ĐẠI HỌC QUỐC GIA HÀ NỘI

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

****

Vương Thị Hải Yến

NGHIÊN CỨU GIẢI PHÁP NÂNG CAO AN TOÀN BẢO MẬT CHO DỮ LIỆU ĐÁM MÂY

CHUYÊN NGÀNH: KỸ THUẬT PHẦN MỀM

MÃ SỐ: 60480103

TÓM TẮT LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC:

TS. LÊ QUANG MINH

Hà Nội - 2017

MỞ ĐẦU

1. Lý do chọn đề tài

Ngày nay, cùng với sự phát triển không ngừng của internet, các dịch vụ lưu trữ đám mây như Google Drive, Dropbox, SugarSync, Amazon Cloud Drive, Box, Mimedia (m) Drive, Skydrive, SpidekOak… cũng đang được sử dụng ngày càng rộng rãi bởi những tính năng sao lưu, lưu trữ dữ liệu trực tuyến với khả năng đồng bộ theo thời gian thực và tự động thực hiện sao lưu chia sẻ toàn bộ thư mục mà mình muốn, nó còn cho phép người sử dụng quay trở lại quá khứ để khôi phục những dữ liệu bị xóa hoặc bị thay đổi… Thêm vào đó, nhà cung cấp thường cho người dùng một số gói miễn phí hoặc với chi phí giá rất rẻ, thuận tiện trong việc cài đặt và sử dụng đối với các cá nhân và đơn vị nhỏ. Vì vậy số lượng sử dụng dịch vụ ngày càng tăng. Điều này đòi hỏi các dịch vụ trên phải tạo lập được uy tín, đảm bảo độ bảo mật và an toàn cho dữ liệu sử lưu trữ trên đó. Tuy nhiên, đây là chương trình lưu trữ tự động trên một máy chủ, tính bảo mật dữ liệu chưa thể khẳng định được, không thể chắc chắn thông tin có bị đánh cắp hoặc lộ bí mật hay không.

Chính vì vậy đề tài ***Nghiên cứu giải pháp nâng cao an toàn bảo mật cho dữ liệu đám mây*** được lựa chọn với mong muốn có thể là một tài liệu bổ ích để có thể giúp người phát triển hiểu kỹ hơn về khái niệm, lợi ích và những vấn đề liên quan đến lưu trữ đám mây. Ngoài ra, đề tài cũng sẽ nghiên cứu và xây dựng một giải pháp nhằm nâng cao tính an toàn bảo mật cho dữ liệu lưu trữ đám mây.

Trên cơ sở các nghiên cứu đã có, luận văn đã tập trung vào các mục tiêu và các vấn đề cần giải quyết sau:

2. Mục tiêu và phạm vi nghiên cứu

Luận văn tập trung nghiên cứu về điện toán đám mây, các vấn đề lưu trữ dữ liệu, an toàn dữ liệu trên điện toán đám mây; chỉ ra, phân tích những mặt ưu nhược điểm của các giải pháp đã được đưa vào sử dụng trong việc bảo vệ dữ liệu đám mây để làm rõ tính cấp thiết của đề tài. Đồng thời trình bày các phương pháp dự phòng nâng cao độ tin cậy của hệ thống. Sau đó, trình bày tổng hợp, phân tích kiến thức xoay quanh cơ chế RAID. RAID đối với bài toán an toàn dữ liệu cho hệ thống máy. Từ đó đề xuất chi tiết

giải pháp nâng cao an toàn dữ liệu lưu trữ trên đám mây, chứng minh độ tin cậy của giải pháp và vận dụng cụ thể vào doanh nghiệp Việt Nam hiện nay.

3. Phương pháp nghiên cứu

Trong luận văn này tôi đã kết hợp nhiều phương pháp nghiên cứu khác nhau phù hợp với yêu cầu của đề tài, bao gồm các phương pháp nghiên cứu sau:

- Phương pháp phân tích, tổng hợp
- Phương pháp nghiên cứu thực tiễn
- Phương pháp thực nghiệm

4. Kết quả đạt được

Từ mục tiêu nghiên cứu giải pháp đảm bảo an toàn dữ liệu lưu trữ trên điện toán đám mây, luận văn đã tập trung làm rõ được những lý thuyết cơ bản về điện toán đám mây, vấn đề bảo vệ dữ liệu trên điện toán đám mây hiện nay, chỉ ra những giải pháp đã được sử dụng trước đó và phân tích rõ những ưu điểm, hạn chế cần phải khắc phục; các phương pháp dự phòng nâng cao độ tin cậy của hệ thống. Sau đó, trình bày tổng hợp, phân tích kiến thức xoay quanh cơ chế RAID. RAID đối với bài toán an toàn dữ liệu cho hệ thống máy.

Đồng thời vận dụng cơ sở lý thuyết RAID vào việc giải quyết bài toán an toàn dữ liệu lưu trữ trên đám mây. Kết quả cuối cùng là luận văn đã đề xuất thành công giải pháp mới nâng cao an toàn dữ liệu lưu trữ trên đám mây, chứng minh thành công tính đúng đắn và hiệu quả và tính khả thi của giải pháp; đưa ra được quy trình cụ thể của việc ứng dụng giải pháp vào thực tiễn vào doanh nghiệp Việt Nam hiện nay.

5. Cấu trúc luận văn

Luận văn được trình bày trong 3 chương, với nội dung chính của mỗi chương như sau:

MỞ ĐẦU

Giới thiệu về lý do chọn đề tài, mục tiêu, phạm vi nghiên cứu và phương pháp nghiên cứu, tóm lược các kết quả đạt được.

Chương 1 - *Tổng quan về các phương pháp bảo vệ dữ liệu lưu trữ trên đám mây*

Chương này trình bày cơ sở lý thuyết về điện toán đám mây. Giới thiệu một số dịch vụ lưu trữ đám mây.

Lập luận dẫn chứng về những vấn đề mất mát dữ liệu, an toàn dữ liệu trong lưu trữ trên dịch vụ đám mây.

Trình bày và phân tích giải pháp đặc trưng mã hóa dữ liệu trong lưu trữ dữ liệu đám mây, ưu nhược điểm, những nhược điểm cần phải khắc phục để đảm bảo độ an toàn bảo mật cho dữ liệu đám mây.

Từ đó rút ra kết luận về tính cấp thiết, ý nghĩa thực tiễn khoa học của luận văn là giải quyết vấn đề bài toán đặt ra: "***Nghiên cứu giải pháp nâng cao an toàn bảo mật cho dữ liệu đám mây***".

Chương 2 - *Phương pháp dự phòng cấu trúc nâng cao độ tin cậy cho hệ thống lưu trữ dữ liệu*

Nêu các phương pháp dự phòng nâng cao độ tin cậy của hệ thống. Sau đó, trình bày tổng hợp, phân tích kiến thức xoay quanh cơ chế RAID. RAID đối với bài toán an toàn dữ liệu cho hệ thống máy.

Chương 3 - *Đề xuất giải pháp nâng cao an toàn bảo mật dữ liệu lưu trữ trên đám mây và ứng dụng vào thực tế doanh nghiệp*

Trình bày chi tiết giải pháp, phát biểu bài toán xác định và mô tả quy trình bài toán thực tế và đưa ra lập luận chứng minh độ tin cậy của giải pháp. Thực tế ứng dụng giải pháp vào doanh nghiệp.

