Virtual Memory II
CSE 351 Winter 2017

Administrivia

- Lab 4 due Friday, how is it going?
- HW3 graded, posted.

- Plan for rest of quarter:
  - HW4 out today, due Mon, Mar 6.
  - Lab 5 out Friday Mar 3, due Mon Mar 13 (cut-off Mar 15).
  - Final: Wed, Mar 15th, **8:30-10:20**, here.
    - Material for the whole quarter. More details next week.
Indirection in Virtual Memory

- Each process gets its own private virtual address space
- Solves: *dealing with small physical memory, memory management, protection, sharing.*

VM and the Memory Hierarchy

- Think of virtual memory as array of \( N = 2^n \) contiguous bytes
- *Pages* of virtual memory are usually stored in physical memory, but sometimes spill to disk
  - Pages are another unit of aligned memory (size is \( P = 2^p \) bytes)
  - Each virtual page can be stored in *any* physical page (no fragmentation!)
Memory Hierarchy: Core 2 Duo

Memory Hierarchy:

- **SRAM** (Static Random Access Memory)
  - L1 I-cache
  - L1 D-cache
  - 32 KB
  - Throughput: 16 B/cycle
  - Latency: 3 cycles

- **L2 unified cache**
  - 4 MB
  - Throughput: 8 B/cycle
  - Latency: 14 cycles

- **Main Memory**
  - ~8 GB
  - Throughput: 2 B/cycle
  - Latency: 100 cycles

- **Disk**
  - ~500 GB
  - Throughput: 1 B/30 cycles
  - Latency: millions

- **SRAM**
  - Static Random Access Memory
- **DRAM**
  - Dynamic Random Access Memory

**Miss Penalty (latency):**
- SRAM: 33x
- DRAM: 10,000x

Virtual Memory (VM)

- Overview and motivation
- VM as a tool for caching
- **Address translation**
- VM as a tool for memory management
- VM as a tool for memory protection
Address Translation

How do we perform the virtual → physical address translation?

Address Translation: Page Tables

- CPU-generated address can be split into:
  - \( n \)-bit address: Virtual Page Number | Page Offset
  - Request is Virtual Address (VA), want Physical Address (PA)
  - Note that Physical Offset = Virtual Offset (page-aligned)

- Use lookup table that we call the page table (PT)
  - Replace Virtual Page Number (VPN) for Physical Page Number (PPN) to generate Physical Address
  - Index PT using VPN: page table entry (PTE) stores the PPN plus management bits (e.g. Valid, Dirty, access rights)
  - Has an entry for every virtual page – why?
Page Table Diagram

- Page tables stored in physical memory
  - Too big to fit elsewhere – managed by MMU & OS
- How many page tables in the system?
  - One per process

Page Table Address Translation

In most cases, the MMU can perform this translation without software assistance
Page Hit

- **Page hit:** VM reference is in physical memory

Page Table (DRAM)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VP 0</td>
<td>PP 0</td>
<td>null</td>
</tr>
<tr>
<td>VP 1</td>
<td>PP 1</td>
<td>null</td>
</tr>
<tr>
<td>VP 2</td>
<td>PP 2</td>
<td>null</td>
</tr>
<tr>
<td>VP 3</td>
<td>PP 3</td>
<td>null</td>
</tr>
<tr>
<td>VP 4</td>
<td>PP 4</td>
<td>null</td>
</tr>
<tr>
<td>VP 5</td>
<td>PP 5</td>
<td>null</td>
</tr>
<tr>
<td>VP 6</td>
<td>PP 6</td>
<td>null</td>
</tr>
<tr>
<td>VP 7</td>
<td>PP 7</td>
<td>null</td>
</tr>
</tbody>
</table>

Virtual memory (disk)

