Peer Lingering Unaccomplished Time Oriented Network (PLUTON) File System

Rahul Raj M
Ilahia College of engineering and technology
Muvattupuzha,Kerala,India

Rosna P Haroon
Ilahia College of engineering and technology
Muvattupuzha,Kerala,India

Abstract - The main aim of any type of network level file system is to share the objects among a number of users. PLUTON is also a network oriented file system works in a peer connected network, which will provide some other benefits more than any common network file system provides. Here we are proposing a network file system, which can provide remote file system access for a specific period of time. The connection is automatically released whenever a threshold time limit is reached and the remote accessed files will be in an encrypted form.

Keywords - PLUTON, DS, ESP, RPC, DHT, FUSE, time to live

I. INTRODUCTION

The increased use of networks is raising the importance of network oriented file systems, also the sharing of objects. Distributed file system not at all the last word of network file systems, there exist more types of file systems having the performance better than the distributed file systems. While dealing with distributed systems two things should take into account. First one is the background tasks and next one is the necessity. In simple words it is the one out of thousand jobs of a distributed system to share an object among multiple users, by doing so it should also take care about so many other things such as the synchronization mechanism to avoid collision, service satisfaction algorithms to avoid starvation[3]. Next thing is to consider about whether a file system such as distributed system is necessary for a LAN connected network. The question is very clear that the distributed system does many operations to preserve data integrity and to avoid conflicts, but in LAN which is an interconnection of a few network, so is it feasible to implement a distributed file system in LAN network? Answer will be always no, because current situation is not requiring that much of advanced techniques.

The use of virtual file system will preserve most of system parameters such as time, work load, memory interactions also increases system performance. But there is always an ambiguity in the use of virtual file system is that whether they provide the same performance of others.

Not only sharing of objects but also storing of objects is also a main concern of network oriented file system. All systems have a fixed amount of memory inbuilt in the system, here we can call it as the local memory, additionally some secondary storage devices canalso be added to the system to increase memory, but if that is not enough to satisfy the requirement we should avail of the memory available in network, that is the user logged in one system can store his/her data to another system. Other than theperfect utilization of available memory the data stored in a network is having high security that the data in a single node due to the effect of strong encryption algorithms.

II. RELATED WORKS

Without a doubt we can say that the distributed system is parallel to PLUTON, also we can say that PLUTON is implementing some properties of the distributed system.

A DS (distributed system) is the way of globalizing physical resources in a local system. Simply a distribute network is a collection of resources connected by a network. Advantage of this scheme is that resources are sharable. Each system can use the free memory space of other networks. Totally effective use of memory can be obtained here [5].

The main design issues coming with distributed systems are the different types of transparencies, flexibility, reliability, performance, scalability, security and fault tolerance. Transparency deals with abstraction of system's implementation, which implies that there are a number of complex operations, are running on the back side of the

system, and user is hidden from those details. Flexibility concerned about the capability of the system to work with different types of platforms. The fault tolerance of the system is directly proportional to the reliability. More the fault tolerance more the reliability of the system can be achieved. Performance is actually the measure of parameters of the DS;At any cost the DS should provide high performance. Scalability is nothing but the easiness to add any number of nodes to the system without interrupting common operations of the system. Security is the first concern in a DS because it shares more number of resources among wide variety of users. The chances for intrusion are more in this case, Also the integrity of resources are also the main concern.

In a distributed network the remote login is not at all a concern because here the file system is distributed among all users. Because of the transparencies' each user seems that there exist only one file system in the network. Considering the local network where a number of computers are connected using LAN cables. Here each user is aware about the existence of individual file systems. A global file system does not exist in this concept. The only provision of users is to send files and folders among each other. In a Linux kernel based operating system (Ubuntu 10.4) there exist a command called 'scpy' who's meaning is system copy. Here there are two arguments in the right side of the command.First argument is the source file to be transferred with absolute path and the second argument is the destination to which the file should be copied, it is also specified with absolute path. The absolute path include the ip address of the system followed by the complete path in the system begins with root of the directory structure (/), But this is not at all considering as a file system operation. Next provision provided in Ubuntu is the remote login facility. Here a user can login to other system by knowing thepassword of the others. This is actually a disadvantage of the mechanism to expose the password to others. The packages such as Virtual network computing is present to provide interfaces foe remote login but the main disadvantage of those are slow working speed, offering fewer features and less security options.