KẾT LUẬN

Tóm lược kết quả chính của luận văn, đánh giá giải pháp mới đưa ra và so sánh với những giải pháp đã được sử dụng.

Chương 1. TỔNG QUAN VỀ CÁC PHƯƠNG PHÁP BẢO VỆ DỮ LIỆU LƯU TRỮ TRÊN ĐÁM MÂY HIỆN NAY

1.1. Khái quát về điện toán đám mây

1.1.1. Khái niệm

Điện toán đám mây (tiếng Anh: *cloud computing*), còn gọi là **điện toán máy chủ ảo**, là mô hình điện toán sử dụng các công nghệ máy tính và phát triển dựa vào mạng Internet. Thuật ngữ "đám mây" ở đây là lối nói ẩn dụ chỉ mạng Internet (dựa vào cách được bố trí của nó trong sơ đồ mạng máy tính) và như một liên tưởng về độ phức tạp của các cơ sở hạ tầng chứa trong nó. Ở mô hình điện toán này, mọi khả năng liên quan đến công nghệ thông tin đều được cung cấp dưới dạng các "dịch vụ", cho phép người sử dụng truy cập các dịch vụ công nghệ từ một nhà cung cấp nào đó "trong đám mây" mà không cần phải có các kiến thức, kinh nghiệm về công nghệ đó, cũng như không cần quan tâm đến các cơ sở hạ tầng phục vụ công nghệ đó [9,12].

1.1.2. Đặc điểm của điện toán đám mây

Những ưu điểm và thế mạnh dưới đây đã góp phần giúp "điện toán đám mây" trở thành mô hình điện toán được áp dụng rộng rãi trên toàn thế giới:

-Tốc độ xử lý nhanh , cung cấp cho người dùng những dịch vụ nhanh chóng và giá thành rẻ dựa trên nền tảng cơ sở hạ tầng tập trung(đám mây).

-Chi phí đầu tư ban đầu về cơ sở hạ tầng, máy móc và nguồn nhân lực của người sử dụng điện toán đám mây được giảm đến mức thấp nhất

-Không còn phụ thuộc vào thiết bị và vị trí địa lý

-Chia sẻ tài nguyên và chi phí trên một địa bàn rộng lớn , mang lại các lợi ích cho người dùng.

-Với độ tin cậy cao, điện toán đám mây còn phù hợp với các yêu cầu cao và liên tục của các công ty kinh doanh và các nghiên cứu khoa học

- Khả năng mở rộng được , giúp cải thiện chất lượng các dịch vụ được cung cấp trên "đám mây".
- Khả năng bảo mật được cải thiện do sự tập trung về dữ liệu
- Các ứng dụng của điện toán đám mây dễ dàng để sửa chữa và cải thiện.

- Tài nguyên sử dụng của điện toán đám mây luôn được quản lý và thống kê trên từng khách hàng và ứng dụng , theo từng ngày, từng tuần, từng tháng. Điều này đảm bảo cho việc định lượng giá cả của mỗi dịch vụ do điện toán đám mây cung cấp để người dùng có thể lựa chọn phù hợp

Như vậy, có thể thấy điện toán đám mây có những ưu điểm vượt trội qua đó đóng vai trò quan trọng và rất hữu ích trong thế giới hiện nay.

1.1.3. Kiến trúc của điện toán đám mây

Điện toán đám mây bao gồm 6 thành phần cơ bản: Cơ sở hạ tầng (Infrastructure), lưu trữ đám mây (Cloud Storage), nền tảng đám mây (Cloud Platform), ứng dụng (Application), dịch vụ (Services), khách hàng (Client)

Các mô hình triển khai điện toán đám mây: Đám mây công cộng (Public cloud), đám mây riêng (Private cloud), đám mây cộng đồng (Community cloud), đám mây lai (Hybird cloud)

1.2. Các nhà cung cấp dịch vụ điện toán đám mây

Các dịch vụ đám mây phổ biến như: EC2 của Amazon, Azure của Microsoft, IBM cung cấp Smart Cloud Enterprise, Google cung cấp App Engine, Redhat cung cấp Redhat’s Openshift, Vmware có Cloud Foundry, Viện Công nghiệp Phần mềm và Nội dung số Việt Nam có iDragon Clouds... Trong đó Google Cloud, Redhat’s Openshift, Vmware Cloud Foundry và NISCI iDragon Clouds là những PaaS mã nguồn mở, cho phép thực thi trên một nền hạ tầng với chi phí thấp và dễ dàng thay thế.

Với công nghệ lưu trữ đám mây chúng ta có thể xử lí dữ liệu dưới dạng cấu trúc và phi cấu trúc, tài liệu đến hình ảnh từ nhiều nguồn khác nhau. Ở đây ta xét 3 nhà cung cấp dịch vụ lưu trữ dữ liệu đám mây lớn là Google drive, dropbox và box.

1.3. Phương pháp bảo vệ dữ liệu lưu trữ trên đám mây

1.3.1. Một số vấn đề thực tế về an toàn dữ liệu trong lưu trữ trên đám mây hiện nay

Trên thực tế lập luận và dẫn chứng, những giải pháp đã được đưa vào sử dụng chưa đảm bảo được tuyệt đối tính an toàn, toàn vẹn dữ liệu lưu trữ trên đám

mây. Dựa trên các nghiên cứu, Cloud Security Alliance (CSA) đã đưa ra những vấn đề có mức độ nguy hại cao nhất trong điện toán đám mây gồm [7]:

- Sử dụng bất hợp pháp dịch vụ.
- API (Application Programming Interfaces) không bảo mật.
- Các lỗ hổng trong chia sẻ dữ liệu.
- Mất dữ liệu.
- Tấn công luồng dữ liệu.
- Những nguy hại từ bên trong: Các mối đe dọa này bao gồm gian lận, phá hỏng dữ liệu, đánh cắp hoặc mất thông tin bí mật do chính người trong cuộc được tin tưởng gây ra.

1.3.2. Các biện pháp bảo vệ dữ liệu lưu trữ trên đám mây được sử dụng hiện nay

Mã hóa dữ liệu:

Mã hóa dữ liệu là công nghệ chuyển hóa dữ liệu này thành 1 dạng dữ liệu mới mà người dùng không thể đọc được hoặc hiểu được nó. Bằng cách sử dụng các thuật toán lồng vào nhau, thường dựa trên 1 khóa (key) để mã hóa dữ liệu.

Hầu hết các hình thức mã hóa đều yêu cầu bạn thiết lập mật khẩu, cho phép bạn mã hóa tập tin và sau đó giải mã nó khi bạn muốn xem lại. Nếu bạn sử dụng mật khẩu yếu, tin tặc có thể phá mã hóa và truy cập tập tin, làm thất bại mục đích của mã hóa.

