ĐẠI HỌC THÁI NGUYÊN

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN&TRUYỀN THÔNG

PHẠM THỊ PHƯỢNG

TÌM HIỂU MÔ HÌNH ĐIỆN TOÁN ĐÁM MÂY VÀ VẤN ĐỀ BẢO MẬT DỮ LIỆU TRONG ĐIỆN TOÁN ĐÁM MÂY

Ngành: Công nghệ thông tin

Chuyên ngành: Khoa học máy tính

Mã số: 60 48 0101

LUẬN VĂN THẠC SĨ CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. HỒ VĂN CANH

Thái Nguyên – 2019

Số hóa bởi Trung tâm Học liệu và Công nghệ thông tin – ĐHTN http://lrc.tnu.edu.vn

LỜI CAM ĐOAN

Học viên xin cam đoan luận văn này là công trình nghiên cứu thực sự của bản thân, dưới sự hướng dẫn khoa học của TS. Hồ Văn Canh.

Các số liệu, kết quả trong luận văn là trung thực và chưa từng được công bố dưới bất cứ hình thức nào. Tất cả các nội dung tham khảo, kế thừa của các tác giả khác đều được trích dẫn đầy đủ.

Em xin chịu trách nhiệm về nghiên cứu của mình.

Tác giả

Phạm Thị Phượng

Số hóa bởi Trung tâm Học liệu và Công nghệ thông tin – ĐHTN http://lrc.tnu.edu.vn

LỜI CẢM ƠN

Học viên trân trọng cảm ơn sự quan tâm, tạo điều kiện và động viên của Lãnh đạo Đại học Thái Nguyên, các thầy cô Khoa Đào tạo sau đại học, các khoa đào tạo và các quý phòng ban Học viện trong suốt thời gian qua.

Học viên xin bày tỏ sự biết ơn sâu sắc tới TS. Hồ Văn Canh đã nhiệt tình định hướng, bồi dưỡng, hướng dẫn học viên thực hiện các nội dung khoa học trong suốt quá trình nghiên cứu, thực hiện luận văn.

Xin chân thành cảm ơn sự động viên, giúp đỡ to lớn từ phía Cơ quan đơn vị, đồng nghiệp và gia đình đã hỗ trợ học viên trong suốt quá trình triển khai các nội dung nghiên cứu.

Mặc dù học viên đã rất cố gắng, tuy nhiên, luận văn không tránh khỏi những thiếu sót. Học viên kính mong nhận được sự đóng góp từ phía Cơ sở đào tạo, quý thầy cô, các nhà khoa học để tiếp tục hoàn thiện và tạo cơ sở cho những nghiên cứu tiếp theo.

Xin trân trọng cảm ơn!

Thái Nguyên, tháng năm 2019

Học viên

Phạm Thị Phượng

Số hóa bởi Trung tâm Học liệu và Công nghệ thông tin – ĐHTN http://lrc.tnu.edu.vn

DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT

DANH MỤC CÁC HÌNH VẼ, BẢNG BIỂU

Số hóa bởi Trung tâm Học liệu và Công nghệ thông tin – ĐHTN http://lrc.tnu.edu.vn

Số hóa bởi Trung tâm Học liệu và Công nghệ thông tin – ĐHTN http://lrc.tnu.edu.vn

MỞ ĐẦU

1. Đặt vấn đề

Điện toán đám mây_ Cloud Computing được hình thành năm 1969 và có sự phát triển mạnh mẽ từ khi có internet băng thông rộng, đã làm thay đổi cách thức hoạt động của điện toán truyền thống. Hiện nay, điện toán đám mây (ĐTĐM) được các quốc gia trên thế giới ứng dụng rộng rãi trong các lĩnh vực hoạt động của đời sống, kinh tế xã hội. Bằng việc tối ưu sử dụng các nguồn tài nguyên hệ thống, điện toán đám mây đem lại nhiều lợi ích, cơ hội mới cho các các cơ quan, tổ chức, doanh nghiệp trong quá trình đẩy mạnh ứng dụng công nghệ thông tin, truyền thông vào hoạt động chuyên ngành [3, 4].

Các hoạt động liên quan tới điện toán đám mây được chính phủ các quốc gia phát triển mang tính chiến lược trên phạm vi toàn thế giới như đám mây Nebula, google moderator của Mỹ, đám mây G-clould của Anh, kasumigaseki của Nhật Bản…bởi vậy điện toán đám mây luôn thu hút nhiều quốc gia, tổ chức, các tập đoàn, công ty và nhà khoa học, các chuyên gia đầu tư nghiên cứu [10, 11, 13].

Ở nước ta hiện nay, hầu hết các tổ chức, doanh nghiệp đã có hiểu biết cơ bản về điện toán đám mây. Nhiều tổ chức, doanh nghiệp đã và đang sử dụng điện toán đám mây theo các mức độ khác nhau. Một số công trình nghiên cứu [3, 6] đã chỉ rõ điện toán đám mây là giải pháp tối ưu để các doanh nghiệp nước ta giảm thiểu chi phí cũng như tăng hiệu suất làm việc ở mức tối đa.

