

# Using Data Stream Management Systems to analyze Electric Power Consumption Data

Talel Abdessalem, Raja Chiky,  
Georges Hébrail and Jean-Louis Vitti  
Ecole Nationale Supérieure des Télécommunications,  
Electricité De France R&D  
March 2007



## Context

- Development of AMM (Automatic Metering Management)
  - Electric power consumption will be measured at a rate up to one index per second.
- Development of Data Stream Management Systems (DSMS)

⇒ Aim:

Using Data Stream Management Systems to analyze  
Electric Power Consumption Data



## Outline

- Data Stream Management Systems
- Experiments
- Example of queries
- Synthesis
- Conclusion



11/05/2007

WDSA'07

3



## DSMS

### -Definitions

- DBMS
  - data stored in **finite**, **persistent** data sets
  - One-time queries
- Data Stream
  - **Ordered**, **infinite** and **continuously** generated sequence of data that can be read only once
  - Near real-time monitoring and analysis is required



11/05/2007

WDSA'07

4



## DSMS

### -Definitions (contd.)

- **Continuous queries:**
  - Queries carried out on streams
  - Persistent
  - Result given as a stream
  - Example :  
Aggregated electric consumption grouped by city over the last 24 hours
- **Windowing technics to handle some blocking operations like aggregation**
  - physically defined window in terms of a time interval
  - logically defined window in terms of the number of tuples
  - Fixed windows, sliding, or with landmark  
Ex.: March 2007, last hour, start at 01/01/2007



11/05/2007

WDSA'07

5



## DSMS

### -Existing prototype systems

- **General purpose DSMS**
  - STREAM : Stanford University
  - TelegraphCQ : Berkeley University
  - Aurora (Medusa, Borealis) : Brandeis, Brown University, MIT
- **Specialized DSMS**
  - Gigascope et Hancock : AT&T (Network monitoring and Telecom streams)
  - NiagaraCQ : University of Wisconsin-Madison (continuous XML query system for dynamic web content)
  - ...



11/05/2007

WDSA'07

6



## Outline

- Data Stream Management Systems
- Experiments
- Example of queries
- Synthesis
- Conclusion



11/05/2007

WDSA'07

7



## Experiments: study and installation of STREAM

- General purpose stream data manager
- Data streams and stored relations
- Windowing :
  - sliding windows
  - logical, physical
  - Partitioned windows
- CQL (continuous query language) for declarative query specification
- Timestamps in streams (integer timestamp)
- Flexible query plan generation
- Resource management:
  - Operator scheduling
  - Graceful approximation: can handle high data rates



11/05/2007

WDSA'07

8



## Experiments: study and installation of TelegraphCQ

- Built as an extension to the PostgreSQL relational DBMS(particular mode of execution)
- Data structure :
  - Relational structure of PostgreSQL
  - Stream structure (CREATE STREAM ...)
- Windowing (physical, sliding, landmark, jumping)
- Each stream has a special time attribute that TelegraphCQ uses as the tuples timestamp for windowed operations
- Queries can be added dynamically when others are being executed



11/05/2007

WDSA'07

9



## Experiments

- Input
  - Standard relations
    - Correspondence between meter, customer and city
    - Standard Electric Consumption hour by hour
  - Data streams:
    - Data streams of several meters observed every 2 seconds  
*Stream index(meter CHAR,index INT, date DATE)*
    - Stream of temperatures recorded each hour for each city
- Some queries for electric power consumption analysis
  - Q1- Consumption of the last 5 minutes -minute by minute- grouped by meter, or by city;
  - Q2- Historical consumption -minute by minute- grouped by meter, or by city, starting from a fixed point;
  - Q3- Alarm -hour by hour- at exceeding a 'standard' consumption depending on the temperature.

meter	index	date
05012606XX	11089624	12/04/2003 07:53:59
10541492YY	11089624	12/04/2003 07:53:59
16381643ZZ	11089624	12/04/2003 07:53:59
05012606XX	11089626	12/04/2003 07:54:01
10541492YY	11089626	12/04/2003 07:54:01



11/05/2007

WDSA'07

10



# Outline

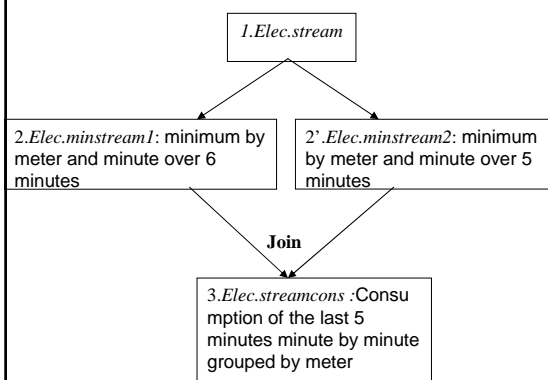
- Data Stream Management Systems
- Experiments
- Example of queries
- Synthesis
- Conclusion



