Sensor Data Management

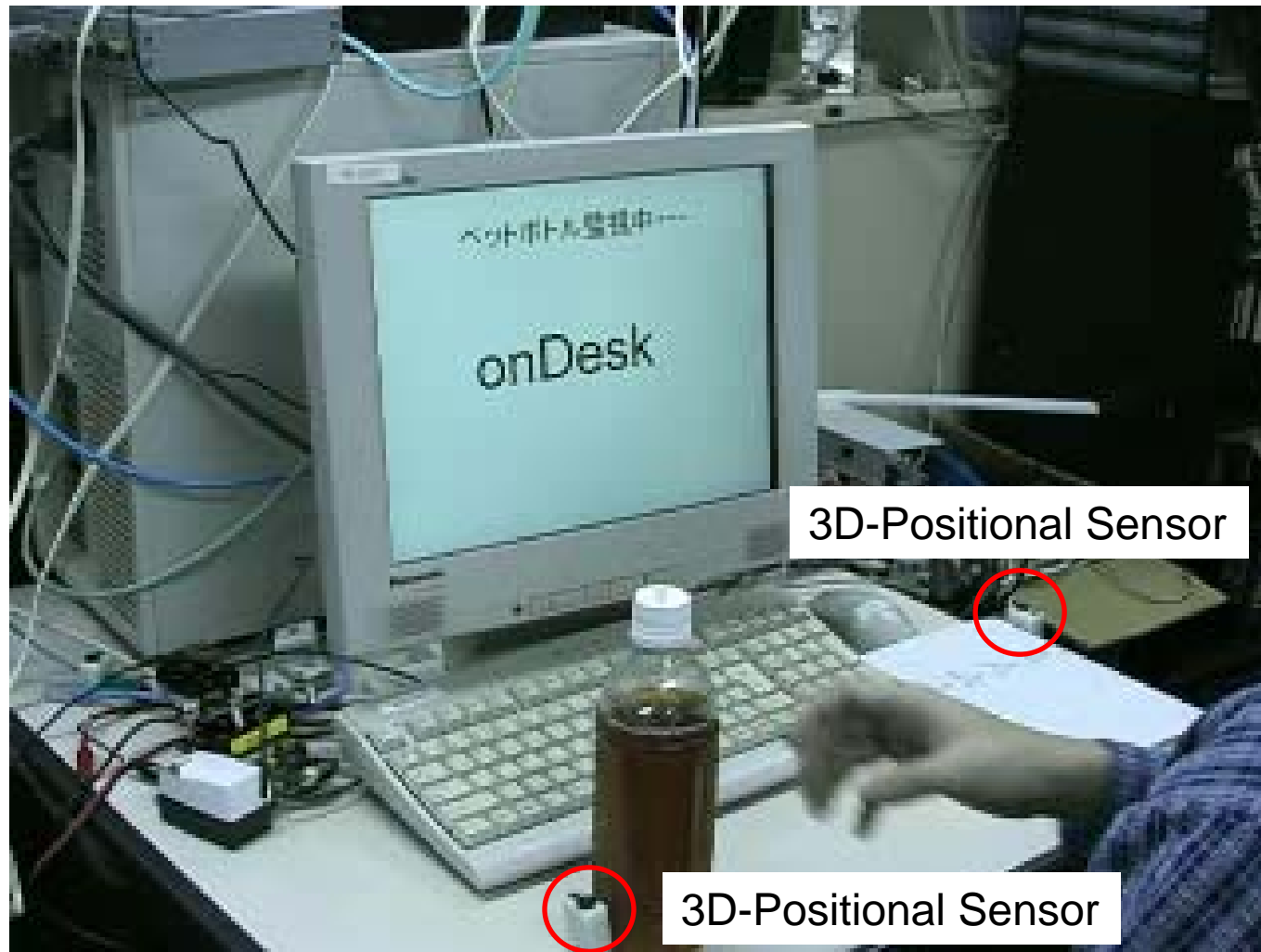## Hideyuki Kawashima
Joined from Feb. 2007〜

