Data Streams

• Continuous, unbounded, rapid, time-varying streams of data elements

• Occur in a variety of modern applications
  – Network monitoring and traffic engineering
  – Sensor networks
  – Telecom call records
  – Financial applications
  – Web logs and click-streams
  – Manufacturing processes

• **DSMS** = Data Stream Management System
DBMS versus DSMS
DBMS versus DSMS

- Persistent relations
- Transient streams (and persistent relations)
**DBMS versus DSMS**

- Persistent relations
- One-time queries
- Transient streams (and persistent relations)
- Continuous queries
The (Simplified) Big Picture

Input streams

Scratch Store

DSMS

Stored Relations

Archive

Register Query

Streamed Result

Stored Result
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)

Total cost of orders fulfilled over the last day by clerk “Sue” for customer “Joe”

Select  Sum(O.cost)
From Orders O, Fulfillments F [Range 1 Day]
Where O.orderID = F.orderID And F.clerk = “Sue”
   And O.customer = “Joe”
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)

Total cost of orders fulfilled over the last day by clerk “Sue” for customer “Joe”

Select Sum(O.cost)
From Orders O, Fulfillments F [Range 1 Day]
Where O.orderID = F.orderID And F.clerk = “Sue”
And O.customer = “Joe”
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)

Total cost of orders fulfilled over the last day by clerk “Sue” for customer “Joe”

Select Sum(O.cost)
From Orders O, Fulfillments F [Range 1 Day]
Where O.orderID = F.orderID
And F.clerk = “Sue”
And O.customer = “Joe”
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)

Total cost of orders fulfilled over the last day by clerk “Sue” for customer “Joe”

Select Sum(O.cost)
From Orders O, Fulfillments F [Range 1 Day]
Where O.orderID = F.orderID 
And F.clerk = “Sue”
And O.customer = “Joe”
Example Query

Two streams, contrived for ease of examples:
Orders (orderID, customer, cost)
Fulfillments (orderID, clerk)

Total cost of orders fulfilled over the last day by clerk “Sue” for customer “Joe”

Select \( \text{Sum}(O \cdot \text{cost}) \)
From Orders O, Fulfillments F [Range 1 Day]
Where O.orderID = F.orderID And F.clerk = “Sue”
And O.customer = “Joe”