Gambit Paged Memory Management Unit

# Features

* 512 entry, 8-way associative TLB
* variable page table depth, subset address translation
* address short-cutting for larger page sizes (16MB)
* 16kB page size
* 52-bit address translations
* 52-bit bus mastering for table lookup
* automatic periodic hardware clearing of page accessed bit

# Operation

The PMMU translates virtual to physical addresses by looking up the translation in a TLB (translation lookaside buffer). Address translation takes two clock cycles to perform. During the first clock cycle it is determined if the access is valid. If the access is a valid access then the input address is transferred to output on the second clock cycle, otherwise an error signal is set active.

The PMMU will automatically walk the page tables on a TLB miss and update the TLB with an address translation. If the translation is not available a page fault is generated. Walking the page tables depends on the number of levels selected by the PTA (page table address) register. More levels to traverse take more memory cycles. Address translation may be shortcut at 16MB pages if indicated in the page table entry.

The PMMU will also periodically reset the accessed bits of page table entries during idle access times, to support the clock algorithm. The update is triggered by an externally provided signal. This signal should be connected to an interval timer (PIT) circuit.

Pages are 16kB in size. Each page contains 1024, 208-bit page table entries or 2048, 104-bit page directory entries.

Since there are only 64 entries at the 4th level of page tables, the lower five bits of the address space identifier may also be used to locate the table. This allow the entire 16kB page to be useful.

