1.264 Final Exam
Fall, 2004

Name: _____

Exam guidelines:

1. 3 hours are allowed to complete the exam.
2. Open notes, open book.
3. No laptop computers or calculators are allowed. You have some arithmetic to do by hand in one question.
4. No cell phones or messaging devices are allowed. Please turn off any that you have brought.
5. Short answer questions: Your answers are limited to a maximum of 2-4 sentences or phrases. Two sentences or phrases are completely adequate to answer any of these questions; 4 sentences are not necessary. Demonstrate that you understand the principles and key points. You will receive full credit for an answer if you make the principal observation(s) that the question is asking for. Details are not necessary.

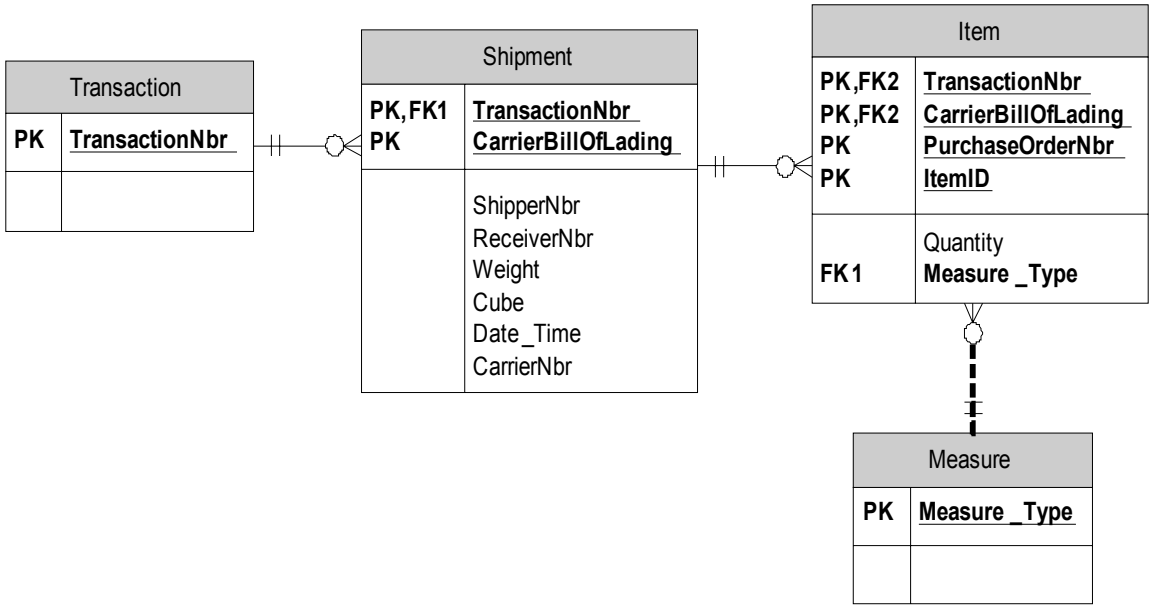
1. **Data exchange between supply chain partners (35 points)**

You wish to exchange shipment notices with a supply chain partner. Each of you sends product to the other, and you have agreed that the contents of a shipment notice are:

1. Transaction:
    a. Transaction number (unique, generated for this transaction)
2. Shipment level detail (transaction may have multiple shipments):
    a. Shipper number (unique identifier)
    b. Receiver number (unique identifier)
    c. Weight (pounds)
    d. Cube (cubic feet)
    e. Date/time shipped
    f. Carrier number (unique identifier)
    g. Carrier bill of lading (shipment) number (unique, generated by carrier)
3. Item level detail within shipment:
    a. Purchase order number (unique, generated by shipper)
    b. Item ID (unique within a purchase order)
    c. Quantity
    d. Measure (item, case, pallet, or container—these are the only choices)

a. Draw a data model for these entities and attributes. (10 points)
    - Draw a box for each entity; give each an appropriate name.
    - List the attributes in the box for each entity.
    - Indicate the primary key for each entity by placing the phrase (PK) next to its name.
    - Draw all relationships between the entities in the model.
    - Indicate foreign keys by placing the phrase (FK) next to attributes that are foreign keys.
    - Indicate the cardinality of each relationship: many-many, many-one or one-one. Use crows foot notation; if you use another notation, define it.
    - Do **not** include tables that would be useful but beyond the scope of the data defined above. For example, do **not** include a table that gives the shipper name associated with the shipper number.

