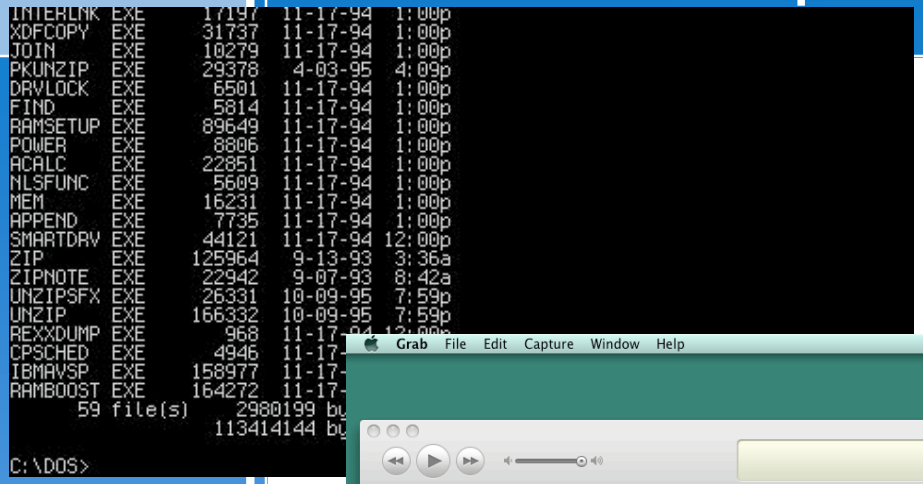
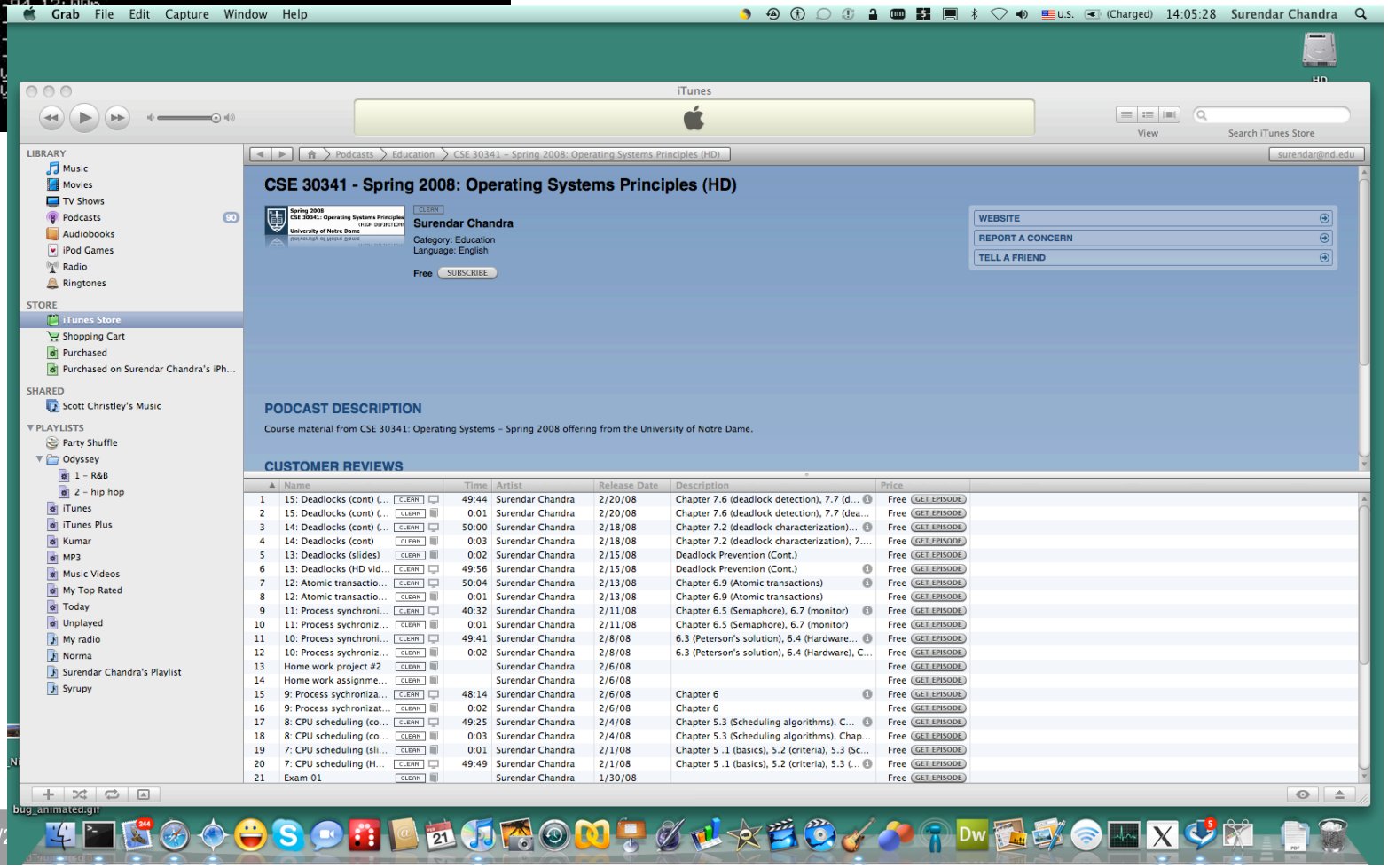
Survey feedback

