International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org ||Volume 7 Issue 5 Ver. II || May 2018 || PP 79-85

Smart Backup and Recovery In Cloud Computing

Dr. Divya Kashyap, Mohammad Amin Samandari
Assistant ProfessorIndian Academy Degree College AutonomousBangalore, India
Master of Computer ScienceIndian Academy Degree College AutonomousBangalore, India
Corresponding Auther: Dr. Divya Kashyap

Abstract—*Data Backup and Recovery is the major point in computing either in pervious (traditional) computing, this functionality was the basement for reliability of system. People are worry of their data they don't want to lose it in anyway specially in case of unexpected natural phenomenon (storm, flood etc.…). Backup and Recovery has increased the use of both traditional and cloud computing users can reclaim their lost data in case of lost or unexpected changes. Different algorithms are being used to provide this functionality.*

Keywords— *Backup, Recovery, Remote Server, Repository, ciphering, encryption Introduction*

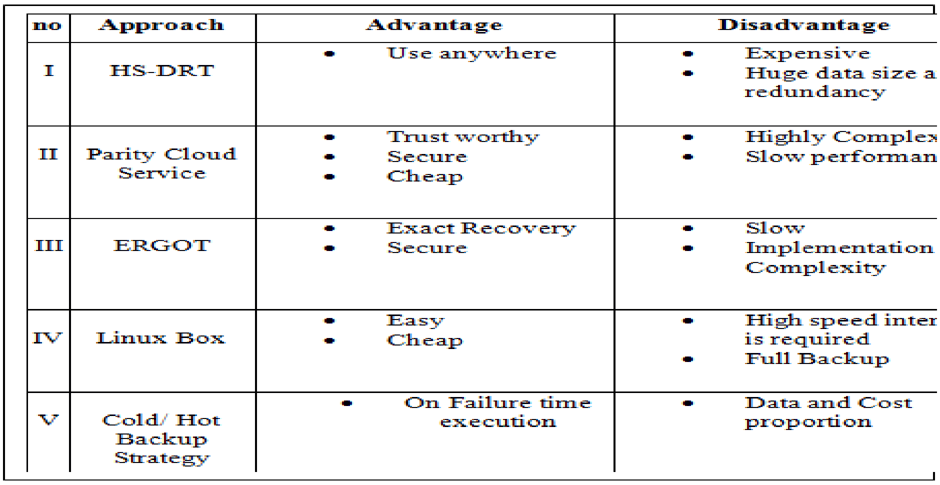
Date of Submission: 07-05-2018 Date of acceptance: 22-05-2018

I. Introduction

Companies are using computer technology as their database and a resource to information the amount of data is increasing day by day, users are unable to curry it with them self, due to the existence of chances of losing it forever or due to its huge amount of data, from other hand, keeping additional servers for backup and recovery may increase the headache of company and it may also increase the expenses Security, Backup Storage and Recovery is on the top of the challenges which companies are facing with respectably[1] . Cloud computing come over this problem as (saving money is the top reason for using the cloud. Cloud offers a variety of benefits such as the ability to do backups more frequently, ease of management, and redundancy, but the top benefit cited by 61% was saving money on storage, followed by 50% who stated savings on administration costs [1]), also providing mobile access to not only data as well as keeping the backup, the availability of data is as important as its safety, the services provider must guaranty this feature. Backup and Recovery is the main point which make cloud computing more reliable, as our data is lost or damage due to different reasons it should be located back to user in minimum time period and data consistency is remained in the services. Cloud computing has the ability to connect the Internet and extensive network in order to use resources that are available remotely, and as a result, present efficient solutions based on pay-per-use [6].

Nowadays almost all business depends on the usage of technology which a company is, and it plays an important role in success. From the data access point of view the cloud computing has come over all the disadvantages of pervious computing system, no need for the knowledge of understanding the entire infrastructure of the network; the only need is the internet but data Security is the most negative effect which denies business to trust in IT, especially in case of putting their data in location which is far from their eyes and physical access of business partners, a secure and easy environment is needed.