Remote Procedure Call (RPC) is the mechanism provided for one process to invoke method in a remote system. Messages are another way of communication among users. They mainly provides two primitives:**send**(*receiver, message*) and **receive**(*sender, message*) [5].

Two semantics provided for message passing are non-blocking and blocking semantics. Blocking semantics block message passing while congestion occurs and non-blocking works in an asynchronous manner [5]. Reliable and unreliable communications are also present. Reliable communicationprovides an acknowledgement mechanism in response of data received in receiver but which is not present in unreliable connections [5].

Consideration of data security in network oriented file system is having high importance. The better way to think on data security is always encryption. In this world of highly growing e-mail technology, the encryption of the same is very important if the mail is critical. The importance of email encryption is explained in the below figure1. There Ann is sending a mail to Carla a friend of Ann. Ann is discussing about the divorce details of her with Carla, at the time of divorce trails are taking place in the court. Here comes the importance of encryption.

Any detail Ann shares with Carla is beneficial to her husband's lawyer because it would help in his case. So here is a possibility for an attack by the lawyer on the mailed content. So the mail server should provide some security measures against attack. The simple idea is to encrypt the email while sending through the network. For that some keys should be provided to encrypt and decrypt the message depending on which encryption scheme is using. It may be public key cryptography or private key cryptography. The email service provider (ESP) takes advantage of self-destructing caches and trashes [2].

The distributed hash table concept is having a limitation that it is applicable to a large system and which should have all the properties of a distributed system. So in simple words it cannot be applied to a simple LAN connected network. Also the complexities due to the implementation of distributed hash table are much higher for a small network just like LAN.

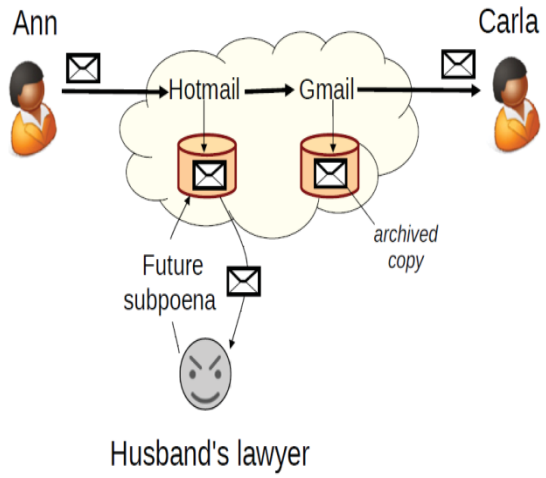

Fig 1: Example scenario

The ESP is responsible to keep the sent messages. If the user needs it once again it should be provided by the ESP, So that caching or extra storage of mails which were sent is an overhead to ESP, but it should be necessary in the view point of a user for the readily availability of data resource. Here we use the self-destruction concept. Clearly, storing all mails from the beginning of the ESP establishment is not economical. So we should provide a destruction mechanism for the data elements. Those mails which expires a particular time limit should be destroyed themselves. For that we need a particular encryption algorithm.

Vanish architecture uses the concept of Distributed Hash Tables (DHT), where the data is stored as key value pair. It is using the method of cryptography. Here data is divided in to N blocks and are encrypted with N keys and storing in different locations. A person who knows all key values can only decrypt the data, which is explained in figure 2 [2].

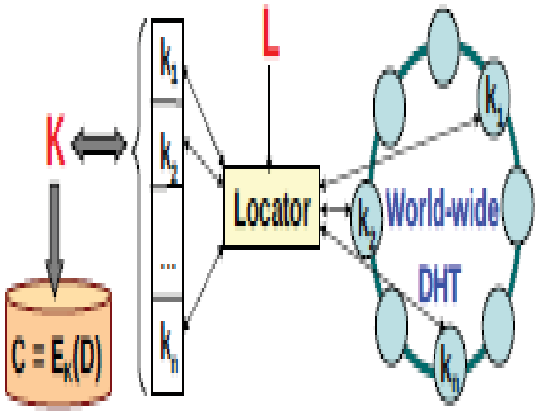

Fig 2: Distributed Hash Table (DHT)

The figure shown above is a distributed hash table where we are using the concept compactable to a distributed network, not for a simple LAN connected one.

