Location Agnostic Data

Arup Nanda

Proligence

It's 11 PM.
Do you know where your data is?

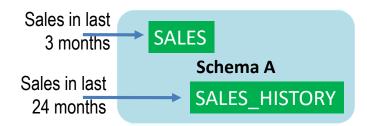




Arup Nanda

Location Agnostic Data

.





SALES

Schema A

SALES_HISTORY

Schema B

SALES_HISTORY





Arup Nanda

Location Agnostic Data

SALES

Schema A

SALES_HISTORY

Schema B

SALES_HISTORY





Datawarehouse

SALES Schema A SALES_HISTORY

Schema B
SALES_HISTORY





HBase
SALES_HISTORY



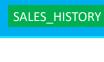
Arup Nanda

Location Agnostic Data

_



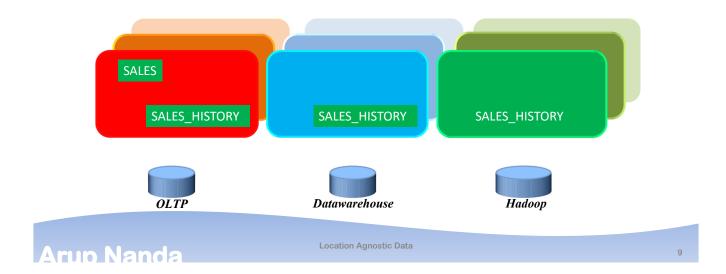


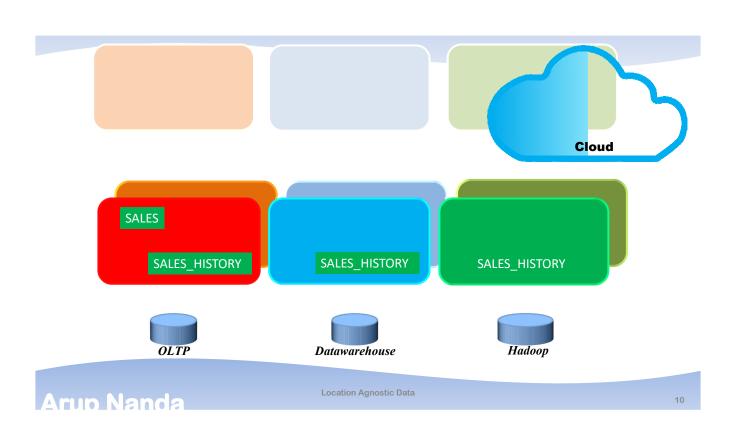














Sharding





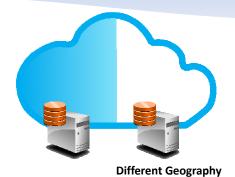




APAC

Rest

LocationNorth AmericaSouth AmericaEMEACustomerVery ImportantImportantSomewhat ImpImportance









Different Location

Different Technology

Arup Nanda

Location Agnostic Data

13

Challenges

- Need to know:
 - The technology
 - Oracle Database, Hadoop?
 - SQL, or Map Reduce Code
 - The location
 - The data availability
 - Do I have the right data in the same DB or got to get from somewhere else?

Solutions

- Put the business logic in the database
- But, which database?
- Which technology

Arup Nanda

Location Agnostic Data

4

Solution, contd.

- Put the logic in the app layer
- · Independent of technology
- But is really technology independent?
- · Definitely inconsistent
- Need to reinvent the wheel
- Still need to know the location, currency, etc.











Different Location

Different Technology

Arup Nanda

Location Agnostic Data

47

Data As a Service

Step 1

- · Gather a list of all high level data
 - Not tables, columns; but data elements such as reservations, sales, etc.

Arup Nanda

Location Agnostic Data

19

Step 2

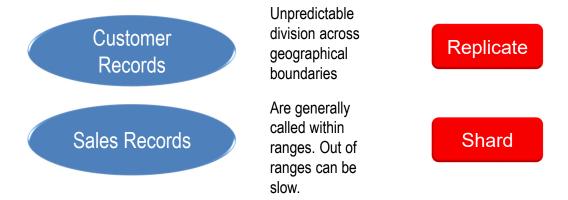
- Map tables, columns, collections, etc. into these data elements
- Define an document business logic for each datastore

Logic	Data store
System of record, less than 3 months	OLTPDB.SchemaA.TableA.ColumnA
Less than 2 years	DWDB.SchemaA.ColumnA
More than 2 years	Hbase

Define all other elements such as how often synced up.