Mã hóa nhằm đảm bảo các yêu cầu sau:

- Tính bí mật (Confidentiality): Dữ liệu không bị xem bởi bên thứ ba.
- Tính toàn vẹn (Integrity): Dữ liệu không bị thay đổi trong quá trình truyền
- Tính không khước từ (Non-repudiation): Là cơ chế người thực hiện hành động không thể chối bỏ việc mình đã làm. Có thể kiểm chứng được nguồn gốc hoặc người đưa tin

Dựa vào thực tế lập luận dẫn chứng, có thể đưa ra những điểm ưu và nhược của phương pháp này như sau:

Ưu điểm:

Với bài toán bảo vệ an toàn dữ liệu lưu trữ trên đám mây đặt ra, giải pháp mã hóa dữ liệu có những mặt tích cực, giải quyết được những vấn đề sau:

- Các hệ mã hóa che dấu được nội dung của văn bản rõ để đảm bảo cho chỉ người chủ hợp pháp của thông tin mới có quyền truy cập thông tin, hay nói cách khác là chống truy cập không đúng quyền hạn. Khi tệp tin bị tin tặc đánh cắp, tin tặc không thể đọc được dữ liệu trong tệp tin đó, điều đó hạn chế được việc tin tặc sử dụng trái phép gây ra những hậu quả không đáng có.
- Tạo các yếu tố xác thực thông tin, đảm bảo thông tin lưu hành trên hệ thống đến người nhận hợp pháp xác thực.
- Tổ chức các sơ đồ chữ ký điện tử, đảm bảo không có hiện tượng giả mạo để gửi thông tin trên mạng.

Nhược điểm:

Không thể phủ nhận những điểm tốt điểm mạnh của việc mã hóa dữ liệu trên đám mây. Tuy nhiên, mã hóa không thể giải quyết hết được vấn đề bảo mật, bởi nguyên nhân của rò rỉ thông tin bao gồm cả việc tồn tại các lỗ hổng như XSS hay SQL injection, cũng như là việc sử dụng mật khẩu khẩu quá ngắn hoặc dễ đoán... Bản thân việc mã hóa không ngăn chặn được việc thông tin bị đánh cắp, việc mất mật khẩu sẽ dẫn tới mất gói dữ liệu, việc này dẫn đến việc không đảm bảo được tính an toàn, toàn vẹn về mặt dự phòng dữ liệu. Bên cạnh đó cũng phải nhấn mạnh rằng, đa phần các công nghệ quản lý khóa mã được sử dụng rộng rãi hiện nay đều tiềm ẩn những rủi ro. Chưa có câu trả lời hoàn hảo cho các câu hỏi như "lưu trữ khóa ở đâu", "phải bảo vệ khóa ra sao", "nhập khóa như thế nào". Nếu số lượng máy ảo trong hệ thống là lớn thì bản thân hệ thống mật mã có thể là nguồn căn của vấn đề. Đó là vì cần có sự phân quyền xem ai được truy cập tới đối tượng nào và tương ứng với nó là việc phân phối khóa.

Để đánh giá một hệ mã hóa ta cần xét đến độ an toàn của thuật toán, tốc độ mã hóa, giải mã và việc phân phối khóa. Có thể thấy "thuật toán nào cũng có thể bị phá vỡ". Các thuật toán khác nhau cung cấp mức độ an toàn khác nhau, phụ thuộc vào độ phức tạp để phá vỡ chúng. Ở đây tôi muốn nhắc đến mức độ phức tạp của hệ mã hóa được sử dụng thấp thì rủi ro gói tin bị phá sẽ cao. Bên cạnh đó, tốc độ xử lý mã hóa, giải mã gói dữ liệu nhanh hay chậm cũng phụ thuộc vào hệ mã hóa có tốt không. Việc bảo mật truy cập, dùng mật khẩu chia quyền cho người sở hữu và người được quyền sử dụng. Ở đây tùy theo người sở hữu có cách chia quyền khác nhau (ví dụ chia quyền cho người sử dụng này chỉ được đọc dữ liệu mà không được chỉnh sửa...), người dùng hoàn toàn tin

tưởng vào nhà cung cấp. Đồng nghĩa với việc không thể đảm bảo tính an toàn bảo mật dữ liệu do những nguyên nhân chủ quan hoặc khách quan xuất phát từ nhà cung cấp dịch vụ mà bạn đang tin tưởng.

Vậy vấn đề cấp thiết đặt ra là tìm ra giải pháp tối ưu làm thế nào giải quyết được vấn đề an toàn bảo mật dữ liệu lưu trữ trên đám mây, khắc phục những hạn chế của giải pháp mã hóa dữ liệu đã áp dụng trước đó mà vẫn đảm bảo được hiệu quả sử dụng cho người dùng.

1.4. Kết luận

Trong chương 1 của luận văn đã hệ thống những lý thuyết cơ bản về điện toán đám mây, vấn đề lưu trữ dữ liệu trên đám mây, bao gồm khái niệm, vai trò, kiến trúc, mô hình dịch vụ, mô hình triển khai điện toán đám mây, những nhà cung cấp dịch vụ điện toán đám mây.

Chương này cũng đưa ra luận điểm những vấn đề còn tồn tại, những lập luận và dẫn chứng về sự thiếu an toàn mất mát dữ liệu. Trình bày những vấn đề có mức độ nguy hại cao nhất trong điện toán đám mây. Trình bày chi tiết và phân tích ưu nhược điểm của giải pháp mã hóa dữ liệu, bảo mật truy cập nhân quyền. Việc bảo vệ an toàn và toàn vẹn cho dữ liệu lưu trữ trên đám mây rất quan trọng tuy nhiên với những con số thống kê thiệt hại và mức độ nguy hại cho thấy tính cấp thiết, ý nghĩa thực tế của việc tìm ra phương pháp nâng cao an toàn bảo mật dữ liệu lưu trữ trên đám mây.

Vậy vấn đề đặt ra ở đây là phương pháp nào nâng cao độ tin cậy cho hệ thống, mức độ tin cậy của hệ thống được đánh giá qua những tiêu chí nào, làm thế nào để xác định độ tin cậy của hệ thống? Câu trả lời được làm rõ trong chương tiếp theo chương 2 của luận văn.

Chương 2. PHƯƠNG PHÁP DỰ PHÒNG CẤU TRÚC NÂNG CAO ĐỘ TIN CẬY CHO HỆ THỐNG LƯU TRỮ DỮ LIỆU

2.1. Tổng quan về phương pháp nâng cao độ tin cậy hệ thống

2.1.1. Một số khái niệm

Hệ thống là một tập hợp gồm nhiều phần tử tương tác, có các mối quan hệ ràng buộc lẫn nhau, tương hỗ nhau và cùng thực hiện hướng tới một mục tiêu nhất định" [3].

"Phần tử là một đối tượng có độ tin cậy độc lập, một bộ phận tạo thành hệ thống mà trong quá trình nghiên cứu độ tin cậy nó được xem như là một đơn vị không chia nhỏ hơn nữa trong hệ thống" [3].

Độ tin cậy của phần tử hoặc hệ thống là xác suất để trong suốt khoảng thời gian khảo sát t, phần tử đó hoặc hệ thống đó vận hành an toàn [3].

2.1.2. Phương pháp đánh giá độ tin cậy của hệ thống qua cấu trúc hệ thống

Vậy làm thế nào để đánh giá độ tin cậy của một hệ thống?

Phần tử bị hư hỏng là một sự kiện ngẫu nhiên xảy ra ở các thời điểm khác nhau nên các chỉ số về độ tin cậy cũng thường tính dưới dạng xác suất.

Độ tin cậy của phần tử giảm dần theo thời gian, để tăng độ tin cậy của hệ thống thì phải thiết kế tăng độ tin cậy của phần tử.

