Ad Serving in Erlang

Bob Ippolito
Mochi Media, Inc.

Erlang User Conference 2008 - Stockholm

November 13, 2008

Author: Bob Ippolito

Date: November 2008

Venue: Erlang User Conference 2008

What's MochiAds?

MochiAds:

- Monetization platform for Flash game ecosystem
- Advertising solution for game developers
- Revenue share and distribution for publishers

Road to Erlang

MochiBot:

- (Originally) Python w/ Twisted
- Fast, but not fast enough (CPU bound)
- Wanted easy multi-node distribution

Previous Experience

Python:

- Non-blocking sockets are tedious
- Threads are too heavy

C/C++:

- Too low-level

Why Erlang?

- Performance
- Concurrency
- Distribution
- Fault Tolerance

(Bad) Benchmarks

ab -c 50 -n 10000 localhost

- Apache: 1x
- Twisted: 1.12x
- mochiweb: 2.5x
- nginx: 3.9x

Erlang at Mochi Media

MochiAds, MochiBot

- High-performance HTTP servers
- Ad targeting
- Real-time analytics
- Social gaming
- Lots of internal use

MochiAds Service

- Front-end
- Data warehouse
- Ad server

What's Not Erlang

Front-end:

- Python and PostgreSQL

  Data warehouse:

- Python and Vertica

Ad Serving Platform

- Juniper Routers
- Cisco Switches
- OpenBSD load balancer
- Nginx HTTP load balancer (Linux)
- Erlang/OTP R12B-3 (Linux)

Ad Server Stack

- Erlang/OTP R12B-3
- mochiweb (http)
- egeoip (geolocation)
- eswf (SWF file format)

Ad Server

- Gather information about client
- Choose an ad
- Log impression data
- Log click data, redirection

Gathering Client Info

- Mostly client-side Flash code
- SharedObject (like cookies)
- Feeds targeting decisions for ad choice

Choose an ad

- Fold over in-memory data structure
- Filter out ads that don't match targeting info
- Weight the rest
- Choose a random number; $0 <= N < \text{sum(Weights)}$

Log impression data

- Validation
- Stream to hourly disk log
- Increment counters in RAM db

Log click data

- Validation
- Stream to hourly disk log
- Increment counters in RAM db
- Redirect (HTTP 302) to destination URL

Cheap Tricks

- Ad request data in URL
- Long-term state in SharedObject

Short-term Feedback Loop

- Client state
- RAM db counters from previous serves that day

Analytics Feedback Loop

- ETL processes adjust per hour
- Campaign weights, budget adjustments, etc.

Lessons Learned (part 1)

- Network partitioning sucks
- Network partitioning sucks
- Network partitioning sucks

Lessons Learned (part 2)

- pg2 is broken
- mnesia is too slow (for us)
- Inter-node distribution protocol can be flaky
- Erlang open source not always robust
- Lists are not a good data type for strings

Favorite Erlang Features

- Module reloading
- Pattern matching
- Binaries
- Lightweight processes
- Concise but not cryptic

More Erlang at Mochi Media

- MochiBot
- MochiCrypt
- MochiScore
- IRC bot
- SVN deployment
- Monitoring system
- Node discovery