Protocol for a Decentralized Mobile File System

Verson 0.1

(C)opyright May 2017 Daniel Smith

This document is the official description of the Decentralized Mobile File System (DMFS) Protocol.

Distribution of this document is unlimited on the condition that any copy produced is identical to the original.

CONTENTS

1. INTRODUCTION

The objective of the DMFS protocol is to exist as an alternative method of cloud storage. It will specifically make improvements regarding resource and security limitations that commonly occur in centralized cloud storage.

2. OVERVIEW

This section will cover information and concepts that universally pertain to the DMFS Protocol.

2.1 TERMINOLOGY

This section lists terms that are specific to a DMFS. The list only includes simple definitions because more detail is provided in the sections below.

DMFS
The acronym for Decentralized Mobile File System

Block
The required universal format for moving stored data in a DMFS.

Remote Device
Any computing device connected to a DMFS.

Routing Device
The device in a DMFS network responsible for facilitating connections between all devices in the network. A Routing Device is still a Remote Device, but it requires its own definition because of its unique role in a DMFS.

Message ID
A whole numerical value passed between Remote Devices that is used to represent specific messages.

Message Type
An organizational group that a Signal can belong to.

2.2 THE BLOCK

This section defines what a Block is. Information about how it's made and handled is included in other sections.

Any data that is to be stored in the DMFS is passed between devices in the format of a Block. A Block contains a piece of a whole file, a header that carries information about the block, and one or more layers of encryption surrounding either the piece of a file or the whole block itself.

Any Block that is not formed properly according to the parameters in this section will not be accepted by the DMFS.

2.2.1 THE BLOCK MODEL

The following diagram shows the configuration of the various sections of a Block.

****

```
                 Block
    _______________________________
   /                               \
   |                               |
   |           Header              |
   |                               |
   |-------------------------------|
   |                               |
   |                               |
   |         File Section          |
   |                               |
   |                               |
   \_______________________________/
```

Figure 1 Standard Block Configuration

Note:
Possible encryption layers are not shown in the diagram because encryption can be applied to any and all sections of a Block

Figure 1 is the standard required configuration of a Block. A Block may contain only one Header and one File Section.

****

2.2.2 THE HEADER______________________________________

The Header is a 432 bit section preceding the File section. It contains the following information in this order:

256 bit ID number for the file that the File Section originated from

32 bit ID number for the block

The 128 bit identifying number of the User the Block belongs to.

The size of the file section in bytes stored by a 16 bit number. If the number is leftat zero then the size of the file section is assumed to be 2 kilobytes.

The Header is required to be exactly 432 bits long.

2.2.3 THE FILE SECTION________________________________

The File Section contains a 256 Kilobyte long section of a file.

2.2.4 ENCRYPTION

A layer of encryption can be placed over any part of the Block, or around the complete Block.

The Remote Deivce that applies the encryption to the Block is responsible for storing the keys to the encryption.

2.3 DMFS SYSTEM MODELS

A DMFS is a peer-to-peer-like network of devices that collectively stores an amount of information. Each connected device will localy store a section of the total information that exists in the DMFS. The minimum requirements to establish a DMFS are at least one Routing Device and at least one Remote Device. There is no maximum limit to the number of devices that can be included in a DMFS.

The following diagrams show some possible configurations for a DMFS network. For the sake of simplicity, many of the processes that take place during the communications between the Remote and Routing devices have been abridged in these diagrams. The exact processes that take place during Block and Info Transfers are detailed in following sections. It's also assumed in the diagrams that the DMFS is already configured and is active.

****

Figure 2 Basic DMFS System Model

```
      ____
     /    \
     |User|
     \____/
        ^
        |
        |File System Operations
        |
        V
      DMFS
```

____

```
/               ^                                       \
|               |                                       |
|               V                                       |
|          --------------                               |
|        _>|Local Device|<______________                |
|       |  --------------               |               |
|       |                              (a)Info Transfer |
|      (b)Block Transfer                |               |
|       |                               |               |
|       |                               |               |
|       |      Network of Remote Devices|               |
|-------|-------------------------------|---------------|
|       |                               |               |
|       |                               V               |
|       |       ---------------------------             |
|       |       |      Routing Device    |             |
|       |       |                         |             |
|       |       ---------------------------             |
|       |         ^                     ^               |
|       |        (c)Info Transfer      (e)Block Transfer|
|       |         |                     |               |
|       V         V                     V               |
|       ---------------              ---------------    |
|       |Remote Device|<----(d)----->|Remote Device|    |
|       |      1      |Block Transfer|      2      |    |
|       ---------------              ---------------    |
\_______________________________________________________/
```

Figure 2 represents a simple DMFS configuration containing four devices: A Local Device operated by the User, a Routing Device, and two Remote Devices. Letters a-e in parentheses indicate instances of communication between the devices. The following scenarios give simplified contexts where these communications may occur:

I. The Local Device needs to allocate a Block

(a) It connects to the Routing Device and requests an IP address of another Remote Device in the network. The Routing Device sends the IP address for Remote Device 1.

