MySQL and Scalable databases

David Axmark
CoFounder

MySQL AB, Creators of MySQL
CHEP 06 Mumbai 2006-02-14

Why MySQL: Some Stats

- Good technology used in many research projects (bioinformatics, physics)

- We estimate that **>100 million copies** of MySQL have been distributed through our Web site & operating system distributions. (40k Day!)

- Lots and **lots of books** on MySQL in for example English (Amazon gives 196 in English), German (118), French (146), Japanese (46), Chinese, Russian, Portuguese, Norwegian, Malay and even Swedish.

- More than 3100 projects on SourceForge that leverage MySQL (runner up has 650)

- Over a thousand partners

Platforms

- Linux
  - RedHat, Suse, Debian, Feodora, Ubuntu, WindRiver, MontaVista
- UNIX
  - Solaris, HPUX, AIX
- Windows
  - NT, Win2k, XP
- MacOS X
- {Free,Open,Net}BSD
- Novell Netware
- OpenVMS
- QNX

- Intel
  - Xeon [IA32]
  - Nacona
  - Itanium [IA64]
- AMD
  - Opteron [64]
  - Athlon [32]
- IBM
  - PowerPC [32 & 64]
- Sun
  - Sparc [32 & 64]

First 64bit MySQL in March 2000. If your code is good 64bit means a recompile!

All compiled from ONE source tree.
Code should be written with portability in mind from the beginning!

Connectors

MySQL AB develops the most important drivers in house:

- Connector/**C**

- Connector/**ODBC**

- Connector/**J** (Java JDBC)

- Connector/**MXJ** (fully embedded Java version)

- Connector/**.NET** (& Mono)

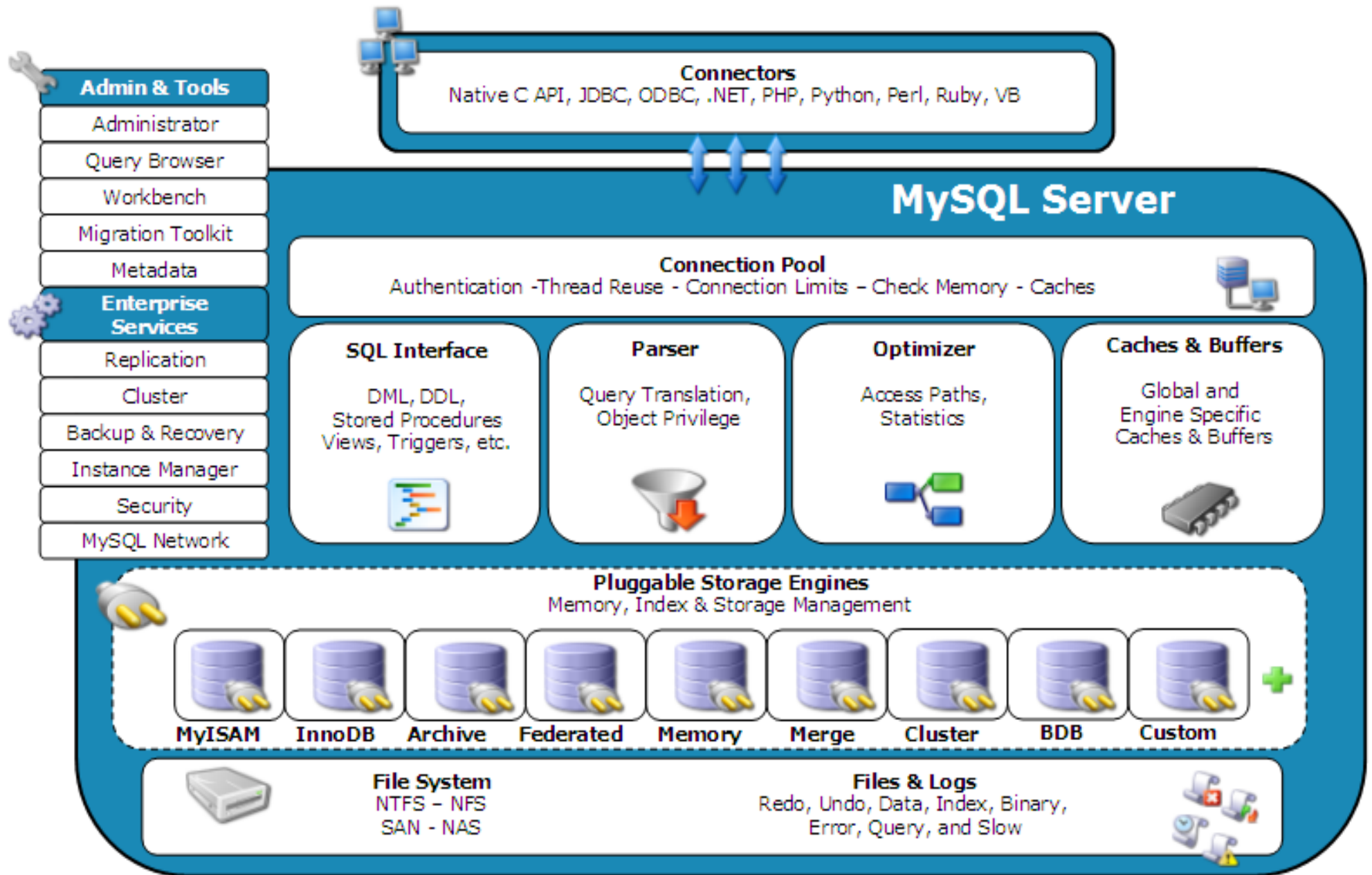- Community with help from MySQL AB:

