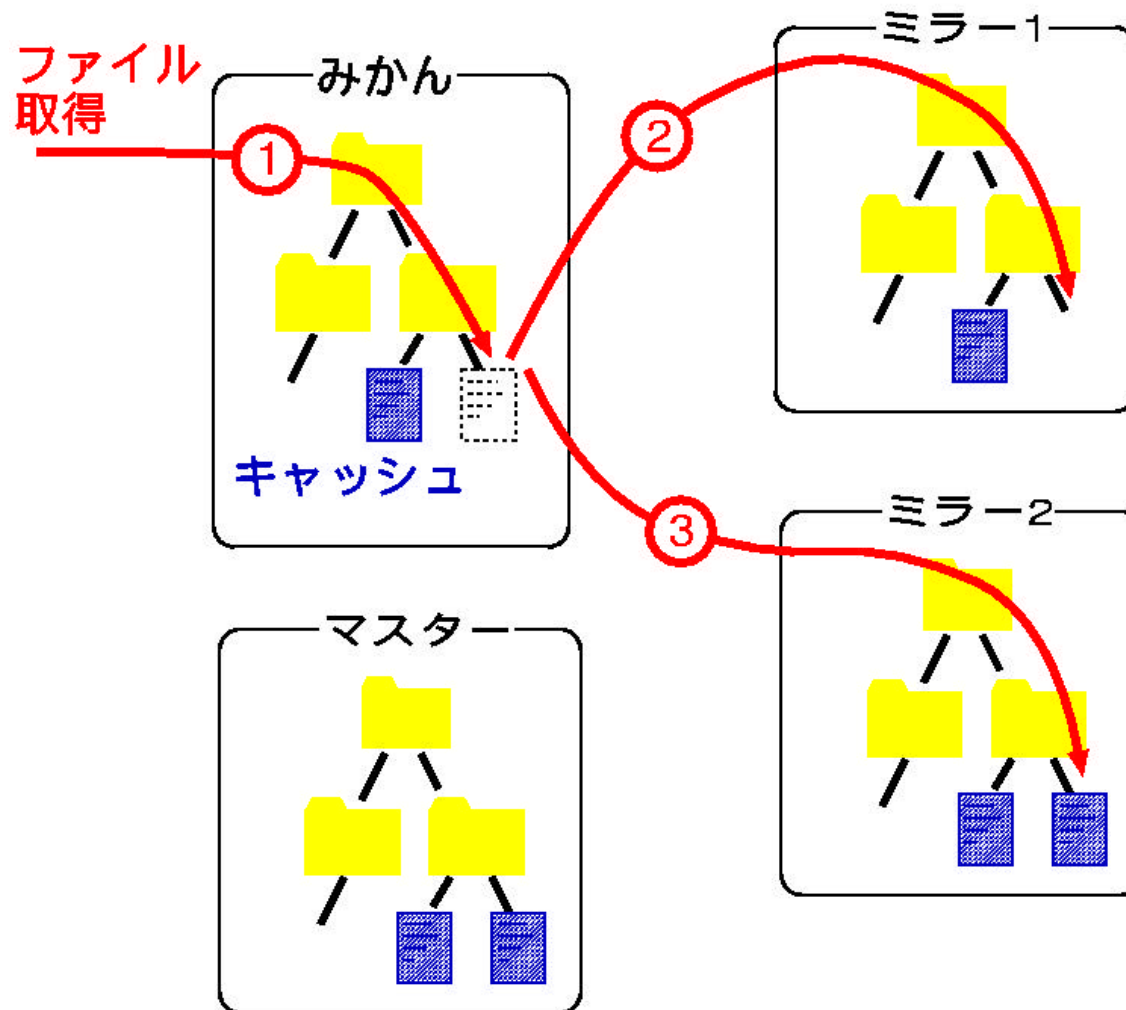
みかんの動作(2)



