

Derby/S A DBMS for Sample-Based Query Answering

Faculty of Computer Science Database Technology Group

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

Scenario

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

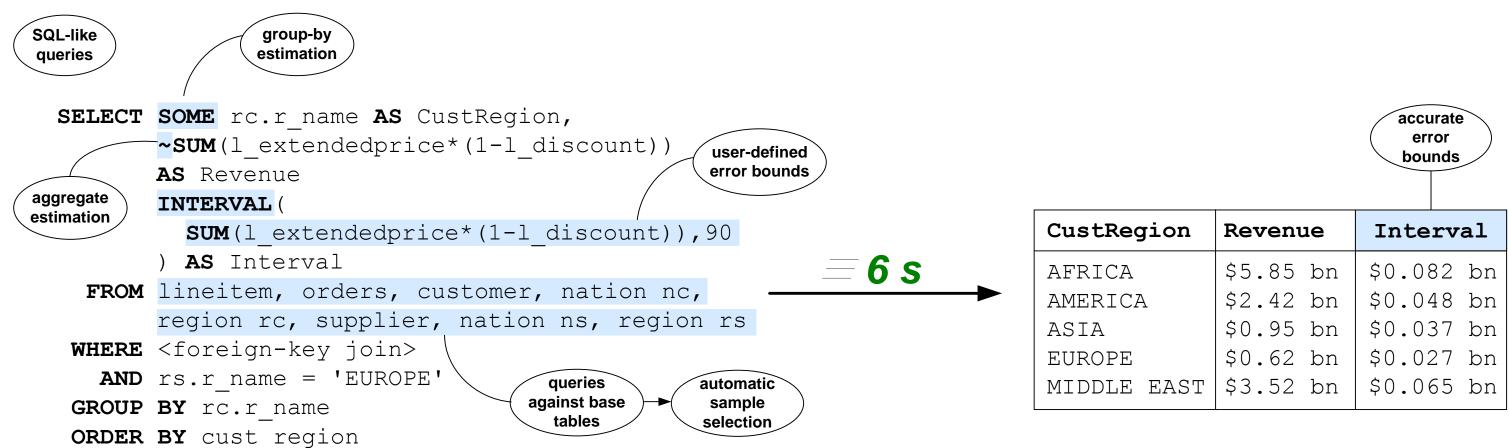
Traditional Query Processing

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** rs.r name = 'EUROPE' GROUP BY rc.r name ORDER BY cust region

60 min

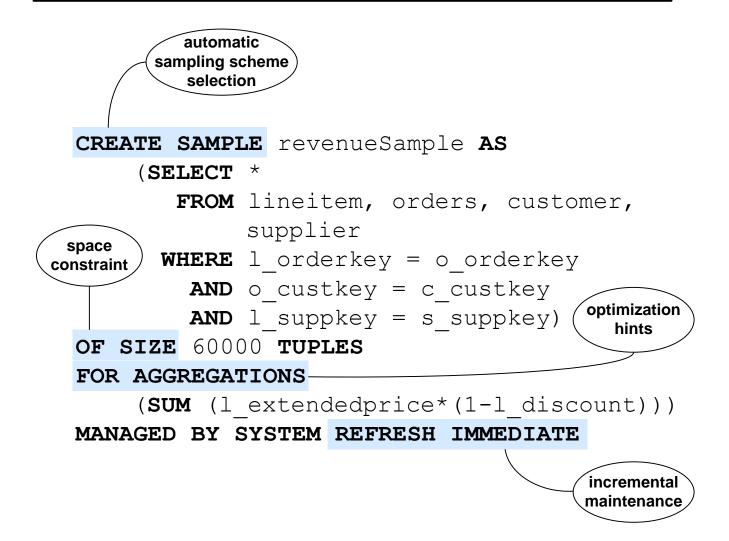
CustRegion	Revenue
AFRICA	\$5.76 bn
AMERICA	\$2.40 bn
ASIA	\$1.00 bn
EUROPE	\$0.66 bn
MIDDLE EAST	\$3.38 bn

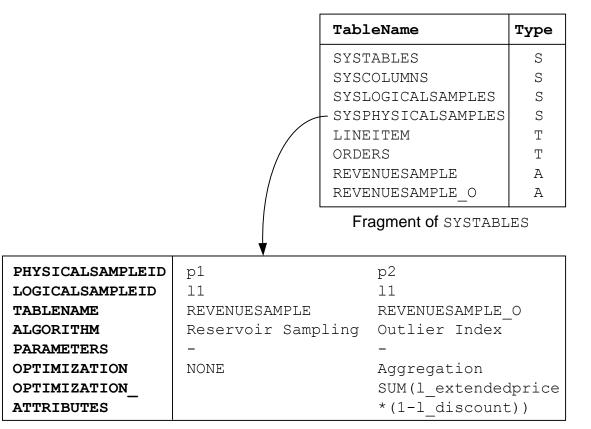
Approximate Query Processing with Derby/S





Sample Repository of Derby/S





Fragment of SYSPHYSICALSAMPLES (pivoted)

http://wwwdb.inf.tu-dresden.de/research/