# Q1 TelegraphCQ

Consumption of the last 5 minutes -minute by minute- grouped by meter

```
CREATE STREAM Elec.stream(meter VARCHAR(12),index INTEGER,tcqtime TIMESTAMP  
TIMESTAMP COLUMN) TYPE ARCHIVED;  
ALTER STREAM Elec.stream ADD WRAPPER csvwrapper;
```



### Example:

```
WITH  
Elec.minstream1 AS  
(SELECT meter, min(index),  
DATE_TRUNC('minute',tcqtime)  
FROM Elec.flux [RANGE BY '6 minutes' SLIDE BY '1  
minute' START AT '2003-12-04 07:50:00']  
GROUP BY meter, DATE_TRUNC('minute',tcqtime)  
ORDER BY DATE_TRUNC('minute',tcqtime))  
Elec.minstream2 AS ...(as Elec.minstream1 with RANGE  
BY 5 minutes)  
(SELECT f1.meter, f1.minindex, f1.tcqtime,  
f2.minindex, f2.minindex - f1.minindex ,f2.tcqtime  
FROM Elec.minstream1 as f1 [RANGE BY '1 minute'  
SLIDE BY '1 minute' START AT '2003-12-04  
07:50:00'], Elec.minstream2 as f2 [RANGE BY '1  
minute' SLIDE BY '1 minute' START AT '2003-12-04  
07:50:00']  
WHERE f1.meter=f2.meter AND f1.tcqtime= (f2.tcqtime -  
interval '1 minute'));
```



## Result of Q1

meter	minindex_b	tcqtime begin	minindex_e	consumption	tcqtime end
05012606XX	11089624	2003-12-04 07:53:00	11089626	2	2003-12-04 07:54:00
05012606XX	11089624	2003-12-04 07:53:00	11089626	2	2003-12-04 07:54:00
05012606XX	11089626	2003-12-04 07:54:00	11089696	70	2003-12-04 07:55:00
05012606XX	11089624	2003-12-04 07:53:00	11089626	2	2003-12-04 07:54:00
05012606XX	11089626	2003-12-04 07:54:00	11089696	70	2003-12-04 07:55:00
05012606XX	11089696	2003-12-04 07:55:00	11089767	71	2003-12-04 07:56:00
05012606XX	11089624	2003-12-04 07:53:00	11089626	2	2003-12-04 07:54:00
05012606XX	11089626	2003-12-04 07:54:00	11089696	70	2003-12-04 07:55:00
05012606XX	11089696	2003-12-04 07:55:00	11089767	71	2003-12-04 07:56:00
05012606XX	11089767	2003-12-04 07:56:00	11089836	69	2003-12-04 07:57:00
05012606XX	11089624	2003-12-04 07:53:00	11089626	2	2003-12-04 07:54:00
05012606XX	11089626	2003-12-04 07:54:00	11089696	70	2003-12-04 07:55:00
05012606XX	11089696	2003-12-04 07:55:00	11089767	71	2003-12-04 07:56:00
05012606XX	11089767	2003-12-04 07:56:00	11089836	69	2003-12-04 07:57:00
05012606XX	11089836	2003-12-04 07:57:00	11089907	71	2003-12-04 07:58:00
05012606XX	11089626	2003-12-04 07:54:00	11089696	70	2003-12-04 07:55:00
05012606XX	11089696	2003-12-04 07:55:00	11089767	71	2003-12-04 07:56:00
05012606XX	11089767	2003-12-04 07:56:00	11089836	69	2003-12-04 07:57:00
05012606XX	11089836	2003-12-04 07:57:00	11089907	71	2003-12-04 07:58:00
05012606XX	11089907	2003-12-04 07:58:00	11089975	68	2003-12-04 07:59:00



11/05/2007

WDSA'07

13



## Outline

- Data Stream Management Systems
- Experiments
- Example of queries
- Synthesis
- Conclusion



11/05/2007

WDSA'07

14



## Synthesis

- Language adequacy
  - Solving queries and processing data “on-the-fly”
  - New and Anti-intuitive logic of queries expression
    - Unborned and transient data
    - ⇒ Windowing + processing + update
- Usability
  - TelegraphCQ:
    - Operational system
    - Queries can be added dynamically when others are being executed
    - Queries result can be re-used as a stream or stored in a file
    - System performance not tested
  - STREAM:
    - CQL definition
    - Queries optimization
    - untimely shutdowns of server during experiments
    - ⇒ Nonoperational prototype



11/05/2007

WDSA'07

15



## Outline

- Data Stream Management Systems
- Experiments
- Example of queries
- Synthesis
- Conclusion



11/05/2007

WDSA'07

16





## Conclusion

- Other logics of queries expression :
  - Example: Aurora
- Intrinsically distributed AMM
  - => Study of distributed DSMS (Borealis)
- Study of commercial follow-up systems:
  - StreamBase, Amalgamated Insight and coral8
- An exact analysis is expensive even impossible: approximation by sampling
  - => panel management in a data stream environment



11/05/2007

WDSA'07

17



*Thank you for your attention*