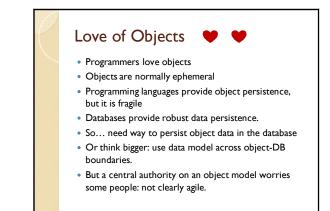
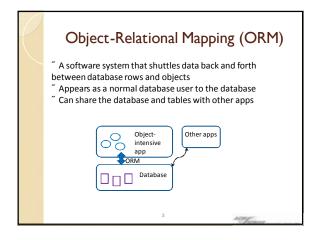
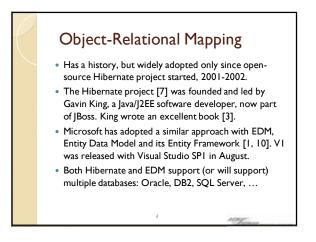
# Object/Relational Mapping 2008: Hibernate and the Entity Data Model (EDM)

Elizabeth (Betty) O'Neil Dept. of Computer Science University of Massachusetts Boston



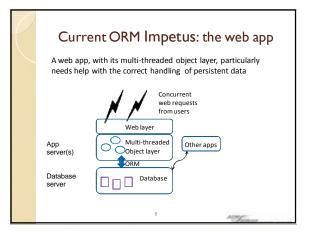


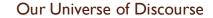


## Java Persistence Architecture (JPA)

- JPA is part of current JEE (previously J2EE), Sun's Java Enterprise Edition
- JPA is a standardized version of the Hibernate architecture
- EJB 3 (current Entity Java Beans) uses JPA for EJB persistence, i.e., persistence of "managed" objects
- JPA and EJB 3 are now available in major application servers: Oracle TopLink 11g, OpenJPA for WebSphere and WebLogic, Hibernate JPA for JBoss
- JPA can be used outside EJB 3 for persistence of ordinary objects, as Hibernate is used in this talk

400





- Object-oriented web apps with database backed data
- Let's consider sites with
  - Possibly many application servers, where the objects live
  - A single database server with plenty of CPUs and disks
- Given today's fast machines and databases, this configuration scales up to many 100s of transactions/second (over 1 M Tx/hour)
- We will concentrate on common characteristics of Hibernate and EDM

#### Outline of Presentation ------Ask questions any time - Schema mapping - Entities and their identity - Relationships - Inheritance - The Pizza Shop Example - Sample code using entity objects

- Development tools, artifacts
- The ORM runtime system
- Transactions, performance

Summary

#### Data Modeling

- Three modeling methodologies:
  - We all know the venerable Chen E-R models for database schemas; the extended E-R models (EER) incorporate inheritance
  - The object modeling approach uses UML class diagrams, somewhat similar to EER
  - The tables of the database define the "physical" schema, itself a model of underlying resources
- The relationship between these models involves schema mapping, covered in last SIGMOD's keynote talk by Phil Bernstein[9]

Q

# Even simple cases need help

- In the simplest case, a program object of class A has fields x, y, z and a table B has columns x, y, z
  - Each instance of A has a row in B and vice versa, via ORM
  - Are we done?
  - $^\circ~$  If x is a unique id, and x, y, and z are simple types, yes.
- -Or some unique id in (x, y, z), possibly composite
- If no unique id in (x, y, z), the object still has its innate identity, but corresponding rows involve duplicate rows, against relational model rules
- So in practice, we add a unique id in this case:
- Class A1 has id, x, y, z and table B1 has id, x, y, z

### Persistent Objects & Identity

- A "persistent object" is an ordinary program object tied via ORM to database data for its long-term state
- The program objects come and go as needed
- Don't confuse this with language-provided persistence (Java/C#), a less robust mechanism
- Persistent objects have field-materialized identity
- It makes sense—Innate object identity depends on memory addresses, a short-lived phenomenon
- So long-lived objects (could be years...) have to be identified this way, it's not the database's fault

ACTIV

11

## Persistent Objects need tracking

- We want only one copy of each unique object in use by an app, a basic idea of OO programming
- Each persistent object has a unique id
- We can no longer can depend on object location in memory to ensure non-duplication
- So we have a harder problem than before—need an active agent tracking objects
- This agent is part of ORM's runtime system
- The ORM uses hashing to keep track of ids, detect duplicates

12

ACTIV

#### **ORM Entities**

- Like E/R entities, ORM entities model collections of realworld objects of interest to the app
- Entities have properties/attributes of database datatypesEntities participate in relationships—see soon (but
- relationships are not "first class" objects, have no attributes)
- Entities have unique ids consisting of one or more properties
- Entity instances (AKA entities) are persistent objects of persistent classes
- Entity instances correspond to database rows of matching unique id

