

株式会社ライブドア クラウドコンピューティングへの 取り組み

2009年7月28日 ライブドア 伊勢 幸一

株式会社ライブドア
livedoor Co., Ltd.

■ 会社沿革

- 1996年 「有限会社オン・ザ・エッジ」として設立
- 1997年 有限会社から株式会社に組織変更
- 2000年 東証マザーズに「株式会社オン・ザ・エッジ」として上場
「DATAHOTEL」データセンタ、ホスティング事業を開始
- 2002年 無料ISP 株式会社ライブドアから営業権を譲受
- 2003年 「エッジ株式会社」に社名変更
旧ライブドアの会員サイトをポータルサイト化し、
CGM型サービス「livedoorBlog」を開始
- 2004年 「株式会社ライブドア」に社名変更
- 2005年 色々ありました……
- 2006年 ものすごい事がありました……
- 2007年 株式会社ライブドアホールディングスに社名変更
メディア事業部、ネットワーク事業部が独立し、
(新)株式会社ライブドアを設立
- 2008年 ネットサービス事業部、情報環境技術研究室を新設

■ 事業内容 メディア事業部

ポータルサイト「livedoor」の運営

「livedoorニュース」 月間1億3000万ページビュー

「livedoorBlog」 開設者数：約290万人

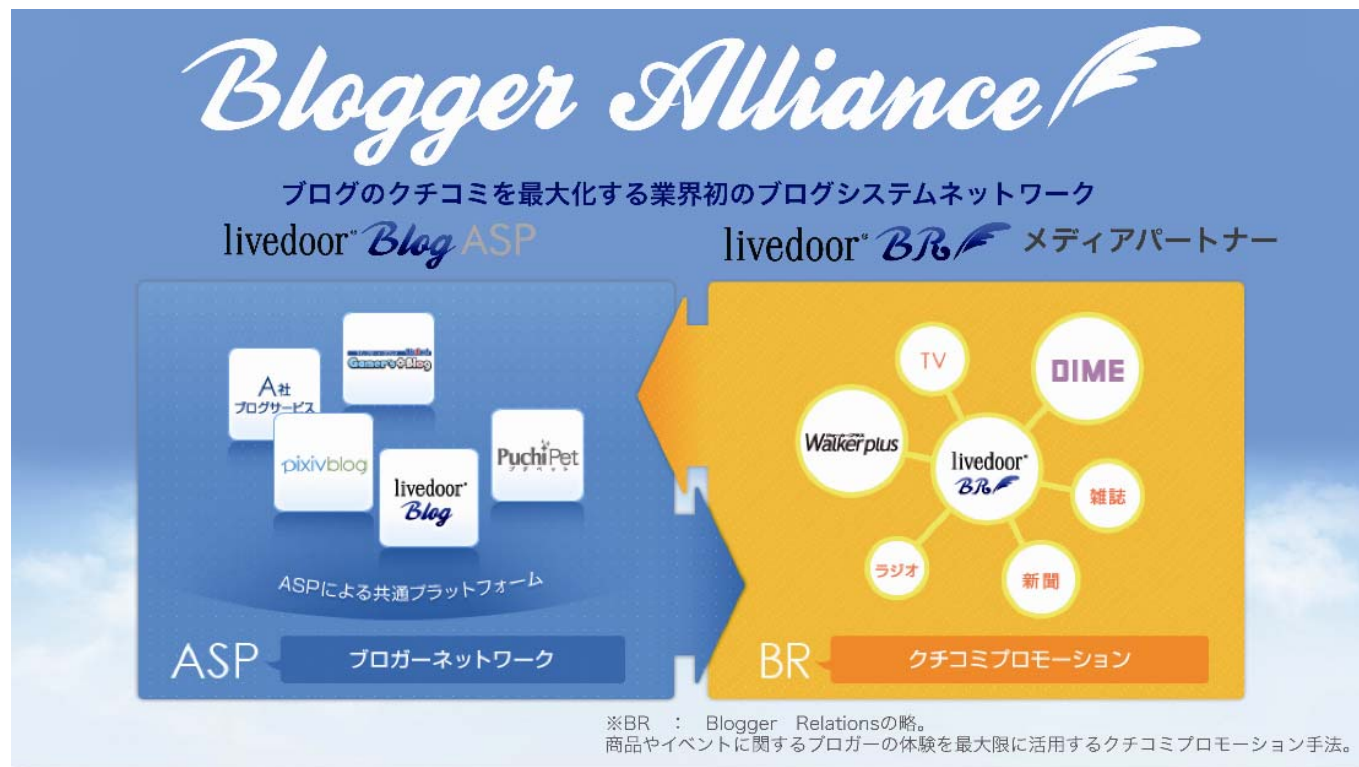
「けいたいlivedoor」 月間9億6000万ページビュー

ページビュー : 約23億PV/月

ユニークユーザー数 : 約2,200万人/月

■ 事業内容 ネットサービス事業部

ポータルサービス「livedoor」で培った プラットフォームをパートナー企業へ提供



pixivblog

ワザップ!ゲーマーズブログ **Wazap!**
Gamer's Blog



法人顧客数約4000社
インフラ事業を中心とした、
インターネットソリューションを提供

- ホスティング事業
- 回線通信サービス事業
- システムインテグレーション事業