  - **PHP**

Languages

- C
- C++
- C#
- Java
- PHP
- Perl
- Python
- Delphi
- Objective C
- Visual Basic
- Smalltalk
- TCL

- Ruby
- Fortran
- Pascal
- ADA
- Lasso
- Pike
- Rexx
- Dylan
- Common Lisp
- Scheme
- Gauche
- Guile

- Mathlab
- Eiffel
- Haskell
- Erlang
- Curl
- Forth
- Slang
- LUA
- OLEDB
- Active X
- ODBC

And even Cobol!

The community are always adding more languages!

Embed MySQL in Java Applications

- Connector/MXJ:

  - Running MySQL native server inside a Java application

  - Deployment, lifecycle and configuration management handled all in Java

  - Deploys platform-specific MySQL binary automatically

MySQL Internal Architecture

Admin & Tools
- Administrator
- Query Browser
- Workbench
- Migration Toolkit
- Metadata

Enterprise Services
- Replication
- Cluster
- Backup & Recovery
- Instance Manager
- Security
- MySQL Network

Connectors
Native C API, JDBC, ODBC, .NET, PHP, Python, Perl, Ruby, VB

MySQL Server

Connection Pool
Authentication -Thread Reuse - Connection Limits – Check Memory - Caches

SQL Interface
DML, DDL, Stored Procedures Views, Triggers, etc.

Parser
Query Translation, Object Privilege

Optimizer
Access Paths, Statistics

Caches & Buffers
Global and Engine Specific Caches & Buffers

Pluggable Storage Engines
Memory, Index & Storage Management

MyISAM	InnoDB	Archive	Federated	Memory	Merge	Cluster	BDB	Custom

File System
NTFS – NFS
SAN - NAS

Files & Logs
Redo, Undo, Data, Index, Binary, Error, Query, and Slow

Standard Storage Engines

- Choose storage engine that matches your application
- Easily mix and match in a single application (but minimize!)

	MyISAM	Memory	InnoDB	NDB/Cluster
Usage	Fastest for read or write mostly apps	In-Memory (HEAP) storage	Fully ACID compliant transactions	High availability cluster, In-memory storage, ACI(D) transactions
Locking	Large-grain table locks Delayed Table Inserts	Fixed length record / covering Indexes	Multi-versioning, Row-level locking	Row-level locking
Indexing	B-tree / R-tree / Full Text / Hash	Hash / B-tree indexes	B-tree indexes	Hash / B*-tree indexes
Storage	Durable / Table Recovery	No disk I/O or persistence	Durable / Log recovery	2x RAM Durable / Log recovery

MySQL Cluster (NDB)

- Distributed in memory storage engine with

    - **Fault Tolerance:** shared nothing architecture

    - **High Availability:** auto failover (five 9's of availability)

    - **Scalability:** Scale by adding more commodity machines to a cluster

    - **High Performance:** Really really fast (many 100000 ops per second) for primary key lookups (mixed read and write). Up to millions of queries per second using low level C API and high end hardware

    - **Simplified applications:** For the application MySQL Cluster is a just a table. In the case of failure you reconnect to another MySQL Server and immediately see the same data

- Archive: For logging data where you don't delete or update.

