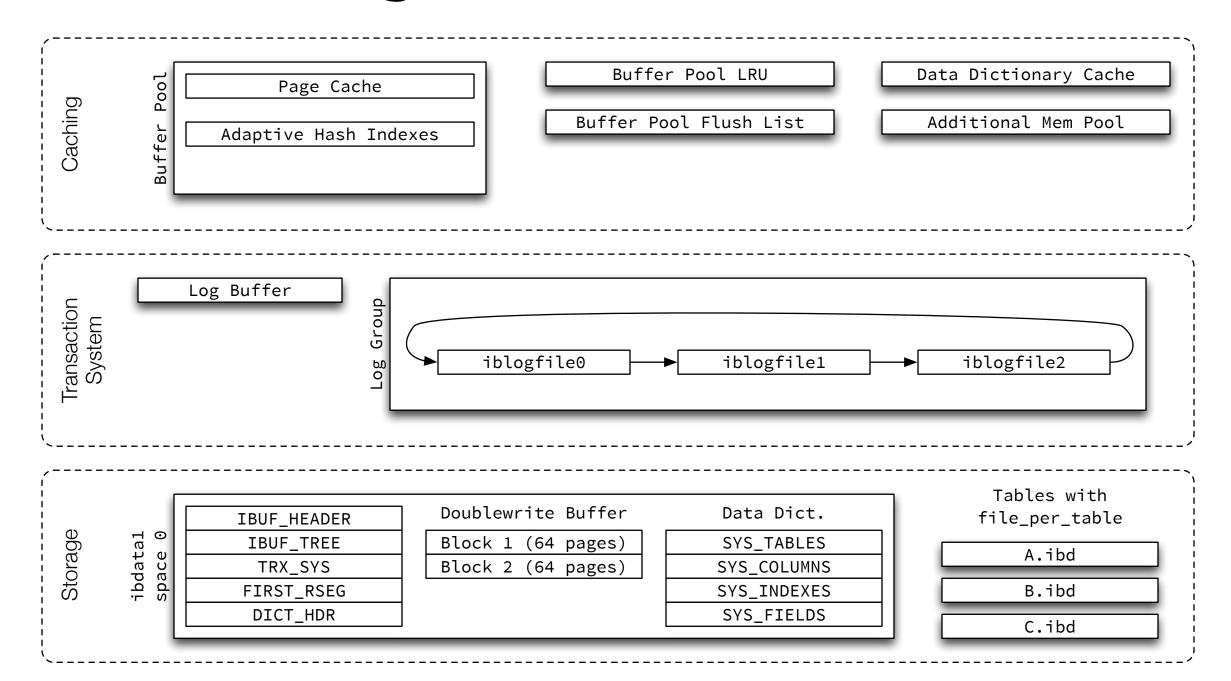
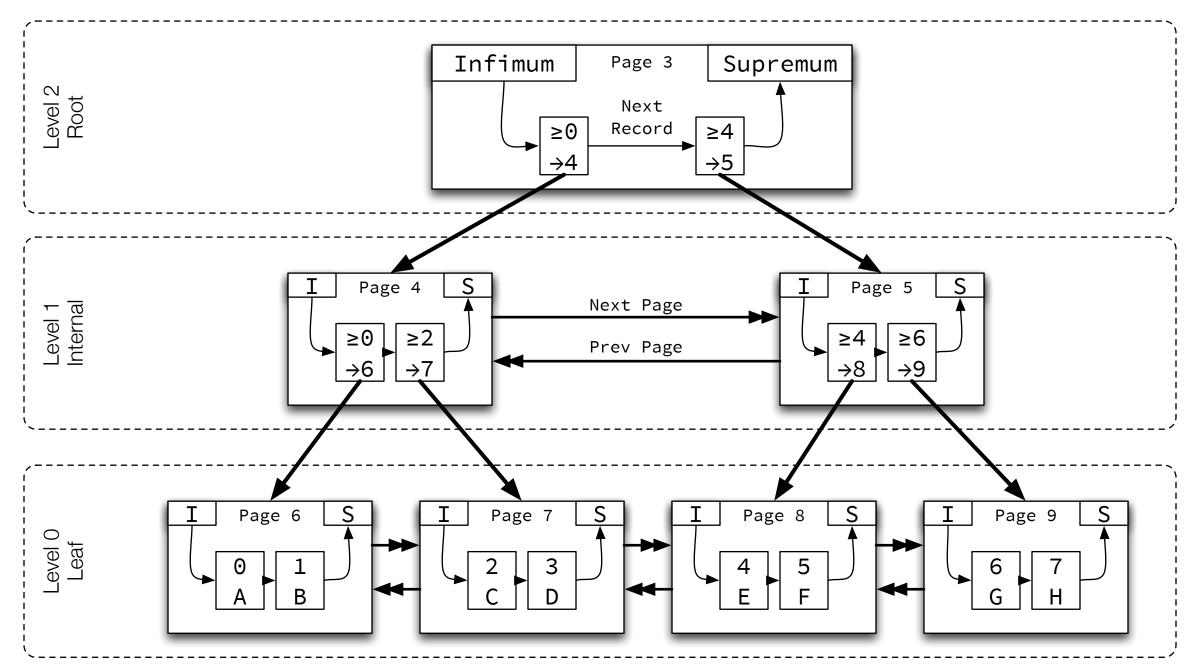
High-level Overview



