Tunebot in the Cloud

Arefin Huq
16 Feb 2010
What is Tunebot?
What is Tunebot?

http://tunebot.cs.northwestern.edu

• Automated online music search engine for query-by-humming (QBH).
• Users sing or hum tunes to search.
• Queries are matched against other sung examples that have been contributed.
• Current DB: 8200 examples of 2700 songs
What is Tunebot?

• A project of the Interactive Audio Lab led by Prof. Bryan Pardo (EECS)
  http://music.cs.northwestern.edu

• Single-machine locally-hosted installation

• PHP/Flash/iPhone front-end running on Apache

• Java/MySQL back-end running as a Tomcat servlet
Tunebot Traffic
(Source: Google Analytics)
Goals

• Improve query response time to < 1 sec
  – Typical query takes 5 seconds to complete
  – Computation is linear in DB size

• Handle larger database
  – DB expected to grow to “critical mass” of 10K

• Adapt to growing and varying load
  – Handle traffic spikes
Considerations

• Contributions should be visible immediately to the contributor.
• Contributions and queries are valuable.
• Deployment and maintenance must be easy.
• Cost
• Revenue
• Research
Proposals

• Fully locally-hosted cluster

• Fully Cloud-based deployment

• Hybrid local-Cloud deployment
  – front-end local, back-end Cloud
  – full local deployment that scales to Cloud with demand (dynamic provisioning)
Project Plan

• Port to Linux (done)
• Load testing framework (in progress)
• Deploy single instance in Cloud (in progress)
• Load-balancing front-end with replicated DBs
  – handle many multiple users simultaneously
• Parallelized queries
  – speed up response for a single user
• Dynamic provisioning
Issues

• Existing code is at advanced prototype level
  – Numerous implementation details have delayed progress to date.
• Tight coupling between front-end and back-end, front-end and database.
• Maintaining database integrity