みかんの実装体系



みかんのファイル取得方法



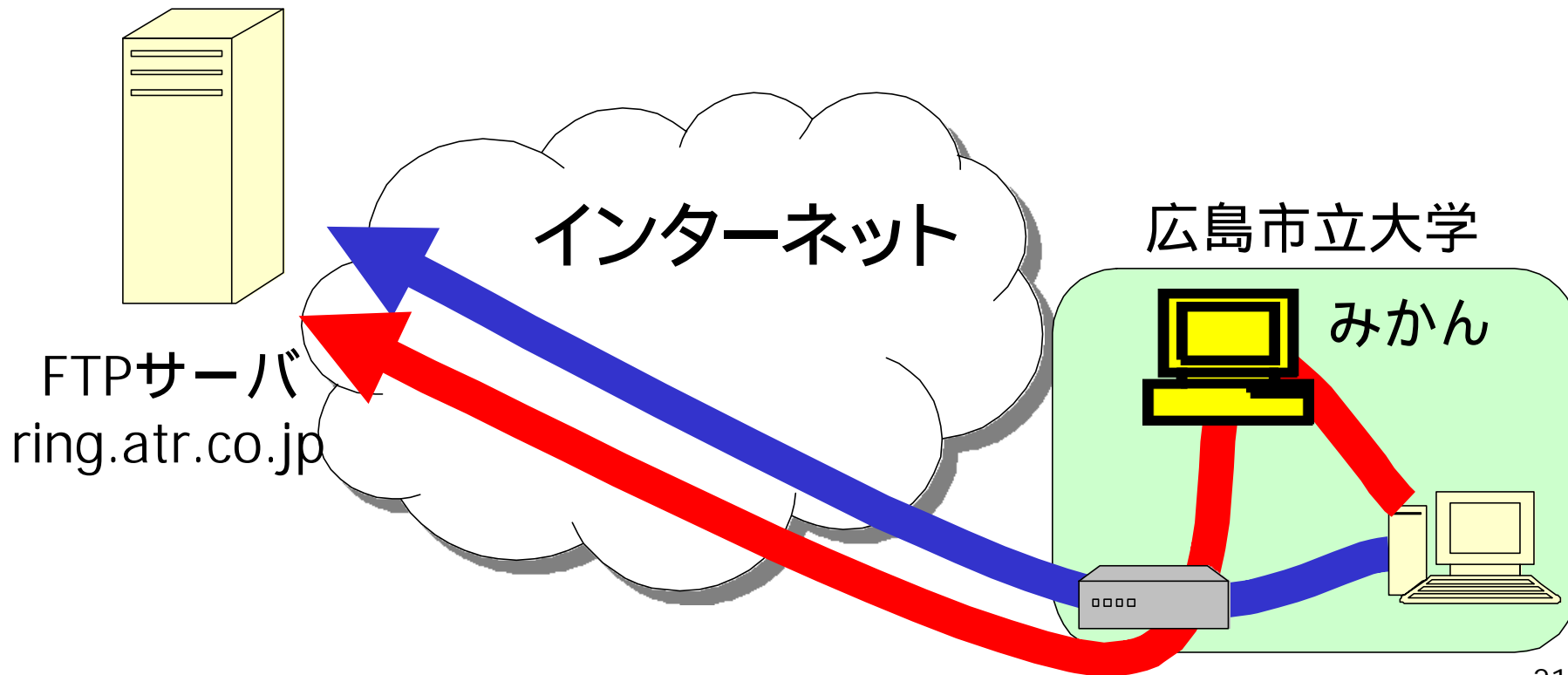


みかんの評価

- オーバーヘッドの評価
- 可用性の評価
- トラフィックの評価
- FTP Mirror Trackerとの比較評価

