

Page Cache

Adrian Huang | Jan, 2022

- * Based on kernel 5.11 (x86_64) – QEMU**
- * SMP (4 CPUs) and 8GB memory**
- * Kernel parameter: nokaslr norandmaps**
- * Userspace: ASLR is disabled**
- * EXT4 file system**
- * Legacy BIOS**

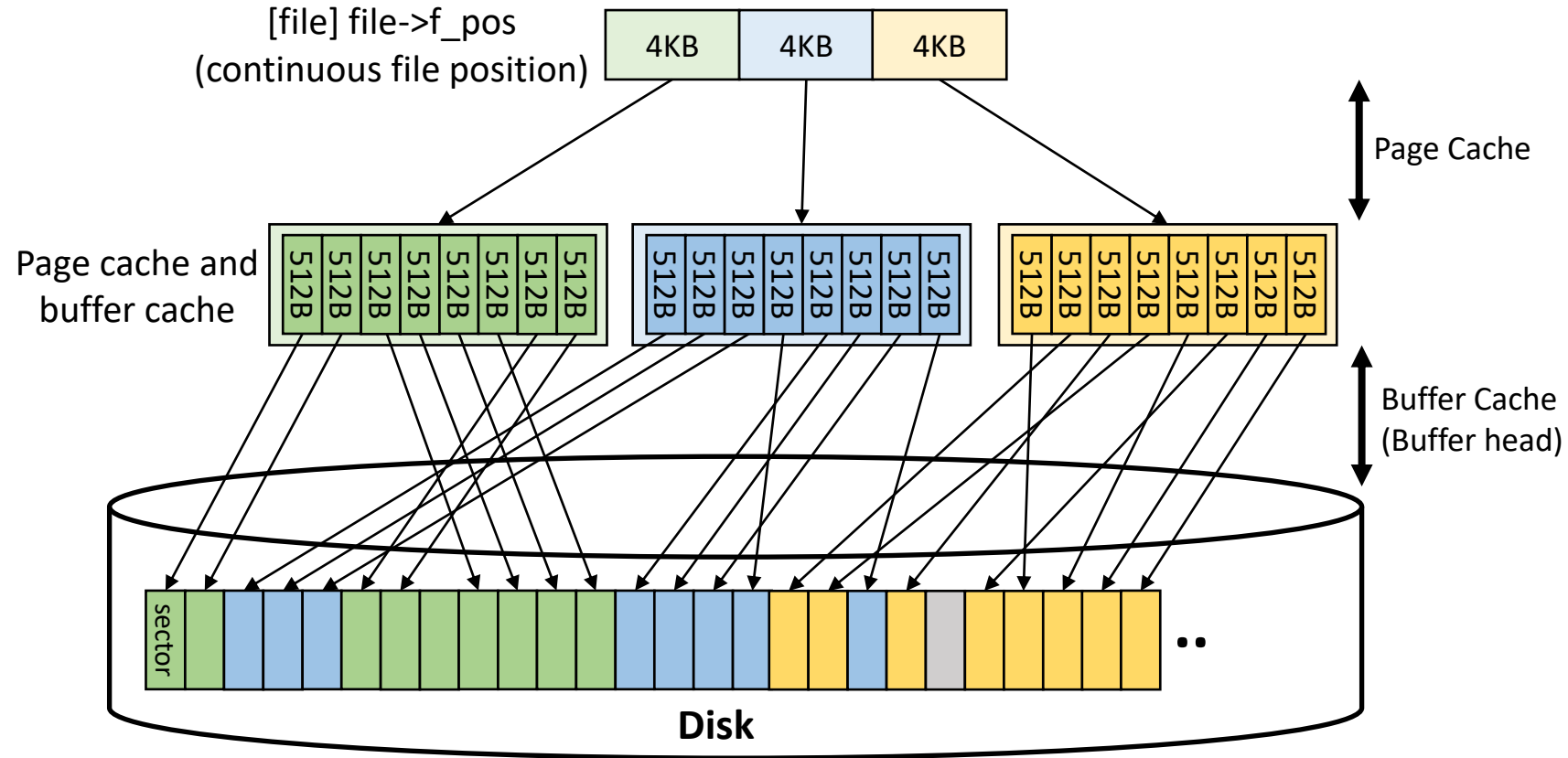
Agenda

- What is page cache?
- Page cache & buffer cache (struct buffer_head)
- How to find an existed page cache?
- Interaction with generic block layer: methods for bio construction
 1. Based on buffer_head
 2. [Without buffer_head] Based on page descriptor & file system
- File system block size & sector size
- [Detail Discussion] With or without buffer_head
 - File system-based IO
 - Block device-based IO

What is page cache?

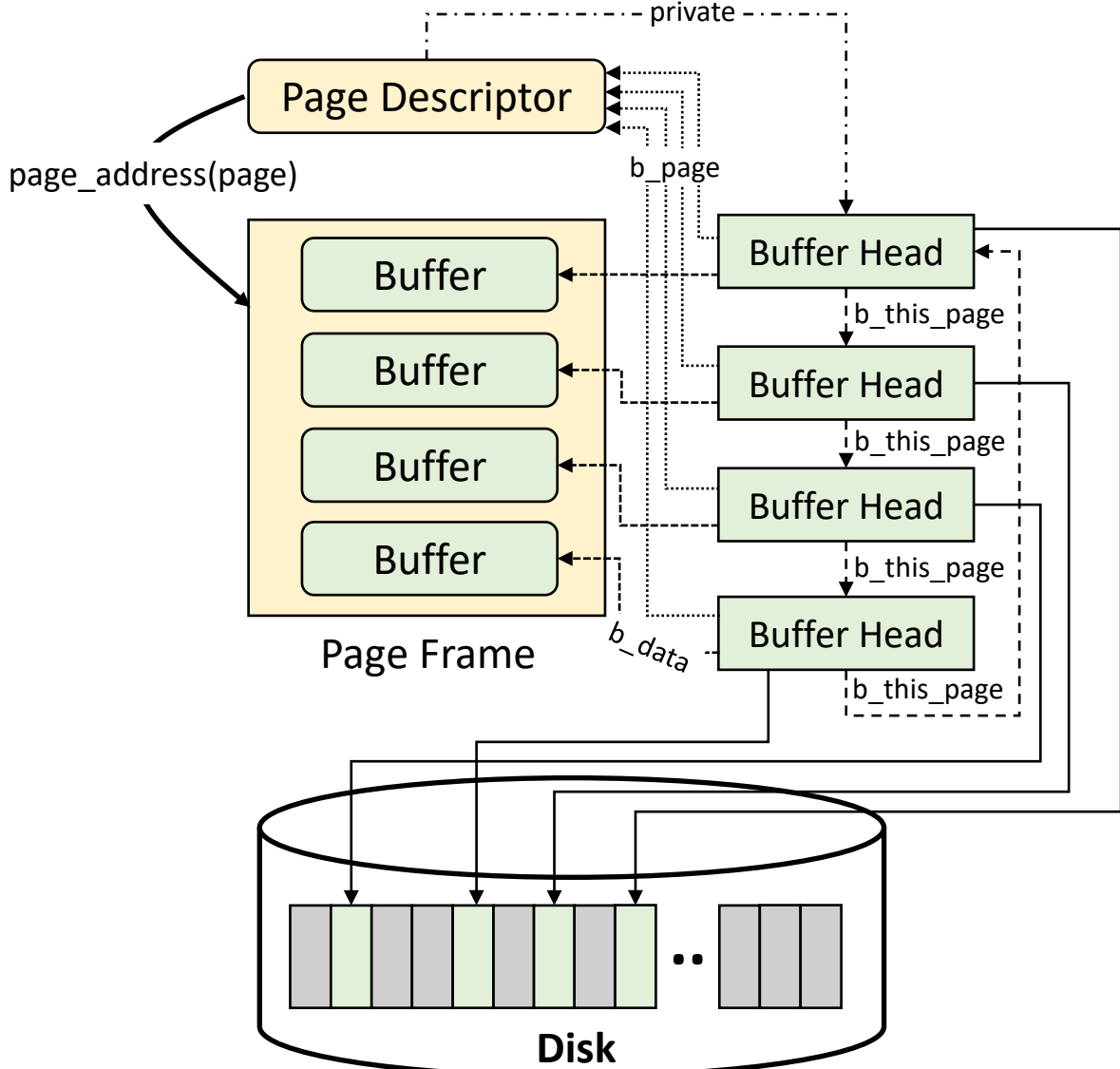
- page cache (stored in physical memory) = cache disk data
 - Speed up disk data access
- Linux kernel refers to the page cache for disk R/W
- If there is enough free memory,
 - the page cache is kept for an indefinite time
 - can be reused by other processes without accessing the disk
- Open a file with the O_DIRECT flag → Bypass page cache
 - Application: Some database applications use their own disk cache algorithm
 - Especially for large data access
 - Example: [Using direct I/O with Oracle](#)

Page cache & buffer cache



1. **Page cache:** Interaction with VFS. (Upper layer)
2. **Buffer cache:** Interaction with the disk. (Lower layer)

Page cache & buffer cache: relationship



[page->flags]

- PG_private: page cache (fs-private data)

buffer_head
b_state
b_this_page
b_page
b_blocknr
b_size = 1024
b_data
b_bdev

Legend

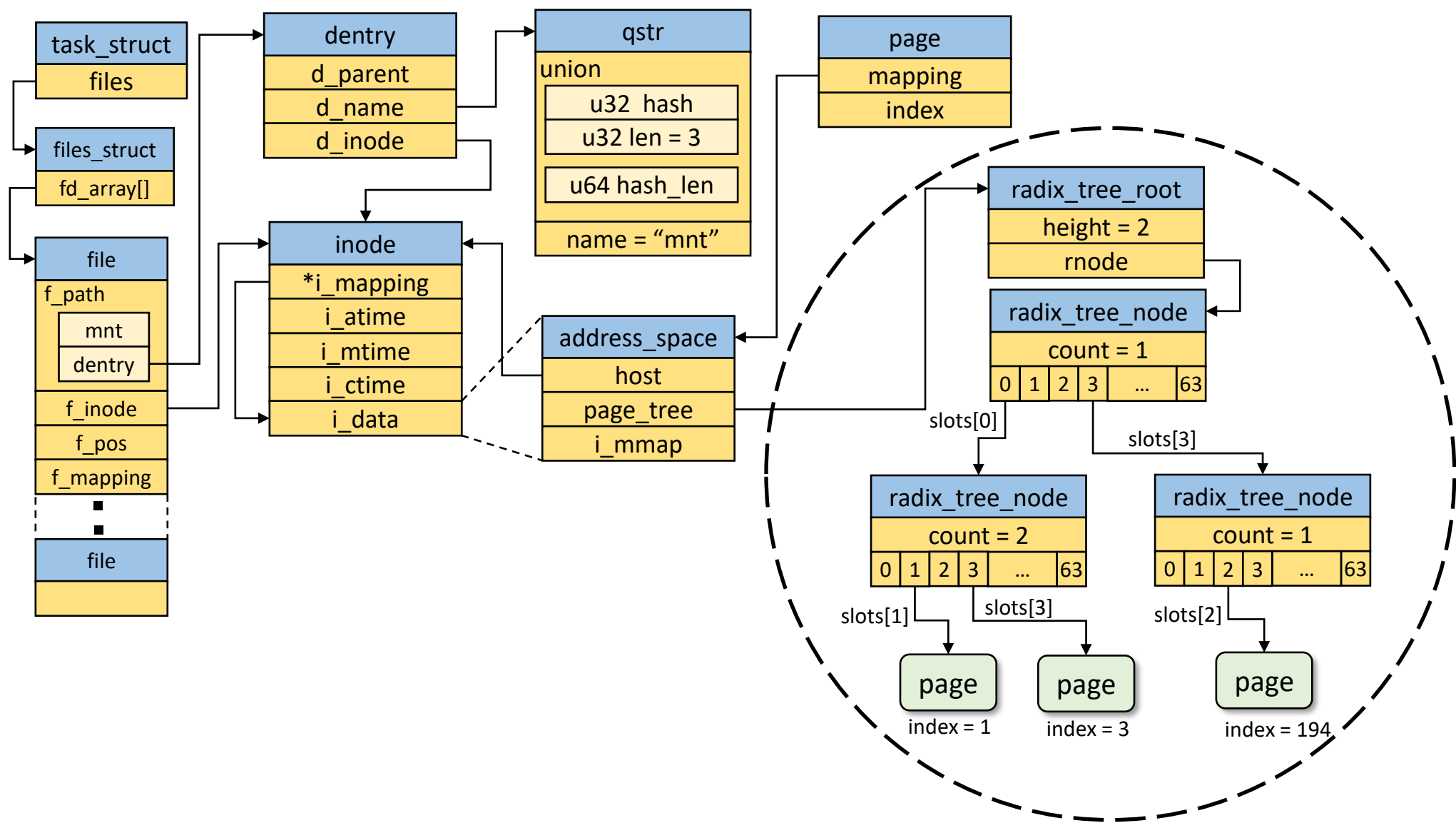
- ```

.....> b_page
-----> b_data
-----> b_this_page
-.....> private
—————> b_dev + b_blocknr:
 submit bh()->submit bio(

```

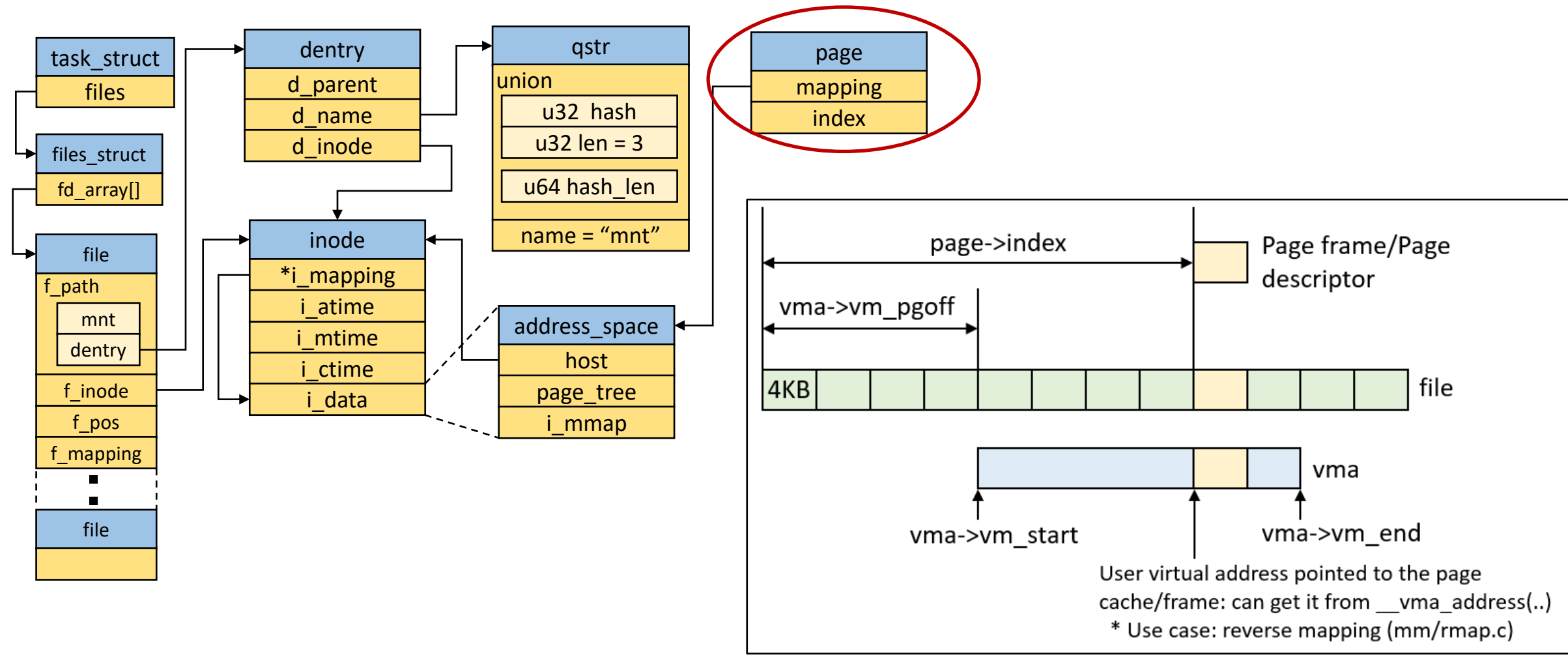
1. Block size = file system-based unit
2. Page cache might **\*NOT\*** include buffer\_head struct. (File system specific: file data)

