

XÂY DỰNG VÀ TRIỂN KHAI NỀN TẢNG ĐIỆN TOÁN Đám Mây ỨNG DỤNG TRONG HỆ THỐNG CUNG CẤP THÔNG TIN KHÍ TƯỢNG THỦY VĂN TRÊN THIẾT BỊ DI ĐỘNG

Ngô Văn Mạnh, Nguyễn Văn Khoa - Trung tâm Thông tin và Dữ liệu Khí tượng Thủy văn
Ban Hà Bằng - Đại học Bách Khoa Hà Nội

Hiện nay, Trung tâm Khí tượng Thủy văn (KTTV) quốc gia đang cung cấp thông tin qua các hình thức như trang thông tin điện tử, fax, điện thoại, truyền hình, truyền thanh và văn bản khi có yêu cầu. Hình thức cung cấp thông tin này có nhiều hạn chế như thời lượng phát sóng có hạn và thiếu tính cập nhật. Để khắc phục nhược điểm này, chúng tôi đã xây dựng hệ thống cung cấp thông tin trên thiết bị di động, cho phép người sử dụng xem thông tin KTTV vào bất kỳ thời điểm nào, gửi cảnh báo đến di động của người sử dụng khi có thiên tai xảy ra. Tuy vậy, để triển khai hệ thống lớn, đáp ứng hàng triệu truy cập đòi hỏi một hạ tầng tính toán ổn định với các tính năng như tự phục vụ người dùng, truy cập trên mọi nền tảng thiết bị, gộp tài nguyên vật lý, khả năng co giãn và đàn hồi tài nguyên. Hiện tại, hạ tầng tính toán đảm bảo được các tính năng này được biết đến là điện toán đám mây. Trong bài báo này, chúng tôi trình bày việc xây dựng và triển khai hạ tầng điện toán đám mây ứng dụng trong hệ thống cung cấp thông tin khí tượng thủy văn trên thiết bị di động.

Từ khóa: điện toán đám mây, hệ thống dự báo khí tượng thủy văn, di động.

1. Mở đầu

Hiện tại, Trung tâm KTTV Quốc gia cung cấp thông tin KTTV qua các hình thức như trang thông tin điện tử website [1], Fax, điện thoại, truyền hình, truyền thanh và văn bản, ... Nói chung, các phương thức truyền thống này mới chỉ tập trung phục vụ cho các cơ quan phòng chống thiên tai và các Bộ ban ngành có liên quan, chưa đáp ứng được hết các nhu cầu mọi lúc, mọi nơi của xã hội. Khi hầu hết người dân đều sử dụng điện thoại di động thì phương thức cung cấp thông tin KTTV thông qua thiết bị di động là một lựa chọn hợp lý trong điều kiện hiện nay.

Như chúng ta biết điện toán đám mây hiện đang là xu thế phát triển tương lai của hầu hết các nhà phát triển ứng dụng. Việc ứng dụng công nghệ điện toán đám mây mang đến rất nhiều lợi ích cho doanh nghiệp, nhà cung ứng dịch vụ và người sử dụng [2, 4, 5, 7, 8]. Trong công trình nghiên cứu trước [13], chúng tôi đã trình bày kiến trúc tổng thể về hệ thống cung cấp thông tin khí tượng thủy văn nhờ thiết bị di động. Trong công trình này, chúng tôi trình bày việc xây dựng và triển khai nền tảng điện toán đám mây ứng dụng cho hệ thống trong [13].

Để triển khai nền tảng điện toán đám mây

chúng ta có một số lựa chọn: 1) thuê một cloud (dịch vụ điện toán đám mây) của một số nhà cung ứng dịch vụ cloud trong và ngoài nước như: amazon, icloud, vnpt, ... Với cách tiếp cận này, chỉ cần bỏ ra một khoản kinh phí là chúng ta sẽ có một cloud với cấu hình mong muốn. Tuy nhiên, chúng ta không thể làm chủ công nghệ cloud theo cách này; 2) tự xây dựng một cloud riêng và làm chủ nó. Theo cách tiếp cận này thì chúng ta có một số lựa chọn như: openstack, cloudstack, dev-stack, ... [4]. Trong công trình này, chúng tôi chọn giải pháp openstack.

2. Ảo hóa và điện toán đám mây

Phần này chúng tôi trình bày lược giản về công nghệ ảo hóa hiện có trên thế giới, khi ảo hóa được xem như bước đầu của việc xây dựng điện toán đám mây, sau đó chúng tôi trình bày về điện toán đám mây theo năm đặc tính, bốn mô hình và ba dịch vụ.