2008年4月 ネットワーク事業部内に
通信技術研究準備室として設立

2008年10月 情報環境技術研究室として新設

研究テーマ

次世代ネットワークアーキテクチャ

P2P配信技術

オーバーレイネットワーク技術

仮想化技術

(電子情報通信学会IA研究会 2009年9月25日発表)

【タイトル】

～クラウドシステムのための分散型URI方式の評価～

DHT技術によるブログ用ストレージの分散化と検索の実証実験

【概要】

従来、インターネット上のリソースを特定するためには、IPアドレスとポート、そしてディレクトリパスという複合的なアソシエーションによって行なわれていた。しかし近年では全てのリソースを直接一意的に識別するURI方式の採用へと移行しつつある。

しかし、リソースの増大に伴い、複数のストレージシステムへデータを分散配置する場合、そのURI論理インデックスと物理ロケーションパスのマッピング情報は膨大な量となり、key-value型の単純な構造をしているとはいえ、従来のRDB型データベースシステムによる検索方式ではデータベースシステムへの負荷集中と過負荷を引き起こし、また単一障害点(Single Point of Failer)の存在を排除できないという問題がある。

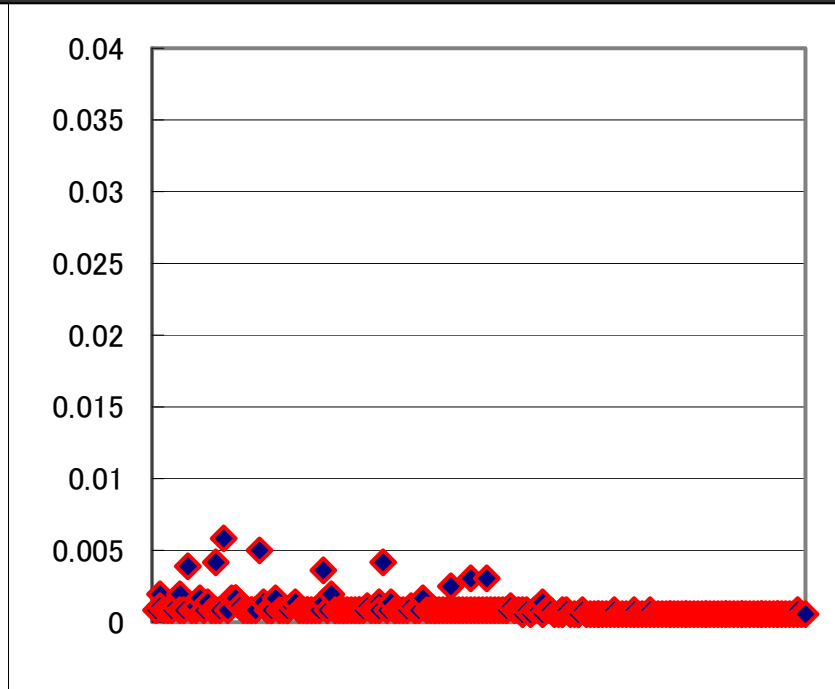
オーバーレイネットワークに用いられる分散ハッシュテーブル(DHT)の手法として数値的な環状探索空間を利用するアルゴリズムにChordがあり、このChord DHTを膨大なURI情報を管理する分散型インデックスとして利用する事によって、ストレージシステムの分散化を容易にし、無尽蔵に巨大化していくWebサービス用ストレージとして高いスケーラビリティを実現することが可能となる。

本論文ではChord DHTの実装の一つであるSOD(Secure Overlay DHT)を利用しブログ用ストレージの分散化を図るため、HTによる分散型URI方式の性能や有効性を検証すると同時に、RDBデータベースによる従来型の方式との性能特性を比較する。

