

*Trường Đại Học Công Nghệ Thông Tin
TPHCM*

ĐỀ TÀI: ĐIỆN TOÁN ĐÁM MÂY

Giảng viên hướng dẫn: Ths.Ngô Hán Chiêu

Sinh viên thực hiện:

Võ Xuân Thiên Phúc 08520290

Trương Thiên Toàn 08520511

Mục lục

n: n: n:

Chương 1. Tổng quan về Cloud Computing.....	2
I. Đặt vấn đề.....	2
II. Định nghĩa	2
III. Mô hình tổng quan.....	3
IV. Các giải pháp.....	5
V. Đặc điểm của cloud computing	6
VI. Xu hướng phát triển.....	7
VII. Các khó khăn, thách thức	8
Chương 2. Hiện thực của Cloud Computing.....	9
I. Hiện thực cloud computing:	9
II. Một số ví dụ về cách tính chi phí trong Cloud	12
III. Market oriented Cloud computing:.....	12
IV. Mô phỏng Cloud:.....	14
Chương 3. So sánh giữa Cloud Computing và Grid Computing	19
I. Tổng quan.....	19
II. So sánh giữa Grids và Clouds.....	21

Chương 1. Tổng quan về Cloud Computing

I. Đặt vấn đề

Ngày nay, đối với các công ty, doanh nghiệp, việc quản lý tốt, hiệu quả dữ liệu của riêng công ty cũng như dữ liệu khách hàng, đối tác là một trong những bài toán được ưu tiên hàng đầu và đang không ngừng gây khó khăn cho họ. Để có thể quản lý được nguồn dữ liệu đó, ban đầu các doanh nghiệp phải đầu tư, tính toán rất nhiều loại chi phí như chi phí cho phần cứng, phần mềm, mạng, chi phí cho quản trị viên, chi phí bảo trì, sửa chữa, ... Ngoài ra họ còn phải tính toán khả năng mở rộng, nâng cấp thiết bị; phải kiểm soát việc bảo mật dữ liệu cũng như tính sẵn sàng cao của dữ liệu.

Từ một bài toán điển hình như vậy, chúng ta thấy được rằng nếu có một nơi tin cậy giúp các doanh nghiệp quản lý tốt nguồn dữ liệu đó, các doanh nghiệp sẽ không còn quan tâm đến cơ sở hạ tầng, công nghệ mà chỉ tập trung chính vào công việc kinh doanh của họ thì sẽ mang lại cho họ hiệu quả và lợi nhuận ngày càng cao hơn.

Thuật ngữ “cloud computing” ra đời bắt nguồn từ một trong những hoàn cảnh như vậy.

Thuật ngữ “cloud computing” còn được bắt nguồn từ ý tưởng đưa tất cả mọi thứ như dữ liệu, phần mềm, tính toán, ... lên trên mạng Internet. Chúng ta sẽ không còn trông thấy các máy PC, máy chủ của riêng các doanh nghiệp để lưu trữ dữ liệu, phần mềm nữa mà chỉ còn một số các “máy chủ ảo” tập trung ở trên mạng. Các “máy chủ ảo” sẽ cung cấp các dịch vụ giúp cho doanh nghiệp có thể quản lý dữ liệu dễ dàng hơn, họ sẽ chỉ trả chi phí cho lượng sử dụng dịch vụ của họ, mà không cần phải đầu tư nhiều vào cơ sở hạ tầng cũng như quan tâm nhiều đến công nghệ. Xu hướng này sẽ giúp nhiều cho các công ty, doanh nghiệp vừa và nhỏ mà không có cơ sở hạ tầng mạng, máy chủ để lưu trữ, quản lý dữ liệu tốt.

Vậy “cloud computing” là gì ? Nó có thể giải quyết bài toán trên như thế nào và có những đặc điểm nổi bật gì ? Chúng ta sẽ đi qua các phần sau để nắm rõ vấn đề này.

