More Durable And Secure Database Design Using Donut Architecture and Blum-Goldwasser Public Encryption

Semuil Tjiharjadi
Program Studi D3 Teknologi Informasi
Fakultas Teknologi Informasi, Universitas Kristen Maranatha
Jl. Prof. Drg. Suria Sumantri No. 65 Bandung 40164
Email: semuil.tjiharjadi@eng.maranatha.edu

Abstrak

Makalah ini menjelaskan tentang perancangan sistem database yang memiliki ketahanan yang lebih baik dari sistem database yang biasa dengan menggunakan arsitektur Donut. Arsitektur Donut ini dapat membagi data menjadi beberapa bagian dan kemudian mengirimkannya setelah data tersebut dienkripsi menggunakan algoritma Blum-Goldwasser. Uniknya sekalipun ada bagian data yang hilang, tapi sistem mampu mengkonstruksi kembali dan mengembalikan ke data asal tanpa kehilangan informasi.

Sistem ini dapat digunakan untuk mengirimkan data secara jarak jauh untuk mengontrol mesin yang sensitif namun memerlukan data akurat untuk melakukan aksinya. Data yang salah atau tidak lengkap dapat menyebabkan masalah, khususnya ketika data yang sudah dikirimkan diganggu oleh hackers, atau koneksi komunikasi yang buruk dan masalah komunikasi lainnya. Di sini diperlukan kemampuan lebih untuk menjamin keakuratan data dan penggunaan sistem arsitektur Donut ini merupakan solusi untuk memecahkan jenis masalah tersebut.

Sebagai hasilnya adalah sebuah sistem yang mampu melakukan enkripsi, pembagian, mengirim dan memulihkan data secara utuh jika faktor kehilangan di bawah 30%. Sistem juga dapat mendeteksi jika terdapat masalah pada validitas data.

Untuk pengembangan yang lebih lanjut adalah bagaimana menggunakan teknik untuk membagi database dan menyebarkannya pada beberapa server di tempat yang berbeda dan bagaimana untuk mengamankannya. Sistem juga harus mampu menggabungkan dan memulihkan data seutuhnya ketika beberapa server rusak dan kehilangan datanya.

***Keywords**: durable, secure, encrypt, integrity*

1. Introduction

Accurate data are needed to control machines. When remotely controlling sensitive machines, it needs a system that can make sure the integrity of perfect data are received from transmission. Beside, the confidential of the data also have to place in highest level. Hacker or cracker can use the weakness of confidential aspect to study data and break them to understand how the system works, and then it is easy for them to control the machine. Hacker or cracker can steal information, change information and create false information. This situation describes the importance of confidential aspect beside the importance of integrity aspect.

Data integrity and confidential are important aspects when sending data remotely to machine. There are two other important aspects when sending remotely, they are Validity and non-repudiation. Both of these aspects detect change of data and where is the data truly come from.

This research describes how to send remote data control that can fulfill confidential, integrity, validity and non repudiation aspects. The system have capability to reconstruct information even system loss up to 30% (parts of the loss information are not located beside it).

2. Donut Architecture

Information file in donut architecture is divide to several parts that overlapping each others. Each part has some information that the other parts also do. This procedure makes each parts can reconstruct new information when some of the information are lost.

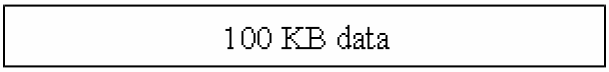
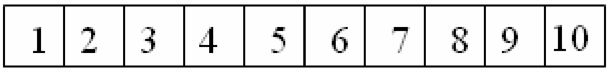


Figure 1: 100 KB information

Before system sends it over network, then the data divide into several pieces, for example 100 KB information divide into 10 pieces, each piece is 10 KB as a part of 100 KB information.

