BIG DATA VÀ GIẢI PHÁP LƯU TRỮ DỮ LIỆU

Nguyễn Mậu Hân

Khoa Công nghệ Thông tin, Trường Đại học Khoa học, Đại học Huế

Email: nmhan2009@gmail.com

Ngày nhận bài: 29/3/2018; ngày hoàn thành phản biện: 26/4/2018; ngày duyệt đăng: 8/6/2018

TÓM TẮT

Big Data được biết như là một giải pháp lý tưởng để xử lý các dữ liệu có cấu trúc, dữ liệu bán cấu trúc hay thậm chí là phi cấu trúc như weblogs, mạng xã hội, e-mail, các dữ liệu cảm biến và các bức ảnh mà có thể được khai thác nhằm tìm ra những thông tin hữu ích. Vấn đề đặt ra là giải pháp nào cho bài toán lưu trữ đối với loại dữ liệu này. Bài báo này đề xuất một giải pháp lưu trữ cho các trung tâm dữ liệu, nơi đang tiếp xúc hằng ngày các loại Big Data, trong khi các phương pháp lưu trữ truyền thống đã bộc lộ nhiều khiếm khuyết.

Từ khóa: Big Data, Dữ liệu có cấu trúc, Haddop, MapReduce.

BIG DATA AND STRUCTURE STORAGE SOLUTION

Nguyen Mau Han

Faculty of Information Technology, University of Sciences, Hue University

Email: nmhan2009@gmail.com

ABSTRACT

Big Data is known as an ideal solution for processing structured data, semi-structured data or even unstructured data such as weblogs, social networks, e-mail, sensitive data, and images that can be exploited to find useful information. The question is any solution to the problem of storage for this kind of data. This article proposes a storage solution for data centers, which are working on the Big Data daily, while traditional storage methods have exposed many defects.

Keywords: Big Data, , Haddop, MapReduce, structured data.

Nguyễn Mậu Hân sinh năm 1957 tại Thừa Thiên Huế. Ông tốt nghiệp cử nhân ngành Toán lý thuyết năm 1981 và thạc sĩ chuyên ngành Khoa học máy tính năm 1998. Ông nhận bằng tiến sĩ tại viện Công nghệ thông tin, Hà Nội năm 2003 và được phong hàm Phó Giáo sư năm 2013. Từ năm 1994 đến nay, ông là giảng viên tại khoa Công nghệ thông tin, Trường Đại học Khoa học, Đại học Huế.

Lĩnh vực nghiên cứu: Xử lý song song và phân tán, tính toán lưới và điện toán đám mây.