Độ tin cậy hay xác suất vận hành an toàn của hệ thống cấu trúc các phần tử song song luôn cao hơn hệ thống cấu trúc các phần tử nối tiếp.

Cấu trúc của một hệ thống dù phức tạp đến đâu thì cũng chỉ quy về 2 dạng là cấu trúc nối tiếp (hệ thống không dự phòng) và cấu trúc song song (hệ thống dự phòng).

2.1.3. Ý nghĩa

Các hệ thống tính toán được tạo ra giúp thay thế hoặc hỗ trợ con người, mang lại rất nhiều ứng dụng và lợi ích cho kinh tế, cuộc sống toàn cầu. Tuy nhiên, nếu không đảm bảo được độ tin cậy thì hệ thống đó coi như không tồn tại, mà từ thực tế cho thấy việc thao tác sai hay sai lầm trong thiết kế chế tạo thiết bị có thể xảy ra bất kỳ lúc nào, điều này dẫn đến những nguy cơ tiềm tàng xảy ra đối với mỗi hệ thống như cấu trúc hệ thống bị phá vỡ, hệ thống hoạt động không chính xác...

Để đưa ra các phương pháp nhằm nâng cao độ tin cậy của các hệ thống thì chúng ta phải dựa trên các thông số đánh giá độ tin cậy của hệ thống. Các phương pháp đánh giá độ tin cậy của hệ thống liên quan đến các vấn đề về sản xuất, lập trình, dự toán, chi phí bảo trì và các chi phí tối thiểu cấu hình hệ thống. Khi biết được các thông tin về độ tin cậy của hệ thống có thể giúp chúng ta có được kế hoạch bảo trì, lập kế hoạch dự phòng để nâng cao độ tin cậy của hệ thống tránh được các lỗi sự cố có thể xảy ra.

2.2. Phương pháp dự phòng cấu trúc nâng cao độ tin cậy cho hệ thống lưu trữ dữ liệu

Phương pháp dự phòng nâng cao độ tin cậy hệ thống bằng cách đưa ra các đối tượng dư thừa là nguồn lực bổ sung và cơ hội cần thiết tối thiểu để các đối tượng có thể thực hiện chức năng, nhiệm vụ của mình. Qua đó đảm bảo hệ thống vẫn hoạt động bình thường khi xuất hiện từ chối của hệ thống trong các thành phần của nó. Có nhiều phương pháp dự phòng: Dự phòng cấu trúc, dự phòng thông tin và dự phòng thời gian [2].

2.3. Khái quát về cơ chế RAID và RAID đối với bài toán an toàn dữ liệu cho hệ thống máy tính

RAID (viết tắt của *Redundant Array of Independent Disks*) là giải pháp lưu trữ dữ liệu sử dụng loạt các ổ đĩa cứng vật lý được ghép lại với nhau thành một hệ thống có chức năng tăng tốc độ truy xuất dữ liệu hoặc bổ sung cơ chế sao lưu, dự phòng dữ liệu cho hệ thống. RAID là hệ thống đĩa được tạo ra nhằm mục đích tăng cường tốc độ truy cập dữ liệu hệ thống lưu trữ, tăng cường độ tin cậy về mặt dữ liệu. Tốc độ chuyển tải dữ liệu tăng lên khi các dữ liệu được chia đều cho các đĩa cứng hoạt động đồng thời.

RAID có tính dự phòng là nhân tố quan trọng nhất trong quá trình phát triển RAID cho môi trường máy chủ (đặc điểm không có ở giải pháp mã hóa). RAID giá thành thấp với mục tiêu là cung cấp bộ nhớ tốt hơn cho hệ thống so với việc sử dụng riêng biệt các ổ đĩa có dung lượng lớn.

Sử dụng phần lý thuyết ở mục trên **2.1. Tổng quan về các phương pháp nâng cao độ tin cậy hệ thống** ta có thể thấy rõ được tầm quan trọng của yếu tố dự phòng đối với việc nâng cao độ tin cậy của một hệ thống bất kỳ. Mặt khác, phương pháp đánh giá độ tin cậy của hệ thống mang tính kinh tế rất cao, nó liên quan đến các vấn đề về sản xuất, lập trình, dự toán, chi phí bảo trì và các chi phí tối thiểu cấu hình hệ thống.

Như vậy, RAID có tính dự phòng và RAID có giá thành hợp lý, có thể sử dụng RAID vào việc giải quyết các bài toán về độ tin cậy, cụ thể xét ở khía cạnh luận văn này, RAID có những đặc điểm rất phù hợp để sử dụng vào giải pháp nâng cao an toàn bảo mật cho dữ liệu được lưu trữ trên hệ thống dịch vụ đám mây.

2.3.1. Các loại RAID

RAID có rất nhiều chuẩn như RAID 0 (striping), RAID 1 (Mirror), RAID 0+1, RAID 10, RAID 5(Parity), RAID 50 … Với các dòng mainboard desktop cao cấp và các main server 1 way và 2 way có hỗ trợ RAID thì thường dùng RAID 0, 1, 4, 5, RAID 0+1, RAID 10.

Các loại Raid thường dùng:

- ***RAID 0***

+ Hoạt động: Dữ liệu sẽ chia làm nhiều phần bằng nhau và được lưu trữ trên từng đĩa cứng. Mỗi đĩa sẽ chứa 1/n dữ liệu.

+ Ưu điểm: Dung lương tăng n lần ổ đĩa đơn vật lý. Tăng tốc độ đọc ghi dữ liệu. Mỗi đĩa chỉ cần đọc/ghi 1/n dữ liệu được yêu cầu. Theo lý thuyết tốc độ sẽ tăng n lần.

+ Khuyết điểm: Tính an toàn dữ liệu thấp, rủi ro cao. Nếu 1 trong bất kỳ ổ đĩa đơn vật lý bị hư thì dữ liệu còn lại ở các ổ đĩa vật lý khác cũng không còn sử dụng được nữa. Xác suất mất dữ liệu sẽ tăng n lần so với dùng ổ đĩa đơn.

- ***RAID 1***

+ Hoạt động: (n=2) Dữ liệu được ghi vào 2 ổ giống hệt nhau (mirroring).

+ Ưu điểm: Tốc độ đọc/ghi và dung lượng lưu trữ bằng ổ cứng đơn. Đảm bảo an toàn dữ liệu. RAID 1 là chuẩn RAID không thể thiếu đối với người quản trị mạng hoặc những ai phải quản lý nhiều thông tin quan trọng.
+ Khuyết điểm: Không phải là lựa chọn cho người say mê tốc độ. Hiệu năng không phải là yếu tố hàng đầu. Không gia tăng dung lượng lưu trữ.

- ***RAID 10***

+ Hoạt động: Là sự kết hợp RAID 0 + RAID 1. Tổng hợp các ưu điểm của đàn anh. Dữ liệu được ghi đồng thời lên 4 đĩa cứng với 2 ổ dạng Striping tăng tốc và 2 ổ dạng Mirroring sao lưu.

+ Ưu điểm: Tăng tốc đọc/ghi. Performance > RAID 5 và RAID 6. Tính an toàn dữ liệu cao hơn RAID 0, RAID 1 và RAID 5. Cho phép tối đa 2 ổ đĩa cùng lúc bị hỏng ở 2 pair.

