Highly available database clusters with JDBC

emmanuel.cecchet@emicnetworks.com
Cons

- Cost
- Scalability limit
Scaling the database tier – Alternative 2 (shared disks)

- Cons
  - still expensive hardware
  - availability

Another well-known database vendor
Database clustering middleware

- scalability
- high availability
- without modifying the client application
- database vendor independent
- on commodity hardware
RAIDb concept

- Redundant Array of Inexpensive Databases
- RAIDb controller
  - gives the view of a single database to the client
  - balance the load on the database backends
- RAIDb levels offers various tradeoff of performance and fault tolerance
**RAIDb levels**

- **RAIDb-0**
  - partitioning
  - no duplication and no fault tolerance
  - at least 2 nodes

SQL requests

**RAIDb controller**

- table 1
- table 2 & 3
- table ...
- table n-1
- table n
RAIDb levels

**RAIDb-1**
- mirroring
- performance bounded by write broadcast
- at least 2 nodes

```
SQL requests

RAIDb controller

Full DB  Full DB  Full DB  Full DB  Full DB
```
**RAIDb levels**

**RAIDb-2**

- partial replication
- at least 2 copies of each table for fault tolerance
- at least 3 nodes

![Diagram of RAIDb controller with SQL requests to Full DB, table x, table y, table x & y, and table z](http://c-jdbc.objectweb.org/ - c-jdbc@objectweb.org)
Middleware implementing RAIDb
- 100% Java implementation
- open source (LGPL)

Two components
- generic JDBC driver (C-JDBC driver)
- C-JDBC Controller

Read-one, Write all approach
- provides eager (strong) consistency

Supports heterogeneous databases
architectural overview

Application server

MySQL JDBC driver

JVM

C-JDBC controller

MySQL database

MySQL

C-JDBC
Inside the C-JDBC Controller

XML configuration file

C-JDBC Controller

Virtual database
- Authentication Manager
  - Recovery Log
  - Scheduler
  - Request Cache
  - Load balancer

Request Manager

Database Backend
- Connection Manager
  - MySQL JDBC driver

MySQL

Virtual database
- Authentication Manager
  - Recovery Log
  - Scheduler
  - Request Cache
  - Load balancer

Request Manager

Database Backend
- Connection Manager
  - Oracle JDBC driver

Oracle

Sockets

Administration console

XML configuration file

JMX

MySQL

MySQL

MySQL
Scaling existing databases

- add open source databases for more fault tolerance and better performances
- rules for on-the-fly query rewriting to handle heterogeneity
TPC-W benchmark
(Amazon.com)

Nearly linear speedups with the shopping mix
Horizontal & Vertical scalability

- adapt and evolve according to user needs
- advanced caching functionalities
- dynamic adding of database backends
- fault tolerant transaction log
- integrated backup and checkpointing
- grid & edge-side servers support

http://c-jdbc.objectweb.org/ - c-jdbc@objectweb.org
Administration

➡ JMX Graphical Console
- cluster administration
- backup/checkpoint mgt
- cluster monitoring
- SQL profiling
- integrated SQL console (iSQL)

➡ Administration API
- integration in any existing administration infrastructure
- JMX notifications for all internal events
- JMX MBeans for monitoring
J2EE end-to-end reliability

- **ObjectWeb exclusivity**
  - no SPOF, transparent failover

- **On-demand availability**
  - scalability, flexibility

- **On-demand scalability**
C-JDBC today

➡️ Web site
- ~200,000 hits/month
- >28,000 downloads
- EU (18 countries) 36%, US 28%, Japan 12%, China 5%, Canada 4%, Australia 4%, India 3%, Brazil 2%, ...

➡️ Community
- 27 committers both industrial & academics
- c-jdbc@objectweb.org: >200 subscribers, 200-300 msgs/month
- translation in japanese, italian, chinese, turkish, french, german

➡️ RPM on JPackage.org
➡️ Professional support by Emic Networks
What’s next?

- **Tribe** (.objectweb.org)
  - replacement for JGroups
  - uniform total order broadcast optimized for clusters

- **LeWYS** (.objectweb.org)
  - hardware and software monitoring
  - monitoring repository

- Distributed query execution

- Optimized support for edge-side servers and interconnected clusters
Demos

Tuead 18

- INRIA booth: 12.00 to 1pm
- ObjectWeb booth: 2pm to 2.30pm

Wednesday 19

- INRIA booth: 10.30am to 11am
- ObjectWeb booth: 1pm to 1.30pm
Q&A

Thanks to all users and contributors ...

http://c-jdbc.objectweb.org
Bonus slides
Current limitations

- JDBC only
- Distributed joins
- Updatable ResultSets
- XA support through XAPool only
- transparent controller failover not supported when using horizontal scalability with JGroups
- network partition/reconciliation not supported