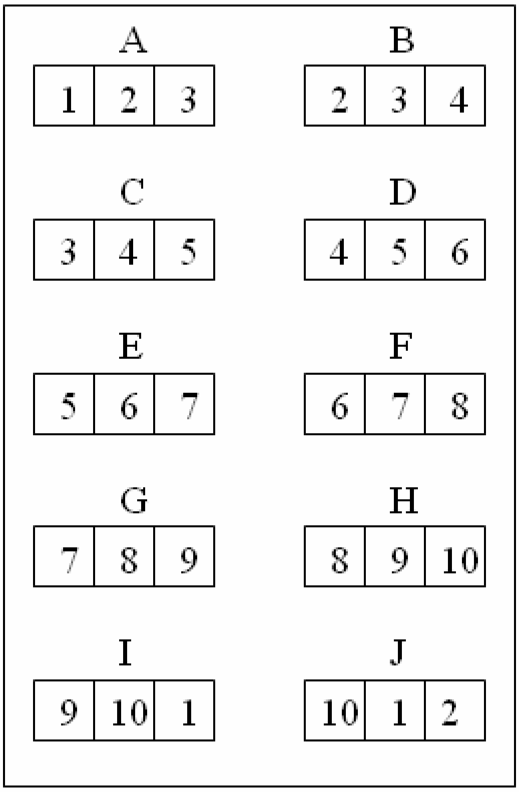


1	2	3	4	5	6	7	8	9	10

Figure 2: New forms of data after it has been divided

All pieces are part of groups that will be sent partly. Each group can be constructed by 2 or more pieces.

For example if each group is constructed by 3 pieces, then it will be like:



A	1	2	3
B	2	3	4
C	3	4	5
D	4	5	6
E	5	6	7
F	6	7	8
G	7	8	9
H	8	9	10
I	9	10	1
J	10	1	2

Figure 3: 3 Parts of information in each 10 Groups

In other way, the system can be described like a donut, each groups has some parts of data of other groups.

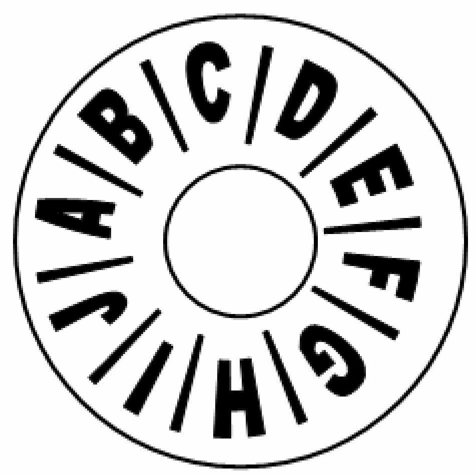


Figure 4: Donut architecture for distributed data

Donut architecture improves reconstruct ability when some of data lost and destruct integrity of their groups. Using donut architecture, system can use their part of data to develop the lost group. This can improve the integrity aspect of the whole information itself.

3. Cryptography

Cryptography is a study of technical mathematics that relevant with information security aspects like data validity, integrity and authentication. Or cryptography can be an art or study to keep security of a messages or information.

There are four main aspect in cryptography, they are :

1. Confidentiality, used to keep security of information from hacker or cracker.
2. Data Integrity, keeping data from illegal change of data. Keeping data integrity, system needs ability to detect manipulation like add, delete, and change of data.
3. Authentication connect with identification, sending information must be authenticated the originality and has been proved from sender and not from someone who pretend as sender.
4. Non-Repudiation, used to anticipate denial of action from sender.

3.1 Message and Encryption

A secret message is decoded by cryptography algorithm. That message is called plaintext and output of the cryptography algorithm is called ciphertext, the process converts plaintext to ciphertext is called encrypt, and the process converts ciphertext to plaintext is called decrypt.

Encrypt and decrypt are mathematical transform function. If message or plaintext is symbolized as M, ciphertext is symbolized as C, encrypt process is symbolized as E, decrypt process is symbolized as D, then mathematical notation of encrypt and decrypt process will be:

Encrypt : $E(M) = C$

Decrypt : $D(C) = D(E(C)) = M$

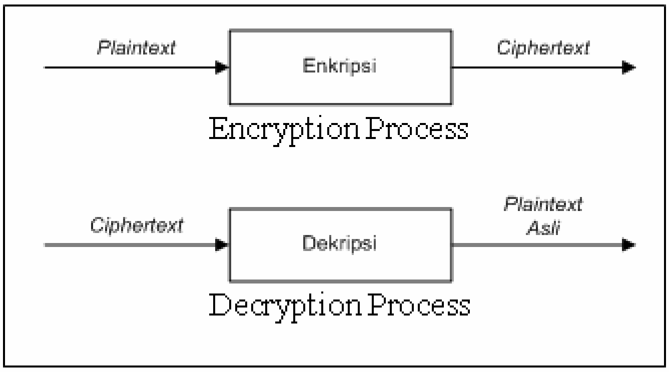


Figure 5: Encryption and Decryption process

3.2 Mathematical Theorem

Blum-Goldwasser Encryption Algorithm using some of mathematical theorems like prime numbers, modulus arithmetic operation, XOR operation exponential modulo n operation, Extended Euclidean algorithm, and quadratic residue modulo n algorithm.