II. Định nghĩa

Theo Wikipedia:

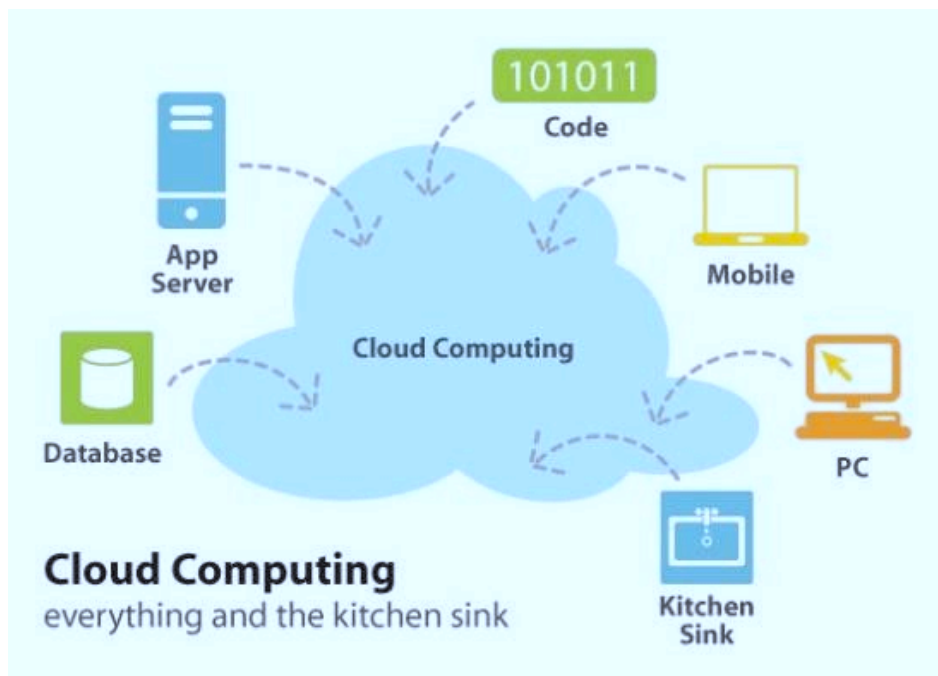
“Điện toán đám mây (cloud computing) là một mô hình điện toán có khả năng co giãn (scalable) linh động và các tài nguyên thường được ảo hóa được cung cấp như một dịch vụ trên mạng Internet”.

Theo Gartner (<http://www.buildingthecloud.co.uk/>) :

“Một mô hình điện toán nơi mà khả năng mở rộng và linh hoạt về công nghệ thông tin được cung cấp như một dịch vụ cho nhiều khách hàng đang sử dụng các công nghệ trên 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 (platform) 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”.

