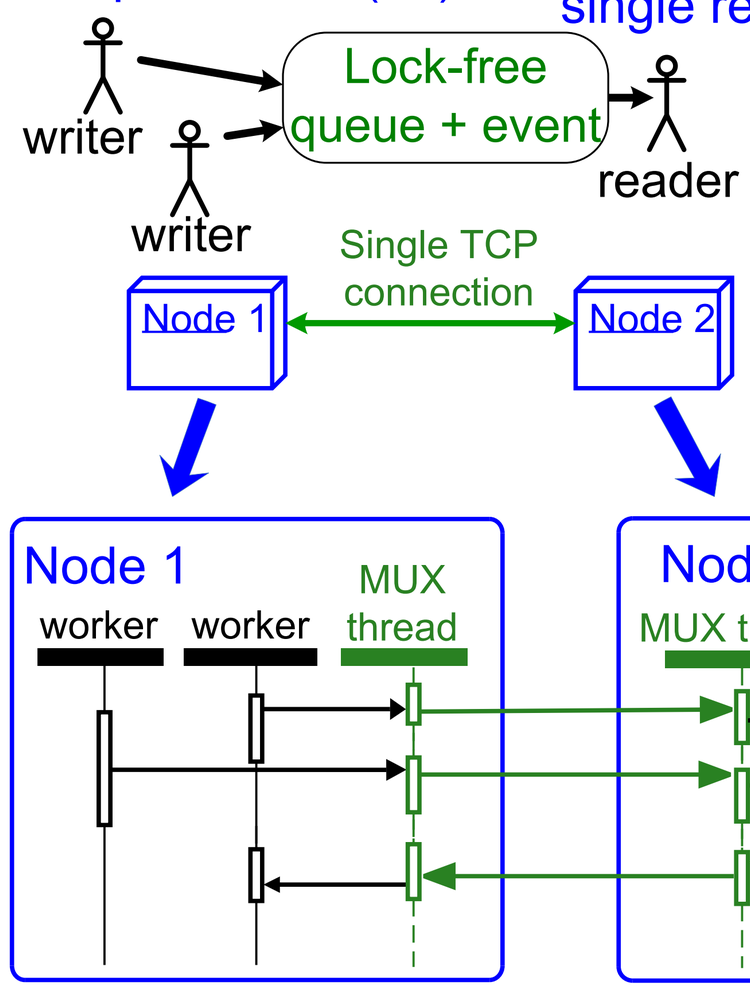
ACM SIGMOD 2010 Programming Contest
Distributed Query Engine

Saint-Petersburg University, Russia

Team SPBU: George Chernishev, Kirill Smirnov

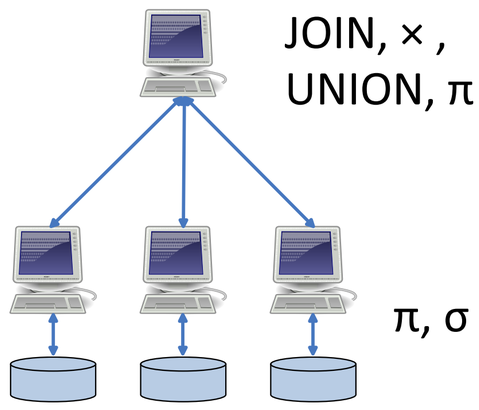
Multiple writers (50)



Event-driven single reader

- Single TCP connection
- Thread pool for indexed lookups
- Lock-free interaction
- Socket/event MUX

High-Level Aspects



- Pipelined query execution
- Master node is responsible for complex operations
- Slave nodes perform only scans and projections

Local Reduction

- Predicate clean-up
- Restriction propagation

Join Operation

- In-operation caching
- Sort one result, probe with another
- Tend to sort full result (to sort once)

Plan Generation

- Try to push down selects with indexed field
- Join indexed field with indexed (to fit in op cache)