2.3.2. Đánh giá độ tin cậy của các hệ thống RAID

Mỗi cấp độ RAID có mẫu khác nhau dẫn đến sự khác nhau về hiệu suất và khả năng dự phòng của từng hệ thống RAID.

RAID có thể nâng cao được tính tin cậy, tốc độ truy cập và dung lượng hệ thống vì RAID sử dụng 2 quy trình để chế tạo là kỹ thuật tạo lát đĩa và kỹ thuật soi gương đĩa. Kỹ thuật tạo lát đĩa tăng được tốc độ đọc ghi. Kỹ thuật soi gương đĩa giúp đạt độ tin cậy cao cho hệ thống lưu trữ.

Với RAID 10 có nhiều ưu điểm, cấu hình phù hợp với máy chủ CSDL, nó là sự kết hợp giữa RAID 1 và 0, giữa tốc độ cao và tính tin cậy cao, vì thế rất phù hợp với các hệ thống máy chủ đòi hỏi tính an toàn cao, hiệu năng lớn như máy chủ CSDL.

2.4. Triển khai RAID

RAID có thể được triển khai ở 2 dạng: RAID cứng và RAID mềm. RAID cứng cung cấp hiệu suất cao nhất cho tất cả các loại RAID, nó cung cấp tính năng chịu lỗi mạnh mẽ hơn đa dạng hơn RAID mềm.

2.5. Kết luận

Trong chương hai của luận văn đã trình bày 2 vấn đề chính:

- Trình bày về các khái niệm liên quan, phương pháp đánh giá độ tin cậy của hệ thống, làm rõ phương pháp dự phòng cấu trúc nâng cao độ tin cậy cho hệ thống lưu trữ dữ liệu và ý nghĩa của việc nâng cao độ tin cậy cho hệ thống.
  Phương pháp dự phòng đóng một vai trò quan trọng trong việc tăng cường hệ thống đáng tin cậy. Một trong những hình thức thường được sử dụng dự phòng là chế độ chờ dự phòng, tôi đã trình bày các phương pháp dự phòng cấu trúc: Hệ thống dự phòng cấu trúc có tải (còn gọi là dự phòng nóng), hệ

thống dự phòng cấu trúc không tải (còn gọi là dự phòng lạnh). Ý tưởng của phương pháp dự phòng này giống như phương pháp dự phòng cấu trúc song song. Để nâng cao độ tin cậy của hệ thống ngoài các phương pháp dự phòng cấu trúc hay còn gọi là dự phòng phần cứng thì chúng ta còn kết hợp sử dụng phương pháp nâng cao độ tin cậy của phần mềm.

- Làm rõ về cơ chế RAID, đặc điểm của từng loại RAID thường dùng và cách triển khai RAID. Đánh giá độ tin cậy của các hệ thống RAID.

Như vậy, với những kiến thức được nêu ở chương này ta nhận thấy rõ tiềm năng của RAID, đặc biệt là tính dự phòng, tính sẵn sàng của dữ liệu. Qua quá trình tìm tòi, nghiên cứu, phân tích và lập luận, tôi đưa ra đề xuất giải pháp nâng cao an toàn bảo mật dữ liệu lưu trữ trên điện toán đám mây sẽ được trình bày chi tiết ở chương sau. Với giải pháp này, ta có thể dụng tuyệt đối những ưu điểm của giải pháp mã hóa dữ liệu và ứng dụng cơ chế RAID vào kết hợp để giải quyết những vấn đề hạn chế còn tồn đọng trong giải pháp mã hóa đám mây, bảo mật truy cập phân quyền.

Chương 3. GIẢI PHÁP NÂNG CAO AN TOÀN BẢO MẬT DỮ LIỆU LƯU TRỮ TRÊN ĐÁM MÂY VÀ ỨNG DỤNG VÀO THỰC TẾ DOANH NGHIỆP

3.1. Giải pháp

3.1.1. Giải pháp

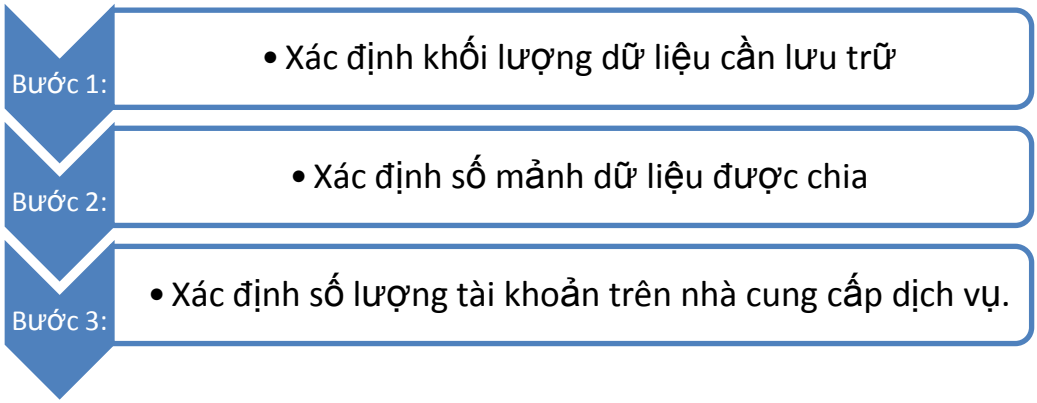
Giải pháp đưa ra là cơ chế lưu trữ dữ liệu trên các dịch vụ cloud, sử dụng các tài khoản miễn phí của các nhà cung cấp dịch vụ như Box, OneDrive, GDrive, Dropbox... Ở đây tận dụng được khả năng linh động của dịch vụ lưu trữ đám mây, kết hợp với cơ chế lưu trữ có dự phòng an toàn của RAID 0,1. Với giải pháp này đã đảm bảo được sự an toàn cho dữ liệu lưu trữ trên đám mây, giải quyết được 2 vấn đề chính đối với bài toán lưu trữ dữ liệu đó là: Tính bảo mật, tính toàn vẹn.

3.1.2. Xây dựng quy trình bài toán thực tế doanh nghiệp:

Một bài toán được đặt ra là: Một doanh nghiệp A ước lượng có một khối lượng thông tin dữ liệu rất lớn cần lưu trữ trên dịch vụ đám mây là n (Gb). Yêu cầu đặt ra là xây dưng quy trình cụ thể để giúp doanh nghiệp A có thể lưu trữ dữ liệu nhanh và hiệu quả nhất mà vẫn đảm bảo được tính an toàn và tính kinh tế.

Xây dựng quy trình bài toán thực tế doanh nghiệp

Quy trình:



Hình 3.1. Quy trình giải quyết bài toán lưu trữ dữ liệu

- Bước 1: Xác định khối lượng dữ liệu cần lưu trữ trên đám mây n (Gb)
- Bước 2: Với khoảng kích thước dữ liệu ước lượng đó xác định ra sẽ băm ra bao nhiêu mảnh dữ liệu để lưu trữ cho hợp lý (đặt là N

mảnh).

- Bước 3: Với N mảnh được băm ra, xác định sẽ có bao nhiêu tài khoản accout trên nhà cung cấp dịch vụ. Dữ liệu đã được phân mảnh, mỗi mảnh sẽ được đưa lên lưu trữ trên các tài khoản khác nhau của các nhà cung cấp dịch vụ.

