RDBMS Losing Workloads in the Cloud
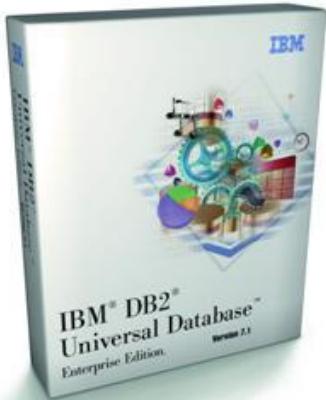
SMDB '09 Closing Panel

James Hamilton, 2009-03-29

VP & Distinguished Engineer, Amazon Web Services

James@amazon.com

web: http://mvdirona.com/jrh/work

blog: http://perspectives.mvdirona.com
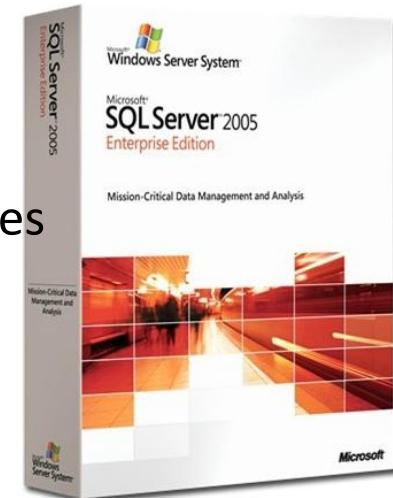
Cloud RDBMS Usage Looks Good

- Hotmail:
  - Over 300M users & >2B (non-spam) messages/day
  - SQL on every BE node of 10k farm
- Facebook:
  - 1,800 MySQL Copies
- Windows Live ID:
  - 420M+ IDs and well over 1B authentications/day
- But … all the complexity is "above" the RDBMS
  - Partitioning & partition management done above RDBMS
  - Many of these workloads could run on simple ISAMS
  - Many new workloads are going non-relational

Many Interesting Workloads Leaving RDBMS or have Left

- Analysis Clusters:
  - Map Reduce (MapReduce, Hadoop, & Cosmos)
  - 2k to 5k node clusters at Google, Yahoo, Microsoft
  - Usage spreading to private farms industry wide (oil/gas, fin, pharma)
  - Cloudera offering MapReduce as a service
- Caching:
  - MemcacheD & countless similar internal solutions
- Scale First, Simple First, & Combinations:
  - BigTable, MemcacheD, Amazon SimpleDB, Facebook Cassandra, Microsfot Azure, BerkelyDB, HBase, Hypertable, CouchDB, MemcacheDB, Scalaris, ...
- For many installations, the most mission critical data-intensive applications don't involve RDBMS

Why the RDBMS Exodus?

- Failure to scale

- Excess administrative complexity

- Resource intensive due to monolithic delivery of un-needed features

- Unpredictable response times

- Opaque failure modes

- Access patterns excessively random

- Slow to evolve to new workload patterns

What to do?

- RDMBS still important but giving many new workloads
- Support simple MemcacheD-like caching models
- Support rich execution models
  - Executing arbitrary analysis code (e.g. MapReduce)
    - E.g. Dewitt's Clustera work
  - No "load first" requirement
- Support profiles that load only code needed for task
- Embrace Recovery Oriented Computing mgmt model
  - Highly monitored to detect failure
  - Partitioned & redundant to scale & operate through failure
  - On failure: restart, reboot, re-image, then replace
- Partitioning, redundancy, & monitoring key to auto-admin

More Information

- These slides:
  - http://mvdirona.com/jrh/talksAndPapers/JamesHamilton_SMDB_Panel.pdf
- Designing & Deploying Internet-Scale Services:
  - http://mvdirona.com/jrh/talksAndPapers/JamesRH_Lisa.pdf
- Recovery-Oriented Computing:
  - http://roc.cs.berkeley.edu/
  - http://www.cs.berkeley.edu/~pattrsn/talks/HPCAkeynote.ppt
  - http://www.sciam.com/article.cfm?articleID=000DAA41-3B4E-1EB7-BDC0809EC588EEDF
- Autopilot: Automatic Data Center Operation:
  - http://research.microsoft.com/users/misard/papers/osr2007.pdf
- Perspectives Blog:
  - http://perspectives.mvdirona.com
- Email:
  - James@amazon.com