CS510 Operating System Foundations

Jonathan Walpole
Virtual Memory (2)
Inverted Page Tables

Problem:
- Page table overhead increases with address space size
- Page tables get too big to fit in memory!

Consider a computer with 64 bit addresses
- Assume 4 Kbyte pages (12 bits for the offset)
- Virtual address space = \(2^{52}\) pages!
- Page table needs \(2^{52}\) entries!
- This page table is much too large for memory!
Inverted Page Tables

How many mappings do we need (maximum) at any time?
Inverted Page Tables

How many mappings do we need (maximum) at any time?

We only need mappings for pages that are in memory!
Inverted Page Tables

An inverted page table
- Has one entry for every resident memory page
- Roughly speaking, one for each frame of memory
- Records which page is in that frame
- Can not be indexed by page number

So how can we search an inverted page table on a TLB miss fault?
Inverted Page Tables

Given a page number (from a faulting address), do we exhaustively search all entries to find its mapping?
Inverted Page Tables

Given a page number (from a faulting address), do we exhaustively search all entries to find its mapping?

- No, that’s too slow!
- A hash table could allow fast access given a page number
- O(1) lookup time with a good hash function
Hash Tables

Data structure for associating a key with a value
- Apply hash function to key to produce a hash
- Hash is a number that is used as an array index
- Each element of the array can be a linked list of entries (to handle collisions)
- The list must be searched to find the required entry for the key (entry’s key matches the search key)
- With a good hash function the average list length will be short
Traditional page table with an entry for each of the $2^{52}$ pages

Indexed by virtual page

$2^{52} - 1$

256-MB physical memory has $2^{16}$ 4-KB page frames

Indexed by hash on virtual page

Hash table

Virtual page

Page frame
Which Page Table Design is Best?

The best choice depends on CPU architecture
64 bit systems need inverted page tables
Some systems use a combination of regular page tables together with segmentation (later)
Memory Protection

Protection though addressability
- If address translation only allows a process to access its own pages, it is implementing memory protection

But what if you want a process to be able to read and execute some pages but not write them?
- eg. the text segment

Or read and write them but not execute them?
- eg. the stack

Can we implement protection based on access type?
Memory Protection

How is protection checking implemented?
- Compare page protection bits with process capabilities and operation types on every load/store
- That sounds expensive!
- Requires hardware support!

How can protection checking be done efficiently?
- Use the TLB as a “protection” look-aside buffer as well as a translation lookaside buffer
- Use special segment registers
Protection Lookaside Buffer

A TLB is often used for more than just “translation”
Memory accesses need to be checked for validity
  - Does the address refer to an allocated segment of
    the address space?
      *If not:* segmentation fault!
  - Is this process allowed to access this memory
    segment?
      *If not:* segmentation/protection fault!
  - Is the type of access valid for this segment?
    Read, write, execute ...?
      *If not:* protection fault!
Protection Checking With a TLB

<table>
<thead>
<tr>
<th>Valid</th>
<th>Virtual page</th>
<th>Modified</th>
<th>Protection</th>
<th>Page frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>140</td>
<td>1</td>
<td>RW</td>
<td>31</td>
</tr>
<tr>
<td>1</td>
<td>20</td>
<td>0</td>
<td>R X</td>
<td>38</td>
</tr>
<tr>
<td>1</td>
<td>130</td>
<td>1</td>
<td>RW</td>
<td>29</td>
</tr>
<tr>
<td>1</td>
<td>129</td>
<td>1</td>
<td>RW</td>
<td>62</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>0</td>
<td>R X</td>
<td>50</td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>0</td>
<td>R X</td>
<td>45</td>
</tr>
<tr>
<td>1</td>
<td>860</td>
<td>1</td>
<td>RW</td>
<td>14</td>
</tr>
<tr>
<td>1</td>
<td>861</td>
<td>1</td>
<td>RW</td>
<td>75</td>
</tr>
</tbody>
</table>
Page Grain Protection

A typical page table entry with support for page grain protection
Memory Protection Granularity

At what granularity should protection be implemented?