# Radix Tree (or XArray): How to find an existed page cache?

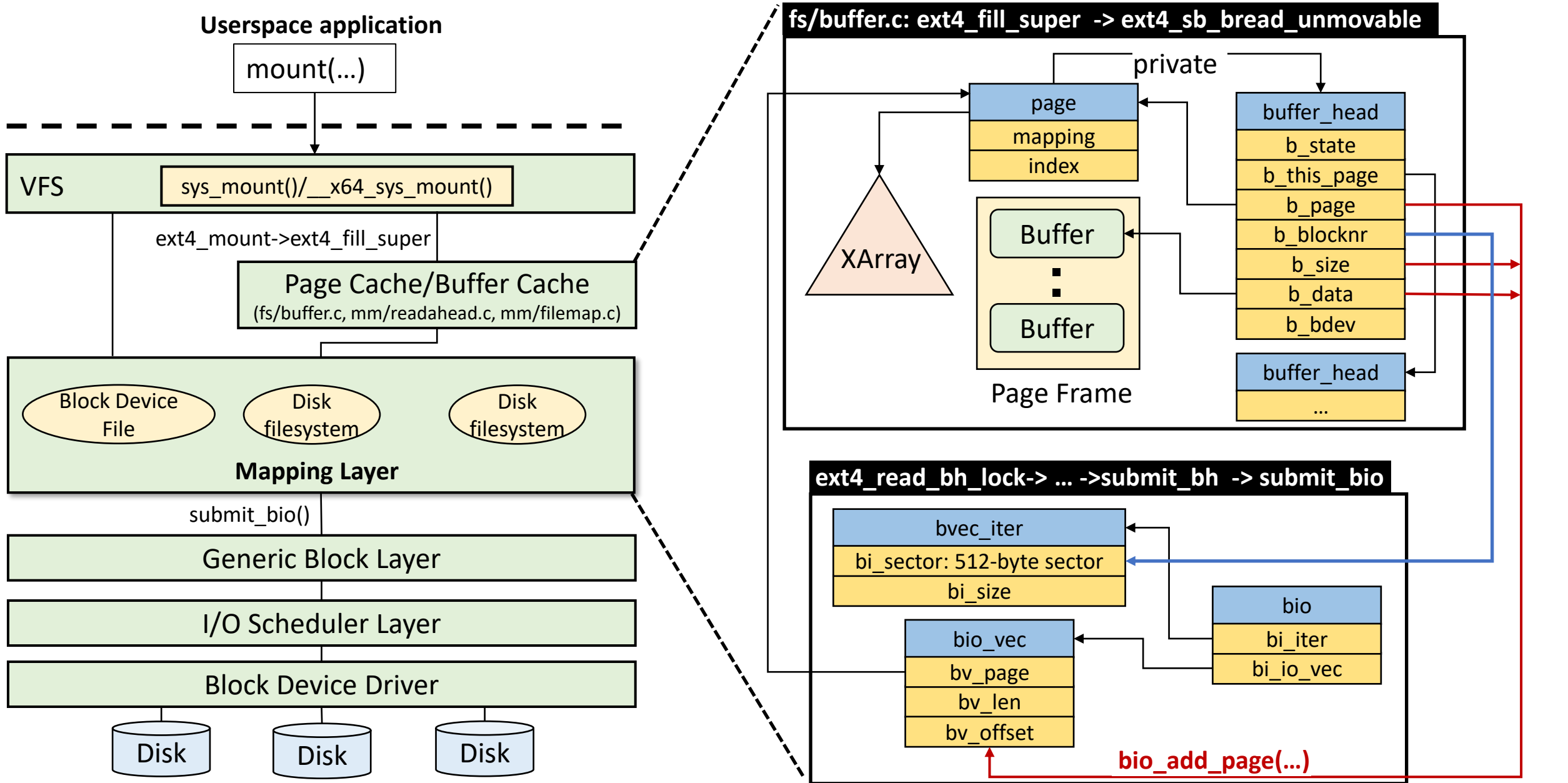


**[v4.20] XArray replaced radix tree**

# Radix Tree (or XArray): How to find an existed page cache?



# Interaction with generic block layer: bio construction based on buffer\_head



# Interaction with generic block layer: buffer\_head

- buffer\_head cache: per-cpu variable

```
#define BH_LRU_SIZE 16

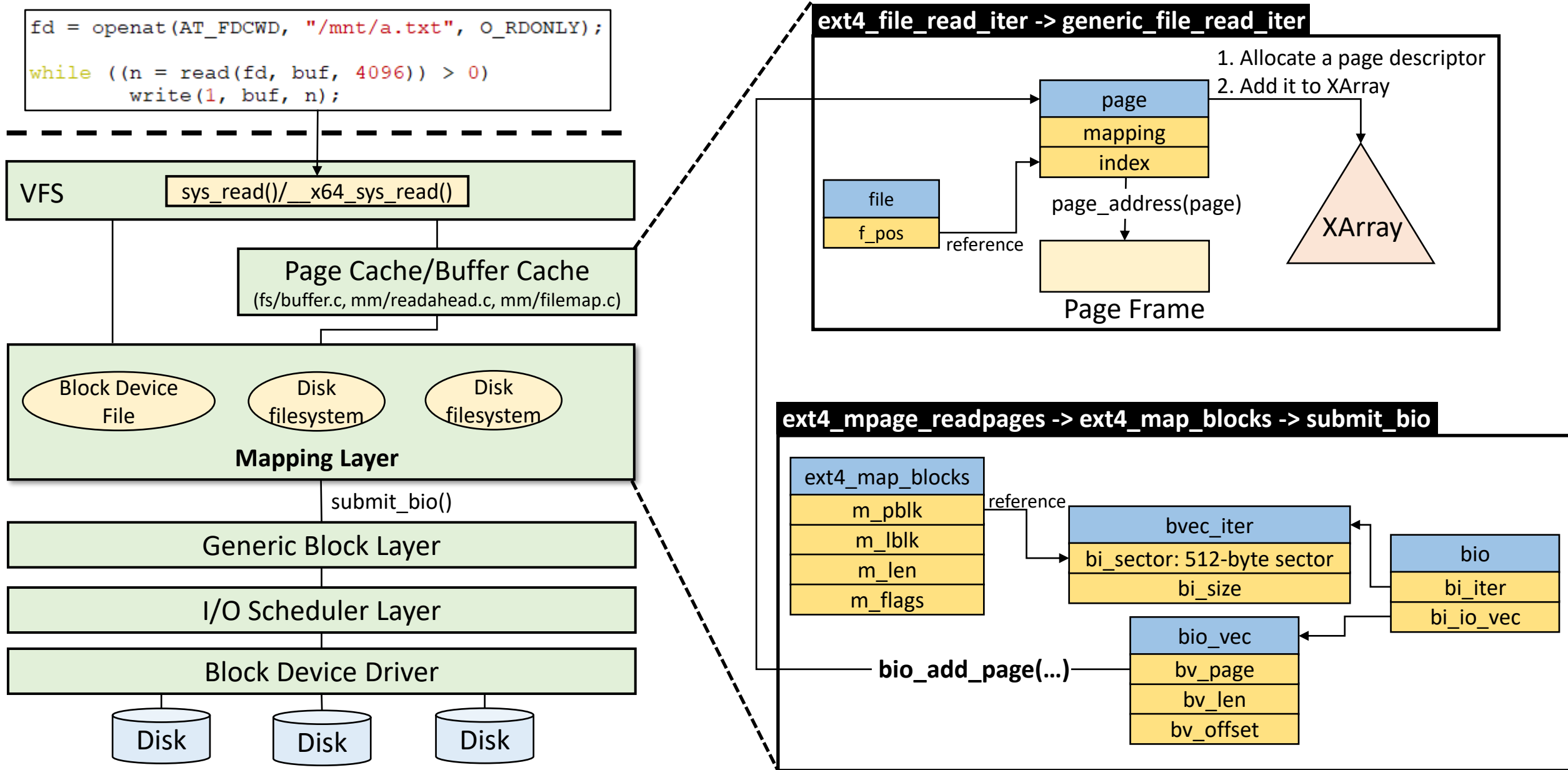
struct bh_lru {
 struct buffer_head *bhs[BH_LRU_SIZE];
};

static DEFINE_PER_CPU(struct bh_lru, bh_lrus) = {{ NULL }};

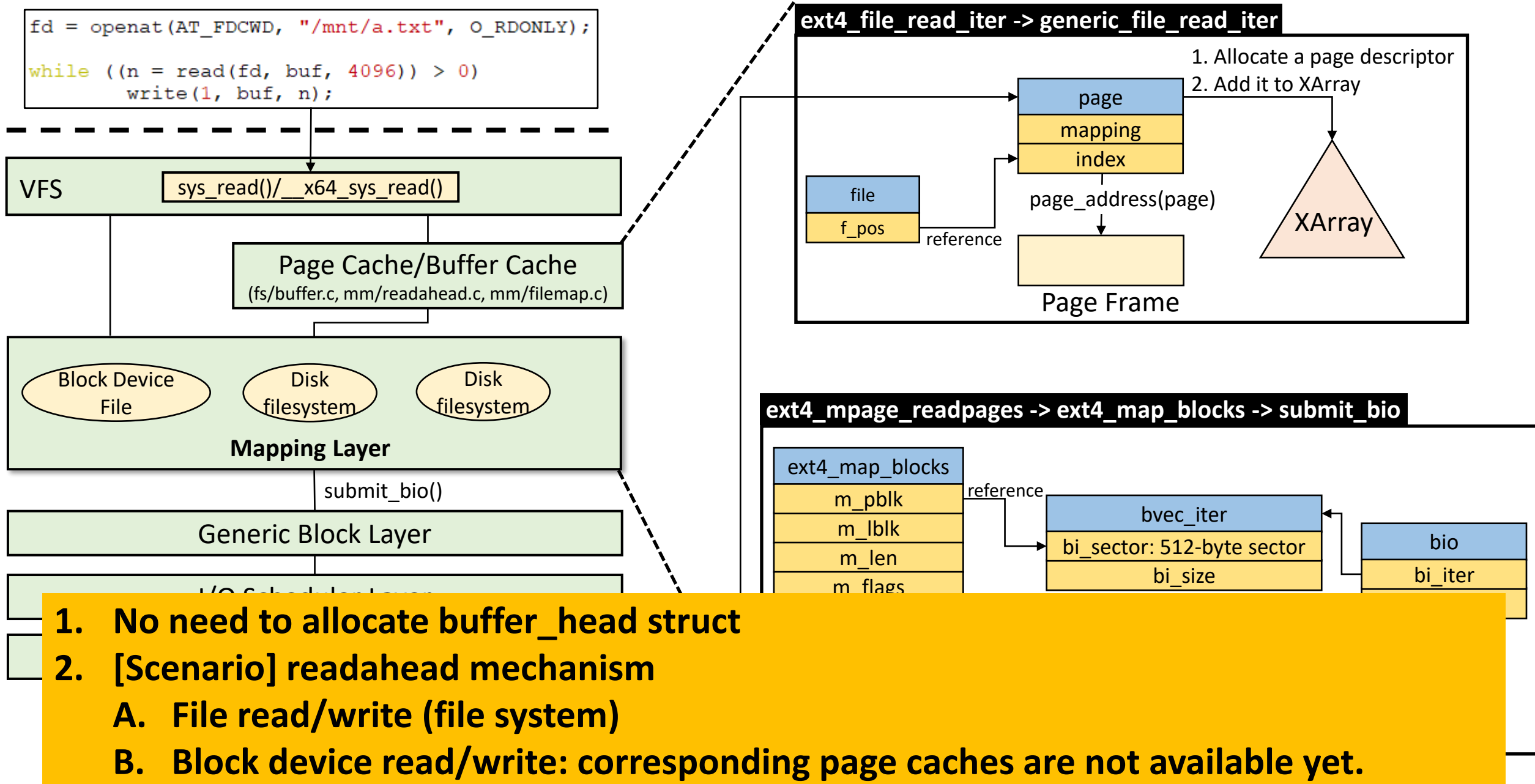
fs/buffer.c
```

- Scenarios
  - File system metadata (file)
    - superblock
    - inode info
    - extent tree
  - File hole
  - Block device: page cache is not up-to-date

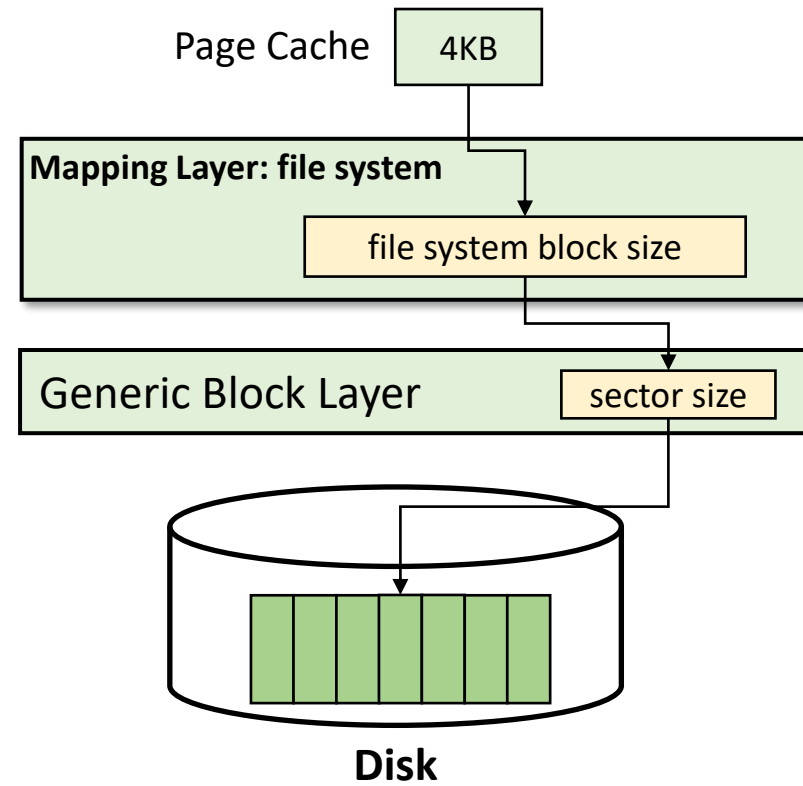
# Interaction with generic block layer: bio construction based on page & file system



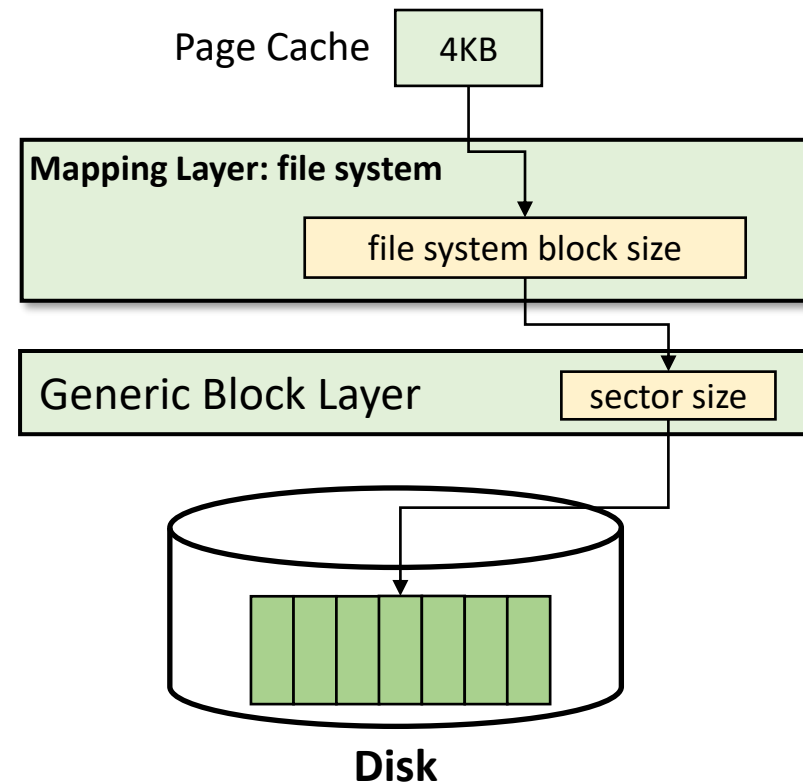
# Interaction with generic block layer: bio construction based on page & file system