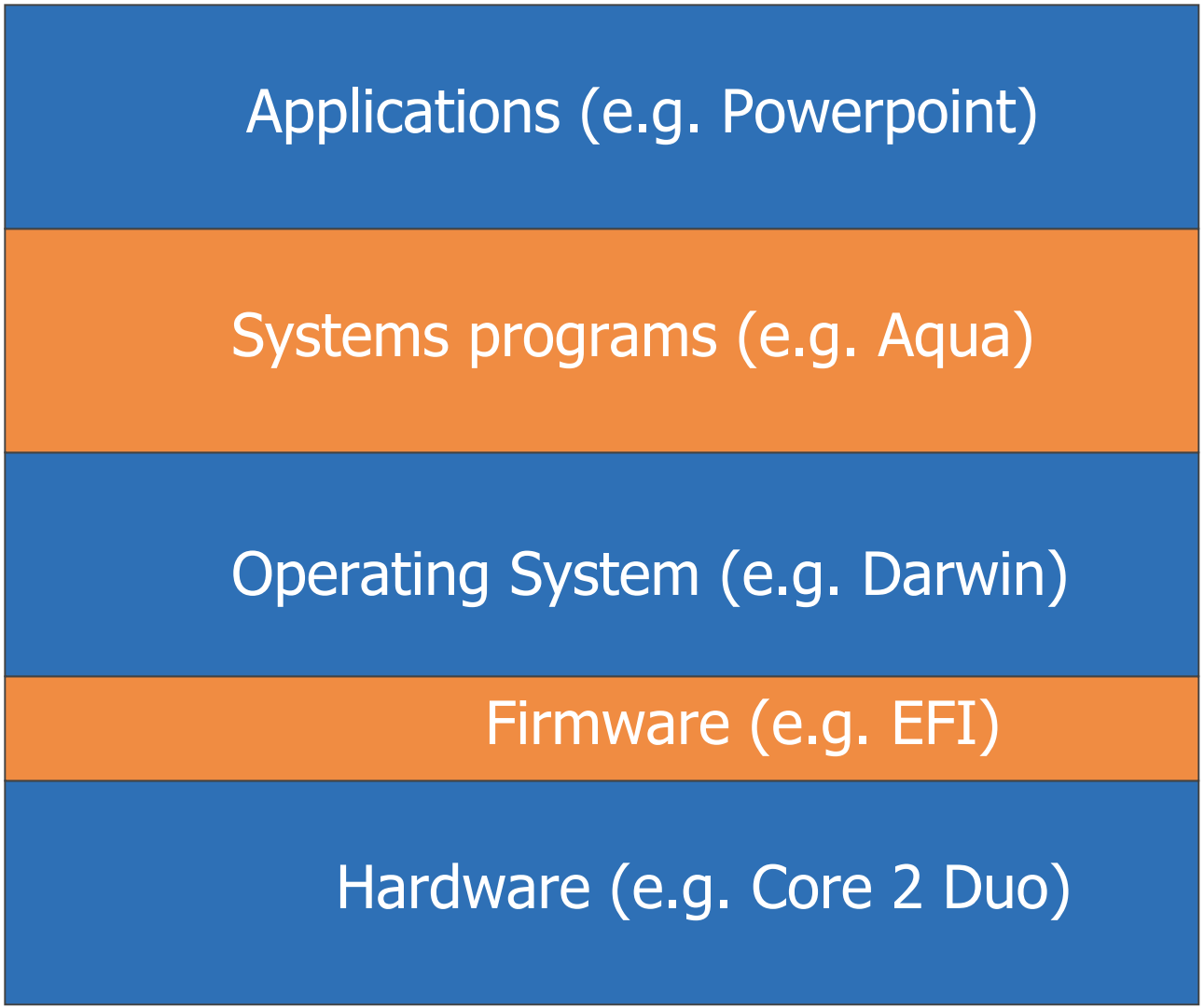
- “I definately agree that we need more code in class. The idea of an operating system for a lot of newbie cse's is what they witness for mac, windows or linux. Very finite and tangible. With the openness of your lectures we get an image of something much more theoretical, so there is an added challenge of connecting the concrete to abstract. I feel that by presenting more (and more 'real', not just a loop segment would help put in context) examples of code we could understand OS's much more clearly and wholly.”

