inst.eecs.berkeley.edu/~cs61c
CS61C : Machine Structures

Lecture 24
Virtual Memory I

2010-08-02

Instructor Paul Pearce

BE CAREFUL WHAT YOU DOWNLOAD…

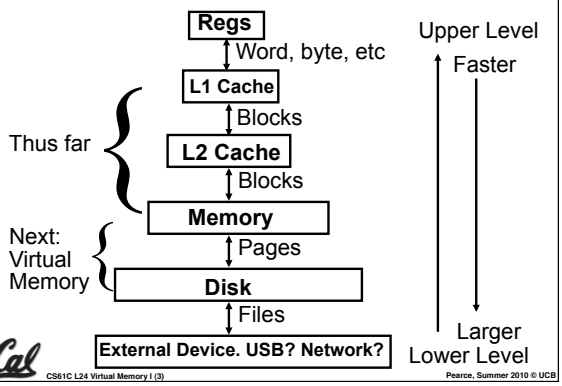
At this year's DefCon security conference, researchers from "Spider Labs" released a proof of concept Android rootkit. This software, once installed, can completely take over an Android system and give the attacker total control over the device, and all of the user's data. Be careful next time you install a wallpaper app from the Android Marketplace…

http://www.reuters.com/article/idUSTRE66T52O20100730

And in Review…

- We've discussed memory caching in detail. Caching in general shows up over and over in computer systems
  - Filesystem cache, Web page cache, Game databases / tablebases, Software memoization, and many more!
- Big idea: if something is expensive but we want to do it repeatedly, do it once and cache the result.
- Cache design choices:
  - Size of cache: speed v. capacity
  - Block size (i.e., cache aspect ratio)
  - Write Policy (Write through v. write back
  - Associativity choice of N (direct-mapped v. set v. fully associative)
  - Block replacement policy
  - 2nd level cache?
  - 3rd level cache?
- Use performance model to pick between choices, depending on programs, technology, budget, ...

Another View of the Memory Hierarchy



Memory Hierarchy Requirements

- **If Principle of Locality allows caches to offer (close to) speed of cache memory with size of DRAM memory, then recursively why not use at next level to give speed of DRAM memory, size of Disk memory?**
- **While we're at it, what other things do we need from our memory system?**

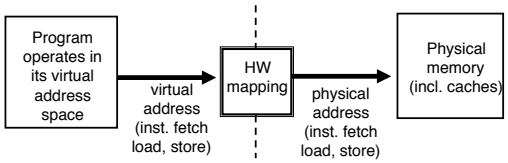
Memory Hierarchy Requirements

- **Allow multiple processes to simultaneously occupy memory and provide protection – don't let one program read/write memory from another**
- **Address space – give each program the illusion that it has its own private memory**
  - **Suppose code starts at address 0x40000000. But different processes have different code, both residing at the same address. So each program has a different view of memory.**

Virtual Memory

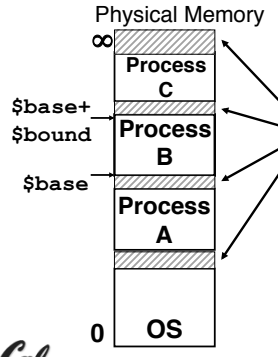
- **Next level in the memory hierarchy:**
  - **Provides program with illusion of a very large main memory:**
  - **Working set of "pages" reside in main memory - others reside on disk.**
- **Also allows OS to share memory, protect programs from each other**
- **Today, more important for protection vs. just another level of memory hierarchy**
- **Each process thinks it has all the memory to itself**
- **(Historically, it predates caches)**

Virtual to Physical Address Translation



- Each program operates in its own virtual address space; ~only program running
- Each is protected from the other
- OS can decide where each goes in memory
- Hardware gives virtual ⇒ physical mapping

Simple Example: Base and Bound Reg



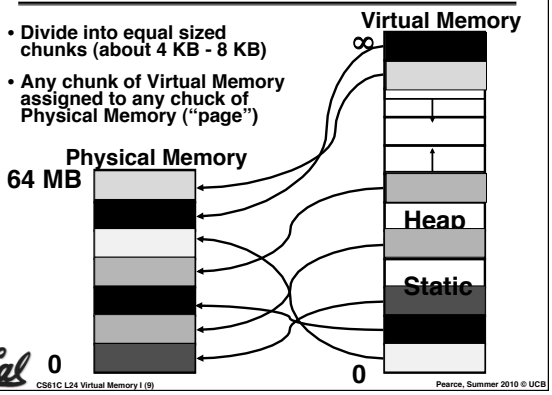
Enough space for Process D, but discontinuous ("fragmentation problem")

- Want:
  - discontinuous mapping
  - Process size >> mem
- Addition not enough!

