

# Tutorial 8

## Intensively Distributed Data

INFS3200/7907  
Advanced Database Systems

### How RFID Technology Works



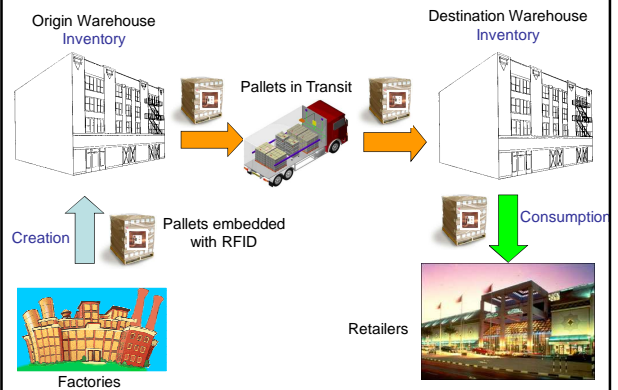
### RFID vs. Barcodes



While **RFID** will probably never completely replace **barcodes**, it has distinct advantages over the **barcode**. For example,

- Human intervention is required to scan a barcode, whereas in most applications an **RFID tag can be detected "hands off."**
- Barcodes must be **visible on the outside of product packaging**. **RFID tags can be placed inside the packaging or even in the product itself.**
- You must have **"line of sight"** to read a barcode. **RFID tagged items can be read even if they are behind other items.**
- The readability of barcodes can be **impaired by dirt, moisture, abrasion, or packaging contours**. **RFID tags are not affected by those conditions.**
- **RFID tags have a longer read range than barcodes.**
- **RFID tags have read/write memory capability; barcodes do not.**
- **More data can be stored in an RFID tag than can be stored on a barcode.**

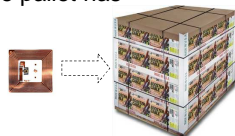
### Scenario



### Q1 (a)

There is an **RFID devices** attached to each **pallet of product**.

- This device records the product's **barcode**, a **pallet ID** and **date of manufacture**.
- The device includes also a **temperature monitor** which is initialized when the product is manufactured, so the RFID also transmits the **highest temperature** the pallet has encountered.



### Q1 (a)

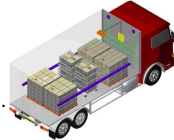
What is the **schema** for the **RFID device**?

- Each RFID device has its own memory for storing the data about itself.
- RFIDDevice (**PalletID**, **ProductCode**, **DateManufactured**, **Highest\_Temp**)

### Q1 (a)

The RFID device is read by one of a number of **readers**:

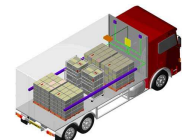
- Some in **warehouses** and
- Some in the **refrigerated trucks** in which the product is transported.
- Each reader adds to the **RFID data** with the **reader's ID, type (Warehouse, therefore fixed; Truck, therefore moving), date and time of reading, and the latitude and longitude** (geographical position) of the reader.
- Readers in **trucks** have **origin and destination warehouse reader IDs** as well.



### Q1 (a)

The schema for a **Reader** ?

- Like the RFID Device, each reader also has its own memory about itself.
- Reader (**ReaderID, Type, Date, Time, Latitude, Longitude, OriginWhseID, DestWhseID**)
- **Type** can be either **Warehouse** or **Truck**.



### Q1 (b)

Assume that when a reader is activated, it reads all pallets at its location and transmits the results (**data stream**) to the information system by initiating a transaction. What is the **schema** for the table of messages (data streams)?

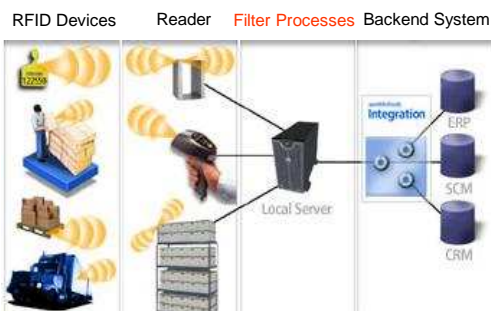
- This schema should be regarded as a format for data streams. As readers are reading messages from RFID devices, the messages (data streams) to the backend system are in this format.
- Readings (**ReaderID, PalletID, Type, Date, Time, Latitude, Longitude, ProductCode, DateManufactured, HighestTemp, OriginWhseID, DestWhseID**)
- The attributes in **blue** and **red** are functionally dependent on **ReaderID** and **PalletID** respectively.

### Q1 (c)

How does it make sense to think of a reader reading an RFID device a **join** between two fragments (**RFIDDevice** and **Reader**)?

- RFIDDevice (**PalletID, ProductCode, DateManufactured, Highest\_Temp**)
- Reader (**ReaderID, Type, Date, Time, Latitude, Longitude, OriginWhseID, DestWhseID**)
- There are no logical join attribute between the both schemas.
- When the reader is reading a pallet, this means they are in the same place. Thus, **Latitude** and **Longitude** are implied as **physical join attributes**.