Data model:


Transaction
PK
TransactionNbr
Shipment
PK, FK1
PK
TransactionNbr
CarrierBillOfLading
ShipperNbr
ReceiverNbr
Weight
Cube
Date_Time
CarrierNbr
Item
PK,FK2
PK,FK2
PK
PK
TransactionNbr
CarrierBillOfLading
PurchaseOrderNbr
ItemID
FK1
Quantity
Measure _Type
Measure
PK
Measure _Type

b. Assume that you are using XML for data transfer. Write out a DTD file that corresponds to your Data Model. This DTD should validate the XML document that will be sent. (10 points)
    - Assume that all fields are mandatory and are parsed character data
    - The DTD must enforce all the business rules contained in your data model.

```
<?xml version="1.0" encoding=”utf-8”?>

<!ELEMENT Transaction (TransactionNbr, Shipment+)>

<!ELEMENT Shipment (CarrierBillOfLading, ShipperNbr,
ReceiverNbr, Weight, Cube, Date_Time, CarrierNbr, Item+)>

<!ELEMENT Item (ItemID, PurchaseOrderNbr, Quantity)>
<!ATTLIST Item
   Measure (Item|Case|Pallet|Container) “Item”>

<!ELEMENT TransactionNbr (#PCDATA)>

<!ELEMENT CarrierBillOfLading (#PCDATA)>
<!ELEMENT ShipperNbr (#PCDATA)>
<!ELEMENT ReceiverNbr (#PCDATA)>
<!ELEMENT Weight (#PCDATA)>
<!ELEMENT Cube (#PCDATA)>
<!ELEMENT Date_Time (#PCDATA)>
<!ELEMENT CarrierNbr (#PCDATA)>

<!ELEMENT PurchaseOrderNbr (#PCDATA)>
<!ELEMENT ItemID (#PCDATA)>
<!ELEMENT Quantity (#PCDATA)>
```

c. Write out an appropriate XML document for the following transaction: (10 points)

- Transaction: AB3554
- Shipment level detail:
    i. Shipper 56444
    ii. Receiver 84532
    iii. Weight 40,000
    iv. Cube 2,000
    v. Date/time: December 15, 2004
    vi. Carrier TRZ33442
    vii. Carrier bill of lading number 8Z222
    viii. Item level detail
        1. Item: A44, purchase order: T554, quantity: 48, measure: cases
        2. Item: G64, purchase order: T554, quantity: 24, measure: cases
        3. Item X77, purchase order: E433, quantity 1, measure: pallets
- Shipment level detail:
    i. Shipper 56444
    ii. Receiver 30495
    iii. Weight 20,000
    iv. Cube 1,000
    v. Date/time: December 14, 2004
    vi. Carrier TRZ33442
    vii. Carrier bill of lading number 99Y74
    viii. Item level detail
        1. Item:GH3, purchase order: T554, quantity: 4, measure: cases
        2. Item: 9UY, purchase order: T554, quantity: 20, measure: items
        3. Item R53, purchase order: F990, quantity 1, measure: items

XML file:

```
<?xml version="1.0" encoding="utf-8"?>
<Transaction>
      <TransactionNbr>AB3554</TransactionNbr>
      <Shipment>
            <CarrierBillOfLading>8Z222</CarrierBillOfLading>
            <ShipperNbr>56444</ShipperNbr>
            <ReceiverNbr>84532</ReceiverNbr>
            <Weight>40000</Weight>
            <Cube>2000</Cube>
            <Date_Time>"DEC-15-2004"</Date_Time>
            <CarrierNbr>TRZ33442</CarrierNbr>
            <Item Measure=”Case”>
                  <ItemID>A44</ItemID>
                  <PurchaseOrderNbr>T554</PurchaseOrderNbr>
                  <Quantity>48</Quantity>
            </Item>
            <Item Measure=”Case”>
                  <ItemID>G64</ItemID>
                  <PurchaseOrderNbr>T554</PurchaseOrderNbr>
                  <Quantity>24</Quantity>
            </Item>
            <Item Measure=”Pallet”>
                  <ItemID>X77</ItemID>
                  <PurchaseOrderNbr>E433</PurchaseOrderNbr>
                  <Quantity>1</Quantity>
            </Item>
      </Shipment>
      <Shipment>
            <CarrierBillOfLading>99Y74</CarrierBillOfLading>
            <ShipperNbr>56444</ShipperNbr>
            <ReceiverNbr>30495</ReceiverNbr>
            <Weight>20000</Weight>
            <Cube>1000</Cube>
            <Date_Time>"DEC-14-2005"</Date_Time>
            <CarrierNbr>TRZ33442</CarrierNbr>
            <Item  Measure=”Case”>
                  <ItemID>GH3</ItemID>
                  <PurchaseOrderNbr>T554</PurchaseOrderNbr>
                  <Quantity>4</Quantity>
            </Item>
            <Item Measure=”Item”>
                  <ItemID>9UY</ItemID>
                  <PurchaseOrderNbr>T554</PurchaseOrderNbr>
                  <Quantity>20</Quantity>
            </Item>
            <Item Measure=”Item”>
                  <ItemID>R53</ItemID>
                  <PurchaseOrderNbr>F990</PurchaseOrderNbr>
                  <Quantity>1</Quantity>
            </Item>
      </Shipment>
</Transaction>
```