オーバーヘッドの評価

- みかんを経由することで、加算される処理時間を評価

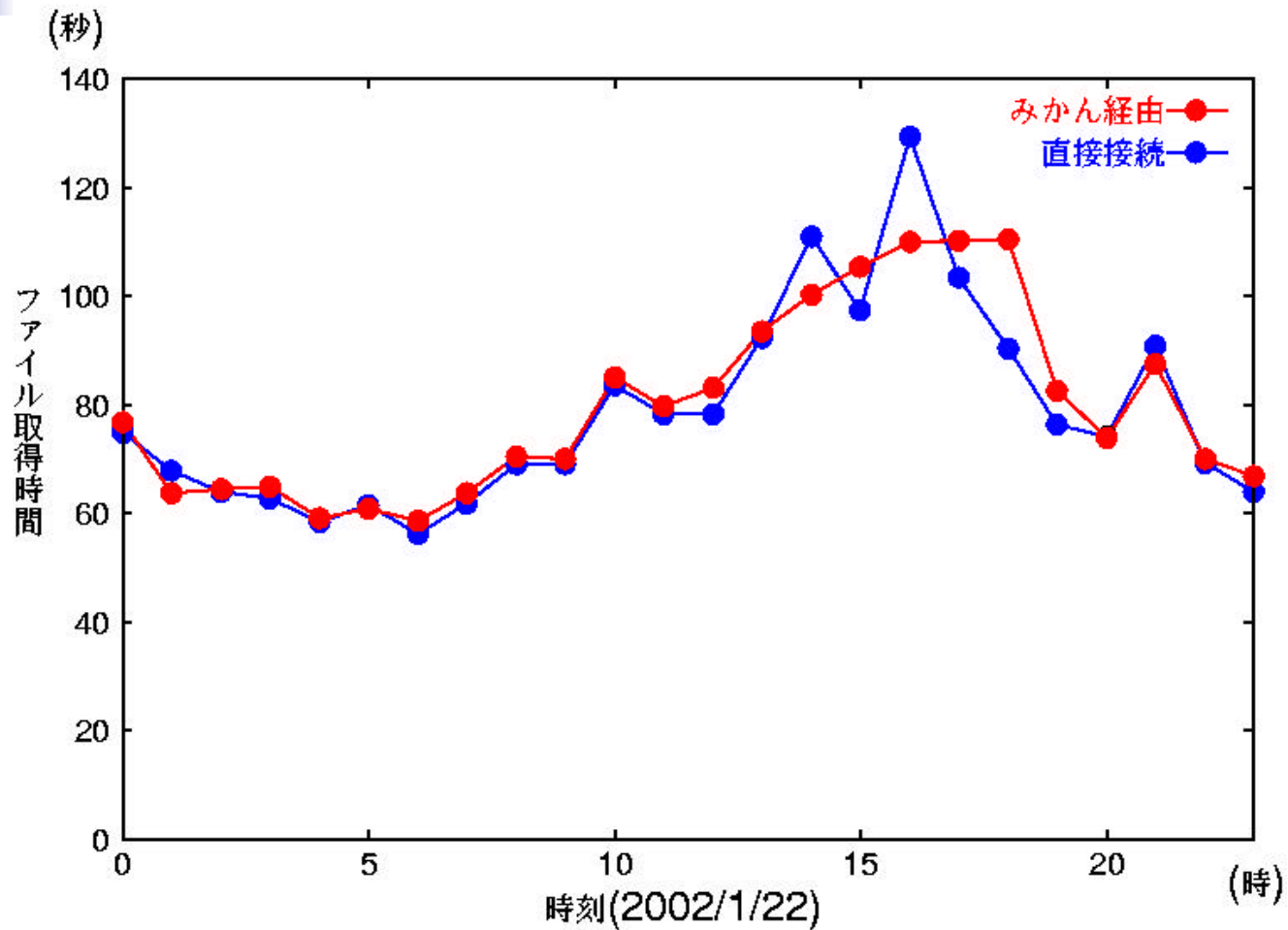




測定方法

- 計15MBの10個のファイルを取得する
- 1時間毎に24時間行い、各測定において10回繰り返す