3.2.1 Prime Numbers

A prime number is an integer more than one and only has 2 division factors, an integer more than one and itself. Cryptography often uses large prime numbers.
Each integer more than two can be created as:

$$n = p_1^{e1} * p_2^{e2} * p_3^{e3} * \ldots * p_k^{ek} \ ,$$

n = integer
p_i = *prime numbers*
e_i = *positif integer*
Example : n = 3458 = 2 * 7 * 13 * 19

3.2.2 Modulus Arithmetic Operation

Most computer programmers are familiar with modulus as a "remainder" operator, usually denoted by "%", which gives the remainder of an integer division instead of the quotient. For example: 27 % 12 = 3

Though the idea is the same, the mechanics here are slightly different from what mathematicians refer to as modulus arithmetic. In essence, modulus arithmetic consists of taking the infinitely long number-line and coiling it around a finite circle. All the numbers that land on the same point along the circle's edge are considered interchangeable, or congruent. Thus, the analogue to the above example in modulus arithmetic would be expressed as:

$$27 = 3 \ (\text{mod } 12),$$

or, in words: 27 is congruent to 3, modulo 12.

3.2.3 Euclidean Algorithm

Euclidean algorithm is used to find gcd (greatest common divisor). The Process of Euclidean algorithm is :

1. Chose two positive integer (a and b), and $a \geq b$
2. When $b \neq 0$ then do mathematical operation like:
   i. $r \leftarrow a \bmod b$
   ii. $a \leftarrow b$

```
   iii.   $b \leftarrow r$
   Repeat operations until $b = 0$
3. Greatest Common Divisor (gcd) is at $a$
```

Repeating process stops until $b = 0$, and result of gcd is at a.

3.2.4 Extended Euclidean Algorithm

Extended Euclidean algorithm is used to find two integers (x and y) that can fulfill this algorithm:

$$ax + by = d\ ,$$

a = integer
b = other integer
d = gcd(a,b)

Extended Euclidean algorithm:

```
1. Chose two positive integer numbers (a and b), and $a \geq b$
2. Count : $d = gcd(a,b)$
3. If $b = 0$ then :
   $d \leftarrow a$          ; $d = gcd(a,b)$
   $x \leftarrow 1$          ; $x = 1$
   $y \leftarrow 0$          ; $y = 0$
4. If $b > 0$ then process:
   i.  $q \leftarrow \lfloor a / b \rfloor$
       $r \leftarrow a - qb$
       $x \leftarrow x_2 - qx_1$
       $y \leftarrow y_2 - qy_1$
   ii. $a \leftarrow b$
       $b \leftarrow r$
       $x_2 \leftarrow x_1$
       $x_1 \leftarrow x$
       $y_2 \leftarrow y_1$
       $y_1 \leftarrow y$
   Process will repeat until $b = 0$.
5. $d \leftarrow a$          ; $d = gcd(a,b)$
   $x \leftarrow x_2$        ; x = value x
   $y \leftarrow y_2$        ; y = value y
```

3.2.5 Quadratic Residue Modulo n

'a' is quadratic residue modulo n when *x* is:

$$x^2 \equiv a \ (mod\ n)$$

With condition $0 < a < n$ and n is prime number or multiply result of two prime numbers.

Quantities of quadratic residue number depend on *n*. If n is prime number then quantity of quadratic residue numbers are:

$$(n - 1) / 2,$$

if *n* is result of p multiply by q then quantity of quadratic residue numbers are

$$(p - 1)\ (q - 1) / 4.$$

Example: $n = 7$

$1^2 = 1 \equiv 1 \ (\text{mod } 7)$

$2^2 = 4 \equiv 4 \ (\text{mod } 7)$

$3^2 = 9 \equiv 2 \ (\text{mod } 7)$

So *quadratic residue* modulo 7 are: { 1, 2, 4 }.

3.3 Blum-Goldwasser Encryption Method

Blum-Goldwasser encryption method was publicized first time in 1984, by its inventors, Manuel Blum and Shafi Goldwasser. Blum-Goldwasser encryption method is a public encryption, and it is using a pair of key. The first one is public key and the other is private key.