2.1. Ảo hóa

Virtualization - ảo hóa [2, 6, 10]: Là việc tạo ra các phiên bản ảo, có thể là các thiết bị phần cứng như CPU, RAM, hệ điều hành, thiết bị lưu trữ, mạng, Bằng cách đưa ra một khái niệm logic về tài nguyên máy tính hơn là một khái niệm vật lý. Ảo hóa cho phép chia sẻ tài nguyên

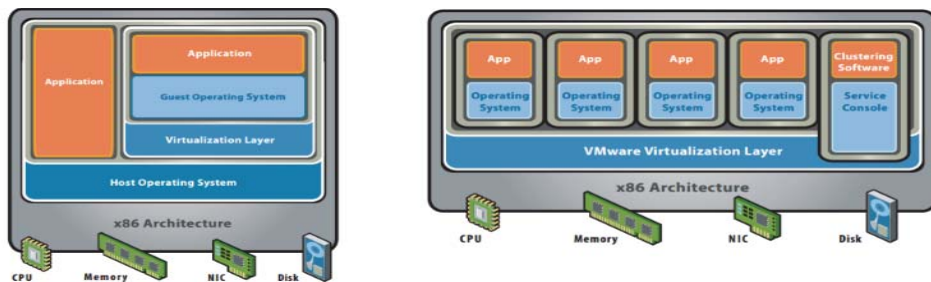
phần cứng, nhờ vậy, hiệu quả sử dụng được tăng lên. Có thể coi ảo hóa là bước đầu tiên của điện toán đám mây. Có thể xem xét sự khác biệt giữa trước và sau ảo hóa trong Bảng 1 [10-12].

- Một số ưu điểm của ảo hóa: giảm thiểu chi phí bảo dưỡng, tương thích với nhiều hệ điều hành đồng thời, tập trung cho kiểm soát và quản trị, dễ dàng sao lưu và phục hồi, khai thác nhiều hơn về công suất và hiệu năng, và là bước đệm để xây dựng điện toán đám mây.

- Phân loại công nghệ ảo hóa: (1) ảo hóa hoàn toàn: Mã mở có KVM, VIR-TUALBOX, KQEMU, trong khi, sản phẩm thương mại có VMWare, Hyper-V, ... (2) ảo hóa một phần có VMWare, Xen, ... (3) ảo hóa ở mức hệ điều hành có OpenVZ, Docker, ... Về hướng tiếp cận ảo hóa thì hiện tại có hai hướng tiếp cận chính: hosted và Bare-metal. Sự khác biệt giữa hai hướng tiếp cận này được thể hiện tại hình 1 [12].

Bảng 1. So sánh sự khác biệt giữa trước và sau ảo hóa

Trước khi ảo hóa	Sau khi ảo hóa
Một hệ điều hành duy nhất trên một máy	Phần cứng độc lập với hệ điều hành và phần mềm
Phần mềm và phần cứng gắn chặt vào nhau	Các máy ảo có thể cung ứng từ mọi hệ thống
Hạ tầng không linh hoạt tốn kém	Có thể quản lý nhiều hệ điều hành và ứng dụng như một đơn vị duy nhất bởi vì chúng đóng gói vào một máy ảo
Chạy nhiều ứng dụng trên một máy dễ xảy ra xung đột	Có thể chạy nhiều ứng dụng trên nhiều máy khác nhau



Hình 1. Sự khác nhau giữa hosted (trái) và bare-metal (phải)

2.2. Điện toán đám mây

Ngày nay, cụm từ “điện toán đám mây” được nhắc đến trên rất nhiều trong thông tin và các phương tiện truyền thông. Rất nhiều các dịch vụ điện toán đám mây ra đời, có thể kể đến như: iCloud, Dropbox, dịch vụ cho thuê máy chủ Amazon EC2, Vậy: “Điện toán đám mây là gì?”. Điện toán đám mây (hay Cloud Computing) hiện vẫn là một mô hình đang tiến hóa (evolving paradism), vì vậy chưa có một định nghĩa thống nhất về mô hình này và các khái niệm hiện nay về điện toán đám mây mới chỉ tạm thời. Các định nghĩa được nêu trong công trình này dựa trên các tài liệu được sử dụng rộng rãi như Wikipedia, hay như định nghĩa của NIST - National Insitute of Standards and Technology.