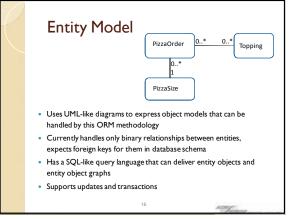
13

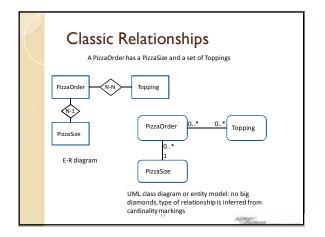
# Value Objects

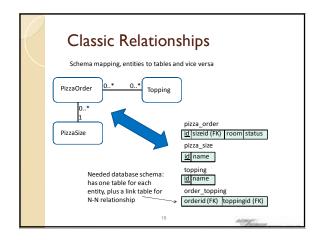
- In fact, persistent objects can be entities or value objects
  Value objects can represent E/R composite attributes and multi-
- valued attributes

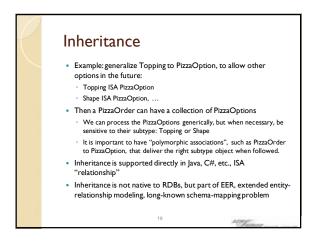
  Example: one address consisting of several address attributes for a customer.
- Programmers want an object for the whole address, hanging off the customer object
- Value objects provide details about some entity, have lifetime tied to their entity, don't need own unique id
- Value objects are called Hibernate "components", EDM "complex types"
- We'll only discuss entities for persistent objects
- For this presentation, persistent object = entity object

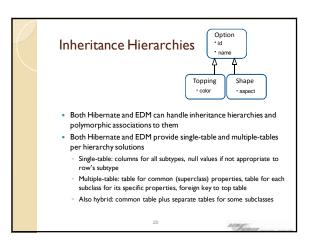
Creating Unique IDs
A new entity object needs a new id, and the database is holding all the old rows, so it is the proper agent to assign it
Note this can't be done with standard SQL insert, which needs predetermined values for all columns
Every production database has a SQL extension to do this
Oracle's sequences
SQL Server's auto-increment datatype
...
The ORM system coordinates with the database to assign the id, in effect standardizing an extension of SQL
Keys obtained this way have no meaning, are called "surrogate keys"
Natural keys can be used instead if they are available.

