

HTTP Meeting 2008
～ Webサービスの明日をみつけよう～
パネルディスカッション
「memcached」

2008/11/26



Confidential

Copyright © 2004-2008 GREE, Inc. All Rights Reserved.

今日のテーマ



memcachedを
どう
使っていきましょう？

1

Confidential

Copyright © 2004-2008 GREE, Inc. All Rights Reserved.

- GREEで使っているもの
 - RDBMS(MySQL)
 - WebDAV
 - ローカルディスク(あとで回収)
- その他よくあるもの
 - RDBMS(MySQL以外のいろいろ)
 - NFS(+drbdしたり)

- GREEで使っているもの
 - memcached
 - MySQL (on memory / on disk)
 - ローカルディスク
- その他よくあるもの
 - memcached互換のいろいろ
 - in processのいろいろ (Javaだとehcacheとか)
 - erlangだと知らないよね、というはなし
- あとはキュー
 - Q4Mとか (GREEではMyISAMでなんとか)

- 半分くらい(場合によってはもっと)key-valueストレージでいいんじゃないか？
 - どのみちパフォーマンス上の問題で複雑なことはできないことが多い
 - 集計などはどのみち別のデータベース
 - プロダクションサーバから(場合によってはtriggerしたり集約しながら)データ回収/レプリケーション
- そこでmemcachedですよ！

- keyにホワイトスペースを使えない
 - binary protocolにすれば大丈夫そう
- サーバ増減|up / down時のコストが高い
 - 運用をミスするといきなり100%キャッシュミスしたり
 - repcachedでがんばればなんとかなる...かも
- たまにはディスクに書きたいんだけど
 - pluggable storageになれば改善されるらしいです
- クライアントライブラリありすぎ
 - しかも実装ばらばら
 - 多言語で利用するときは結構こまります
- 最大接続数くらいはSIGHUPで変えたい...

- とにかく気楽に使える
 - キャッシュとしてなら (memcachedだし)
 - サイズがあふれるとかあまり気にしない
 - たまには気にする
 - MySQLをon memoryで動かしていると結構気を遣う
- 速い
 - qpsはまだまだ改善の余地がありそう (cf. TT)
- 実績あり
- クライアントライブラリいっぱい

- セッション情報
- CSRFトークン
 - その他リクエストをまたぐトークンなど
- 各種新着情報 (GREEなら各ユーザのホーム)
- memcached for MySQL
- その他SQLの結果を適当に
 - などなど「信頼性は落ちるが適当に使えるキャッシュクラスタ」として無しでは生きていけません

- レプリケーションしたくなります
 - -> repcached
- ディスクに書きたくなります
 - -> Tokyo Tyrant
 - -> memcachedb
- なんかもっと
 - -> flare
- それぞれの特性に応じた用途を

- 昨日(11/25)リリースしました
- コンセプトは「heavy weight memcached」です
 - なんかもっと...
- 99.9%(くらい)memcachedプロトコル互換です
 - flush_allのexpireのみ未サポート

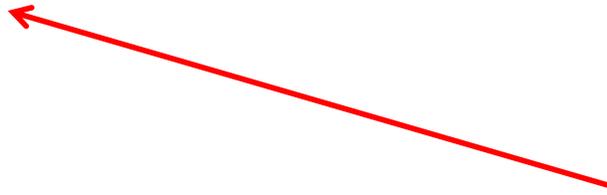
- 1台だと普通のmemcached(但しdisk based)
 - thx to Tokyo Cabinet



flare



クライアント



- 接続が増えてきたらslaveを追加できます(動かしながら) それってrepcached



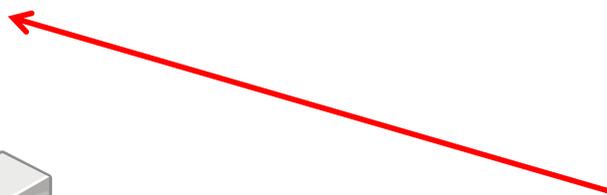
flare(master)



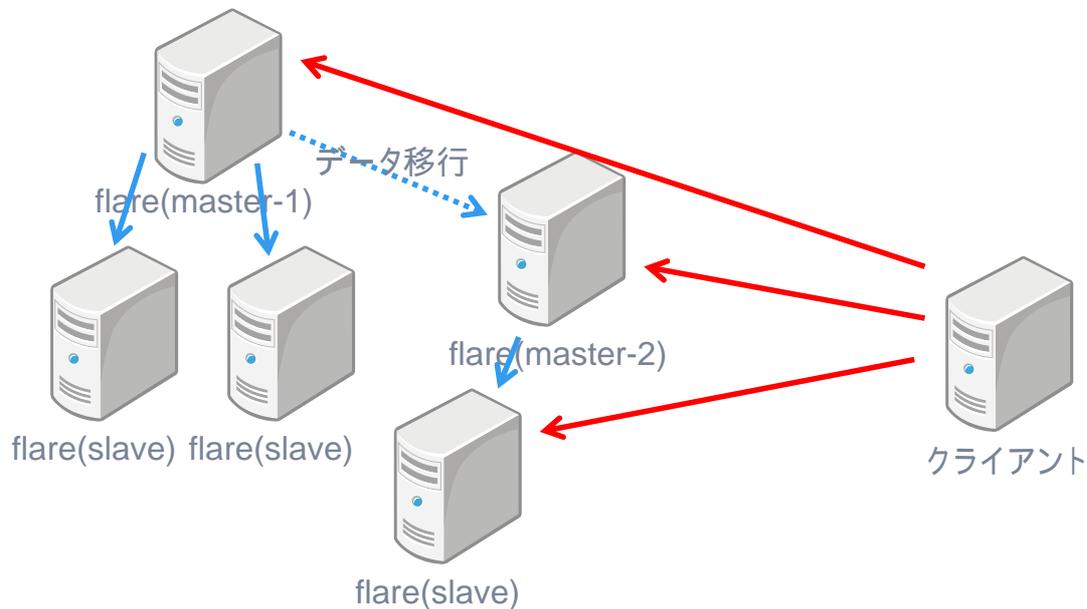
flare(slave) flare(slave)



クライアント



- データが増えてきたらmasterを追加できます (動かしながら) - オリジナル



12

- その他いろいろ
 - ノード管理します (落ちたら外したり、slaveをmasterに移行したり)
 - リクエストプロキシします
 - どこにつないでも同じ結果が返ってきます
 - プロキシするだけのノードのもあります
 - パフォーマンスはmemcachedより悪いです
 - 実績はこれから☺
 - GREEでも使ってみます

13

- MySQLをslaveにもちたい(binlogを出力させたい)
 - 集計がすごい楽になりそう
- FUSEでファイルシステムを実装してみるとよいのではないか？
 - B+Treeならいけそうな気がする
 - (実装にもよりますが)信頼性は少し落ちるけど、細かいファイルを大量におきたい場合などに使えそう

Thank You