Title: Multicast Tree Mirroring(MTM)

Author: Patcharee Basu (yoo@soi.ne.jp)

Date: January 31, 2005

Overview

In the distributed education environment, file distribution system is an important mean to deliver knowledge in forms of lecture materials and archive lecture from lecturer to students in different geographical areas. Despite the fact that new form of distributed education system is based on one-to-many model on a shared medium, i.e. satellite link, most of the file distribution systems today are still based on one-to-one model which creates redundant traffics on shared paths between sender and a number of receivers.

Multicast Tree Mirroring system<MTM> is designed to reliably deliver file system in form of a directory tree from sender machine to a number of receivers on top of IP Multicast. With a given list of receivers, MTM makes sure that all of them correctly get data by keeping its attempts to retransmit in any failure cases without timeout restraint.

MTM Architecture

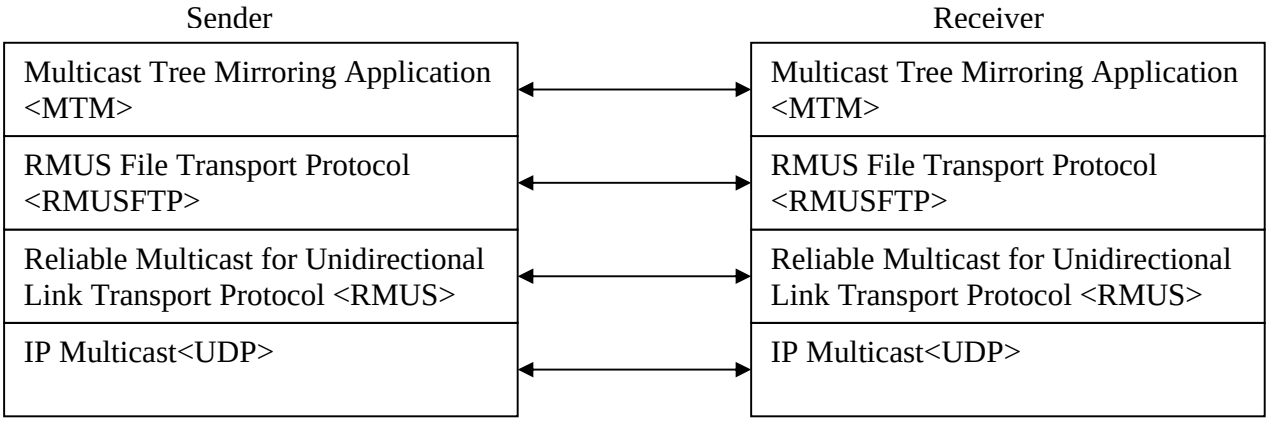


Figure 1. MTM Architecture

Layerly classified in figure 1, IP multicast is the bottom most layer providing broadcasting capability on the Internet with minimum bandwidth consumption, but it does not provide reliable data transfer and does not provide any congestion control to the transfer session. Therefore, RMUS , the upper layer protocol works as transport protocol providing error control and congestion control. RMUS is specially designed for UDLR satellite link. The application layer starts from RMUSFTP layer which is multicast file transfer protocol making use of RMUS to ensure correctness of multicast file transfer.

Like other FTP applications, RMUSFTP has a session timeout that receiver who could not join in time will fail for the transfer. This characteristic does not fit the nature of mirroring application which wants a guarantee that all mirror servers will get files although they could not join in time due to some reasons like server down, network disconnected. MTM is designed to run on top of RMUSFTP providing a guarantee of data integrity and reliability for multicast directory transfer. In conclusion, MTM is a big scale multicast transfer application which reliably deliver a directory tree from a master machine to a list of receivers machine on top of IP multicast with guarantee that all them will eventually receive expected data.