- VP 0
- VP 1
- VP 2
- VP 3
- VP 4
- VP 5
- VP 6
- VP 7

**Example:** Page size = 4 KiB

Virtual Addr: `0x00740b`

VPN: `7`

PPN: `2`

Page Fault

- **Page fault:** VM reference is NOT in physical memory

Page Table (DRAM)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VP 0</td>
<td>PP 0</td>
<td>null</td>
</tr>
<tr>
<td>VP 1</td>
<td>PP 1</td>
<td>null</td>
</tr>
<tr>
<td>VP 2</td>
<td>PP 2</td>
<td>null</td>
</tr>
<tr>
<td>VP 3</td>
<td>PP 3</td>
<td>null</td>
</tr>
<tr>
<td>VP 4</td>
<td>PP 4</td>
<td>null</td>
</tr>
<tr>
<td>VP 5</td>
<td>PP 5</td>
<td>null</td>
</tr>
<tr>
<td>VP 6</td>
<td>PP 6</td>
<td>null</td>
</tr>
<tr>
<td>VP 7</td>
<td>PP 7</td>
<td>null</td>
</tr>
</tbody>
</table>

Virtual memory (disk)

- VP 0
- VP 1
- VP 2
- VP 3
- VP 4
- VP 5
- VP 6
- VP 7

**Example:** Page size = 4 KiB

Provide a virtual address request (in hex) that results in this particular page fault:

Virtual Addr: `???`
Page Fault Exception

- User writes to memory location
- That portion (page) of user’s memory is currently on disk

```
int a[1000];
int main ()
{
    a[500] = 13;
}
```

Page fault handler must load page into physical memory
- Returns to faulting instruction: `mov` is executed again!
  - Successful on second try

Handling a Page Fault

- Page miss causes page fault (an exception)
Handling a Page Fault

- Page miss causes page fault (an exception)
- Page fault handler selects a victim to be evicted (here VP 4)
Handling a Page Fault

- Page miss causes page fault (an exception)
- Page fault handler selects a *victim* to be evicted (here VP 4)
- Offending instruction is restarted: page hit!

Peer Instruction Question

- How many bits wide are the following fields?
  - 16 KiB pages
  - 48-bit virtual addresses
  - 16 GiB physical memory

- 16 KiB pages  
  - 16
- 48-bit virtual addresses  
  - 48
- 16 GiB physical memory  
  - 48
Summary

- Virtual memory provides:
  - Ability to use limited memory (RAM) across multiple processes
  - Illusion of contiguous virtual address space for each process
  - Protection and sharing amongst processes
- Indirection via address mapping by page tables
  - Part of memory management unit and stored in memory
  - Use virtual page number as index into lookup table that holds physical page number, disk address, or NULL (unallocated page)
  - On page fault, throw exception and move page from swap space (disk) to main memory

Virtual Memory (VM)

- Overview and motivation
- VM as a tool for caching
- Address translation
- VM as a tool for memory management
- VM as a tool for memory protection
Review: Terminology

- **Context switch**
  - Switch between processes on the same CPU

- **Page in**
  - Move pages of virtual memory from disk to physical memory

- **Page out**
  - Move pages of virtual memory from physical memory to disk

- **Thrashing**
  - Total working set size of processes is larger than physical memory and causes excessive paging in and out instead of doing useful computation

VM for Managing Multiple Processes

- **Key abstraction:** each process has its own virtual address space
  - It can view memory as a simple linear array

- **With virtual memory,** this simple linear virtual address space **need not be contiguous in physical memory**
  - Process needs to store data in another VP? Just map it to any PP!

![Virtual Address Space Diagram](image)
Simplifying Linking and Loading

- **Linking**
  - Each program has similar virtual address space
  - Code, Data, and Heap always start at the same addresses

- **Loading**
  - `execve` allocates virtual pages for `.text` and `.data` sections & creates PTEs marked as invalid
  - The `.text` and `.data` sections are copied, page by page, on demand by the virtual memory system

---

VM for Protection and Sharing

- The mapping of VPs to PPs provides a simple mechanism to protect memory and to share memory between processes
  - **Sharing**: map virtual pages in separate address spaces to the same physical page (here: PP 6)
  - **Protection**: process can’t access physical pages to which none of its virtual pages are mapped (here: Process 2 can’t access PP 2)
Memory Protection Within Process

- VM implements read/write/execute permissions
  - Extend page table entries with permission bits
  - MMU checks these permission bits on every memory access
    - If violated, raises exception and OS sends SIGSEGV signal to process (segmentation fault)

### Address Translation: Page Hit

1) Processor sends virtual address to MMU (memory management unit)

2-3) MMU fetches PTE from page table in cache/memory
   (Uses PTBR to find beginning of page table for current process)