III. PROPOSED SYSTEM

The name PLUTON explained as Peer Lingering Unaccomplished Time Oriented Network. The first word Peer stands for Peer connected network, because here we are considering a peer connected network. A peer network is one in which each system can work as both client and server. Simply while acting as a server which can access the file system of other nodes too. The word lingering stands for avoiding deficiencies of a peer system, because most of the peer systems haven't good security mechanisms. Here to avoid that we are providing a self-destruction

provision to the files. That means the inabilities of security deficiencies of a peer system are overcome here by the encryption algorithm.That is we can overcome the security deficiencies of a peer system by implementing some encryption techniques.Also importance of the term PLUTON is that the copied data in the system is having the property of self-destruction. Plutonium is a radioactive element which is also having that same property. So the name PLUTON is apt for a self- destructing file system.

A. System Architecture

PLUTON is virtual file system working in a peer connected LAN environment. Here each user in the file system is permitted to connect with another system in the network, through which each user is permitted to view the file system of another system. While connecting with other systems the requested one is termed as client and the request satisfaction node is called the server. Even though it is a peer connected network for conveniencewe are designing the architecture in a client server manner.During the connection period, if the client is copying any of the server files,after a particular threshold time limit files will be automatically encrypted using AES Algorithm.

The basic working of PLUTON is explained with the below fig 3. The main components of the system are server and client. There is no confusion with the peer network because there is no such a concept of client and server in peer network because here each system is client and server, but for convenience here we are introducing the concept of client and server. It is only for convenience. A client is requesting for connection through the network by specifying the ip address of the server. An internal authentication should be provided for the identification of each client system. It is actually the advantage of PLUTON, because in normal mechanisms like remote login we need some explicit authentication schemes such as password of the server system or individual identification mechanism for each client, but instead of that PLUTON is occupied with the ip addresses of the clients which are connected in the network. The point to be noted here is the number of peers to be included in the system. It is having two concerns. First one is to add nodes to the system as less as possible, because it is a network oriented system with less number of properties adapted from the distributed system.

