



# Mining for insight

Osma Ahvenlampi, CTO, Sulake

Implementing business intelligence for Habbo

**sulake**

everyone can play™

**HABBO**<sup>®</sup>

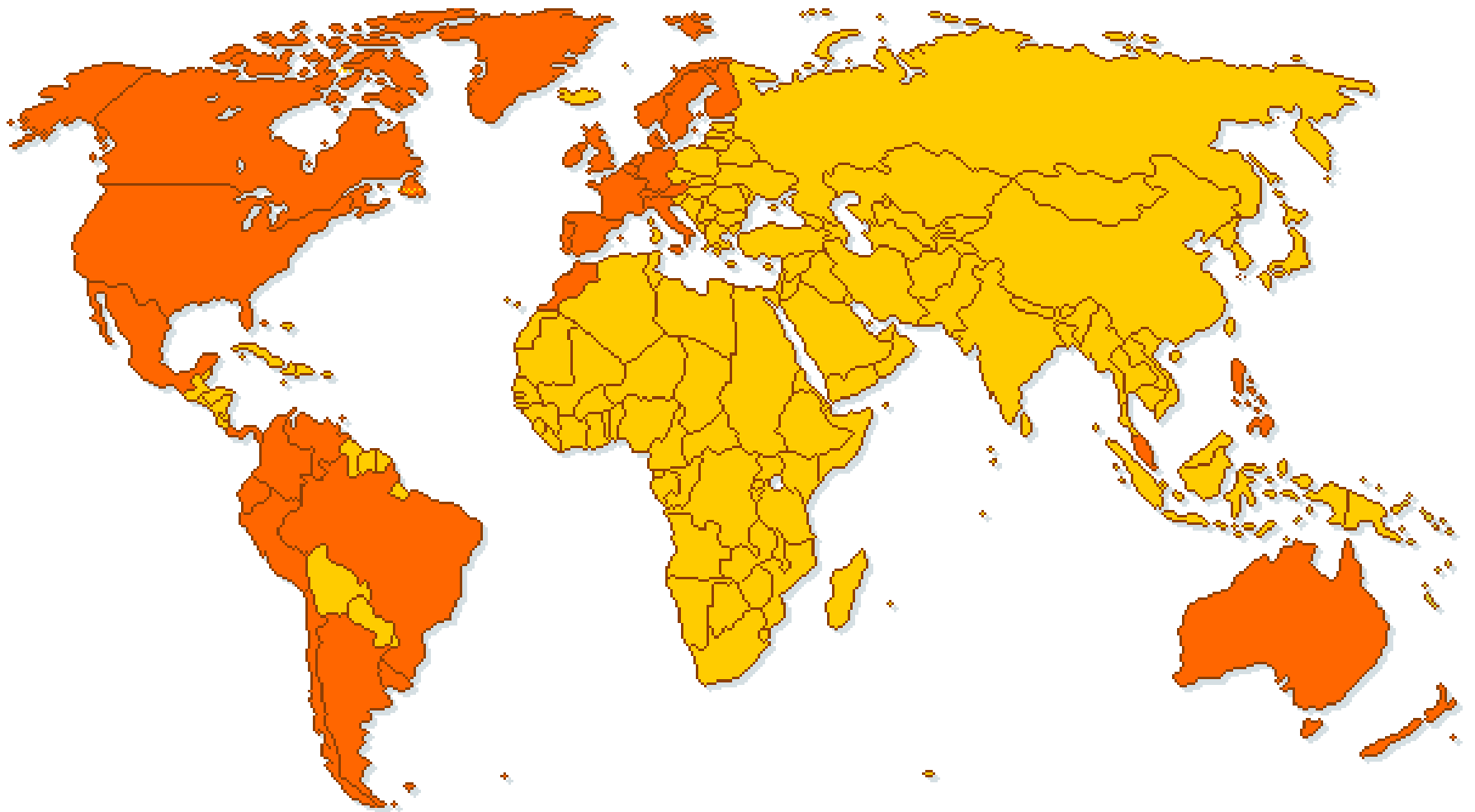
sulake

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Virtual world





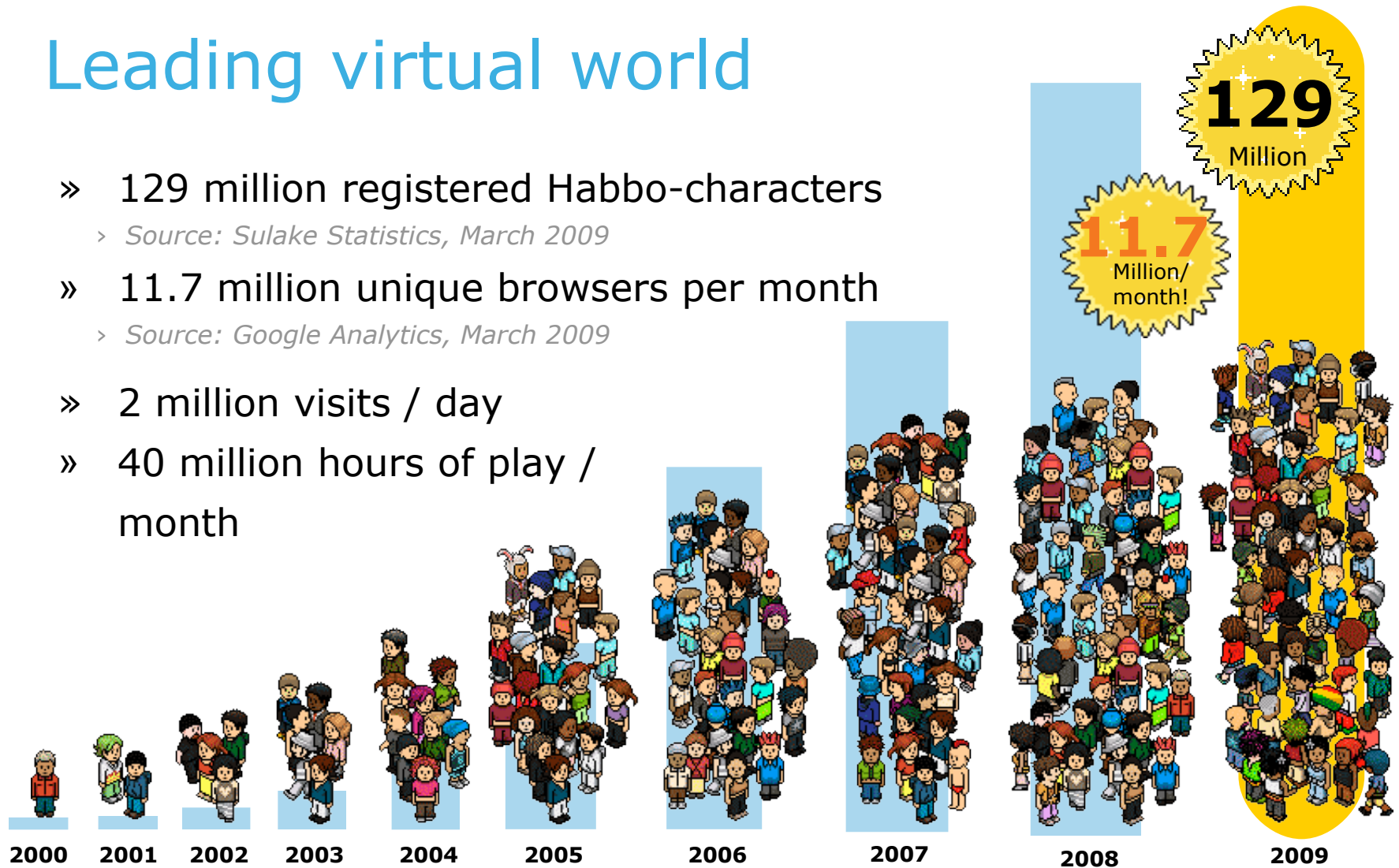
sulake

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

5  
everyone can play™

# Leading virtual world

- » 129 million registered Habbo-characters
  - > Source: Sulake Statistics, March 2009
- » 11.7 million unique browsers per month
  - > Source: Google Analytics, March 2009
- » 2 million visits / day
- » 40 million hours of play / month



# Overview

- Analytics approach and objectives
- Types of data processing
- Description of a solution for scaling event storage and analysis
- Observations about Infobright technology