Người dùng đẩy dữ liệu lên đám mây để lưu trữ thì sẽ có nhu cầu lấy dữ liệu về máy để sử dụng lúc cần thiết. Lúc này các mảnh dữ liệu cần dùng sẽ được gộp lại và tại về máy người dùng.

Với công nghệ phân mảnh, gộp dữ liệu này, hiện nay có nhiều phần mềm hỗ trợ tiện dùng, với giao diện đơn giản, thân thiện, dễ sử dụng như: HJSplit và File Splitter & Joiner.

3.2. Cơ chế lưu trữ dữ liệu

Giải pháp này sử dụng các tài khoản trên các nhà cung cấp dịch vụ cloud hiện nay như: Google drive, OneDrive, Dropbox, Box… để lưu trữ dữ liệu. Những tài khoản miễn phí này có thể được tạo ra đơn giản với địa chỉ email của người dùng. Để đảm bảo tính toàn vẹn cho dữ liệu khi lưu trữ, sẽ sử dụng tối thiểu 3 nhà cung cấp dịch vụ cloud và tối thiểu n (n>=2) tài khoản trên mỗi dịch vụ, do đó số tài khoản dùng để lưu trữ sẽ là 3*n tài khoản [4].

Dữ liệu sẽ được lưu trữ trong các dịch vụ đám mây bằng cách làm theo các quy trình sau: Phân chia dữ liệu của người dùng và mã hóa một phần chúng sau đó lưu trữ dữ liệu vào các tài khoản khác nhau bằng phương pháp tương tự như mô hình RAID 10. Ví dụ: Dữ liệu được chia thành 9 khối và được lưu trữ trong 3 tài khoản trên mỗi dịch vụ đám mây.

Quá trình lưu trữ dữ liệu trên các tài khoản cloud được thực hiện như sau: Với mỗi tập tin người dùng cần lưu trữ, sẽ phân mảnh thành các phần và tiến hành lưu trữ các phần đó trên các tài khoản giống như cơ chế RAID.

Dữ liệu tập tin được lưu vào các tài khoản cloud theo quy tắc:

- Các tài khoản của cùng 1 nhà cung cấp dịch vụ được đặt xen kẽ nhau, theo quy tắc n*i+m (trong đó n là số tài khoản trên cùng 1 dịch vụ, i là số lượt, m là thứ tự tài khoản).

- Trên mỗi tài khoản sẽ lưu trữ 2 mảnh dữ liệu kề nhau theo thứ tự đã phân mảnh.
- Mảnh đầu tiên và cuối cùng sẽ được lưu trên cùng 1 tài khoản.

Với cách phân chia các mảnh vào các tài khoản và thứ tự sắp xếp các tài khoản như vậy sẽ có các ưu điểm là:

- Khi 1 tài khoản bất kì bị mất hoặc không truy cập được, dữ liệu có thể được lấy từ 2 tài khoản lân cận.
- Khi tất cả các tài khoản của cùng một nhà cung cấp dịch vụ bị mất (trường hợp này hiếm xảy ra hơn), dữ liệu của các mảnh vẫn khôi phục được từ các tài khoản khác trên các dịch vụ khác.
- Nếu 2 tài khoản liên tiếp trong danh sách bị mất dữ liệu (trường hợp này có thể xảy ra), dữ liệu không khôi phục được.
- Nếu 2 nhà cung cấp dịch vụ cùng ngừng hoạt động, dữ liệu cũng không khôi phục lại được.

Giải pháp này khuyến khích để dung lượng tối đa cho tập tin tải lên là 200MB. Sau khi phân mảnh, sẽ thêm vào các mảnh dữ liệu này phần header chứa

Total package	Order Package	Next Storage	Filesize	Data...

các thông tin để quản lý như sau:

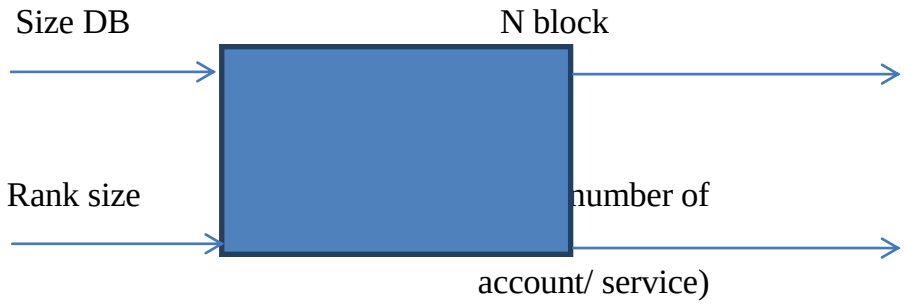
Hình 3.4. Cấu trúc header của các phần

Do được phân mảnh và được lưu trữ phân tán trên các tài khoản của các kho dữ liệu khác nhau, nên dữ liệu của mỗi mảnh trong trường hợp bị truy cập trái phép cũng không thể hiện được nội dung của toàn bộ tài liệu.

3.3. Mô hình bài toán dựa trên lý thuyết xác suất và độ tin cậy của hệ thống

Cho dữ liệu vào là tệp của một người dùng được lưu trữ trong dịch vụ lưu trữ đám mây với kích thước cụ thể.

Mô hình hoạt động:



Hình 3.5. Mô hình hoạt động

Dữ liệu được chia thành các khối N và phân phối đều cho số tài khoản trên mỗi dịch vụ (M). Giả sử M = N.

Gọi P là độ tin cậy ban đầu của hệ thống (P_s).

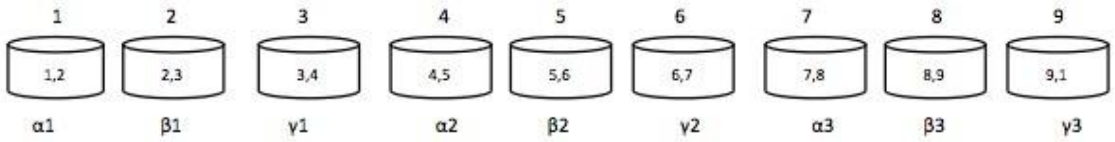
Giả sử một tập tin được phân thành 9 mảnh và sử dụng 3 tài khoản cloud trên mỗi dịch vụ

$\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3$: là độ tin cậy của mỗi mảnh . Khi đó độ tin cậy của hệ thống là $\boldsymbol{P_s = \alpha_1 * \alpha_2 * \alpha_3 * \beta_1 * \beta_2 * \beta_3 * \gamma_1 * \gamma_2 * \gamma_3}$**. (3.1)**

Vậy nên ta giả sử: $\alpha_1 = \alpha_2 = \alpha_3, \beta_1 = \beta_2 = \beta_3, \gamma_1 = \gamma_2 = \gamma_3$

thì độ tin cậy ban đầu của hệ thống là: $\boldsymbol{P_s = \alpha^3 * \beta^3 * \gamma^3}$. **(3.2)**

- Trường hợp 1: Mỗi tài khoản lưu trữ 2 mảnh dữ liệu.