A network basically provides three main functionalities a) availability b) integrity c) confidentiality. Our preferred algorithm assures availability and integrity in a very accurate manner without putting any load to any node of the network and to maintain confidentiality for security issues the service provider must use the latest and strongest data encryption method. We have explained a useful type of encryption called Advance Encryption Standards (AES) as an eliminator of security issues. Different algorithms like Linux Box, Cold/Hot Backup Strategy etc. are being used for backup and recovery purpose in cloud services, we'll see further some of them with their advantages and disadvantages, all of these algorithms are providing Backup and security but they are not 100% fulfillment of user's side as well as the provider's side.



no	Approach	Advantage	Disadvantage
I	HS-DRT	• Use anywhere	• Expensive • Huge data size a redundancy
II	Parity Cloud Service	• Trust worthy • Secure • Cheap	• Highly Complex • Slow performan
III	ERGOT	• Exact Recovery • Secure	• Slow • Implementation Complexity
IV	Linux Box	• Easy • Cheap	• High speed inter is required • Full Backup
V	Cold/ Hot Backup Strategy	• On Failure time execution	• Data and Cost proportion

Table (1): Comparing different Approaches of Cloud Backup and Recovery

II. CURRENT SYSTEM

In the existing system data owner sends data to the cloud server using bidirectional way. To access the data from the Cloud server, the users have to be registered with the cloud server. So that the user have to register their details like username, password and a set of random numbers. This information will store in the database for the future authentication. Once the Data Owner registered in cloud server, the space will be allocated to the data owner. Providing a high level security for Cloud server is optional. If at all security exists, the third party auditor may access the entire data packets for verification. Third party auditors are used by clients and providers to determine the security of the cloud implementation. The main drawbacks can be summarized as:

1. There is no automatic process of backup at system, either the service provider or the client must maintain a separate backup manually. It is very difficult to manage large backup data manually.
2. If anyone wants to take backup for data, he/she have to take a full backup of overall data which is not efficient in-terms of network speed and storage cost.
3. By default there is no security for data travel on channel so there is a chance of leakage of data.

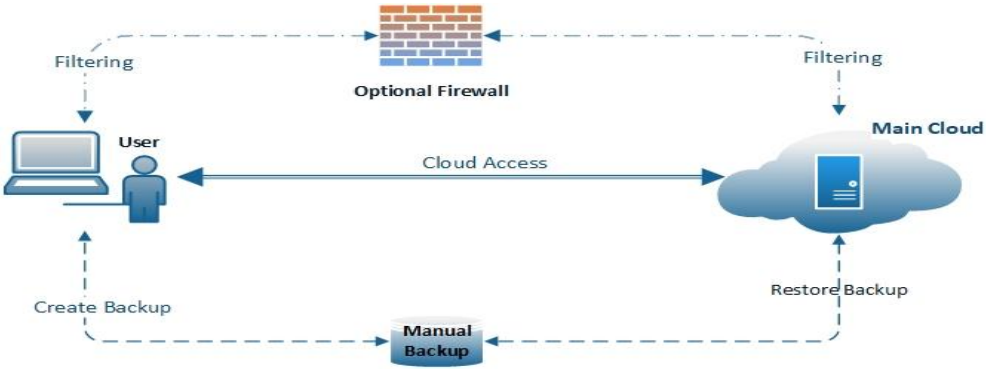


Figure (1): show the existing Cloud System

III. PROPOSED SYSTEM

Our suggested system is using deferential backup and recovery, a differential backup, backs up only data that has changed or newly created since previous backup. Differential backups are cumulative, which aggregates all changes from previous backup. The data size on backup media after a differential backup grows day by day. In case of restoration, incremental backup requires each and every backup media conversely differential backup requires only the latest differential backup media for a full restoration [3][4]. From the user point of the view the primary cloud is the primary cloud backup, if the primary server is offline or damaged, the remote server (backup) contains all the states of the primary server and it'll act as primary server. As the main goal of cloud is to provide remote access for its users so remote server is being access by users which is also called as Remote Data Backup or Main Repository of the system.

As mentioned above we suggest using SBA for Backup and recovery and SEA for data encryption so we can come over the draw bags of current system in cloud system as:

1. The system will have automatic backup and recovery without any interruption of user of service provider which is accessible globally.
2. The data will be backup from the last changes have been done by user or only new of altered data will be backed up or will be recovered.
3. SEA can encrypt data while travelling through network and the algorithm make our network more secure and there is no need for any third party security system like: firewall.
4. Our proposed system can be applied in a very low cost and to use it doesn't need any network professional from the client side.