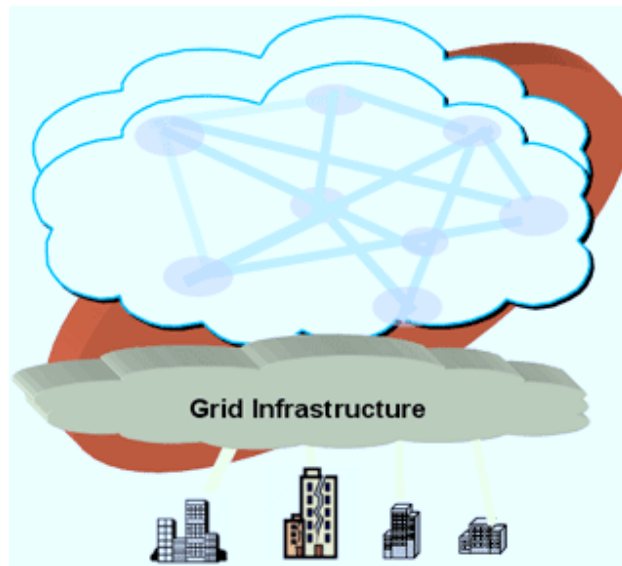


Hình 1: Mọi thứ đều tập trung vào đám mây

III. Mô hình tổng quan

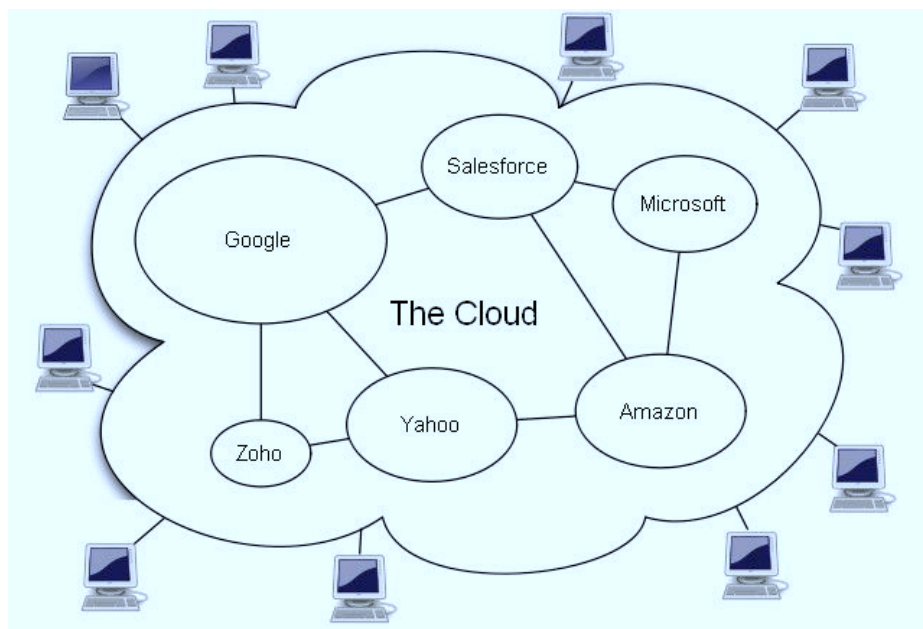
Theo định nghĩa, các nguồn điện toán khổng lồ như phần mềm, dịch vụ ... sẽ nằm tại các máy chủ ảo (đám mây) trên Internet thay vì trong máy tính gia đình và văn phòng (trên mặt đất) để mọi người kết nối và sử dụng mỗi khi họ cần.

CLOUD COMPUTING



Hình 2: Minh họa về cloud computing

Hiện nay, các nhà cung cấp đưa ra nhiều dịch vụ của cloud computing theo nhiều hướng khác nhau, đưa ra các chuẩn riêng cũng như cách thức hoạt động khác nhau. Do đó, việc tích hợp các cloud để giải quyết một bài toán lớn của khách hàng vẫn còn là một vấn đề khó khăn. Chính vì vậy, các nhà cung cấp dịch vụ đang có xu hướng tích hợp các cloud lại với nhau thành “sky computing”, đưa ra các chuẩn chung để giải quyết các bài toán lớn của khách hàng.



Hình 3: Mô hình tổng quan của cloud computing

IV. Các giải pháp

Cloud Computing ra đời để giải quyết các vấn đề sau:

Vấn đề về lưu trữ dữ liệu:

Dữ liệu được lưu trữ tập trung ở các kho dữ liệu khổng lồ. Các công ty lớn như Microsoft, Google có hàng chục kho dữ liệu trung tâm nằm rải rác khắp nơi trên thế giới. Các công ty lớn này sẽ cung cấp các dịch vụ cho phép doanh nghiệp có thể lưu trữ và quản lý dữ liệu của họ trên các kho lưu trữ trung tâm.

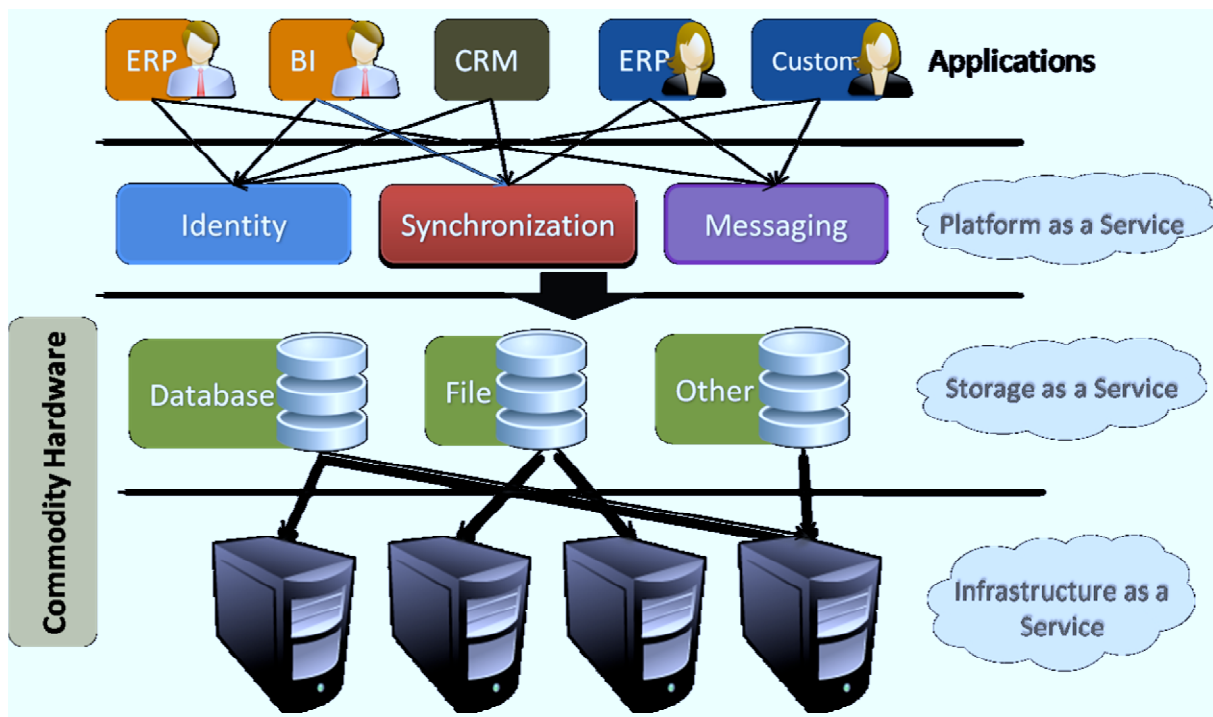
Vấn đề về sức mạnh tính toán:

