









IBM Software Group DB2 Information Management Software Information Integration Challenges			
 Variety Composite applications need data from multiple sources Consumers expect holistic, personalized, and value-added content Critical business information can be in relational, XML, packaged applications, content repositories, file systems 	 Velocity Increasing emphasis on current data Real-time analytics Business performance management for real- time or near real-time business operations 	 Volume Petabytes will be measure of available online data All client interactions are important, e.g., faxes, audio records, e-mail, Internet and intranet content used in analyses and management decisions 	



Distributed access	Consolidated access
 Primary requirements: Very current data Dynamic joining of data Structured and unstructured data Mixed relational and non-relational data Not practical to copy data Small amounts of data in result set 	 Primary requirements: Local performance Structured data Extract, Transform and Load Extensive transformations User metadata Large volumes of data Replication Small amount of changed data Up to near real time updates



















IBM Software Group DB2 Information Manageme	ent Software	IEM
 Heterogeneous Caching Improve query performance and availability Administrator defines <u>Materialized Query Table</u> Pre-computed or frequently used values Any data from the federated relational sources 	WebS	Sphere
 Cached in relational or XML format Implicit use if enabled (automatic routing) Else explicit use only Developer enables cache use If enabled, reads are handled from the cache, writes passed through to the source If not, reads and writes passed through to source Cache refresh managed: Manually By replication Various refresh strategies under design 		Read n Integrator B C Microsoft













	IBM Software Group DB2 Information Management Software	IBM
•	Queries (Conceptual)	
Query 1:	Find an interesting partkey.	
Query 2:	Find the cheapest supplier in a given nation (e.g., "GERMANY") a given partkey (e.g., "28")	for
Query 3:	Reward those customers who placed large orders over a given period of time	
Query 4:	Find customers of one division (one pre-merged company) in select countries who placed very large orders from another division	
Query 5:	Find the average cost of orders that qualifying customers of or division (one pre-merged company) placed throughout the enti (merged) company. Qualifying customers are those that resid in a given country (Japan), belong to a given market segment (household), & have low account balances	ire



IBM Software Group DB2 Information Management Software	IBM
Servlet Architecture	
•With II	
Connect to federated DBMS	
SELECT from view(s), nicknames	
 Disconnect from federated DBMS Without II 	
Connect to N data sources	
 Retrieve data from each req'd table (minimum 1 SELECT per data source) 	
Insert results into a local auxiliary table	
SELECT from auxiliary table	
Clean up (DELETE) contents of auxiliary table	
Disconnect from N data sources	







IBM	Software Group DB2 Informa	ation Management Software	IBM
Performan	ice Comparison		
Elapsed times of our queries (servlet implementation)			
Query	Direct Access	DB2 II Access	
1	3.4 sec	3.5 sec	
2	0.18 sec	0.25 sec	
3	170.1 sec	44.5 sec	
4	79.9 sec	4.5 sec	
5	9.9 sec	15.1 sec	