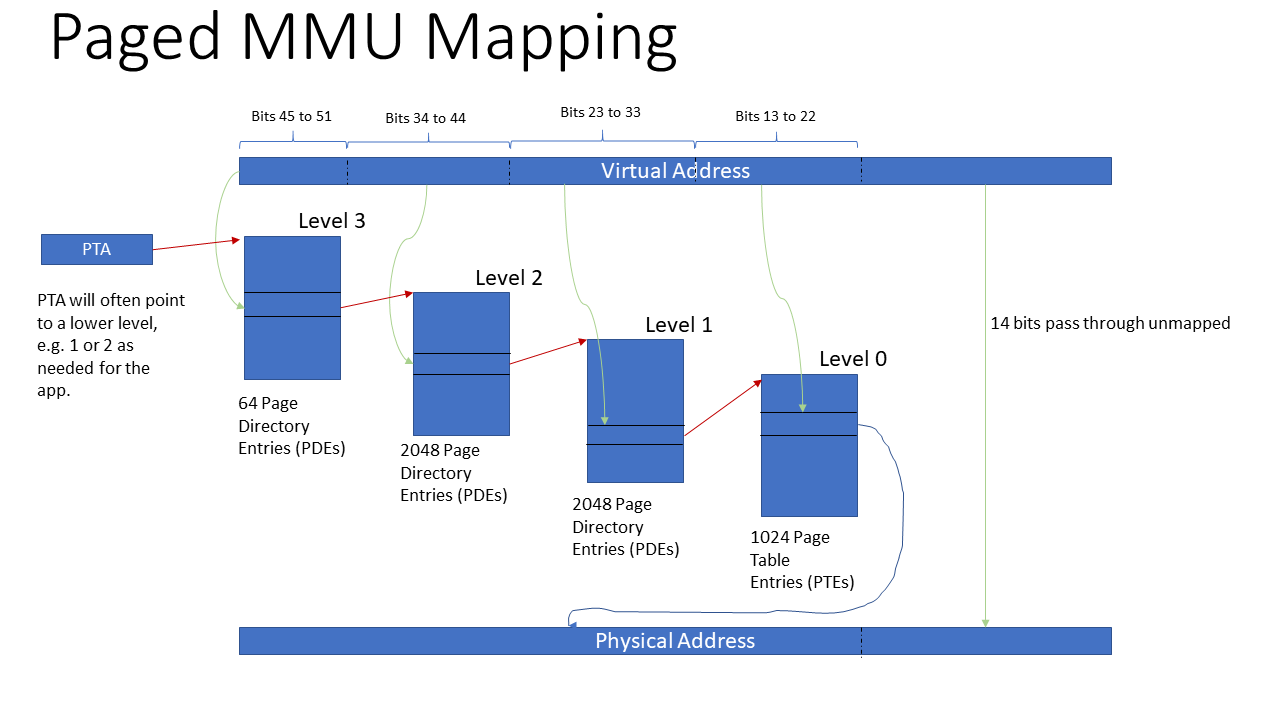
# Protection Mechanisms

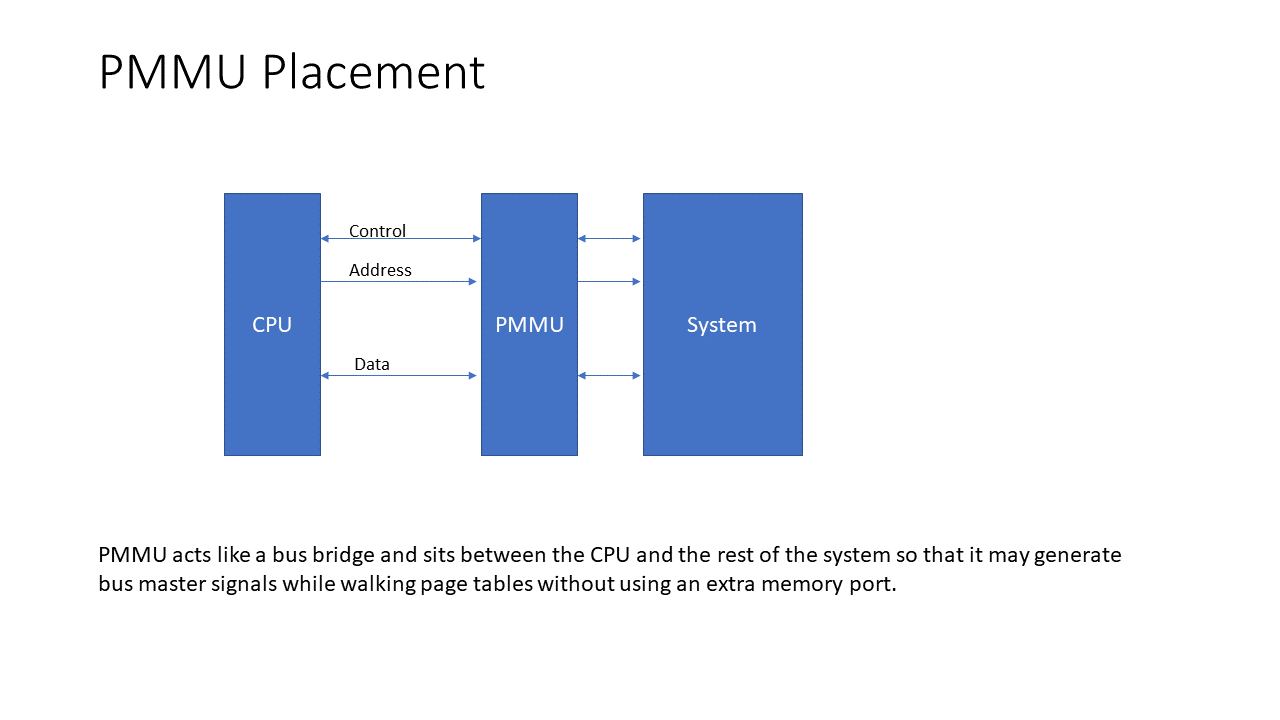
A protection key of all zeros will allow any process to access the page. Otherwise the process must contain a key matching the key specified for the page or a privilege violation exception will occur.

Read, write, or execute accessibility is set independently for each processor operating level.

For executable pages the privilege level of the page must match the processor’s current privilege level or a privilege violation will occur. For data pages the privilege level must be lower or equal to (numerically greater than or equal) to the processor’s privilege level.

Read, write or execute violations will occur if the access type is not appropriate for the current operating level.