Outline

- A DBMS for Signal Streams
- Future Research Direction

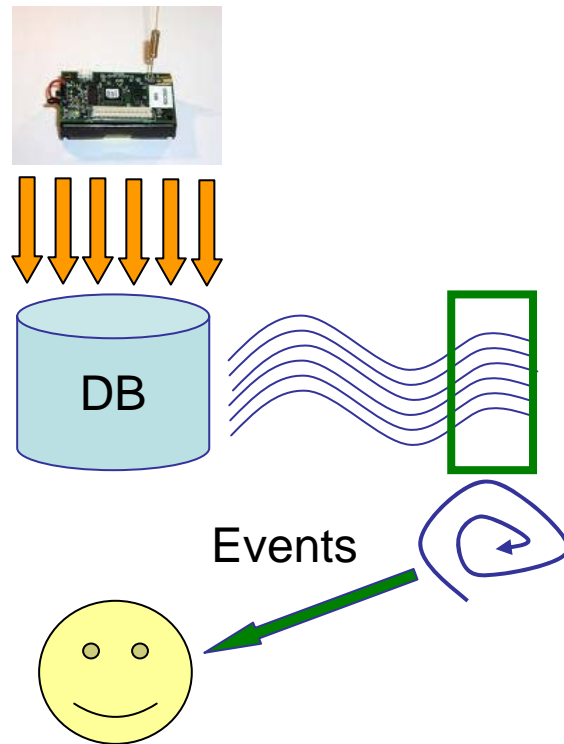# Ubi. Sensor N/W Apps.
MeT: Searching Physical Objects

Supporting Development
of Ubi. Apps by DBMS

Ideal

DB

Events

Devil's Choice

Ans.1：DBMS ＋ Signal Libs.

- Impedance mismatch
- Not general libs

Ans.2: All handmade

- Much cost to develop DBMS
- Cannot reuse for others

Related Work

	Data Store	Cont. Query	Fast Insert	Pattern Recog.
Object Relational DBMS (Informix TimeSeries)	◯			◯
Stream Processing Engine (TelegraphCQ, Boleralis)	◯	◯		◯
KRAFT	◯	◯	◯	◯

KRAFT Data Model:
Abstracted Data Type (ADT) for sensor

ID	Sensor Type	Sensor Value											
1	Accel.	T	T	T	T	T	T	T	T	T	T	T	T
		89	12	40	20	89	31	50	89	71	22	89	2
		76	45	12	30	77	33	76	82	29	62	76	31
		23	90	78	23	11	59	98	24	49	8	55	23
		21	10	15	21	1	4	29	11	48	79	21	43
2	Temp.												
3	Humid.												

INTEGER   TEXT   SENSOR (Multi-Dimensional Time Series)

