



CORAL8, INC.
MAKE IT CONTINUOUS

Teaching an Old Dog New Tricks or Adapting SQL to Event Processing

Event Processing Symposium

March 14th, 2006

Mark Tsimelzon
President & CTO
mark@coral8.com

WWW.CORAL8.COM



Agenda

- Why Start from SQL?
- Query Examples
- What the Vendor Community Can Do
- The Event Processing Design Patterns





Coral8, Inc. – Basic Facts

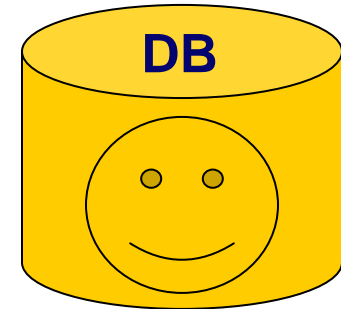
- Founded in 2002
- CEO Terry Cunningham, Founder of Crystal Decisions
- Leverages research from Stanford's Project STREAM
- Headquartered in Mountain View, California
- 30 Employees
- Product: Coral8 Engine
 - First Beta in September 2004
 - First Release in November 2005





Databases: The Good

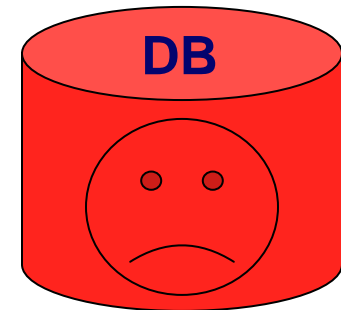
- Relational Model
- High-Level Query Language (SQL)
- Well-understood by millions of developers





Databases: The Bad

- Not designed for continuous incremental querying
- Must store all data on disk before querying
- Optimized for slow disk access
- Do not directly support time dimension
- Hard to scale to hundreds of servers





The Big Idea

- Get the best from databases:
 - Relational Model
 - SQL syntax and semantic
- Replace request-response queries with continuous incremental queries
- Add the time dimension and other extensions

**Arrive at the SQL-like
Continuous Computation Language (CCL)**





Relational Model: Streams Are Like Tables

StreamInRfidReadings - ccl://localhost:6789/Stream/Default/Rfid_warehous...

File Edit Debug Tools Help

▶ || 🔍 ◀▶▶

Timestamp ▼	ReaderID	TagID	TagType
2006/03/03 12:56:26.843750	reader_40167	tag_002497	case
2006/03/03 12:56:26.843750	reader_40715	tag_005853	case
2006/03/03 12:56:26.843750	reader_40548	tag_000178	case
2006/03/03 12:56:26.843750	reader_10114	tag_001487	case
2006/03/03 12:56:26.843750	reader_60071	tag_006448	palette
2006/03/03 12:56:26.843750	reader_30118	tag_000712	case
2006/03/03 12:56:26.843750	reader_20008	tag_002564	case
2006/03/03 12:56:26.859375	reader_40645	tag_000106	case
2006/03/03 12:56:26.859375	reader_40561	tag_007086	case
2006/03/03 12:56:26.859375	reader_40487	tag_003002	case
2006/03/03 12:56:26.859375	reader_30917	tag_007904	palette
2006/03/03 12:56:26.859375	reader_60051	tag_006478	case
2006/03/03 12:56:26.859375	reader_60186	tag_009191	palette
2006/03/03 12:56:26.859375	reader_60075	tag_000916	case
2006/03/03 12:56:26.859375	reader_10052	tag_001875	case
• 2006/03/03 12:56:26.875000	reader_30853	tag_002485	case
• 2006/03/03 12:56:26.875000	reader_20000	tag_009530	case
• 2006/03/03 12:56:26.875000	reader_20008	tag_005745	case
• 2006/03/03 12:56:26.875000	reader_10017	tag_007178	case
• 2006/03/03 12:56:26.875000	reader_40449	tag_003895	case
• 2006/03/03 12:56:26.875000	reader_10152	tag_004532	case
• 2006/03/03 12:56:26.875000	reader_60042	tag_007689	case
• 2006/03/03 12:56:26.875000	reader_10056	tag_000419	case

705.88 messages/sec Total: 62024 messages





Separate Data Model from Complexities of Real World

- Delayed messages
- Out-of-order message
- Lost messages
- Server vs. application timestamp

The screenshot shows the Coral8 Studio interface for configuring a stream. The window title is "Rfid_warehouse - Default - Coral8 Studio". The menu bar includes File, Edit, Run, Tools, and Help. The Explorer pane on the left shows a tree view of streams, with "StreamInRfidReadings" selected. The Properties pane on the right shows the configuration for "Stream 'StreamInRfidReadings'".