# File system block size & sector size



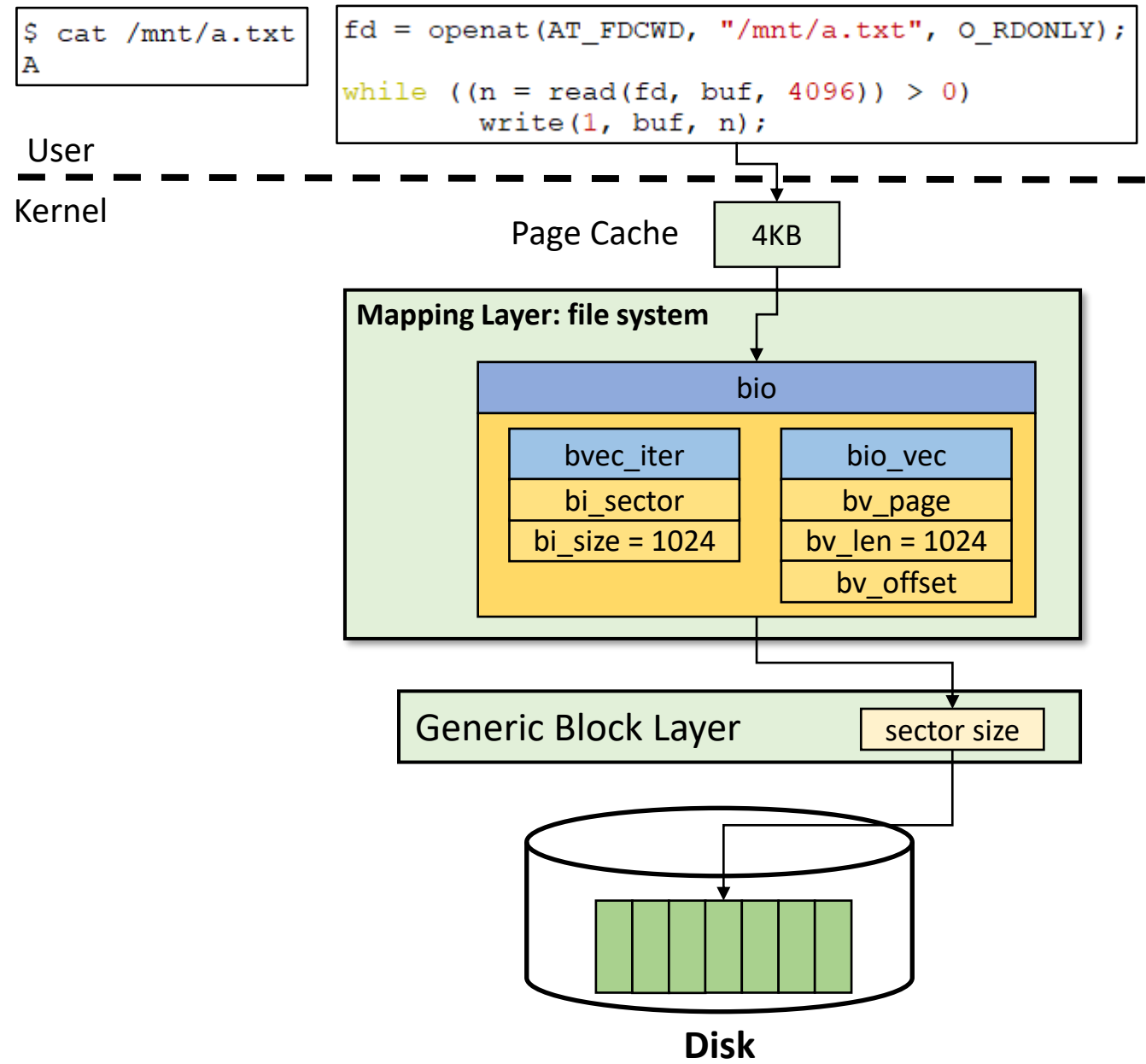
# File system block size & sector size



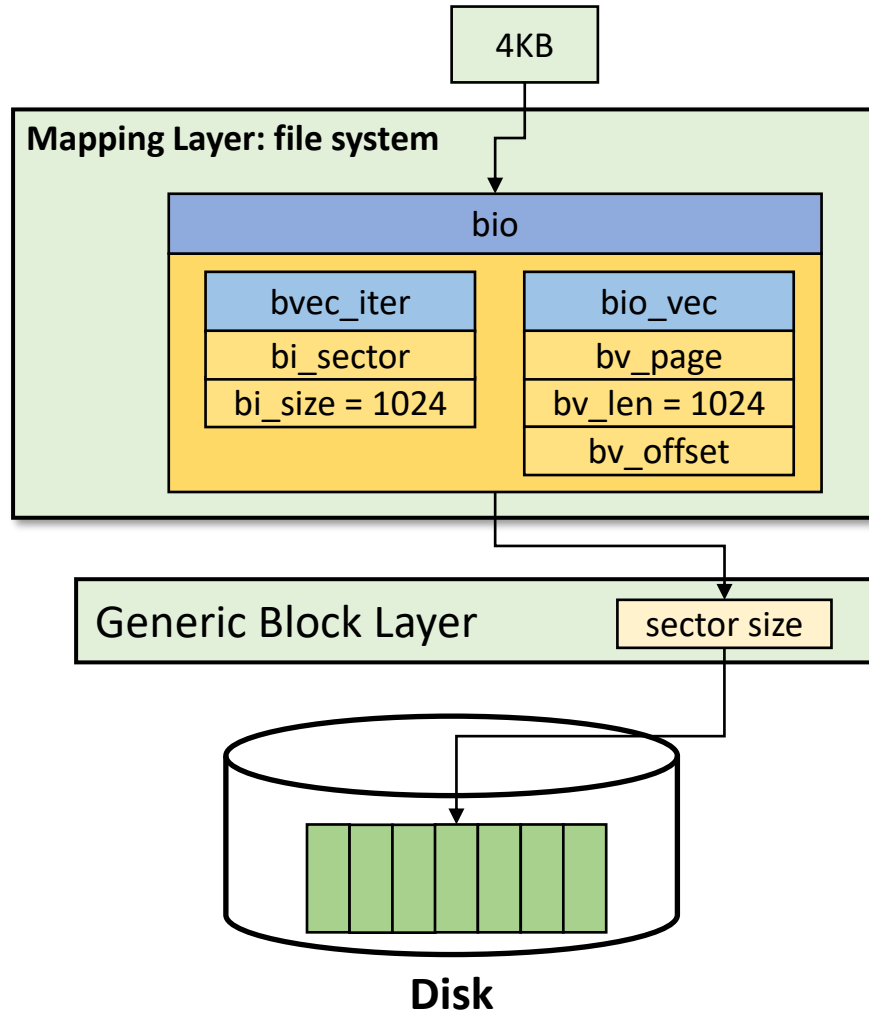
```
adrian@adrian-ubuntu:img$ dumpe2fs src.img | grep "Block size:"
dumpe2fs 1.45.5 (07-Jan-2020)
Block size: 4096
```

```
adrian@adrian-ubuntu:img$ dumpe2fs src-1024-byte-bsz-8MB.img | grep "Block size:"
"
dumpe2fs 1.45.5 (07-Jan-2020)
Block size: 1024
```

# File system block size & sector size: file system block size = 1024



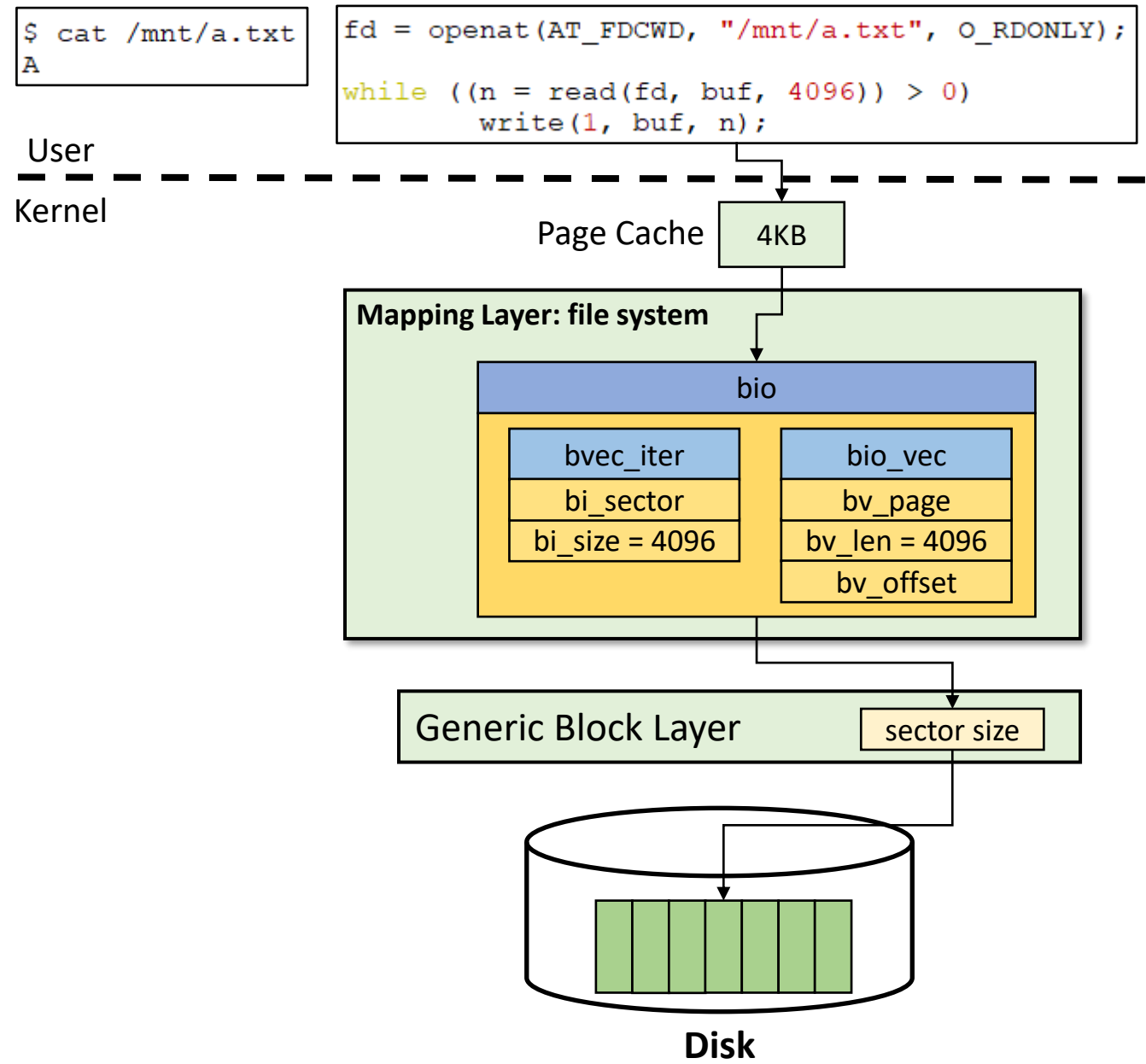
# File system block size & sector size: file system block size = 1024



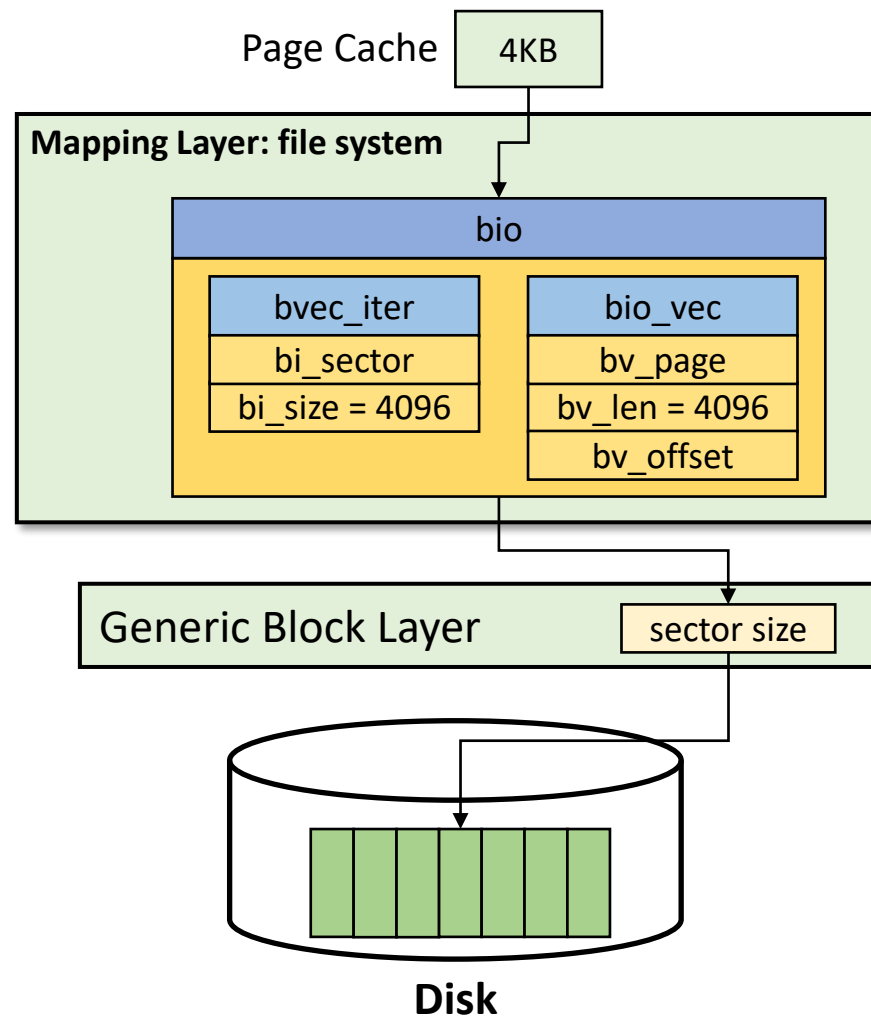
```
(gdb) bt
#0 0xffffffff811a9f8c in __bio_add_page (bio=bio@entry=0xffff8881047136c0, page=
=<optimized out>, len=len@entry=4096, off=off@entry=0) at /home/adrian/git-repo/
gdb-linux-real-mode/src/linux-5.11/arch/x86/include/asm/bitops.h:206
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff8881047136c0, page=p
age@entry=0xffffea0007e04440, len=len@entry=4096, offset=offset@entry=0) at /hom
e/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff811804a6 in ext4_mpage_readpages (inode=<optimized out>, rac=<opti
mized out>, page=<optimized out>, page@entry=0x0 <fixed_percpu_data>) at /home/a
drian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage.c:387
#3 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>) at /home/adrian/g
it-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:3261
```

```
(gdb) p bio->bi_iter
$7 = {
 bi_sector = 12828,
 bi_size = 1024,
 bi_idx = 0,
 bi_bvec_done = 0
}
(gdb) p *bio->bi_io_vec
$8 = {
 bv_page = 0xffffea0007e01180,
 bv_len = 1024,
 bv_offset = 0
}
```

# File system block size & sector size: file system block size = 4096



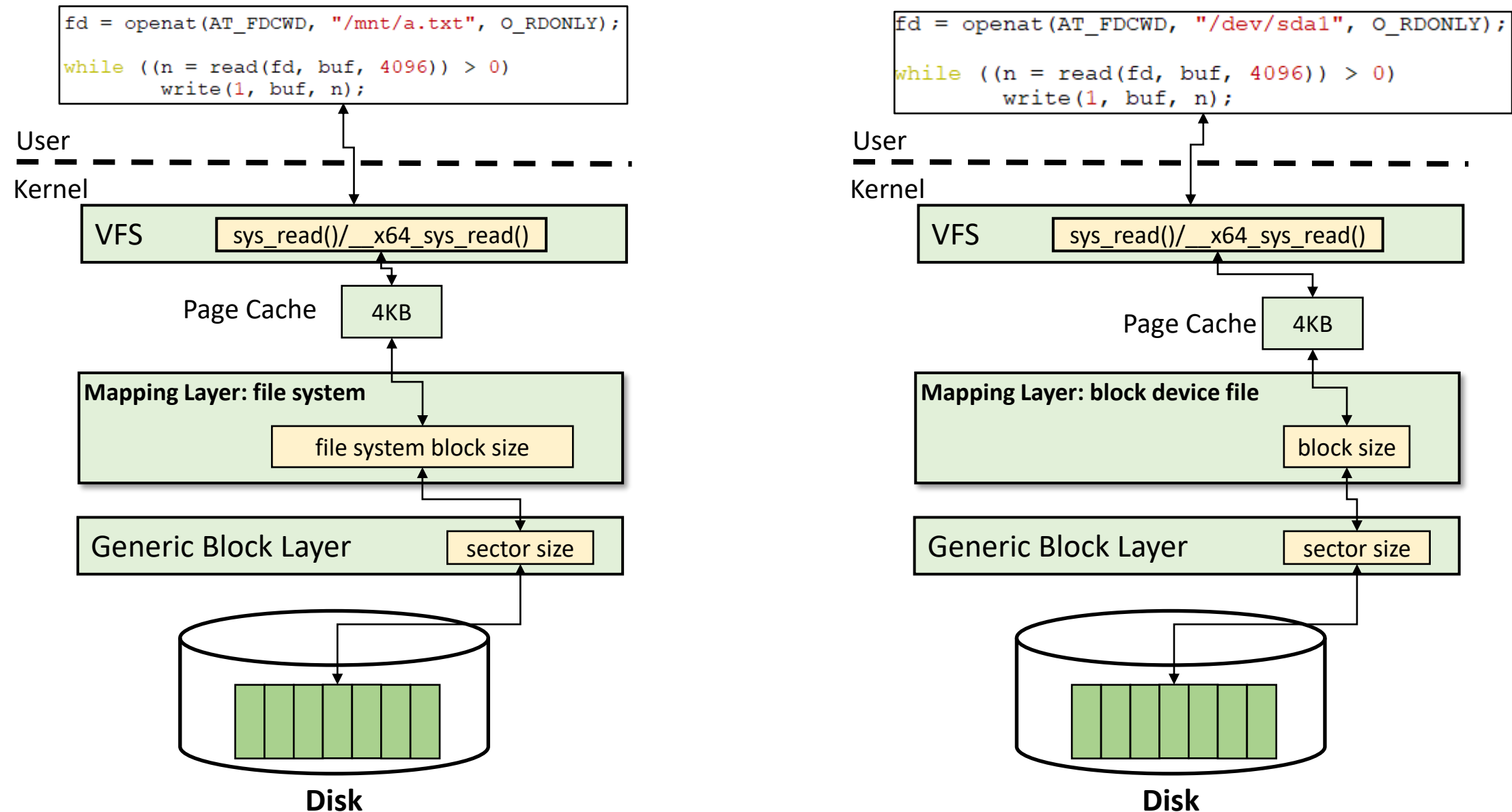
# File system block size & sector size: file system block size = 4096



```
(gdb) bt
#0 0xffffffff811a9f8c in __bio_add_page (bio=bio@entry=0xfffff8881047136c0, page=
=<optimized out>, len=len@entry=4096, off=off@entry=0) at /home/adrian/git-repo/
gdb-linux-real-mode/src/linux-5.11/arch/x86/include/asm/bitops.h:206
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xfffff8881047136c0, page=p
age@entry=0xffffea0007e04440, len=len@entry=4096, offset=offset@entry=0) at /hom
e/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff811804a6 in ext4_mpage_readpages (inode=<optimized out>, rac=<opti
mized out>, page=<optimized out>, page@entry=0x0 <fixed_percpu_data>) at /home/a
drian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage.c:387
#3 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>) at /home/adrian/g
it-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:3261
```