IV. Remote Server

A server which contains all the data of main cloud and acts as main server in case the main cloud lost its data or is unavailable, it's located in a remote location far from main zone. When this Backup server is at remote location (i.e. Far away from the main server) and having the complete state of the main cloud, then this remote location server is termed as Remote Data Backup Server [14]. If a main cloud i.e. if central repository lost its data then it uses the information from remote repository, the very clear objective is to help clients to collect information from remote repository if network connectivity is not available or the main cloud is unable to provide data to the clients [7]. It also called as disaster recovery; a disaster is an unexpected event in a system lifetime. It can be made by nature (like the tsunami and earthquake), hardware/software failures cloud-based DR solution is an increasing trend because of its ability to tolerate disasters and to achieve the reliability and availability [10]. Disaster recoveries as a service is free or pay on use offer. When incompatibilities are occurred due to software changes then breaking of DRaaS in cloud may occur. [13]

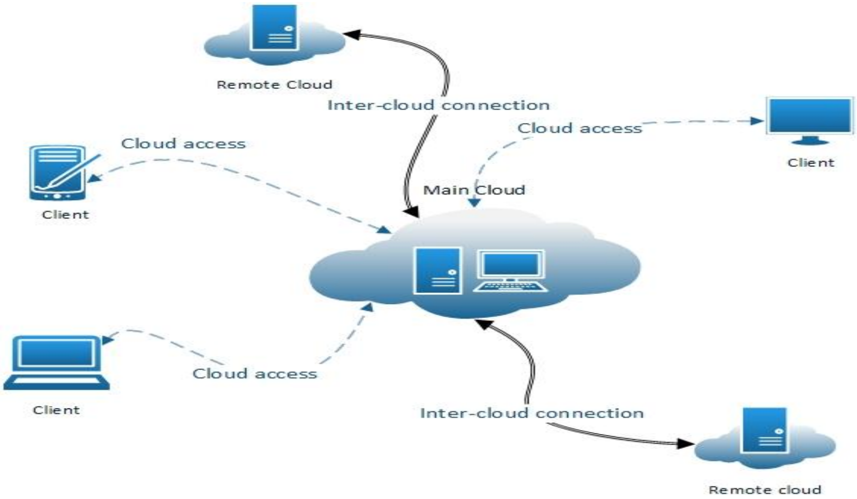


Figure (2): shows the structure of Remote backup servers or DR

Also the suggested system can help as to maintain the following characteristic of a network:

- Integrity: all data which is being transformed to the remote backup must be the exact one in the primary server. Also the verification of data is being checked by integrity.
- Security: Remote backup server is the full responsibility of providing the security for the data, its very import to allow only those users who is fully authorized to the system.
- Confidentially: if the no. of users whose are using cloud is being increased, the accessed data must be hide from the rest of user, we can come over this issue by using different techniques like data encryption, restricting users to access only some specific folders etc….
- Cost efficiency: cloud computing is not limited to only some lightweight no of data, increasing the amount of data should not decrease the efficiency of the services. Taking backup and recovering it must be as possible as easy.[5][8][11].

V.Algorithm and Architecture

When you develop a comprehensive strategy for backup and restoring data, you must first identify the failure or disaster situations that can occur and their potential business impact. In some industries, you must consider regulatory requirements for data security, privacy, and records retention [2]. This algorithm focus on simple mechanism of the back-up and recovery process there is no need for such a complex analysis to apply. It mainly uses the idea of Exclusive–OR (XOR) method of the computing world. E.g. - If suppose there are two data files: P and Q. When we XOR P and Q it produced X. i.e. X =P Ex-OR Q ($X = P \oplus Q$) If suppose only a part P data file get deleted and we want that exact data file back then we are able to get only that exact data file from a huge amount with the help of Q and X data file .i.e. P = X Ex-OR Q (P = X⊕Q) [5].