Answer the following questions. Give short answers only:

d. What must be added to the XML file that you created in step c above to ensure that it is validated using the DTD file you created in step b? (2 points)

Add as the second line:

```
<!DOCTYPE TRANSACTION SYSTEM "YourDTDFilename.dtd">
```

e. Compare and contrast a DTD file with an XML Schema (xsl) file. List the important differences. (3 points)

A DTD file specifies a data model and defines the tags in an XML file. It is used to validate an XML file. A DTD is itself not an XML file.

An XSL file also specifies a data model and defines the tags in an XML file, and is used to validate an XML file. It can also direct how an XML file can be transformed into another XML file or for presentation (fonts, colors, etc.). An XSL file is itself an XML file defined according to a specific schema. It has greater capabilities than a DTD file and is a later, more capable standard.

2. **Security (25 points)**

If you were to exchange XML documents regularly with your supply chain partner for advance shipment notices, as you designed in question 1, please outline the security steps that are required. Describe the most serious attacks that could be attempted on your supply chain data exchange.

a. Who would attack, and for what purpose? Describe at least two possibilities. (5 points)

i. **Thieves wishing to identify high-value shipments or containers to rob.**
ii. **Thieves wishing to manipulate financial records, to generate false billings, for example**
iii. **Blackmailers threatening to shut down key operations unless they are paid protection money or other such**
iv. **Spies for competitors trying to obtain confidential information**

b. List at least 5 attacks, in order of seriousness. (5 points)

i. **Theft of sensitive data from databases**
ii. **Alteration of sensitive data by unauthorized access on server**
iii. **Interception of data going across the network between supply chain partners**
iv. **A denial of service attack to disrupt communications, or the threat of one**
v. **Damage to server or client software through intrusion, to disrupt operations or to install Trojans, worms, etc.**

c. What effect would each one have on your supply chain if it succeeded? Give one example for each attack. (5 points)

i. **Obtain shipment data for theft**
ii. **Cover up theft or fraud**
iii. **Determine if thefts have been detected and if they are suspects**
iv. **Disruption of shipments could result in substantial supply chain expense or worse, if factory operations must shut down**
v. **Longer term continuing problems with theft, fraud, disruptions**

d. What would you need to do to prevent or recover from each? (5 points)

i. **Secure servers, databases; review employee involvement; improve passwords, etc.**
ii. **Same as 1.**
iii. **Use encryption on network traffic**
iv. **Have backup nodes with different IP addresses, URLs. Use software to combat DOS attacks**
v. **Operating system and application reinstallations, virus checkers, passwords, security procedure improvements more generally**

e. Would your security measures be different if monetary information (shipment value, contract terms, freight cost, etc.) were being exchanged via the network between you and your supply chain partner? Why or why not? (3 points)

The financial data is present in your databases, which are likely to be the most vulnerable point of attack. Encryption protects these across the network. You would probably take greater care that your partner's nodes (servers, databases, etc.) are well secured if they are storing your financial data, but you would probably do nothing different across the network.

f. Based on your analysis of risks, what protocol(s) would you use to secure the network link between your database and your supply chain partner's? (2 points)

SSL, using at least 128 bit symmetric keys and at least 1024 bit public keys. 256 and 2048, respectively, if possible.