# Background

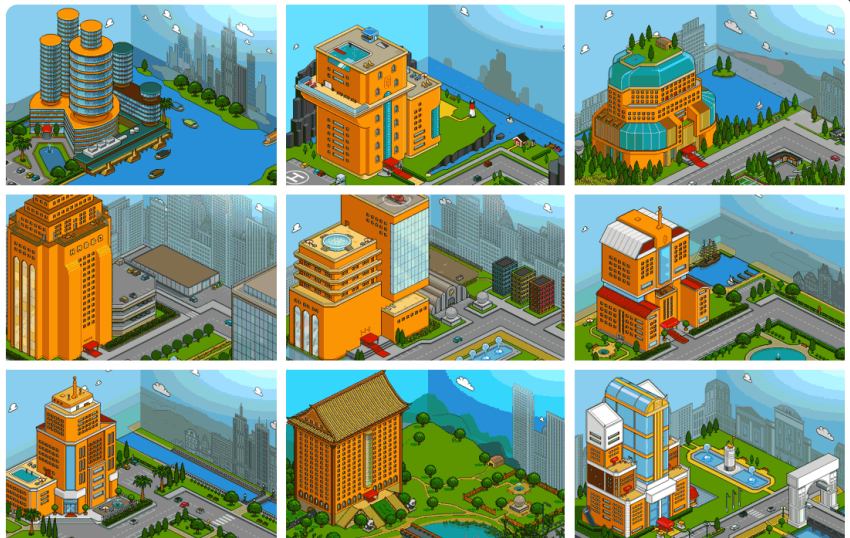
**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Scaling a virtual world

- Java code is “easy” to scale
  - Clustered, load-balanced process model on J2SE + open source stack
- MySQL not so much
- Local communities provide natural shards



# Data management in Habbo

- Several dozen DB servers
- Close to a hundred MySQL processes
  - MySQL on many-core hardware!
- Terabytes of managed data
  - Fragmented all over the place
  
- 3 million new user accounts monthly
- 2 million visits daily, average ~40 minutes
- Hundreds of interactions every visit
- Hundreds of millions of user-created “rooms”

# Analytics

**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

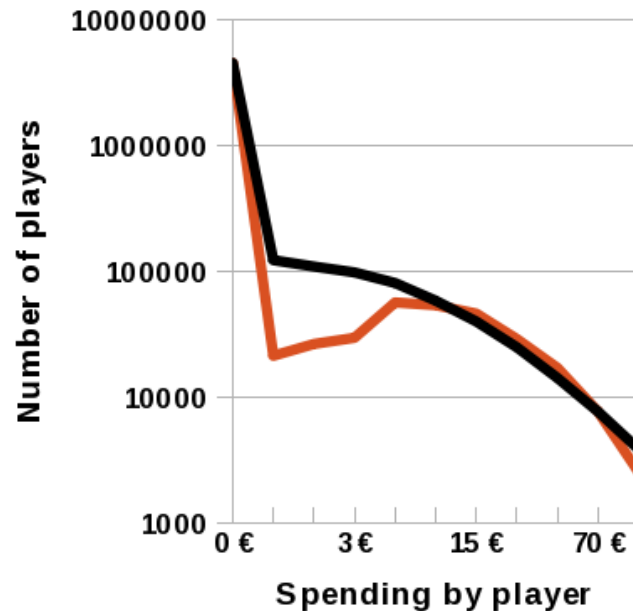
everyone can play™

# Business objective

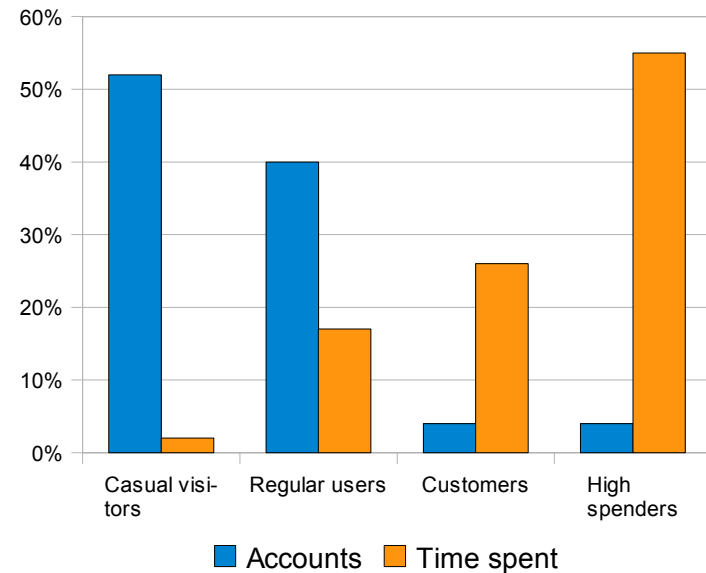
- Development process is iterative
  - Benefits from constant learning
- Requires data
  - Users, visits, meetings, purchases, trading, friends, events, activities, achievements, places, items and so on..
- Up-to-date management information
- Virtual worlds == virtual economies
  - Economies require oversight