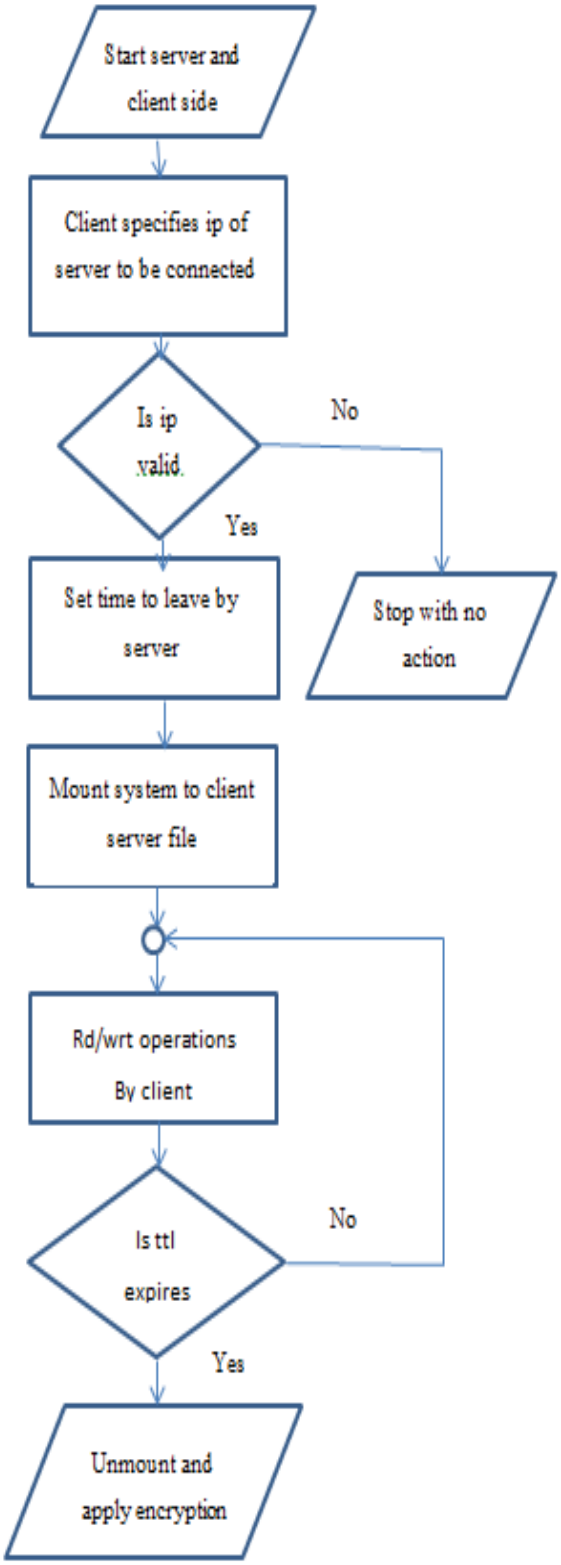

Fig3: System architecture

If you are trying to adapt allproperties of a distributed system then PLUTON becomes more complex and whose performance will reduce. The second concern is adding the reliable nodes to the system. The word reliability is not related to the system but related to the persons who are using the system, that is each peer in the system should be reliable and there should not be at least one node which is unnecessary. This can be explained with an example.

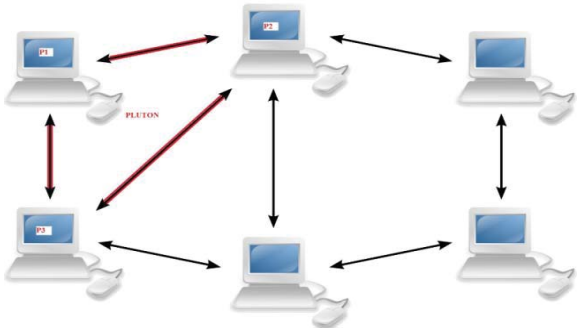

Figure 5:System architecture

The above figure shows the example of PLUTON existing in a network. It is very clear that PLUTON is a connection of computers existing in a LAN connection. The network may be a peer connected or not but the PLUTON is having the nodes which are peer connected. Simply PLUTON is a peer network inside a network, which implies that this connection is a secure peer connection in which each node is really need a connection with other node and each of them are trustworthy. Consider a university in which a number of computers are connected with LAN, out of that the department heads should have something more to communicate, for that need we can implements the concept of PLUTON. The computes of department heads can be connected to form a PLUTON. This satisfies the first concern. There must be less number of nodes should be in the network, because only a few nodes are in the PLUTON. It is the only necessity to connect the department head's computers each other because they have something more to share other that common employees shares. The next concern is also satisfied here. Each department heads are trustworthy. It is actually based on the belief that those are in the reputed positions of the university and won't venture for any mal practice. The main confusion occurs here is that what is the need of the consideration of security? As stated earlier here we are using internal authentication, which means each node get access to another node without specifying any password or other authentication mechanisms. This is on the assumption that all nodes connected in PLUTON are the most reliable nodes. No any other node is allowed to enter in this network.

After verification of IP server allow the client to access its file system. The client can perform any of the operation which can be performed in a local file system such as creation of a file/directory, delete, edit cut copy paste etc. One thing must be noted that for the beginning of every connection server is specifying a time limit. We can call it as time to leave (ttl) time constraint. Why it is called as a constraint is that it is limiting the disclosure of server data resources to client. It implicitly says that whenever client can poses the files of server for a specified time. Whenever client is copying any of the server file to its system, it will be noted by the server by storing that time to somewhere and which must be passes to client, and then it is client's duty that to check continuously whether time is exceeded or not? This constraining time is found by adding the time when the file is copied and the value of ttl set by the server. While crossing the time the system will show a warning message to the client that "your files will loss within few minutes".

The next consideration is about what we should do after expiring the time? The normal way of thinking is that to delete the files permanently which are copied to the server, because it is server's need is to keep all its files in inside of the system permanently. So after expiring the time the file should be unavailable to client. To achieve this, removal of file from the entire system can be adopted, that is files may be deleted after use, but it is not at all an efficient mechanism. In Linux kernel based operating system it is easy to recover deleted files and folders even if they were permanently deleted. Slut kit is such type of a package which is used to recover deleted files. The next best way is encryption. While expiring the time we can be encrypt that file to an unreadable format using a strong encryption algorithm. Here we are using Rijndael algorithm for that purpose, which is known as the strong encryption algorithm at present. The other enhancement which can provide is the repeated encryption that is the encryption algorithm is placed in a loop with a few minutes delay so repeated encryption will be applied to the files till the connection of server releases.

