

Sailing in Uncharted Waters

A story of implementing Apple extensions for SMB2 in EMC Isilon clustered storage

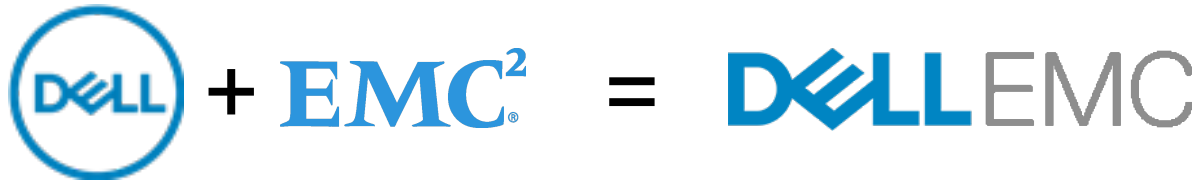
Rafal Szczesniak

Sumaithri Mukkamalla

EMC Isilon

Apologies (from Rafal)

First, I'd like to apologise for not being here in person. Here's what has happened:



The merger has closed soon before the conference and caused a severe disruption in my travel plans. Many thanks to Sumaithri for helping to carry the torch!

Agenda

- ❑ Apple OSX and SMB
- ❑ Apple Extensions via Create Context
- ❑ ReadDirAttr
- ❑ OsxCopyFile
- ❑ Credits – observations

Apple OSX and SMB

- ❑ Apple started shipping their client and server even back in the days of SMB1
- ❑ No changes to the protocol itself though
- ❑ After SMB2 saw the light of day Apple started shipping their own proprietary implementation
- ❑ A few features optimising the protocol for OSX platform have been introduced as part of that effort

Why is it so important now?

- ❑ Starting from OSX 10.9 (*Mavericks*) SMB2 has become the preferred file sharing protocol
- ❑ AFP (Apple File Protocol) is being deprecated
- ❑ Media & Entertainment industry (primarily) is a heavyweight user of OSX, so they need all the features it provides

What is different in SMB2 traffic for OSX?

- ❑ Apple-specific create context
- ❑ Compound requests (multiple requests in one SMB2 packet)
- ❑ Metadata saved in Alternate Data Streams (ADS)
- ❑ Supported OSX features advertised via the create context

What is SMB2 Create Context?

- ❑ A part of a variable-length chain of key/value pairs sent in SMB2 request opening a file
- ❑ It can also be sent in the response
- ❑ The key is a string describing what the value is
- ❑ The value can be basically anything pre-defined

What is SMB2 Create Context?

| | | | | | | |
|-----|-----------|-------------|-------------|------|-----|------------------------|
| 829 | 92.834675 | 192.168.2.6 | 10.7.223.31 | SMB2 | 454 | Create Request File: t |
| 830 | 92.835379 | 192.168.2.6 | 10.7.223.31 | SMB2 | 262 | Create Request File: t |
| 831 | 92.858835 | 10.7.223.31 | 192.168.2.6 | SMB2 | 578 | Create Response File: |
| 833 | 92.859053 | 10.7.223.31 | 192.168.2.6 | SMB2 | 223 | Create Response Error |


```
Allocation Size: 4096
End Of File: 533
> File Attributes: 0x00000020
> GUID handle File: test dir 01\file01.html
▼ ExtraInfo SMB2_CREATE_QUERY_MAXIMAL_ACCESS_REQUEST
  Offset: 0x00000098
  Length: 32
  ▼ Chain Element: SMB2_CREATE_QUERY_MAXIMAL_ACCESS_REQUEST "MxAc"
    Chain Offset: 0x00000000
    > Tag: MxAc
    ▼ Data: MxAc INFO
      Offset: 0x00000018
      Length: 8
      > MxAc INFO
> SMB2 (Server Message Block Protocol version 2)
> SMB2 (Server Message Block Protocol version 2)
```

| | | | | |
|------|-------------------------|-------------------|-------|-----------|
| 00d0 | 00 00 00 00 00 00 98 00 | 00 00 20 00 00 00 | 00 00 | |
| 00e0 | 00 00 10 00 04 00 00 00 | 18 00 08 00 00 00 | 4d 78 |Mx |
| 00f0 | 41 63 00 4f d1 01 00 00 | 00 00 9f 01 1f 00 | fe 53 | Ac.0....S |

AAPL Create Context

- ❑ “AAPL” context enables the first step towards OSX compatibility
- ❑ The first step is AAPL ServerQuery context
- ❑ Request – a wish list requested by the client
- ❑ Response – server responds “this is what I support”

AAPL Create Context (contd.)

