International Journal of Enhanced Research in Management & Computer Applications, ISSN: 2319-7471
Vol. 3 Issue 11, November-2014, pp: (29-34), Impact Factor: 1.296, Available online at: www.erpublications.com

Data Theft protection by File Splitting and Multiple Encryption

Prafful Agarwal
Associate Software Quality Engineer, EMC2, Bangalore, India

Abstract: Washington 21st April 2009, a$300 Billion Joint Strike Fighter Project's most important information concerning design and electronic system was stolen from Pentagon. This not only hacked one of the most secured systems of the world but also challenged the security of highly confidential data. So to overcome the above kind of abnormalities in modern systems which store highly confidential data, in this paper we are trying a new approach to protect the most important of data on the servers, by dividing it into chunks, then encrypt ting it and sending to different servers across the net work when ever system is affected by an attack. The System has been distributed into 6 phases:

a)-Security breach warning system.
b)-First level file encryption.
c)-File breakage into chunks
d)-Second Level file encryption.
e)-Sending file over the net work to different servers.
f)-Hacker tracking system, after successful attack. After the completion of fifth phase, the files will be retrieved on another server, will be joined and decrypted.

Keywords: Blow-Fish, crypto co-processor, FGPA, Data theft, multiple encryption, file splitting.

INTRODUCTION

In this paper, our concern is about the security of highly confidential data using the concepts of high performance computing. As in the abstract part we have already listed six phases, now let's have a brief discussion of each phase one by one.

1). Security breach warning system- In this phase we de fine the protocols needed to alarm the administrator whenever an attack is detected by the system, moreover this system also invokes the second phase.

2). First level file encryption- This one focuses on faster file encryption algorithms.

3). File breakage in to chunks-This one focuses on the methods of faster file splitting, causing usage of high performance computing.

4). Second Level file encryption- This is the most important phase, as the maximum percentage of security of system relies on this phase. In this, we encrypt each splitted part of the file with different algorithm and different techniques that makes the encryption process faster.

5). Sending file over the net work to different servers-In this we discuss about the ways of sending encrypted parts of the file to different servers present in different net works plus an acknowledgement to administrator, which defines which part is present on which server with what encryption technique.

6). Hacker tracking system, after successful attack-This phase assumes if somehow attacker is able to retrieve all the files, then our system will trace the location of the hacker.

Phase i: Security breach Warning System

There is no such publicly available resource which tells us about these curity breach system for military or for highly confidential data. But the system is easily adaptable to any security breach warning system as it would be initiation step for next phase to begin which is the first phase encryption of file.

[10][9] Specialized Warning Systems, this phase aim to provide information about a specific environment. Their aim is

therefore to aid specialists in their task. All the investigated systems were described in theory. No practical, ongoing implementation was found to be publicly available. On a lower technical level, reference number, as in

[1] defines a system to predict future attacks through analysis of existing data. Variable length Markov models are used together with the assumption that an Attacker follows a certain sequential logic. The results show that the intended prediction is possible. Kimet al. reference number, as in

[2] suggest a system based on real time data. It consists of three analysis engines to minimize false alerts. The result is a forecast about the immediate future of the observed it environment. On a higher level, reference number, as in

[3] defines threat alert levels for the North American electricity sector. The Threats that have to be Present for a specific alert level are described in generic terms, such as "general threat of increased cyber(…) activity" and high level countermeasures are given. The MITRE Corporation suggested the Cyber Prep frame work reference number, as in[4]. It defines threat levels for an organization based on the intent ion and cap abilities of the potential attacker. This allows for a targeted and effective response to the suspected threats.

Phase ii: First level file encryption

Blowfish is a symmetric encryption algorithm, meaning that it uses the same secret key to both encrypt and decrypt messages. [12]Blowfish is also a block cipher, meaning that it divides a message up into fixed length blocks during encryption and decryption. The block length for Blowfish is 64 bits; messages that aren't a multiple of eight bytes in size must be padded. A graphical representation of the Blowfish algorithm appears in Figure 1.Inthis description, a64-bit plain text message is first divided in to32 bits. [7]The"left"32 bits are XO Red with the first element of a P-array to create a value I'll call P', run-through a trans formation function called F, then XO Red withthe"right"32 bits of the message to produce anew value I'll call F'.F' then replaces the "left" half of the message and P' replaces the" right" half, and the process is repeated 15 more times with successive members of the P-array. The resulting P' and F' are then XO Red with the last two entries in the P-array (entries17and 18), and recombined to produce the 64-bit cipher text.