Hình 3.6. Độ tin cậy của hệ thống trong trường hợp 1

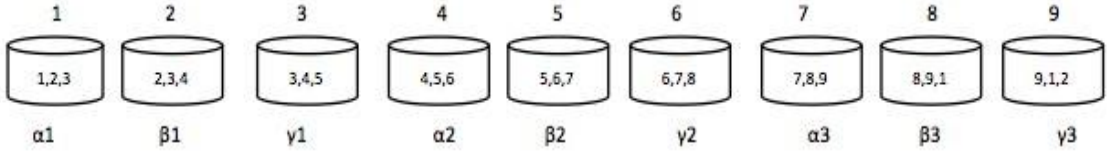
Trong trường hợp này, khi các tài khoản xen kẽ bị mất, dữ liệu có thể được khôi phục từ các tài khoản xung quanh. Nếu nhà cung cấp chấm dứt dịch vụ, dữ liệu sẽ được an toàn nhờ các tài khoản lân cận.

Nếu (Ps1) là độ tin cậy trong Trường hợp 1, khi đó:

$$P_{s1} = \alpha^3*\beta^3*\gamma^3 + (1-\alpha)^3*\beta^3*\gamma^3 + \alpha^3*(1-\beta)^3*\gamma^3 + \alpha^3*\beta^3*(1-\gamma)^3 + \alpha*\beta^2*\gamma^2*(1-\alpha)^2*(1-\beta)*(1-\gamma) + \alpha^2*\beta^2*\gamma*(1-\alpha)*(1-\beta)*(1-\gamma)^2 + \alpha^2*\beta*\gamma^2*(1-\alpha)*(1-\beta)^2*(1-\gamma). \quad (3.3)$$

Áp dụng lý thuyết xác suất thống kê ta xây dựng nên công thức tính P_{s1} [1].

- Trường hợp 2: Mỗi tài khoản lưu trữ 3 mảnh dữ liệu



Hình 3.7. Độ tin cậy của hệ thống trong trường hợp 2

Trong Trường hợp 2, nếu có 2 tài khoản liền kề bị mất hoặc không thể tiếp cận được thì có thể lấy dữ liệu từ các tài khoản lân cận. Nếu một nhà cung cấp chấm dứt dịch vụ, bạn có thể sử dụng dữ liệu từ các tài khoản lân cận. Nếu (Ps2) là độ tin cậy trong Trường hợp 2, khi đó:

$$P_{s2}=\alpha^3*\beta^3*\gamma^3+(1-\alpha)^3*\beta^3*\gamma^3+\alpha^3*(1-\beta)^3*\gamma^3+\alpha^3*\beta^3*(1-\gamma)^3+(1-\alpha)^3*(1-\beta)^3*\gamma^3+\alpha^3*(1-\beta)^3*(1-\gamma)^3+(1-\alpha)^3*\beta^3*(1-\gamma)^3. \quad (3.4)$$

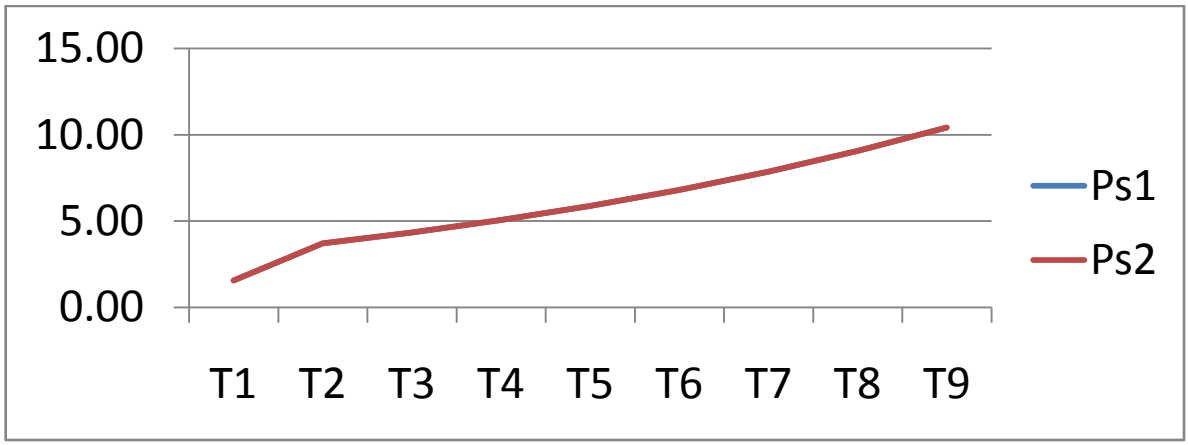
Áp dụng lý thuyết xác suất thống kê ta xây dựng nên công thức tính P_{s2} [1].

Khi một nhà cung cấp gặp sự cố

Bảng 1. Bảng so sánh độ tăng độ tin cậy của trường hợp 1

A	β	γ	P_S	P_{S1}	P_{S2}	*Độ tin cậy tăng % (TH1)*	*Độ tin cậy tăng % (TH2)*
0.9999	***0.9999***	***0.8***	*0.511693*	*0.519688*	*0.519688*	*1.5625*	*1.5625*
0.99999	***0.99999***	***0.75***	*0.42185*	*0.437474*	*0.437474*	*3.703704*	*3.703704*
0.99999	***0.99999***	***0.74***	*0.4052*	*0.422775*	*0.422775*	*4.337354*	*4.337354*
0.99999	***0.99999***	***0.73***	*0.388994*	*0.408675*	*0.408675*	*5.059676*	*5.059676*
0.99999	***0.99999***	***0.72***	*0.373226*	*0.395176*	*0.395176*	*5.881344*	*5.881344*
0.99999	***0.99999***	***0.71***	*0.35789*	*0.382277*	*0.382277*	*6.814264*	*6.814264*
0.99999	***0.99999***	***0.7***	*0.342979*	*0.369978*	*0.369978*	*7.87172*	*7.87172*
0.99999	***0.99999***	***0.69***	*0.328489*	*0.358279*	*0.358279*	*9.068549*	*9.068549*
0.99999	***0.99999***	***0.68***	*0.314413*	*0.347179*	*0.347179*	*10.42133*	*10.42133*

Với trường hợp này ta đưa vào số liệu giả sử với độ tin cậy của α = β = 0.9999 gần như tuyệt đối và giá trị γ nhỏ hơn thay đổi nhỏ dần để làm rõ độ tăng của độ tin cậy trước và sau khi sử dụng giải pháp trong trường hợp một nhà cung cấp gặp sự cố.



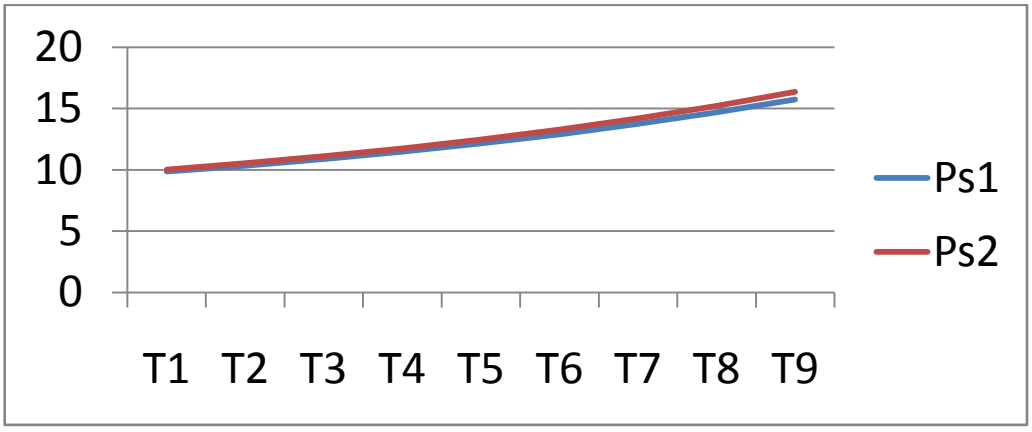
Hình 3.8. Biểu đồ hiển thị độ tăng của độ tin cậy ở trường hợp 1

