NOTES ON THE STRUCTURE OF THE VFAT FILESYSTEM

By Galen C. Hunt

(This documentation was provided by Galen C. Hunt gchunt@cs.rochester.edu and lightly annotated by Gordon Chaffee).

This document presents a very rough, technical overview of my knowledge of the extended FAT file system used in Windows NT 3.5 and Windows 95. I don't guarantee that any of the following is correct, but it appears to be so.

The extended FAT file system is almost identical to the FAT file system used in DOS versions up to and including 6.223410239847 :-). The significant change has been the addition of long file names. Theses names support up to 255 characters including spaces and lower case characters as opposed to the traditional 8.3 short names.

Here is the description of the traditional FAT entry in the current Windows 95 filesystem:

```
struct directory { // Short 8.3 names
        unsigned char name[8];          // file name
        unsigned char ext[3];           // file extension
        unsigned char attr;             // attribute byte
        unsigned char lcase;            // Case for base and extension
        unsigned char ctime_ms;         // Creation time, milliseconds
        unsigned char ctime[2];         // Creation time
        unsigned char cdate[2];         // Creation date
        unsigned char adate[2];         // Last access date
        unsigned char reserved[2];      // reserved values (ignored)
        unsigned char time[2];          // time stamp
        unsigned char date[2];          // date stamp
        unsigned char start[2];         // starting cluster number
        unsigned char size[4];          // size of the file
};
```

The lcase field specifies if the base and/or the extension of an 8.3 name should be capitalized. This field does not seem to be used by Windows 95 but it is used by Windows NT. The case of filenames is not completely compatible from Windows NT to Windows 95. It is not completely compatible in the reverse direction, however. Filenames that fit in the 8.3 namespace and are written on Windows NT to be lowercase will show up as uppercase on Windows 95.

Note that the "start" and "size" values are actually little endian integer values. The descriptions of the fields in this structure are public knowledge and can be found elsewhere.

With the extended FAT system, Microsoft has inserted extra directory entries for any files with extended names. (Any name which legally fits within the old 8.3 encoding scheme does not have extra entries.) I call these extra entries slots. Basically, a slot is a specially formatted directory entry which holds up to 13 characters of a files extended name. Think of slots as additional labeling for the directory entry of the file to which they correspond. Microsoft prefers to refer to the 8.3 entry for a file as its alias and the extended slot directory entries as the file name.

The C structure for a slot directory entry follows:

```
struct slot { // Up to 13 characters of a long name
```

```
    unsigned char id;              // sequence number for slot
    unsigned char name0_4[10];     // first 5 characters in name
    unsigned char attr;            // attribute byte
    unsigned char reserved;        // always 0
    unsigned char alias_checksum;  // checksum for 8.3 alias
    unsigned char name5_10[12];    // 6 more characters in name
    unsigned char start[2];        // starting cluster number
    unsigned char name11_12[4];    // last 2 characters in name
};
```

If the layout of the slots looks a little odd, it's only because of Microsoft's efforts to maintain compatibility with old software. The slots must be disguised to prevent old software from panicing. To this end, a number of measures are taken:

1. The attribute byte for a slot directory entry is always set to 0x0f. This corresponds to an old directory entry with attributes of "hidden", "system", "read-only", and "volume label". Most old software will ignore any directory entries with the "volume label" bit set. Real volume label entries don't have the other three bits set.

2. The starting cluster is always set to 0, an impossible value for a DOS file.

Because the extended FAT system is backward compatible, it is possible for old software to modify directory entries. Measures must be taken to insure the validity of slots. An extended FAT system can verify that a slot does in fact belong to an 8.3 directory entry by the following:

1. Positioning. Slots for a file always immediately proceed their corresponding 8.3 directory entry. In addition, each slot has an id which marks its order in the extended file name. Here is a very abbreviated view of an 8.3 directory entry and its corresponding long name slots for the file "My Big File.Extension which is long":

```
<proceeding files...>
<slot #3, id = 0x43, characters = "h is long">
<slot #2, id = 0x02, characters = "xtension whic">
<slot #1, id = 0x01, characters = "My Big File.E">
<directory entry, name = "MYBIGFIL.EXT">
```

   Note that the slots are stored from last to first. Slots are numbered from 1 to N. The Nth slot is or'ed with 0x40 to mark it as the last one.