### How RFID Technology Works



### Q1 (d)

There are a number of **filter processes** for data streams:-

- CreationConsumption
- RefrigerationFailure
- WarehouseInventory
- PalletsInTransit

Why are the filter processes necessary ?

- Usually, a fraction of the inflow data streams is valuable to organizations, so the useful data is extracted through the **filter processes**.

## Q1 (d)

What other tables (in the backend system) are needed for each of the filter processes? Show schemas.

- CreationConsumption
  - Pallets (PalletID, DateCreated, DateConsumed)
- RefrigerationFailure
  - RefFailure (ReaderID, PalletID, Type, Date, Time, Latitude, Longitude, ProductCode, DateManufactured, HighestTemp, OriginWhseID, DestWhseID)

## Q1 (d)

- WarehouseInventory
  - Pallets (PalletID, DateCreated, DateConsumed, WhseReaderID, DateEntered, TimeEntered, InTruckReaderID, DateLeft, TimeLeft, OutTruckReaderID, DestWhseReaderID, CreateQ, ConsumeQ)
  - Where CreateQ and ConsumeQ are both Booleans
- PalletsInTransit
  - Pallets (PalletID, DateCreated, DateConsumed, WhseReaderID, DateEntered, TimeEntered, InTruckReaderID, DateLeft, TimeLeft, OutTruckReaderID, DestWhseReaderID, CreateQ, ConsumeQ, InTransitQ)
  - Where InTransitQ is a Boolean

## Q1 (e)

Give an SQL query implementing each filter process.

- First time a reading for this pallet is made.
- CreationConsumption:

```
INSERT INTO Pallets (PalletID, DateCreated, DateConsumed)
SELECT PalletID, Date, Null FROM Readings
WHERE Reading.PalletID NOT IN SELECT PalletID
FROM Pallets.
```

## Q1 (e)

- RefrigerationFailure
  - We concern the highest temperature over a limit.
  - Readings (ReaderID, PalletID, Type, Date, Time, Latitude, Longitude, ProductCode, DateManufactured, HighestTemp, OriginWhseID, DestWhseID)

```
INSERT INTO RefFailure (*)
SELECT * FROM Reading
WHERE HighestTemp > Threshold
```

## Q1 (e)

- WarehouseInventory
- Pallets (PalletID, DateCreated, DateConsumed, WhseReaderID, DateEntered, TimeEntered, InTruckReaderID, DateLeft, TimeLeft, OutTruckReaderID, DestWhseReaderID, CreateQ, ConsumeQ, InTransitQ)

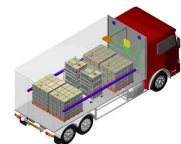
```
//Pallet coming into a warehouse
UPDATE Pallets
SET WhseReaderID=Readings.ReaderID,
    DateEntered=Readings.Date,
    TimeEntered=Readings.Time
WHERE PalletID IN
SELECT PalletID FROM Readings, Pallets
WHERE Readings.Type='Warehouse'
//Pallet is in a warehouse
AND Pallets.WhseReaderID<=>Readings.ReaderID
//But pallet not recorded as being in that warehouse
```



## Q1 (e)

- Pallets (PalletID, DateCreated, DateConsumed, WhseReaderID, DateEntered, TimeEntered, InTruckReaderID, DateLeft, TimeLeft, OutTruckReaderID, DestWhseReaderID, CreateQ, ConsumeQ, InTransitQ)

```
//Pallet leaving a warehouse
- UPDATE Pallets
  SET OutTruckReaderID=Readings.ReaderID,
      DateLeft=Readings.Date,
      TimeLeft=Readings.Time
  WHERE PalletID IN
  SELECT PalletID FROM Readings, Pallets
  WHERE Readings.PalletID = Pallets.PalletID
  AND Readings.Type= 'Truck'
  //Pallet is in a truck
  AND Pallets.DateLeft=NULL
  //But Pallet not recorded as having left the warehouse
- PalletsInTransit: InTransitQ updated after pallets leaving
warehouse event.
```



### Q1 (f)

Which of the filter processes requires permanent storage of the reader data (data streams from readings)?

- None of the processes require permanent storage of the reader data.
- Each process is only concerned with an event that can be detected from the stream of readings, e.g. dateCreated for Pallets.

### Q1 (g)

For the filter processes that don't require permanent storage of the reader data, what is permanently stored as a result of the filter?

- They all contain a filter for the detection of an event.
- What they store is record of that event having occurred.

### Q1 (h)

Why is what is stored in Q1 (g) not a replica of the reader data (data streams – readings) not permanently stored?

- What is stored is not the reading itself.
- The fact that the reading has triggered a significant event.
- That a significant has occurred is determined by the process, not by the reading.