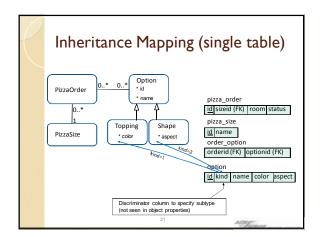


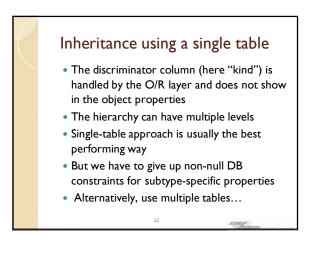


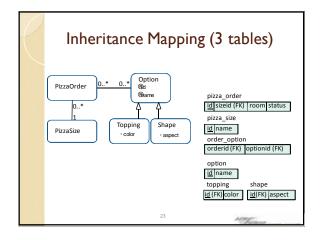


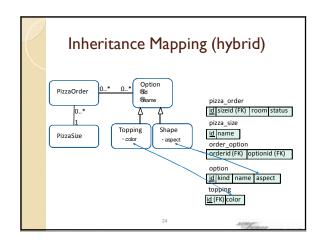


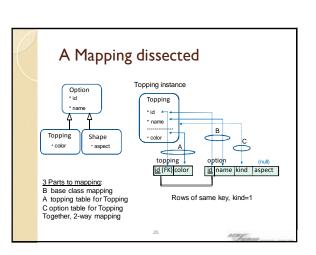


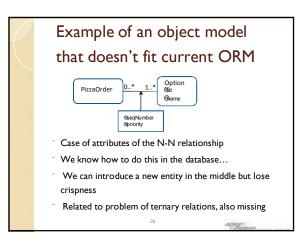


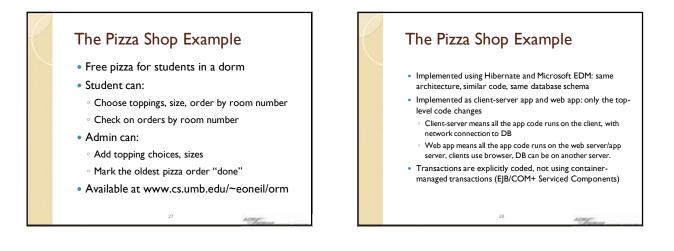


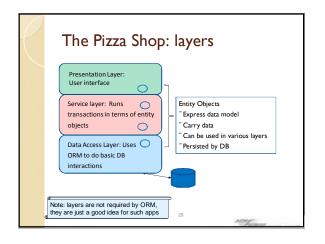


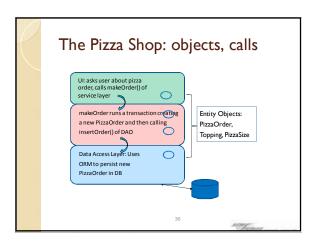


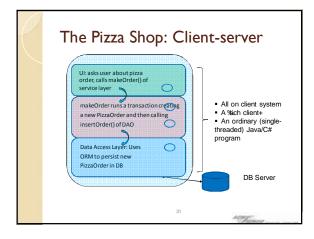


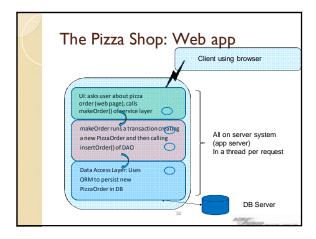


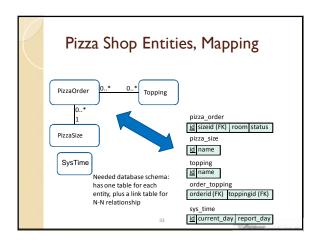


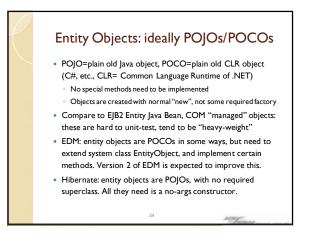


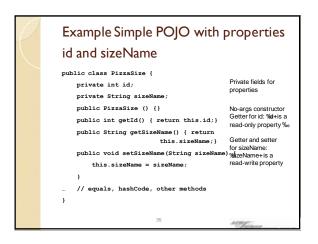


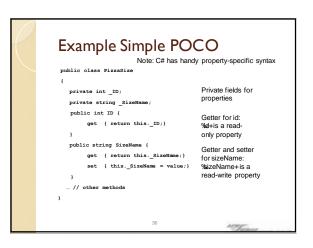


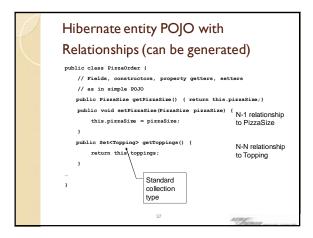


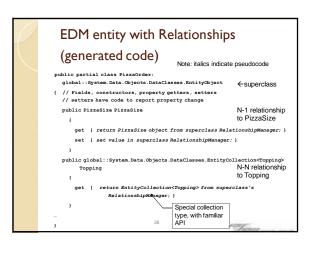


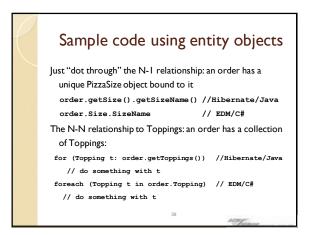














Hibernate: session.persist(order); Later, after commit, or optional earlier session.flush(), order id is valid EDM:

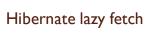
context.AddObject("PizzaOrder", order);

Later, after context. SaveChanges(), order id is valid

40

# Sample code for a "finder"

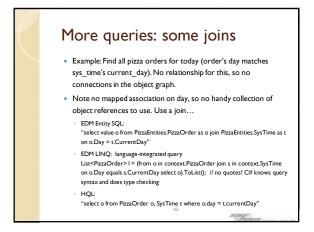
To get FizzaOrder objects for a certain room and day, including available Toppings (and PizzaSize) for each Hibernate HQL: Toppings available in a lazy way List<PizzaOrder> orders = session.createQuery("from PizzaOrder o where o.roomNumber = "+ roomNumber + " and o.day = " + day ).list(); EDM Entity SQL: Toppings available by explicit request here: List<PizzaOrder> orders = new ObjectQuery<PizzaOrder>("select value o from PizzaEntities.PizzaOrder as o where o.RoomNumber = " + roomNumber + " and o.Day = " + day, context). Include("Topping").Include("PizzaSize").ToList();



- In the finder query, Hibernate returns PizzaOrder objects with a "proxy" for the associated PizzaSize and a "collection wrapper" for the Toppings collection
- As long as the runtime system is still alive, first access to such an association results in a DB hit for the actual data
- First access after the runtime is shut down results in an exception: it's too late to be lazy
- This default strategy can be overridden in the mapping file: lazy="false" for PizzaOrder's Toppings, for example.
- EDM: no implicit database access, so need to code what you need