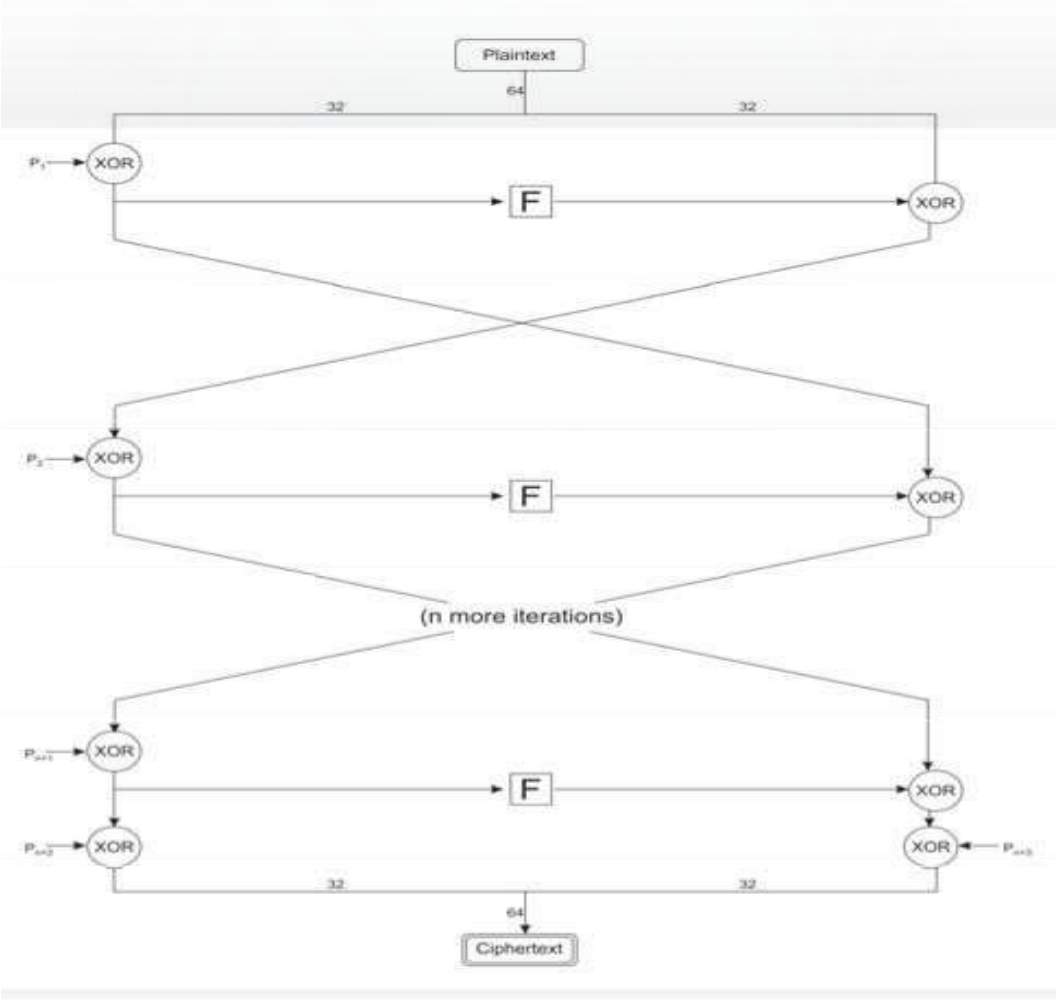


Figure-1: Graphical Representation Blowfish Algorithm

A graphical representation of F appears in Figure 2. The function divides a 32-bit input into four parts and uses those as indices into an S-array. The lookup results are then added and XO Red together to produce the output. Because Blowfish is a symmetric algorithm, the same procedure is used for decryption as well as encryption. The only difference is that the input to the encryption is plaintext; for decryption, the input is cipher text. The P-array and S-array values used by Blowfish are recomputed based on the user's key. In effect, the user's key is trans formed into the P-array and S-array; [12]the key itself may be discarded after the trans formation. The P-array and S-array need not be recomputed (as long as the key doesn't change), but must remain secret. I'll refer you to the source code for computing the P and S arrays and only briefly summarize the procedure as follows:

- P is an array of eighteen 32-bit integers.
- S is a two-dimensional array of 32-bitintegerofdimension 4x256.
- Both arrays are initialized with constants, which happen to be the hexadecimal digits of π(a pretty decent random number source).
- The key is divided up into 32-bit blocks and XO Red with the initial elements of the P and S arrays. The results are written back into the array.
- A message of all zeros is encrypted; the results of the encryption are written back to the P and S arrays. The P and S arrays are now ready for use.