  – Compressed data on disk

  – High throughput for insert (> 10000/sec on a normal machine) & table scans

  – Useful for logging and auditing (there changes should not be allowed)

- [Your custom Storage-Engine goes here]

MySQL 5.0

- Stored Procedures

- Triggers

- Views

- XA (distributed transactions)

- Cursors (read only, forward scrolling)

- Strict Mode (classical DB error handling)

- Information_Schema (Data Dictionary)

- Precision Math (56 digits of precision, ext with recompile)

- Many additions to out optimizer

5.0: 'Small' things also gets added!

Extension to LOAD DATA for doing transformation/calculation at the time you load the data.

```
LOAD DATA INFILE 'file.txt' INTO TABLE t1
    (col1, @var1, @var2)
  SET col2 = @var1-@var2, col3 = @var2;
```

Upcoming Features

- **Partitioning** (in the current alpha release)

    - Parallel query (later)

- MySQL Cluster with **Disk data** (for non indexed columns)

- **Replication**

    - **Row based** (physical) replication (MySQL default is logical)

    - Multi source replication

- Hash & Merge joins

- XPath (XML) support

- Global Backup API

- Even better optimisation of sub-queries etc

MySQL Graphical (GUI) Tools

- ## Available Now!

  - MySQL Administrator (Win, Linux, MacOS)

  - MySQL Query Browser (Win, Linux, MacOS)

  - MySQL Migration Toolkit (Win, Linux)

    - Plug-in Architecture for sources (Java/JDBC)

- ## In Development

  - MySQL Workbench

  - MySQL Cluster & Replication Manager

MySQL Administrator

MySQL Query Browser

MySQL Migration Tool

MySQL Workbench

- Good bug reports since bugs gets fixed

    - **Repeatable bug reports are as valuable as code!**

- Lots of testing of **all** code. All features available for all!

- Freedom & Independence!

    - You have the ultimate documentation, the source!

- Security is not by obscurity, No hidden hooks in the code
- Lots of Eco system code gets written by the community
- We can hire people who already know the code
- Result: **Low Total Cost of Ownership**

A few example references

- **Cox Communications**
  - Data warehouse for >6.3 million cable users. > 3,600 tables with > 4 billion rows. Insert speed 4 million inserts every 2 hours.

- **Yahoo**!
  - Yahoo has over 5000 MySQL server to run over 200 web properties

- **Google**
  - MySQL is used heavily in the ad system

- **Sabre** (travelocity.com and other travel booking sites)
  - Runs 45 4CPU servers with 16GB RAM for MySQL. Saved a few million USD.

- **Los Alamos National Labs, USA**
  - Searching 55 million scientific journal articles (7Tb)

- **SLAC** is researching MySQL for their multi peta byte database

Some more References

- Rakuten.co.jp, Slashdot.org, Freshmeat.net (Web)

- Cisco, Nokia, Ericsson, HP, Intel, Motorola (Embedded)

- Lloyds TSB Bank, HypoVereinsbank

- Human Genome Project, Sanger Institute, Cambridge, UK

- Sahana (disaster recovery system for the tsunami), Ensembl.org and Human Genome Project (used for cancer research), Wikipedia, Bugzilla, Craigslist, Feedster, Flickr, Freshmeat, LiveJournal, Neopets, Slashdot, SugarCRM, Technorati, Wordpress, Citysearch, CERN's ATLAS Experiment

No Software Patents!

- Software Patents are a threat against Free Software, Software innovation and Developing countries

- In Europe our side was successful and the SW Pat proposal was thrown out (a real thriller!)

- MySQL has been spending lots of cash and lots of Management time (CEO,VPs & Founders) fighting Software Patents

    – Other backers included RedHat

Amira   Maria