Blum-Blum-Shub generator was used to make pseudorandom bit sequence, then the process will be continued with XOR plaintext operation. The result of this process is ciphertext and it will be transmitted with pseudorandom bit by the sender.

Receiver will make pseudorandom bit sequence back using pseudorandom bit and private key, then it will process XOR operation with ciphertext to get message back.

Blum-Goldwasser encryption algorithm have 3 parts, they are: key generator algorithm, encryption algorithm, and decryption algorithm.

3.4 Key Generator Algorithm

Blum-Goldwasser encryption method has some keys, one public key and four private keys.

Key generator of Blum-Goldwasser encryption method process is:

- Choosing two different prime numbers randomly and each number must be congruent to *3 mod 4* operation. First number is notated as *p* and second number is notated as *q*.

$$p \equiv 3 \bmod 4$$

$$q \equiv 3 \bmod 4$$

- Counting *p* multiplying with *q*, and result is notated as *n.*

$$n = p * q$$

- Counting 2 integer using *Extended Euclidean algorithm*, first number is notated as *a* and second number is notated as *b*, and counting will continue until they can find the result as:

$$ap + bq = 1$$

n is public key that will tell to sender, and *p*, *q*, *a* and *b* are private key that will be kept secretly by receiver.

3.5 Blum-Goldwasser Encryption Algorithm

Encryption process need public key (*n*), sequence of the process is:

- Input *n*
- Count $k = log_2 n$
- Count $h = log_2 k$
- Show sending message (*plaintext*) as sequence of *binary string*:

$$m = m_1m_2...m_t$$

- Chose x_0 that is a *random quadratic residue modulo n*
- Use $(i = 1 \to t)$ and count :
  - $x_i = x_{i-1} \bmod n$
  - p_i, p_i is *h least significant bits* from x_i.
  - $c_i = p_i \oplus m_i$
- Count $x_{t+1} = x_t^2 \bmod n$
- Ciphertext is sequence of c that is added by x_{t+1}