Features of KRAFT

- Fast Data Insertion
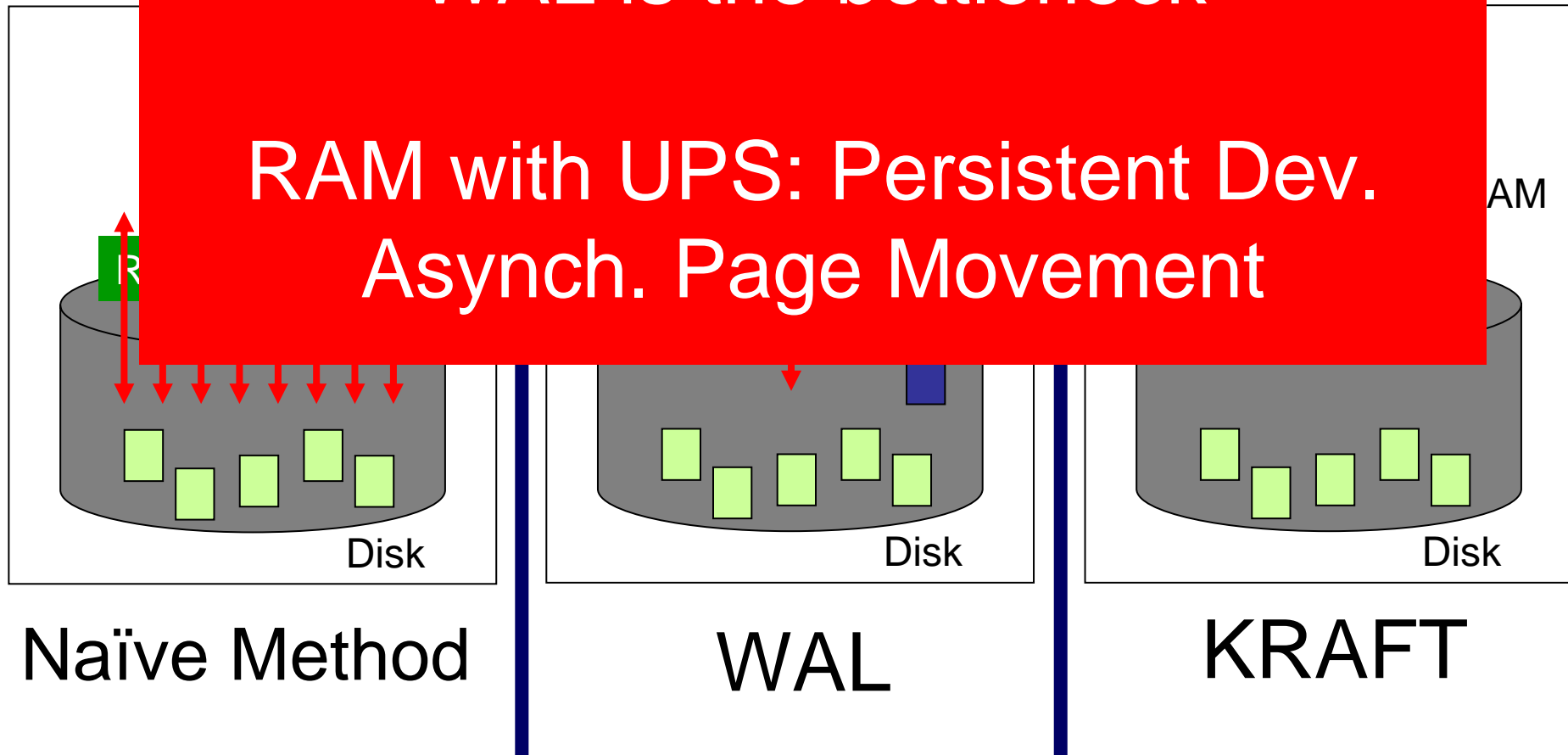- Time Series Analysis
- Periodic Query Execution