Theo NIST [3]: “Điện toán đám mây là một mô hình cho phép truy cập mạng dễ dàng, theo yêu cầu tới một tài nguyên điện toán được chia

sẽ (như mạng, máy chủ, không gian lưu trữ, ứng dụng, và dịch vụ). Các tài nguyên này có thể được cung cấp nhanh chóng và thu hồi lại với chi phí quản lý và tương tác với nhà cung cấp dịch vụ tối thiểu.”

Nói chung thì một mô hình được gọi là điện toán đám mây nếu chúng tuân thủ theo năm đặc tính, bốn mô hình triển khai, và ba mô hình dịch vụ. Sau đây, chúng tôi sẽ trình bày sơ lược về các đặc điểm này [12]:

- 05 đặc tính:
 - + On-demand self-service: Khả năng tự phục vụ của người dùng, chủ động khởi tạo, tạm dừng dịch vụ,
 - + Broad network access: Khả năng truy cập trên mọi nền tảng thiết bị, mọi loại hạ tầng về mạng, khu vực địa lý.
 - + Resource pooling: Khả năng gộp - gom tài nguyên vật lý, sau đó, phân bổ một cách tự động

cho người sử dụng dựa theo nhu cầu.

+ Rapid elasticity: Khả năng co giãn và đàn hồi tài nguyên một cách nhanh chóng thuận tiện.

+ Measured service: Khả năng đo lường dịch vụ để kiểm soát thời gian sử dụng, từ đó tính toán chi phí theo mức độ sử dụng dịch vụ.

- 04 mô hình triển khai:

+ Private cloud: Các cơ sở hạ tầng đám mây được cung cấp độc quyền cho riêng một cá nhân hoặc tổ chức cụ thể. Các đám mây này được sở hữu, quản lý và điều hành bởi chính các khách hàng đó, từ một bên thứ ba, hoặc cả hai.

+ Community Cloud – đám mây cho cộng đồng: Các cơ sở hạ tầng đám mây được cung cấp độc quyền cho một cộng đồng khách hàng cụ thể có cùng các mối quan tâm (ví dụ như yêu cầu an ninh...). Các đám mây này được sở hữu và điều hành bởi một hoặc nhiều tổ chức trong cộng đồng, từ một bên thứ ba, hoặc cả hai.

+ Public Cloud - Đám mây công cộng: Với mô hình triển khai này, các hạ tầng điện toán

đám mây được cung cấp cho mọi đối tượng khách hàng, cho mọi mục đích khác nhau. Các đám mây này được quản trị bởi nhà cung cấp dịch vụ điện toán đám mây.

+ Hybrid Cloud - Đám mây lai: Mô hình triển khai này là tổ hợp của hai hay nhiều mô hình triển khai ở trên (private, community hoặc public cloud).

- 03 mô hình dịch vụ:

+ Software as a Service (SaaS): Cung cấp các dịch vụ về phần mềm, bán hoặc cho thuê lâu dài.

+ Platform as a Service (PaaS): Cung cấp dịch vụ nền tảng như database, môi trường phát triển chương trình, ...

+ Infrastructure as a Service (IaaS): Cung cấp dịch vụ về hạ tầng, các máy chủ, tài nguyên như Ram, CPU, storage, ...

Hiện tại một số platform (nền tảng) có thể được sử dụng để triển khai cloud computing như: euca-lyptus, cloudstack, openstack. Trong phần này, chúng tôi so sánh chúng theo một số tiêu chí sau:

Storage (Lưu trữ):

	OpenStack	CloudStack	Eucalyptus	OpenNebula
Disk image	Yes	Yes	Yes	Yes
Block device	Yes	Yes	Yes	Yes
Fault tolerance	Yes	Yes	Yes	Yes

VM image (Virtual Machine image):

	OpenStack	CloudStack	Eucalyptus	OpenNebula
Image service	Yes	Yes	Yes	No
Seft service	Yes	Yes	No	No
Amazon API	Yes	Yes	Yes	No

Networking (Mạng):

	OpenStack	CloudStack	Eucalyptus	OpenNebula
Auto allocation	Yes	Yes	Yes	No
Floating IP	Yes	No	Yes	No
Layer 2	Yes	Yes	Yes	No

Thông thường, khi lựa chọn platform phải quan tâm đến: Kích thước của hệ thống triển khai trên đó, độ phức tạp hệ thống, khả năng chịu lỗi, sự thích ứng với các đám mây khác. Trong công trình này, chúng tôi lựa chọn Openstack vì các lý do sau:

- Openstack còn “non trẻ” và còn tiến hóa được thêm nữa.
- Có sự ủng hộ các công ty lớn: IBM, Google, Amazone, ...
- Sử dụng một ngôn ngữ duy nhất python.
- Triển khai với quy mô lớn.
- Cung cấp toàn bộ API.
- Cộng đồng phát triển lớn.