Tuy nhiên trong quá trình nghiên cứu và ứng dụng cho thấy có nhiều vấn đề về nguy cơ an ninh an toàn thông tin đang đặt ra hiện nay đối với việc lưu trữ dữ liệu trên đám mây [16, 24]. Do vậy, tình hình sử dụng công nghệ đám mây còn gặp phải một số khó khăn nhất định, hiệu quả ứng dụng chưa phát huy tối đa tính ưu việt của các dịch vụ. Trước những yêu cầu cấp bách đó, đòi hỏi cần có những nghiên cứu, giải pháp tăng tính an toàn cho đám mây cũng như việc bảo mật thông tin, dữ liệu lưu trữ.

Xuất phát từ thực tiễn đó, luận văn *"Tìm hiểu điện toán đám mây và vấn đề bảo mật dữ liệu trong điện toán đám mây"* mang tính cấp thiết, thực sự có ý nghĩa khoa học và thực tiễn.

2. Đối tượng và phạm vi nghiên cứu

Nghiên cứu tìm hiểu về điện toán đám mây, kiến trúc, mô hình, ưu nhược điểm và giới thiệu một số nhà cung cấp dịch vụ điện toán đám mây

- Nghiên cứu một số vấn đề bảo mật dữ liệu trong điện toán đám mây và phương pháp khắc phục. Từ đó đi sâu tìm hiểu phương pháp bảo vệ dữ liệu đã lưu trữ bằng các thuật toán mã hóa AES và RSA.

- Nghiên cứu và cài đặt, thử nghiệm hệ thống máy chủ lưu trữ ownCloud.

3. Hướng nghiên cứu của luận văn

Nghiên cứu tổng quan mô hình điện toán đám mây, một số vấn đề bảo mật dữ liệu trong điện toán đám mây và phương pháp khắc phục. Từ đó đi sâu nghiên cứu phương pháp bảo vệ dữ liệu đã lưu trữ bằng các thuật toán mã hóa trên máy chủ ownCloud. Nghiên cứu xây giải pháp mã hóa dữ liệu an toàn từ phía người dùng và ổ chức cài đặt, thực nghiệm, đánh giá các kết quả nghiên cứu đạt được.

4. Những nội dung nghiên cứu chính

Chương 1: Tổng quan về điện toán đám mây

Nghiên cứu về tổng quan khái niệm, lịch sử hình thành và phát triển của điện toán đám mây, kiến trúc và một số mô hình của điện toán đám mây. Đồng thời phân tích chỉ ra những ưu, nhược điểm, tình hình triển khai nghiên cứu ứng dụng và sử dụng công nghệ điện toán đám mây thế giới và tại Việt Nam.

Chương 2: Bảo vệ thông tin trong điện toán đám mây

Nội dung Chương 2 nghiên cứu tìm hiểu vấn đề an ninh thông tin, một số tiêu chuẩn về an ninh thông tin, phân loại an ninh thông tin trong điện tóa đám mây, vấn đề an ninh dữ liệu trong điện toán đám mây và giải pháp. Trên cơ sở đó, tập trung phân tích hai thuật toán mã hóa dữ liệu lưu trữ cho điện toán đám mây là RSA và AES.

Chương 3: Ứng dụng bảo vệ thông tin trong điện toán đám mây

Nghiên cứu và xây dựng điện toán đám mây riêng tích hợp công cụ thu thập thông tin tự động dựa trên phần mềm mã nguồn mở ownCloud và Apache Nutch. Nghiên cứu phân tích thuật toán mã hóa dữ liệu phía ownCloud và đề xuất xây dựng giải pháp mã hóa dữ liệu an toàn phía client sử dụng RSA kết hợp AES 256. Tiến hành cài đặt, thực nghiệm và rút ra những kết luận, đề xuất.

5. Phương pháp nghiên cứu

- Nghiên cứu các bài báo khoa học trong nước và quốc tế.

- Nghiên cứu một số vấn đề bảo mật dữ liệu trong điện toán đám mây và phương pháp khắc phục. Từ đó đi sâu tìm hiểu phương pháp bảo vệ dữ liệu đã lưu trữ bằng các thuật toán mã hóa AES và RSA.

- Cài đặt ứng dụng thử nghiệm và đánh giá.

6. Ý nghĩa khoa học của luận văn

Nghiên cứu vấn đề bảo mật dữ liệu trong điện toán đám mây có ý nghĩa và vai trò to lớn trong việc vệ an ninh thông tin. Đây là vấn đề đang được quan tâm, thu hút nhiều quốc gia, tổ chức, cá nhân đầu tư nghiên cứu. Luận văn đã kết hợp hai kỹ thuật sử dụng các search engine để xây dựng đám mây thu tin tự động và kỹ thuật mã hóa để bảo mật dữ liệu. Do vậy, luận văn có tính khoa học và ứng dụng thực tiễn.

