Virtual Memory
- Implemented by hardware, MMU - memory management unit
- Often in OS core or as LRU
- Some issues with caching, but a miss ("page fault")
  has a penalty of millions of cycles potentially,
- Smart choice of fault by large block sizes
  "page" size - 1K-16K bytes
  Typical: 1K-4K bytes
- Willing to pay price for better replacement strategies (LRU)

Introduction to paging
- Physical memory too large to use associative lookup
  - Use page table, one page table per process
  - Use a page table register
  - Pointing to "current" page table

![Diagram of virtual memory addressing]
Pages are always "write back", only written to disk when removed from memory.

- Need "dirty" bit in page table to indicate page has been modified.

Page Replacement Algorithm:

- Common one is an approximation to LRU.
- Each page has "reference" bit indicating page has been accessed.
- All reference bits are cleared periodically.
- Replace page w/ a clear reference bit.

When are page tables kept?

- Size is Address Space/Page size, e.g. $2^{32}/2^{20} = \frac{2^{12}}{}$ entries.
- If too large, can be paged as well.
- Also, multi-level page tables, or page tables that grow.
But, you don't want to access a page table in memory cache for every memory access (instructions or data).

- Obvious idea: cache the parts of the page table currently in use.
- This cache is generally called the "translation lookaside buffer" (TLB).

TLB looks like:

- TLB is treated like other caches
  - Can be set associative, direct mapped, or fully associative
  - Need very simple & quick replacement alg.
  - Generally random
  - Block size can be several contiguous page table entries.
Relationship between MMU & cache hierarchy

Virtual address

\[ \text{V. Page} \rightarrow \text{Offset} \]

\[ \downarrow \]

\[ \text{TLB} \rightarrow \text{Page Table} \]

\[ \text{TLB hit} \rightarrow \text{Page table hit} \]

\[ \text{Physical Address} \rightarrow \text{A page} \rightarrow \text{Offset} \]

\[ \downarrow \]

\[ \text{Cache} \rightarrow \text{Memory} \]

\[ \text{Cache hit} \rightarrow \text{Cache miss} \rightarrow \text{Page table hit} \]

\[ \text{DATA} \rightarrow \text{Block to memory} \]

Worst case: TLB miss, page fault, cache miss

= 2 memory references + disk access

\[ \text{Yikes!} \]