

Internet Week 2010

S5 HTTP Meeting ~ソーシャルアプリ over HTTP~

~ソーシャルアプリケーションの現状~

GREE株式会社 藤本真樹



藤本 真樹 (取締役執行役員CTO)

- 31歳 / 独身
- 2005/06 ~ グリー株式会社取締役
 - GREE開発全体(特に低レイヤ)のマネジメント
 - GREE基幹部分/ミドルウェア開発
 - Ethna (ウェブアプリケーションフレームワーク)
 - Flare (分散Key Value Store)
 - ...

GREEの紹介 (1)



SNS + ソーシャルゲーム + プラットフォーム



数字 (2010/09)

- 会員数 = 2246万人
- 月間ページビュー = 465.2億ページビュー
 - 日次 = 15.5億ページビュー
- サーバ台数 = ? 台
 - ほとんどが普通のx86サーバです
 - SSDなどはこれから
- 最大同時接続数 = > ? (画像除く)

ソーシャルアプリケーションの現状



ソーシャルアプリケーションの現状 (日本)



	Facebook	GREE	mixi	モバゲー
Interface	Facebook API	OpenSocial	OpenSocial	OpenSocial
フィーチャーフォン	×			
スマートフォン		×		×
PC		×		

- ソーシャルアプリケーション (実質)メジャーSNSプラットフォーム上で動作するアプリケーション
- フィーチャーフォン上のアプリケーションが多くのトラフィック
- 「プラットフォーム」としてはフィーチャーフォン上のアプリケーションは徐々に成熟(?)
- スマートフォン上のプラットフォーム/アプリケーションは大いにこれから

ソーシャルアプリケーションの現状 (日本)



• ソーシャルアプリケーション ソーシャルゲーム

	GREE	mixi	モバゲー
1	ドラゴンコレクション	サンシャイン牧場	怪盗ロワイヤル
2	100万人の三国志	まちつく! Mixi版	戦国ロワイヤル
3	恋人は専属SPforGEE	みんなでケンテイ	農園ホッコリーナ
4	君とのナイショの...	バンドやろうよ!	スペースデブリーズ
5	コーデマニア	星空バータウン	ポッピンシティ
6	恋愛上等イケメン学園	アニマルパラダイス	アクアスクエア
7	しろつく	三文堂のひよこ生活	海賊トレジャー
8	ダービーズキングの伝説	みんなで会社ごっこ2	モバ7
9	おみせやさんforGREE	ドラゴンハンターズ	四人打ち麻雀
10	ヒメこい	Disneyご当地コレクション	100万人の信長の野望

• たかがゲーム、されどゲーム

- 億単位のPV / 億単位の売上 (がある、らしい、です)

• アプリケーションの広がりはまだまだこれから

- Another game model, non-gaming, etc... + スマートフォンの普及

ソーシャルアプリケーションの現状 (海外)



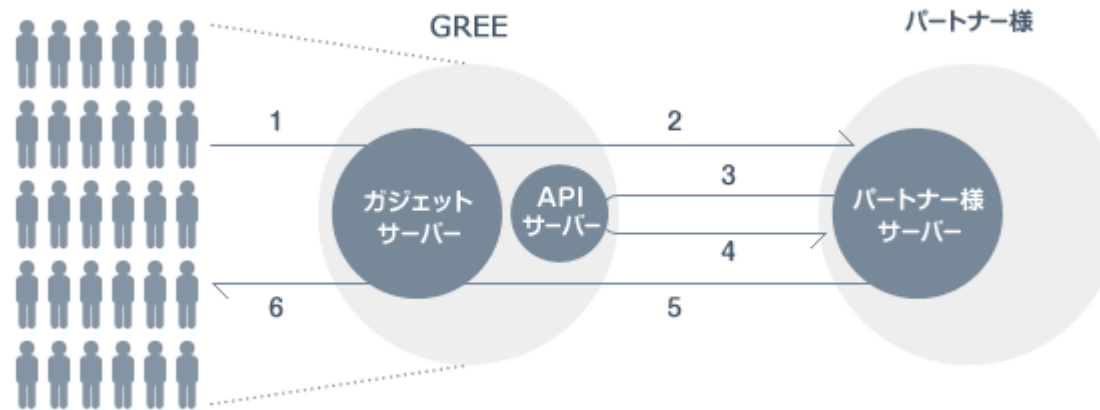
- Top 6/10 Apps are Games
 - Also has simple communication apps
- Mobile (= iPhone/Android)
 - Still has no dominant platform

