Virtual Memory: Concepts

15-213 / 18-213: Introduction to Computer Systems
16\textsuperscript{th} Lecture, Mar. 19, 2013

Instructors:
Anthony Rowe, Seth Goldstein, and Gregory Kesden
Today

- VM Motivation and Address spaces
- VM as a tool for caching
- VM as a tool for memory management
- VM as a tool for memory protection
- Address translation
Virtual Memory Abstraction

- Programs refer to virtual memory addresses
  - movl (%ecx),%eax
  - Conceptually very large array of bytes
  - Each byte has its own address
  - Actually implemented with hierarchy of different memory types
  - System provides address space private to particular “process”

- Allocation: Compiler and run-time system
  - Where different program objects should be stored
  - All allocation within single virtual address space

- But why virtual memory?
- Why not physical memory?
Problem 1: How Does Everything Fit?

64-bit addresses: 16 Exabyte

Physical main memory: Few Gigabytes

And there are many processes ....
Problem 2: Memory Management

Physical main memory

Process 1
Process 2
Process 3
...
Process n

stack
heap
.text
.data
...

What goes where?
Problem 3: How To Protect

Physical main memory

Process i

Process j

Problem 4: How To Share?

Physical main memory

Process i

Process j
Solution: Level Of Indirection

- Each process gets its own private memory space
- Solves the previous problems

Diagram:
- Virtual memory
- Physical memory
- Mapping

Process 1

Process n
One simple trick solves all of these problems

- Each process gets its own private image of memory
  - appears to be a full-sized private memory range
- This fixes “how to choose” and “others shouldn’t mess w/ yours”
  - in addition to “making everything fit”
- Implementation: translate addresses transparently
  - add a mapping function
    - to map private (i.e. “virtual”) addresses to physical addresses
    - do the mapping on every load or store
- This mapping trick is the heart of virtual memory
Address Spaces

- **Linear address space**: Ordered set of contiguous non-negative integer addresses:
  \{0, 1, 2, 3 \ldots \}

- **Virtual address space**: Set of $N = 2^n$ virtual addresses
  \{0, 1, 2, 3, \ldots, N-1\}

- **Physical address space**: Set of $M = 2^m$ physical addresses
  \{0, 1, 2, 3, \ldots, M-1\}

- Clean distinction between data (bytes) and their attributes (addresses)
- Each datum can now have multiple addresses
- Every byte in main memory:
  one physical address, one (or more) virtual addresses
A System Using Physical Addressing

- Used in “simple” systems like embedded microcontrollers in devices like cars, elevators, and digital picture frames
A System Using Virtual Addressing

- Used in all modern servers, desktops, and laptops
- One of the great ideas in computer science
Why Virtual Memory (summary)?

- **Uses main memory (RAM) efficiently**
  - Use DRAM as a cache for the parts of a virtual address space

- **Simplifies memory management**
  - Each process gets the same uniform linear address space

- **Isolates address spaces**
  - One process can’t interfere with another’s memory
  - User program cannot access privileged kernel information
Today

- VM Motivation and Address spaces
- (1) VM as a tool for caching
- (2) VM as a tool for memory management
- (3) VM as a tool for memory protection
- Address translation
(1) VM as a Tool for Caching

- **Virtual memory** is an array of N contiguous bytes stored on disk.
- The contents of the array on disk are cached in **physical memory (DRAM cache)**
  - These cache blocks are called **pages** (size is $P = 2^p$ bytes)

![Diagram of Virtual and Physical Memory](image.png)
Enabling data structure: Page Table

- A **page table** is an array of page table entries (PTEs) that maps virtual pages to physical pages.
  - Per-process kernel data structure in DRAM
**Page Hit**

- **Page hit:** reference to VM word that is in physical memory (DRAM cache hit)

<table>
<thead>
<tr>
<th>Virtual address</th>
<th>Physical page number or disk address</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTE 0</td>
<td>Valid:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>PTE 7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Physical memory (DRAM): VP 1, VP 2, VP 7, VP 4

Virtual memory (disk): VP 1, VP 2, VP 3, VP 4, VP 6, VP 7

Memory resident page table (DRAM): PP 0, PP 3

Diagram:
- Virtual address
- Physical page number or disk address
- Physical memory (DRAM)
- Virtual memory (disk)
- Memory resident page table (DRAM)
Page Fault

- **Page fault**: reference to VM word that is not in physical memory (DRAM cache miss)
Handling Page Fault

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

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

- Page miss causes page fault (an exception)
- Page fault handler selects a victim to be evicted (here VP 4)
Handling 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!
Locality to the Rescue Again!