REQUEST

RESPONSE

SMB2_APPLE_SERVER_QUERY

SMB2_APPLE_SERVER_QUERY

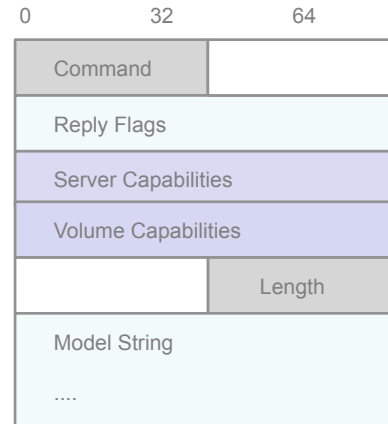
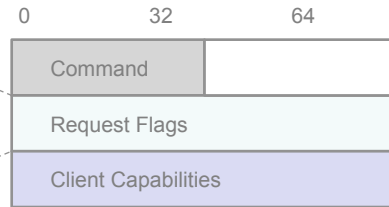
SMB2_APPLE_SERVER_CAPS
SMB2_APPLE_VOLUME_CAPS
SMB2_APPLE_MODEL_INFO

SMB2_APPLE_SERVER_CAPS
SMB2_APPLE_VOLUME_CAPS
SMB2_APPLE_MODEL_INFO

SMB2_APPLE_CAPS_READ_DIR_ATTR
SMB2_APPLE_CAPS_OSX_COPYFILE
SMB2_APPLE_CAPS_UNIX_BASED
SMB2_APPLE_CAPS_NFS_ACE

SMB2_APPLE_CAPS_READ_DIR_ATTR
SMB2_APPLE_CAPS_OSX_COPYFILE
SMB2_APPLE_CAPS_UNIX_BASED
SMB2_APPLE_CAPS_NFS_ACE

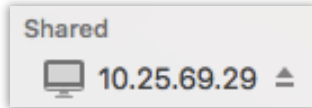
"OneFS"



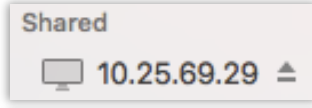
AAPL Create Context (contd.)

- ❑ What does Isilon support?
 - ❑ ReadDirAttr extension
 - ❑ OsxCopyFile extension
 - ❑ Case sensitive name lookups
 - ❑ We are unix-based
- ❑ What's interesting, Model String can affect the way the server is shown in the Apple Finder window

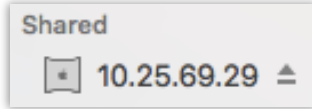
Whom would you like to be today?



iMac

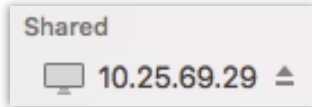


MacBook



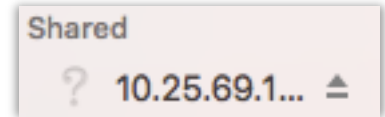
MacPro

or....



MacOneFS

We have to be “MacSomething” because the alternative is:



AAPL Create Context (contd.)

| | | | | | |
|-----|-----------|-------------|-------------|------|----------------------|
| 137 | 29.008685 | 10.7.223.31 | 192.168.2.6 | SMB2 | 650 Create Response |
| 139 | 29.009001 | 192.168.2.6 | 10.7.223.31 | SMB2 | 414 Create Request F |

```
Allocation Size: 0
End Of File: 0
> File Attributes: 0x00000010
> GUID handle File:
v ExtraInfo <invalid> SMB2_CREATE_QUERY_MAXIMAL_ACCESS_REQUEST
  Offset: 0x00000098
  Length: 120
  v Chain Element: <invalid> "AAPL"
    Chain Offset: 0x00000058
    v Tag: AAPL
      Offset: 0x00000010
      Length: 4
    v Data
      Offset: 0x00000018
      Length: 64
  v Chain Element: SMB2_CREATE_QUERY_MAXIMAL_ACCESS_REQUEST "MxAc"
    Chain Offset: 0x00000000
```