$$c = (c1,c2,...,ct,x_{t+1})$$

3.6 Blum-Goldwasser Decryption algorithm

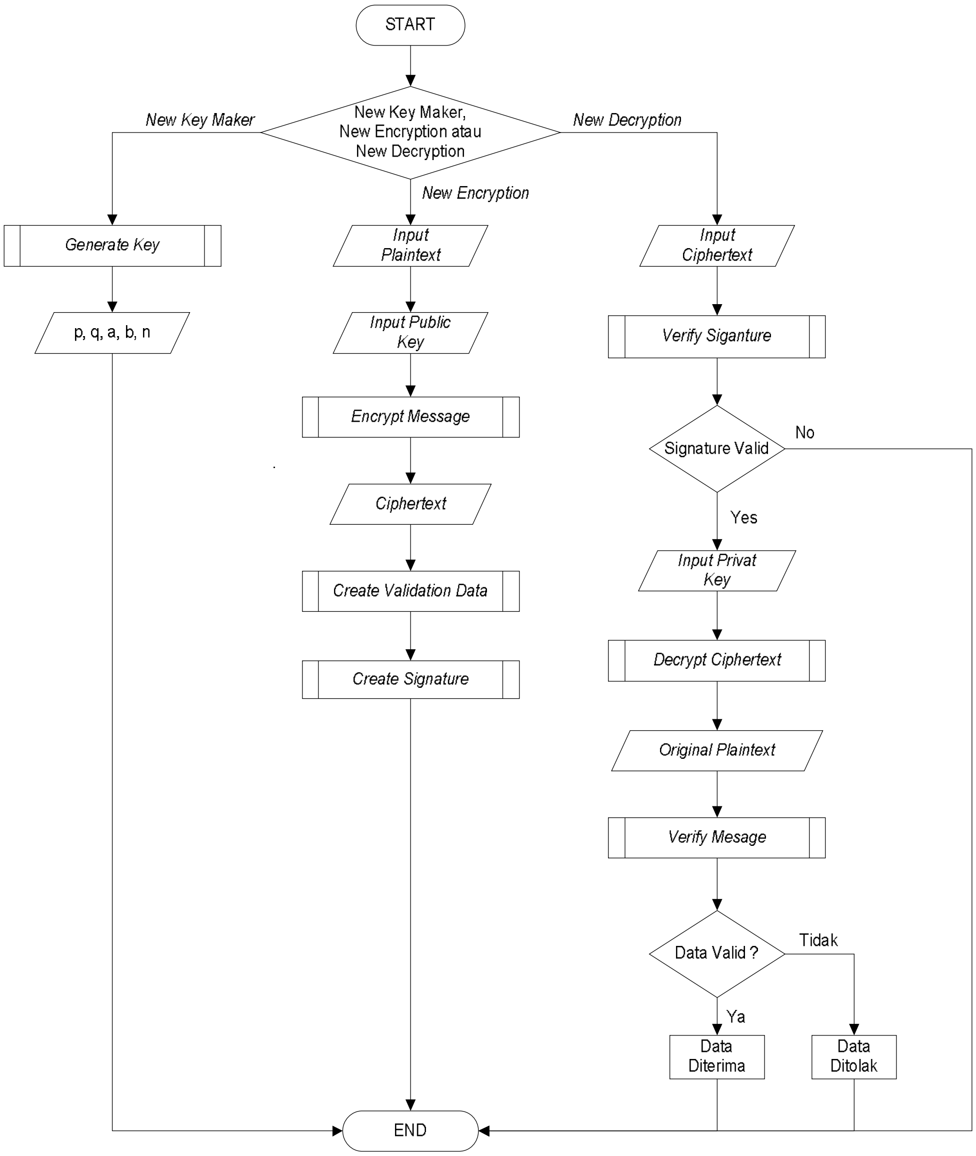


Figure 6: Blum Goldwasser Flowchart

Decryption process needs private keys (p, q, a and b), sequence of this process is:

- Input p, q, a and b.
- Count

$$d_1 = \left(\frac{(p+1)}{4}\right)^{t+1} \bmod(p-1)$$

- Count

$$d_2 = \left(\frac{(q+1)}{4} \right)^{t+1} \mathrm{mod}(q-1)$$

- Count **u** $= x_{t+1}^{d1} \bmod p$
- Count **v** $= x_{t+1}^{d2} \bmod q$
- Count $x_0 = (vap + ubq) \bmod n)$
- Use $(i = 1 \rightarrow t)$, and count :
- $x_i = x_{i-1}$ *mod n*
- p_i, p_i is *h least significant bits* from x_i.
- $m_i = p_i \oplus c_i$

After decryption process finishes then the output is plaintext of the message.

3.7 Program

The Program has 9 subprograms, which are:

1. Donut architecture program
2. Key generator program
3. Encryption program
4. Validation making program
5. Digital Signature program
6. Digital Signature check program
7. Decryption program
8. Validation check program
9. Data remote control system.

4. Analysis

Analysis process for donut architecture has been done by using 3 database file: suppliers.db, products.db, and customers.db. Suppliers.db was divided into 3 files: suppliers1.dds, suppliers2.dds, and suppliers3.dds. products.db was divided into 4 files: products1.dds, products2.dds, products3.dds, and products4.dds. Customers.db was divided into 5 files : customers1.dds, customers2.dds, customers3.dds, customers4.dds and customers5.dds.

4.1 Dividing Process Analysis

Dividing processes have been done to 3 files: Suppliers.db, Products.db, and also customers.db, and all of them have succeed

Table 1 Dividing process Analysis Table

No	Source file	Size (kb)	Result dividing file	Size (kb)	Process status
1	Suppliers.db	6,243	Suppliers1.dds	3,148	Success
			Suppliers2.dds	3,132	Success
			Suppliers3.dds	3,128	Success
2	Products.db	345	Products1.dds	121	Success
			Products2.dds	118	Success
			Products3.dds	117	Success
			Products4.dds	118	Success
3	Customers.db	12,345	Customers1.dds	4,192	Success
			Customers2.dds	4,186	Success
			Customers3.dds	4,188	Success
			Customers4.dds	4,186	Success
			Customers5.dds	4,184	Success

4.2 Durability Process Analysis

4.2.1 Durability process Analysis to reform Suppliers.db

Table 2. Durability process Analysis to reform Suppliers.db

Test number	Suppliers1.dds	Suppliers2.dds	Suppliers3.dds	Status
1	unavailable	available	available	success
2	available	Unavailable	available	success
3	available	available	Unavailable	success
4	Unavailable	Unavailable	available	Fail
5	Unavailable	available	Unavailable	Fail
6	available	Unavailable	Unavailable	Fail

Systems need at least 2 files to reform suppliers.db, so its durability ratio = 66%