2. Checksum. Each slot has an "alias_checksum" value. The checksum is calculated from the 8.3 name using the following algorithm:

```
for (sum = i = 0; i < 11; i++) {
    sum = (((sum&1)<<7)|((sum&0xfe)>>1)) + name[i]
}
```

3. If there is in the final slot, a Unicode NULL (0x0000) is stored after the final character. After that, all unused characters in the final slot are set to Unicode 0xFFFF.

Finally, note that the extended name is stored in Unicode. Each Unicode character takes two bytes.

NOTES ON UNICODE TRANSLATION IN VFAT FILESYSTEM

(Information provided by Steve Searle <steve@mgu.bath.ac.uk>)

Char used as filename	Char(s) used in shortname	Char(s) used in longname slot	Entries which have been corrected
0x80 (128)	0x80	0xC7	
0x81 (129)	0x9A	0xFC	
0x82 (130)	0x90	0xE9	E
0x83 (131)	0xB6	0xE2	E
0x84 (132)	0x8E	0xE4	E
0x85 (133)	0xB7	0xE0	E
0x86 (134)	0x8F	0xE5	E
0x87 (135)	0x80	0xE7	E
0x88 (136)	0xD2	0xEA	E
0x89 (137)	0xD3	0xEB	E
0x8A (138)	0xD4	0xE8	E
0x8B (139)	0xD8	0xEF	E
0x8C (140)	0xD7	0xEE	E
0x8D (141)	0xDE	0xEC	E
0x8E (142)	0x8E	0xC4	E
0x8F (143)	0x8F	0xC5	E
0x90 (144)	0x90	0xC9	E
0x91 (145)	0x92	0xE6	E
0x92 (146)	0x92	0xC6	E
0x93 (147)	0xE2	0xF4	E
0x94 (148)	0x99	0xF6	
0x95 (149)	0xE3	0xF2	
0x96 (150)	0xEA	0xFB	
0x97 (151)	0xEB	0xF9	
0x98 (152)	"_~1"	0xFF	
0x99 (153)	0x99	0xD6	
0x9A (154)	0x9A	0xDC	
0x9B (155)	0x9D	0xF8	
0x9C (156)	0x9C	0xA3	
0x9D (157)	0x9D	0xD8	
0x9E (158)	0x9E	0xD7	
0x9F (159)	0x9F	0x92	
0xA0 (160)	0xB5	0xE1	
0xA1 (161)	0xD6	0xE0	
0xA2 (162)	0xE0	0xF3	
0xA3 (163)	0xE9	0xFA	
0xA4 (164)	0xA5	0xF1	
0xA5 (165)	0xA5	0xD1	
0xA6 (166)	0xA6	0xAA	
0xA7 (167)	0xA7	0xBA	
0xA8 (168)	0xA8	0xBF	
0xA9 (169)	0xA9	0xAE	
0xAA (170)	0xAA	0xAC	
0xAB (171)	0xAB	0xBD	
0xAC (172)	0xAC	0xBC	