Magic of Acceleration
-- Data Insertion Mechanism --
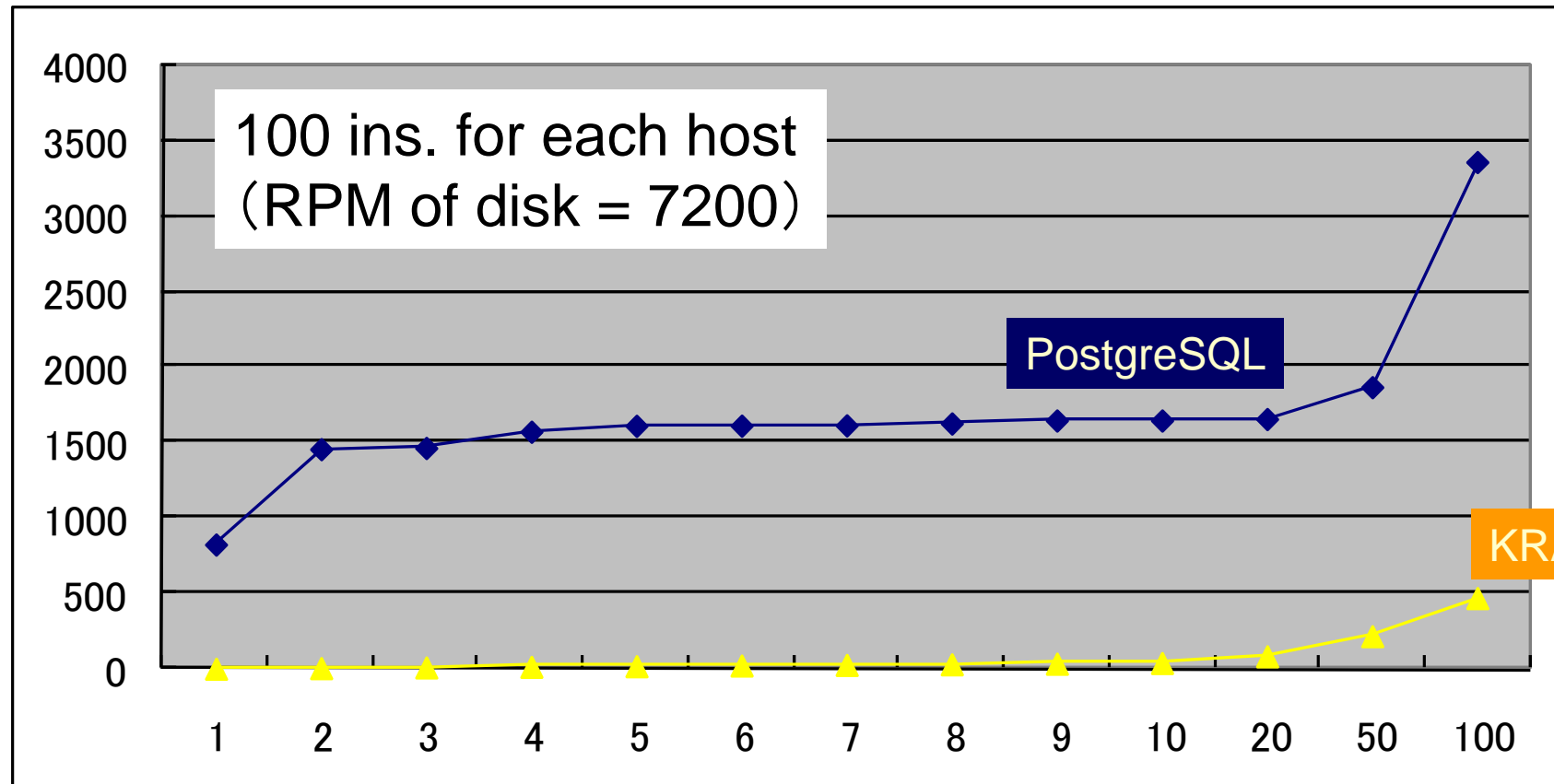

WAL is the bottleneck

RAM with UPS: Persistent Dev.
Asynch. Page Movement

Naïve Method

WAL

KRAFT

Data Insertion Speed

Mili Sec

100 ins. for each host
（RPM of disk = 7200）

PostgreSQL

KRAFT

4000
3500
3000
2500
2000
1500
1000
500
0

1   2   3   4   5   6   7   8   9   10   20   50   100

# of Concurrent Client Accesses

Features of KRAFT

- Fast Data Insertion
- Time Series Analysis
- Periodic Query Execution

Data Model (again)