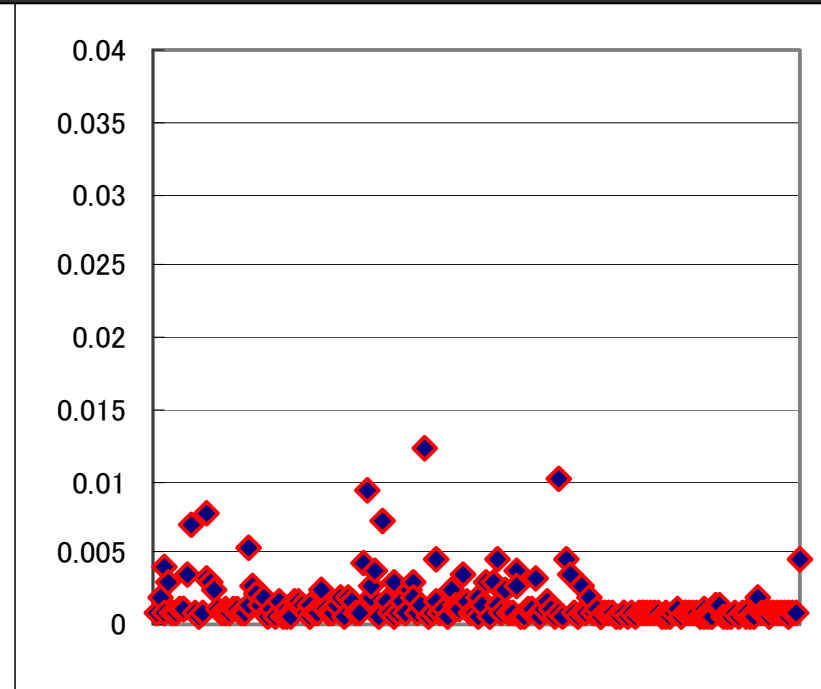
■ 研究室紹介 DHT性能評価

【DHT】 ノード数:増減 | 時間当たりのリクエスト数:200

【NODE数】 400 【リクエスト数】 200



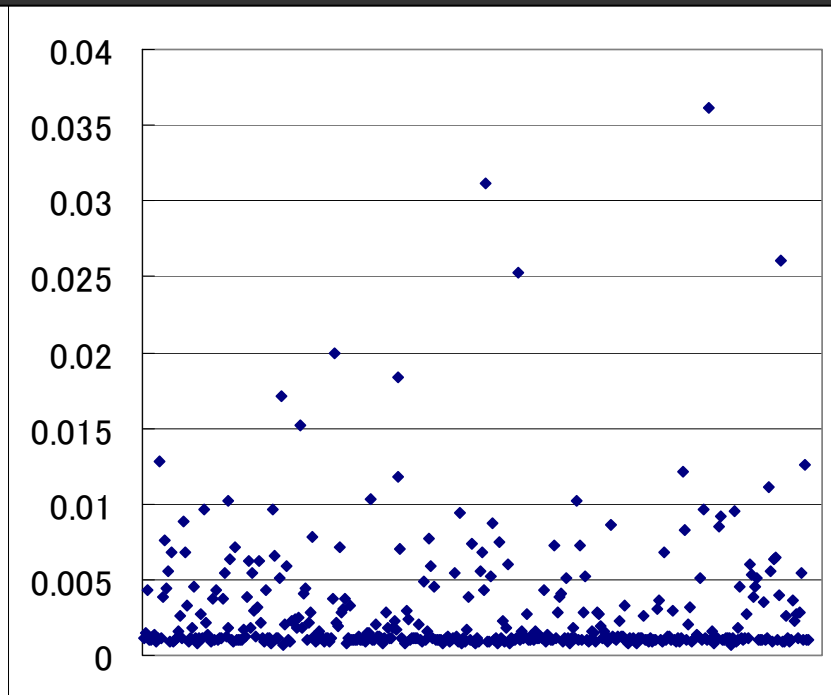
【NODE数】 1.200 【リクエスト数】 200



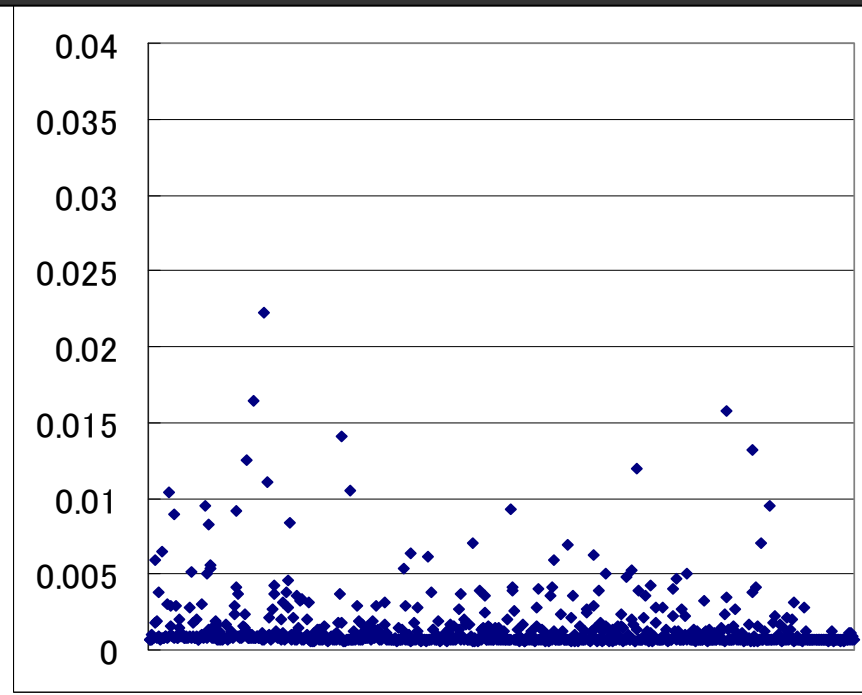