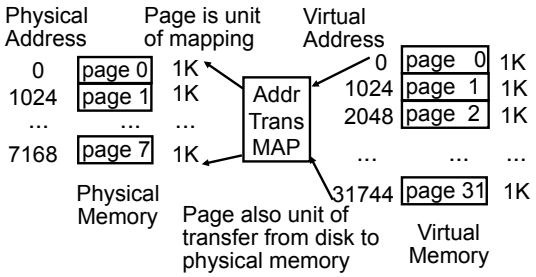
⇒ use Indirection!

Mapping Virtual Memory to Physical Memory

- Divide into equal sized chunks (about 4 KB - 8 KB)
- Any chunk of Virtual Memory assigned to any chuck of Physical Memory ("page")

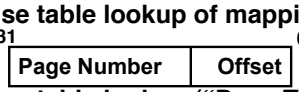


Paging Organization (assume 1 KB pages)



Virtual Memory Mapping Function

- Cannot have simple function to predict arbitrary mapping
- Use table lookup of mappings

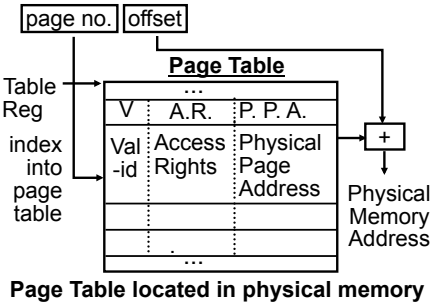


- Use table lookup ("Page Table") for mappings: Page number is index
- Virtual Memory Mapping Function
  - Physical Offset = Virtual Offset
  - Physical Page Number = PageTable[Virtual Page Number]

(P.P.N. also called "Page Frame")

Address Mapping: Page Table

Virtual Address:



Page Table located in physical memory

Page Table

- A page table is an operating system structure which contains the mapping of virtual addresses to physical locations
  - There are several different ways, all up to the operating system, to keep this data around
- Each process running in the operating system has its own page table
  - "State" of process is PC, all registers, plus page table
  - OS changes page tables by changing contents of Page Table Base Register

Administrivia

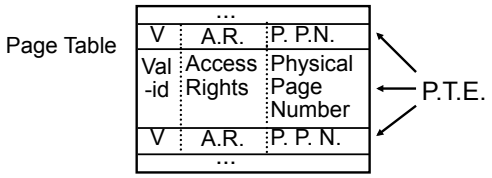
- Project 2 due tonight at midnight.
- Paul has OH today from 12pm-2pm in 511 Soda
  - Noah also has extra OH scheduled. Check your email.
- Grading for Project 2 will be done face to face, Tuesday – Friday from 1pm to 5pm in the lab.
  - You must pick a time that both you AND your partner can make
  - Signups will be first come first serve
  - Here is the URL. Because you came to lecture, you get first crack at it =)
  - http://inst.eecs.berkeley.edu/~cs61c/su10/signups/signup.cgi
- PROJECT 3 IS SOLO! No partner!
- Final exam is Thursday, August 12th, 8am-11am in 10 Evans.

Requirements revisited

- Remember the motivation for VM:
- Sharing memory with protection
  - Different physical pages can be allocated to different processes (sharing)
  - A process can only touch pages in its own page table (protection)
- Separate address spaces
  - Since programs work only with virtual addresses, different programs can have different data/code at the same address!
- What about the memory hierarchy?

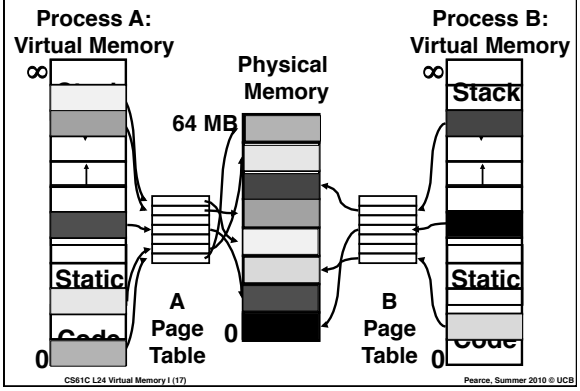
Page Table Entry (PTE) Format

- Contains either Physical Page Number or indication not in Main Memory
- OS maps to disk if Not Valid (V = 0)



- If valid, also check if have permission to use page: Access Rights (A.R.) may be Read Only, Read/Write, Executable