B+Tree Structure



Levels are numbered starting from 0 at the leaf pages, incrementing up the tree.

Pages on each level are doubly-linked with previous and next pointers in ascending order by key.

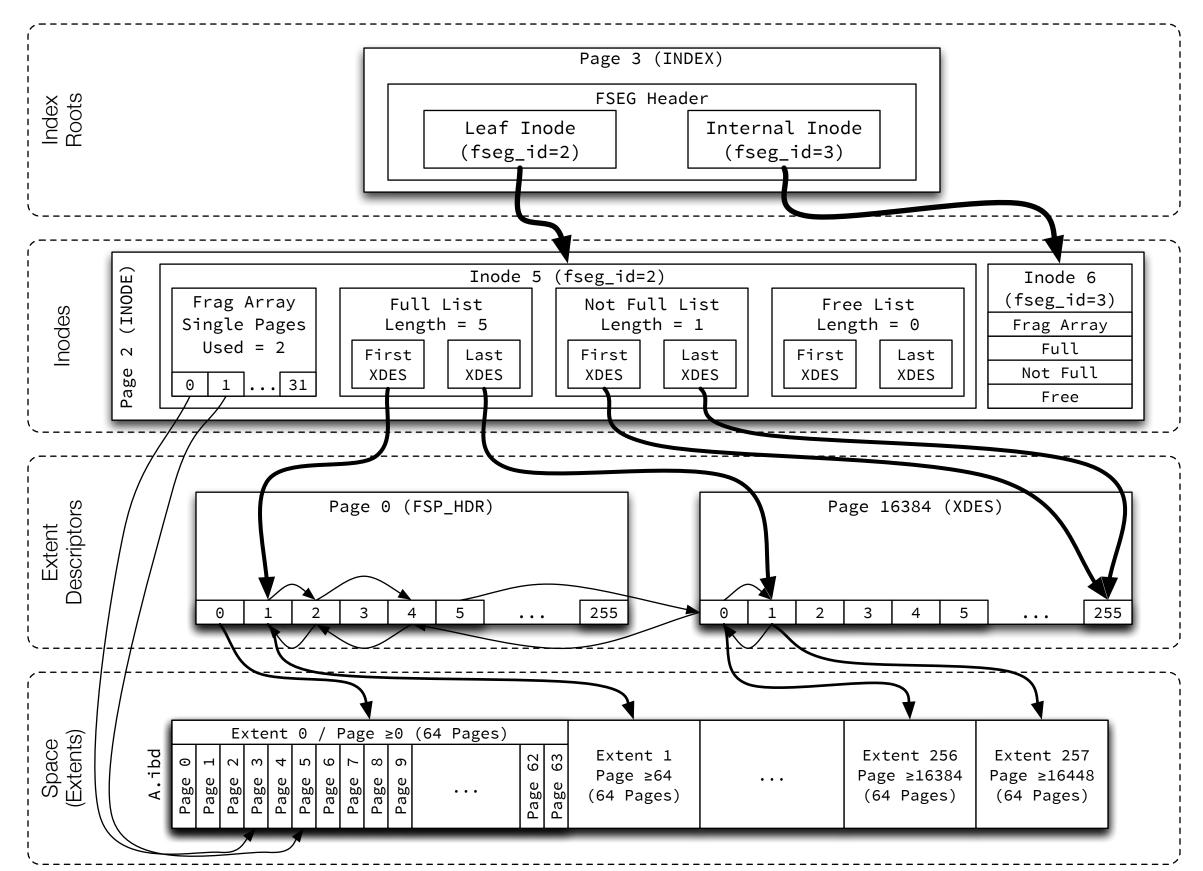
Records within a page are singly-linked with a next pointer in ascending order by key.