■ 研究室紹介 DHT性能評価

【DHT】 ノード数:1,300 | 時間当たりのリクエスト数の変動

【NODE数】 1.300 【リクエスト数】 400

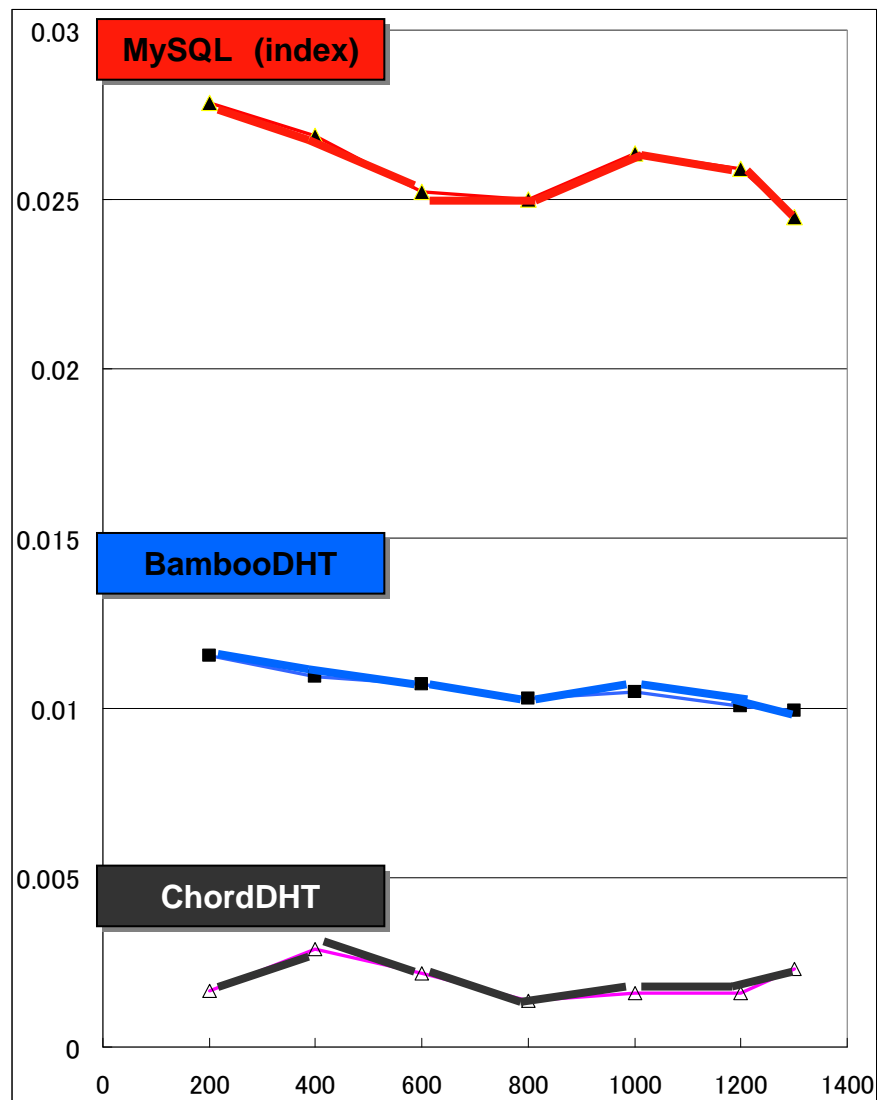


【NODE数】 1.300 【リクエスト数】 1,200



研究室紹介 DHT性能評価

2つのDHTアルゴリズムと
MySQLによるクエリ解決速度