Paging/Virtual Memory Multiple Processes



Analogy

- Book title like virtual address
- Library of Congress call number like physical address
- Card catalogue like page table, mapping from book title to call #
- On card for book, in local library vs. in another branch like valid bit indicating in main memory vs. on disk
- On card, available for 2-hour in library use (vs. 2-week checkout) like access rights

Comparing the 2 levels of hierarchy

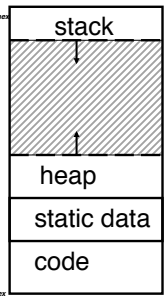
Cache version	Virtual Memory vers.
Block or Line	Page
Miss	Page Fault
Block Size: 32-64B	Page Size: 4K-8KB
Placement: Direct Mapped, N-way Set Associative	Fully Associative
Replacement: LRU or Random	Least Recently Used (LRU)
Write Thru or Back	Write Back
Cache out	Page out / Swap out

Notes on Page Table

- Solves Fragmentation problem: all chunks same size, so all holes can be used
- OS must reserve "Swap Space" on disk for each process
- To grow a process, ask Operating System
  - If unused pages, OS uses them first
  - If not, OS swaps some old pages to disk
  - (Least Recently Used to pick pages to swap)
- Will add details, but Page Table is essence of Virtual Memory

Why would a process need to "grow"?

- A program's *address space* contains 4 regions:
  - stack: local variables, grows downward
  - heap: space requested for pointers via `malloc()`; resizes dynamically, grows upward
  - static data: variables declared outside main, does not grow or shrink
  - code: loaded when program starts, does not change



For now, OS somehow prevents accesses between stack and heap (gray hash lines).

That time is now!

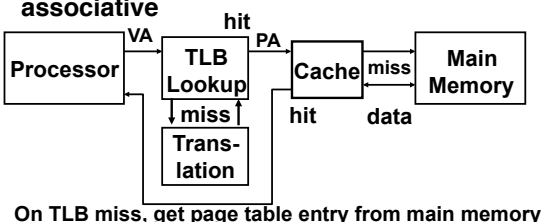
- Before, we stated:
  - For now, OS somehow prevents accesses between stack and heap (gray hash lines).
- How does the OS accomplish this?
- We will mark the bottom of the stack by creating a "guard" page. This would be a page that is simply marked as invalid in it's PTE.
  - Should the heap grow into the "guard" page, a page fault will occur. Since the OS handles page faults, it will recognize that a collision occurred, and generate an error!
  - For more brutal details, ask Paul later

Virtual Memory Problem #1

- Map every address ⇒ 1 indirection via Page Table in memory per virtual address ⇒ 1 virtual memory accesses = 2 physical memory accesses ⇒ SLOW!
- Observation: since locality in pages of data, there must be locality in virtual address translations of those pages
- Since small is fast, why not use a small cache of virtual to physical address translations to make translation fast?
- For historical reasons, cache is called a Translation Lookaside Buffer, or TLB

Translation Look-Aside Buffers (TLBs)

- TLBs usually small, typically 128 - 256 entries
- Like any other cache, the TLB can be direct mapped, set associative, or fully associative



On TLB miss, get page table entry from main memory

More on this tomorrow.

Virtual Memory Problem #2

- If each page is 1KB, we need 10 bits for offset
- This means for a 32-bit address space, there are 22 bits for the virtual page number
  - → 2^{22} page table entries (PTEs)!
- If each PTE is 4 bytes (realistic)…
- That means our page table is 2^{24} bytes, or 16MB.
  - And that's just for 1 process! How many processes do you have running right now? 50? 100?

Virtual Memory Problem #2

- Solution: Multi-level page tables!
  - Modern systems use multiple levels of page tables to deal with such problems.
  - Now the page table can be broken up into pieces just like memory pages.
  - For example, 32-bit x86 uses a 2 level page table scheme You'll learn more about this in CS 162. For now, we'll stick with our simplified model.
  - Checkout Wikipedia's "Page Table" entry if you can't wait.
- Want to know how all this works for Virtual Machines? Take CS 262A!

Peer Instruction

1) Locality is important yet different for cache and virtual memory (VM): temporal locality for caches but spatial locality for VM
2) VM helps both with security and cost

	12
a)	FF
b)	FT
c)	TF
d)	TT

And in conclusion…

- Manage memory to disk? Treat as cache
  - Included protection as bonus, now critical
  - Use Page Table of mappings for each process vs. tag/data in cache
  - TLB is cache of Virtual ⇒ Physical addr trans
- Virtual Memory allows protected sharing of memory between processes
- Spatial Locality means Working Set of Pages is all that must be in memory for process to run fairly well