Stream URI: `ccl://localhost:6789/Stream/Default/Rfid_warehouse/`

Http URI:

Maximum size of the queue:

Maximum row age (microseconds):

External Streams

Stream receives external data

Maximum delay (microseconds):

Messages can come out of order

Max Synchronizer Delay

Uses Server Timestamp

Output

Ln: 1 Col: 1 Ch: 1





Simple CCL Query: When We Say SQL We Mean SQL

This is a registered continuous query!

```
INSERT INTO
  StreamLargeOrders
SELECT
  CustomerID, OrderID, OrderTotal
FROM
  StreamOrders
WHERE
  OrderTotal > 100000
```





Main Extension: Windows

Compute Volume-Weighted Average Price
over a 5 minute jumping window

```
INSERT INTO
  StreamVWAP
SELECT
  Symbol, SUM(Volume*Price) / SUM(Volume)
FROM
  StreamTrades KEEP EVERY 5 minutes
GROUP BY
  Symbol
```





Joins are EASY

A Join between two windows

```
INSERT INTO
  StreamSeriousAlerts
SELECT
  `Common IP`, StreamIDSAlerts.IP
FROM
  StreamIDSAlerts KEEP 10 minutes,
  StreamVirusCheckAlerts KEEP 10 minutes
WHERE
  StreamIDSAlerts.IP = StreamVirusCheckAlerts.IP
```





DB Subquery

A join between a stream and the results of an SQL query:

```
INSERT INTO
  StreamOut
SELECT
  InTrades.Symbol, InTrades.Price, Closing.Price
FROM
  InTrades,
  (Database "OracleDb"
  [[
    SELECT price FROM price_history ph
    WHERE ph.symbol = InTrades.Symbol AND
    ph.closing_date = current_date-1
  ]])
) as Closing
```





Event Pattern Matching

```
INSERT INTO
  StreamAlerts
SELECT
  a.id, b.amount, c.amount
FROM
  StreamA a, StreamB b, StreamC c, StreamD d
MATCHING
  [10 seconds: a, b || c, !d]
ON
  a.id = b.id = c.id = d.id
WHERE
  b.amount > 5000 AND c.amount < 10000
```





Developers Demand Testing and Debugging Features

- Easily switch between live and recorded data for debugging and back-testing
- Examine the contents of input, output, and intermediate streams
- Examine the contents of every window
- Single-step, forward and backward





Enterprise Features

- XML Support
- State Persistence
- Clustering for High Availability
- Clustering for Performance
- Guaranteed Message Processing
- Vision: Grids for Stream Processing

Grid Vision White Paper:

<http://www.coral8.com/downloads/Coral8GridOperationNonNDA.pdf>





What the Vendor Community Can Do

- **Develop Common Terminology**
 - Stream, query, message, event, CEP, ESP, etc.
- **Create Benchmarks**
 - Define a precise set of language-independent benchmarks
- **Explore Advanced Use Cases**
 - Move beyond VWAP
 - Explain design patterns that can be implemented with ESP technology





Ten Design Patterns (1-5):

1. **Simple Computation over Windows**
Moving Avg, Count, Max, Min, StdDev...
2. **Complex Mathematical Computation**
Time Series, Vector & Matrix computation
3. **Correlation across Multiple Streams**
Joins across streams and windows
4. **Correlation between Streams and Static Data**
Access to historical and reference data
5. **Funnel in Front of a Database**
Reduce the amount of noise the database sees





Ten Design Patterns (6-10):

6. Event Pattern Matching

Fraud detection, security violations, etc.

7. Complex Logic

State machines of all kinds

8. Process and Exception Tracking

Report deviations from the plan, exceptions, missing steps, unfinished steps, etc.

9. Request-Response Queries

Query the current state

10. Subscription Queries

Users parameterize and instantiate their own queries





Conclusion and Q&A

SQL-based platforms are great
for
Event Processing Applications

Coral8 Engine: Easy to Program, Easy to Deploy, Easy to Scale

www.coral8.com