- Object Relational（ADT for sensor）

ID	Name (text)	Sensor
0	Kawashima	

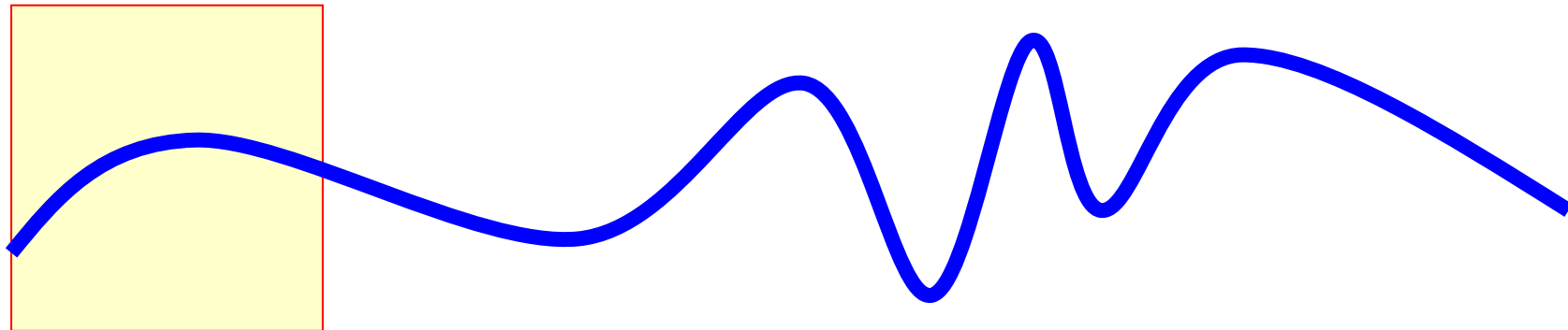| Arrival time |
|--------------|
| Data Value |

Select * from TABLE where id = 1;

Insert into TABLE values …

Append into TABLE values …

Delete from TABLE where id = X

Funcs for SENSOR ADT

- Aggregation (as usual DBMS)
  - MAX, MIN, CNT, AVG, SUM,
- Signal Processing
  - FFT, Wavelet
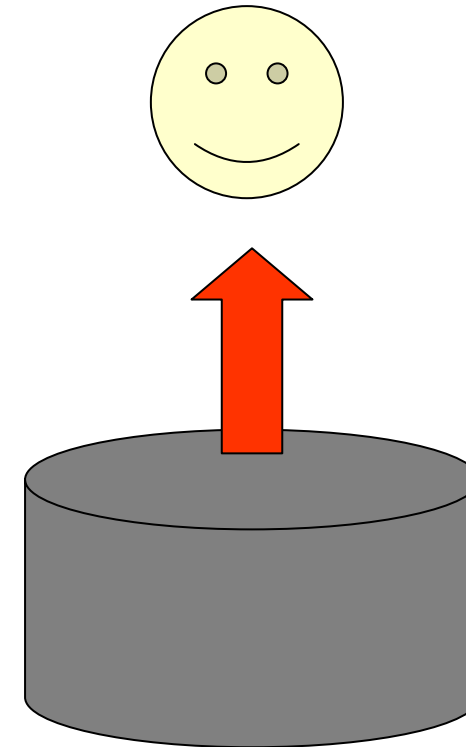- Similar Sequence Retrieval
  - DTW, Euclid
- Sliding Window

Features of KRAFT

- Fast Data Insertion
- Time Series Analysis
- Periodic Query Execution

Periodic Query Execution (Continual Query)
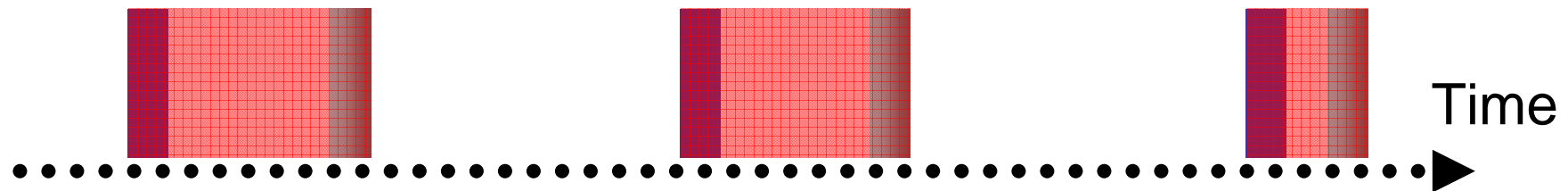
- After query registration, answers are periodically returned