- 測定した10回の平均をとる

測定結果





可用性の評価

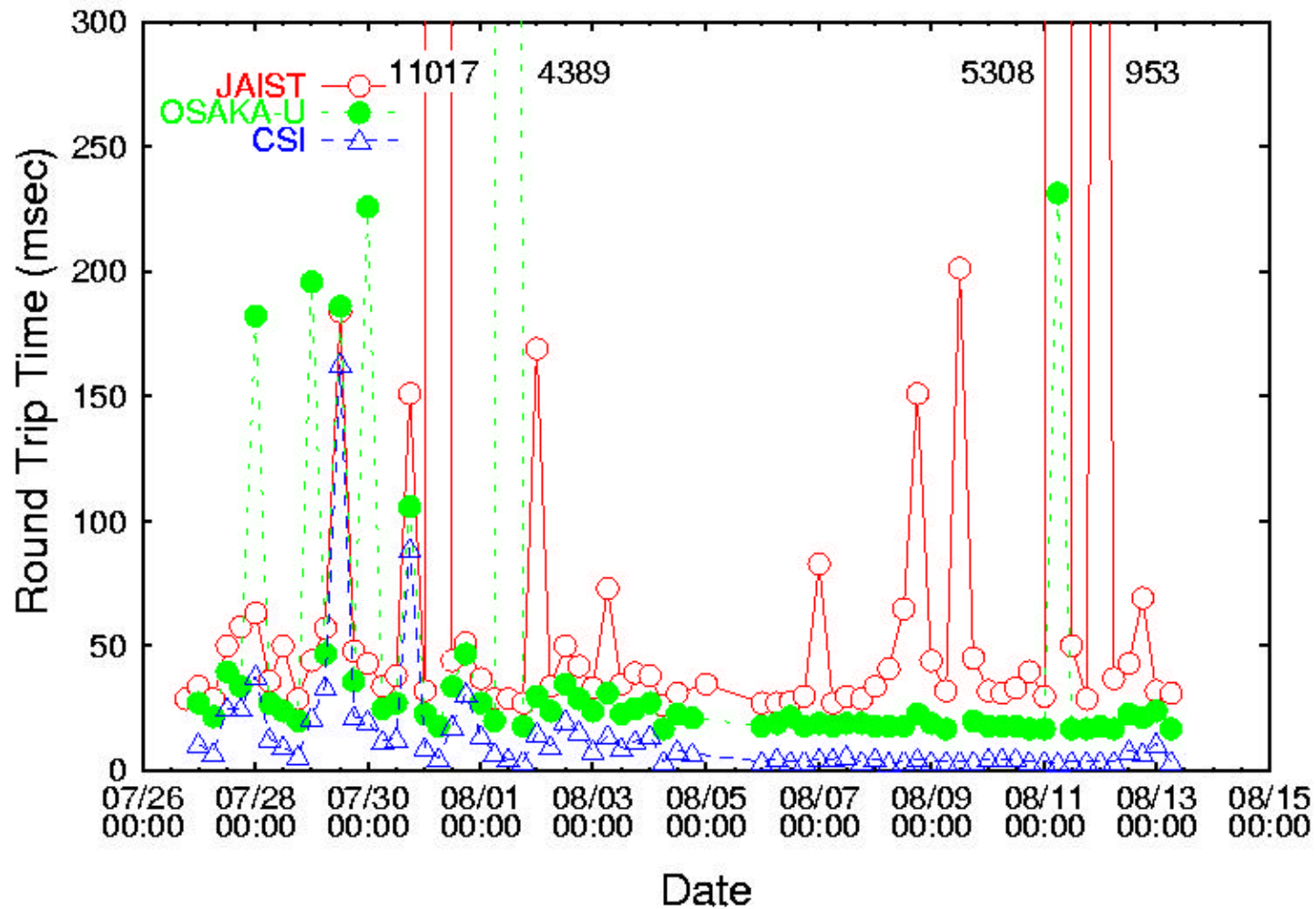
- みかんを運用した場合の効果について評価
 - ネットワーク状況の測定
 - ファイル取得時間の測定