P	Q	X
1	0	1
0	1	1
1	1	0
0	0	0

Table (2): XORing and its output [8]

The seed block algorithm providing simple and fast backup recovery process in backend, as shown in Fig:2 Main-cloud, Remote-server and client is making the echo system of our concept, by registering of each client they are getting a SeedBlock (a SeedBlock is the XORed of a unique id of a particular client and a random number), the SeedBlock is then being saved in remote server, When clients upload data in cloud its being saved in Main-cloud, then the uploaded data is being XOR with SeedBlock and the result is being saved in remote server as a file' (file dash), this file' is then used to recover the damaged files from remote server by XORing it back with SeedBlock of particular client.
The architecture representation of the Seed Block Algorithm is shown in the Fig.3.

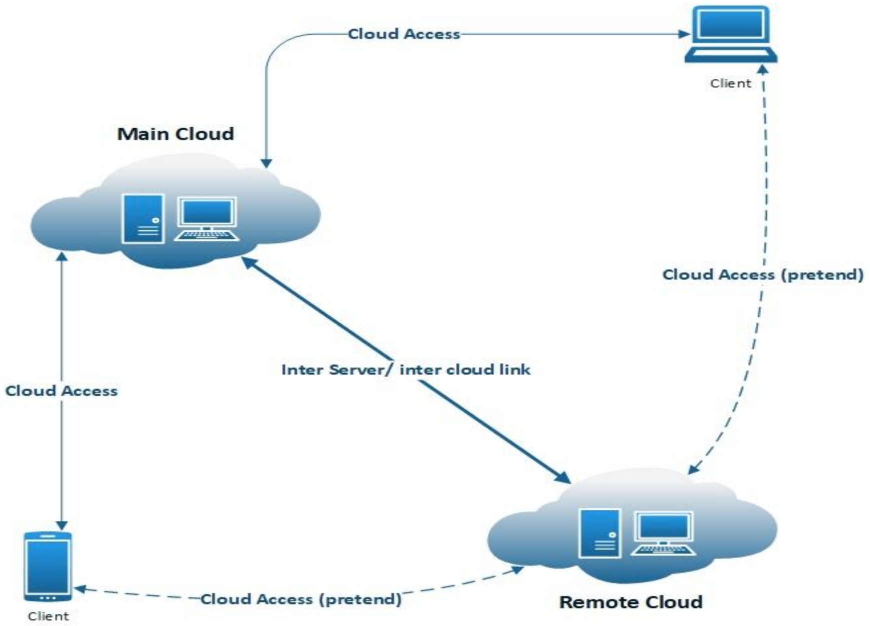


Fig (3): Shows the Structure of SBA Backup

ADVANTAGES OF PROPOSED SYSTEM
1. Recovery of same sized data
2. Privacy is the primary objective
3. Low cost implementation

VI. Used Algorithm

Initialization:
Main Cloud: M_C
Remote Server: RS
Clients of Main Cloud: Ci
Files: a1 and a'1
Seed block: Si
Random Number: r
Client's ID: Client ID
Input: a1 created by Ci, r is generated at M_C
Output: Recovered file a1 after deletion at M_C
Given Authenticated clients could allow uploading, Downloading and do modification on its own the files only.
Step 1:

- Generate a random number.

```
•  int r = rand();
       Step 2:
•  create a seed Block Si for each Ci and Store  Si at RS
•  Si = r⊕Client_ID.
•  (Repeat step 2 for all clients)
       Step 3:
•  If Ci/Admin creates/modifies a and  stores a1 at MC then
•  a'1 create as a'1=a1⊕Si
       Step 4:
•   Store a'1 at RS
       Step 5:
•  If server crashes a1 deleted from MC then
     o  we do EXOR to retrieve the original a1 as: a1 = a'1⊕Si
   Step 6:
     •  Return a1 to Ci
   Step 7: STOP. [9][12]
```

