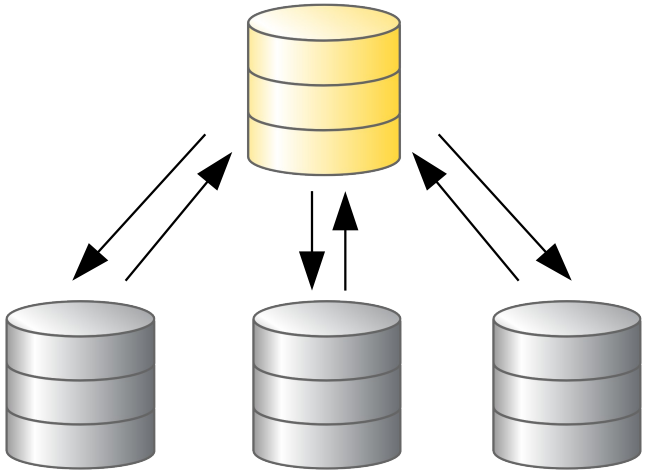
Distributed databases with MariaDB and Spider engine



Developments from E-Science @ AIP Potsdam

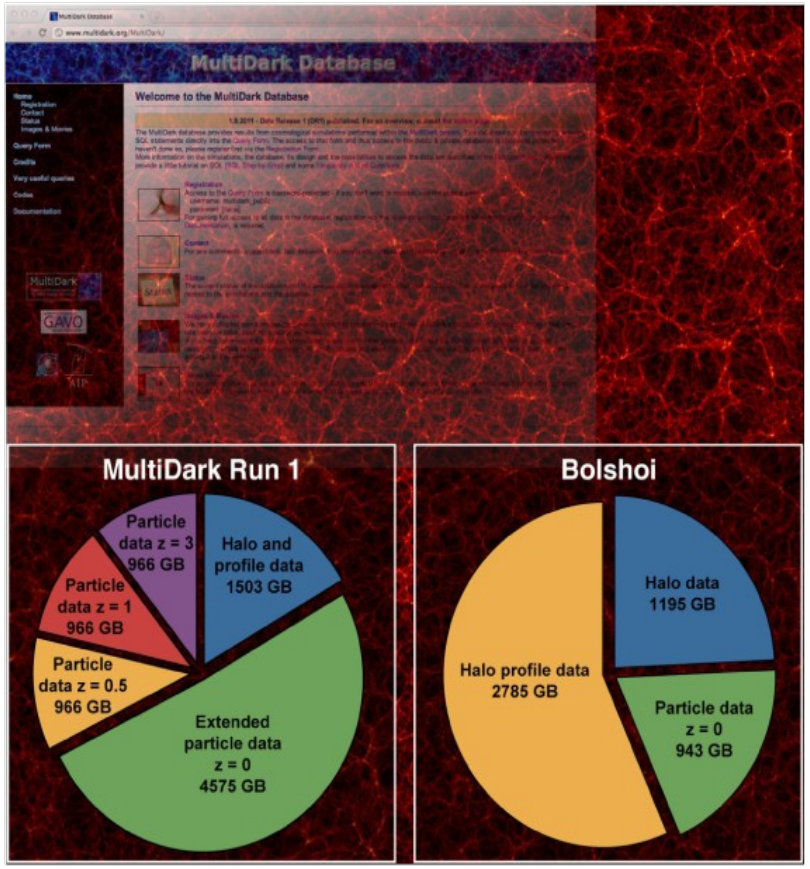
Kristin Riebe

German Astrophysical Virtual Observatory

Example: MultiDark Database

- Collaboration with Spanish MultiDark project
- cosmological simulations in a database
- 2 simulations uploaded (14 TB, 1.5 10^{11} rows)
- Webinterface: www.multidark.org

- > 150 registered users
  > 1.5 million queries in 3 years
  > 6 TB downloaded



Database server

- Current setup:
  - like (first) Millennium DB
  - 1 Microsoft SQL Server

- Issues:
  - **retrieval** times slow on full table scans (~ 30-40 min), cannot have index for every possible query
  - **index** on particle data (~ 10^{10} particles) takes ~ 1 week
  - **transaction logs** take time … (but useful for data integrity)
  - if multiple servers: need to buy expensive **license** (unless you know the right people ;-))

Database server

- Goal:
  - **speed-up** queries involving full table scans
  - want to serve simulations with even **more particles** (at least factor 10)
  - use only **open source** software (enable mirroring services without expensive licenses)