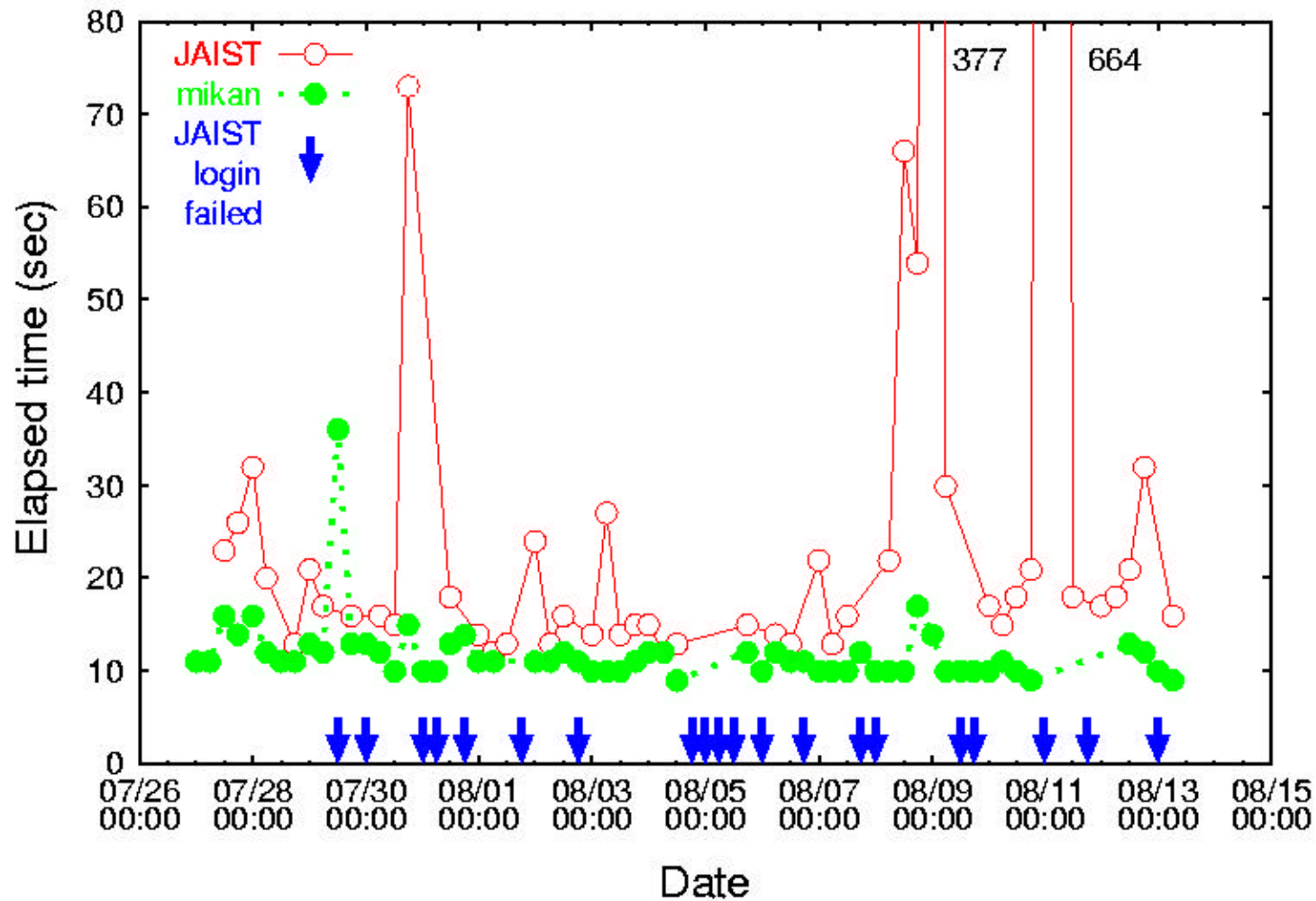
評価方法

- 以下のサーバに対して測定を行う
 - JAIST(北陸先端大)
 - OSAKA-U(阪大)
 - CSI(中国・四国インターネット協議会)
- ネットワーク状況の測定
 - pingによりRTTを測定
- ファイル取得時間の測定
 - 計7MBの10個のファイルを取得する
 - 6時間おきに計測する
 - みかん経由でサーバ選択を行った場合と直接FTPサーバから取得した場合を測定する

