



開発物語

オープンソースでスケールアウト

2005年9月20日
バタラ・ケスマ
株式会社イー・マーキュリー

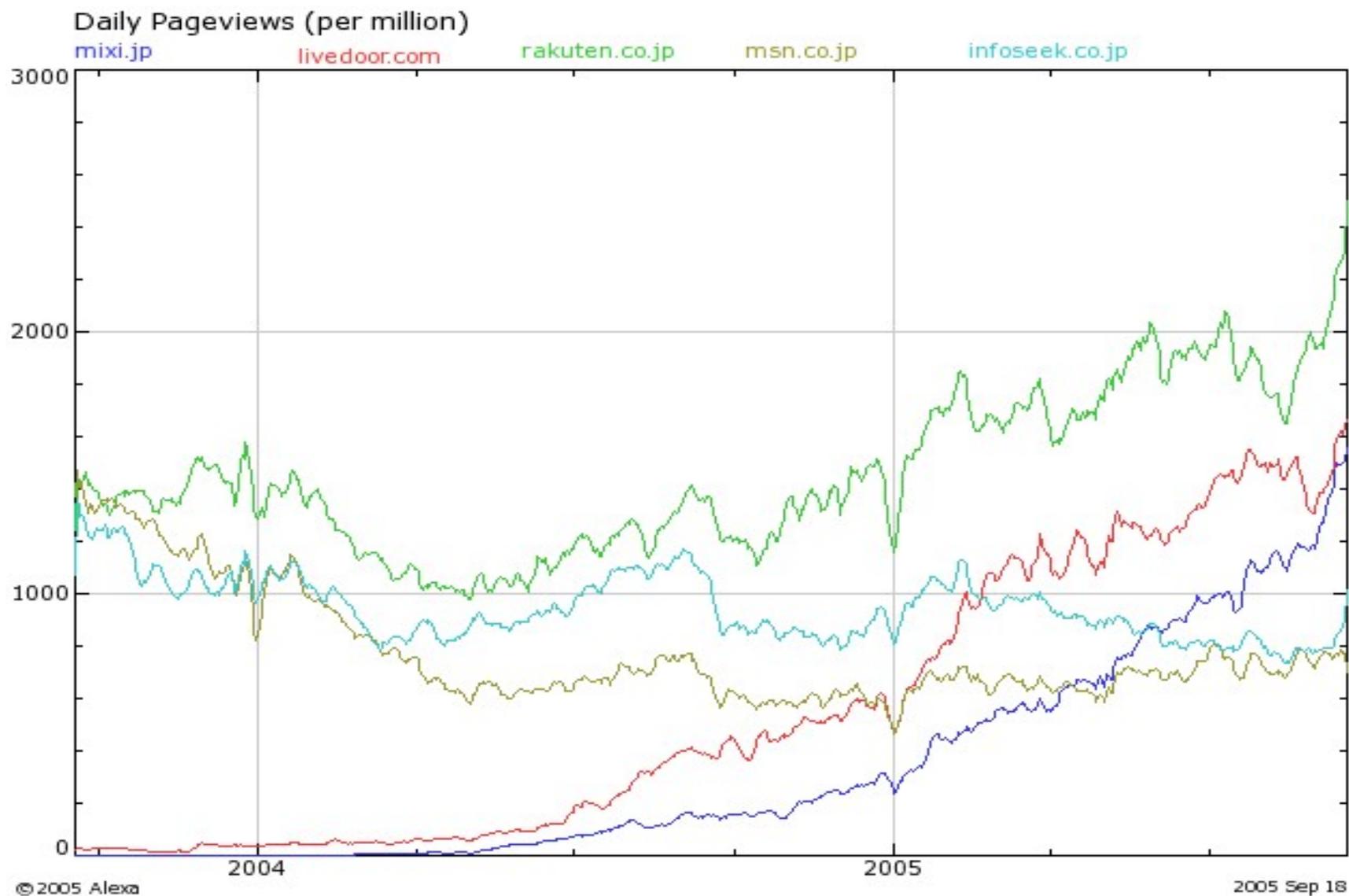
mixiの歴史

- 2003年12月から開発開始
- 2004年2月にオープン
- 最初の6ヶ月間ぐらい設計、コーディングは1人
- 2ヶ月で1万ユーザ、60万ページビュー／日

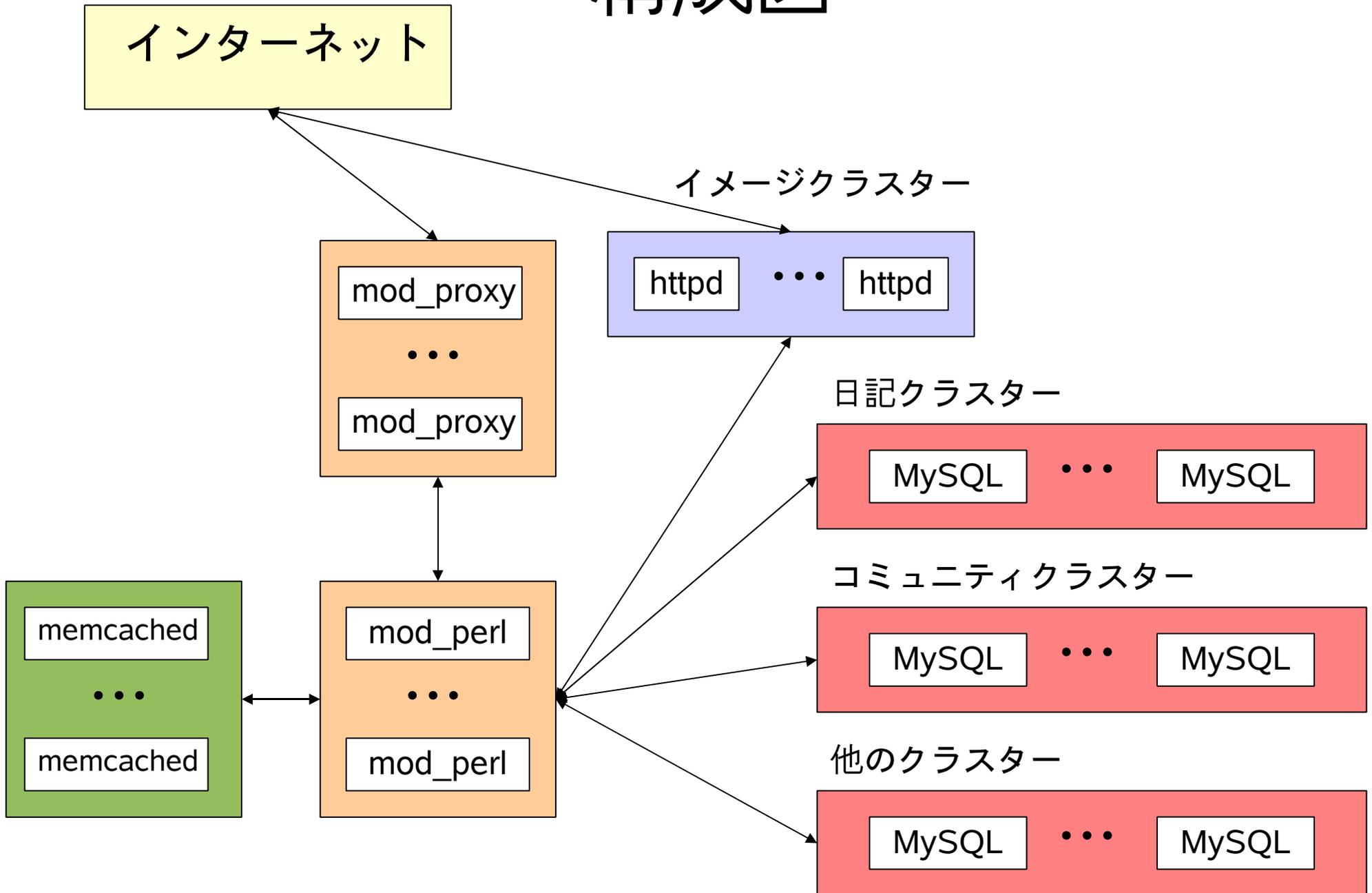
現在のmixi

- 1日は6000万ページビュー
- ユーザ数135万人 (1日約7000人)
- アクティブユーザ70% (3日以内にログイン)
- 作成された日記3950万件 (1日約24万件)
- Alexaランキング国内で10位
- 約150台のサーバ
- 5人のエンジニア

Alexaのページビュー



構成図



mod_proxy

- DNSラウンドロビンで負荷分散
- mod_proxyとmod_rewriteの組み合わせ
- .plのリクエストのみ
- ディスクにログ
 - 各ホストからログを集約して解析

mod_perl

- Apache 2のpreforkとmod_perl 2
- ModPerl::Registry
 - 一部でHandlerで処理
- HTML::Template メモリーにキャッシュ



MySQL

- MySQL 4.0と4.1
- ログ系はMyISAM、データはInnoDB
- Master-Master, Master-Slave, 色々な構成
- PerlからApache::DBIとDBIでコネクト
- connectはとっても速い、persistent connectionは要らない
- skip-grant-tablesでさらに負荷軽減
 - インターネットから隔離されたLAN

DBスケールアウト

- 最初はMaster-Slaveレプリケーション
- 負荷が高くなると、Slaveを増やす
- mixiはreadとwriteの多いサイト
- Writeが分散されない