4) MMU sends physical address to cache/memory requesting data

5) Cache/memory sends data to processor

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA</td>
<td>Virtual Address</td>
</tr>
<tr>
<td>PTEA</td>
<td>Page Table Entry Address</td>
</tr>
<tr>
<td>PTE</td>
<td>Page Table Entry</td>
</tr>
<tr>
<td>PA</td>
<td>Physical Address</td>
</tr>
<tr>
<td>Data</td>
<td>Contents of memory stored at VA originally requested by CPU</td>
</tr>
</tbody>
</table>
Address Translation: Page Fault

1) Processor sends virtual address to MMU
2-3) MMU fetches PTE from page table in cache/memory
4) Valid bit is zero, so MMU triggers page fault exception
5) Handler identifies victim (and, if dirty, pages it out to disk)
6) Handler pages in new page and updates PTE in memory
7) Handler returns to original process, restarting faulting instruction

Hmm... Translation Sounds Slow

- The MMU accesses memory twice: once to get the PTE for translation, and then again for the actual memory request
  - The PTEs may be cached in L1 like any other memory word
    - But they may be evicted by other data references
    - And a hit in the L1 cache still requires 1-3 cycles

- What can we do to make this faster?
  - Solution: add another cache!
Speeding up Translation with a TLB

- **Translation Lookaside Buffer (TLB):**
  - Small hardware cache in MMU
  - Maps virtual page numbers to physical page numbers
  - Contains complete *page table entries* for small number of pages
    - Modern Intel processors have 128 or 256 entries in TLB
  - Much faster than a page table lookup in cache/memory

    ![TLB Diagram](image)

A TLB hit eliminates a memory access!

---

**TLB Hit**

1. CPU VA
2. MMU PA
3. Cache/ Memory
4. TLB PTE
5. Data
A TLB miss incurs an additional memory access (the PTE)
- Fortunately, TLB misses are rare

**Fetching Data on a Memory Read**

1) **Check TLB**
- **Input**: VPN, **Output**: PPN
- **TLB Hit**: Fetch translation, return PPN
- **TLB Miss**: Check page table (in memory)
  - **Page Table Hit**: Load page table entry into TLB
  - **Page Fault**: Fetch page from disk to memory, update corresponding page table entry, then load entry into TLB

2) **Check cache**
- **Input**: physical address, **Output**: data
- **Cache Hit**: Return data value to processor
- **Cache Miss**: Fetch data value from memory, store it in cache, return it to processor
Address Translation

**Virtual Address**

![Diagram showing address translation process]

1. TLB Lookup
2. TLB Hit
3. TLB Miss
   - Page Table "Walk"
   - Protection Check
      - PTE
      - Access Denied
      - Access Permitted

**Protection Fault**
- Physical Address
  - SIGSEGV
  - Check cache

**Summary of Address Translation Symbols**

- **Basic Parameters**
  - $N = 2^n$ Number of addresses in virtual address space
  - $M = 2^m$ Number of addresses in physical address space
  - $P = 2^p$ Page size (bytes)

- **Components of the virtual address (VA)**
  - VPO Virtual page offset
  - VPN Virtual page number
  - TLBI TLB index
  - TLBT TLB tag

- **Components of the physical address (PA)**
  - PPO Physical page offset (same as VPO)
  - PPN Physical page number
Simple Memory System Example (small)

- **Addressing**
  - 14-bit virtual addresses
  - 12-bit physical address
  - Page size = 64 bytes

![Diagram of Virtual Memory Example](image)

Simple Memory System: Page Table

- Only showing first 16 entries (out of ____)
  - **Note:** showing 2 hex digits for PPN even though only 6 bits

<table>
<thead>
<tr>
<th>VPN</th>
<th>PPN</th>
<th>Valid</th>
<th>VPN</th>
<th>PPN</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>28</td>
<td>1</td>
<td>8</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>–</td>
<td>0</td>
<td>9</td>
<td>17</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>1</td>
<td>A</td>
<td>09</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>02</td>
<td>1</td>
<td>B</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>–</td>
<td>0</td>
<td>C</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>1</td>
<td>D</td>
<td>2D</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>–</td>
<td>0</td>
<td>E</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>–</td>
<td>0</td>
<td>F</td>
<td>0D</td>
<td>1</td>
</tr>
</tbody>
</table>
Simple Memory System: TLB

- 16 entries total
- 4-way set associative

Why does the TLB ignore the page offset?