Infimum represents a value lower than any key on the page, and is always the first record in the singly-linked list of records.

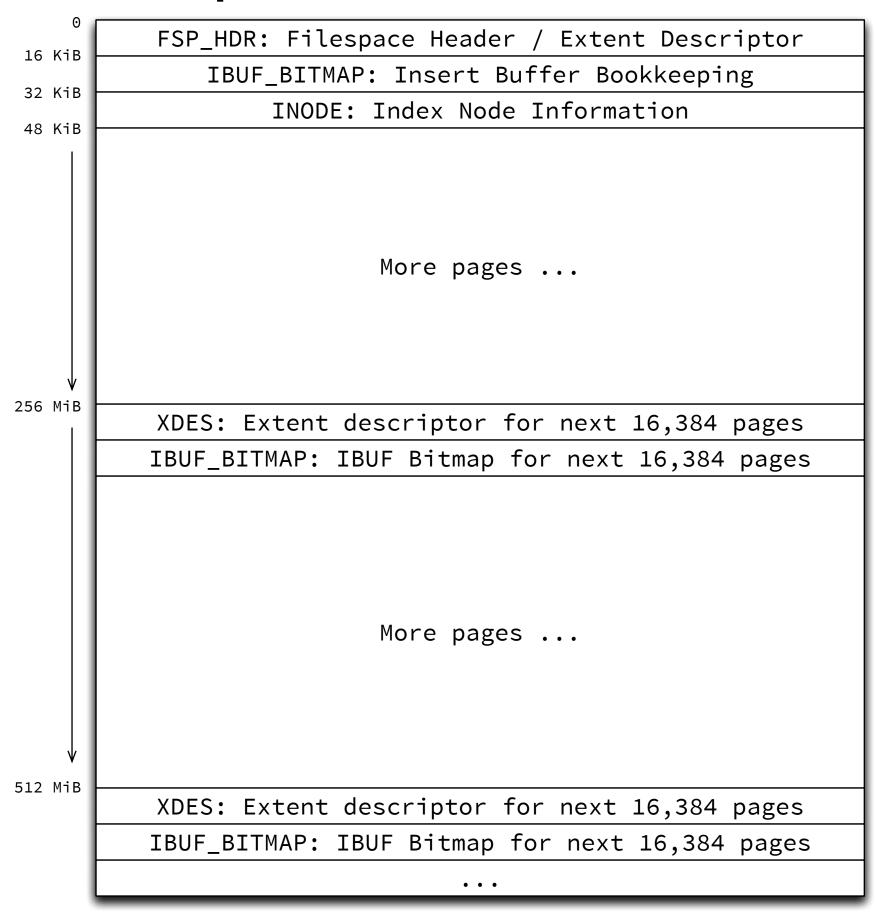
Supremum represents a value higher than any key on the page, and is always the last record in the singly-linked list of records.