- Virtual memory works because of locality

- At any point in time, programs tend to access a set of active virtual pages called the *working set*
  - Programs with better temporal locality will have smaller working sets

- If (working set size < main memory size)
  - Good performance for one process after compulsory misses

- If (SUM(working set sizes) > main memory size)
  - *Thrashing*: Performance meltdown where pages are moved (copied) in and out continuously
Today

- VM Motivation and Address spaces
- (1) VM as a tool for caching
- (2) VM as a tool for memory management
- (3) VM as a tool for memory protection
- Address translation
(2) VM as a Tool for Memory Management

- Key idea: each process has its own virtual address space
  - It can view memory as a simple linear array
  - Mapping function scatters addresses through physical memory
    - Well chosen mappings simplify memory allocation and management
Simplifying allocation and sharing

- **Memory allocation**
  - Each virtual page can be mapped to any physical page
  - A virtual page can be stored in different physical pages at different times

- **Sharing code and data among processes**
  - Map multiple virtual pages to the same physical page (here: PP 6)

![Virtual Address Space for Process 1:

Virtual Address Space for Process 2:

Address translation

Physical Address Space (DRAM)

(e.g., read-only library code)
Simplifying Linking and Loading

**Linking**
- Each program has similar virtual address space
- Code, stack, and shared libraries always start at the same address

**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

<table>
<thead>
<tr>
<th>Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x08048000</td>
<td>Unused</td>
</tr>
<tr>
<td>0x40000000</td>
<td>Memory-mapped region for shared libraries</td>
</tr>
<tr>
<td>0x04000000</td>
<td>Run-time heap (created by malloc)</td>
</tr>
<tr>
<td>0x00000000</td>
<td>User stack (created at runtime)</td>
</tr>
<tr>
<td>Memory invisible to user code</td>
<td>%esp (stack pointer)</td>
</tr>
<tr>
<td></td>
<td>brk</td>
</tr>
<tr>
<td></td>
<td>Loaded from the executable file</td>
</tr>
</tbody>
</table>

---

Kernel virtual memory
Today

- VM Motivation and Address spaces
- (1) VM as a tool for caching
- (2) VM as a tool for memory management
- (3) VM as a tool for memory protection
- Address translation
VM as a Tool for Memory Protection

- Extend PTEs with permission bits
- Page fault handler checks these before remapping
  - If violated, send process SIGSEGV (segmentation fault)
Today

- VM Motivation and Address spaces
- (1) VM as a tool for caching
- (2) VM as a tool for memory management
- (3) VM as a tool for memory protection
- Address translation
VM Address Translation

- **Virtual Address Space**
  - $V = \{0, 1, \ldots, N-1\}$

- **Physical Address Space**
  - $P = \{0, 1, \ldots, M-1\}$

- **Address Translation**
  - $MAP: V \rightarrow P \cup \{\emptyset\}$
  - For virtual address $a$:
    - $MAP(a) = a'$ if data at virtual address $a$ is at physical address $a'$ in $P$
    - $MAP(a) = \emptyset$ if data at virtual address $a$ is not in physical memory
      - Either invalid or stored on disk
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
  - CO: Byte offset within cache line
  - CI: Cache index
  - CT: Cache tag
Address Translation With a Page Table

Virtual address

Virtual page number (VPN)  Virtual page offset (VPO)

Physical address

Physical page number (PPN)  Physical page offset (PPO)

Valid bit = 0: page not in memory (page fault)

Page table address for process

Page table base register (PTBR)

Valid

Physical page number (PPN)

n-1

p  p-1

m-1

0

0  p-1  p 0
Address Translation: Page Hit

1) Processor sends virtual address to MMU
2-3) MMU fetches PTE from page table in memory
4) MMU sends physical address to cache/memory
5) Cache/memory sends data word to processor
Address Translation: Page Fault

1) Processor sends virtual address to MMU
2-3) MMU fetches PTE from page table in 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
Views of virtual memory

- **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 programming
  - Simplifies protection by providing a convenient interpositioning point to check permissions
Integrating VM and Cache

VA: virtual address, PA: physical address, PTE: page table entry, PTEA = PTE address
Speeding up Translation with a TLB

- Page table entries (PTEs) are cached in L1 like any other memory word
  - PTEs may be evicted by other data references
  - PTE hit still requires a small L1 delay

- Solution: *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
TLB Hit

A TLB hit eliminates a memory access
TLB Miss

A TLB miss incurs an additional memory access (the PTE)
Fortunately, TLB misses are rare. Why?
Conclusions

(1) VM allows efficient use of limited main memory (RAM)
   - Use RAM as a cache for the parts of a virtual address space
     - some non-cached parts stored on disk
     - some (unallocated) non-cached parts stored nowhere
   - Keep only active areas of virtual address space in memory
     - transfer data back and forth as needed

(2) VM simplifies memory management for programmers
   - Each process gets a full, private linear address space

(3) VM isolates address spaces
   - One process can’t interfere with another’s memory
     - because they operate in different address spaces
   - User process cannot access privileged information
     - different sections of address spaces have different permissions