CHƯƠNG 1: TỔNG QUAN VỀ ĐIỆN TOÁN ĐÁM MÂY

1.1 Điện toán đám mây

Điện toán đám mây - Cloud Computing (sau đây có thể gọi tắt là đám mây) là mô hình điện toán đang tiến tới hoàn chỉnh, mỗi tổ chức tiêu chuẩn, mỗi hãng công nghệ đang đưa ra những định nghĩa và cách nhìn của riêng mình.

Theo Wikipedia: *"Điện toán đám mây là một mô hình điện toán có khả năng co giãn linh động, các tài nguyên thường được ảo hóa và được cung cấp như một dịch vụ trên mạng Internet"*.

Theo Ian Foster: *"Một mô hình điện toán phân tán có tính co giãn lớn mà hướng theo co giãn về mặt kinh tế, là nơi chứa các sức mạnh tính toán, kho lưu trữ, các nền tảng và các dịch vụ được trực quan, ảo hóa và co giãn linh động, sẽ được phân phối theo nhu cầu cho các khách hàng bên ngoài thông qua Internet"*.

Một số định nghĩa thì cho rằng điện toán đám mây là điện toán máy chủ ảo, tuy nhiên, định nghĩa này chưa thực sự đầy đủ và chính xác, máy chủ ảo không phải là thành phần thiết yếu của một đám mây. Nó chỉ là thành phần chủ chốt để một vài loại đám mây hoạt động.

Hiện tại, định nghĩa của Viện tiêu chuẩn và công nghệ quốc gia Mỹ - NIST (National Institute of Science and Technology) được cho là thể hiện rõ nhất bản chất của điện toán đám mây [14]: *điện toán đám mây là mô hình điện toán cho phép truy cập qua mạng để lựa chọn và sử dụng tài nguyên tính toán (mạng, máy chủ, lưu trữ, ứng dụng và dịch vụ…) theo nhu cầu một cách thuận tiện và nhanh chóng. Đồng thời, điện toán đám mây cũng cho phép kết thúc sử dụng dịch vụ, giải phóng tài nguyên dễ dàng, giảm thiểu các tương tác với nhà cung cấp.*

Như vậy, điện toán đám mây có thể coi là bước tiếp theo của ảo hóa, bao gồm ảo hóa phần cứng và ứng dụng, là thành phần quản lý, tổ chức, vận hành các hệ thống ảo hóa trước đó.

Điện toán đám mây có năm đặc điểm chính như sau:

Tự phục vụ theo nhu cầu (On-deman self-service): Người sử dụng có thể tự cung cấp các tài nguyên như máy chủ ảo, tài khoản email… mà không cần có

người tương tác với nhân viên của nhà cung cấp dịch vụ (nhân viên công nghệ thông tin).

Mạng lưới truy cập rộng lớn (Broad Network Access): Khách hàng có thể truy cập tài nguyên qua mạng máy tính (như mạng Internet) từ nhiều thiết bị khác nhau (điện thoại thông minh, máy tính bảng, máy tính xách tay…).

Tài nguyên được chia sẻ (Resource Pooling): Tài nguyên của các nhà cung cấp dịch vụ được chia sẻ tới nhiều khách hàng. Thông thường, các công nghệ ảo hóa được sử dụng để cho nhiều bên cùng thuê và cho phép tài nguyên được cấp phát động dựa theo nhu cầu của khách hàng.

Tính linh hoạt nhanh (Rapid elasticity): Tài nguyên có thể được cung cấp và giải phóng nhanh, tự động dựa trên nhu cầu. Khách hàng có thể tăng hoặc giảm việc sử dụng dịch vụ đám mây một cách dễ dàng theo nhu cầu hiện tại của mình.

Ước lượng dịch vụ (Measured service): Khách hàng chỉ chi trả cho tài nguyên thực tế họ đã sử dụng. Thông thường, nhà cung cấp dịch vụ sẽ cung cấp cho khách hàng bảng điều khiển (dashboard) để họ có thể theo dõi việc sử dụng dịch vụ của họ.

Điện toán đám mây đã khắc phục được yếu điểm quan trọng của điện toán truyền thống về khả năng mở rộng và độ linh hoạt. Các công ty, tổ chức có thể triển khai ứng dụng và dịch vụ nhanh chóng, giảm chi phí và ít rủi ro về đầu tư ban đầu.

1.2 Lịch sử hình thành và phát triển của điện toán đám mây

Điện toán đám mây thường được mọi người biết đến như một công nghệ mới được phát triển trong những năm gần đây. Tuy nhiên, khái niệm này không mới như ta vẫn nghĩ. Điện toán đám mây đã bắt đầu được hình thành vào khoảng giữa thế kỷ 20, khi có sự ra đời của các máy tính mainframe. Dưới đây là một số mốc phát triển quan trọng của điện toán đám mây [20]:

Năm 1969, J.C.R Liicklider là người chịu trách nhiệm tạo điều kiện cho sự phát triển của APANET trong cuốn Advanced Research Project Agency Network đã nêu ý tưởng về mạng máy tính giữa các thiên hà, có vẻ giống với điện toán đám mây.