Lecture 22
Concurrent Programming

14th November 2002

Leaders and followers pattern

Leaders and followers is similar to the bag-of-tasks pattern discussed in the textbook.

Problem: Clients periodically present service requests which requires a significant amount of work that along the way involves waiting for something else to happen. Example - on-line transaction processing.

Forking server

An approach you may have seen before:

```
serversocket = socket(...)
rc = bind(serversocket)
rc = listen(serversocket)
while true {
   newsocket = accept(serversocket)
   if (!fork()) {
      close(serversocket);
      serviceoneclient(newsocket);
   }
}
```

This is the approach used in many Unix servers – web, ftp, telnet, ssh, etc. The problem is that the fork() operation is quite costly both in time and space. It takes a lot of work to construct a new address space, and the new process occupies a significant amount of memory (even if the code and much of the data is implemented using shared memory techniques).

The forking server approach was the only one used in the Apache web server until quite recently.

Threaded server

Using threads instead of processes can reduce the cost. In many threading systems, the memory cost is reduced to a separate, pre-allocated stack per thread, and the time cost is lessened because less address space manipulation is needed to create the thread.

Examples: recent Apache web servers, the AOLServer web server, many services written in Java.

Leaders and followers

Instead of a thread per client, the LF pattern allocates a fixed number of threads, just a sufficient number to keep the server busy. At any time, one thread is designated as the leader and is the only thread allowed to read a client request. Since it doesn't know which client will make the request it uses a selecting operator to wait for any existing client or a new client. When a request arrives, the leader removes that client from the selectable client list, lets one of the follower threads become the new leader and goes off to process the request it received.

```
/* initialization */
serversocket = socket(...)
rc = bind(serversocket)
rc = listen(serversocket)
selectset = emptyset;
setadd(selectset, serversocket);
... create n server threads ...
/* each server thread */
while true {
   if iAmLeader() {
      rc = select(selectset)
      readyfd = findreadyfd()
      if readyfd==serversocket {
         newfd = accept(serversocket)
         selectset = union(selectset, {newfd})
      } else {
         selectset = setremove(selectset,readyfd) /* why? */
         leader = null;
         retireAsLeader();
         readAndServiceOneRequest(newfd);
      }
   else {
      waitToBecomeLeader();
   }
```

```
}

synchronized boolean iAmLeader () {
   return leader==myId;
}

synchronized waitToBecomeLeader() {
   while (leader!= null) {wait()}
   leader = myId;
}

synchronized retireAsLeader() {
   leader = null;
   notify();
}
```

Example1: Akamai

Akamai is a company that sells the service of web caching. Akamai has thousands of servers worldwide on thousands of network. The goal is to have servers that are close in network distance to all internet users. Akamai uses DNS manipulation to direct clients to nearby servers. Once a client request reaches an Akamai server, satisfying the request may involve interaction with Akamai's customer's servers. (Akamai's customer in this case is the company that owns the web site; the client might be that company's customer).

An Akamai-like server might have thousands of connected clients simultaneously, only some of which have an outstanding request. The ones with outstanding requests will have interactions with the customer servers which need to be managed over several interactions – a thread is a natural way to manage the state evolution of these. But we can't afford to tie up a thread for each connected client.

Example 2: NFS server

Before the Solaris had a multithreaded kernel, and still today in Linux, the leaders and followers scheme is used to implement the NFS server. All of the work of an NFS server can happen in the kernel – request packets arrive first in the kernel, file reads and writes can only get to the disk by being handled by the kernel. Thus, although NFS can be implemented in a user-level process, doing so adds overhead that could be avoided by a kernel implementation.

The problem is that the OS kernel in these systems is an asynchronous system, and as we've seen, handling complex state evolution in an asynchronous system over long periods adds a lot of complexity. What the NFS implementors did

was in essence create a fixed-size pool of user processes that do nothing but make a single system call that waits in the kernel for NFS requests. As each request arrives it is assigned to one of these processes which goes through the sequence of steps necessary to satisfy that request. Once the request is satisfied the process is ready to accept the next request.

In this system the NFS server can be working on a fixed maximum number of requests at a time.

Once the kernel evolved to include kernel-level threads subterfuge of user-level processes was not needed any more, and threads could be created as needed to handle the load. If I were designing the system, I would certainly still use a thread pool so that I didn't have to create a thread for each request. But with kernel level threads available I could make the size of the thread pool change dynamically.