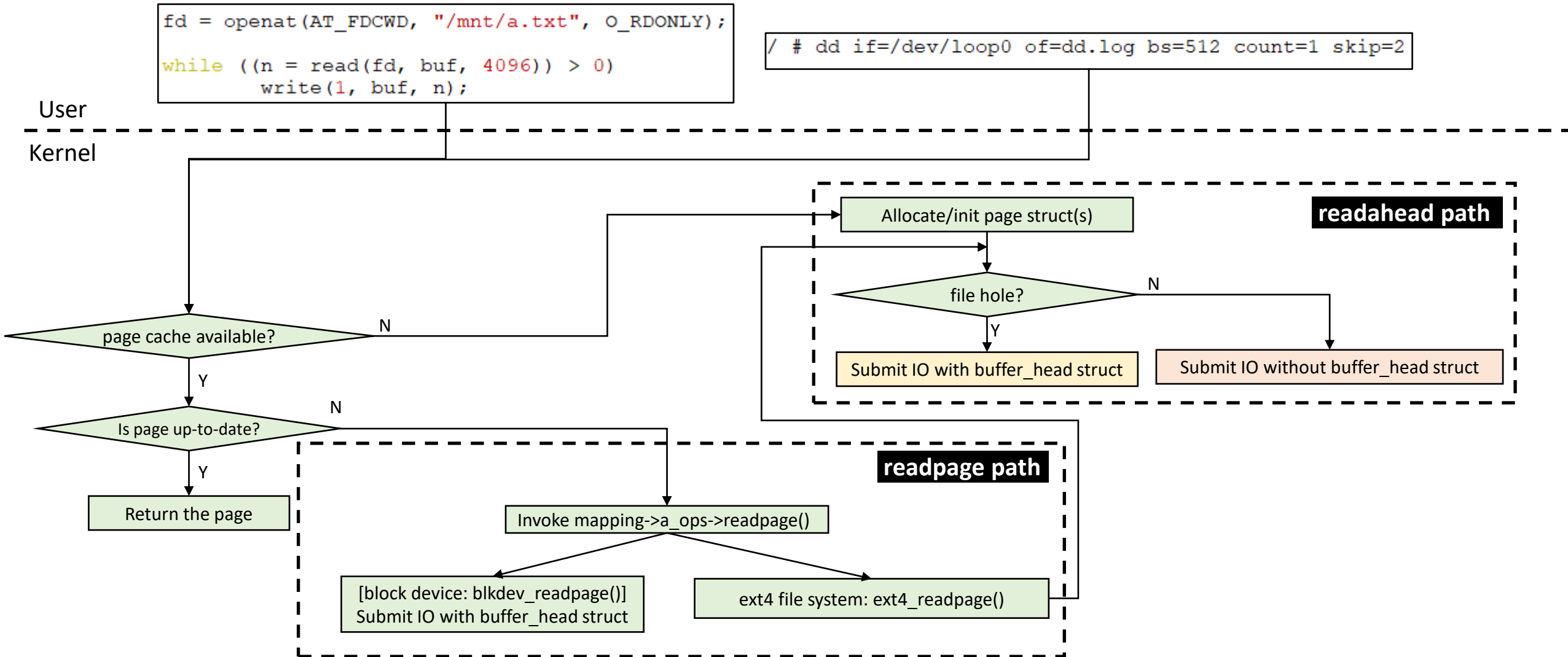
```
(gdb) p bio->bi_iter
$4 = {
 bi_sector = 48920,
 bi_size = 4096,
 bi_idx = 0,
 bi_bvec_done = 0
}
(gdb) p *bio->bi_io_vec
$5 = {
 bv_page = 0xffffea0007e05f60,
 bv_len = 4096,
 bv_offset = 0
}
```

# File system access & raw block device access

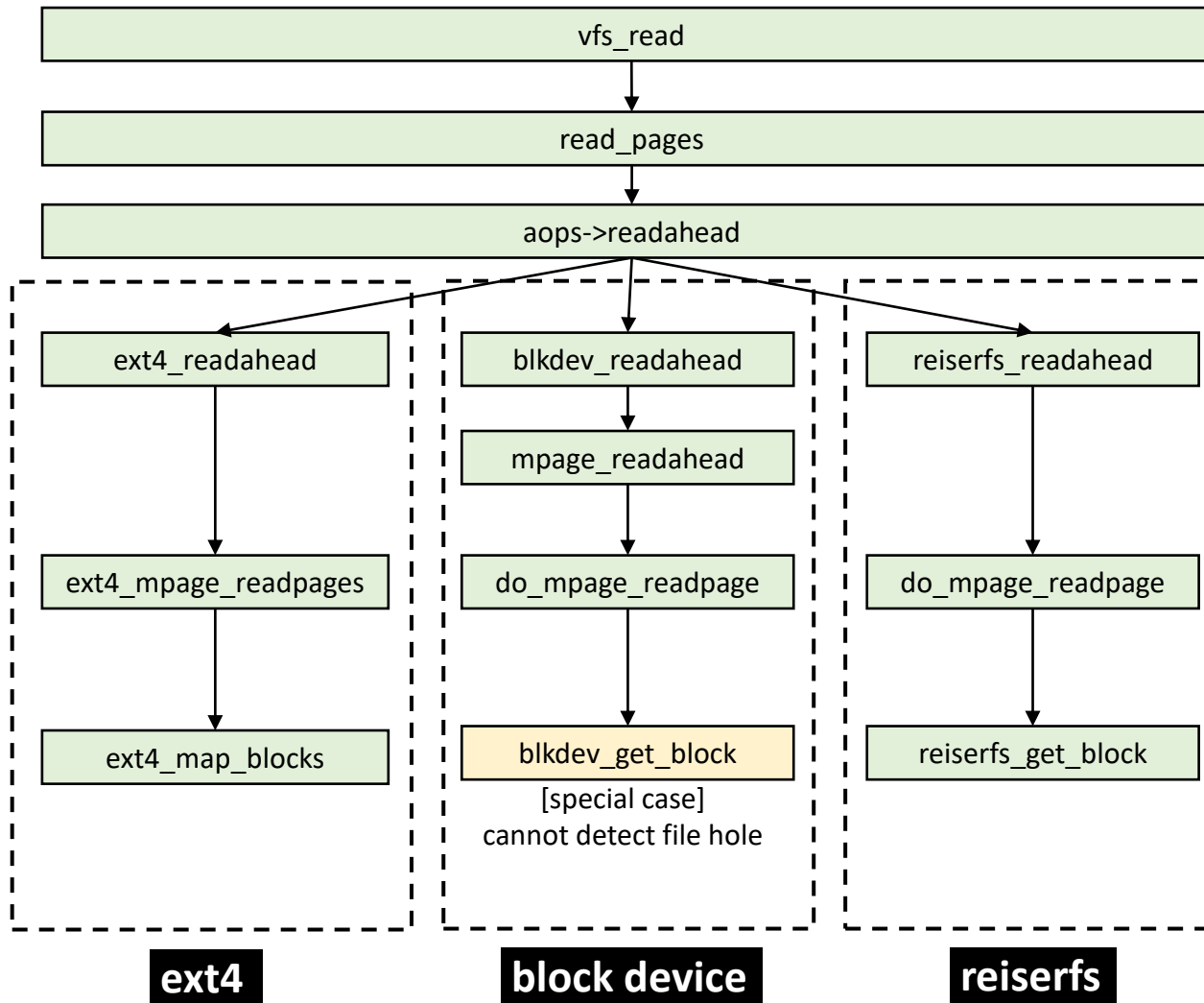


Submit IO with or without `buffer_head` struct?

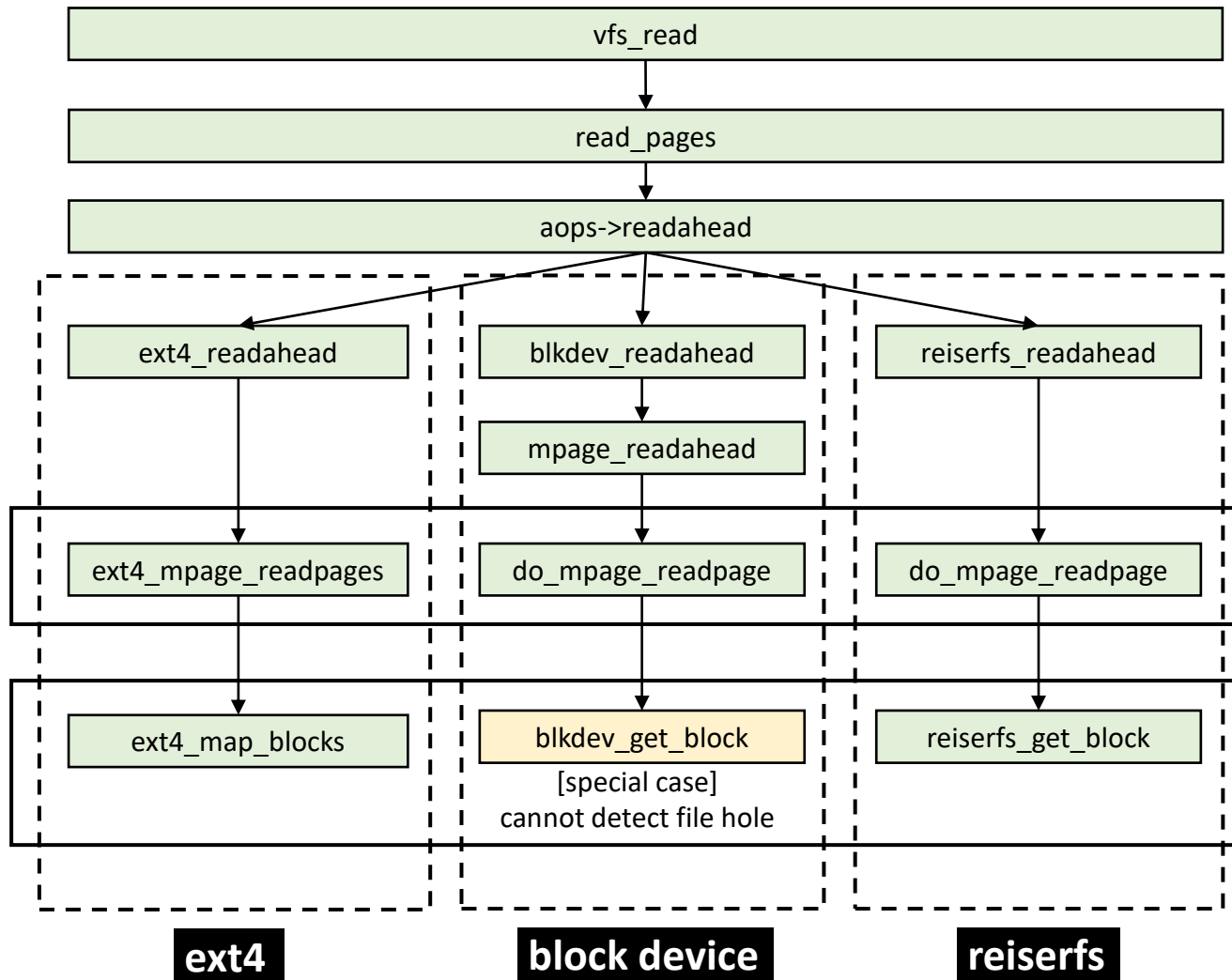
# With or without buffer\_head



# File hole detection – file system implementation



# File hole detection – file system implementation



```
static int
blkdev_get_block(struct inode *inode, sector_t iblock,
 struct buffer_head *bh, int create)
{
 bh->b_bdev = I_BDEV(inode);
 bh->b_blocknr = iblock;
 set_buffer_mapped(bh);
 return 0;
}

fs/block_dev.c
```

Detect file hole

Check file hole for specific blocks. Do not set  
MAPPED flag if blocks are holes

# With or without buffer\_head: system configuration

## Test Configuration

- ext4 file system
  - block size: 1024 bytes
- sector size: 512 bytes

## ext4 file system: block size

```
$ dumpe2fs src-1024-byte-bsz-8MB.img | grep "Block size"
Block size: 1024
```

## mount command

```
/ # mount /adrian/img/src-1024-byte-bsz-8MB.img /adrian/mnt/
```

## Sector Size

```
/ # mount
none on / type rootfs (rw)
none on /proc type proc (rw,relatime)
none on /sys type sysfs (rw,relatime)
none on /sys/kernel/debug type debugfs (rw,relatime)
/dev/loop0 on /adrian/mnt type ext4 (rw,relatime)
/ # blockdev --getss /dev/loop0
512
```

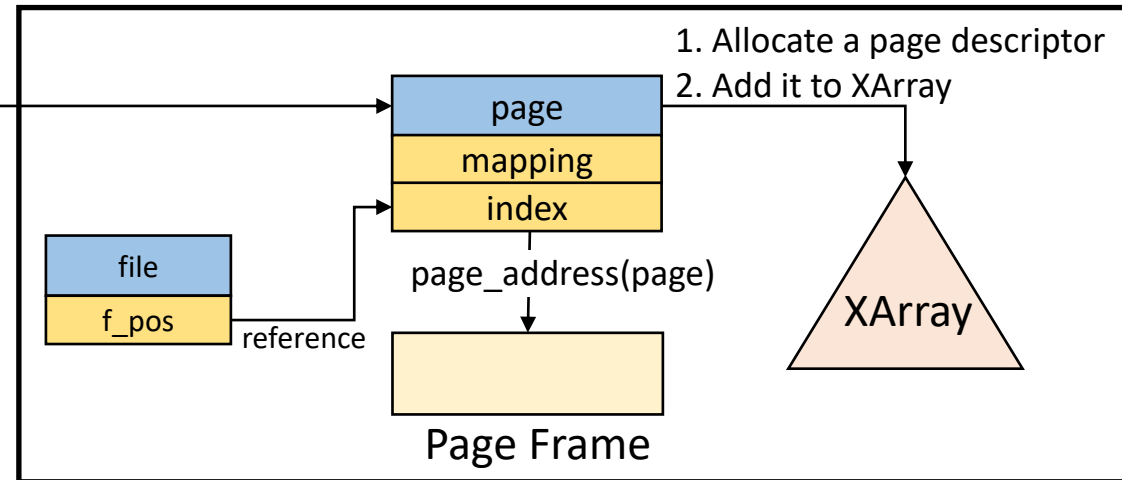
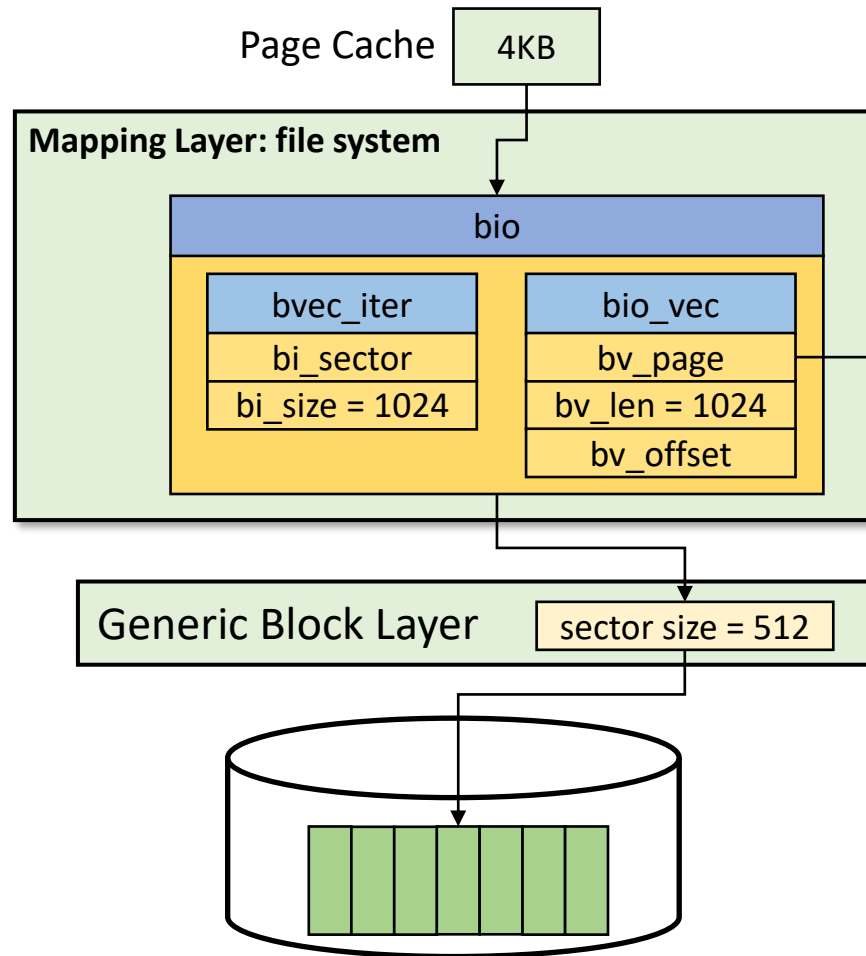
# File size = 1023 bytes