MTM Components

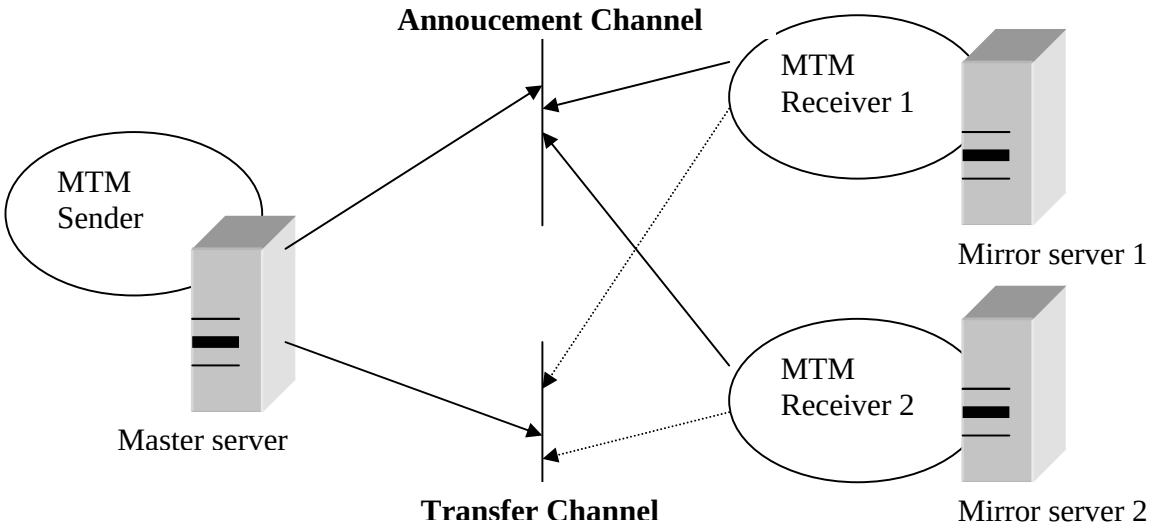


Figure 2. MTM Components

MTM has two main components as shown in figure 2, MTM sender and MTM receiver. MTM sender is running at the machine which contains the master copy of a directory to be distributed. MTM receiver is running on each server that will mirror contents from master server. These components communicate through an announcement channel which is an agreed multicast group. MTM receiver will be joining to the announcement channel at all time and listen for the file transfer announcement from MTM sender, then MTM receiver will join a file transfer at a transfer channel which is another multicast group told in the announcement message.