| | | | |
|------|-------------------------|-------------------------|-------------------|
| 00d0 | 00 00 00 00 00 00 98 00 | 00 00 78 00 00 00 58 00 |x...X. |
| 00e0 | 00 00 10 00 04 00 00 00 | 18 00 40 00 00 00 41 41 |@...AA |
| 00f0 | 50 4c 00 00 00 00 01 00 | 00 00 00 00 00 00 07 00 | PL..... |
| 0100 | 00 00 00 00 00 00 0f 00 | 00 00 00 00 00 00 01 00 | |
| 0110 | 00 00 00 00 00 00 00 00 | 00 00 12 00 00 00 4d 00 |M. |
| 0120 | 61 00 63 00 50 00 72 00 | 6f 00 36 00 2c 00 31 00 | a.c.P.r. o.6.,.1. |
| 0130 | 00 00 39 38 45 32 00 00 | 00 00 10 00 04 00 00 00 | ..98E2.. |

ReadDirAttr extension

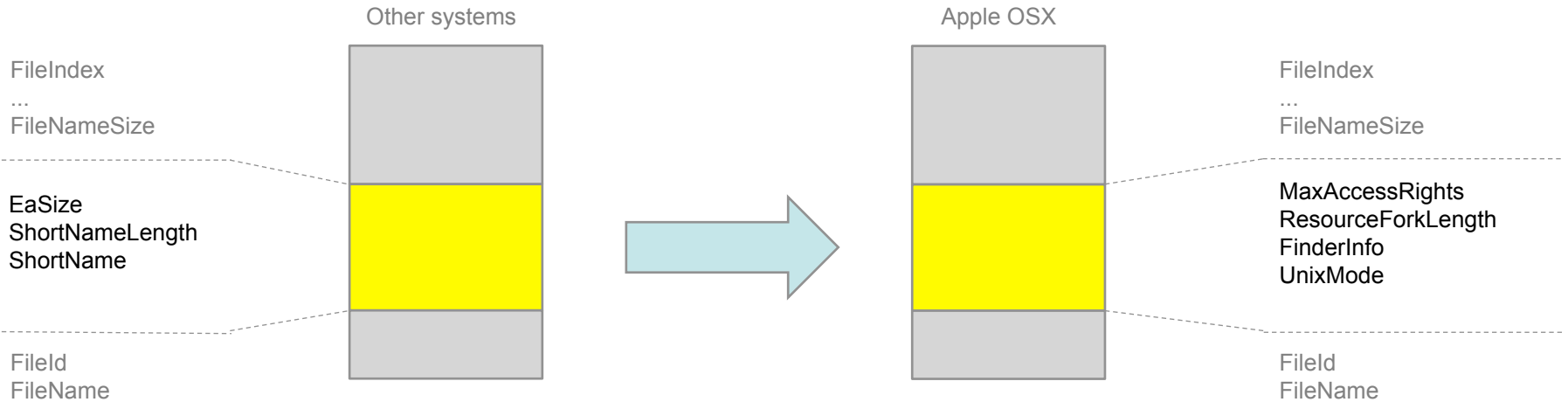
- ❑ Comes into play when listing directories with SMB2 Find request
- ❑ The request and response utilise various levels of information about individual entries (like files) in the directory

ReadDirAttr extension (contd.)

- ❑ FileIdBothDirInformation level has been modified to accommodate OSX-specific data:
 - ❑ Finder Info
 - ❑ Resource Fork Length
 - ❑ Access rights, Unix mode

ReadDirAttr extension (contd.)

