



# GPU TECHNOLOGY CONFERENCE

**Empulse GmbH**  
**Michael Hummel - Managing Director**  
**ParStream - a parallel database on GPUs**

GTC, San Jose Convention Center, CA | Sept. 20-23, 2010

PRESENTED BY  NVIDIA.

## Huge demand for mass-data analysis

*The demand for analysis of structured mass data is growing tremendously in all industries.*

### eCommerce

- Web analytics
- SEO
- Affiliate nets

### Social Networks

- Ad serving
- Profiling
- Targeting

### Finance

- Algo trading
- Trend analysis
- Fraud detection

### Telco

- Profiling
- Targeting
- Billing

### Energy

- Smart metering
- Smart grids
- Wind parks

### Many more

- Science
- Transportation
- Security

**Billions of records need to be analysed quickly - an \$8 billion market**

## Current databases are not engineered for mass-data

*“Existing databases are outdated”*

Several Statements on VLDB2010

---

## Many new approaches moving in the wrong direction

*“MapReduce ... we got rid of it.”*

Lipkovitz, Senior Director of Engineering, Google - 09/09/10

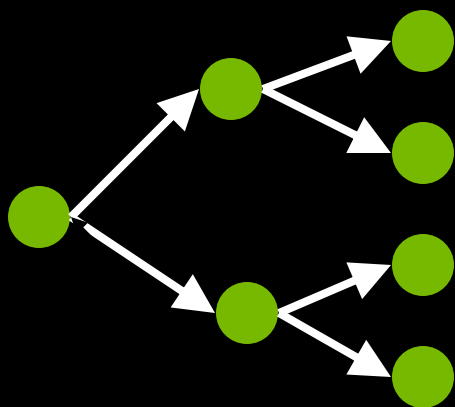
## ParStream - a revolutionary parallel database for high-performance mass-data analysis



- Unique index technology
- Efficient parallel processing
- Ultra-fast, even with billions of records
- Scales linearly up to petabytes
- Real-time analysis, continuous import
- Cost & energy efficient

## ParStream's unique index enables parallel processing

### Conventional database



- Only predefined analysis
- Sequentially

decompression

**small**

very  
large

- Flexible analysis
- Frequent & costly decompression

### ParStream's approach



Immediate  
massiv-parallel  
analysis

- Flexible analysis
- No decompression required
- Fits into GPU memory

## Experts are impressed

- *An extremely innovative idea*
- *Convinced me completely*
- *2 years ahead of competition*

Prof. Dr. Markl, leading database expert at the TU Berlin

## Customers are delighted

- *Outperforms all databases by a factor of at least 35*
- *Reduced our response time from 3 minutes to 15 ms (x12000)*
- *Scales linearly*
- *Runs stable in production*

Leading web-analytics company

## Today: Market-ready

- *Successful running in productive customer environments*
- *Complete, market-ready product*
- *empulse is privately held, no capital raised*
- *Supported by German Ministry of Economics*
- *Research partner TU Berlin*



## In 5 years: \$200 Mio TO

- *International company with industry specific solutions*
- *BI market \$8 billion p.a.*
- *Expected market share >2%*
- *Turnover \$200 Mio p.a. with licenses, services & consulting*

Thank you for your attention!



A revolutionary parallel database  
for online-analysis of structured mass-data

[www.parstream.com](http://www.parstream.com)





**GPU**  
TECHNOLOGY  
CONFERENCE



EMERGING  
COMPANIES SUMMIT

PRESENTED BY  NVIDIA.

**Backup**

PRESENTED BY  NVIDIA.

## Customer projects - some examples:

Web analytics provider	Concurrent User Calculation	1 billion records in 15 ms
SEO Consultant	Data mining	1 billion records joined together within < 1 second
Market research	Flexible multi-column filtering & grouping	5000 queries, >1000 columns per online-analysis
Climate research	Filter & geo-clustering	3 billion records in 100 ms to scale up to 3 petabyte
Best travel offer search	High throughput for complex filter & sort	1000 queries per second on 1 billion travel offers

## Our main challenge is fast growth

Establish strong  
sales & consulting

- Partner with vendors, integrators & solution providers
- Recruit experienced market specialists

Go international

- Invest in market entry & structures in Europe & US

Stay in the lead

- Speed up growth through financing
- Develop industry specific solutions

Grow organisation

- Hire IT, sales and market experts

Complexity

- Sell, plug & play appliance / establish database services

## Current databases are not engineered for mass-data

*“Existing Database Architectures are 20-30 years old and are not able to cope with current data sizes.”*

Several Statements on VLDB2010

## Many new approaches moving in the wrong direction

*“MapReduce isn't suited to calculations that need to occur in near real-time. You can't do anything that takes a relatively short amount of time, so we got rid of it.”*

Lipkovitz, Senior Director of Engineering, Google - 09/09/10

## Academics are impressed...



Prof. Dr. Markl  
TU Berlin -  
IBM-DB2 Expert

- “ParStream is an **extremely innovative** idea for processing data in parallel.”
- “The fundamental concept of **ParStream**, as well as its innovative aspects, have **convinced me completely**.”
- “Based on the advancement of empulse ... it is **unlikely that competitors will be able to develop a comparable product within the next two years**.”

## Customers are delighted...

- “On average, **ParStream outperforms** competing commercial, column-oriented **databases by a factor of 35**”
- “**ParStream scales linearly**: The response times grow as expected for large data volumes. No decrease in performance could be observed.”
- “Over the last few months, **ParStream has been running stable** in production with multiple instances running on different servers.”

## Benefits at a glance:

- ULTRA-FAST: get analytical results in fractions of a second
- SCALABLE: easily store & analyse billions of records
- UNIQUE: perform analytics that were previously not feasible
- REALTIME: simultaneously import and analyse data
- EFFICIENT: get more effective use of your business framework
- COST-EFFECTIVE: save on infrastructure costs

# GPU

TECHNOLOGY  
CONFERENCE

# EMERGING COMPANIES SUMMIT

PRESENTED BY  NVIDIA.

## ParStream - a complete parallel database on GPUs