	Facebook
1	Phrases
2	FarmVille
3	Texas HoldEm Poker
4	Frontier Ville
5	Mafia Wars Game
6	Causes
7	Café World
8	Quiz Planet
9	Are YOU Interested
10	Give Hearts

ソーシャルアプリケーションの 技術

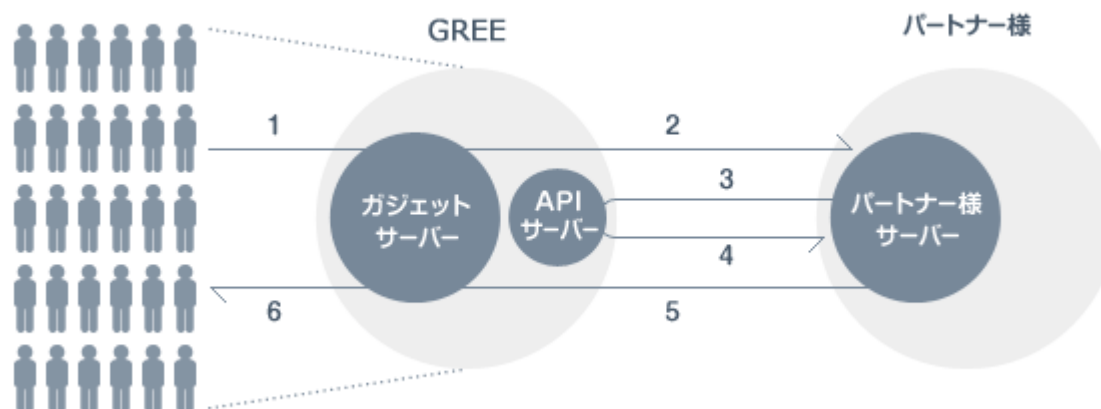


基本的には通常のウェブアプリケーションと同様



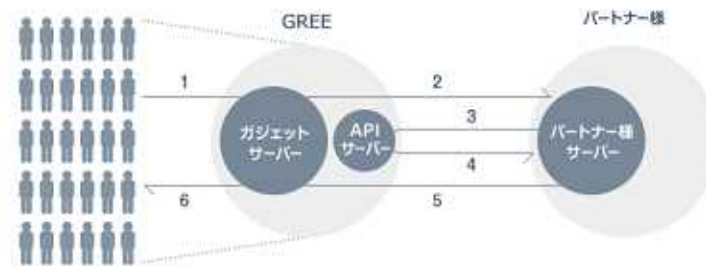
- 上記はモバイル版拡張 (基本は通常のOpen Social Applicationと同様)
- 単純なアプリケーションならOAuthの処理を追加するだけ
 - (多くの)フィーチャーフォンでJavaScriptが使えないことによる代替案
 - iframeの代わりに、ガジェットサーバによるHTTPリクエスト
 - JavaScript APIの代わりにRestful API
 - PCブラウザベースでも多くのガジェットは似たような振る舞い

これにより発生する技術的な問題



- 全てのリクエストがガジェットサーバを経由する (当然ながら画像も含む)
 - トラフィックの増大
 - 考えようによっては無駄遣い (経路による)
 - ガジェットサーバのCPU負荷
 - 「5」 「6」で多くの処理を行う
 - TCPポート問題 (特にNAPT時)