MTM work flow

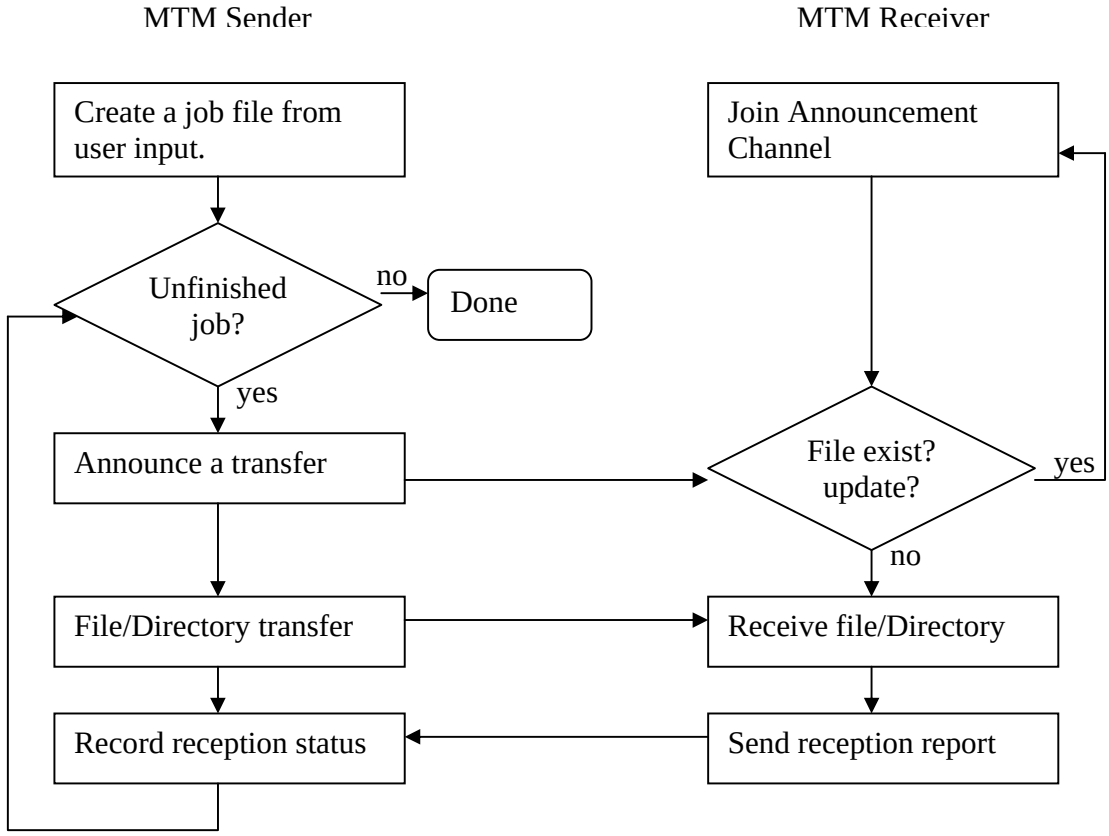


Figure 3. MTM Work Flow

Once user wants to distribute a directory from master server, user will give an input of which directory to be sent and which receivers to be guarantee. Then traces the directory tree for files and directories and creates a job file which keeps the information of file/directory list and receiving status of all given receivers as shown in table below. (All the working status of MTM sender will be kept in to hard drive so that in any cases of failures, MTM can recover and continue its work to finish all the intended transfers)

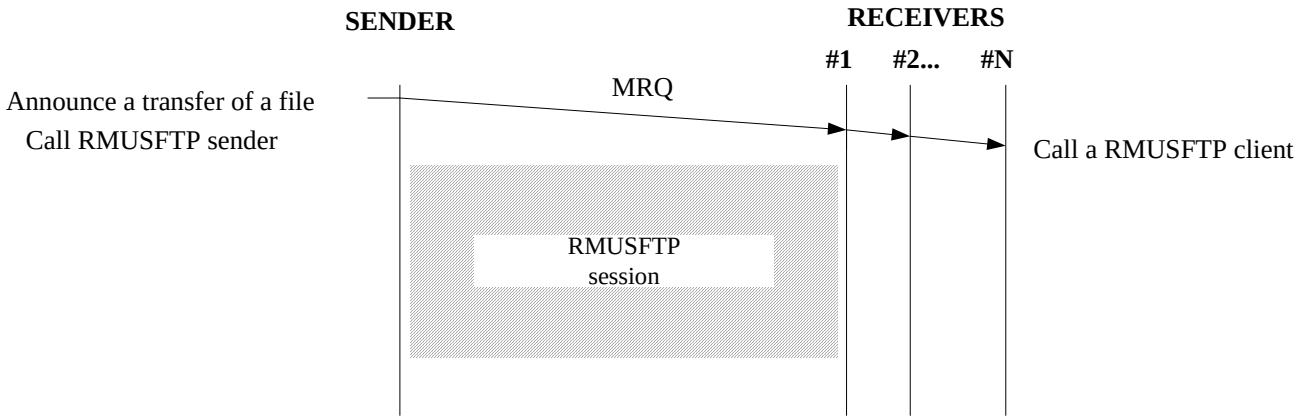
Directory1	Receiver1[STATUS],Receiver2[STATUS],
Directory1.File1	Receiver1[STATUS],Receiver2[STATUS],
Directory1.File2	Receiver1[STATUS],Receiver2[STATUS],
Directory1.File…	Receiver1[STATUS],Receiver2[STATUS],
Directory1.FileN	Receiver1[STATUS],Receiver2[STATUS],
Directory2	Receiver1[STATUS],Receiver2[STATUS],
Directory2.File1	Receiver1[STATUS],Receiver2[STATUS],

Directory2.File2 Receiver1[STATUS],Receiver2[STATUS], Directory2.File… Receiver1[STATUS],Receiver2[STATUS], Directory2.FileM Receiver1[STATUS],Receiver2[STATUS],