## read.c

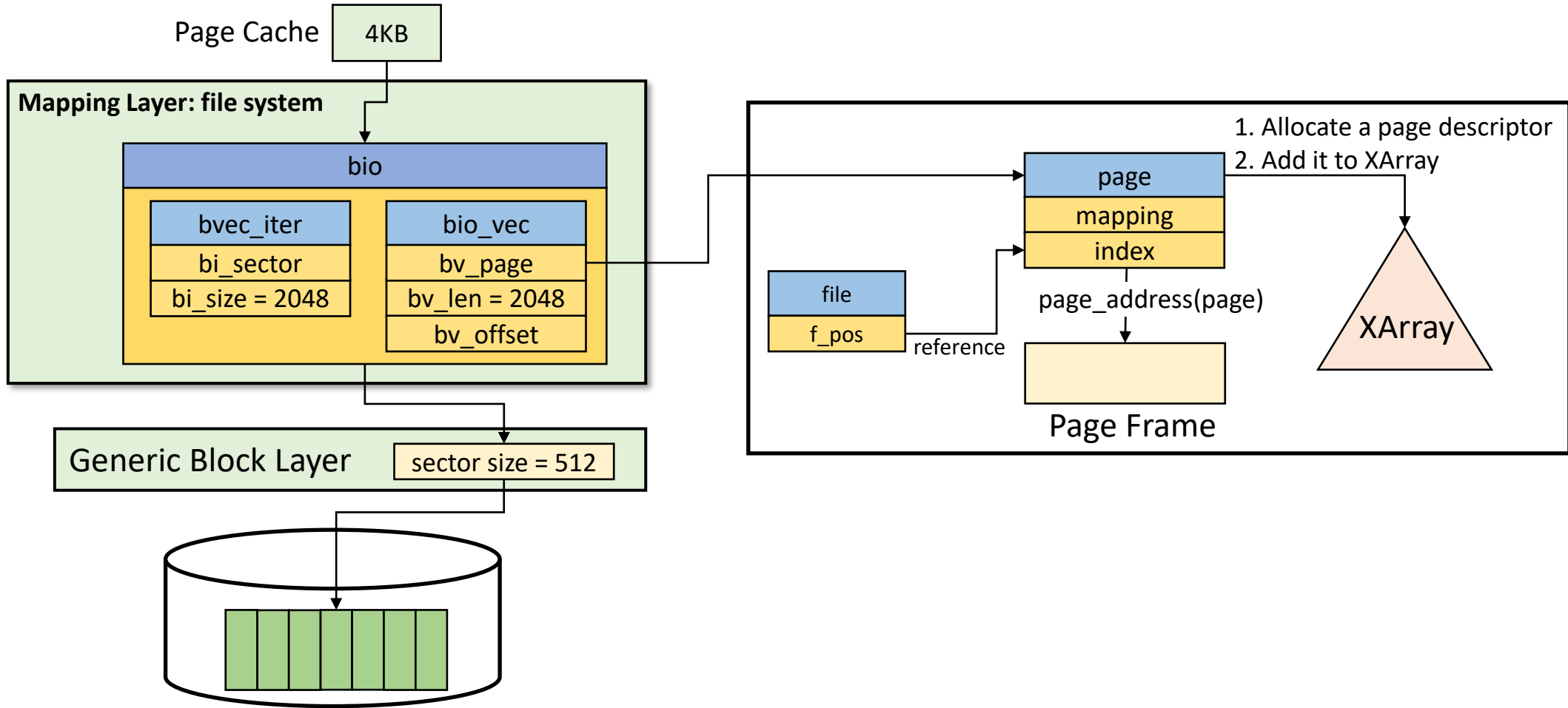
```
fd = openat(AT_FDCWD, argv[1], O_RDONLY);

while ((n = read(fd, buf, atoi(argv[2]))) > 0)
 write(1, buf, n);
```

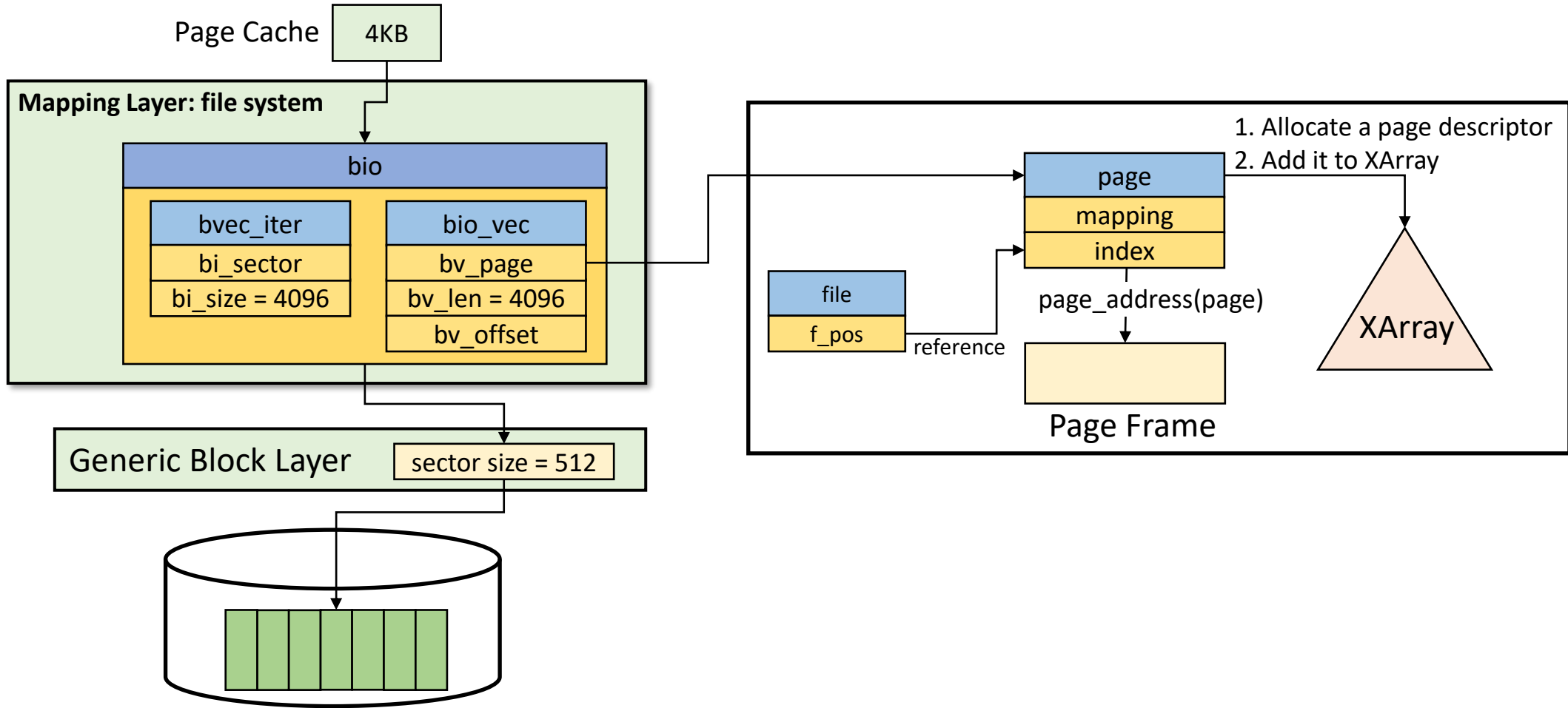
```
/ # stat /adrian/mnt/files/1023.txt
File: /adrian/mnt/files/1023.txt
Size: 1023 Blocks: 2 IO Block: 1024 regular file
Device: 700h/1792d Inode: 128 Links: 1
Access: (0644/-rw-r--r--) Uid: (0/ UNKNOWN) Gid: (0/ UNKNOWN)
Access: 2022-01-18 09:23:06.000000000 +0000
Modify: 2022-01-18 09:17:15.000000000 +0000
Change: 2022-01-18 09:17:15.000000000 +0000
/ # /read /adrian/mnt/files/1023.txt 512
```



# File size = 2047 bytes

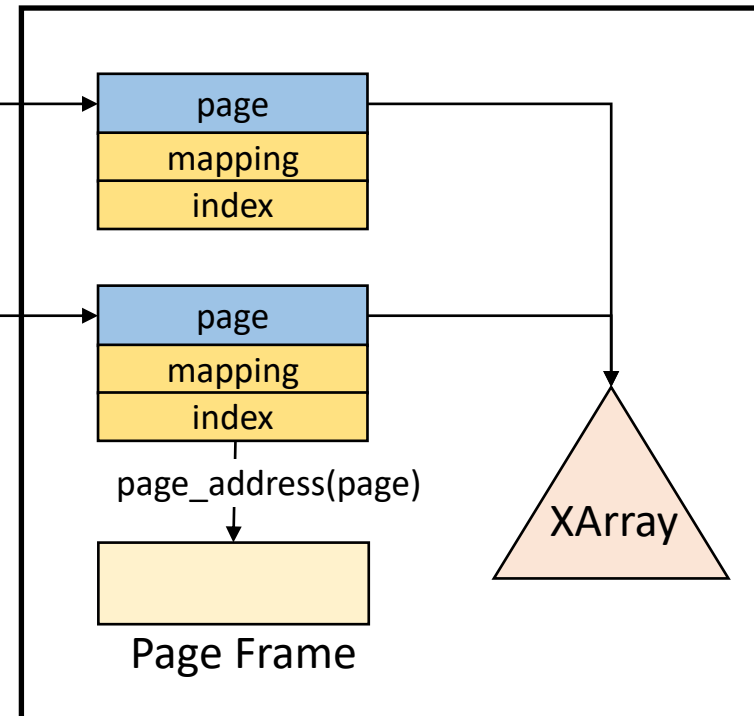
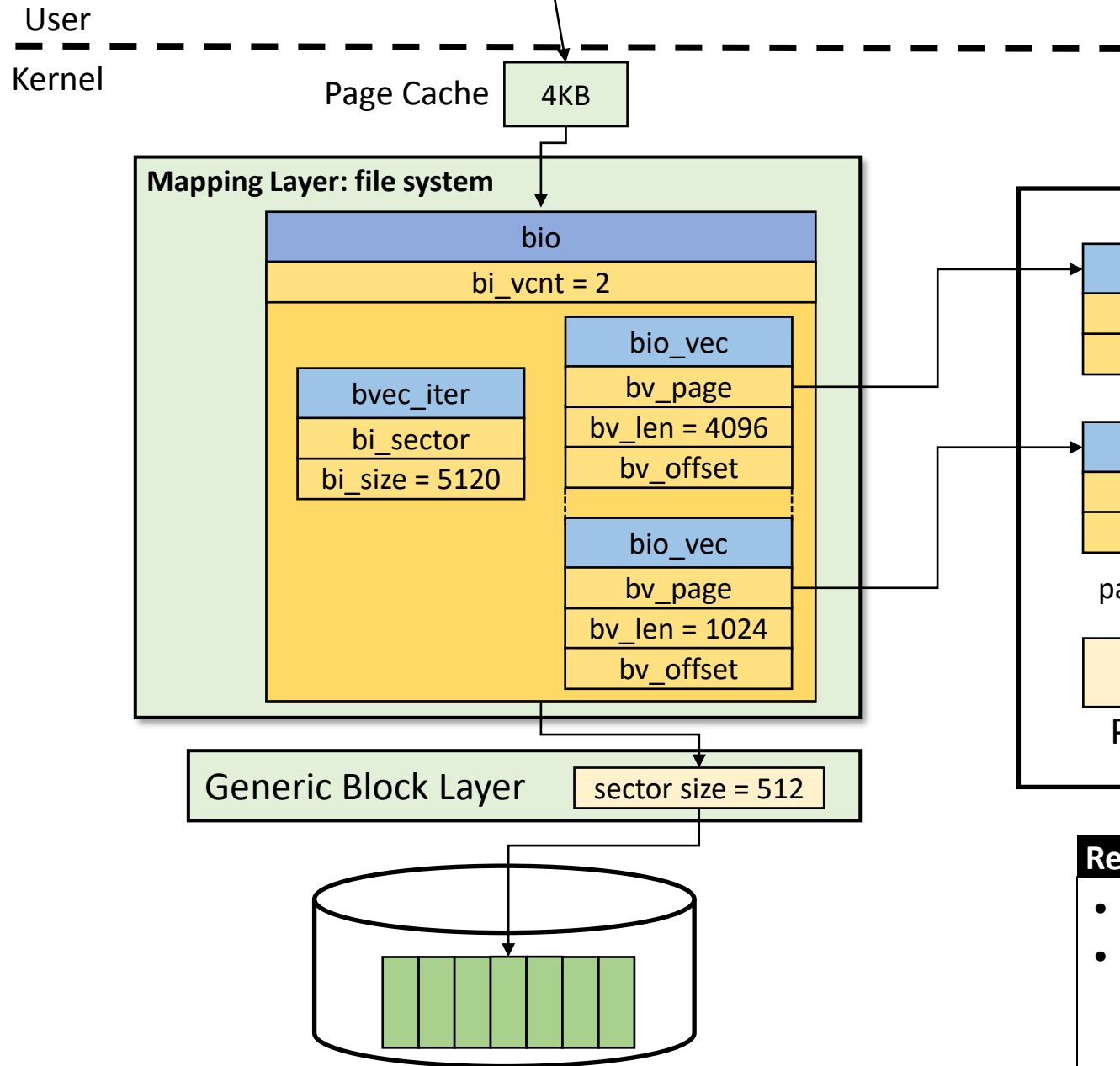


# File size = 4095 bytes



/ # /read /adrian/mnt/files/5119.txt 512

File size = 5119 bytes



### Readahead mechanism

- Spatial locality
- Default readahead pages: 32
  - If file size < “default readahead pages”, read the number of pages of the file.

# Readahead mechanism: default pages = 32

```
#define VM_READAHEAD_PAGES (SZ_128K / PAGE_SIZE)
```

```
#include/linux/pagemap.h
```

```
struct backing_dev_info *bdi_alloc(int node_id)
{
 struct backing_dev_info *bdi;

 bdi = kzalloc_node(sizeof(*bdi), GFP_KERNEL, node_id);
 if (!bdi)
 return NULL;

 if (bdi_init(bdi)) {
 kfree(bdi);
 return NULL;
 }
 bdi->capabilities = BDI_CAP_WRITEBACK | BDI_CAP_WRITEBACK_ACCT;
 bdi->ra_pages = VM_READAHEAD_PAGES;
 bdi->io_pages = VM_READAHEAD_PAGES;
 return bdi;
}
EXPORT_SYMBOL(bdi_alloc);
```

```
mm/backing-dev.c
```

```
728,0-1
```

```
int super_setup_bdi_name(struct super_block *sb, char *fmt, ...)
{
 struct backing_dev_info *bdi;
 int err;
 va_list args;

 bdi = bdi_alloc(NUMA_NO_NODE);
 if (!bdi)
 return -ENOMEM;

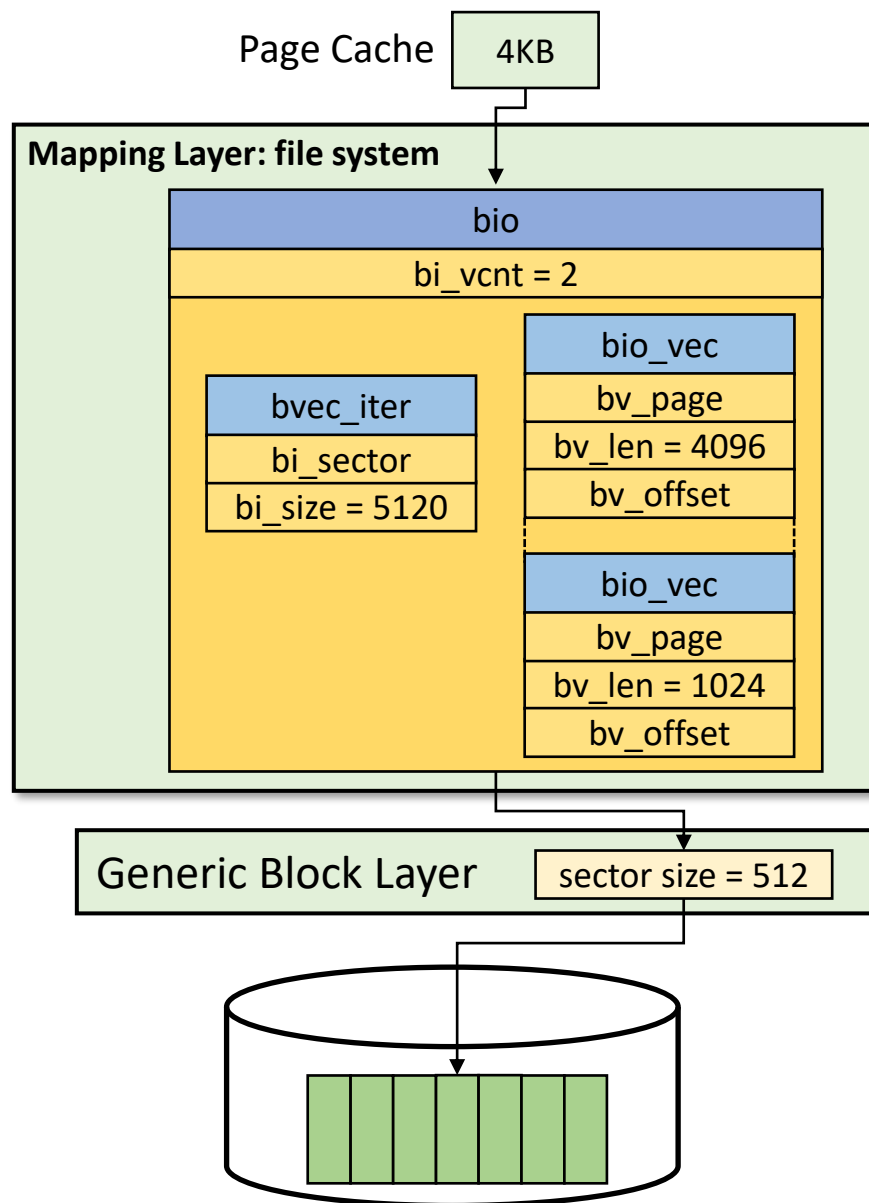
 +-- 8 lines: va_start(args, fmt);-----
 sb->s_bdi = bdi;

 return 0;
}
```

```
fs/super.c
```

```
1534,1-8
```