Non-leaf pages contain the minimum key of the child page and the child page number, called a "node pointer".

Index File Segment Structure



Space File Overview



ibdata1 File Overview

0	
o	FSP_HDR: Filespace Header / Extent Descriptor
. KiB —	IBUF_BITMAP: Insert Buffer Bookkeeping
	INODE: Index Node Information
KiB —	SYS: Insert Buffer Header
KiB —	INDEX: Insert Buffer Root
	TRX_SYS: Transaction System Header
	SYS: First Rollback Segment
	SYS: Data Dictionary Header
ш	More pages
↓	
e 64	Double Write Buffer Block 1 (64 pages)
128	Double Write Buffer Block 2 (64 pages)
192	
- 11	
- 11	
- 11	
- 11	
- 11	
- 11	More pages
- 11	- F-6
- 11	
- 11	
- 11	
* L	

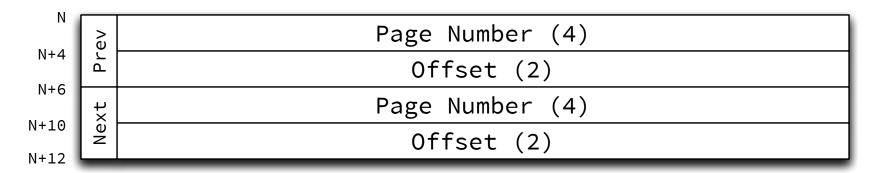
IBD File Overview

FSP_HDR: Filespace Header / Extent Descriptor 16 KiB IBUF_BITMAP: Insert Buffer Bookkeeping 32 KiB INODE: Index Node Information 48 KiB INDEX: Root page of first index 64 KiB INDEX: Root page of second index INDEX: Node pages ... INDEX: Leaf pages ... ALLOCATED: Reserved but unused pages ... More pages ...