42

ACTOR



# • Group by, having, order by

- Parameterized queries
- Pairs of objects returned, etc.
- Scalars and aggregates
- Build up queries using methods
- Hibernate: direct SQL queries
- stored procedures
- Control of fetch strategies

#### Entity objects in two layers Service layer DAO Create context and transaction Call DAO to get entity objects to • Run query for objects or work on scalar results • Or Call DAO to add objects Add new objects • Or Call objects' own methods • EDM: context.SaveChanges()/ Hibernate: session.flush() (can be done automatically) Commit transaction, drop context

# Entity Objects can do more

44

- So far, entity objects carry data to/from database
- i.e, represent persistent data
- But objects should have related behavior too
- No problem: add methods to entity classes
- Suppose app needs to compute optimal ordering of toppings for building pizza
- List<Topping>x.getToppingsInBuildOrder()
- This should be method of PizzaOrder, an entity

46

## Adding business methods to the

#### entity classes

- Hibernate: relatively simple entity classes, can expand as needed
- EDM: generated code for entity classes: how can we add to it?
- Partial classes of C# come to the rescue:
  - Generated code for PizzaOrder provides data methods in one partial class for PizzaOrder
  - We code business methods in another partial class for PizzaOrder, in another source file.

47

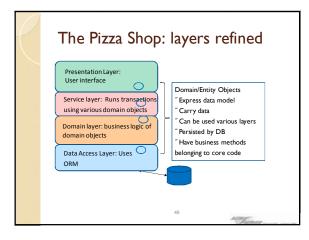
• Compiler puts them together

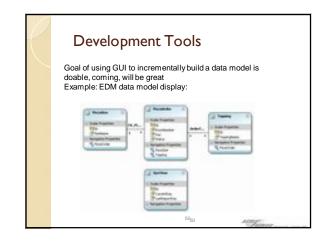
#### "Rich" Domain model

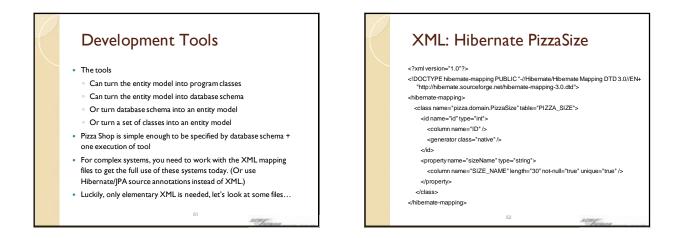
- Domain classes (entities) manage their persistent data, and "rich" ones also provide app-related actions on their data
- Idea of DDD, domain-driven design (Fowler[13], Evans[14], 2004)
- Service layer coordinates actions *between* entities as needed for transactional actions

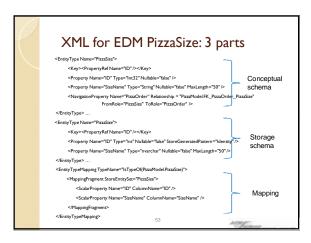
48

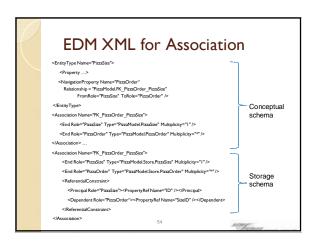
- Service layer should be thin, delimiting transactions and calling on domain classes for most of the work
- Data-only entities dubbed "anemic domain model"

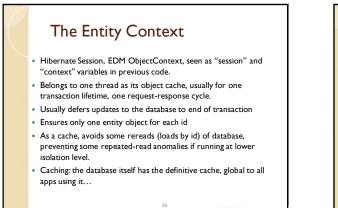


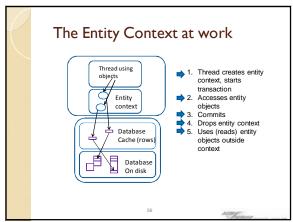


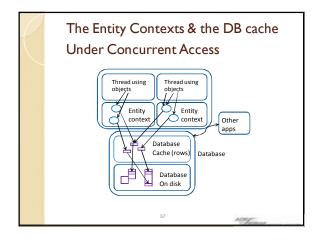


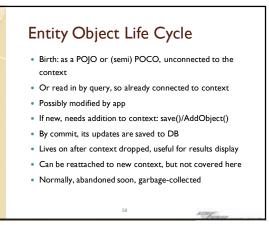












## Queries and the object cache

- When a query delivers entity objects, the id's may already be in the cache
- Need to avoid duplicates, preserve app's changes
- Hibernate flushes changes to DB before query by default
- EDM, by default, preserves the older object of a certain id, avoiding some DB writes at this point

ACTIV

• But tricky: need to flush changes affecting search conditions to maintain query correctness.

