# Carnegie Mellon Univ.
# Dept. of Computer Science
# 15-415/615 - DB Applications

*C. Faloutsos – A. Pavlo*

Lecture#27: Final Review



## Administrivia – Final Exam

- **Who:** You
- **What:** http://cmudb.io/f15-final
- **When:** Friday Dec 18th 1:00pm- 4:00pm
- **Where:** GHC 4401
- **Why:** Because you really want to stay on campus until the last day.





## Administrivia – Final Exam

- **What to bring:**
  - CMU ID
  - Calculator
  - Two pages of notes (double-sided)
- **What not to bring:**
  - Live animals





## Administrivia – Course Evals

- Your feedback is strongly needed:
  - https://cmu.smartevals.com
- Things that we want feedback on:
  - Homework Assignments
  - Reading Materials
  - Lectures

# Extended Office Hours

- **Andy**:
  - Monday December 14 @ 12:00pm-1:00pm
  - Wednesday December 16 @ 12:00pm-1:00pm
- **Christos**:
  - Tuesday December 15 @ 11:00am-12:00pm
  - Wednesday December 16 @ 11:00am-12:00pm





# Stuff Before Mid-Term

- SQL
- Sorting





# Query Optimization & Evaluation

- Operator Algorithms:
  - **Selections**: Access paths
  - **Projections & Group Bys**: Hashing vs. Sorting
  - **Joins**: Nested Loop, Index Nested Loop, Sort-Merge, Grace Hash
- Cost Estimations





# Schema Refinement

- Functional Dependencies
  - Armstrong's Axioms
  - Closures
  - Canonical Covers
  - Super Key vs. Candidate Key

# Normalization

- Decomposition:
  - Loseless Joins
  - Dependency Preserving
  - Redundancy Avoidance
- Normal Forms
  - 1NF, 3NF, BCNF





# Database Design & Tuning

- Index Selection & Clustering
- Denormalization
- Decomposition





# Transactions

- ACID
- Conflict Serializability:
  - How to check?
  - How to ensure?
- View Serializability





# Transactions

- Two-Phase Locking
  - Strict vs. Non-Strict
  - Deadlock Detection & Prevention
- Multiple Granularity Locking
  - Intention locks
- B+Tree Latch Crabbing
- Isolation Levels / Anomalies

# Transactions

- Timestamp Ordering Concurrency Control
  - Thomas Write Rule
- Optimistic Concurrency Control
  - Read Phase
  - Validation Phase
  - Write Phase
- Multi-Version Concurrency Control





# Crash Recovery

- Buffer Pool Policies:
  - STEAL vs. NO-STEAL
  - FORCE vs. NO-FORCE
- Write-Ahead Logging
- Logging Schemes
- ARIES Recovery





# Data Warehouses + Mining

- Data cubes
  - CUBE BY
  - ROLAP vs. MOLAP
- Data Mining
  - Supervised Learning (Decision Trees)
  - Unsupervised Learning (Assoc. Rules)





# Column Stores

- DSM vs. NSM
  - Advantages
  - Disadvantages
- Compression Schemes

# Distributed Databases

- Partitioning Schemes
- Two-Phase Commit