これにより発生する技術的な問題



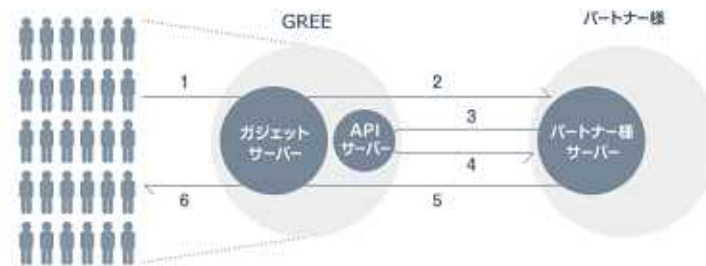
• トラフィックの増大

- ガジェットサーバがHTTP Proxyサーバのような振舞い
- ガジェットサーバ <-> サービスプロバイダ間のトラフィックが新たに発生

• Solution

- とりあえずdeflate
- CDNの提供 (Content-Type: Image/*はデフォルトでキャッシュ)
- Peering

これにより発生する技術的な問題



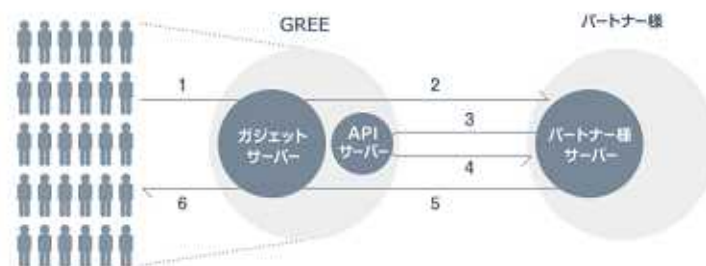
• ガジェットサーバーのCPU負荷増大

- HTMLのFull Parse
 - ヘッダ/フッタの挿入
 - 禁止タグの削除
 - HTMLの正当性チェック
 - 属性に記述されているURL変換
 - etc...

• Solution

- C/C++で高速なものを書く
- GREEではPHP Extensionで実装 (必要に応じてPHP側のメソッドをコールバック)

これにより発生する技術的な問題



• TCPポート問題

- ガジェットサーバ <-> サービスプロバイダは、しばしば(ほとんど)グローバルアドレス間での通信
- 一方で通常アプリケーションサーバはプライベートアドレス
- 通信量は膨大 + 短いセッションが多数
 - NAPTセッションがあふれる
 - Destinationは往々にして1IPアドレスで80番ポート固定
 - Source側のポートがあふれる

• Solution

- できるだけ多くのグローバルアドレスを確保する (この時代に...)

ソーシャルアプリケーション プロバイダの技術



基本は通常のウェブアプリケーションと同様...



- 通常のウェブアプリケーションとの違い

- Open Social (and related specs)
 - OAuth
 - Lifecycle Events
 - Payment Related
- トラフィックパターン
 - 急に伸びることもあれば、急に落ちることもあり
 - パフォーマンスダウンでアプリケーション提供一時停止
- 多くのアプリケーションがWrite Heavy
 - 可能なら、出来る限り多くのアクティビティを共有したい + リアルタイムに反映させたい
 - ユーザ多くない時期からDisk I/Oに悩まされることは多い

1. Writeを分散

- sharding (後述)