B. Solution Methodology

The layered structure of the system is given in fig5. The layering is based on the interaction among different components. Here the common resources of a computer system such as main memory cpu are not specified,

instead of that what are the additional components to be added with a system to archive a PLUTON connection is shown in figure.

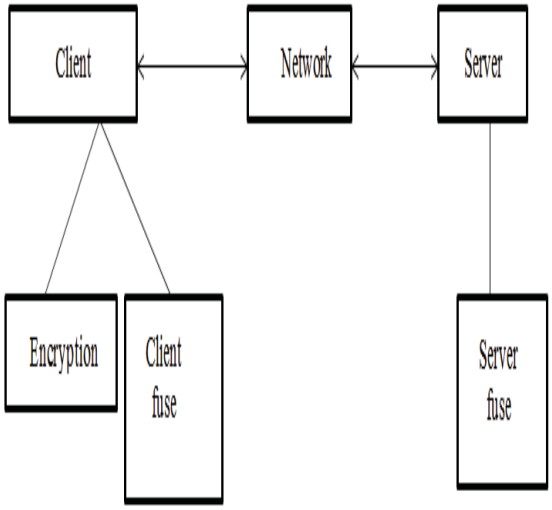

Fig5: Block diagram of PLUTON

Table1: Rijndael key description

Key Length	Number of Keys	Rounds	Subkeys
128	2^128=3.4x10^8	10	44
192	2^192=6.2x10^57	12	52
256	2^256=1.1x10^77	14	60

The block diagram is provided to understand different components of the system. The essential parts of the system are client and server. As stated earlier client send requests and server is to send responses back. Network portion is dealing with the network management, which is not in our concern because it is purely socket programming. While passing a request network checks for the availability of the requested port if the port is available allocate it to the requested party else showing a port busy message.

The main subpart of the client is the encryption subsystem. Actually this encryption subsystem includes two things. One is the time check mechanism and another is the encryption algorithm AES itself. The time check mechanism is same as stated earlier which checks whether
the provided time expires or not, if yes it will call the AES algorithm to encrypt the copied file. The question is about why we are using AES algorithm? We are using AES algorithm because it the strongest algorithm available in till this time. AES is not broken by any attacker; also it is the most widely using algorithm especially by the US government for many of the sophisticated data storage such as navy usage, military purposes etc.

Advanced Encryption Standard (AES) is actually a conference conducted by NIST. There were there AES conferenced conducted repeatedly at last by a complex method of elimination process they chose Rijndael as the new encryption algorithm, invented by Joan Daemen, Vincent Rijmen. It is having several advantages which are listing below:

1. Rijndael is not a fiestel cipher. The property of fiestel cipher is that whose each round look like same so crypt analysis won't be easy[6].
2. All of the rounds are not equal. The Rijndael is having 10, 12 or 14 rounds according to the key size also in each round there are four different sub rounds which areSubBytes, ShiftRows, MixColums, AddRoundKey. First round include only AddRoundKey step, last round excludes AddRoundKey and intermediate steps includes all the for steps. This gives an advantage that all of the steps are not same so easy crypt analysis is not possible.[6]
3. The possible key numbers are high so no brute force analysis impossible.[6].

The strength of any of the encryption algorithm is its key management. Below there is a table showing the possible number of keys and rounds that can be selected by the user while encryption.The blocks common in client and server are the client and server FUSE. Actually both are common but gave different names due to their existing environments. When it is present n server side then call it as server FUSE when it is in client side call it as client FUSE, anyway function of this subsystem is to provide necessary file system operations in any system.

Expansion of FUSE is File system in User SpacE. As the name implies FUSE is user level file system, which is used to perform of all kernel level operation from user level. Simply we can say that FUSE is a frame work for performing file operations from user side. The existence of FUSE is given in the below figure.

The basic architecture of FUSE is given in figure above. The first question comes is what is the importance of FUSE in a file system oriented operations. The answer is very clear. All the operations in file system such as creating file/ folder, deleting a file/ folder, renaming file/ folder etc are implemented in kernel level. It is very difficult to change kernel level codes by going deep in to the file system implementation, and another point to be noted is that once we change the kernel level properties of the file system which will reflect on each and every operations in the entire file system. To avoid these difficulties we go for use level file system. Here we are just modifying the necessary operations of the file system when needed. The primitive components of FUSE are some predefined functions, which are used for creating, deleting, renaming file and folders. Our aim is to change the working of these functions and for that we will alias there functions and give new definitions for these. Aliasing is the same concept of implementing interfaces in java. The functions are actually having declarations and no definitions, we are aliasing these files by specifying new names and then modifying the code. While doing operations we can declare so many predefined structures and can use them for different operations.