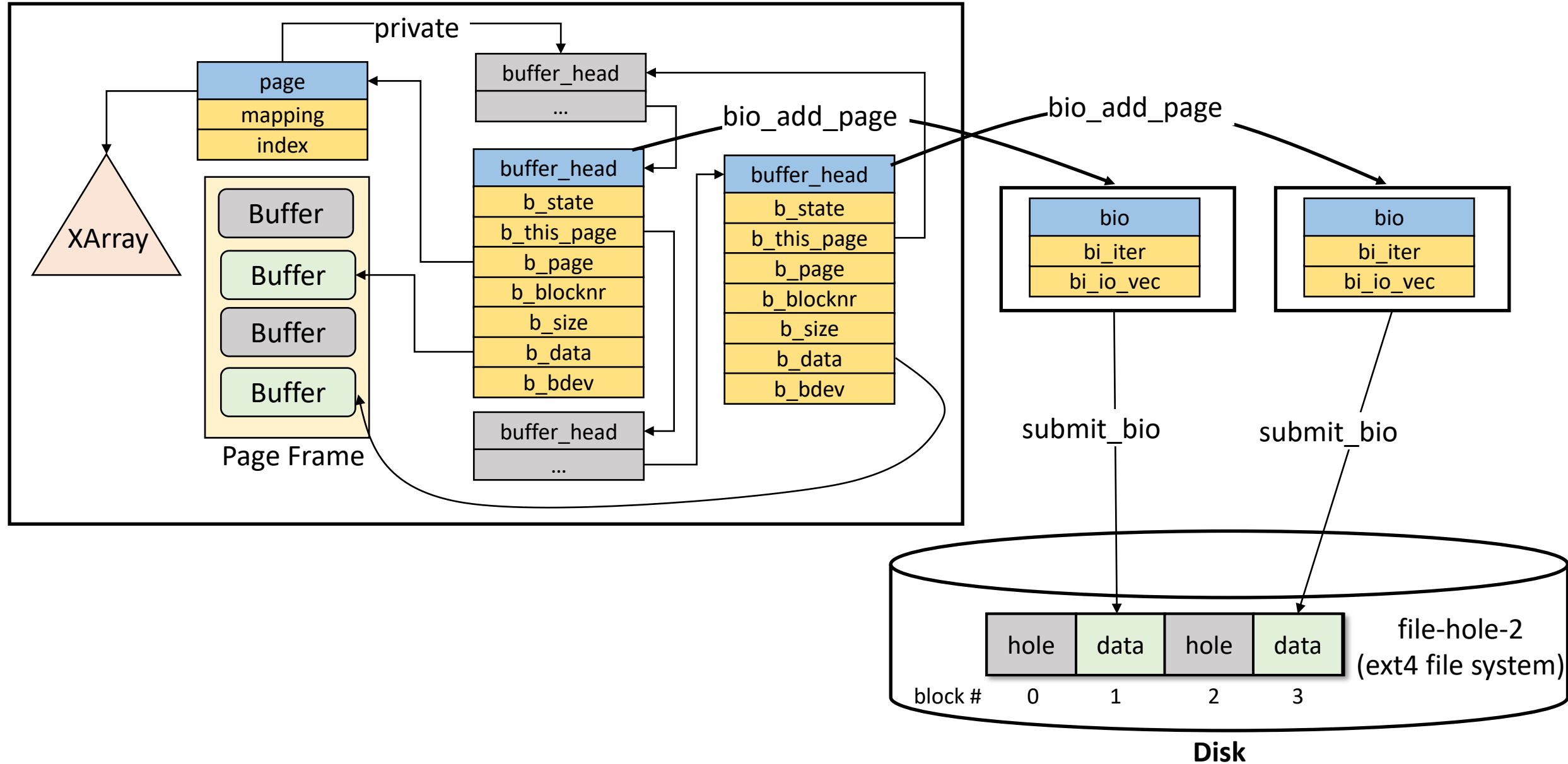
# File size = 5119 bytes



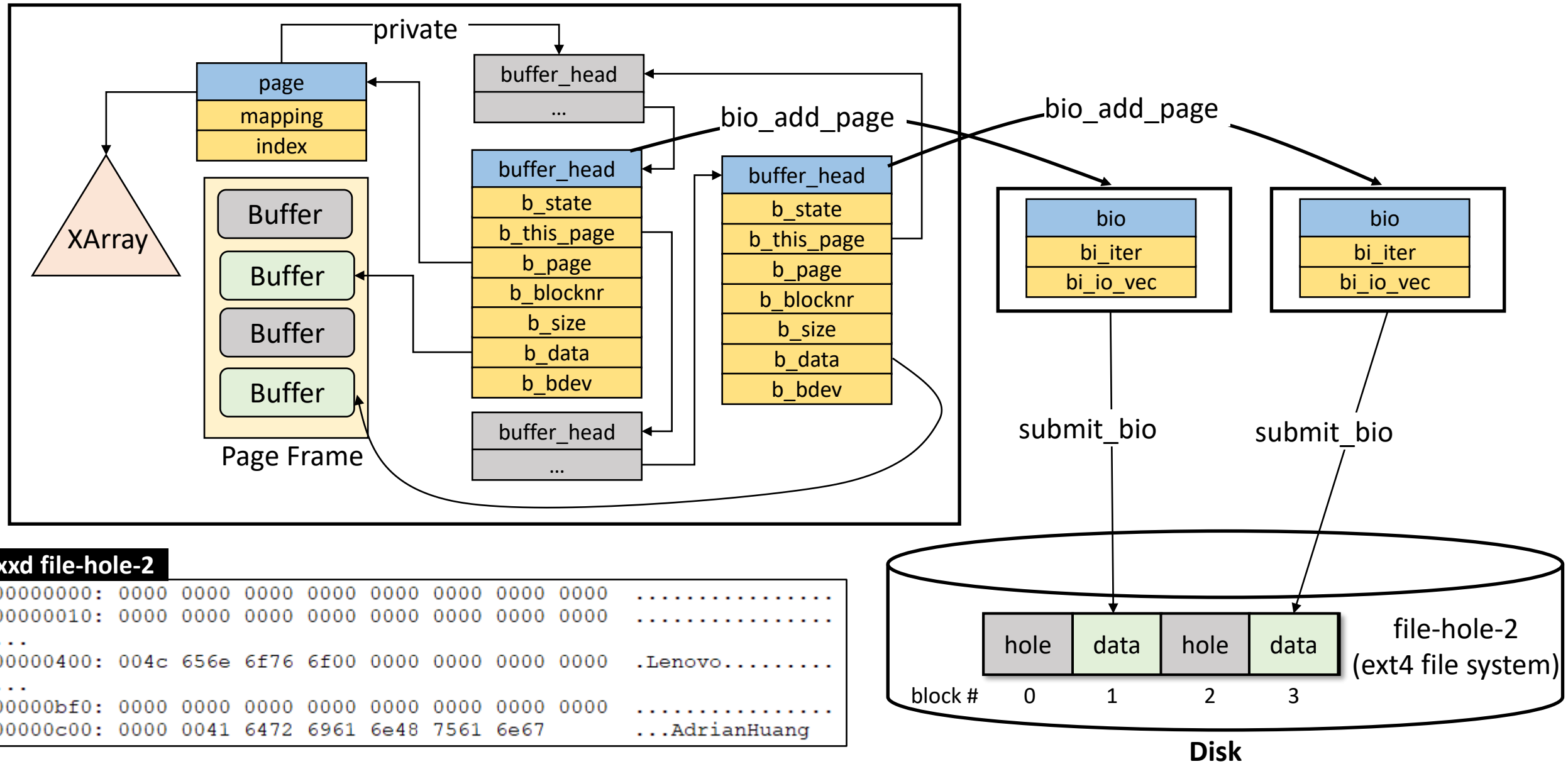
```
Thread 2 hit Breakpoint 7, submit_bio (bio=bio@entry=0xffff888104714900)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/bi
o.h:61
61 if (bio &&
(gdb) p bio
$30 = (struct bio *) 0xffff888104714900
(gdb) p *bio
$31 = {
 bi_next = 0x0 <fixed_percpu_data>,
 bi_disk = 0xffff8881046e0c00,
 bi_opf = 524288,
 bi_flags = 0,
 bi_ioprio = 0,
 bi_write_hint = 0,
 bi_status = 0 '\000',
 bi_partno = 0 '\000',
 __bi_remaining = {
 counter = 1
 },
 bi_iter = {
 bi_sector = 5300,
 bi_size = 5120,
 bi_idx = 0,
 bi_bvec_done = 0
 },
 bi_end_io = 0xffffffff8117ff70 <mpage_end_io>,
 bi_private = 0x0 <fixed_percpu_data>,
 {
 bi_integrity = 0x0 <fixed_percpu_data>
 },
 bi_vcnt = 2,
 bi_max_vecs = 2,
 __bi_cnt = {
 counter = 1
 },
 bi_io_vec = 0xffff888104714968,
 bi_pool = 0xffffffff81bc3be0 <fs_bio_set>,
 bi_inline_vecs = 0xffff888104714968
}
```

```
(gdb) p bio->bi_io_vec[0]
$33 = {
 bv_page = 0xffffea0007e00d58,
 bv_len = 4096,
 bv_offset = 0
}
(gdb) p bio->bi_io_vec[1]
$34 = {
 bv_page = 0xffffea0007e00dc8,
 bv_len = 1024,
 bv_offset = 0
}
```

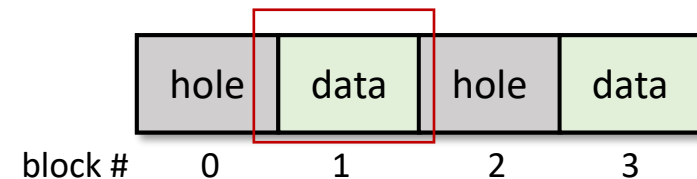
# File hole – Use buffer\_head



# File hole – Use buffer\_head



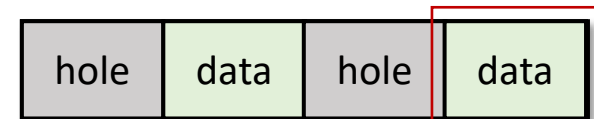
# File hole – Use buffer\_head



```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xfffff8881047120c0, page=0xfffffea0007e02bf8,
 len=len@entry=1024, off=off@entry=1024)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xfffff8881047120c0,
 page=<optimized out>, len=1024, offset=1024)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0,
 op_flags=op_flags@entry=0, bh=0xfffff8881004aad00,
 write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
 wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3034
#3 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#4 block_read_full_page (page=page@entry=0xfffffea0007e02bf8,
 get_block=0xffffffff81164860 <ext4_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#5 0xffffffff81180697 in ext4_mpage_readpages (inode=<optimized out>,
 rac=<optimized out>, page=<optimized out>,
 page@entry=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage
.c:404
#6 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:
3261
#7 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc900002cbc90,
 pages=pages@entry=0xffffc900002cbb70, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:1
30
#8 0xffffffff810abcd3 in page_cache_ra_unbounded (
 ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=1,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
38
#9 0xffffffff810abd88 in do_page_cache_ra (
 ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=<optimized out>,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
67
```

```
(gdb) bt
#0 submit_bio (bio=bio@entry=0xfffff8881047120c0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/bi
o.h:61
#1 0xffffffff8111e8ec in submit_bh_wbc (op=op@entry=0,
 op_flags=<optimized out>, op_flags@entry=0, bh=0xfffff8881004aad00,
 write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
 wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3054
#2 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#3 block_read_full_page (page=page@entry=0xfffffea0007e02bf8,
 get_block=0xffffffff81164860 <ext4_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#4 0xffffffff81180697 in ext4_mpage_readpages (inode=<optimized out>,
 rac=<optimized out>, page=<optimized out>,
 page@entry=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage
.c:404
#5 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:
3261
#6 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc900002cbc90,
 pages=pages@entry=0xffffc900002cbb70, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:1
30
#7 0xffffffff810abcd3 in page_cache_ra_unbounded (
 ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=1,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
38
#8 0xffffffff810abd88 in do_page_cache_ra (
 ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=<optimized out>,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
67
#9 0xffffffff810abf08 in ondemand_readahead (ractl=0xffffc900002cbc90,
 ra=0xfffff888100320398,
 hit_readahead_marker=hit_readahead_marker@entry=false,
 req_size=req_size@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:5
49
```

# File hole – Use buffer\_head



block # 0 1 2 3

```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff888104712240, page=0xffffea0007e02bf8,
len=len@entry=1024, off=off@entry=3072)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff888104712240,
page=<optimized out>, len=1024, offset=3072)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0,
op_flags=op_flags@entry=0, bh=0xffff8881004aa138,
write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
wbc=0x0 <fixed_percpu_data>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3034
#3 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#4 block_read_full_page (page=page@entry=0xffffea0007e02bf8,
get_block=0xffffffff81164860 <ext4_get_block>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#5 0xffffffff81180697 in ext4_mpage_readpages (inode=<optimized out>,
rac=<optimized out>, page=<optimized out>,
page@entry=0x0 <fixed_percpu_data>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage
.c:404
#6 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:
3261
#7 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc900002cbc90,
pages=pages@entry=0xffffc900002cbb70, skip_page=skip_page@entry=false)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:1
30
#8 0xffffffff810abcd3 in page_cache_ra_unbounded (
ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=1,
lookahead_size=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
38
#9 0xffffffff810abd88 in do_page_cache_ra (
ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=<optimized out>,
lookahead_size=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
67
```

```
(gdb) bt
#0 submit_bio (bio=bio@entry=0xffff888104712240)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/bi
o.h:61
#1 0xffffffff8111e8ec in submit_bh_wbc (op=op@entry=0,
op_flags=<optimized out>, op_flags@entry=0, bh=0xffff8881004aa138,
write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
wbc=0x0 <fixed_percpu_data>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3054
#2 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#3 block_read_full_page (page=page@entry=0xffffea0007e02bf8,
get_block=0xffffffff81164860 <ext4_get_block>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#4 0xffffffff81180697 in ext4_mpage_readpages (inode=<optimized out>,
rac=<optimized out>, page=<optimized out>,
page@entry=0x0 <fixed_percpu_data>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage
.c:404
#5 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:
3261
#6 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc900002cbc90,
pages=pages@entry=0xffffc900002cbb70, skip_page=skip_page@entry=false)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:1
30
#7 0xffffffff810abcd3 in page_cache_ra_unbounded (
ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=1,
lookahead_size=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
38
#8 0xffffffff810abd88 in do_page_cache_ra (
ractl=ractl@entry=0xffffc900002cbc90, nr_to_read=<optimized out>,
lookahead_size=<optimized out>)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:2
67
#9 0xffffffff810abf08 in ondemand_readahead (ractl=0xffffc900002cbc90,
ra=0xffff888100320398,
hit_readahead_marker=hit_readahead_marker@entry=false,
req_size=req_size@entry=1)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:5
49
```