- Human passive, DBMS active (usually, human active, DBMS passive)

- A PQ is realized as a thread

Problem on
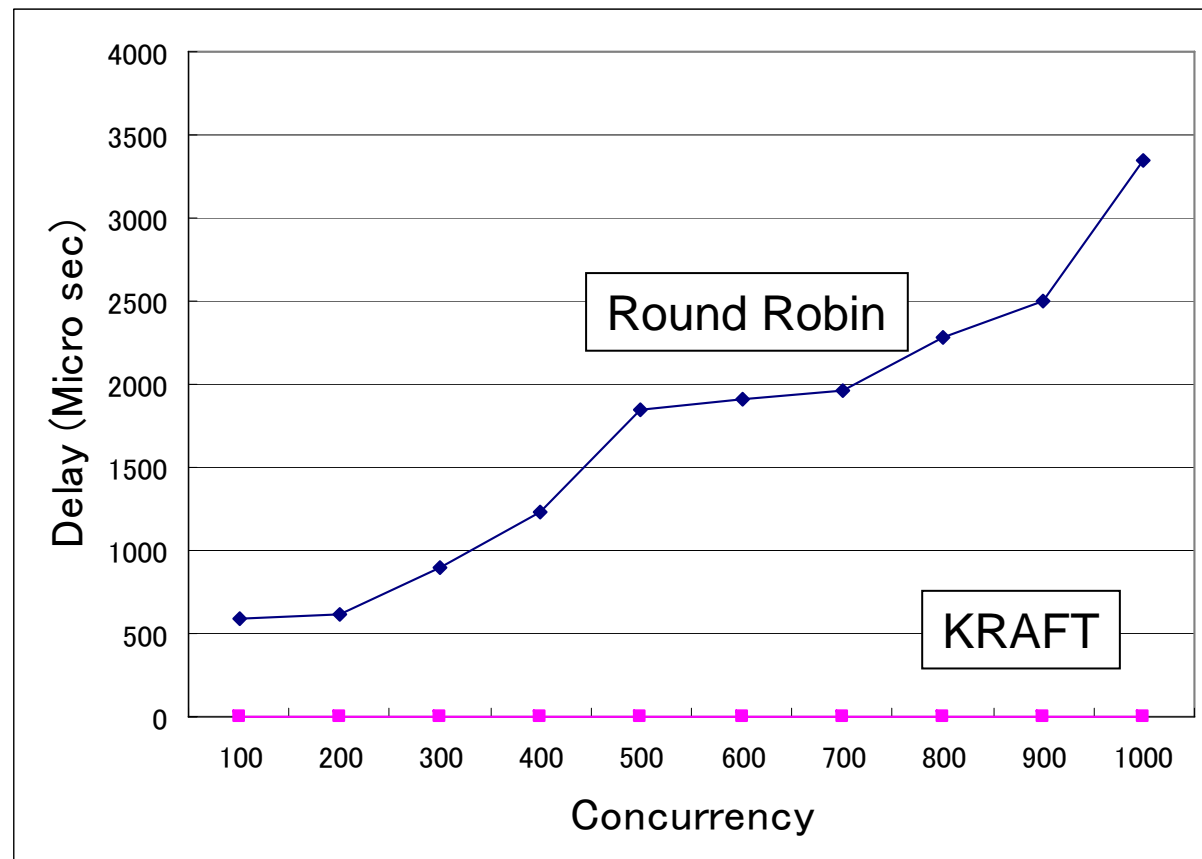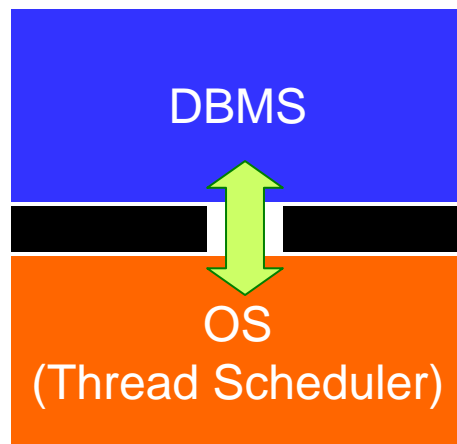# Periodic Data Acquisition