(b) The Local Device then connects to Remote Device 1 and transfers the Block.

II. Suppose Remote Device 1 has more than one Block it needs to allocate.

(c) It connects to the Routing Device and requests more than one IP address. The RoutingDevice returns the IP adresses for the Local

Device and Remote Device 2.

(b) Remote Device 1 then connects to the Local device and stores one of the Blocks.

(d) Then it connects to Remote Device 2 and stores the second block.

III. Suppose the Routing Device has a Block it needs to store.

(e) Since it already knows the IP addresses of all the Remote Devices in the DMFS, it connects to Remote Device 2 and stores the Block there.

****

****

Figure 3 Complex DMFS System Model

```
                 ____
                /    \
                |User|
                \____/
                  ^
                  |
                  V             DMFS
         _______________________________________________________
        /               ^                                       \
        |               |                                       |
        |               V                                       |
        |       --------------                                  |
        |       |Local Device|                                  |
        |       --------------                                  |
        |               ^                                       |
        |               |  Network of Remote Devices            |
        |---------------|---------------------------------------|
        |               |                 ---------------       |
        |               |        __(b)___>|Remote Device|       |
        |              (a)      |         |      1      |       |
        |               |       |         ---------------       |
        |               |       |                 ^             |
        |               |       |                 |             |
        |               |       |                (c)            |
        |  -------------|-------|--   ------------|-----------  |
        |  |            V       | |   |           V          |  |
        |  |   ---------------- | |   |   ----------------   |  |
        |  |   |Routing Device|<- |   |   |Routing Device|   |  |
        |  |   |       1      |   |   |   |       2      |   |  |
        |  |   ----------------   |   |   ----------------   |  |
```

```
|  |            ^         |   |           ^         |   |
|  |           (d)        |   |          (e)        |   |
|  |            V         |   |           V         |   |
|  |   ---------------    |   |   ---------------   |   |
|  |   |Remote Device|    |   |   |Remote Device|   |   |
|  |   |      2      |    |   |   |      3      |   |   |
|  |   ---------------    |   |   ---------------   |   |
|  |                      |   |                     |   |
|    ----------------------   -----------------------   |
|       File System 1             File System 2         |
|                                                       |
\_______________________________________________________/
```

Figure 3 represents a more complexe DMFS network. This network has one more Routing Device and Remote device than Figure 1. In Figure 2, a single DMFS network holds two parallel file systems. Routing Devices 1 and 2 both identify as part of the same DMFS, but they also identify as managing different file systems. When a DMFS has parallel file systems, Remote Devices can use one or more of the file systems. Connections a-e are used to illustrate the orientation of each Device inside the network.

(a) & (d) The Local Device and Remote Device 2 are included in File System 1.

(b) & (c) Remote Device 1 is included in both File System 1 and File System 2.

(e) Remote Device 3 is only included in File System 2.

****

3. REMOTE DEVICE NETWORK FUNCTIONS____________________________________

All communications between devices in a DMFS are performed using a TCP connection. The act of communication happens through the use of Message IDs. These are whole numerical values sent from one Remote Device to another throuth the TCP connection. They're used to represent specific messages that would be sent through the network.

A Message ID can represent the state of a Remote Device, a request for data, or an error. A Signal cannot be used to pass data that's stored in the DMFS. Message IDs are grouped into the following types:

DEVICE STATE

    100 - Message Recieved

Returned if the previous signal has been
successfully processed.

101 - declares itself a Routing Device
Sent in response to Signal 202 if device is a
Routing Device

102 - declares itself a Remote Device
Sent in response to Signal 202 if device is a
Remote Device

103 - resources available
this message is sent to tell the recieving
Remote Device that the sending Device has
resources available for the recieving Device
to use.

... - EMPTY
The following values have no messages assigned

REQUEST

200 - request for Block
This signal is sent followed by the ID number
of the block. Only one ID number can follow
this Signal as it represents a request for
only one block.

201 - request for multiple blocks
This signal is followed by any amount of ID
numbers.