# Call path with/without buffer\_head struct

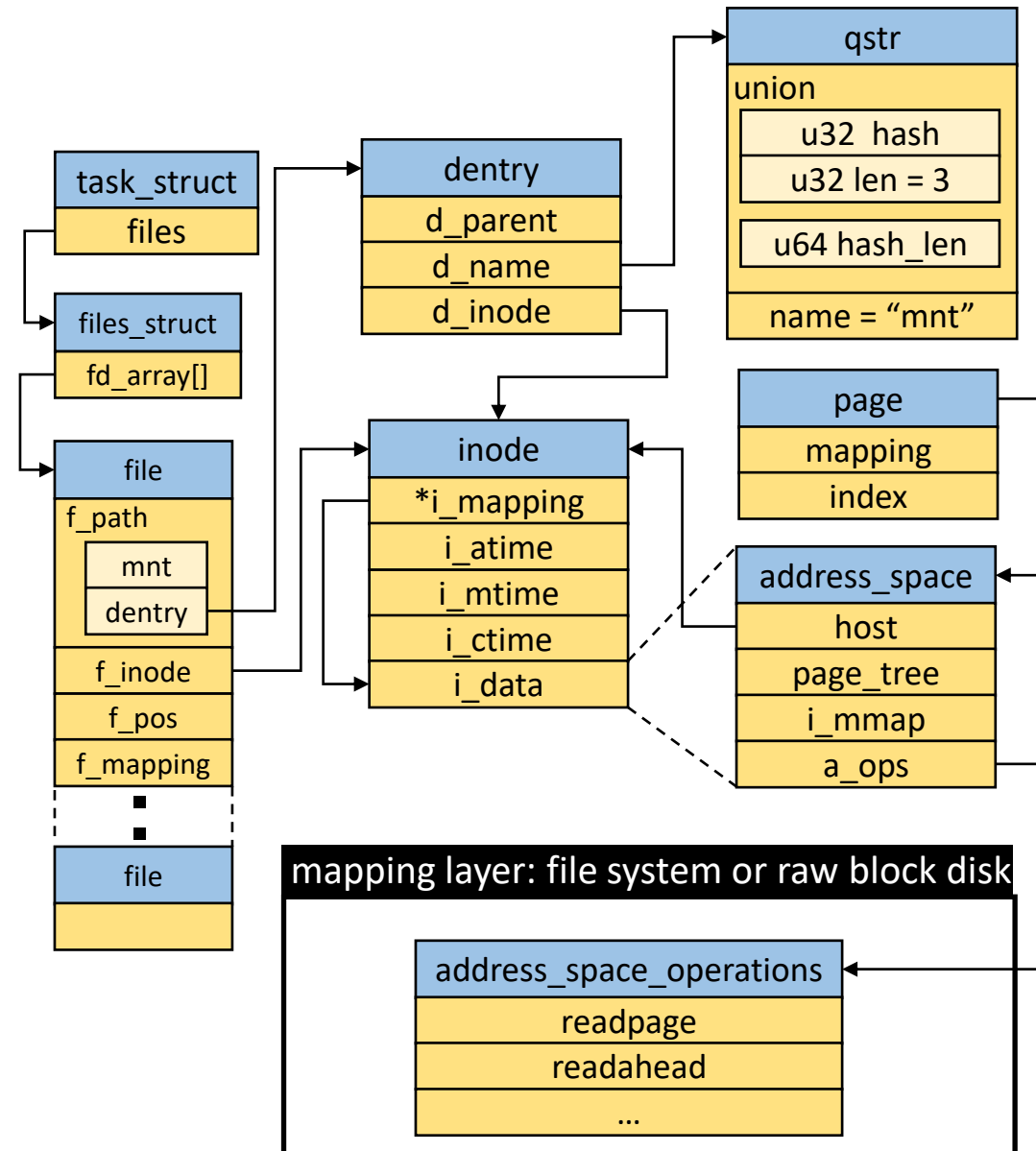
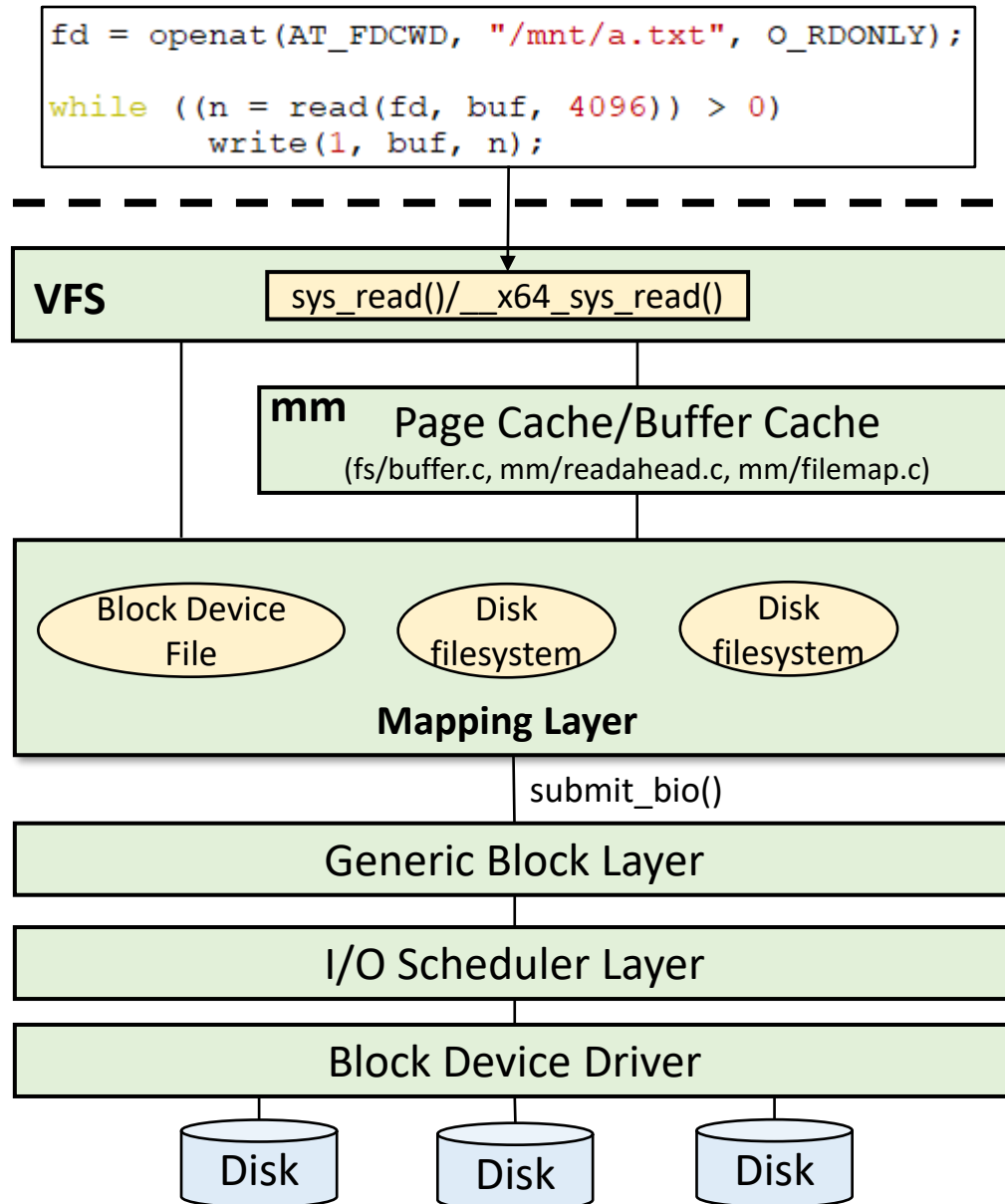
## Call path without buffer\_head

```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff888104714840, page=0xffffea0007e06350,
 len=len@entry=1024, off=off@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff888104714840,
 page=page@entry=0xffffea0007e06350, len=len@entry=1024,
 offset=offset@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff811804a6 in ext4_mpage_readpages (inode=<optimized out>,
 rac=<optimized out>, page=<optimized out>,
 page@entry=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage.c:387
#3 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:3261
#4 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc90000273c90,
 pages=pages@entry=0xffffc90000273b70, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:130
#5 0xffffffff810abcd3 in page_cache_ra_unbounded (
 ractl=ractl@entry=0xffffc90000273c90, nr_to_read=1,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:238
#6 0xffffffff810abd88 in do_page_cache_ra (
 ractl=ractl@entry=0xffffc90000273c90, nr_to_read=<optimized out>,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:267
#7 0xffffffff810abf08 in ondemand_readahead (ractl=0xffffc90000273c90,
 ra=0xffff8881001fbe98, hit_readahead_marker@entry=false,
 req_size=req_size@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:549
#8 0xffffffff810ac1b0 in page_cache_sync_ra (
 ractl=ractl@entry=0xffffc90000273c90, ra=ra@entry=0xffff8881001fbe98,
 req_count=req_count@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:577
```

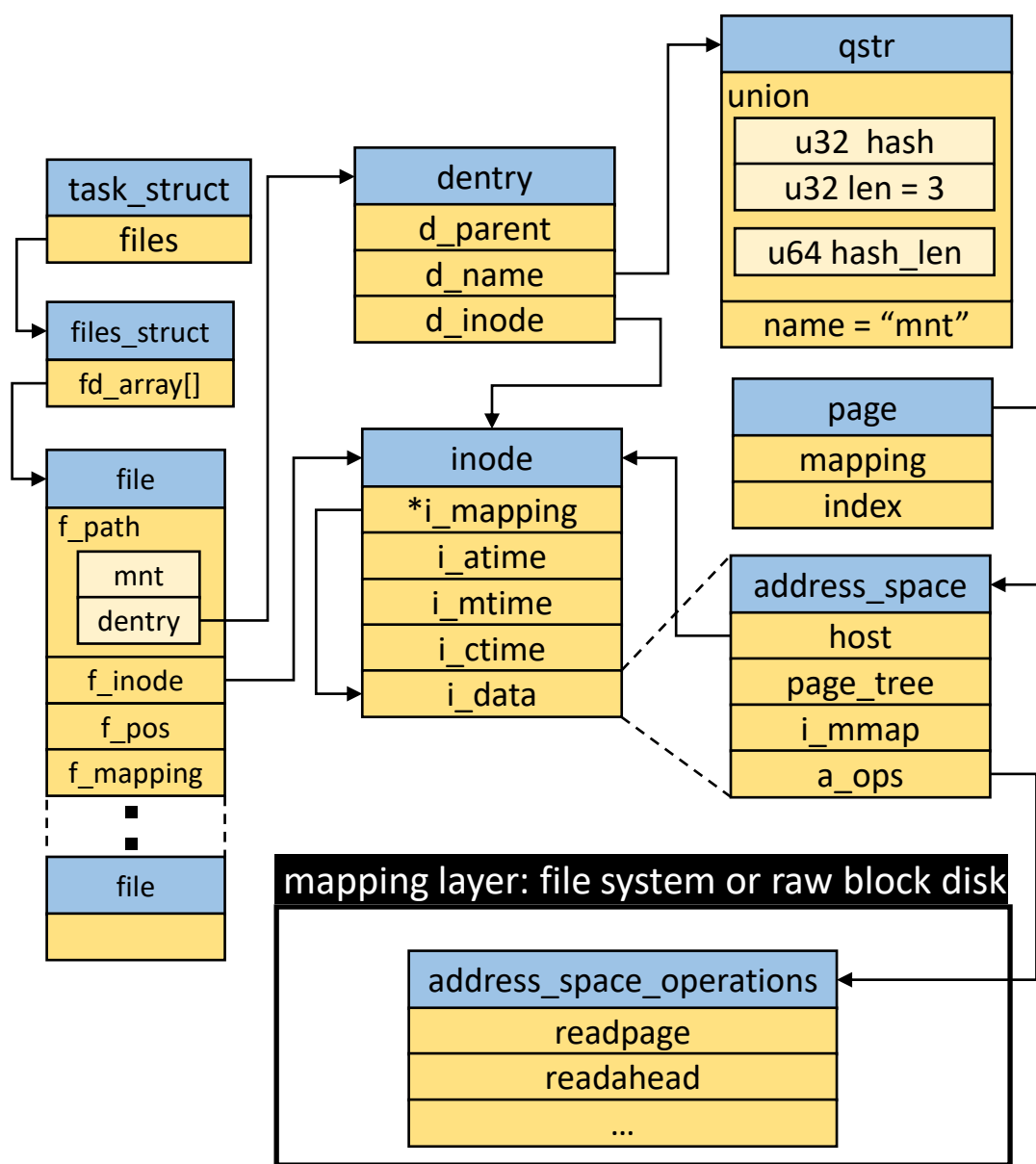
## Call path with buffer\_head

```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff888104714180, page=0xffffea0007e00d20,
 len=len@entry=1024, off=off@entry=1024)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff888104714180,
 page=<optimized out>, len=1024, offset=1024)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0,
 op_flags=op_flags@entry=0, bh=0xffff888100543000,
 write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
 wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3034
#3 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#4 block_read_full_page (page=page@entry=0xffffea0007e00d20,
 get_block=0xffffffff81164860 <ext4_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#5 0xffffffff81180697 in ext4_mpage_readpages (inode=<optimized out>,
 rac=<optimized out>, page=<optimized out>,
 page@entry=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage.c:404
#6 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:3261
#7 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc90000177c90,
 pages=pages@entry=0xffffc90000177b70, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:130
#8 0xffffffff810abcd3 in page_cache_ra_unbounded (
 ractl=ractl@entry=0xffffc90000177c90, nr_to_read=1,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:238
#9 0xffffffff810abd88 in do_page_cache_ra (
 ractl=ractl@entry=0xffffc90000177c90, nr_to_read=<optimized out>,
 lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:267
```

# Interaction between VFS/mm and mapping layer



# Interaction between VFS/mm and mapping layer



```
static const struct address_space_operations ext4_aops = {
 .readpage = ext4_readpage,
 .readahead = ext4_readahead,
 .writepage = ext4_writepage,
 .writepages = ext4_writepages,
 .write_begin = ext4_write_begin,
 .write_end = ext4_write_end,
 .set_page_dirty = ext4_set_page_dirty,
 .bmap = ext4_bmap,
 .invalidatepage = ext4_invalidatepage,
 .releasepage = ext4_releasepage,
 .direct_IO = noop_direct_IO,
 .migratepage = buffer_migrate_page,
 .is_partially_uptodate = block_is_partially_uptodate,
 .error_remove_page = generic_error_remove_page,
 .swap_activate = ext4_iomap_swap_activate,
};
```

fs/ext4/inode.c

3647,1

```
static const struct address_space_operations def_blk_aops = {
 .readpage = blkdev_readpage,
 .readahead = blkdev_readahead,
 .writepage = blkdev_writepage,
 .write_begin = blkdev_write_begin,
 .write_end = blkdev_write_end,
 .writepages = blkdev_writepages,
 .releasepage = blkdev_releasepage,
 .direct_IO = blkdev_direct_IO,
 .migratepage = buffer_migrate_page_norefs,
 .is_dirty_writeback = buffer_check_dirty_writeback,
};
```

fs/block\_dev.c

# Block device access – without buffer\_head (full-page access)

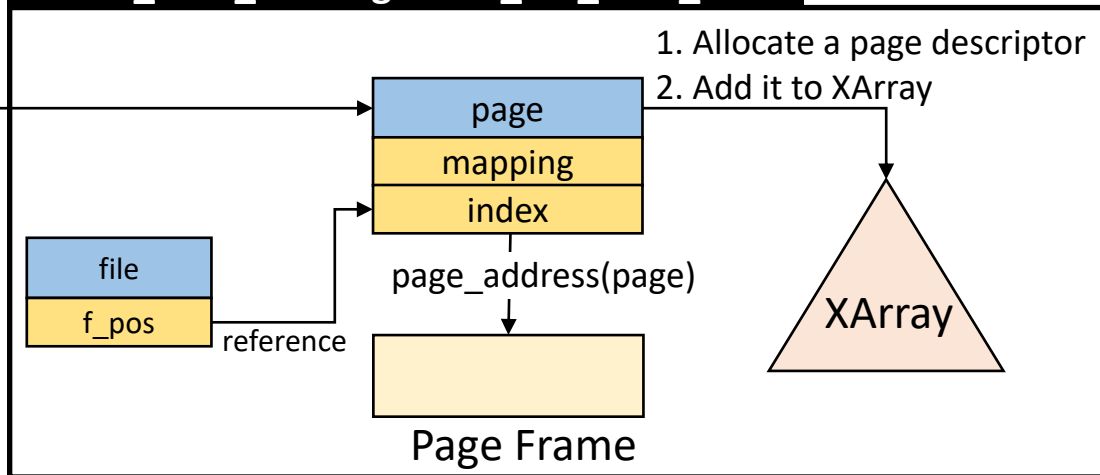
```
/ # mount /adrian/img/src-1024-byte-bsz-8MB.img /adrian/mnt/
loop0: detected capacity change from 16384 to 0
EXT4-fs (loop0): mounted filesystem with ordered data mode. Opts: (null). Quota.
ext4 filesystem being mounted at /adrian/mnt supports timestamps until 2038 (0x)

/ # blockdev --getss /dev/loop0
512

/ # blockdev --getbsz /dev/loop0
1024
```

```
/ # dd if=/dev/loop0 of=/dd.log bs=512 count=4 skip=512
```

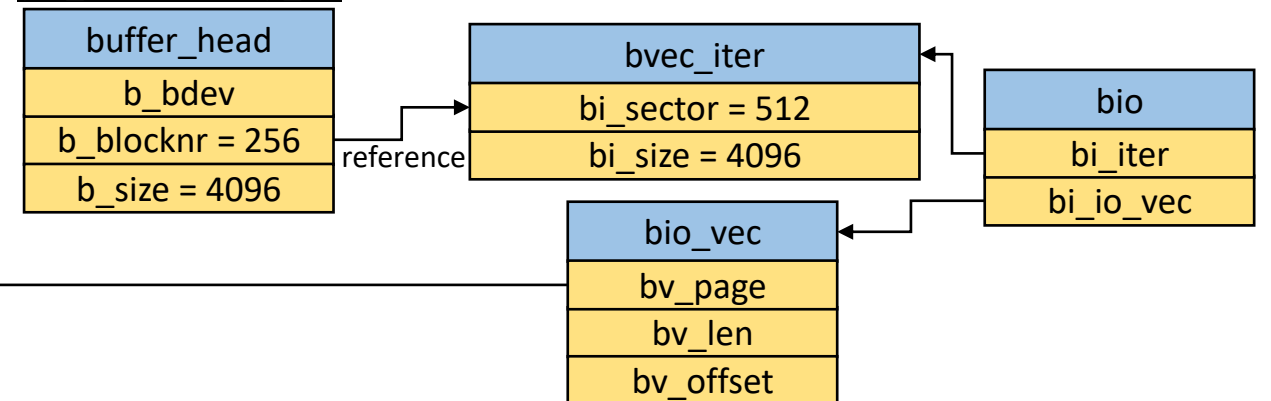
## blkdev\_read\_iter -> generic\_file\_read\_iter