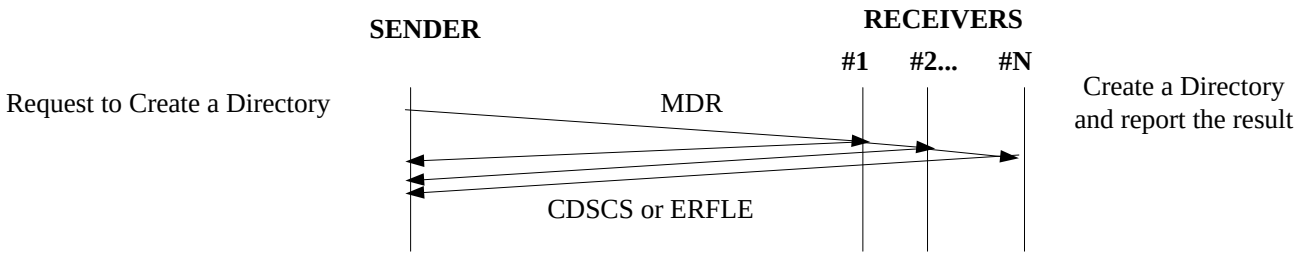
For each entry in the job file, MTM sender will check if not all receivers successfully get the file or directory, it will announce a new transfer. This announcement is going through the shared channel telling which file to be sent and which IP multicast address is the transfer channel. MTM receiver will get announcement message and check if it has an update copy of that file or directory on its server or not. If yes, it ignores this announcement and listen for next one. If not, MTM receiver will join the transfer channel which RMUSFTP will take place to transfer a file. At the end of transfer, MTM receiver will report the reception result and MTM sender records it into the job file. MTM sender keep retransmission if some receiver still could not get the file.

MTM protocol

Scenario 1



Scenario 2



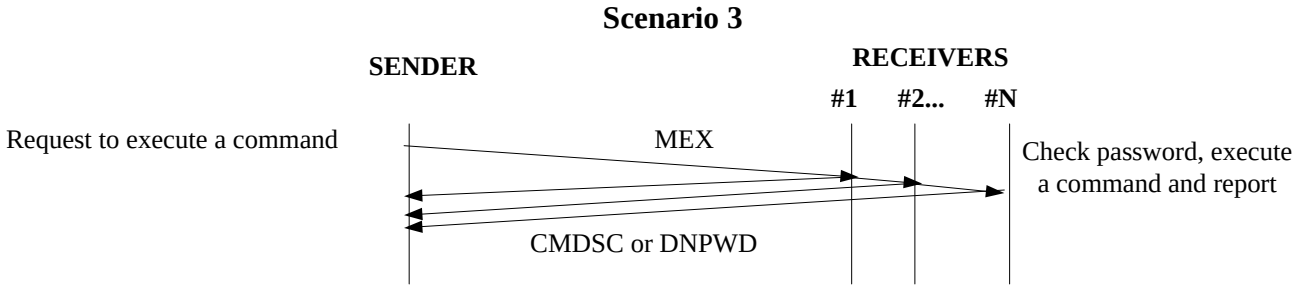


Figure 4. Scenarios of MTM protocol

MTM protocol is designed so that sender has full control of operations to create a directory tree on receiver machines. As shown in figure 4, there is no specific sequential flow of the protocol but the sender will send a request to receiver to perform an operation and then take the report back if the operation success or not.

Three types of requests are defined since they are basic operations to make a mirror tree at receiver site.

File transfer request(scenario 1)
Sender sends a request packet[MRQ] containing
- filename
- address/port where RMUSFTP session will take place
- list of receivers allowed to be join this transmission.

Receiver then join the transfer by RMUSFTP, there is no acknowledgement in MTM layer since the RMUSFTP layer will get all the report.

Directory creation request(scenario 2)
Sender sends a request packet[MDR] containing
- directory path
- directory attributes < access time, modified time, mode>
- feedback port
- list of receivers allowed to do this operation.

Receiver will create the directory, set directory status according to given attributes. If creation success, receiver reports CDSCS packet. if not, it reports ERFLE packet.

Command execution request(scenario 3)
Sender sends a request packet[MEX] containing
- Command to be executed
- Password
- list of receivers allowed to do this operation

And since this operation will be performed only when the tree mirror success. Sender gives more information with MEX packet so that receiver can log the mirror work status

- root directory name
- root directory description

Receiver checks the password. If not correct, it sends back DNPWD packet. If correct, it executes the command and write a log file to record a completion of tree mirror.

MTM Deployment

MTM has been used to distribute the lecture materials and archive course for SOI Asia courses since September 2004. MTM has been working well with 12 MTM receivers currently installed in AIT, AYF, BUET, CRMA, ITB, ITC, MUST, PSU, UNIBRAW, UNHAS, UNSRAT. MTM has been developed to support IPv6 multicast and will be deployed to SOI Asia environment in the near future.