Page-level?
- A lot of overhead for storing protection information for non-resident pages

Segment level?
- Coarser grain than pages
- Makes sense if contiguous groups of pages share the same protection status
Segment-Granularity Protection

All pages within a segment usually share the same protection status
  - So we should be able to batch the protection information

Then why not just use segment-size pages?
Segment-Granularity Protection

Segments vary in size from process to process
Segments change size dynamically (stack, heap)

Need to manage addressability, access-based protection and memory allocation separately

Coarse-grain protection can be implemented simply through addressability
Segmented Address Spaces

*Traditional Virtual Address Space*
- “flat” address space (1 dimensional)

*Segmented Address Space*
- Program made of several pieces or “segments”
- Each segment is its own mini-address space
- Addresses within a segment start at zero
- The program must always say which segment it means
  - either embed a segment id in an address
  - or load a value into a segment register and refer to the register
- Addresses:
  - Segment + Offset
- Each segment can grow independently of others
Segmented Address Spaces

**Example:** A compiler
Segmented Memory

Segment 0
Symbol table

Segment 1
Source text

Segment 2
Constants

Segment 3
Parse tree

Segment 4
Call stack
Instruction and Data Spaces

* One address space
* Separate I and D spaces
# Pure Paging vs. Pure Segmentation

<table>
<thead>
<tr>
<th>Consideration</th>
<th>Paging</th>
<th>Segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need the programmer be aware that this technique is being used?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>How many linear address spaces are there?</td>
<td>1</td>
<td>Many</td>
</tr>
<tr>
<td>Can the total address space exceed the size of physical memory?</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Can procedures and data be distinguished and separately protected?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Can tables whose size fluctuates be accommodated easily?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Is sharing of procedures between users facilitated?</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Why was this technique invented?</td>
<td>To get a large linear address space without having to buy more physical memory</td>
<td>To allow programs and data to be broken up into logically independent address spaces and to aid sharing and protection</td>
</tr>
</tbody>
</table>
Segmentation vs. Paging

Do we need to choose one or the other?
Why not use both together
- Paged segments
- Paging for memory allocation
- Segmentation for maintaining protection information at a coarse granularity
- Segmentation and paging in combination for translation
Paged Segments in MULTICS

Each segment is divided up into pages.
Each segment descriptor points to a page table.
# Paged Segments in MULTICS

Each entry in segment table

<p>| | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main memory address of the page table</td>
<td>18</td>
</tr>
<tr>
<td>Segment length (in pages)</td>
<td>9</td>
</tr>
<tr>
<td>Page size:</td>
<td></td>
</tr>
<tr>
<td>0 = 1024 words</td>
<td></td>
</tr>
<tr>
<td>1 = 64 words</td>
<td></td>
</tr>
<tr>
<td>0 = segment is paged</td>
<td></td>
</tr>
<tr>
<td>1 = segment is not paged</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous bits</td>
<td></td>
</tr>
<tr>
<td>Protection bits</td>
<td></td>
</tr>
</tbody>
</table>
Paged Segments in MULTICS

Conversion of a 2-part MULTICS address into a main memory address
The MULTICS TLB

Simplified version of the MULTICS TLB
Existence of 2 page sizes makes actual TLB more complicated
Spare Slides
Pentium Segmentation & Paging

Conversion of a (selector, offset) pair to a linear address
Pentium Segmentation & Paging

Pentium segment descriptor

- Base 24-31
- G
- D
- 0
- Limit 16-19
- P
- DPL
- S
- Type
- Base 16-23
- 0
- 4

- Base 0-15
- Limit 0-15

- 32 Bits

- Relative address

- 0: 16-Bit segment
- 1: 32-Bit segment

- 0: Li is in bytes
- 1: Li is in pages

- 0: Segment is absent from memory
- 1: Segment is present in memory

- Privilege level (0-3)
  - 0: System
  - 1: Application

- Segment type and protection