Team dbis

DBIS Datenbanken und Informationssysteme

Technische Universität Kaiserslautern

Heuristics
- predicate pushdown
- operator selection

Statistics
- value distribution
- min/max
- duplicates

Partitioning
- number of partitions
- involved nodes

SELECT .. FROM .. WHERE

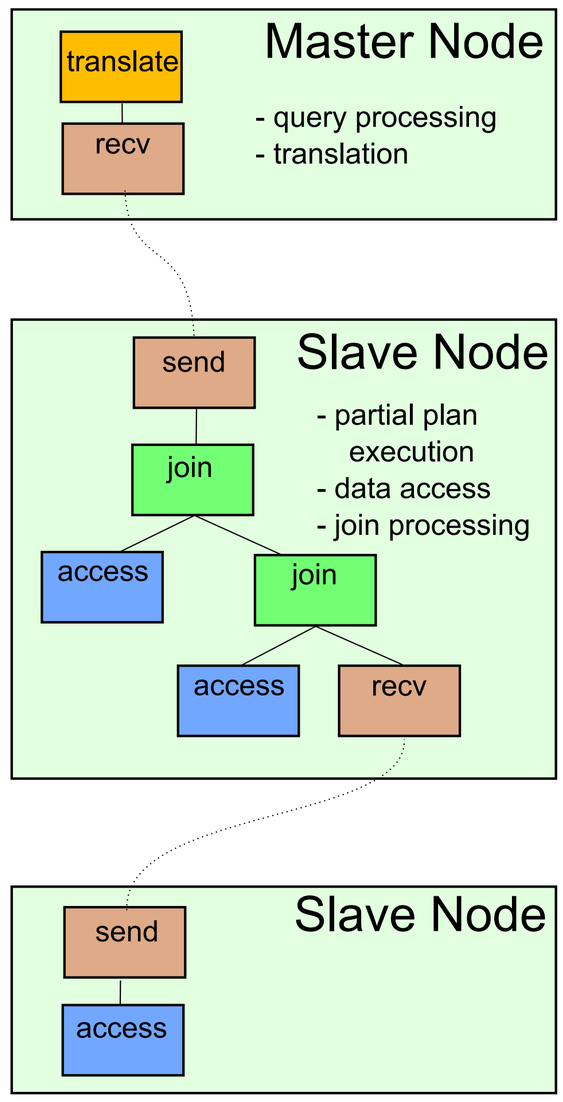
Search for "good" query plans

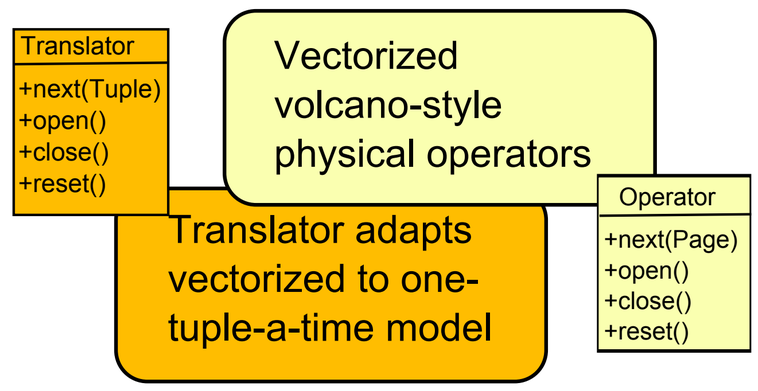
several candidates

Cost-based plan selection

best plan

Plan distribution and execution





- Block NestedLoop Join
  default if no other join fits
- PartitionedHash Join
  uses the hash-based index
- Merge Join
  if already sorted

- TBScan
  plain old table scanner
- TIDGrabber
  fetches rows from file by Tuple IDentifier
- IXScan
  evaluates a single index
- MultiIXScan
  handles multiple indexes in a single run

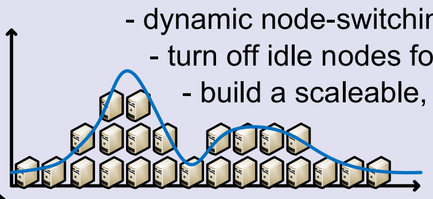
TCP sockets tuned for performance: NODELAY, TCP_CORK

Black Arts

a lot of nasty compile-time optimizations

Future steps

- move from read-only to write-supported DBMS
- recycle code-base for future project:
  an Energy-Proportional DBMS running on multiple nodes
  - dynamic node-switching wrt. performance needs
  - turn off idle nodes for power savings
  - build a scaleable, energy-aware database cluster



Daniel Schall
Volker Hudlet

Database and Information Systems
University of Kaiserslautern

{schall, hudlet}@cs.uni-kl.de
http://wwwlgis.cs.uni-kl.de

Contact

Supported by

NSF Microsoft amazon web services™ INRIA exalead YAHOO!