202 - request what type of device
Sent to determine if the recieving device is a
Routing or Remote Device.

203 - request memory for storage
this messsage is sent to let the recieving
Remote Device know that the sending Device
wants to store a block.
This message communicates the intent to store
only a single Block.

... - EMPTY
The follwoing values have no messages assigned

ERROR

300 - generic "process failed" error
A catch all Signal that is returned when an
unknown error has occurred when a Remote
Device is attempting to satisfy a request.

301 - Block not found
Returned and followed by the ID number of the missing Block in response to Signals 200 and 201. This should not prevent the Remote Device from returning Blocks that are present.

302 - insufficient permission
Returned when a Remote Device cannot satisfy a request because the requesting Remote Device is not the owner of the Block and the Block's permissions do not support the Remote Device's access.

303 - request signal out of order
Returned when there is missing data or communication that should have preceded the request.

304 - not enough resources
Returned when the Remote Device cannot satisfy the request because it lacks computing resources

305 - not enough memory
Returned when a Remote Device cannot store the data given to it because it lacks the memory.

... - EMPTY
The following values have no messages assigned

ROUTING

400 - request IP address
Sent to the Routing Device. The Routing Device sould return a single IP address of a Remote Device in the DMFS.

401 - request multiple IP addresses
Sent to the Routing Device followed by the number of IP addresses being requested. The Routing Device should return that number of IP addresses.

402 - request authentication
send by the Routing Device to a Remote Device in order to allow a certain action

403 - request to join DMFS
sent to the Routing Device by a Remote Device that wants to join the DMFS.

404 - authenticate this device

this message is used by a Remote Device to request that the Routing Device verify the user account of a third Remote Device. After this message is sent, the first Remote Device will supply the ip address of the third Remote Device that needs authentication.

405 - Device is verified

this message is sent in response to Message 404 if the Remote Device in question is an authentic user.

... - EMPTY

The following values have no messages assigned

Note that not all numbers within the Signal Types have been used. This allows more signals to be added in future versions of the DMFS Protocol.

If a Signal has already been assigned a number, it cannot be assigned to a new number. Additionally, no other signals can be assigned to the same number. If a new signal is added to the Protocol, it should be added to the next consecutive available number in its Type.

3.1 COMMUNICATION FUNCTIONS

A Remote Device may accept as many connections to other Remote Devices as its system can handle.

After recieving any message or data, the Remote Device should return Message 100 to the sending Device. If the recieving Device doesn't return Message 100, the sending Device should repeat sending the data or message until it recieves Message 100. Message 100 should never be returned in response to another Message 100.

3.1.1 DATA REQUESTING

This is the simplest function. A requesting Device can send any Message from the 200 group and the holding Device will send the information requested.

3.1.2 BLOCK SENDING AND RECIEVING

In order to send a Block to another Device, the sending Device needs to first send a request for resources from the recieving

Device. The recieving Device can either respond saying resources are available or giving an error that resources aren't available.
If the recieving Device gave an affirmative, the sending Device will send a single Block.

When requesting a Block the process of asking for resources is skipped. The requesting Device requests a Block from the holding Device. The holding Device then sends the Block.

3.1.3 AUTHENTICATION___________________________________

There are certain operations that Remote Devices can perform that require the User Account be verified, such as removing a Block.

A Remote Device that encounters a request that needs authentication will send Message 404 to the Routing Device and the IP address of the second Remote Device that needs to be verified. The Routing Device will then send Message 402 to the second Remote Device. If the second Remote Device provides the correct credentials, the Routing Device will return an affirmative message to the first Remote Device and the operation will continue.

3.2 ROUTING FUNCTIONS____

A Routing Device can be included in a DMFS network that's managed by a seperate Routing Device. It can also include seperate Routing Devices in its network.

Routing Devices are capaable of acting as a traditional internet server. It will support a Domain Name and will accept http traffic. Domain Names for Routing Devices must be stored as an SRV record.

3.2.1 ADDING DEVICES__________________________________

A Remote Device that is not included in a DMFS can send Message 403 to a Routing Device. The Routing Device will expect the Remote Device to then send a Username and a Password. They should be sent to the Routing Device seperately.

3.2.2 REQUESTING IP ADDRESSES

When a Remote Device needs to communicate with a second Remote Device, it asks the Routing Device for one or more IP addresses.

3.3 ALTERNATIVE CONNECTION METHODS

The default path of interaction between Remote Devices is through an internet connection. Other remote technologies can be used as well such as a Bluetooth or Cellular connection.