0xAD (173)	0xAD	0xA1	
0xAE (174)	0xAE	0xAB	
0xAF (175)	0xAF	0xBB	
0xB0 (176)	0xB0	0x91 0x25	
0xB1 (177)	0xB1	0x92 0x25	
0xB2 (178)	0xB2	0x93 0x25	
0xB3 (179)	0xB3	0x02 0x25	
0xB4 (180)	0xB4	0x24 0x25	
0xB5 (181)	0xB5	0xC1	
0xB6 (182)	0xB6	0xC2	
0xB7 (183)	0xB7	0xC0	
0xB8 (184)	0xB8	0xA9	
0xB9 (185)	0xB9	0x63 0x25	
0xBA (186)	0xBA	0x51 0x25	
0xBB (187)	0xBB	0x57 0x25	
0xBC (188)	0xBC	0x5D 0x25	
0xBD (189)	0xBD	0xA2	
0xBE (190)	0xBE	0xA5	
0xBF (191)	0xBF	0x10 0x25	
0xC0 (192)	0xC0	0x14 0x25	
0xC1 (193)	0xC1	0x34 0x25	
0xC2 (194)	0xC2	0x2C 0x25	
0xC3 (195)	0xC3	0x1C 0x25	
0xC4 (196)	0xC4	0x00 0x25	
0xC5 (197)	0xC5	0x3C 0x25	
0xC6 (198)	0xC7	0xE3	E
0xC7 (199)	0xC7	0xC3	
0xC8 (200)	0xC8	0x5A 0x25	E
0xC9 (201)	0xC9	0x54 0x25	E
0xCA (202)	0xCA	0x69 0x25	E
0xCB (203)	0xCB	0x66 0x25	E
0xCC (204)	0xCC	0x60 0x25	E
0xCD (205)	0xCD	0x50 0x25	E
0xCE (206)	0xCE	0x6C 0x25	E
0xCF (207)	0xCF	0xA4	E
0xD0 (208)	0xD1	0xF0	
0xD1 (209)	0xD1	0xD0	
0xD2 (210)	0xD2	0xCA	
0xD3 (211)	0xD3	0xCB	
0xD4 (212)	0xD4	0xC8	
0xD5 (213)	0x49	0x31 0x01	
0xD6 (214)	0xD6	0xCD	
0xD7 (215)	0xD7	0xCE	
0xD8 (216)	0xD8	0xCF	
0xD9 (217)	0xD9	0x18 0x25	
0xDA (218)	0xDA	0x0C 0x25	
0xDB (219)	0xDB	0x88 0x25	
0xDC (220)	0xDC	0x84 0x25	
0xDD (221)	0xDD	0xA6	
0xDE (222)	0xDE	0xCC	
0xDF (223)	0xDF	0x80 0x25	
0xE0 (224)	0xE0	0xD3	
0xE1 (225)	0xE1	0xDF	
0xE2 (226)	0xE2	0xD4	
0xE3 (227)	0xE3	0xD2	

0xE4 (228)	0x05	0xF5
0xE5 (229)	0x05	0xD5
0xE6 (230)	0xE6	0xB5
0xE7 (231)	0xE8	0xFE
0xE8 (232)	0xE8	0xDE
0xE9 (233)	0xE9	0xDA
0xEA (234)	0xEA	0xDB
0xEB (235)	0xEB	0xD9
0xEC (236)	0xED	0xFD
0xED (237)	0xED	0xDD
0xEE (238)	0xEE	0xAF
0xEF (239)	0xEF	0xB4
0xF0 (240)	0xF0	0xAD
0xF1 (241)	0xF1	0xB1
0xF2 (242)	0xF2	0x17 0x20
0xF3 (243)	0xF3	0xBE
0xF4 (244)	0xF4	0xB6
0xF5 (245)	0xF5	0xA7
0xF6 (246)	0xF6	0xF7
0xF7 (247)	0xF7	0xB8
0xF8 (248)	0xF8	0xB0
0xF9 (249)	0xF9	0xA8
0xFA (250)	0xFA	0xB7
0xFB (251)	0xFB	0xB9
0xFC (252)	0xFC	0xB3
0xFD (253)	0xFD	0xB2
0xFE (254)	0xFE	0xA0 0x25
0xFF (255)	0xFF	0xA0