4.2.2 Durability process Analysis to reform Products.db

Table 3. Durability process Analysis to reform Products.db

Test number	Products1. dds	Products2. dds	Products3. dds	Products4. dds	Status
1	Unavailable	available	available	available	success
2	available	Unavailable	available	available	success
3	available	available	Unavailable	available	success
4	available	available	available	Unavailable	success
5	Unavailable	Unavailable	available	available	Fail
6	Unavailable	available	Unavailable	available	Fail
7	Unavailable	available	available	Unavailable	Fail
8	available	Unavailable	Unavailable	available	Fail
9	available	Unavailable	available	Unavailable	Fail
10	available	available	Unavailable	Unavailable	Fail

Systems need at least 3 files to reform products.db, so its durability ratio = 75%

4.2.3. Durability Process Analysis to Reform Customers.db

Table 4. Durability process Analysis to reform Customers.db

Test number	Customers 1.dds	Customers 2.dds	Customers 3.dds	Customers 4.dds	Customers 5.dds	Status
1	unavailable	available	available	available	available	success
2	available	unavailable	available	available	available	success
3	available	available	unavailable	available	available	success
4	available	available	available	unavailable	available	success
5	available	available	available	available	unavailable	success
6	unavailable	unavailable	available	available	available	success
7	unavailable	available	unavailable	available	available	success
8	unavailable	available	available	unavailable	available	success
9	unavailable	available	available	available	unavailable	success
10	available	unavailable	unavailable	available	available	success
11	available	unavailable	available	unavailable	available	success
12	available	unavailable	available	available	unavailable	success
13	available	available	unavailable	unavailable	available	success
14	available	available	unavailable	available	unavailable	success
15	available	available	available	unavailable	unavailable	success
16	unavailable	unavailable	unavailable	available	available	Fail
17	unavailable	unavailable	available	unavailable	available	Fail
18	unavailable	unavailable	available	available	unavailable	Fail
19	unavailable	available	unavailable	unavailable	available	Fail
20	unavailable	available	unavailable	available	unavailable	Fail
21	unavailable	available	available	unavailable	unavailable	Fail
22	available	unavailable	unavailable	unavailable	available	Fail
23	available	unavailable	unavailable	available	unavailable	Fail
24	available	unavailable	available	unavailable	unavailable	Fail
25	available	available	unavailable	unavailable	unavailable	Fail
26	unavailable	unavailable	unavailable	unavailable	available	Fail
27	unavailable	unavailable	unavailable	available	unavailable	Fail
28	unavailable	unavailable	available	unavailable	unavailable	Fail
29	unavailable	available	unavailable	unavailable	unavailable	Fail
30	available	unavailable	unavailable	unavailable	unavailable	Fail

Systems need at least 3 files to reform Customers.db, so its durability ratio = 80%

5. Conclusions

This article presented donut architecture as a good method to increase data integrity and extremely improve survival ability of data information.

Blum-Goldwasser encryption and decryption program able to send data that fulfill confidential factor, validity factor and integrity factor.

There are some unexplained characters of Blum-Characteristic Encryption and Decryption Method like a character in plaintext can be encrypted several times but will have different ciphertext. The other is size of public key in Blum-Goldwasser encryption and decryption method is not effect size of ciphertext.

References

[Bel92] Bell, David dan Jane Frimson. (1992) *Distributed Database Systems.* Addison-Weslye Publishing Company.

[Dou96] Douba, Salim. (1996). *Networking UNIX.* Sams Publishing.

[Gan01] Ganger, Gregory R. (2001). *Survivable Strorage System.* Carnegie Mellon University.

[Men96] A.Menezes, P.van Oorschot, and S.Vanstone. (1996) *Handbook of Applied Cryptography.* CRC Press.

[Ros03] Rosen, Kenneth H. (2003). Discrete Matehematics and Its Applications. 5th edition. McGraw-Hill.

[Sch96] Schneier, Bruce. (1996). *Applied Cryptography second edition.* John Wiley & Sons. Inc.

[Str97] Strunk, John D. (1997). *Self-Securing Storage: Protecting Data in Compromised Systems.* Carnegie Mellon University. October 2000.Pressman. Roger S.. *Software Engineering*. edisi keempat. McGraw Hill.1997.

[Wyl01] Wylie, Jay J. (2001). Selecting the Right Data Distribution Scheme for a Survivable Strorage System. Carnegie Mellon University. Mei 2001.