■ Webサービスの特徴と課題

- ・ユーザ参加型のサービスはデータが無限に増加
- ・ユーザあたりの収益率は低い
- ・リクエストパターン Read が殆ど

【CAP定理】

Consistency 整合性、一貫性

Availability 可用性

Partition Tolerance 拡張性

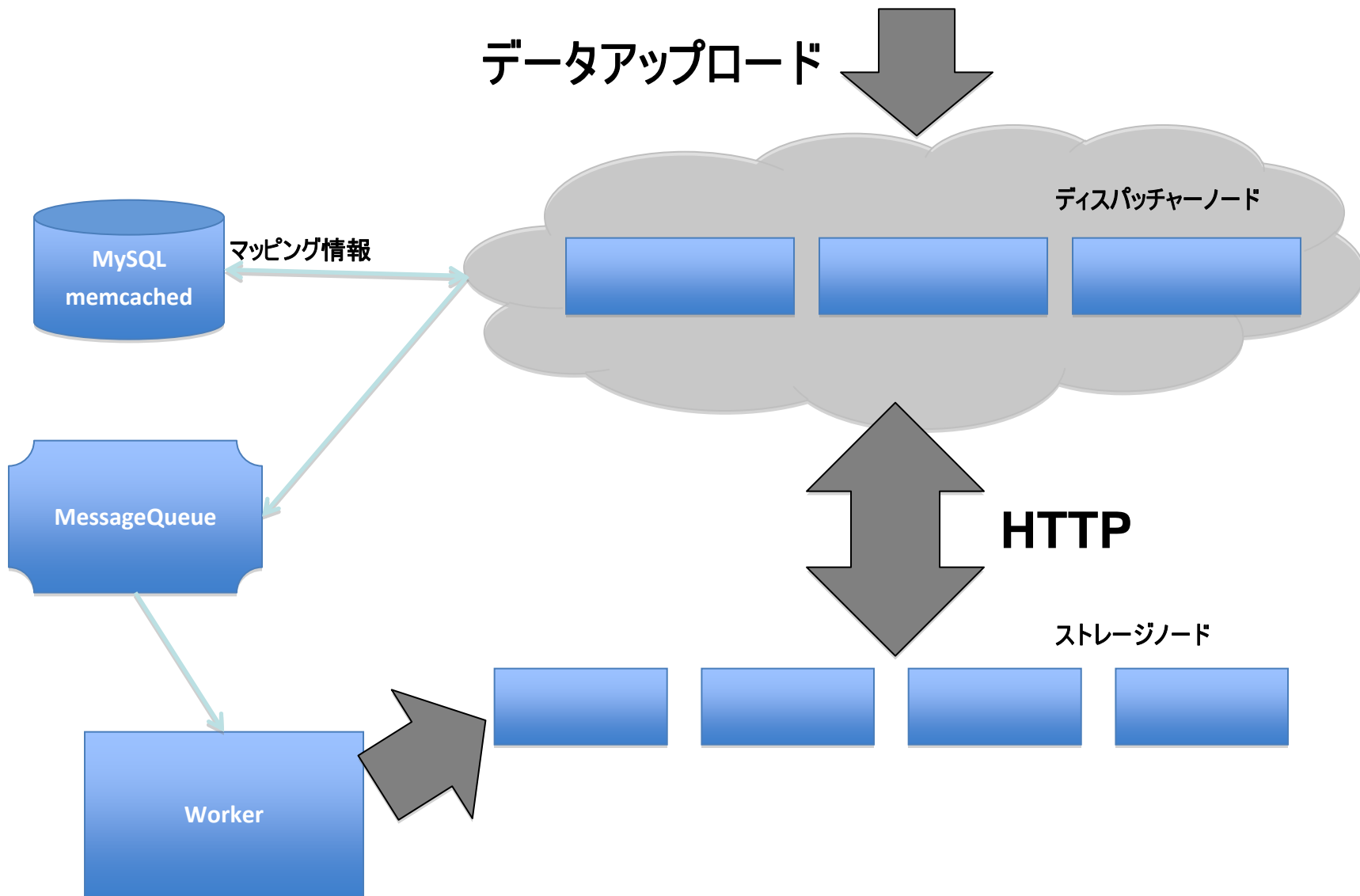
(同時に満たせるのは2つまで)

