OS Services as User Apps

Version control
Application program
Application program
File system A
File system B
Line discipline
TCP/IP
user mode

Microkernel
privileged mode
Why?

• It’s cool …
• Assume that OS coders are incompetent, malicious, or both …
  – OS components run as protected user-level applications
• Extensibility
  – easier to add, modify, and extend user-level components than kernel components
Implementation Issues

- What are the building blocks?
- What is run in privileged mode?
Mach

• Developed at CMU, then Utah
• Early versions shared kernel with Unix
  – basis of NeXT OS
    - basis of Macintosh OS X
• Later versions still shared kernel with Unix
  – basis of OSF/1
• Even later versions actually functioned as working microkernel
  – basis of GNU/HURD project
    - HURD: HIRD of Unix-replacing daemons
    - HIRD: HURD of interfaces representing depth
Mach’s Building Blocks

- **Tasks**
  - represent services/objects
  - holders of access rights
- **Threads**
  - represent virtual processors
- **Ports**
  - communication channels and access rights
- **Messages**
  - carriers of data and access rights

Tasks may be used to construct objects. A task itself is an object, as is a thread. There are additional objects implemented by the kernel, including device objects.
Mach Ports (1)

- Access rights

Client -> Port

Send rights

Port -> Server

Receive rights
Mach Ports (2)

- Communication construct

Client

Request Message

Request Port

Response Port

Server
One may provide a “send-once right” to a port, allowing the recipient to send just one message. This is useful for response ports.
Method Invocation

- Tasks implement objects
- Ports are object references
- Messages are method invocations and responses
We'll take up the topic of optimizing message transfer across address-space boundaries when we discuss virtual memory.
A port reference (or port descriptor) is implemented very much like a file descriptor.
The fat lines represent receive rights, the skinny lines represent send rights.
Note that all threads of a task have access to all the port rights owned by the task. Operations on threads via the thread port include suspending and resuming the thread, setting its register context, getting its register context, and terminating it. Creating a thread is an operation on the task port. Operations on task ports include creating a new task and terminating a task. A thread can set up a method to handle messages on its exception port, which are sent in response to exceptions by the kernel. If a thread hasn’t set up such a method, or what it has setup doesn’t handle a particular exception, the message is sent to the task’s exception port.

A task’s bootstrap port is used to obtain other port rights from some other object (the one that holds its receive rights).
Virtual Memory

- Memory cache objects
  - implemented in kernel
  - represent what's in real memory
- Memory objects
  - implemented in kernel or as user tasks
  - represent what's mapped into real memory
Devices

- Device master port exported by kernel
- Tasks holding send rights may request access to any device
  - send rights given for device port
Successful Microkernel Systems

•
•
•...

Operating Systems In Depth
Attempts

• Windows NT 3.1
  – graphics subsystem ran as user-level process
  – moved to kernel in 4.0 for performance reasons

• Macintosh OS X
  – based on Mach
  – all services in kernel for performance reasons

• HURD
  – based on Mach
  – services implemented as user processes
  – no one uses it, for performance reasons …