What is an operating system?

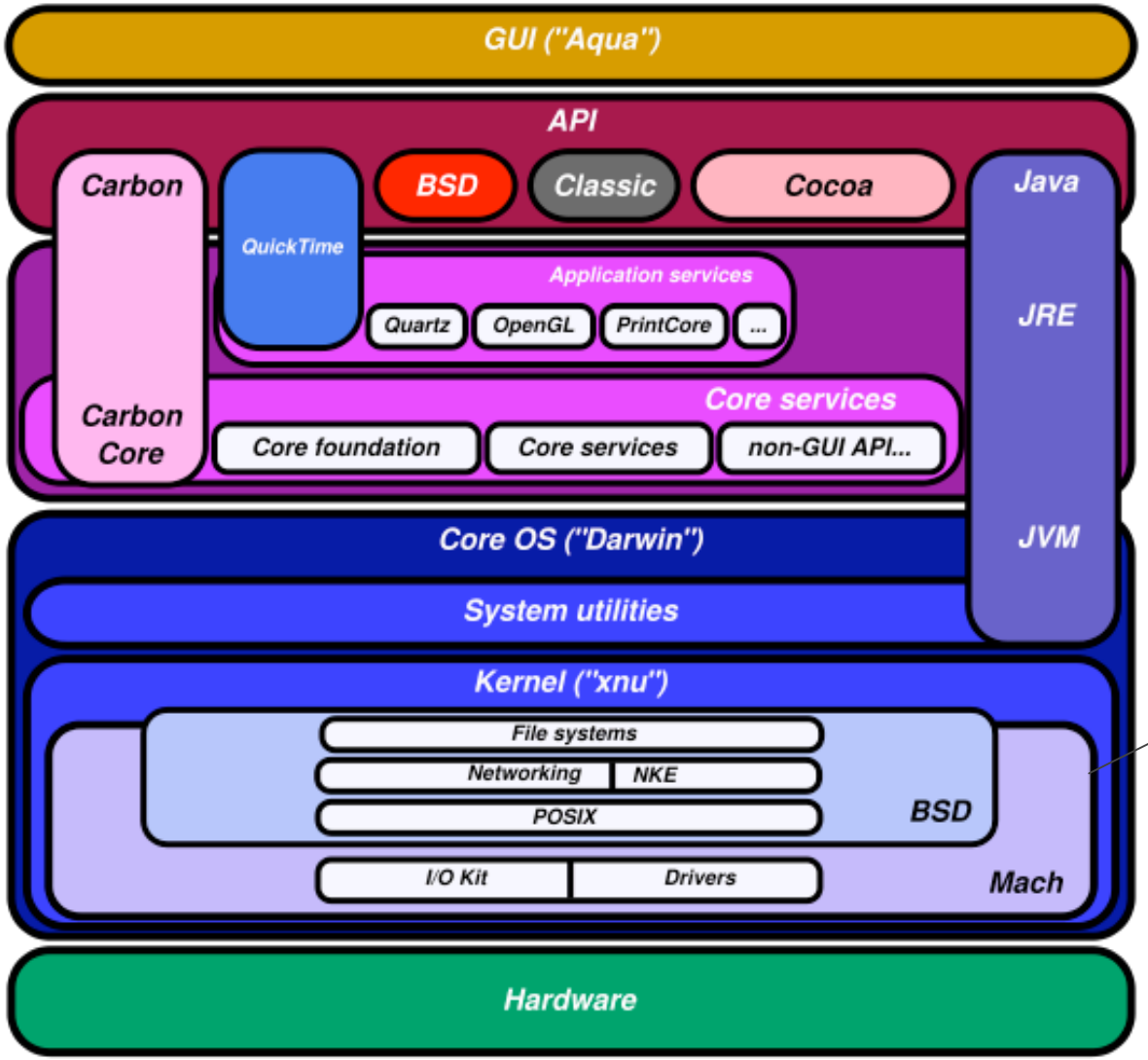




What is an operating systems?



Mac OSX (from Wikipedia)



Our focus

Recap

- Module 1:
  - Process is an abstraction for a program in execution
  - Threads is a way to assign multiple processors to a program
  - PCB internally represents a process to the OS
  - System calls are a way for applications (user level) to communicate and get services from OS (kernel level)
  - Scheduling algorithm (OS service) decides which thread to run and for how long

Module 2

- Challenge: Multiple threads running inside a process can cause race conditions for some applications
  - Threads that did not share data were fine
  - Threads that shared data can give unpredictable results
    - One solution is to force each thread to behave in a predictable fashion - significant slowdown because you cannot use unexpected slack (extra CPU)
    - Another solution is to force some partial order which is flexible enough and yet gives you good performance
      - Notion of conflict serializability

Module 2: Critical sections

- We defined notions of critical sections - to be used by the programmer.
  - need hardware support (TestandSet, Swap)
  - it is sufficiently important that you get OS support
    - With OS support, you can implement block vs spinlock
- The whole module is concerned with the problem faced by application program that exploit a feature provided by the OS (threads) and the solutions. A bit of the solution involves the Operating System. Some of the other services that could be provided by the OS (deadlock prevention/detection) is too hard that no OS actually implements it

Kernel mutex lock

```
__mutex_fastpath_lock(&lock->count, __mutex_lock_slowpath);

#define __mutex_fastpath_lock(v, fail_fn)
do {
        unsigned long dummy;

        typecheck(atomic_t *, v);
         typecheck_fn(void (*)(atomic_t *), fail_fn);
        __asm__ __volatile__(
                LOCK_PREFIX "   decl (%%rdi)    \n"