Problem：Delay of query activation

Time

Ideal Start Time (PST)

Real Start Time (RST)

Approach

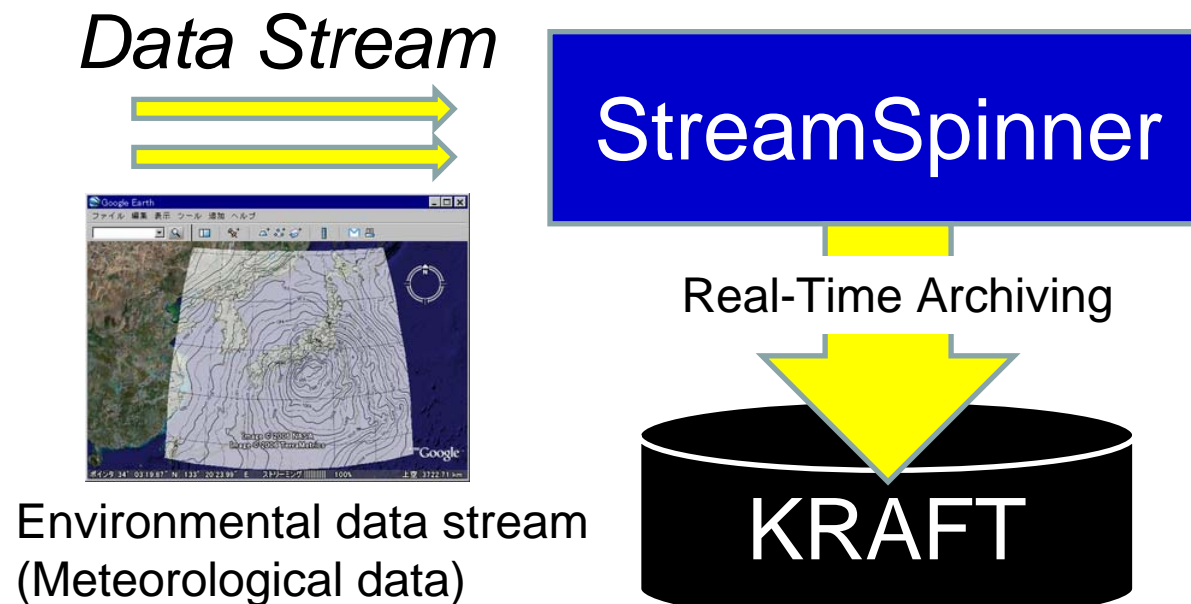- Cooperation of CQ and OS scheduler

Outline

- A DBMS for Signal Streams
- Future Research Direction

# Research Direction (1/2)

- Storage for *StreamSpinner* (Stream Processing Engine, explained later)
  - From toy apps. to real sensor apps.