VII. ADVANCED ENCRYPTION STANDARD ALGORITHM (AES):

AES is a non-Feistel cipher, that is, AES uses only invertible components. A component is the plaintext has the corresponding component is the cipher. It is a symmetric-key block cipher which uses a single secret key for both encryption and decryption, AES is designed as substitution cipher to be resistant to exhaustive search attack. it has defined versions with 10, 12 and 14 rounds each version uses a different cipher key size 128, 192 and 256 but the round keys are always 128 bits (A round is a recursive cipher where each recursion is a combination of transposition units, substitution units and other components used to achieve diffusion and confusion). Any observation/repetition of patterns in the plaintext can be eliminated/avoided with the presence of strong diffusion and confusion provided by AES. The algorithm is more advantageous compared to DES, due to lack of weak keys. it uses certain types of transformation to provide higher level of security: substitution, permutation and mixing & key adding. We have 3 different AES versions AES-128, AES-192, AES-256 to enhance security mechanism. In the versions AES 128, 192 and 256 the words are generated in group of 4, 6 and 8 respectively, this key expansion `mechanism in AES has designed to provide feature that block the cryptanalyst. The algorithm is a follow:

Initialization:

```
1.  String encrypt (String Data)
2.  Key key=generate key();
3.  Cipher c=Cipher get Instance (ALGO);
4.  c.init(cipher,encrypt.MODE,Key);
5.  byte[] enval=c.dofinal(Data.getbytes());
6.  String encryptedValue= new BASE64Encode().
7.  encode(encVal);
8.  Return encrypted Value
9.  Key generatekey key=new Secret keyspec(keyValue.ALGO)
10. returnkey;
```

VIII. IMPLEMENTATION AND RESULT:

During the implementation, we observed that the size of original file uploaded by the client over the cloud is exactly same as size of backup file which is stored at the remote server and is same about the different types of file. So we can say that sba is capable to maintain the size of recovered file compare to the original file. Sba recovered the file without data loss.

Type	**Original file size**	**Backup Size**	**Recovered file size**
Text	300KB	300KB	300KB
	5MB	5MB	5MB
Image	630KB	630KB	630KB
	15MB	15MB	15MB

Table (2): shows the performance of SBA

Also we have analyzed the speed of proposed system:

Implementation:

As suggested algorithm to implement it on a system we need the following:

1. Data owner
2. Main cloud server
3. Data splitting and encryption
4. Key server
5. Replica server
6. Parity bit addition

1) Data Owner

A person who is going to put data in the cloud is called the owner and for doing this a user must register him/her self to the cloud to allot the space for him/her self

2) Main cloud server

It's the main server which is going to serve the clients as per there requests. This server has the responsibility of forwarding the request of client to the data owner and it also it has the responsibility of maintaining the information related to data owner and the users/clients.

3) Data splitting and encryption

Putting all data into a single location maybe harmful for the owner of the data in case of hacking the system the hacker should not reach to full data and along with that if he found any part of data, encryption is another firewall of getting data.

Combination of these both functionality is creating a good security of our cloud.

4) Key server

In our encryption algorithm a key is the magic of everything like backup, recovery and encryption, to do so a fast and specific server is needed for quick transactions, this type of provider is called key server.

5) Replica Server

Replica server is also called as remote server or disaster recovery server, the need of the server is come in practice when the main cloud lose the data and we have to find our exact data from this server.

6) Parity bit addition

One data has been uploaded to main cloud and a key has been generated for it and saved in key server, a parity bite is also being add with data as a checkpoint of last changes and for future backup and recovery.

IX. Conclusion

As per study of deferent algorithms for backup and recovery in cloud computing we found that, it's a big pressure on network as well as on user to replace all existing data in server after each change also its time consuming. To overcome the problem, we have suggested and algorithm which only need to backup the new altered data and in case of file damage it recovers only those part of data which is lost. Data security is another problem to which the major challenge of all network types in is general, to overcome the problem of spending a lot of budget on network security tools (e.g: Firewall) we the encryption of data while traveling in the network using SEA encryption algorithm to provide data security and maintain confidentially. Also, the need of putting an external backup server is removed.

References:

[1]. ChristineMurray,"Cloud_Backup_And_Disaster_Recovery_Meets_NextGeneration_Database_Demands", A Forrester Consulting thought leadership paper, Commissioned by Microsoft, March'2014.