3.3.1 ALTERNATE CONNECTION BEHAVIOR

Because other connections may not give as easy access to all Remote Devices in a DMFS as an internet connection can, a Remote Device using a supported alternative connection will have to modify how it connects to the DMFS.

[explain modified behavior in Version 0.2]

3.4 CONNECTION ERRORS

If two Remote Devices are interacting and the connection is suddenly lost, the Remote Device that initiated the interrupted interaction will be responsible for re-initiating the interaction again.

3.4.1 ROUTING DEVICE CHANGES IP ADDRESSES

If the Routing Device(s) for the DMFS have a domain name, then the Remote Devices can simply interact through the URL associated with the DMFS regardless of the individual addresses of the Routing Devices.

If the Routing Device(s) changes addresses and the Remote Devices were relying on it's IP address to interact with it, there is no good way to recover the DMFS network.

4. DATA STORAGE FUNCTIONS

4.1 UPLOADING AND DOWNLOADING PROCESSES

This section describes the processes of moving information into a DMFS.

Before information is stored in a DMFS, it must be converted into Blocks. The following diagrams show the conversion processes for uploading and downloading data from a DMFS.

****

Figure 1 Converting to a Block

```
  ___________
 /           \
 |           |
 |    File   |
 |           |
 \___________/

     ^
     |
     V

  _________________
 /                 \
 |   ___________   |
 |  /           \  |
 |  |           |  |
 |  |    File   |  |
 |  |           |  |
 |  \___________/  |
 |                 |
 |   Encryption    |
 \_________________/


     ^
     |
     V

 |^^^^^^^^^^^^^^^^|
 |                |
 |128 Kilobyte    |
 |section of      |
 |encrypted data  |
 |                |
 |                |
 vVvVvVvVvVvVvVvVvV

     ^
     |
```

```
                     V


                ______________
               /              \
               |              |
               |    Header    |
               |______________|
              (                )
              ------------------
              |^^^^^^^^^^^^^^^^|
              |                |
              |128 Kilobyte    |
              |section of      |
              |encrypted data  |
              |                |
              vVvVvVvVvVvVvVvVvV
```

When a file is converted into Blocks, it is first encrypted. Then the file is divided into 128 Kilobyte sections (referred to as a File Section). A Header is then prepended to each section. These Headers each contain the information described in section 2.2.2.

When converting a file to Blocks, the Block ID numbers given directly correspond to the position of the File Section in the file it originated from. This means that when a file is divided into 10 Blocks, the Block containing the first File Section of the converted file will recieve the number 0 and the Block containing the last File Section will recieve the number 9.

****

4.2 ALLOCATION____

When Blocks are being stored in a DMFS, Remote Devices that spend more time on average connected will recieve more Blocks. This increases the stability of the DFMS as it increases the accessiblity of the information stored in the DMFS.

5. RESOURCE MANAGEMENT____

All Remote Devices preform resource management tasks.

5.1 STORAGE MANAGEMENT____

A DMFS works by distributing the task of storing information to all the devices in the network. Each Remote Device must make at least one Gigabyte of its storage permanently available for use by the DMFS. This reserved storage is where the Remote Device must store all Blocks that it accepts. This portion of storage can be released from use when the Remote Device permanently ceases to be a part of any DMFS.

If a Remote Device has already reserved storage for a DMFS and then joins another DMFS, it is not required to enlarge its current portion of storage or make a new portion. Remote Devices can use the same portion of storage for multiple DMFS networks.

When a Remote Device has used all its portion of storage, it must notify the Routing Devices of all DMFS networks it's connected to that it has no storage available. The Remote Device has the choice to enlarge its portion of storage or wait until Blocks are removed and it has storage available again. In the event that the Remote Device is able to store more Blocks again, it must sent a message to the Routing Devices of all the DMFS networks that it's connected to that it has resources available.

5.2 MANAGING DEVICES

[This section may or may not be useful; keep for now]

5.2.1 ACCESS CONTROL

Routing Devices are responsible for access management.

The default implementation for access control is to use the system of user accounts described below. Other implementations can be used in place of or alongside the default system as long as their use does not require any special modifications to Remote Device behavior.

Each Remote Device will store a username and a password for each user that interacts with a DMFS. When a Remote Device first connects with a DMFS, it shares the usernames and password hashes with the Remote Device.

Associated with the username and password is a 128 bit unique ID number.

If a DMFS is connected in a hierarchy, the job of assigning an ID number is given the the Routing Device at the top of the higherarchy. The new ID number is then passed back to the Routing Device overseeing the DMFS with the new user.