# Examples of analysis

- Spending patterns  
<http://bit.ly/B8sg>



- Behavioral segments



# Managing data

How to operate, collect, and analyse data  
at Habbo's scale

**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Three types of data processing

## Real-time shared state

- In-memory data structures
- Game (business) logic
- Try to keep data footprint small

## OLTP

- Transactional integrity
- Persistent customer state

## Log files & analytics

- High-volume events
- Post-processing

# Event logs and analytics w/ MySQL

Recap of methods we've used over time



Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Large analytics tables and MyISAM

- Fast writes, as long as you don't maintain a lot of indexes
- Fast reads, but only if you do maintain a lot of indexes
- Terrible crash behavior
  - Have you ever tried to myisamchk a 1 TB table?
- Good for interim or throw-away buffers only

# Large analytics tables and InnoDB

- Ok for OLTP
- Crash-safe
- Pretty slow for batch loads, even after lots of tuning
  - Google and Percona patches help!
- Not good for complex tables with lots of indexes
- Horrible if you ever need to change your table schema
  - analysis databases change constantly!

## Take 2

- Must be able to load millions of rows every day
- Retaining billions of rows
- Schema evolves with new features and improving analysis
- Can't afford days of downtime for changes or maintenance

# Columnar databases

- Turn storage by 90 degrees
- Enables very wide tables and rapid access to narrow sets of columns
- Compresses well
- Perfect for data warehousing
- Not a new field, but enjoying a comeback
  - Expensive MPP solutions; Vertica, ParAccel..
- For MySQL, there's at least Infobright

# Applications for columnar storage

- Data Warehousing / Business Intelligence
  - Next-day results (typically)
  - Big storage, complex data models, lots of repetition
  - Analytic query performance
- Event log management
  - Not realtime, but as close to it as possible
  - Very high volumes, simple content
  - Long-term storage issue