3. **System architecture and data communications (20 points)**

You and your supply chain partner currently exchange approximately 100 advance shipment notices per day, and you both handle them using HTTP servers communicating via the public Internet via a DSL connection to an Internet service provider. Each shipment contains approximately 20,000 individual items, packed in a small number of cases or pallets. You and your partner agree to use RFID (radio frequency identification) technology starting in 2005 and must assess whether your Web servers, databases and communications links can handle the volume of data.

The situation is:

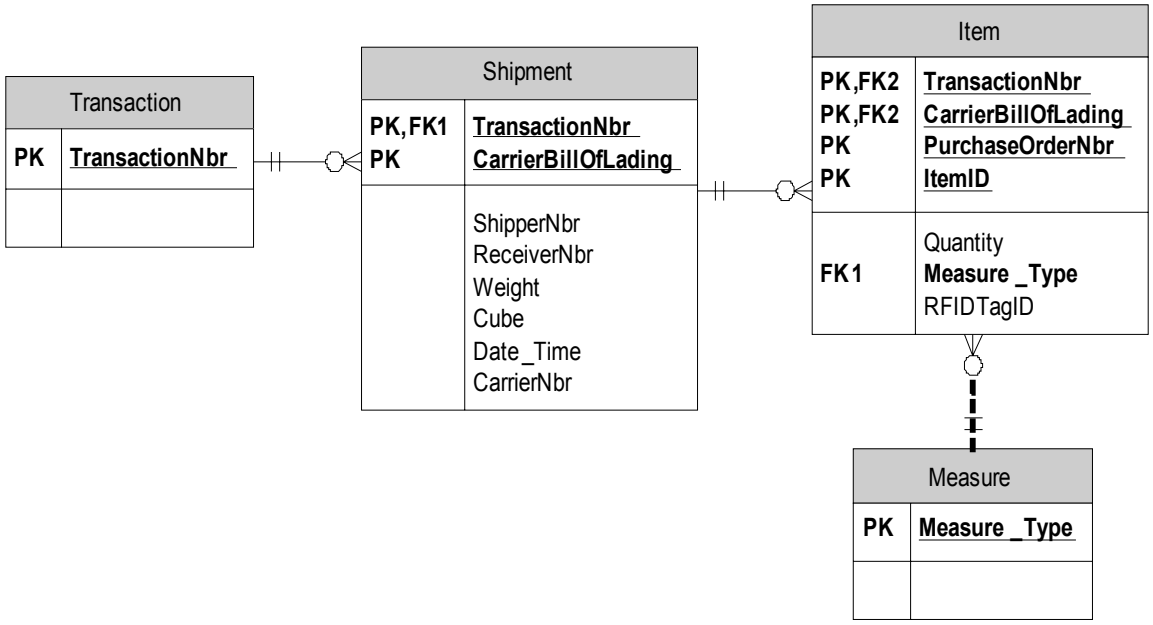
1. Each item must have an RFID tag. The RFID tag contains 1,024 bits of data.
2. Each shipment contains 20,000 individual items
3. There are 100 shipments per day. One shipment at a time is scanned.
4. The shipment is scanned at the origin and destination, and the supply chain partners exchange this data: the shipper sends the origin scan to the receiver, and the receiver sends the destination scan to the shipper.
5. The scanner reads all the data in 60 seconds, and it must be written to the database in that time, because the scanner must then flush all that data from its buffers to be ready to read the next shipment, which may be waiting to be scanned.

Questions:

a. What entities or attributes must be added to the data model you created in question 1a to model the RFID tags? Each tag has a unique serial number and is associated with a unique item ID, which is associated with a purchase order. There are no tags on cases, pallets, containers, etc. (5 points)
   - Redraw the data model from question 1a below, adding the necessary entities, attributes, keys, relationships, and cardinality in the same way as question 1a. (This is very simple; little change is needed.)

The only change is to add an RFIDTagID attribute to the Item table. The RFIDTag is unique and required, but you don't need to show that in the data model.

Data model including RFID tags:



b. How many disks are required on the database server at each node to handle the I/O rate? (5 points)
   - Assume that writing each RFID tag scan requires 5 reads and 7 writes to disk, since the product item ID must be checked for validity, the log is written and an index is written.
   - Each disk can handle a maximum of 140 input/output (I/O, or read/write) operations per second, though you must limit your I/O rate to 60% of the maximum to maintain acceptable performance. (Use 80 I/O ops/second.)
   - Do not consider the number of disks required for the log, temporary tables, rollback buffers, etc. Just compute the number of disks required to accommodate the raw I/O rate.