`bio_add_page(...)`

## read\_pages -> blkdev\_readahead -> mpage\_readahead -> do\_mpage\_readpage

### get\_block -> map\_bh

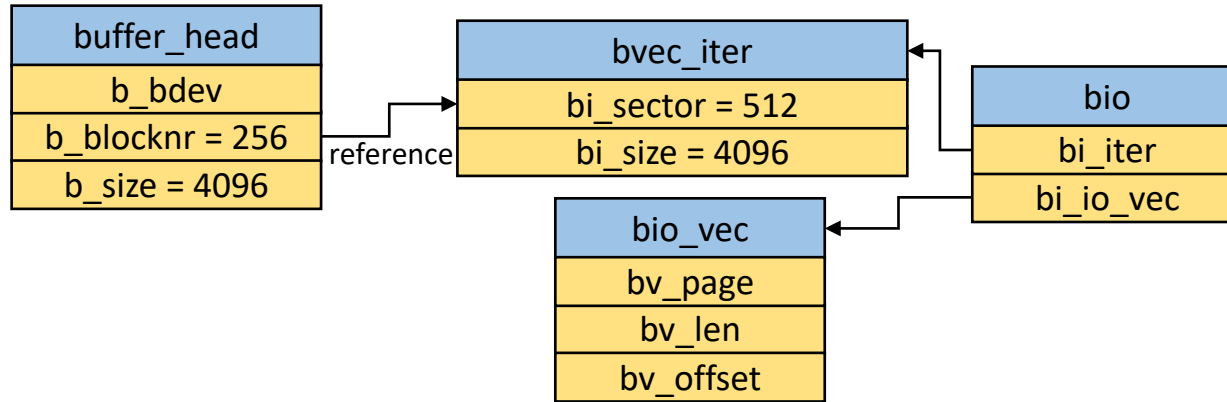


# Block device access – without buffer\_head (full-page access)

```
/ # dd if=/dev/loop0 of=/dd.log bs=512 count=4 skip=512
```

read\_pages -> blkdev\_readahead -> mpage\_readahead ->  
do\_mpage\_readpage

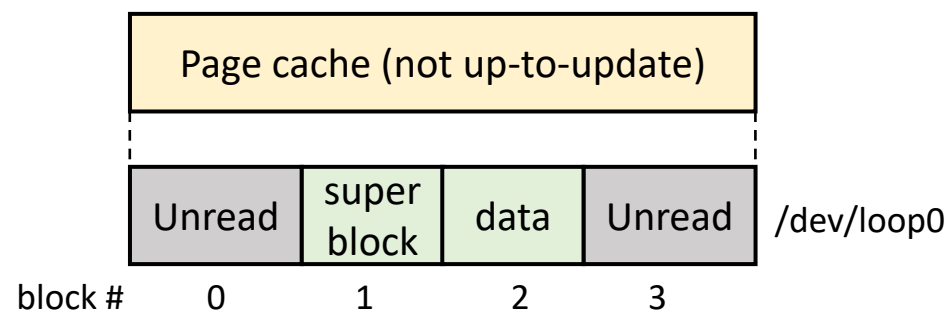
get\_block -> map\_bh



```
(gdb) f
#0 submit_bio (bio=bio@entry=0xfffff888104a1b840)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/mpage.c:395
o.h:61
61 if (bio &&
(gdb) p bio->bi_iter
$51 = {
 bi_sector = 512,
 bi_size = 4096,
 bi_idx = 0,
 bi_bvec_done = 0
}
(gdb) p bio->bi_vcnt
$52 = 1
(gdb) p bio->bi_io_vec[0]
$53 = {
 bv_page = 0xfffffea0003903498,
 bv_len = 4096,
 bv_offset = 0
}
```

```
(gdb) f
#2 mpage_readahead (rac=0xffffc90000333cb8,
 get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/mpage.c:395
395 mpage_bio_submit(REQ_OP_READ, REQ_RAHEAD, args.bio);
(gdb) p args.map_bh
$56 = {
 b_state = 0,
 b_this_page = 0x0 <fixed_percpu_data>,
 b_page = 0xfffffea0003903498,
 b_blocknr = 256,
 b_size = 4096,
 b_data = 0x0 <fixed_percpu_data>,
 b_bdev = 0xfffff8881004a8000,
 b_end_io = 0x0 <fixed_percpu_data>,
 b_private = 0x0 <fixed_percpu_data>,
 b_assoc_buffers = {
 next = 0x0 <fixed_percpu_data>,
 prev = 0x0 <fixed_percpu_data>
 },
 b_assoc_map = 0x0 <fixed_percpu_data>,
 b_count = {
 counter = 0
 },
 b_uptodate_lock = {
 {
 rlock = {
 raw_lock = {
 {
 val = {
 counter = 0
 },
 {
 locked = 0 '\000',
 pending = 0 '\000'
 },
 {
 locked_pending = 0,
 tail = 0
 }
 }
 }
 }
 }
 }
}
```

# Block device access – with buffer\_head



```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff888104712180, page=0xffffea0003904f48,
 len=len@entry=1024, off=off@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff888104712180,
 page=<optimized out>, len=1024, offset=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0,
 op_flags=op_flags@entry=0, bh=0xffff8881004aa618,
 write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
 wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#3 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op=op@entry=0,
 op_flags=op_flags@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#4 block_read_full_page (page=<optimized out>,
 get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#5 0xffffffff81123523 in blkdev_readpage (file=<optimized out>,
 page=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:6
31
```

```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff8881047120c0, page=0xffffea0003904f48,
 len=len@entry=1024, off=off@entry=3072)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff8881047120c0,
 page=<optimized out>, len=1024, offset=3072)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0,
 op_flags=op_flags@entry=0, bh=0xffff8881004aaa90,
 write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
 wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3034
#3 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op=op@entry=0,
 op_flags=op_flags@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#4 block_read_full_page (page=<optimized out>,
 get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#5 0xffffffff81123523 in blkdev_readpage (file=<optimized out>,
 page=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:6
31
```

Backup

# [File system] page cache without buffer\_head

```
(gdb) bt
#0 __bio_add_page (bio=bio@entry=0xffff888104714840, page=0xffffea0007e06350, len=len@entry=1024, off=off@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:906
#1 0xffffffff811a9fde in bio_add_page (bio=bio@entry=0xffff888104714840, page=page@entry=0xffffea0007e06350, len=len@entry=1024, offset=offset@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:941
#2 0xffffffff811804a6 in ext4_mpage_readpages (inode=<optimized out>, rac=<optimized out>, page=<optimized out>, page@entry=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/readpage.c:387
#3 0xffffffff81161d5e in ext4_readahead (rac=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/inode.c:3261
#4 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc90000273c90, pages=pages@entry=0xffffc90000273b70, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:130
#5 0xffffffff810abcd3 in page_cache_ra_unbounded (ractl=ractl@entry=0xffffc90000273c90, nr_to_read=1, lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:238
#6 0xffffffff810abd88 in do_page_cache_ra (ractl=ractl@entry=0xffffc90000273c90, nr_to_read=<optimized out>, lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:267
#7 0xffffffff810abf08 in ondemand_readahead (ractl=0xffffc90000273c90, ra=0xffff8881001f8e98, hit_readahead_marker=hit_readahead_marker@entry=false, req_size=req_size@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:549
#8 0xffffffff810ac1b0 in page_cache_sync_ra (ractl=ractl@entry=0xffffc90000273c90, ra=ra@entry=0xffff8881001f8e98, req_count=req_count@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:577
#9 0xffffffff810a57b9 in page_cache_sync_readahead (req_count=1, index=0, file=0xffff8881001f8e00, ra=0xffff8881001f8e98, mapping=0xffff88810044cc00)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/pagemap.h:839
#10 generic_file_buffered_read_get_pages (iocb=iocb@entry=0xffffc90000273e98, iter=iter@entry=0xffffc90000273d30, nr=nr@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2365
#11 0xffffffff810a60cf in generic_file_buffered_read (iocb=iocb@entry=0xffffc90000273e98, iter=iter@entry=0xffffc90000273e70, written=written@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2485
#12 0xffffffff810a648a in generic_file_read_iter (iocb=iocb@entry=0xffffc90000273e98, iter=iter@entry=0xffffc90000273e70)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2637
#13 0xffffffff8115716f in ext4_file_read_iter (to=0xffffc90000273e70, iocb=0xffffc90000273e98) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/file.c:131
#14 ext4_file_read_iter (iocb=0xffffc90000273e98, to=0xffffc90000273e70) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/ext4/file.c:114
#15 0xffffffff810edfd3 in call_read_iter (file=0xffff8881001f8e00, iter=0xffffc90000273e70, kio=0xffffc90000273e98)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/fs.h:1895
#16 new_sync_read (filp=filp@entry=0xffff8881001f8e00, buf=buf@entry=0x4c3300 "", len=<optimized out>, ppos=ppos@entry=0xffffc90000273ef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:415
#17 0xffffffff810ef2a7 in vfs_read (file=file@entry=0xffff8881001f8e00, buf=buf@entry=0x4c3300 "", count=count@entry=512, pos=pos@entry=0xffffc90000273ef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:496
#18 0xffffffff810ef7b3 in ksys_read (fd=<optimized out>, buf=0x4c3300 "", count=512) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:634
#19 0xffffffff810ef845 in __do_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:644
#20 __se_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#21 __x64_sys_read (regs=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#22 0xffffffff812b05f2 in do_syscall_64 (nr=<optimized out>, regs=0xffffc90000273f58) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/common.c:46
#23 0xffffffff8140007c in entry_SYSCALL_64 () at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/entry_64.S:120
```

# [Block device] page cache with buffer\_head

```
/ # dd if=/dev/loop0 of=dd.log bs=512 count=4
```

```
(gdb) bt
#0 bio_add_page (bio=bio@entry=0xffff888100246f00, page=0xffffea00038f87c8, len=1024, offset=0) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:936
#1 0xffffffff8111e89b in submit_bh_wbc (op=op@entry=0, op_flags=op_flags@entry=0, bh=0xffff8881004c44e0, write_hint=write_hint@entry=WRITE_LIFE_NOT_SET, wbc=0x0 <fixed_percpu_data>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3034
#2 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#3 block_read_full_page (page=<optimized out>, get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#4 0xffffffff81123523 in blkdev_readpage (file=<optimized out>, page=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:631
#5 0xffffffff810a3695 in generic_file_buffered_read_readpage (iocb=iocb@entry=0xffffc9000019fe98, filp=filp@entry=0xffff88810484ad00, page=page@entry=0xffffea00038f87c8,
 mapping=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2206
#6 0xffffffff810a5bad in generic_file_buffered_read_pagenotuptodate (count=<optimized out>, pos=0, page=0xffffea00038f87c8, iter=0xffffc9000019fe70, filp=0xffff88810484ad00,
 iocb=0xffffc9000019fe98) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2307
#7 generic_file_buffered_read_get_pages (iocb=iocb@entry=0xffffc9000019fe98, iter=iter@entry=0xffffc9000019fe70, pages=pages@entry=0xffffc9000019fd58, nr=nr@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2405
#8 0xffffffff810a60cf in generic_file_buffered_read (iocb=iocb@entry=0xffffc9000019fe98, iter=iter@entry=0xffffc9000019fe70, written=written@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2485
#9 0xffffffff810a648a in generic_file_read_iter (iocb=0xffffc9000019fe98, iter=0xffffc9000019fe70) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2637
#10 0xffffffff81123580 in blkdev_read_iter (iocb=<optimized out>, to=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:1714
#11 0xffffffff810edfd3 in call_read_iter (file=0xffff88810484ad00, iter=0xffffc9000019fe70, kio=0xffffc9000019fe98)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/fs.h:1895
#12 new_sync_read (filp=filp@entry=0xffff88810484ad00, buf=buf@entry=0x694bb0 "", len=<optimized out>, ppos=ppos@entry=0xffffc9000019fef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:415
#13 0xffffffff810ef2a7 in vfs_read (file=file@entry=0xffff88810484ad00, buf=buf@entry=0x694bb0 "", count=count@entry=1024, pos=pos@entry=0xffffc9000019fef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:496
#14 0xffffffff810ef7b3 in ksys_read (fd=<optimized out>, buf=0x694bb0 "", count=1024) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:634
#15 0xffffffff810ef845 in __do_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:644
#16 __se_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#17 __x64_sys_read (regs=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#18 0xffffffff812b05f2 in do_syscall_64 (nr=<optimized out>, regs=0xffffc9000019ff58) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/common.c:46
#19 0xffffffff8140007c in entry_SYSCALL_64 () at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/entry_64.S:120
```

# [Block device] page cache with buffer\_head

```
/ # dd if=/dev/loop0 of=dd.log bs=512 count=4
```

```
(gdb) bt
#0 submit_bio (bio=bio@entry=0xffff888104712000) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/bio.h:61
#1 0xffffffff8111e8ec in submit_bh_wbc (op=op@entry=0, op_flags=<optimized out>, op_flags@entry=0, bh=0xffff8881004aa680, write_hint=write_hint@entry=WRITE_LIFE_NOT_SET,
wbc=0x0 <fixed percpu data>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3054
#2 0xffffffff81121e51 in submit_bh (bh=<optimized out>, op_flags=0, op=0) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:3060
#3 block_read_full_page (page=<optimized out>, get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/buffer.c:2339
#4 0xffffffff81123523 in blkdev_readpage (file=<optimized out>, page=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:631
#5 0xffffffff810a3695 in generic_file_buffered_read_readpage (iocb=iocb@entry=0xffffc90000197e98, filp=filp@entry=0xffff88810023fc00, page=page@entry=0xffffea00038083e8,
mapping=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2206
#6 0xffffffff810a5bad in generic_file_buffered_read_pagenotuptodate (count=<optimized out>, pos=0, page=0xffffea00038083e8, iter=0xffffc90000197e70, filp=0xffff88810023fc00,
iocb=0xffffc90000197e98) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2307
#7 generic_file_buffered_read_get_pages (iocb=iocb@entry=0xffffc90000197e98, iter=iter@entry=0xffffc90000197e70, pages=pages@entry=0xffffc90000197d58, nr=nr@entry=1)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2405
#8 0xffffffff810a60cf in generic_file_buffered_read (iocb=iocb@entry=0xffffc90000197e98, iter=iter@entry=0xffffc90000197e70, written=written@entry=0)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2485
#9 0xffffffff810a648a in generic_file_read_iter (iocb=0xffffc90000197e98, iter=0xffffc90000197e70) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2637
#10 0xffffffff81123580 in blkdev_read_iter (iocb=<optimized out>, to=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:1714
#11 0xffffffff810edfd3 in call_read_iter (file=0xffff88810023fc00, iter=0xffffc90000197e70, kio=0xffffc90000197e98)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/fs.h:1895
#12 new_sync_read (filp=filp@entry=0xffff88810023fc00, buf=buf@entry=0x694bb0 "", len=<optimized out>, ppos=ppos@entry=0xffffc90000197ef8)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:415
#13 0xffffffff810ef2a7 in vfs_read (file=file@entry=0xffff88810023fc00, buf=buf@entry=0x694bb0 "", count=count@entry=512, pos=pos@entry=0xffffc90000197ef8)
at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:496
#14 0xffffffff810ef7b3 in ksys_read (fd=<optimized out>, buf=0x694bb0 "", count=512) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:634
#15 0xffffffff810ef845 in __do_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:644
#16 __se_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#17 __x64_sys_read (regs=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#18 0xffffffff812b05f2 in do_syscall_64 (nr=<optimized out>, regs=0xffffc90000197f58) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/common.c:46
#19 0xffffffff8140007c in entry_SYSCALL_64 () at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/entry_64.S:120
#20 0x0000000000000000 in ?? ()
```

# [Block device] page cache without buffer\_head

```
(gdb) bt
#0 bio_add_page (bio=0xffff888104713780, page=page@entry=0xffffea0003904ed8, len=len@entry=4096, offset=offset@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/block/bio.c:936
#1 0xffffffff811285ad in do_mpage_readpage (args=args@entry=0xffffc90000373a40) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/mpage.c:315
#2 0xffffffff81128b0b in mpage_readahead (rac=0xffffc90000373cb8, get_block=get_block@entry=0xffffffff81123130 <blkdev_get_block>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/mpage.c:391
#3 0xffffffff811234f0 in blkdev_readahead (rac=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:636
#4 0xffffffff810ab8e0 in read_pages (rac=rac@entry=0xffffc90000373cb8, pages=pages@entry=0xffffc90000373b98, skip_page=skip_page@entry=false)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:130
#5 0xffffffff810abcd3 in page_cache_ra_unbounded (ractl=ractl@entry=0xffffc90000373cb8, nr_to_read=3, lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:238
#6 0xffffffff810abd88 in do_page_cache_ra (ractl=ractl@entry=0xffffc90000373cb8, nr_to_read=<optimized out>, lookahead_size=<optimized out>)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:267
#7 0xffffffff810abf08 in ondemand_readahead (ractl=0xffffc90000373cb8, ra=0xffff888100258598, hit_readahead_marker=hit_readahead_marker@entry=false, req_size=req_size@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:549
#8 0xffffffff810ac1b0 in page_cache_sync_ra (ractl=ractl@entry=0xffffc90000373cb8, ra=ra@entry=0xffff888100258598, req_count=req_count@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/readahead.c:577
#9 0xffffffff810a57b9 in page_cache_sync_readahead (req_count=1, index=66, file=0xffff888100258500, ra=0xffff888100258598, mapping=0xffff8881004a83d8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/pagemap.h:839
#10 generic_file_buffered_read_get_pages (iocb=iocb@entry=0xffffc90000373e98, iter=iter@entry=0xffffc90000373e70, pages=pages@entry=0xffffc90000373d58, nr=nr@entry=1)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2365
#11 0xffffffff810a60cf in generic_file_buffered_read (iocb=iocb@entry=0xffffc90000373e98, iter=iter@entry=0xffffc90000373e70, written=written@entry=0)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2485
#12 0xffffffff810a648a in generic_file_read_iter (iocb=0xffffc90000373e98, iter=0xffffc90000373e70) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/mm/filemap.c:2637
#13 0xffffffff81123580 in blkdev_read_iter (iocb=<optimized out>, to=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/block_dev.c:1714
#14 0xffffffff810edfd3 in call_read_iter (file=0xffff888100258500, iter=0xffffc90000373e70, kio=0xffffc90000373e98)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/include/linux/fs.h:1895
#15 new_sync_read (filp=filp@entry=0xffff888100258500, buf=buf@entry=0x694bb0 "", len=<optimized out>, ppos=ppos@entry=0xffffc90000373ef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:415
#16 0xffffffff810ef2a7 in vfs_read (file=file@entry=0xffff888100258500, buf=buf@entry=0x694bb0 "", count=count@entry=512, pos=pos@entry=0xffffc90000373ef8)
 at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:496
#17 0xffffffff810ef7b3 in ksys_read (fd=<optimized out>, buf=0x694bb0 "", count=512) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:634
#18 0xffffffff810ef845 in __do_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:644
#19 __se_sys_read (count=<optimized out>, buf=<optimized out>, fd=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#20 __x64_sys_read (regs=<optimized out>) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/fs/read_write.c:642
#21 0xffffffff812b05f2 in do_syscall_64 (nr=<optimized out>, regs=0xffffc90000373f58) at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/common.c:46
#22 0xffffffff8140007c in entry_SYSCALL_64 () at /home/adrian/git-repo/gdb-linux-real-mode/src/linux-5.11/arch/x86/entry/entry_64.S:120
#23 0x0000000000000000 in ?? ()
```