- データの遅延が起きる
 - データクラスターに移行

データクラスター

- JOINが少ないテーブルを別のDBに分散
- レベル1：機能別の分散（日記クラスター、コミュニティクラスター、など）
- レベル2：さらにID別で分散（1～100番の日記クラスター、など）
 - アルゴリズムでクラスターを選択
 - マッピングテーブルでクラスターを選択
- データによって、構成を調整する

画像サーバ

- SquidでDNSラウンドロビン
 - 数が少ないが、表示が多い (プロフィール写真)
- イメージクラスター
 - 数が多いが、表示が少ない (日記の写真)



Squid

- ストレージにファイルをアップロード
- 複数のSquidで配信
- SquidのSiblingでストレージの負荷を軽減
- ただし総数が多く、配信頻度が低い画像には向かない
 - 日記の写真など
 - 画像が増え続けるが、古い画像にはアクセスが(とっても)少ない

イメージクラスター

- mixi用に開発
- ストレージと配信を同じホストで
- アップロードされた画像を複数のホストに分散して配置
 - ホスト単位の負荷を軽減
 - 冗長性を確保
 - 画像の場所はMySQLで管理
- ホスト間でファイルを簡単に移動・複製できる
 - 専用クラスターに古い画像を集める
 - 配信サーバの追加

memcached

- メモリーが余っているマシンに設定
- ホットデータをキャッシュ
- 分散のアルゴリズムでサーバ追加時、ほとんどのキャッシュが消える
 - サーバリストの変更あまりしない方がいい
 - しょうがない？
 - Siblingからキャッシュを取るコードを追加すべきか？

スペック

- Linux 2.6
- MySQL 4.0と4.1
- Apache 2
- mod_perl 2
- Perl 5.8.x
- Squid
- memcached
- 約5万行のソースコード
- 約570枚のテンプレート
ト

ありがとうございます

- 質問は**bkesuma@emercury.co.jp**
- 現在エンジニア募集中！