2. Diskにかかない

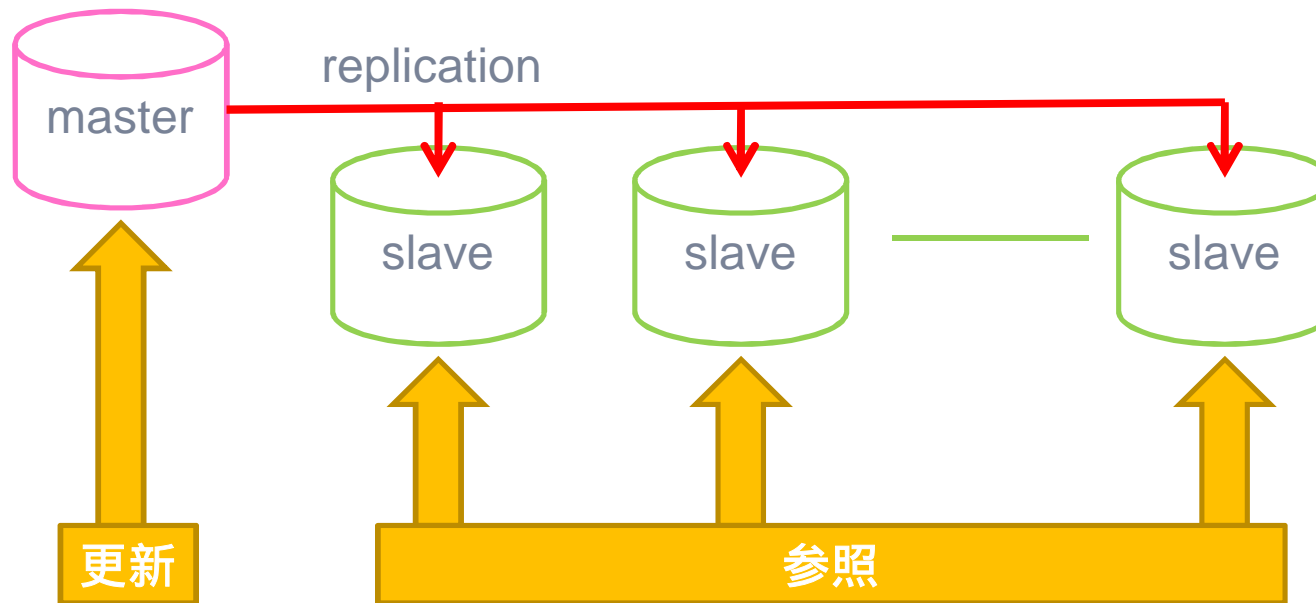
- オンメモリファイルシステムで運用 (e.g. tmpfs)

3. Diskにあまりかかない

- キャッシュ(memcached)や中間データ(KVS)の利用

- Read (Writeの前に)

- MySQLではレプリケーション機能を使って参照専用のslaveを増やせる
- 基本的には台数を増やすだけで対応可能
- *サーバ単位でのWriteの量は変わらない点に注意*