   Time to read the data and write it to disk = 60s
   I/O operations per scan= 12
   I/O operations per shipment = 20,000 * 12 = 240,000
   I/O rate = 240,000 / 60 = 4,000 ops/s
   Acceptable I/O rate of each disk = 80 ops/s
   Minimum number of disks needed = 4,000 / 80 = 50

c. Would the existing DSL connection running at 1.5Mbps be sufficient to handle the data transfer between the two partners if the RFID tag scan data had to be exchanged within 20 seconds of it being present in the database at the location it was scanned? (5 points)
   - Compute the bandwidth required. Show your calculations briefly. Each tag read has 1024 bits of data that must be transferred. Nothing else is being sent on the network. Assume only one shipment is sent in the 20 second interval. Assume a maximum network utilization of 70%, which is achievable in uninterrupted bulk data transfers such as these.

     Data to be transferred in 20s = 20,000*1024 bits
     Bandwidth required = 20,000*1024/20 bits/s = 1.024 Mbps
     Capacity of DSL = 1.5 Mbps * 0.7 max utilization= 1.05 Mbps
     1.5Mbps DSL is barely sufficient for the data transfer, but is the correct choice.

d. If the database server at each node (yours and your partner's) handles only the RFID and advance shipment notice XML traffic. An XML shipment notice is sent for each shipment, and the RFID tag scans are sent for each shipment. The XML shipment notice includes the RFID tag number, which is 8 bytes long, but no other RFID information. Will the XML traffic or the RFID traffic place a greater load on the database server? How much more traffic does the higher one impose on the server than the lower one? Or are they essentially equal? Explain briefly (5 points).

An RFID tag read is sent for every item, containing 1024 bits or 128 bytes of information. An XML shipment notice is sent for each shipment, which also includes every item. Each item has approximately 256 bytes of information also, looking at the XML file from question 1c, including the 8 bytes for the RFID tag number. (The exact amount depends on the length of the tag names, etc.) There is some additional information on shipments and transactions in the XML file, but it is very small relative to the item information. Thus, the loads are essentially equal.