安価に容量を拡大出来る

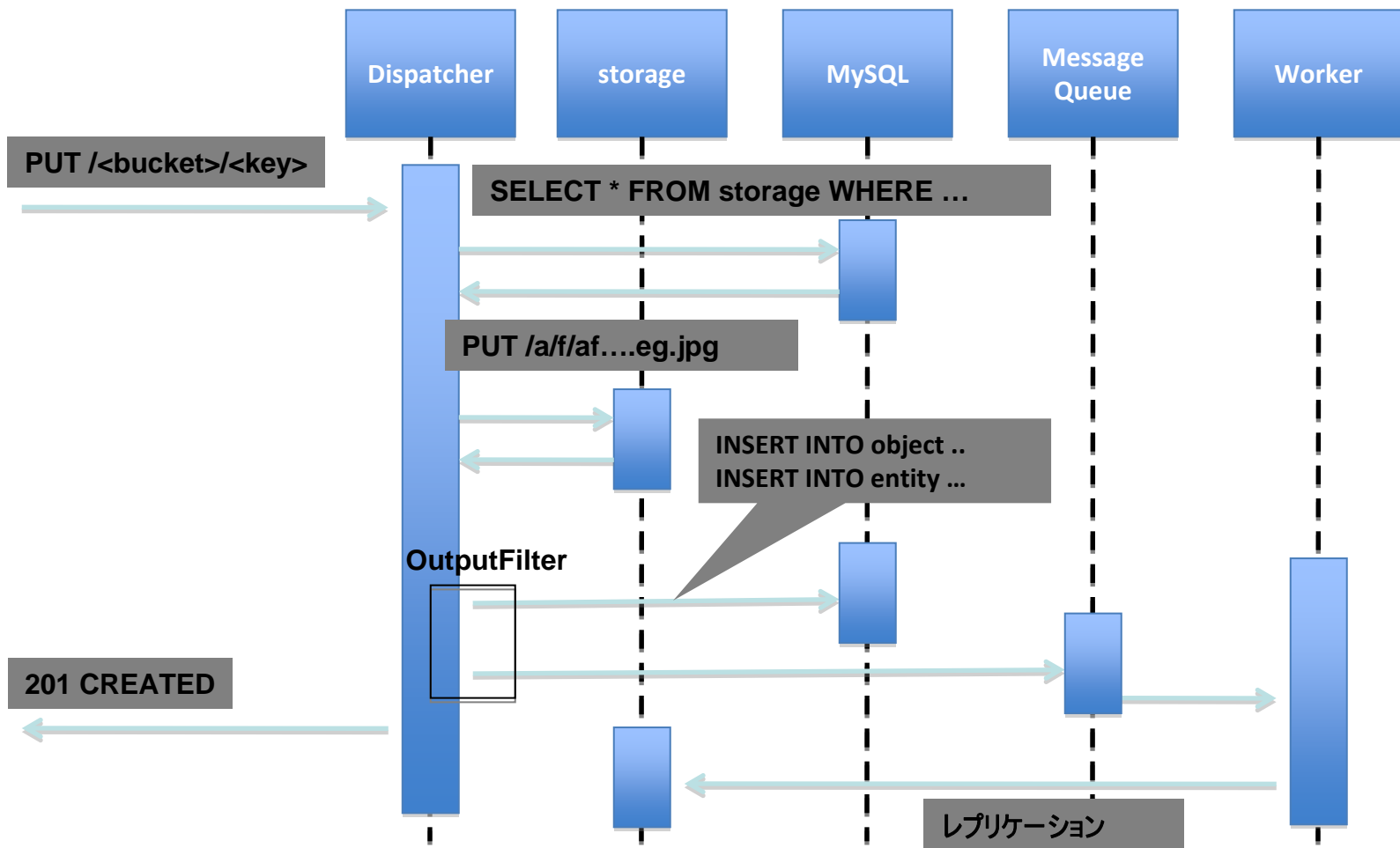
クライアントからは一つに見える

過剰な機能/信頼性より価格

クラウドストレージシステム STF (Storage Farm)