[2]. Ruth, Andy - Backup_and_Recovery_Approaches_Using_AWS, https://aws.amazon.com/backup-recovery/, 2016

[3]. "What is differential backup? - Definition from WhatIs.com," SearchDataBackup. [Online]. Available: http://searchdatabackup.techtarget.com/definition/differential-backup. [Accessed: 03-Sep-2015]

[4]. Akash Kaveti - Effectiveness of Backup and Disaster Recovery in Cloud - http://www.diva-portal.se/smash/get/diva2:861846/FULLTEXT01.pdf,

[5]. Ms. Ashwini S. Gharde, Ms. Kamini Ghaormare, "Study of Seed Block Algorithm in Cloud Computing Environment", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727 - PP 37-39

[6]. Mostafa Ghobaei-Arani, "Fault-Tolerance Techniques in Cloud Storage: A Survey ,International Journal of Database Theory and Application", Vol.8, No.4 (2015), pp.183-190

[7]. Ms. Ashwini S. Gharde, Ms. Kamini Ghaormare, "Study of Seed Block Algorithm in Cloud Computing Environment", IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727 - PP 37-39

[8]. Ms. Kruti Sharma - Prof. Kavita R Singh, "Seed Block Algorithm: A Remote Smart Data Back-up Technique for Cloud Computing", International Conference on Communication Systems and Network Technologies by IEEE, 2013

[9]. Ms. Kruti Sharma - Prof. Kavita R Singh, "Seed Block Algorithm: A Remote Smart Data Back-up Technique for Cloud Computing", International Conference on Communication Systems and Network Technologies by IEEE, 2013

[10]. Mohammad Ali Khoshkholghi1, Azizol Abdullah1, Rohaya Latip1, Shamala Subramaniam1 & Mohamed Othman1, "Disaster Recovery in Cloud Computing: A Survey", Computer and Information Science; Vol. 7, No. 4; 2014 ISSN 1913-8989 E-ISSN 1913-8997 Published by Canadian Center of Science and Education,

[11]. Vijayalaxmi V Kadlimatti1, Ramesh Kumar H K2, “SEED BLOCK ALGORITHM: A REMOTE SMART DATA BACK-UP TECHNIQUE FOR CLOUD COMPUTING”, International Journal of Science, Technology & Management Volume No 04, Special Issue No. 01, April 2015 ISSN (online): 2394-1537

[12]. Kakad Umesh, Kankhar Mahesh ,Mysore Ajay , Nitin Rathee, “BACKUP AND RECOVERY SYSTEM USING SEED BLOCK ALGORITHM”, International Journal of Advanced Research in Computer Engineering& Technology (IJARCET) Volume 4 Issue 1, January 2015 pp 125-128

[13]. Mr.Akshay A. Gharat, Mr. Devendra E. Mhamunkar, “Disaster Recovery in Cloud Computing”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 5, May 2015 pp 1796-1801

[14]. Vedashree N, Praveen Kumar KC, Anilkumar G, “Data Recovery in Cloud Environment Using Seed Block Algorithm”, Vedashree N et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (5) , 2015, 4593-4598, ISSN: 0975-9646 pp 4593-4598

[15]. Kolipaka Kiran1, Janapati Venkata Krishna 2, “Smart Data Back-up Technique for Cloud

[16]. Computing using Secure Erasure Coding”, International Journal of Computer Trends and Technology (IJCTT) – volume 16 number 3 – Oct 2014, ISSN: 2231-2803 pp 85-89

[17]. Karishma Nadha, Sushma Somani, “Backup of real time data and Recovery using Cloud computing”, International Journal of Engineering Development and Research, Volume 4, Issue 2, ISSN: 2321-9939, pp: 138-143, 2016.

[18]. 1Karishma Nadhe, 2Sushma Somani, “Backup of real time data and Recovery using cloud computing”, International Journal of Engineering Development and Research - © 2016 IJEDR | Volume 4, Issue 2 | ISSN: 2321-9939, pp: 138-143.

[19]. Kruti Sharma, Kavita R Singh, “Online Data Backup and Disaster Recovery Techniques in cloud computing: A Review”, International Journal of Engineering and Innovative Technology (IJEIT) Volume 2, Issue 5, November 2012 - ISSN: 2277-3754, ISO 9001:2008 Certified, pp: 249 -254.

Dr. Divya Kashyap "Smart Backup and Recovery In Cloud Computing "International Journal of Engineering Science Invention (IJESI), vol. 07, no. 05, 2018, pp 79-85