3. Xây dựng và triển khai Openstack

3.1. Giới thiệu Openstack

Openstack là một nền tảng mã nguồn mở, được thành lập bởi NASA và Rack-space Hosting vào năm 2010, được sử dụng để xây dựng private cloud (dịch vụ đám mây nội bộ) hay public cloud (dịch vụ đám mây công cộng).

Đặc điểm của openstack là được thiết kế theo tính module và mở về thiết kế với 99.99% được viết bằng python. Sau khi ra đời, openstack được cộng đồng phát triển mạnh mẽ. Tính đến 04/2014, cộng đồng opestack có mặt trên 200 nước với gần 19 nghìn người tham gia [11, 12]. Openstack được phát triển nhiều phiên bản, trung bình cứ 06 tháng thì có một phiên bản mới ra đời. Tính đến nay phiên bản mới nhất với tên là Kilo.

3.2. Kiến trúc Openstack

Openstack được xây dựng theo tính module và có thể lựa chọn module để triển khai. Hiện tại

bản Juno có tất cả 11 projects (xem hình 2):

- Horizon: Cung cấp giao diện cho người dùng - tương tác với Openstack và tương tác với API của dịch vụ;
- Keystone: Xác thực và ủy quyền trong Openstack (tạo, xóa, sửa, xác thực người dùng);
- Nova: Tạo sửa, xóa, quản lý vòng đời của máy ảo. Hỗ trợ nhiều công nghệ ảo hóa như KVM, Vware, QEMU, ...
- Glance: Lưu trữ và truy vấn các disk image, hỗ trợ nhiều định dạng như vmdk, vhd, qcow, ...
- Swift: Đọc và ghi các đối tượng (chính là các file) thông qua HTTP.
- Neutron: Cung cấp các giải pháp mạng trong Openstack
- Cinder: Cung cấp volume trong Openstack
- Ceilometer: Thống kê, giám sát tài nguyên mà người sử dụng đã sử dụng.

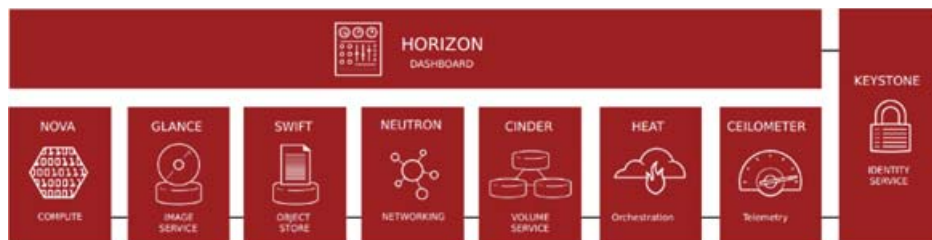
3.3. Triển khai Openstack

Openstack Juno được cài đặt từ trang chủ của Openstack [11]. Mô hình cài đặt trên ba node như hình 3 với công nghệ ảo hóa KVM, trên hệ điều hành Ubuntu 14.04. Trong phần này để đơn giản, chỉ thực hiện trên một node compute. Tuy nhiên, có thể mở rộng dễ dàng bằng việc bổ sung thêm các node compute.

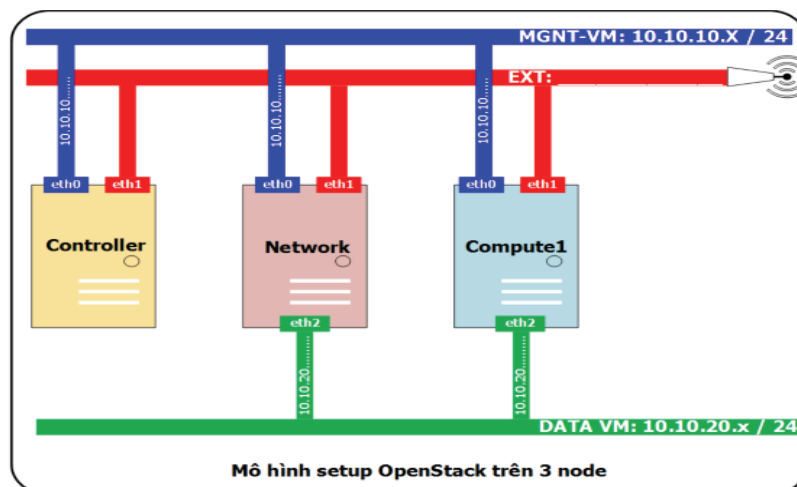
Trong hình 3, controller đóng vai trò điều khiển hoạt động của các node trong quá trình tính toán. Trong khi network đóng vai trò điều khiển hoạt động dịch vụ mạng (networking) phục vụ giao tiếp giữa bên ngoài và trong hệ thống. Cuối cùng, các node compute đóng vai trò tính toán, đây chính là tài nguyên tính toán trong Openstack. Theo mô hình chuẩn của Openstack, các gói cài đặt trên các node sẽ như hình 4.