Khi hai nhà cung cấp gặp sự cố

Bảng 2. Bảng so sánh độ tăng độ tin cậy ở trường hợp 2

A	β	γ	P_S	P_{S1}	P_{S2}	Độ tin cậy tăng % (TH1)	Độ tin cậy tăng % (TH2)
0.99999	**0.699**	**0.79**	0.168384	0.184992	0.185244	9.863302	10.01321
0.99999	**0.698**	**0.78**	0.161375	0.178067	0.17836	10.34332	10.52497
0.99999	**0.697**	**0.77**	0.154581	0.171401	0.171739	10.8806	11.09945
0.99999	**0.696**	**0.76**	0.147998	0.164991	0.165379	11.48209	11.7444
0.99999	**0.695**	**0.75**	0.14162	0.158835	0.159278	12.15554	12.46845
0.99999	**0.694**	**0.74**	0.135444	0.152929	0.153433	12.90955	13.28123
0.99999	**0.693**	**0.3**	0.008986	0.123918	0.133842	1279.067	1389.509
0.99999	**0.692**	**0.72**	0.123681	0.14186	0.142502	14.69875	15.21718
0.99999	**0.691**	**0.71**	0.118085	0.136692	0.137411	15.75656	16.36574

Tương tự với trường hợp này ta đưa vào số liệu giả sử với độ tin cậy của α gần như tuyệt đối và giá trị β, γ nhỏ hơn thay đổi nhỏ dần để làm rõ độ tăng của độ tin cậy trước và sau khi sử dụng giải pháp trong trường hợp hai nhà cung cấp gặp sự cố.



Hình 3.9. Biểu đồ hiển thị độ tăng của độ tin cậy ở trường hợp 2

Giả sử hai trường hợp này có độ tin cậy cao, nhưng nếu một dịch vụ thay đổi chính sách bảo mật của mình hoặc vào thời điểm đó các hacker đang khai thác lỗ hổng, độ tin cậy của hệ thống sẽ giảm đáng kể. Khả năng dự phòng của hai trường hợp này sẽ trở nên khả thi hơn bằng cách nâng cao độ tin cậy lên 1,56% và 3,7%.

3.4. Đánh giá và so sánh với giải pháp khác

Đối với vấn đề bảo vệ an toàn dữ liệu lưu trữ điện toán đám mây từ trước đến nay được các nhà cung cấp dịch vụ đám quan tâm nghiên cứu. Tuy nhiên mức độ hiệu quả các giải pháp đã được sử dụng như thế nào hay giải pháp trong luận văn này đưa ra có những ưu điểm gì?

Mã hóa có thể là một giải pháp cho vấn đề "Làm sao có thể ngăn chặn truy cập bất hợp pháp tới dữ liệu của người dùng khi mật khẩu của họ đang bị đánh cắp" vì đơn giản chỉ cần mã hóa các tập tin trước khi gửi lên các dịch vụ cloud sẽ ngăn chặn thông tin rò rỉ từ các tập tin bị đánh cắp. Khi đó nếu mật khẩu bị đánh cắp, bên thứ 3 vẫn sẽ có quyền truy cập đến dữ liệu, nhưng họ sẽ không có khả năng giải mã để xem dữ liệu. Hiện nay một số phần mềm Credeoncp, Spideroak, BoxCryptor đã được phát triển dựa trên nguyên lý mã hoá dữ liệu của người dùng trước khi đưa lên cloud.

Các giải pháp để nâng cao tính bảo mật cho các dịch vụ lưu trữ cloud hiện nay đa phần đều ứng dụng cơ chế mã hoá dữ liệu, điều này hạn chế được việc lộ dữ liệu bí mật và truy cập bất hợp pháp. Tuy nhiên, cần nhận định rằng, những điều cam kết về quyền riêng tư của người dùng từ các nhà cung cấp dịch vụ chỉ là

tương đối, và chúng ta chưa thể khẳng định được do hạ tầng và giải pháp của họ là hoàn toàn đóng.

Bên cạnh đó, yếu tố đảm bảo tính toàn vẹn dữ liệu chưa được đề cập nhiều, dịch vụ cloud có thể dừng bất cứ khi nào do nhiều nguyên nhân, khi đó dữ liệu của người dùng sẽ không thể khôi phục được. Với đề xuất về giải pháp này đã tính đến yếu tố bảo mật dữ liệu và tính dự phòng cho việc khôi phục trong trường hợp bị mất mát.

KẾT LUẬN

Công nghệ điện toán đám mây đang phát triển nhanh chóng và trở thành một nền tảng được sử dụng rộng rãi cho các ứng dụng tính toán phức tạp và hình thành cụm lưu trữ dữ liệu. Vấn đề an ninh và an toàn dữ liệu luôn là điều được quan tâm và thu hút nhiều nghiên cứu của các nhà khoa học.

Sau thời gian tìm hiểu, nghiên cứu tài liệu và làm luận văn dưới sự hướng dẫn của thầy TS. Lê Quang Minh tôi đã hoàn thành luận văn với đề tài *" **Nghiên cứu giải pháp nâng cao an toàn bảo mật cho dữ liệu đám mây"***. Luận văn đã đạt được kết quả sau:

- Tìm hiểu, nghiên cứu những lý thuyết tổng quan xoay quanh dịch vụ lưu trữ đám mây. Đưa ra luận điểm những vấn đề còn tồn tại, những lập luận và dẫn chứng về sự thiếu an toàn mất mát dữ liệu. Trình bày những vấn đề có mức độ nguy hại cao nhất trong điện toán đám mây. Trình bày chi tiết và phân tích ưu nhược điểm của giải pháp mã hóa dữ liệu, bảo mật truy cập nhân quyền. Qua đó làm nổi bật lên tính cấp thiết, ý nghĩa thực tiễn của chủ đề luận văn thực hiện.
- Tìm hiểu về các phương pháp dự phòng nâng cao độ tin cậy của hệ thống. Sau đó, trình bày tổng hợp, phân tích kiến thức xoay quanh cơ chế RAID, triển khai RAID. RAID đối với bài toán an toàn dữ liệu cho hệ thống máy.

- Dựa trên những cơ sở lý thuyết nêu trên, luận văn đã đưa ra giải pháp nâng cao an toàn dữ liệu lưu trữ trên đám mây, giải pháp này đã giải quyết vấn đề chính còn tồn tại ở những dịch vụ lưu trữ trên cloud hiện nay đó là: Tính bảo mật và toàn vẹn cho dữ liệu người dùng.

Xây dựng thành công quy trình giải quyết bài toán thực tế (đặc biệt sử dụng vào việc lưu trữ dữ liệu cho cá nhân tổ chức doanh nghiệp). Sử dụng toán học vào chứng minh được độ tin cậy của giải pháp, đưa ra bảng số liệu tính toán và đồ thị để so sánh làm rõ mức độ cải thiện lớn về độ tin cậy của hệ thống khi sử dụng giải pháp.