Access unit: sector (512 bytes)
Max. request: $2^n$ sectors

Access unit: word (4~8 bytes)
Management unit: page (4KB)

Split large requests

Support sub-page cache

Access unit: block ($2^n$ sectors)
Max. request: $\infty$
Cache management

• Unit of caching?
  – Page cache: efficiency of memory management
  – Buffer cache: support for block unit I/Os

• Linux implementation
  – Earlier linux: buffer cache only
  – Recent linux: buffer cache over page cache
I/O subsystem architecture

- Read/write (device, addr, offset, nr sectors)
- Hard to manipulate
Page cache

- In-memory cache of disk blocks
  - Page unit
  - Retrieval: (inode, offset) \rightarrow \text{struct page *}

Efficient I/Os in a page unit
• Uses memory pages in storage I/O operations
  – No address translation: can use unmapped pages

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>block_device</td>
<td>bi_bdev</td>
</tr>
<tr>
<td>sector_t</td>
<td>bi_sector</td>
</tr>
<tr>
<td>struct bio *</td>
<td>next</td>
</tr>
<tr>
<td>long</td>
<td>bi_rw</td>
</tr>
<tr>
<td>int</td>
<td>bi_size</td>
</tr>
<tr>
<td>short</td>
<td>bi_vcnt</td>
</tr>
<tr>
<td>struct bio_vec*</td>
<td>bi_io_vec</td>
</tr>
<tr>
<td>bio_end_io_t</td>
<td>bi_end_io</td>
</tr>
<tr>
<td>void *</td>
<td>bi_private</td>
</tr>
</tbody>
</table>

**bi_idx:** For repetitive processing
From v3.14, struct bio changed
- For easier repetitive operations

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>block_device</td>
<td>bi_bdev</td>
</tr>
<tr>
<td>struct bio *</td>
<td>next</td>
</tr>
<tr>
<td>long</td>
<td>bi_rw</td>
</tr>
<tr>
<td>struct bvec_iter</td>
<td>bi_iter</td>
</tr>
<tr>
<td>short</td>
<td>bi_vcnt</td>
</tr>
<tr>
<td>struct bio_vec *</td>
<td>bi_io_vec</td>
</tr>
<tr>
<td>bio_end_io_t</td>
<td>bi_end_io</td>
</tr>
<tr>
<td>void *</td>
<td>bi_private</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>sector_t</td>
<td>bi_sector</td>
</tr>
<tr>
<td>int</td>
<td>bi_size</td>
</tr>
<tr>
<td>int</td>
<td>bi_idx</td>
</tr>
<tr>
<td>int</td>
<td>bi_bvec_done</td>
</tr>
</tbody>
</table>
• A block-unit cache: struct buffer_head
  – For sub-page I/O requests