FileIdBothDirInformation



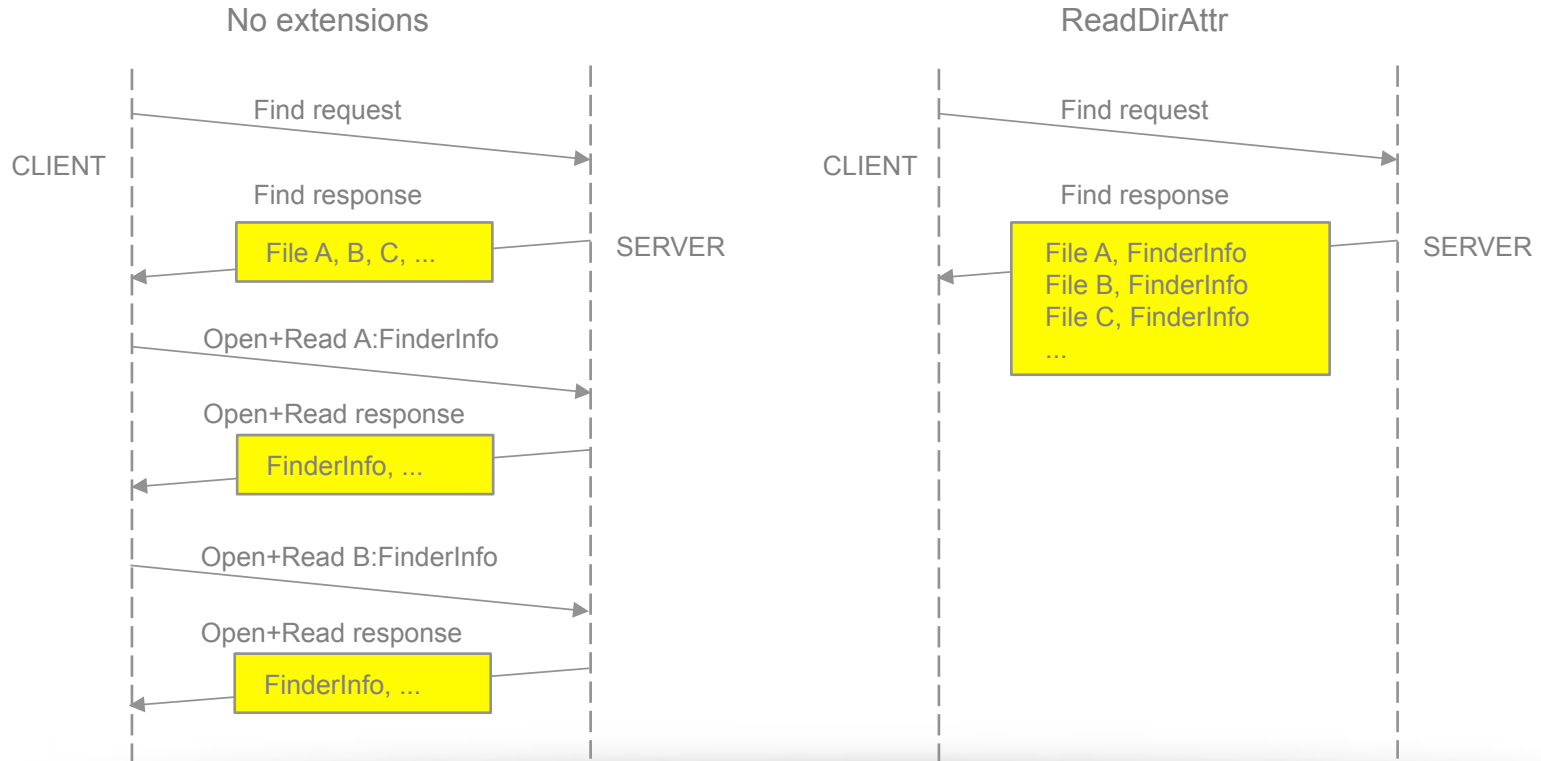
ReadDirAttr exceptions

- ❑ Symbolic links is where we're no longer acting special
- ❑ EaSize (or MaxAccessRights) is set to `IO_REPARSE_TAG_SYMLINK (0xA000000C)`
- ❑ Client is expected to get the link information (FsCtl)

Benefits (ReadDirAttr)

- ❑ The basic directory listing takes one request and just one (potentially large) response
- ❑ Previously, it would take many round trips to read all the attributes before anything was shown in the client's Finder window
- ❑ “Basic” listing doesn't show everything yet (like icons) but the user can start traversing the directory tree right away