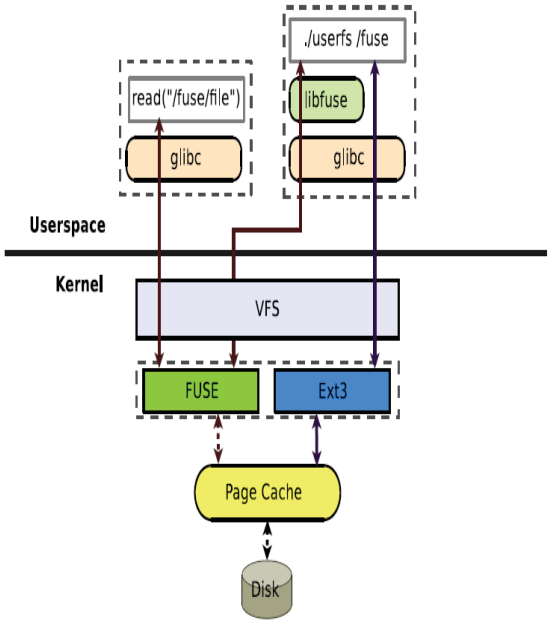

Fig6: FUSE architecture

IV. CONCLUSION

PLUTON is a network file system distributed over a few number of computers compared to distributed network. While considering the security features,this will work better than any of the distributed systems, because

of the use of encryption algorithm. Normally peer connected system doesn't consider about the security features, but here we are proposing a client server architecture within the peer system and providing more security to the server.

V. FUTURE WORKS

A number of future enhancements for the system can be proposed. There are more number of properties of DS can be implemented in PLUTON which provide better performances and capabilities of the system. The time to live (ttl) constraint is also having a limitation in this system. It allocate only one ttl time for all files, but there are different files exist in the system with variable sizes. So it will be inefficient to provide same ttl to file 10 lines and another file with 1000 lines. There for an alternate mechanism should be provided to calculate and allocate ttl to each files individually based on the file size.

ACKNOWLEDGEMENT

The completion of any work is not possible without expressing gratitude to all humble minds those who helped sincerely and selflessly for the preparation of this paper. First of all I would like to express my gratitude towardsDr:BabuKurian the principal of Ilahiacollege of Engineering and technology for his immense support. Next I would like to thankMrs: Rosna.P.Haroon, Asst.ProfessorIlahiacollege of engineering and technology who guided me to write this paper very well. I would like to thank all of the staffs of CSE Dept in Ilahia College of engineering and technology who were in the evaluation committee of major project of 8th semester btech degree computer science, because I have developed this work for the major project presentation as described by the syllabus of Mahatma Gandhi University Kottayam. At last am expressing my gratitude to all friends of mine who gave me encouragement for the completion of this paper.

REFERENCES

[1] LingfangZeng , Shibin Chen , Qingsong Wei , and Dan FengWuhan National Laboratory for Optoelectronics, School of Computers, Huazhong University of Science and Technology,430074 ChinaData Storage Institute SeDas: A Self-Destructing Data System Based on Active Storage Framework)

[2] RoxanaGeambasu, TadayoshiKohnoAmitA.LevyHenryM.Levy2011Vanish:IncreasingDataPrivacywithSelf-DestructingData University of Washington.

[3] Mr. Mahesh Maurya, Mr.ChitvanOza, Prof. KetanShah,Assistant Professor, Associate Professor, Student-BTech Computer Science MPSTME, SVKM's NMIMS (Deemed-to be University)A Review of Distributed File Systems

[4] AdityaRajgarhia, Stanford University Stanford, CA 94305, US AshishGehani SRI International Menlo Park, CA 94025, USA Performance and Extension of User Space File Systems

[5] Goscinski School of computing and Mathematics Deakin University Geelong, J. Silcock and Australia Message Passing, Remote Procedure Calls and Distributed Shared Memory as Communication Paradigms for Distributed Systems

[6] Julia Juremi Ramlan MahmodSalasiahSulaiman, JazrinRamli 2012 Enhancing Advanced Encryption Standard S-Box Generation Based on Round KeyUniversity Putra Malaysia