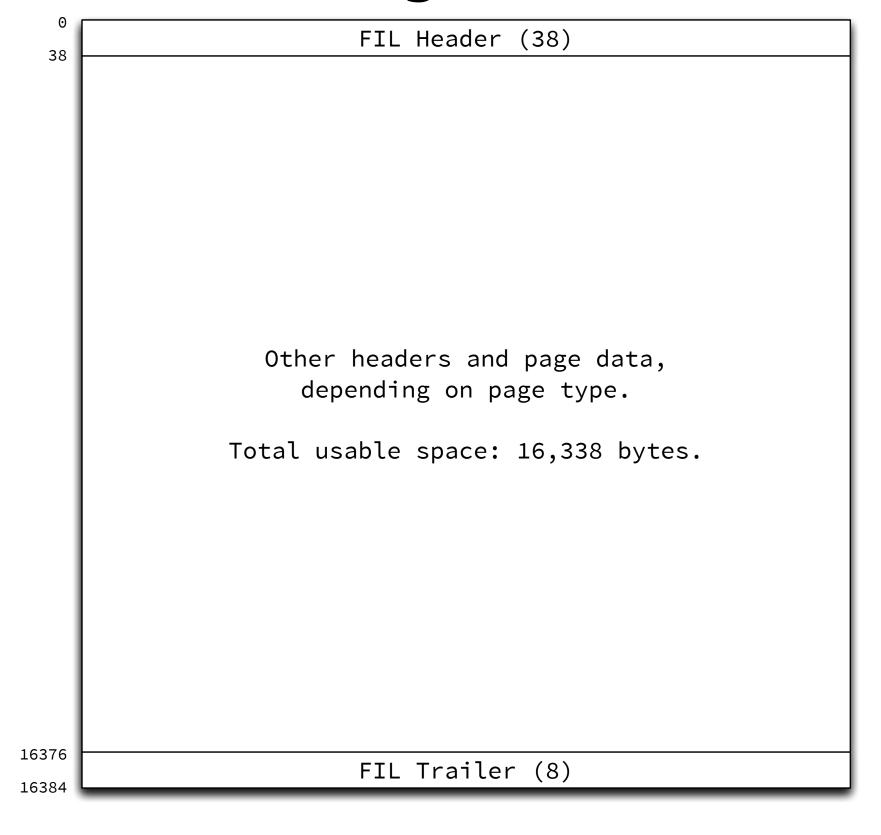
List Base Node

Nι		י
		List Length (4)
N+4	st	Page Number (4)
N+8	Fir	Offset (2)
N+10 N+14	st	Page Number (4)
N+14 N+16	La	Offset (2)

List Node



Basic Page Overview



FIL Header/Trailer

_	
0	Checksum (4)
4	Offset (Page Number) (4)
8	
12	Previous Page (4)
16	Next Page (4)
	LSN for last page modification (8)
24	Page Type (2)
26	Flush LSN (0 except space 0 page 0) (8)
34	
38	Space ID (4)
	•••
16376	Old-style Checksum (4)
16380	
16384	Low 32 bits of LSN (4)

FSP_HDR/XDES Overview

0 38	FIL Header (38)
150	FSP Header (zero-filled for XDES pages) (112)
190	XDES Entry 0 (pages 0- 63) (40)
230	XDES Entry 1 (pages 64- 127) (40)
270	XDES Entry 2 (pages 128- 191) (40)
310	XDES Entry 3 (pages 192- 255) (40)
V	
10310	XDES Entry 254 (pages 16256-16319) (40)
10350	XDES Entry 255 (pages 16320-16383) (40)
10390	
	(Empty Space, 5,986 bytes)
16376	FIL Trailer (8)
16384	TIL Haitel (6)

FSP Header

20	
38	Space ID (4)
42	· · · · · · · · · · · · · · · · · · ·
46	(Unused) (4)
- 1	Highest page number in file (size) (4)
50	Highest page number initialized (free limit) (4)
54	Flags (4)
58	Number of pages used in "FREE_FRAG" list (4)
62	List base node for "FREE" list (16)
78	List base node for "FREE_FRAG" list (16)
94	List base node for "FULL_FRAG" list (16)
110	Next Unused Segment ID (8)
118	List base node for "FULL_INODES" list (16)
134 150	List base node for "FREE_INODES" list (16)

XDES Entry

N	
111	File Segment ID (8)
N+8	List node for XDES list (12)
N+20	State (4)
N+24	Page State Bitmap (16)
N+40	2 bits per page, 1=free, 2=clean

INODE Overview

FIL Header (38)
List node for INODE Page list (12)
INODE 0 (192)
INODE 1 (192)
INODE 2 (192)
=::001 = (=0=)
•••
TNODE 93 (103)
INODE 83 (192)
INODE 84 (192)
(Empty Space, 6 bytes)
FIL Trailer (8)

INODE Entry

N	FSEG ID (8)
N+8	Number of used pages in "NOT_FULL" list (4)
N+12	List base node for "FREE" list (16)
N+28	List base node for "NOT_FULL" list (16)
N+44	List base node for "FULL" list (16)
N+60	Magic Number = 97937874 (4)
N+64	Fragment Array Entry 0 (4)
N+68	
	•••
N+188 N+192	Fragment Array Entry 31 (4)
INTIDZ	

INDEX Overview

0 5	
38	FIL Header (38)
74	INDEX Header (36)
- 1	FSEG Header (20)
94	System Records (26)
	User Records
	Records are un-ordered physically but singly-linked to each other via "next" pointers to the byte offset of the next record in ascending order.
Heap Top	Free Space
	Page Directory
	The page directory grows downwards from the FIL trailer in ascending order by key. The number of entries is stored in the INDEX header.
.6376	FIL Trailer (8)
.6384	