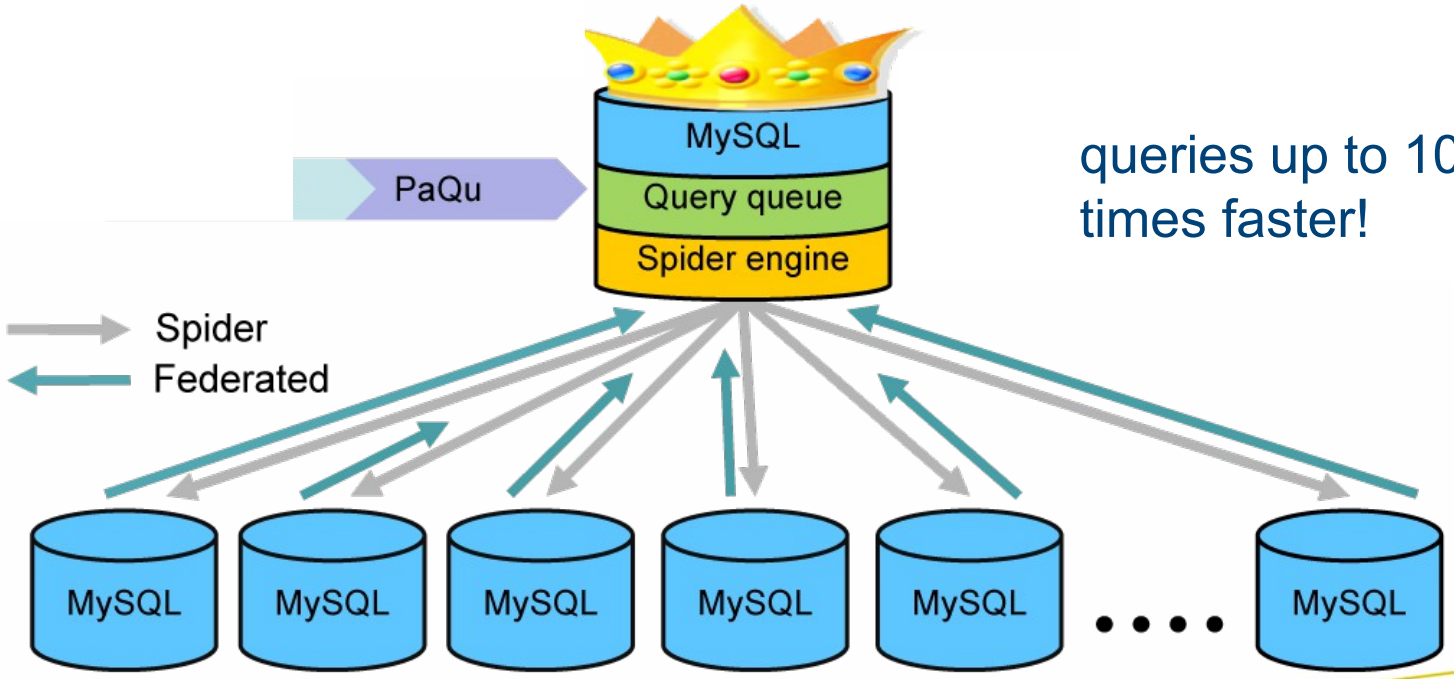
- Solution:
  - distributed data over multiple servers with MariaDB/MySQL + Spider engine

MySQL/MariaDB

- MySQL:
  - open source, plugin-system (C)
  - free choice of storage engine
  - MyISAM engine: no transactions
    (need fast select, data changes are rare)
- MariaDB:
  - spin-off of MySQL
  - developed by original MySQL-developers
    (left MySQL after it was taken over by Sun/Oracle)
  - “An enhanced, drop-in replacement for MySQL.“ (https://mariadb.org/)
  - => no difference in interface, just exchange the sources
  - advantage:
    - Spider engine by Kentoku Shiba included (for distributed data)
    - more community driven, support for community developments

Spider engine

- data tables partitioned, distributed over 10 nodes using Spider engine
- PaQu reformulates queries, head node sends them to nodes
- head node collects data via federated table



queries up to 10 times faster!

Additional developments

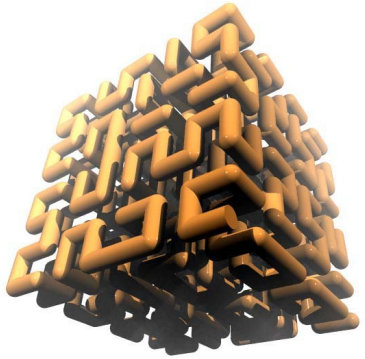
- PaQu:
  - reformulates queries, based on Shard-Query
  - e.g.: aggregate function: count
    - count on each node
    - sum on head node

- QueryQueue:
  - allow asynchronous jobs
  - plugin for MySQL, supports priorities
  - control number of executing jobs on server
  - jobs stored in user table for later retrieval