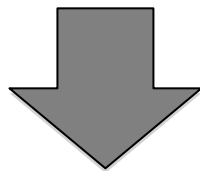


クラウドストレージシステム STF (Storage Farm)



■ STFの問題点

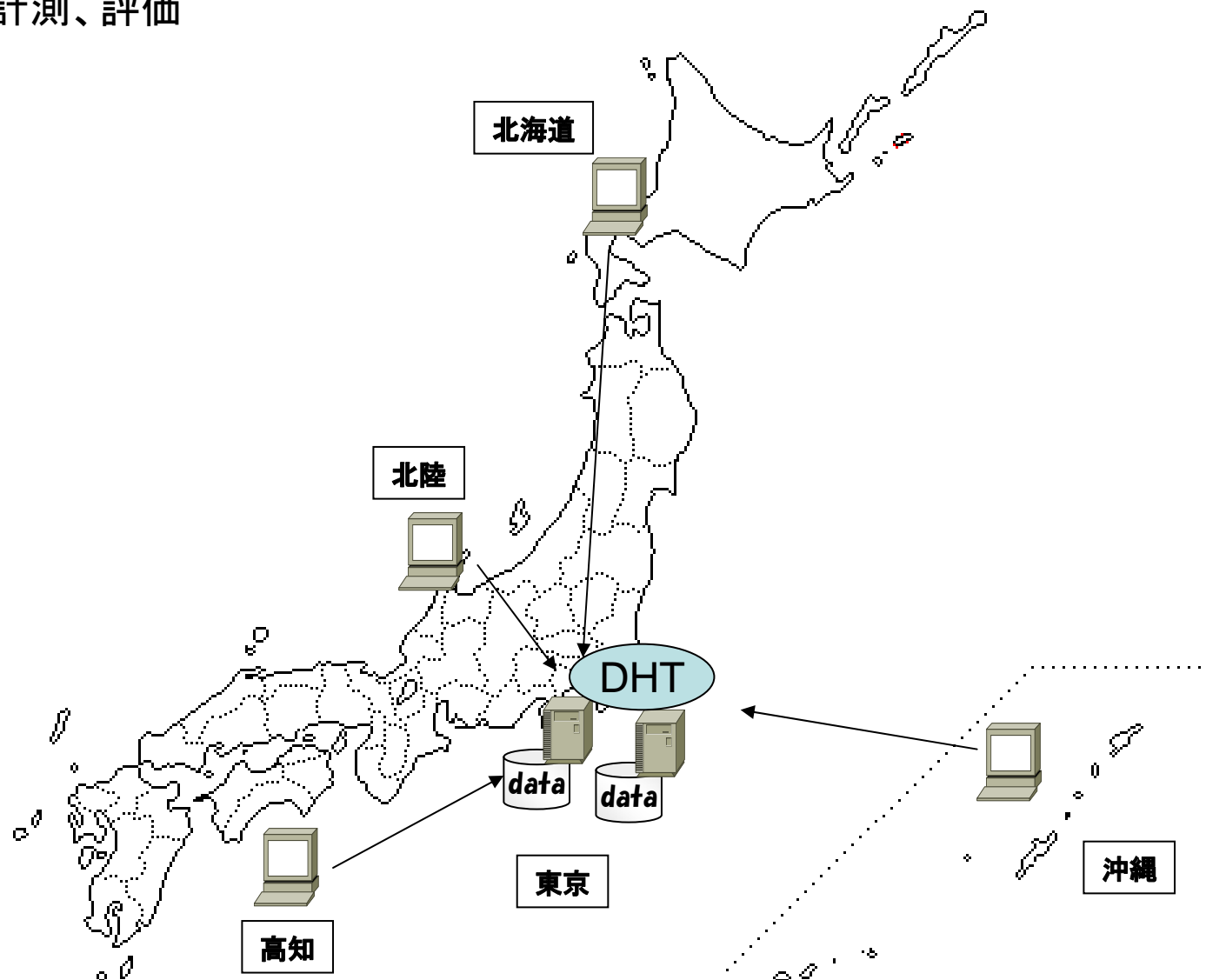
- mySQLはパーティショニング性に乏しい。
- スケールアウトしにくい。
- 一箇所にホスティングする必要性がある。