# An approach to logging

## High-performance data collection

**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

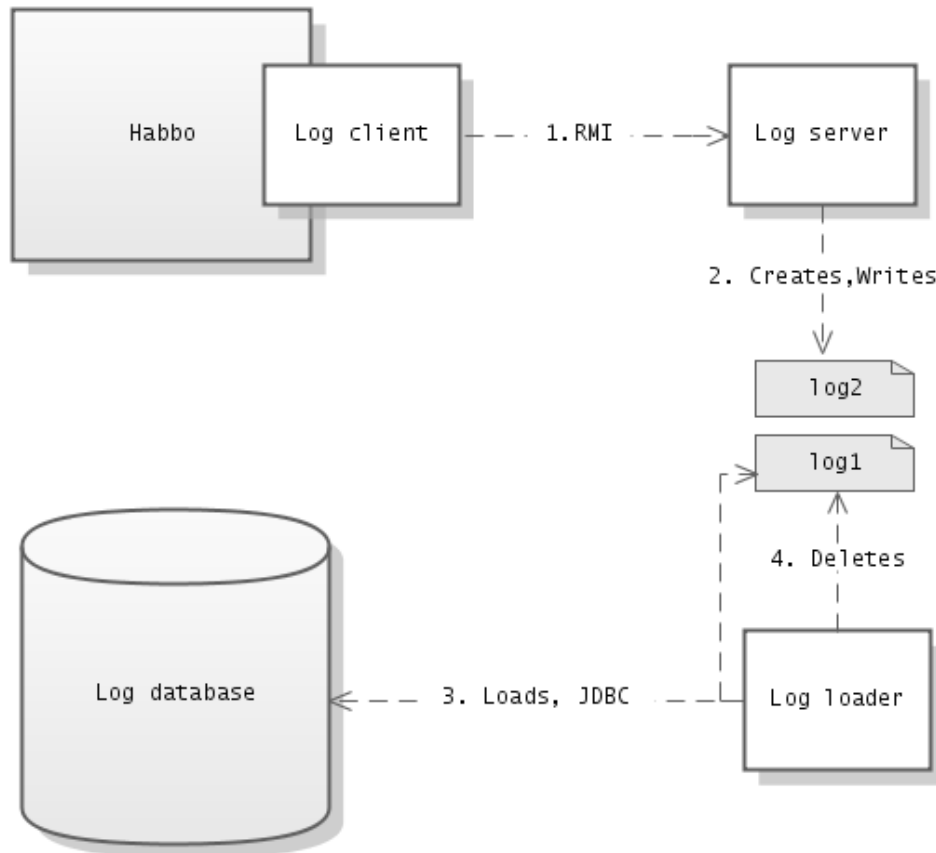
everyone can play™

# Structures of logfiles

- “Clickstream” event dataflows
  - Logins, logouts, messages, actions
  - For unstructured or semi-structured-data, Hadoop
- Simple structure, don't even try to be relational
  - Immediate output
  - Traditionally on small scale with text files or MyISAM
- Direct analysis is tricky
  - Long-lasting activities split to “begin” and “end”
  - What if there's an interruption?
  - Related events scattered around

# Overall architecture

<http://bit.ly/ice-logs>



1. Java clients buffer locally and sends batch data over RMI to log server
2. Log servers buffers to local interim flat file
3. Log loader takes new files and loads them to an Infobright ICE database
4. Files are then removed

Single-thread performance using low-end hardware:  
**100,000 processed log entries per second**

# Keep it simple

- Don't complicate event processing by making logging structures more complex
- Reasons to keep the format simple:
  - It's simple to implement
  - Eventually, scale will require it anyway
- Process the data to a richer structure asynchronously
  - Eg, using Hadoop

# From logging to data warehousing

Collection is nice, but using the data is nicer

**sulake**

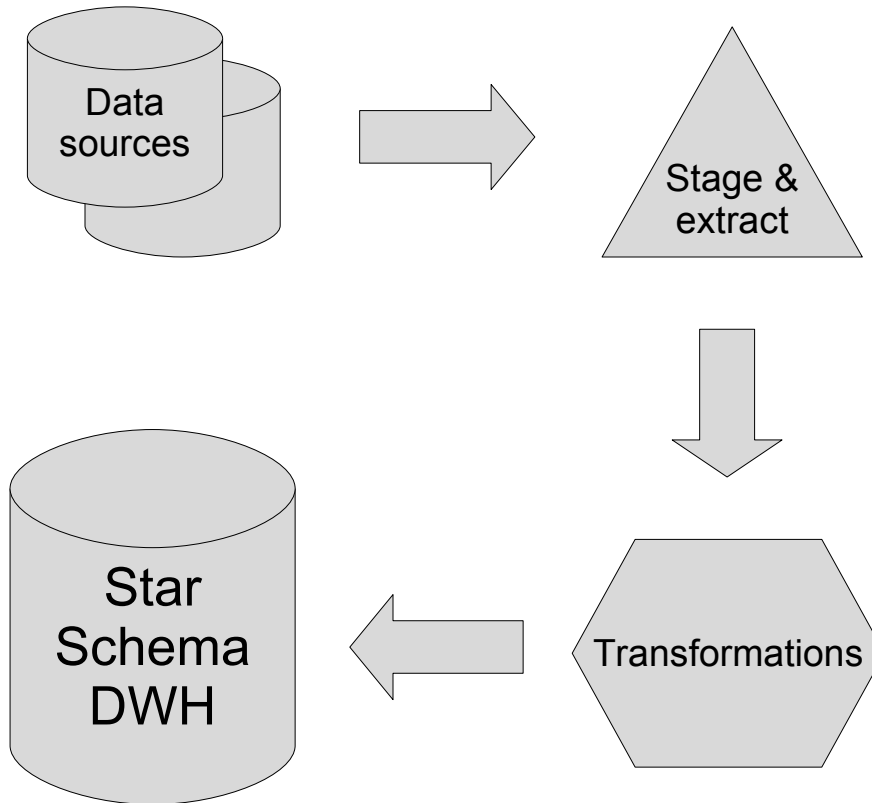
Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Analysis tools

- Raw event streams are difficult to use as-is
- Multiple sources for information(!)
- Postprocess and integrate
  - Combine or link related events
  - Calculate value(s)
- Store in a schema which facilitates dimensional reporting (eg, star or snowflake)