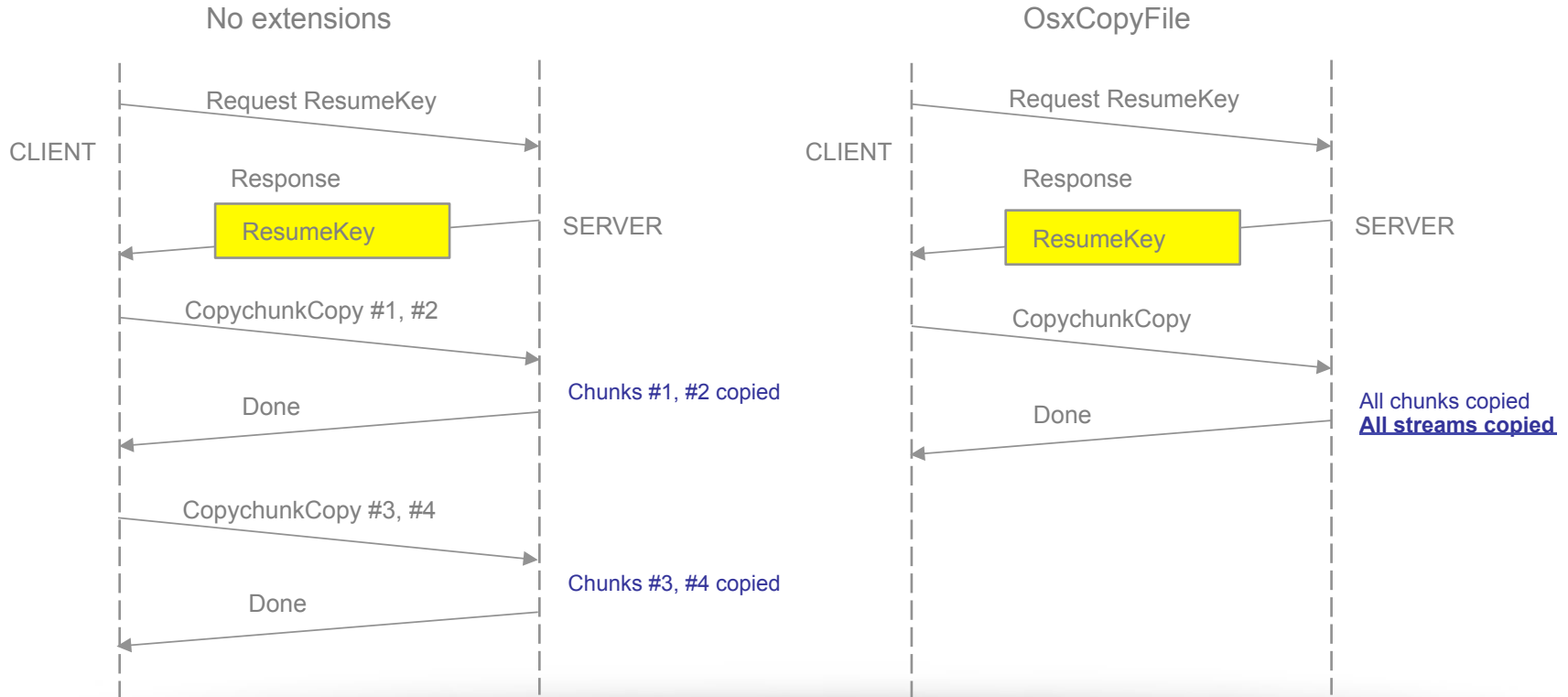
Benefits (ReadDirAttr) (contd.)



OsxCopyFile extension

- ❑ Changes the semantics of a standard Server-Side Copy operation
- ❑ File gets copied in one single shot instead of requesting multiple small chunks being copied one-by-one
- ❑ All the ADSes are copied along with the file (never forget the metadata...)

Benefits (OsxCopyFile)



Credits

- ❑ OSX client requests a lot of credits (256 per each initial Create request)
- ❑ Most servers are conservative in granting them but you can try to be more liberal (like OSX server)
- ❑ The demand seems satisfied around 65k collected
- ❑ The biggest consumers are Read and Write requests when transferring large files

Credits (contd.)

If we're liberal enough we may be sent this:

```
▶Transmission Control Protocol, Src Port: 56273 (56273), Dst Port: microsoft-ds (445), Seq: 11775, Ack: 19296, Len: 496
▶NetBIOS Session Service
▼SMB2 (Server Message Block Protocol version 2)
  ▶SMB2 Header
  ▶Read Request (0x08)
▶NetBIOS Session Service
▼SMB2 (Server Message Block Protocol version 2)
  ▶SMB2 Header
  ▶Read Request (0x08)
▶NetBIOS Session Service
▼SMB2 (Server Message Block Protocol version 2)
  ▶SMB2 Header
  ▶Read Request (0x08)
▶NetBIOS Session Service
▼SMB2 (Server Message Block Protocol version 2)
  ▶SMB2 Header
  ▶Read Request (0x08)
```

Multiple NetBios frames with one request each.

Unsupported so far

- ❑ Resolved extension:
 - ❑ The other AAPL create context command
 - ❑ Allows OSX client to reverse-resolve a unique id to the file name
 - ❑ Needs more effort put in the file system driver in order to do it right

Thank you!

Rafal Szczesniak
rafal.szczesniak@isilon.com

Sumaithri Mukkamalla
sumaithri.mukkamalla@isilon.com