59

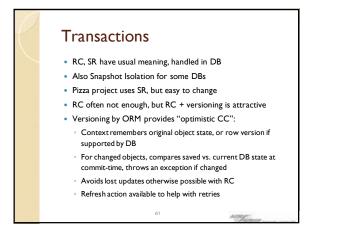
#### Transactions

- Hibernate and EDM are designed for transactional apps
- Both support transactions involving single or multiple
- DBs/resource managers, via JTA or DTC for distributed case (JTA=Java Transaction API, DTC= Microsoft's Distributed Transaction Coordinator)
- Both support both explicit transactions and container-managed transactions
- We're considering simple case of single DB, explicit transactions
- Still have choice in isolation level, mainly:
  - Read-committed (RC)
  - Read-committed with ORM-coordinated versioning

60

ACTOR

Serializable (SR)



#### Conversations

- So far, each request has had one transaction, good enough for Pizza Shop
- Some actions perceived as a unit to the user are made up of several requests
- Example: Read a current bid amount, let user decide on new bid, then make the new bid
- Two DB transactions here, since no DB transaction should span the think time
- The two transactions are related: a "conversation" or "session" or "business transaction" with one user

62

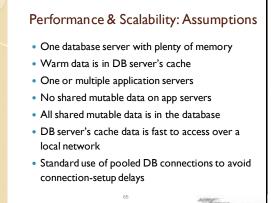
#### Conversations

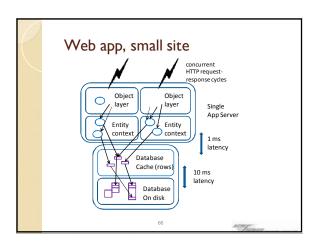
- Example Conversation: Look at bid, think, update bid
- Someone else can slip in a bid update between my look and update
- One solution: make my bid update contingent on the bid amount still being what I saw before, abort second transaction if not
- This is versioning again, now used across multiple system transactions in the same context

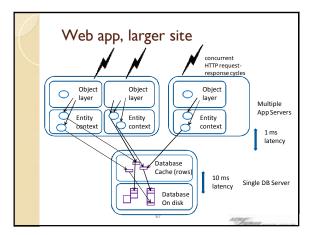
63

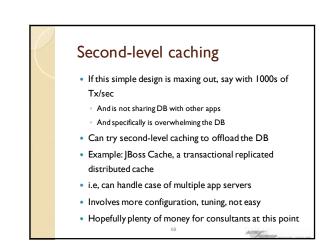
# Conversations and Sessions

- We're thinking about a context spanning several requests of a conversation so it can do version checking
  - ----|--look----|--update--|---Tx1:10 ms 2 min Tx2:10 ms
- Expensive in memory, however, since the context must be kept alive between requests, while the user thinks: above 20 ms vs 2020 ms, factor of 100
- Rather than holding a whole context for a conversation, we can condense it down to a usually-small dataset as part of "session data", save nearly a factor of 100
- This can be held in the common database, but as unshared data, has other possibilities too









#### Summary

- Relational technology continues to prove its worth, and ORM is using it in full
- Only one deficiency of SQL 92 uncovered: standard way to generate new unique id
- The object-relational impedance mismatch has been largely overcome
- No textbooks yet, little academic work: Hibernate semantics are "example-driven"...

69

 See sources, run web app (Hibernate version) at www.cs.umb.edu/~eoneil/orm

# Bibliography

#### In proceedings (subset):

- Adya, Atul, Blakely, Jose, Melnik, Sergey, Meralidhar, S., and the ADO.NET Team, 2007, Anatomy of the ADO.NET Entity Framework, In Proceedings of SIGMOD 2007, ACM Press, New York, NY.
- [3] Bauer, Christian, King, Gavin 2006 Java Persistence with Hibernate, Manning
- [5] Bernstein, Phil, Melnik, Sergey. 2007 Model Management 2.0—Manipulating Richer Mappings. In Proceedings of SIGMOD 2007, ACM Press, New York, NY, 1-12.
   [7] Hibernate, http://www.hibernate.org

#### [10] MSDN Library, 2006 The ADO.NET Entity Framework Overview,

http://msdn2.microsoft.com/en-us/library/aa697427(VS.80).aspx

#### Added:

[13] Fowler, Martin 2003 Patterns of Enterprise Application Architecture, Addison-Wesley

[14] Evans, Eric 2004 Domain-Driven Design, Addison-Wesley 70