Page 0

0x80 (128)	0x00
0x81 (129)	0x00
0x82 (130)	0x00
0x83 (131)	0x00
0x84 (132)	0x00
0x85 (133)	0x00
0x86 (134)	0x00
0x87 (135)	0x00
0x88 (136)	0x00
0x89 (137)	0x00
0x8A (138)	0x00
0x8B (139)	0x00
0x8C (140)	0x00
0x8D (141)	0x00
0x8E (142)	0x00
0x8F (143)	0x00
0x90 (144)	0x00
0x91 (145)	0x00
0x92 (146)	0x00
0x93 (147)	0x00
0x94 (148)	0x00
0x95 (149)	0x00
0x96 (150)	0x00
0x97 (151)	0x00

0x98 (152) 0x00
0x99 (153) 0x00
0x9A (154) 0x00
0x9B (155) 0x00
0x9C (156) 0x00
0x9D (157) 0x00
0x9E (158) 0x00
0x9F (159) 0x92
0xA0 (160) 0xFF
0xA1 (161) 0xAD
0xA2 (162) 0xBD
0xA3 (163) 0x9C
0xA4 (164) 0xCF
0xA5 (165) 0xBE
0xA6 (166) 0xDD
0xA7 (167) 0xF5
0xA8 (168) 0xF9
0xA9 (169) 0xB8
0xAA (170) 0x00
0xAB (171) 0xAE
0xAC (172) 0xAA
0xAD (173) 0xF0
0xAE (174) 0x00
0xAF (175) 0xEE
0xB0 (176) 0xF8
0xB1 (177) 0xF1
0xB2 (178) 0xFD
0xB3 (179) 0xFC
0xB4 (180) 0xEF
0xB5 (181) 0xE6
0xB6 (182) 0xF4
0xB7 (183) 0xFA
0xB8 (184) 0xF7
0xB9 (185) 0xFB
0xBA (186) 0x00
0xBB (187) 0xAF
0xBC (188) 0xAC
0xBD (189) 0xAB
0xBE (190) 0xF3
0xBF (191) 0x00
0xC0 (192) 0xB7
0xC1 (193) 0xB5
0xC2 (194) 0xB6
0xC3 (195) 0xC7
0xC4 (196) 0x8E
0xC5 (197) 0x8F
0xC6 (198) 0x92
0xC7 (199) 0x80
0xC8 (200) 0xD4
0xC9 (201) 0x90
0xCA (202) 0xD2
0xCB (203) 0xD3
0xCC (204) 0xDE
0xCD (205) 0xD6
0xCE (206) 0xD7

0xCF (207)	0xD8
0xD0 (208)	0x00
0xD1 (209)	0xA5
0xD2 (210)	0xE3
0xD3 (211)	0xE0
0xD4 (212)	0xE2
0xD5 (213)	0xE5
0xD6 (214)	0x99
0xD7 (215)	0x9E
0xD8 (216)	0x9D
0xD9 (217)	0xEB
0xDA (218)	0xE9
0xDB (219)	0xEA
0xDC (220)	0x9A
0xDD (221)	0xED
0xDE (222)	0xE8
0xDF (223)	0xE1
0xE0 (224)	0x85, 0xA1
0xE1 (225)	0xA0
0xE2 (226)	0x83
0xE3 (227)	0xC6
0xE4 (228)	0x84
0xE5 (229)	0x86
0xE6 (230)	0x91
0xE7 (231)	0x87
0xE8 (232)	0x8A
0xE9 (233)	0x82
0xEA (234)	0x88
0xEB (235)	0x89
0xEC (236)	0x8D
0xED (237)	0x00
0xEE (238)	0x8C
0xEF (239)	0x8B
0xF0 (240)	0xD0
0xF1 (241)	0xA4
0xF2 (242)	0x95
0xF3 (243)	0xA2
0xF4 (244)	0x93
0xF5 (245)	0xE4
0xF6 (246)	0x94
0xF7 (247)	0xF6
0xF8 (248)	0x9B
0xF9 (249)	0x97
0xFA (250)	0xA3
0xFB (251)	0x96
0xFC (252)	0x81
0xFD (253)	0xEC
0xFE (254)	0xE7
0xFF (255)	0x98