Step 3

• Identify datastores that can't tolerate latency

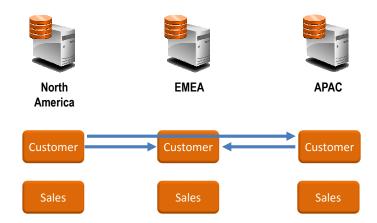


Arup Nanda

Location Agnostic Data

21

Example



Sharded Data Strategies

Use DB Links
 create database link newyork
 connect to gwuser/gwuser
 using 'newyork';

Arup Nanda

Location Agnostic Data

22

Sharded Data, contd.

· Create View SALES as

```
select * from sales
union all
select * from sales@newyork
union all
select * from sales@london
```

Arup Nanda

Location Agnostic Data

Sharded Data, contd.

Create INSTEAD OF triggers
 create trigger
 instead of INSER on sales
 for eac! row
 if country"code in #'\$S'%'&''%() t!en
 INSER into sales@newyork
 (

Arup Nanda

Location Agnostic Data

25

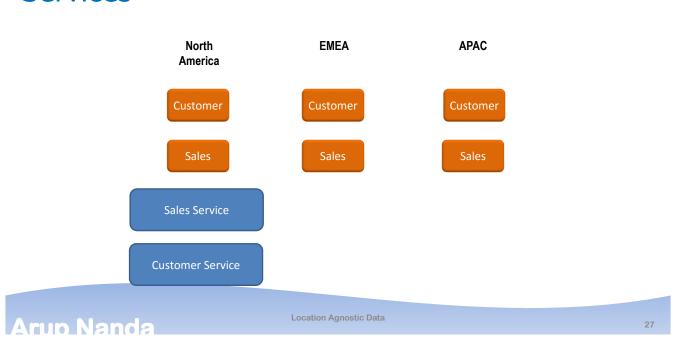
Step 4

- · Write APIs to serve data elements
- · Could be:
 - Java Webservice
 - Oracle Packages
 - RESTful APIs

Arup Nanda

Location Agnostic Data

Services



Service

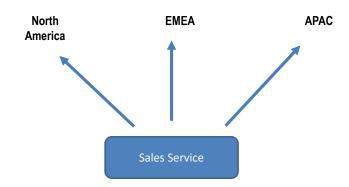
- Has a fixed "catalog" of requests it can address
- · Has knowledge about location, data dispersion, technology
- Example
 - Accept StoreID
 - Lookup country of the store
 - Is data replicated?
 - If so, get local data
 - If not, locate the DB for that country
 - Get the data from that DB

Arup Nanda

Location Agnostic Data

28

Services

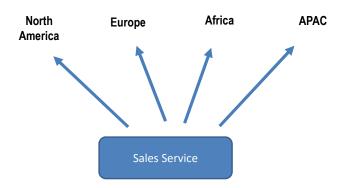


Arup Nanda

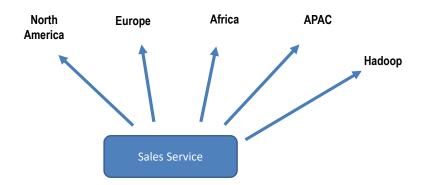
Location Agnostic Data

29

Services



Services

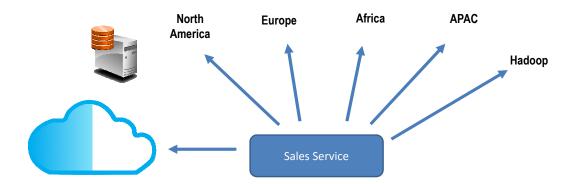


Arup Nanda

Location Agnostic Data

31

Services



Suitability

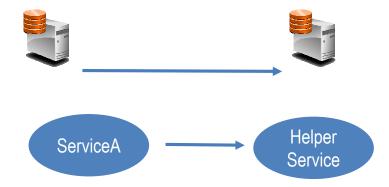
- 1. Reference data—country details, store details
 - Static or almost static
- 2. Low variability—customer preferences
- 3. Latency tolerated—European customers generally call the European call centers
- 4. Read type of access

Arup Nanda

Location Agnostic Data

2

Bridging Technologies



Gateways



Arup Nanda

Location Agnostic Data

25

Summary

- Data will continue to exist in many places
- · Single source of truth is next to impossible to attain
- Architect to live with diverse data, sometimes competing versions
- Decide which data is to be manually replicated and sharded
- Use gateways or services to connect from one database to the other allowing location and technology transparency.
- In Oracle databases, DB Links, Views and Instead of Triggers can provide the location transparencies
- · Gateways can bridge different technologies

Thank You!

Blog: arup.blogspot.com Tweeter: @ArupNanda Facebook.com/ArupKNanda Google Plus: +ArupNanda

Location Agnostic Data

37