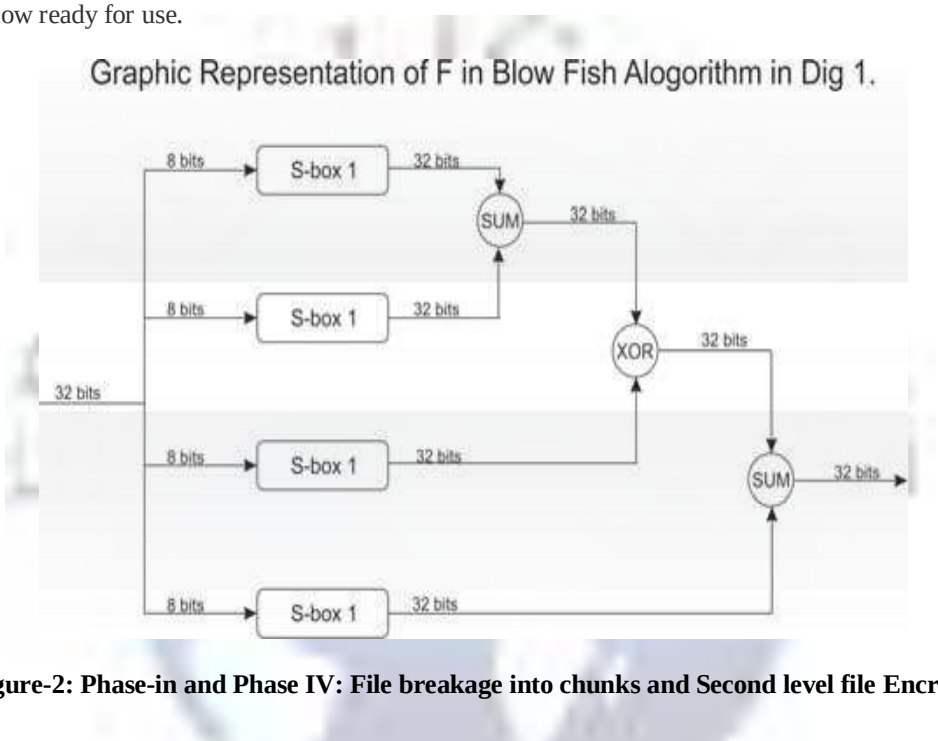
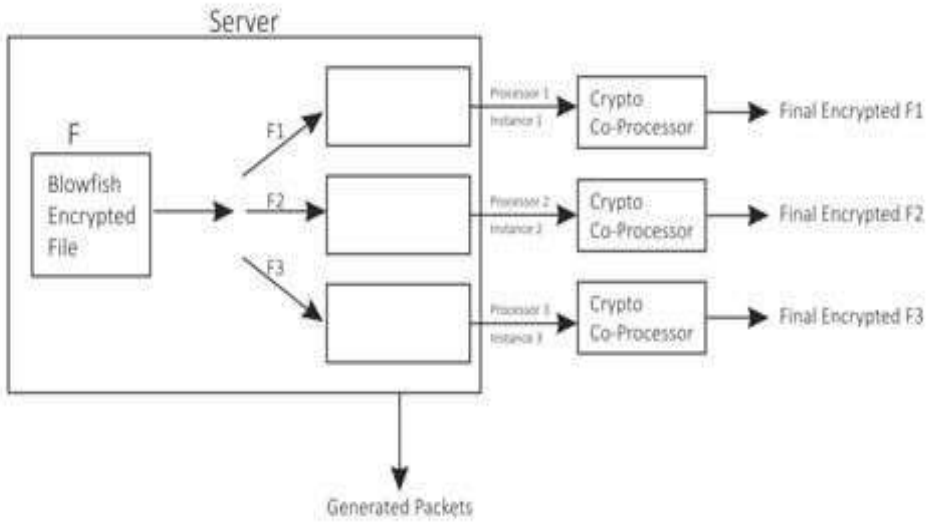