ネットワーク状況の測定結果



ファイル取得時間の測定

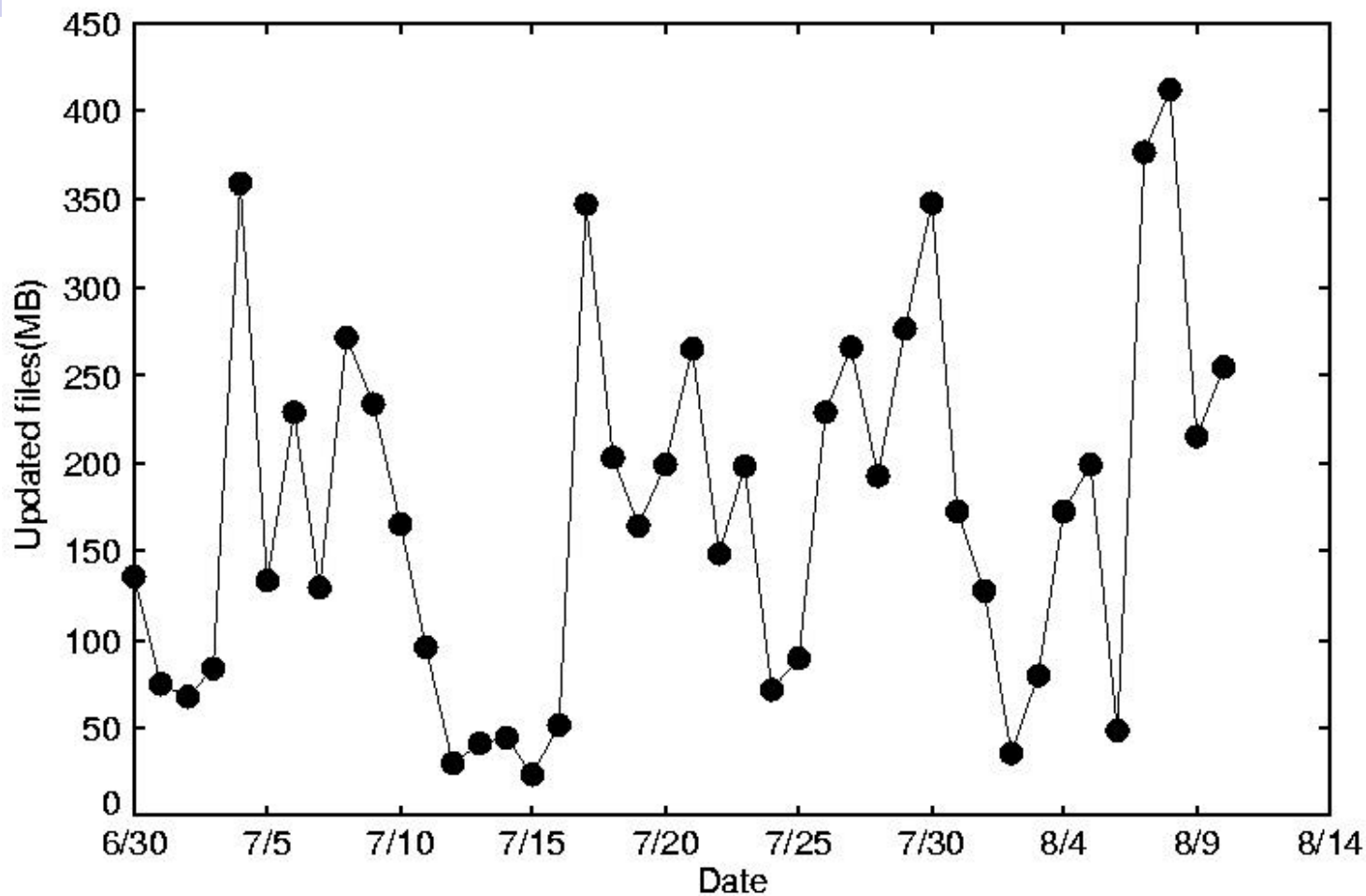




トラフィックの評価

- モニタリングにかかるトラフィックの評価
- Vine Linuxのディレクトリツリーの更新がどの程度行われているのかを測定
- ftp.jaist.ac.jpからVine Linuxのファイルリストを1日毎に取得
- ファイルリストの比較を行い更新量を測定

FTPサーバの更新状況





ミラーサーバとの比較

- 組織内に新規ミラーサーバを構築
 - 膨大な資源が必要(ディスク、帯域)
 - トラフィック大・最新性確認頻度低
- 提案するみかん
 - 必要最低限の資源
 - トラフィック小・最新性確認頻度高

Vine Linux: ファイルリスト(ls -lR) 5.4Mbyte程度
ファイル総容量 31Gbyte程度



FTP Mirror Trackerとの比較評価

- 選択単位の違い
 - ファイル単位(提案方式)
VineSeedディレクトリのサーバ候補・・・10
 - ディレクトリ単位(FTP Mirror Tracker)
VineSeedディレクトリのサーバ候補・・・2
- 対象とするファイルの違い
 - 提案方式はLinuxディストリビューション程度



既知の問題点

- ファイル取得要求の誘導方法
 - ミラーサーバを直接指定された場合は、代理FTPサーバの機能を利用することができない
- モニタリングによるネットワークトラフィックの増加
(適切な更新頻度の見きわめが必要)
- 近いサーバの選択方法



みかんの利点

- ファイル単位のサーバ選択
- 従来クライアントでアクセス可能
- 省資源
- キャッシュ機能



TODO

- まともなキャッシュ機能
- より賢いサーバ選択機構
- 分割ダウンロード
- ミラーサーバのリストをネットワーク上から取得
- 更新状況配送
- アクセス制御
- 普通のFTPサーバ



ご清聴ありがとうございました

みかんの1次配布先

<http://www.nets.ce.hiroshima-cu.ac.jp/~masato/mikan/>