# Data warehousing process



1. Extract data from all sources, whether external, OLTP databases, or event logs
2. Identify common dimensions, related data, transform structure and reorganize schema
3. Load to the final data warehouse

# Tools for data integration

- Could be scripted, but maintenance is a killer
  - Documentation
  - Data processing nasty to deal with in script form
- Choose a tool from the start
  - Pentaho/Kettle
  - Talend
  - BO DataIntegrator
  - Informatica
  - Etc

# Reporting and analysis

- Handwritten SQL
  - Expressive but cumbersome
- OLAP cubes
  - Rapid but memory limited, require constant reloads
- Query builders
  - With a good UI and DB schema, it's what I'd choose
- Specialized tools where available
  - For web traffic analysis, building anything seems mad when tools like Google Analytics available for free

# Columnar storage in MySQL

Observations on our Infobright solution

**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™

# Infobright's Brighthouse

- Columnar engine for MySQL
  - Its own server (5.1 based), not a pluggable engine
- No indexing of data required
  - Data is packed per-column per 64k row values
  - Engine maintains summary data per each 64k values
  - Queries target each pack where summary matches
  - Joins are supported by additional pack-to-pack data
  - Ideal for numeric data
- Fast data loader
- Most MySQL-compatible tools “just work”

# Immediate benefits

- 4 times faster loads (without tuning rest of out toolchain)
- 1/8<sup>th</sup> of the disk space needed, so we could reallocate terabytes of storage to other uses
- No time spent worrying about ad-hoc business report needing a nonexistent index
- Typical query performance increase significant

# Queries which benefit a lot

- Summarizing of big tables w/ GROUP BY
  - Historically a horrible area for MySQL (requires disk-based temporary tables and sorts)
  - Infobright often executes from the knowledge grid information only, without even accessing the actual data!
- These can be tens or hundreds of times faster

```
select t2.c3,sum(t1.c1) from t1 join t2 using (c2)
where c2 between 1 and 1000000 group by c3
```

# Unsuitable applications

- Single-row selects aren't Infobright's natural domain
  - Row-based engine would fetch by index
  - Infobright needs to unpack 64k rows from each columnar datapack
- Typically execute in about 1 second (instead of milliseconds)

```
select * from t1 where pk=?
```

- Text searching

# Limitations

- You might be used to functionality which doesn't work with the engine
  - No DML in community edition (insert,update,delete)
  - Practical restrictions to DML also in enterprise edition
- No constraints, primary keys or auto increment
- Table changes still require a full reload
- Diagnostics are under-developed
  - No useful query plan “explain”
  - Query states give little insight (always “init”)
  - No engine status internals exposed

# Differences in query use

- Mostly, the same query strategies as with normal MySQL, however some notable differences
- Joins perform well when pack-to-pack metadata is available
  - Eg, `FROM t1 JOIN t2 ON t1.column1=t2.column2`
  - No arithmetic operations in join conditions
  - No BETWEEN or inequality operators (in joins!)
  - LEFT JOIN is sometimes fast, many times slow
- On the other hand, WHERE .. IN (subselect) usually is very quick
- Every column selected adds cost (more I/O)

# Sizing and scaling considerations

- Table-level locking impacts loading strategy
  - SELECT vs INSERT/LOAD starvation issues similar to MyISAM (w/ “low priority updates”)
- Parallel loader, but no parallelizing select
- Can't replicate or partition over multiple servers
- Practical scaling assumptions:
  - one host server
  - 500 times main memory in data storage (less actual disk utilized thanks to compression)
  - # of cores equal to # of users
  - Eg; 30 TB on a single HP DL 585 compressed to 6 TB

# Different tech for different purposes

## Real-time shared state

- In-memory data structures
- Java clusters, Memcached, etc
- Future: MySQL Cluster NDB?

## OLTP

- Transactional integrity
- InnoDB

## Log files & analytics

- High-volume events
- Infobright ICE / IEE

# Thank you!

[www.sulake.com](http://www.sulake.com)  
[www.habbo.com](http://www.habbo.com)

[osma@sulake.com](mailto:osma@sulake.com)  
[www.fishpool.org](http://www.fishpool.org)  
[twitter.com/osma](https://twitter.com/osma)

**sulake**

Osma Ahvenlampi - MySQL Conf 2009 - Mining for insight

everyone can play™