Figure-2: Phase-in and Phase IV: File breakage into chunks and Second level file Encrypt



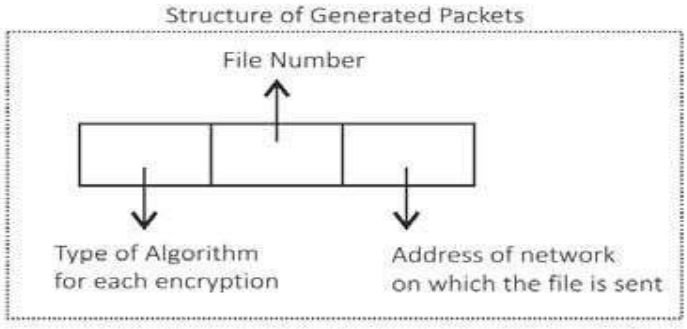


Figure-3: File Breakage and Double Encryption

Till second phase we have encrypted the file with blow fish algorithm, now in this upcoming phase we are going to split the file and then encrypt Each splitted part again with different algorithm with The help of crypto co-processor. The second phase encrypted file will be divided into different parts according to the required needs, and then we will have to encrypt tall the splitted files as can be seen in Figure-3.Forthatwearegoingtomixtheapproaches suggested by reference number, as in [5] for multi processing and reference number, as in [6][13] for co-processor cryp to graph y, so that multiple encryption with very high speed can be achieved. These are the certain points showing the explanation of third phase:

1)-After the main file will be encryp ted using BLOWFISH, it will be splitted into three parts and saved on the server at different locations.

2)-As we know that un like Processors, co-processors cannot fetch instructions from memory, so it is the duty of main server to provide the file stocryp to co-processor.

3)-To make the above process fast, we will use the multi processing system and provide the work of getting the splitted file from the main server to the crypto co-processor to different cores of the main server.

4)-The multiprocessing system will encrypt all the three files parallel, making the multiple encryptions faster.
5)-To encrypt the file main server will select a particular algorithm from the algorithm core (AES or TWOFISH) present in the crypto co-processor.

6)-Based on the selected algorithm the operational controller instructs the co-processor to download the algorithm from the core and programs the FGPA.

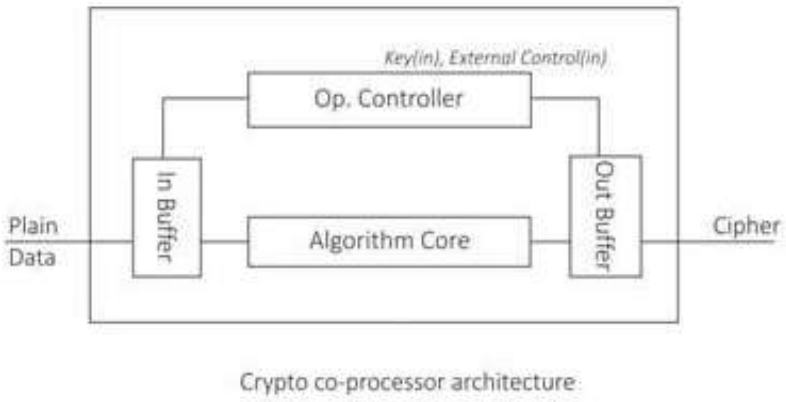


Figure-4: Cryptoco- processor architecture

7)-Now Files will be passed to the crypto co-processor for encryption (architecture of crypto co-processor can be seen in Figure-4.

8)-As soon as any of the three files gets encrypted, an acknowledgement signal will be passed to the main server.

9)-In response, the main server will generate a packet containing-type of the algorithm used for encrypting this particular file, the sequence number of file (in reference of joining), the information about the network on which the file will be sent.

PHASE V: SENDING FILE TO DIFFERENT NETWORKS

In the third phase as soon as the encryption of splitted file is done, the server generates a packet containing-

1)-Type of algorithm used in the encryption of that particular file.

2)-File number in sequence of joining.