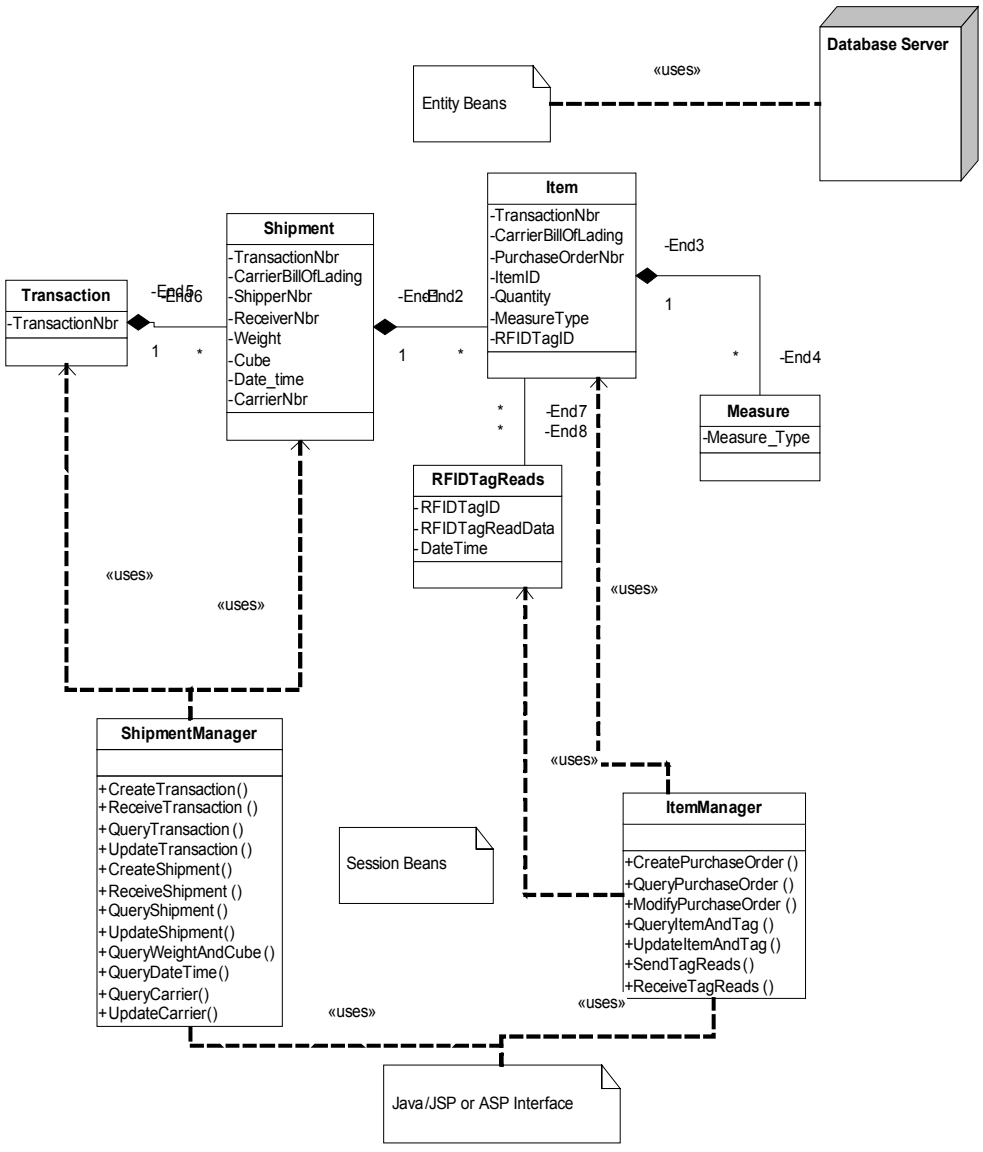
4. **UML (20 points)**

   Draw a UML class diagram, similar to the one shown in class in the system architecture section, that includes:

   a. The classes (entities) that correspond to the data model. (Sorry for making you draw this three times!) These are also called 'entity beans'. Include the XML and RFID scan entities and relationships. Remember to represent the RFID tag reads that are being sent; they are data too. Please use the notation used in the lecture example. (5 points)
   b. The database, as a component, similar to the lecture example. (This is very simple.) (1 point)
   c. The business logic, or 'session beans'. List the elements of business logic (methods) that are required to view and edit the transaction, shipment and item data, and to exchange advance shipment notices and RFID tag scans. Remember that the business logic or session beans contain a list of the processes that are present. These processes are simple; they are fragments or small parts of overall processes. List at least 10 processes and group them into two or more session beans containing similar processes. (10 points)
   d. Assume there is a very simple user interface, implemented in Java or Active Server Pages, for a human user to monitor and query the RFID and XML exchanges. Draw it on the UML diagram. It will use 'session beans' or a set of business logic to support it as well; add methods to the session beans in part c above if needed. (4 points)

You must draw a UML model that includes all the elements listed. Follow these steps. You only need to turn in one drawing that includes all the elements listed in steps a-d.

UML model:


Database Server
«uses»
Entity Beans
Item
-TransactionNbr
-CarrierBillOfLading
-PurchaseOrderNbr
-ItemID
-Quantity
-MeasureType
-RFIDTagID
-End3
1
*
-End4
Shipment
-TransactionNbr
-CarrierBillOfLading
-ShipperNbr
-ReceiverNbr
-Weight
-Cube
-Date_time
-CarrierNbr
Transaction
-TransactionNbr
-End5
-End6
1
*
-End1
-End2
1
*
Measure
-Measure_Type
*
*
-End7
-End8
RFIDTagReads
-RFIDTagID
-RFIDTagReadData
-DateTime
«uses»
«uses»
«uses»
«uses»
ShipmentManager
+CreateTransaction()
+ReceiveTransaction ()
+QueryTransaction ()
+UpdateTransaction ()
+CreateShipment()
+ReceiveShipment ()
+QueryShipment ()
+UpdateShipment()
+QueryWeightAndCube ()
+QueryDateTime()
+QueryCarrier()
+UpdateCarrier()
Session Beans
ItemManager
+CreatePurchaseOrder ()
+QueryPurchaseOrder ()
+ModifyPurchaseOrder ()
+QueryItemAndTag ()
+UpdateItemAndTag ()
+SendTagReads ()
+ReceiveTagReads ()
«uses»
«uses»
Java/JSP or ASP Interface