Chi tiết các bước cài đặt có thể tham khảo tại [11, 12]. Trong bài báo này chúng tôi không trình bày chi tiết lại mà chỉ thực hiện triển khai Openstack ứng dụng trong hệ thống cung cấp thông tin KTTV cho thiết bị di động. Sử dụng project có tên là dashboard để tạo môi trường điện toán đám mây trong Openstack (xem hình 5 và 6). Trong giao diện dashboard, chúng ta có thể tạo các network, instance, volume, Tại đây, tạo một web server trên đám mây chạy hệ thống cung cấp thông tin KTTV. Thông số khởi tạo một webserver và lưu trữ dữ liệu trên máy ảo với cấu hình: CPU 8 nhân, RAM 26 GB, ổ cứng 500GB. Máy ảo được cài đặt hệ điều hành Window server 2008 (hình 5 và hình 6).

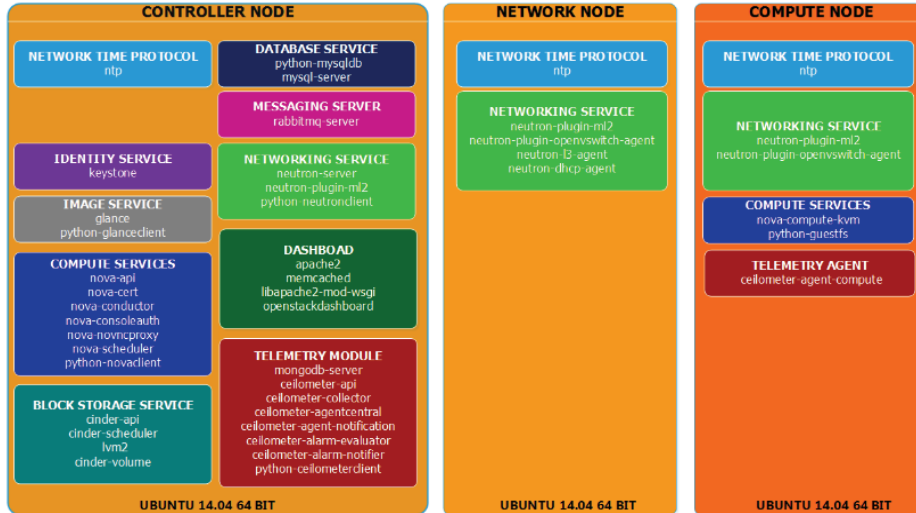
Sau khi thiết lập xong môi trường điện toán đám mây, chúng ta sẽ triển khai hệ thống thông tin KTTV trên đó.



Hình 2. Kiến trúc tổng thể của Openstack



Hình 3. Mô hình cài đặt Openstack trên 3 node.