                        "   jns 1f              \n"
                        "   call "#fail_fn"     \n"
                        "1:"

                :"=D" (dummy)
                : "D" (v)
                : "rax", "rsi", "rdx", "rcx",
                  "r8", "r9", "r10", "r11", "memory");
} while (0)
```

Review

- Critical section problem: Primarily a problem for threaded application that share some data. There is a need to ensure that only one thread gets to be inside the critical section. There is a need to be fair
- Semaphore, monitors are good programming abstraction
- Typical problem with threaded, shared data
  - Bounded buffer
  - Reader-writer: More than one reader inside CS
  - Dining philosophers: solutions dead-lock prone

Atomic transactions

- Notions of Databases
  - Transaction, commit, abort, rollback/roll forward, logs
  - Serializability, conflict serializability

- Deadlocks:
  - A phenomenon faced by applications that used multiple mutually exclusive resources
  - OS can prevent deadlocks during allocation or detect deadlocks after they have happened
    - Applications can use these techniques while requesting locks
  - Most current OS's do not do any of them

Survey concern

- "Programming examples would be good for help in identifying critical sections for multithreaded processes, and how to handle them"
- Identifying critical sections requires a deep understanding of your program. The finer grain you have, the more performance you achieve
  - Steps: Identify all shared variables
    - Lock all accesses to them
  - Deeply understand the program and only lock accesses that are likely to cause a problem
    - Miss a variable that should've been protected - tragedy

Code example

```
Func()
{
    for (int I=start_x; I < end_x; I++)
        for (int j=start_y; j < end_y; j++)
            matrix[I][j]=10;
}
```

thread_create func() with start_x, end_x, start_y, end_y equalling (0,10,0,10), (11,20,11,20)

might not require locks