



NoSQL Databases

University of California, Berkeley
School of Information
IS 257: Database Management

Announcements



- Questions?
- Assignment 3 has been released (short due date)
 - Going to work on data importing today in lab
- Assignment 4 (Final) to be released shortly
 - Don't forget the web app portion of the final
- Today lecture is pulled from 3 sources

Lecture Outline



- NoSQL Databases
- MongoDB
- MongoDB Atlas Data Import Lab



NoSQL Introduction



- Switch to Presentation from Keith Hare
- NoSQL Database Architecture

NoSQL Database Architectures

NoSQL Database



- NoSQL databases use a variety of file structures and access methods for their operation
- There is very little commonality across the different NoSQL DBs in terms of file storage
- We will look at a couple of examples
 - BerkeleyDB – the grand-daddy of NoSQL DBs
 - MongoDB – One of the best known NoSQL DBs

BerkeleyDB Architecture

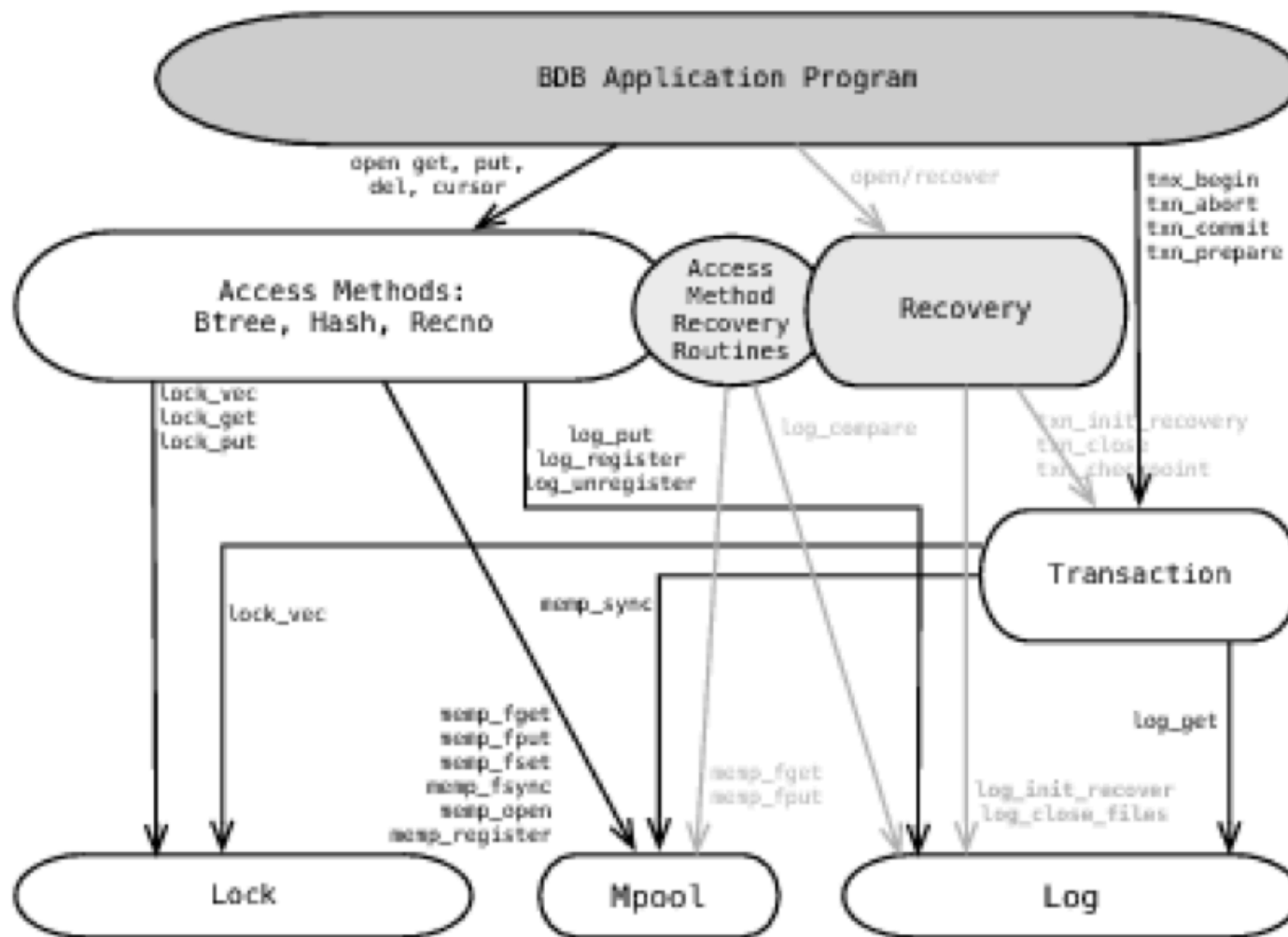


Figure 4.3: Actual Berkeley DB 2.0.6 Architecture.



- The system provides three types of underlying file DBMS access methods
 - RecordID
 - Btree
 - Hashed
- RecordID is a simple numeric record lookup
- Btree uses clever caching to keep the frequently used and higher tree levels in memory
- Hash uses extensible hashing

MongoDB Storage



- MongoDB uses memory-mapped files for data storage
- A memory-mapped file is a file with data that the operating system places in memory by way of the `mmap()` system call. `mmap()` thus maps the file to a region of virtual memory.
- Memory-mapped files are the critical piece of the storage engine in MongoDB.
- By using memory mapped files MongoDB can treat the contents of its data files *as if they were in memory*.

MongoDB Storage



- This provides MongoDB with an extremely fast and simple method for accessing and manipulating data.
- Memory mapping assigns files to a block of virtual memory with a direct byte-for-byte correlation. Once mapped, the relationship between file and memory allows MongoDB to interact with the data in the file as if it were memory.

MongoDB Storage



- How does MongoDB work with memory mapped files?
- MongoDB uses memory mapped files for managing and interacting with all data.
- MongoDB memory maps data files to memory as it accesses documents.
- Data that isn't accessed is not mapped to memory.

Deeper Dive on MongoDB

