**Scenario**

Data: TPC-H (1 GB)
Query: Total revenue from European suppliers per region of customer

**Traditional Query Processing**

```sql
SELECT rc.r_name AS CustRegion,
       SUM(l_extendedprice*(1-l_discount)) AS Revenue
FROM lineitem, orders, customer, nation nc,
     region rc, supplier, nation ns, region rs
WHERE <foreign-key join>
AND rc.r_name = 'EUROPE'
GROUP BY rc.r_name
ORDER BY cust_region
```

![Diagram of Traditional Query Processing]

- **60 min**

**Approximate Query Processing with Derby/S**

```sql
SELECT SOME rc.r_name AS CustRegion,
       ~SUM(l_extendedprice*(1-l_discount)) AS Revenue
FROM lineitem, orders, customer, nation nc, region rc, supplier, nation ns, region rs
WHERE <foreign-key join>
AND rs.r_name = 'EUROPE'
GROUP BY rc.r_name
ORDER BY cust_region
```

![Diagram of Approximate Query Processing]

- **≈ 6 s**

**Sample Repository of Derby/S**

```sql
CREATE SAMPLE revenueSample AS
  (SELECT *
   FROM lineitem, orders, customer, supplier
   WHERE l_orderkey = o_orderkey
     AND c_custkey = c_custkey
     AND l_suppkey = s_suppkey)
OF SIZE 60000 TUPLES
FOR AGGREGATIONS
  (SUM (l_extendedprice*(1-l_discount)))
MANAGED BY SYSTEM REFRESH IMMEDIATE
```

![Diagram of Sample Repository]

**Derby/S**

A DBMS for Sample-Based Query Answering

Anja Klein
anja.klein@sap.com
Rainer Gemulla
gemulla@inf.tu-dresden.de
Philipp Rösch
philipp.roesch@tu-dresden.de
Wolfgang Lehner
wolfgang.lehner@tu-dresden.de

http://www.db.in.tu-dresden.de/research/