## Page Table Entry

The following layout shows the page table entry structure as stored in memory.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 51 39 | 38 26 | 25 | 24 | 23 20 | 19 16 | 15 12 | | 11 8 | | 7 4 | 2 1 0 | | |
| Physical Page Number38 | | | | | | | DUA3 | | ~8 | | | T2 | P |
| Virtual Page Number38 | | | | | | |  | |  | | |  |  |
| Share Count13 | Privilege Level13 | U | S | CRWX | CRWX | CRWX | | CRWX | | CRWX | CRWX | | |
| Reference Counter26 | | ~6 | | | Protection Key20 | | | | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Word | Bit |  |  |  |  |
| 0 | 0 |  | P | Page present | 1 = page present in memory |
| 1 to 2 |  | T | Entry type | 1 = page directory entry, 2 = page table entry |
| 3 to 10 |  |  | reserved |  |
| 11 |  | A | 1 = accessed |  |
| 12 |  | U | undefined usage | available for use by OS |
| 13 |  | D | 1 = dirty | set if the page is written to |
| 14 to 51 |  | PPN | physical page number |  |
| 1 | 0 to 13 |  |  | reserved |  |
| 14 to 51 |  | VPN | virtual page number |  |
| 2 | 0 | User | Xu | 1 = executable |  |
| 1 | Wu | 1 = page writeable |
| 2 | Ru | 1 = readable |
| 3 | Cu | 1 = cachable | Ignored for executable pages which are always cached |
| 4 to 7 | Other | … | supervisor |  |
| 8 to 11 | … | supervisor |  |
| 12 to 15 | … | supervisor |  |
| 16 to 19 | … | hypervisor |  |
| 20 | Machine | Xm |  |  |
| 21 | Wm |  |  |
| 22 | Rm |  |  |
| 23 | Cm |  |  |
| 24 |  | S | 1 = shortcut translation | Translation shortcut bit eg (16MiB pages) |
| 25 |  | U | undefined usage | available for use by OS |
| 26 to 38 |  | PL | Privilege Level |  |
| 39 to 51 |  | SC | Share Count | number of times page is shared |
| 3 | 0 to 19 |  | PK | Protection Key | process must have a matching key in its collection for access |
| 20 to 25 |  |  | reserved |  |
| 26 to 51 |  | RC | Reference Counter |  |
|  |  |  |  |  |
|  |  |  |  |  |

## Page Directory Entry

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 51 14 | 13 11 | 10 3 | 2 1 | 0 |
| Page Number38 | DUA3 | ~8 | T2 | P |
| Virtual Page Number38 |  |  |  |  |

# Hardware Interface

|  |  |  |  |
| --- | --- | --- | --- |
| Port | Bits | Dir. | Description |
| rst\_i | 1 | I | resets the core |
| clk\_i | 1 | I | clock input |
| clock\_tick\_i | 1 | I | This input is meant to be driven by a timing source and is used for hardware clearing of the accessed bit in page table entries. |
| cyc\_o | 1 | O | bus master cycle valid strobe |
| stb\_o | 1 | O | bus master data strobe |
| lock\_o | 1 | O | not used |
| ack\_i | 1 | I | bus master transfer acknowledge |
| we\_o | 1 | O | bus master write transaction enable - may be masked off if page is read-only |
| sel\_o | 8 | O | bus master byte lane select |
| padr\_o | 52 | O | bus master address output (physical address) |
| dat\_i | 104 | I | bus master data input |
| dat\_o | 104 | O | bus master data output |
| invalidate | 1 | I |  |
| invalidate\_all | 1 | I | invalidates all entries in the TLB |
| pta | 52 | I | physical address of page table in memory. The page table should be 16kB aligned. Low order bits select table depth. Each level covers 2048 times the lower level memory range.   |  |  |  |  | | --- | --- | --- | --- | | PTA bits [9:8] | Address Space |  |  | | 0 | 16MB | root page table only |  | | 1 | 32GB | two level table |  | | 2 | 64TB | three table levels |  | | 3 | 128XB | four table level |  | |
| asid\_i | 8 | I | address space identifier |
| page\_fault | 1 | O | activated when a page is not present |
| pl\_i | 13 | I | privilege level input |
| ol\_i | 3 | I | operator level input (machine level (0) bypasses address mapping) |
| icl\_i | 1 | I | indicates an instruction cache load is taking place |
| cyc\_i | 1 | I | indicates a bus cycle is active |
| stb\_i | 1 | I | indicates data is strobed |
| ack\_o | 1 | O | acknowledge |
| we\_i | 1 | I | indicates a write cycle is taking place |
| sel\_i | 8 | I | byte lane selects |
| vadr\_i | 52 | I | virtual address to translate |
| cac\_o | 1 | O | page is not cachable |
| prv\_o | 1 | O | privilege violation |
| exv\_o | 1 | O | execute violation |
| rdv\_o | 1 | O | read violation |
| wrv\_o | 1 | O | write violation |
| keys\_i | 20\*8 | I | process keys |