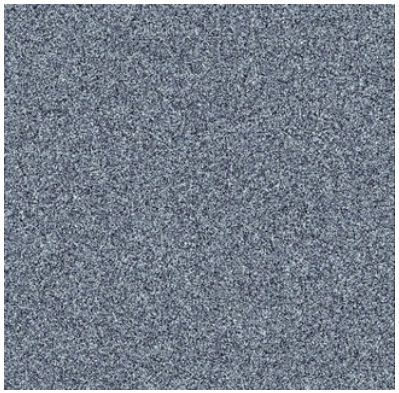
see https://github.com/adrpar/

Further MySQL plugins

- C-library libhilbert
  - For creating indexes of space-filling Peano-Hilbert curve in up to 20 dimensions



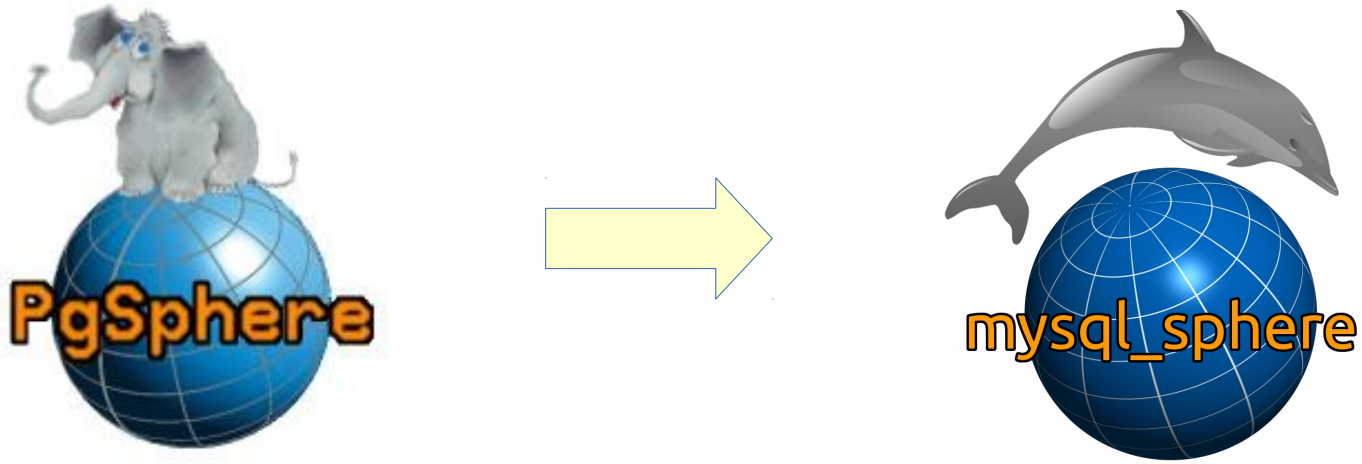
- MySQL sprng
  - Based on SPRNG library (www.sprng.org)
  - Implements several random number generators
  - Better random sampling for large numbers than with built-in function



see https://github.com/adrpar/

mysql_sphere

- Functions of pgSphere converted to mysql_sphere
- Allows queries on a spherical surface (cut outs, angle-ranges
- Especially important for observational databases



- … now also ported to SQLite!

see https://github.com/adrpar/

Data download: VOTable dump

- fork of mysqldump
- dumps VOTable format 1.3, ASCII or binary format, directly from MySQL database tables
- => especially useful for large tables, no additional conversion on server needed
- for ucds, units: a json-like comment string is required

```
CREATE TABLE foo (
     x DOUBLE COMMENT 'DQIMETA={"unit":"Mpc","ucd":"pos.cartesian.x"}',
     y DOUBLE COMMENT 'DQIMETA={"unit":"Mpc","ucd":"pos.cartesian.y"}',
     ...);
```

- Download from https://github.com/adrpar/mysqldump-vo

New portal: www.cosmosim.org

- with Spider nodes in background, PaQu, QueryQueue

Web application: Daiquiri

- Developed by Jochen Klar und Adrian Partl
- http://escience.aip.de/daiquiri/
- Web application for publishing data
- Modular, highly customizable
- Using PHP, Zend-framework
- Modern interface using bootstrap, jQuery
- Authentication, Query Interface
- Wordpress integration
- One code base to serve most needs, open source, (easily) extendable
- supports SAMP and UWS

Summary

- need to speed-up database queries
- solution using distributed data with MariaDB and Spider engine is working
- => queries scale nicely
- plugin development for MySQL/MariaDB in C possible, could even write own storage engine etc.
- => everything adjustable, open source
- => MySQL/MariaDB is an alternative to commercial databases that shouldn't be ignored