INDEX Header

20	
38	Number of Directory Slots (2)
40	Heap Top Position (2)
42	Number of Heap Records / Format Flag (2)
44	Free Space (2)
46	Garbage Space (2)
48	Last Insert Position (2)
50	Page Direction (2)
52	Page Direction (2)
	Number of Inserts in Page Direction (2)
54	Number of Records (2)
56	Maximum Transaction ID (0)
64	Maximum Transaction ID (8)
66	Page Level (2)
	Index ID (4)
74	` '

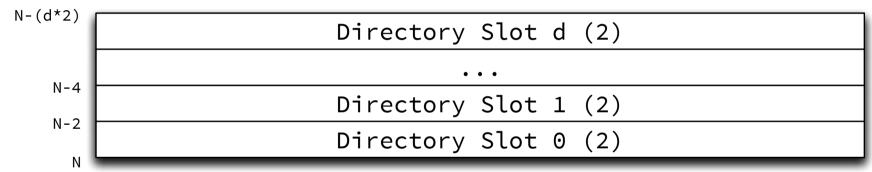
FSEG Header

7 4	
74 78	Leaf Pages List Space ID (4)
82	Leaf Pages List Page Number (4)
84	Leaf Pages List Offset (2)
	Internal (non-leaf) List Space ID (4)
88	Internal (non-leaf) List Page Number (4)
92	Internal (non-leaf) List Offset (2)
94	· · · · · · · · · · · · · · · · · · ·

INDEX System Records

0.4	
94	Info Flags (4 bits)
0.5	Number of Records Owned (4 bits)
95	Order (13 bits)
	Record Type (3 bits)
97	Next Record Offset (2)
99	"infimum\0" (8)
107	Info Flags (4 bits)
108	Number of Records Owned (4 bits)
100	Order (13 bits)
110	Record Type (3 bits)
112	Next Record Offset (2)
120	"supremum" (8)

INDEX Page Directory



Clustered Key Record Format Leaf Pages

, , , [Variable field lengths (1-2 bytes per var. field)
N-5	Info Flags (4 bits)
N-4	Number of Records Owned (4 bits)
	Order (13 bits)
N-2	Record Type (3 bits)
N L	Next Record Offset (2)
N+k	Cluster Key Fields (k)
N+k+6	Transaction ID (6)
N+k+13	Roll Pointer (7)
N+k+13+j	Non-Key Fields (j)

Clustered Key Record Format Node Pages

N 5	Variable field lengths (1-2 bytes per var. field)
N-5	Info Flags (4 bits)
N-4	Number of Records Owned (4 bits)
	Order (13 bits)
N-2	Record Type (3 bits)
N	Next Record Offset (2)
N+k	Cluster Key Min. Key on Child Page (k)
N+k+4	Child Page Number (4)

Secondary Key Record Format Leaf Pages

	Variable field lengths (1-2 bytes per var. field)
N =	Nullable field bitmap (1 bit per nullable field)
N-5	Info Flags (4 bits)
N-4	Number of Records Owned (4 bits)
	Order (13 bits)
N-2	Record Type (3 bits)
N	Next Record Offset (2)
N+k	Secondary Key Fields (k)
N+k+i	Cluster Key Fields (j)

Secondary Key Record Format Node Pages

	Variable field lengths (1-2 bytes per var. field)
N E	Nullable field bitmap (1 bit per nullable field)
N-5	Info Flags (4 bits)
N-4	Number of Records Owned (4 bits)
	Order (13 bits)
N-2	Record Type (3 bits)
N	Next Record Offset (2)
N+k	Secondary Key Min. Key on Child Page (k)
N+k+j	Cluster Key Fields (j)
N+k+j+4	Child Page Number (4)