Simple Memory System: Cache

- Direct-mapped with $K = 4$ B, $C/K = 16$
- Physically addressed

Note: It is just coincidence that the PPN is the same width as the cache Tag
Current State of Memory System

TLB:

<table>
<thead>
<tr>
<th>Set</th>
<th>Tag</th>
<th>V</th>
<th>PPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>03</td>
<td>-</td>
<td>01</td>
</tr>
<tr>
<td>1</td>
<td>03</td>
<td>2D</td>
<td>02</td>
</tr>
<tr>
<td>2</td>
<td>02</td>
<td>-</td>
<td>08</td>
</tr>
<tr>
<td>3</td>
<td>07</td>
<td>-</td>
<td>03</td>
</tr>
</tbody>
</table>

Page table (partial):

<table>
<thead>
<tr>
<th>VPN</th>
<th>PPN</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>02</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

Cache:

<table>
<thead>
<tr>
<th>Index</th>
<th>Tag</th>
<th>V</th>
<th>B0</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>19</td>
<td>1</td>
<td>99</td>
<td>11</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>1</td>
<td>00</td>
<td>02</td>
<td>04</td>
<td>08</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>1</td>
<td>43</td>
<td>6D</td>
<td>09</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0D</td>
<td>1</td>
<td>36</td>
<td>72</td>
<td>F0</td>
<td>1D</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>16</td>
<td>1</td>
<td>11</td>
<td>C2</td>
<td>DF</td>
<td>03</td>
</tr>
</tbody>
</table>

Memory Request Example #1

- **Virtual Address:** 0x03D4

- **Physical Address:**

Note: It is just coincidence that the PPN is the same width as the cache Tag.
Memory Request Example #2

- Virtual Address: 0x038F

![Diagram of TLB and PPN with VPN, TLBT, TLBI, VPO, VPN, TLB Hit, Page Fault, PPN]

- Physical Address:

![Diagram of CT, CI, CO, PPN, PPO, CT, CI, CO, Cache Hit, Data (byte)]

Note: It is just coincidence that the PPN is the same width as the cache Tag

Memory Request Example #3

- Virtual Address: 0x0020

![Diagram of TLB and PPN with VPN, TLBT, TLBI, VPO, VPN, TLB Hit, Page Fault, PPN]

- Physical Address:

![Diagram of CT, CI, CO, PPN, PPO, CT, CI, CO, Cache Hit, Data (byte)]

Note: It is just coincidence that the PPN is the same width as the cache Tag
Memory Request Example #4

- **Virtual Address:** $0x036B$

![TLB Diagram]

- **Physical Address:**

![Cache Diagram]

**Note:** It is just coincidence that the PPN is the same width as the cache Tag

Virtual Memory Summary

- **Programmer’s view of virtual memory**
  - Each process has its own private linear address space
  - Cannot be corrupted by other processes

- **System view of virtual memory**
  - Uses memory efficiently by caching virtual memory pages
    - Efficient only because of locality
  - Simplifies memory management and sharing
  - Simplifies protection by providing permissions checking
Memory System Summary

- Memory Caches (L1/L2/L3)
  - Purely a speed-up technique
  - Behavior invisible to application programmer and (mostly) OS
  - Implemented totally in hardware

- Virtual Memory
  - Supports many OS-related functions
    - Process creation, task switching, protection
  - Operating System (software)
    - Allocates/shares physical memory among processes
    - Maintains high-level tables tracking memory type, source, sharing
    - Handles exceptions, fills in hardware-defined mapping tables
  - Hardware
    - Translates virtual addresses via mapping tables, enforcing permissions
    - Accelerates mapping via translation cache (TLB)

Memory System – Who controls what?

- Memory Caches (L1/L2/L3)
  - Controlled by hardware
  - Programmer cannot control it
  - Programmer can write code to take advantage of it

- Virtual Memory
  - Controlled by OS and hardware
  - Programmer cannot control mapping to physical memory
  - Programmer can control sharing and some protection
    - via OS functions (not in CSE 351)