3)-The information about the network on which the file will be sent.

The packets generated will then be sent to the administrator, so that the administrator later on can recover all the files from all the networks, can decrypt them using proper algorithm and also join them in proper sequence.

PHASE VI: HACKER TRACING SYSTEM AFTER SUCCESSFUL ATTACK

Considering the worst case scenario i. e. the hacker is able to gather all files of the s stem then this phase provides us with additional feature of the system in which when hacker is trying to decrypt the file (in our case we are assuming hacker is trying to decrypt the file while being connected to the internet) which are encrypted by different algorithms each one has a specific decryption key if hacker tries to use multiple keys to decrypt a single file then a packet will be generated(asseeninFigure-5) containing the information about the IP of the hacker with the information of which network it is using for accessing internet so we can easily trace back the hacker and take required action. It is of strategic importance for our cyberspace Security to be able to trace back to the origin of an Internet Attack. However, it is particularly challenging due to the Evading techniques that attackers use: [8] IP spoofing and attacking across stepping stones. These methods are being mentioned in the paper [8] which has been used as a reference.

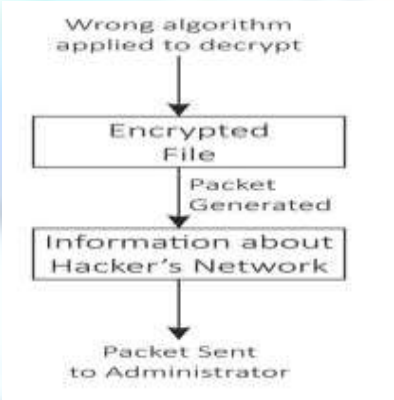


Figure-5: Packet Generated on Hacker's System

CONCLUSIONS

Data theft is always been a concerning factor in any field. This technique helps data protection in this method major concern is our confidential data should not be reached to hacker or attacker. By this methodology, when system tracks a security breach it under goes a secure method of splitting the file into different parts and this file initial under goes blow fish algorithm and then another phase of encryption which is chosen from random. These encrypted files are sent over different networks for protection. Con side ring a worst cases scenario even if attacker gets all file if even anyone of the file is decrypted by a wrong method the nth is would send in formation of the attacker to the main system.

REFERENCES

[1]. D. Schnetzer Fava, Characterization of Cyber Attacks Through Variable Length Markov Models, Rochester Institute of Technology, Rochester, NewYork, 2007.
[2]. S.Kim, S.Shin, H.Kim, K.H.K won, and Y.Han, Hybrid Intrusion Forecasting Framework for Early Warning System, IEICE, vol. E91D, 2008.
[3]. North American Electric Reliability Council, Threat Alert System and Cyber Response Guidelines for the Electricity Sector,

USA, 2002.
[4]. D. Bodeau, J. Fabius Greene, and R. Graubart, How Do You Assess Your Organization's Cyber Threat Level?, The MITRE Corporation ,USA, 2010.
[5]. PrabinRanjan Sahoo Tata Consultancy Services Prabin.Sahoo@tcs.com and Chetan Phalak Tata Consultancy Services Chetan.Phalak@tcs.com
[6]. **High Speed file splitter: A solution to address data split need for Parallel Processing**
[7]. Praveen Ram C1, Sreenivaasan G2 Department of Computer Science Engineering Rajalakshmi Engineering College, Anna University Chennai, India. Security as a Service (SasS).
[8]. Encrypting data with the Blow fish algorithm by Bill Gatliff Consultant, Gdbstubs Library
[9]. Internet Attack Traceback Cross-validation and Pebble Tracing by FangYu and David Lee
[10]. Study on the development of early warning model for Cyber Attack By So Jeong KIM Policy & Research Planning Department Electronics and Telecommunications Research Institute Daejeon, ROK NedMoran, "A CyberEarly Warning Model," Inside Cyber Warfare: Mapping the Cyber Underworld, O'REILLY, 2010, p.179
[11]. http://en.wikipedia.org/wiki/Blowfish_(cipher)
[12]. http://www.design-reuse.com/articles/5922/encrypting-data-with-the-blowfish-algorithm.html
[13]. Peter Gutmann-AnOpen-source Cryptographic Coprocessor University of Auckland, Auckland, New Zealandpgut 001@cs.auckland.ac.nz