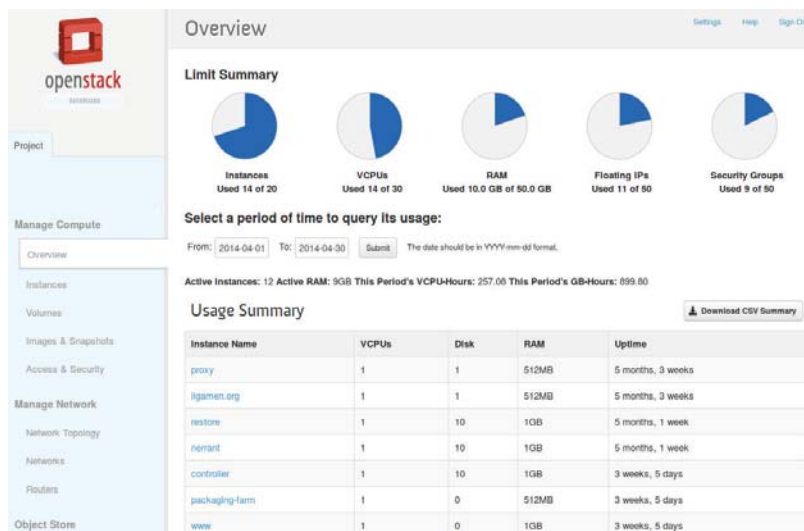


Hình 4. Các gói cài đặt trên từng node trong openstack juno

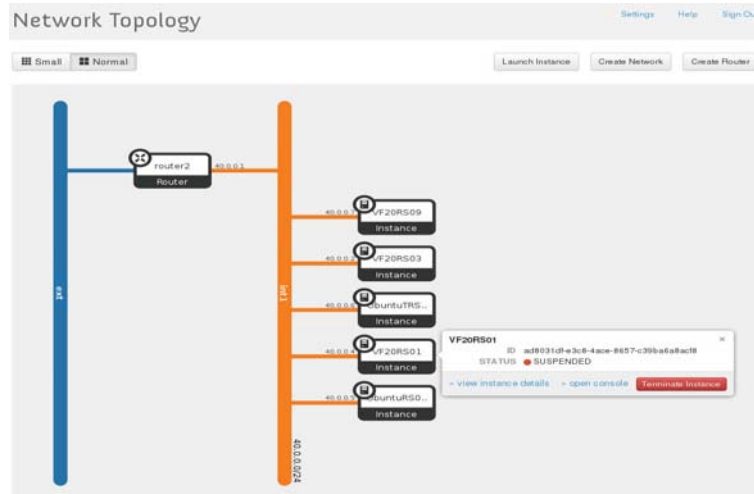
4. Kết luận

Bài báo giới thiệu về việc xây dựng và triển khai mô hình điện toán đám mây ứng dụng trong hệ thống cung cấp thông tin khí tượng thủy văn cho thiết bị di động. Hiện tại, có một số lựa chọn để triển khai điện toán đám mây như: openstack,

clouds-tack, dev-stack, Trong công trình này, chúng tôi chọn giải pháp openstack. Các lần thực nghiệm hệ thống trên nền tảng Openstack cho thấy hệ thống chạy ổn định, cân bằng tải tốt, và có thể mở rộng tài nguyên tính một cách dễ dàng.



Hình 5. Giao diện dashboard của Openstack



Hình 6. Các mạng instance ảo được tạo ra trong điện toán đám mây

Tài liệu tham khảo

1. Website KTTV <http://www.nchmf.gov.vn>, <http://www.jma.go.jp>, <http://www.hko.gov.hk>
2. <http://en.wikipedia.org/wiki/Virtualization>
3. <http://www.csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
4. <http://www.slideshare.net/OReillyOSCON/comparing-open-sourceprivate-cloud-platforms>
5. <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
6. <http://www.vmware.com/pdf/virtualization.pdf>
7. <https://www.gartner.com/doc/1572031/road-map-virtualization-cloudcomputing>
8. <http://www.slideshare.net/ProfEdge/introduction-to-cloud-computing-23970527>
9. <http://www.slideshare.net/OReillyOSCON/comparing-open-source-privatecloud-Platforms>
10. <http://www.slideshare.net/bligneri/comparison-of-open-source-virtualizationtechnology>
11. <http://docs.openstack.org/juno/install-guide/install/apt/content/>
12. <https://vietstack.wordpress.com/>
13. <https://sites.google.com/site/hethongcungcaphongtinktvtv/>.

DEPLOY CLOUD COMPUTING FOR HYDRO-METEOROLOGICAL INFORMATION SUPPLYING SYSTEM BASED ON MOBILE DEVICES

Ngo Van Manh, Nguyen Van Khoa -HydroMeteorological Data Center
Ban Ha Bang - Ha Noi University of Science and Technology

At present, National Hydrometeorological Center has been providing information through traditional many ways such as email, fax, telephones, televisions, radios and documents at request. These methods have several disadvantages such as limited broadcasting time and lack of updated information. To overcome mentioned draw-backs, we built a system to provide information in mobilephone. The system allows users to view Hydrometeorological information anytime, sends early warnings to users' mobile when disasters occur. However, to deploy the system which meets millions of visits requires a stable computing infrastructure such as on-demand self-service, broad network access, resource pooling, rapid elasticity. Currently, the computing infrastructure which supports these features are known as cloud computing. In this paper, we deploy cloud computing infrastructure for hydro-meteorological information supplying system.

Keywords: cloud computing, hydro-meteorological forecasting system, mobile