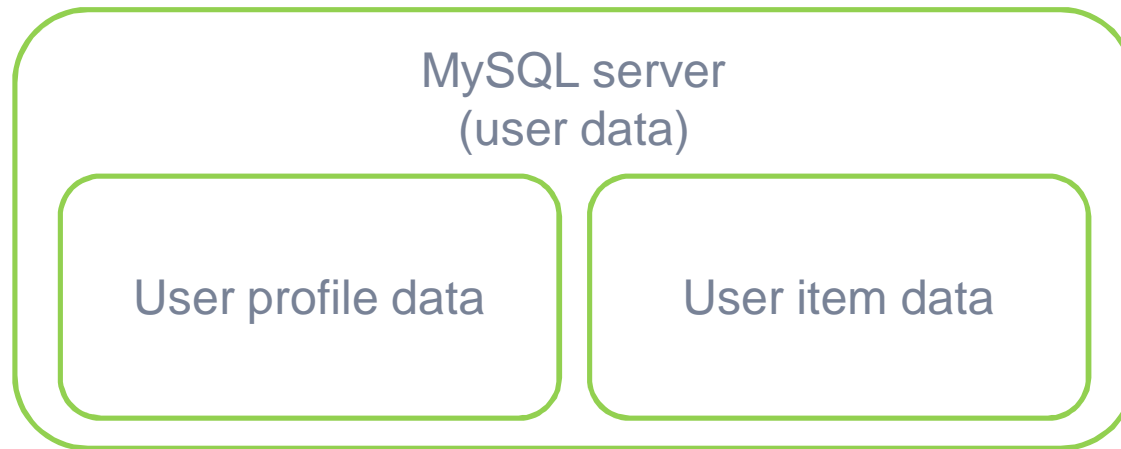
- Write

- マスタースレーブ構成では、マスターは1台
- 書き込みの限界は低い

分散

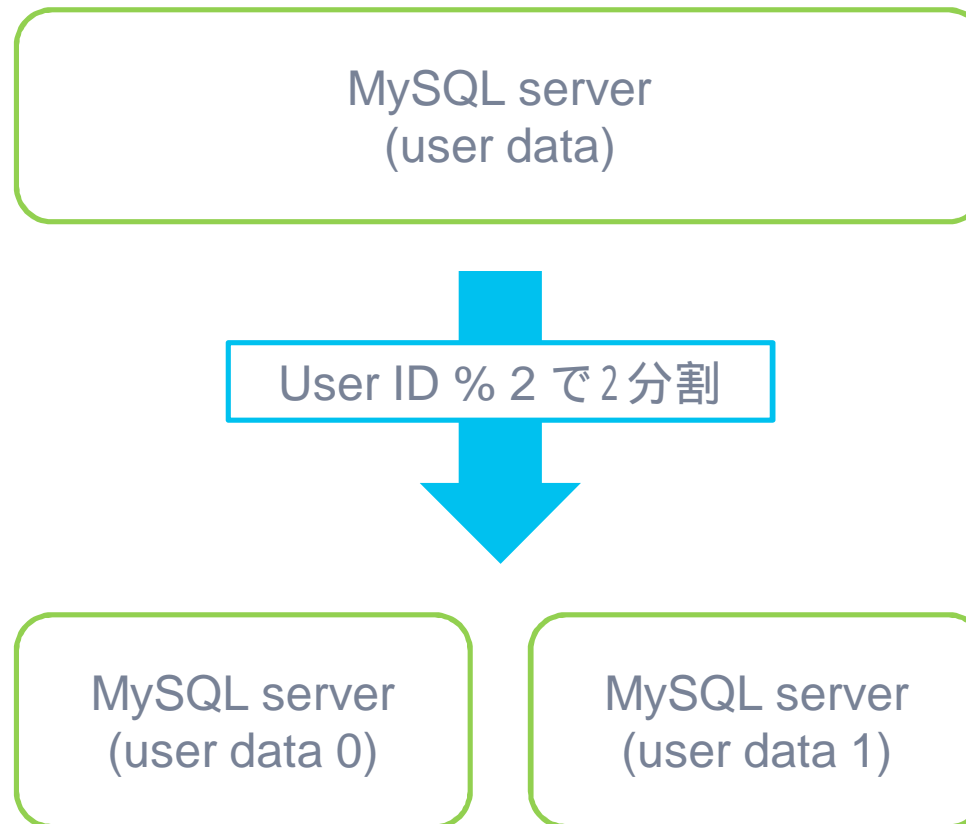
- (垂直分割)
 - 機能ごとにテーブル/データベースを分ける
- 水平分割 (Sharding)
 - 同一テーブルを範囲ごとに分割する
 - Range
 - Hash / Modular
 - Index

Writeを分散:Sharding (3)



→ JOINが制約

Writeを分散:Sharding (4)



→ User ID以外のConditionが制約
(たとえばAgeなどで検索したい場合は?)

- tmpfs(Linuxの場合)に全データを置いてしまう
 - ハイパフォーマンス (I/Oがボトルネックになることはほぼなくなる 昔ながらの手法)
 - 注意点
 - サーバダウンでデータロス
 - ラック単位 / DC単位で分散
 - 適宜snapshot
 - トラフィック (e.g. 12.5k x 1万qps = 1Gbps)
 - Swap領域
 - データサイズ

Diskにあまりかかない



- I/OはKVS [Key Value Store] (GREEでは Flare(自作)) に対して行う
 - パフォーマンスがよい (<> RDBMS)
 - 運用コストが低い
 - Memcachedよりも高いデータ信頼性
 - 必要ならオンメモリファイルシステムで運用
- + 適当なタイミング(e.g. Level Up)でRDBMSへ Flush
 - 集計 / Valueでの検索 / バックアップ



サマリ



- 現状ソーシャルアプリケーション ソーシャルゲーム
 - 新たなアプリケーションの余地も大いに
- プラットフォームプロバイダで発生するいくつかの技術的問題 (特にフィーチャーフォン上)
 - トラフィック / ガジェットサーバの負荷 / TCPポート
- アプリケーションプロバイダで発生するいくつかの問題
 - OAuth / Lifecycle Events / ...
 - トラフィックパターン
 - Disk Writes

Thank You
&
Happy Hacking
<fujimoto@php.net>