mySQLによるマッパーをDHTにリプレイス。

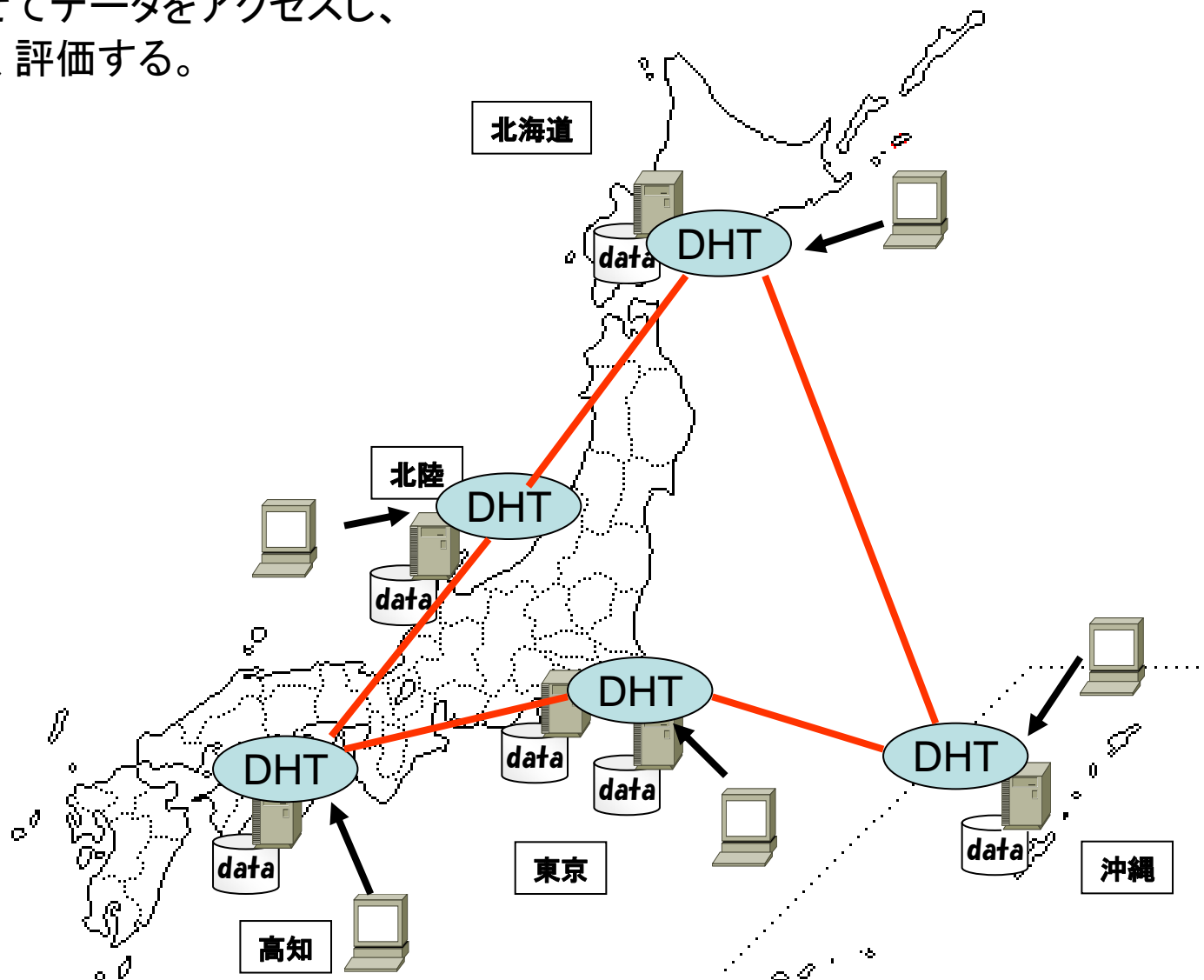
■ 広域分散型クラウドストレージ実証実験

1. 都市部に集約したDHT方式ストレージにリプレイスし
レスポンス性能を計測、評価



広域分散型クラウドストレージ実証実験

2.全国にDHTとデータを分散し、Webサーバは最寄のDHTに問い合わせデータをアクセスし、レスポンスを測定、評価する。



■ 広域分散型クラウドストレージのメリット

1. ロケーション(都心部)の呪縛からの開放

不動産コストメリット

地域経済活性化

人材雇用促進

データセンタマイグレーション

2. 地発電地消費

自然エネルギー発電を利用

グリーンIT化の推進

以上