Data Stream

StreamSpinner

Real-Time Archiving

KRAFT

Environmental data stream
(Meteorological data)

Research Direction (2/2)

- ## Data Mining Meteorological Data
  - Classification of Weather Map by Pressure

Weather Map Database

Except for Winter Type

Winter Type （西高東低）
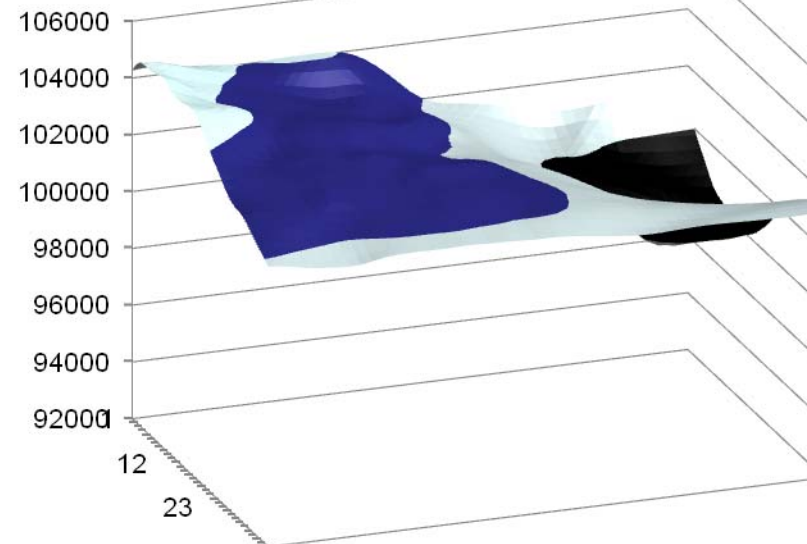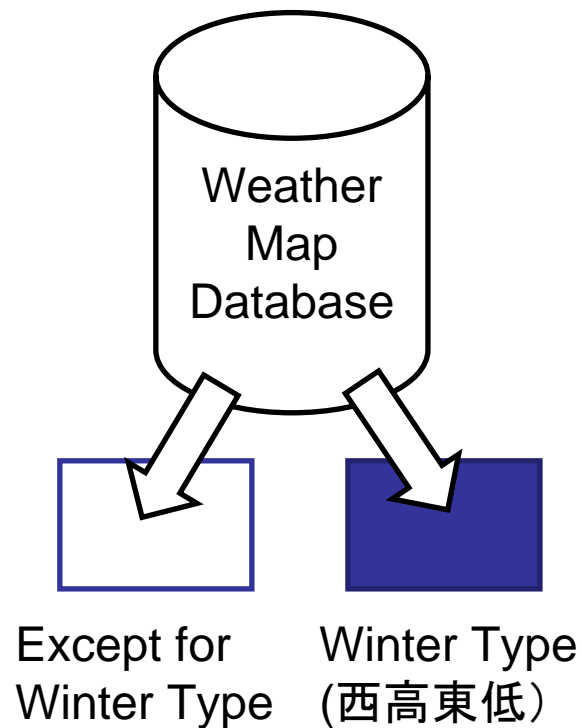
Summary

- KRAFT: a DBMS kernel for signal streams
  - Fast data insertion
  - Signal processing functions
  - Periodic query execution
  - http://sourceforge.jp/projects/kraft
- Future Direction
  - Integrating StreamSpinner
  - Data Mining for Scientific Data

Funding and Outputs

- Funding
  - MEXT Grant-in-Aid for Young Scientists (B)
  - IPA Exploratory Software Project
- Publications
  - Hideyuki Kawashima, "KRAFT: A Real-Time Active DBMS for Signal Streams", Proc. of 4th International Conference on Networked Sensing Systems (INSS'07) pp. 163-166, June 2007
  - Hideyuki Kawashima, Yutaka Hirota, Michita Imai, "MeT: A Real World Oriented Metadata Management System for Semantic Sensor Networks", Proc. of 3rd International Workshop on Data Management for Sensor Networks (DMSN'06), pp. 13-18, September 2006.
  - Hideyuki Kawashima, Michita Imai, Yuiciro Anzai, "Providing Persistence for Sensor Data Stream by Remote WAL", Proc. of 8th International Conference on Data Warehousing and Knowledge Discovery (DaWaK'06), Lecture Notes in Computer Science, Vol. 4081, pp. 524-533, September 2006.