Có 2 giải pháp chính:

- Sử dụng các siêu máy tính (super-computer) để xử lý tính toán.
- Sử dụng các hệ thống tính toán song song, phân tán, tính toán lưới (grid computing).

Vấn đề về cung cấp tài nguyên, phần mềm:

Cung cấp các dịch vụ như IaaS (infrastructure as a service), PaaS (platform as a service), SaaS (software as a service). Chương 2 sẽ đi sâu hơn về các dịch vụ này.



Hình 4: Minh họa về các dịch vụ

V. Đặc điểm của cloud computing

Nhanh chóng cải thiện với người dùng có khả năng cung cấp sẵn các tài nguyên cơ sở hạ tầng công nghệ một cách nhanh chóng và ít tốn kém.

Chi phí được giảm đáng kể và chi phí vốn đầu tư được chuyển sang hoạt động chi tiêu. Điều này làm giảm rào cản cho việc tiếp nhận, chẳng hạn như cơ sở hạ tầng được cung cấp bởi đối tác thứ 3 và không cần phải mua để dùng cho các tác vụ tính toán thực hiện 1 lần hay chuyên sâu mà không thường xuyên. Việc định giá dựa trên cơ sở tính toán theo nhu cầu thì tốt đối với những tùy chọn dựa trên việc sử dụng và các kỹ năng IT được đòi hỏi tối thiểu (hay không được đòi hỏi) cho việc thực thi.

Sự độc lập giữa thiết bị và vị trí làm cho người dùng có thể truy cập hệ thống bằng cách sử dụng trình duyệt web mà không quan tâm đến vị trí của họ hay thiết bị nào mà họ đang dùng, ví dụ như PC, mobile. Vì cơ sở hạ tầng off-site (được cung cấp bởi đối tác thứ 3) và được truy cập thông qua Internet, do đó người dùng có thể kết nối từ bất kỳ nơi nào.

Việc cho thuê nhiều để có thể chia sẻ tài nguyên và chi phí giữa một phạm vi lớn người dùng, cho phép:

- Tập trung hóa cơ sở hạ tầng trong các lĩnh vực với chi phí thấp hơn (chẳng hạn như bất động sản, điện, v.v.)
- Khả năng chịu tải nâng cao (người dùng không cần kỹ sư cho các mức tải cao nhất có thể).
- Cải thiện việc sử dụng và hiệu quả cho các hệ thống mà thường chỉ 10-20% được sử dụng.

Độ tin cậy cải thiện thông qua việc sử dụng các site có nhiều dư thừa, làm nó thích hợp cho tính liên tục trong kinh doanh và khôi phục thất bại. Tuy nhiên, phần lớn các dịch vụ của cloud computing có những lúc thiếu hụt và người giám đốc kinh doanh, IT phải làm cho nó ít đi.

Tính co giãn linh động (“theo nhu cầu”) cung cấp tài nguyên trên một cơ sở mịn, tự bản thân dịch vụ và gần thời gian thực, không cần người dùng phải có kỹ sư cho chịu tải. Hiệu suất hoạt động được quan sát và các kiến trúc nhất quán, kết nối lỏng lẻo được cấu trúc dùng web service như giao tiếp hệ thống.

Việc bảo mật cải thiện nhờ vào tập trung hóa dữ liệu, các tài nguyên chú trọng bảo mật, v.v... nhưng cũng nâng cao mối quan tâm về việc mất quyền điều khiển dữ liệu nhạy cảm. Bảo mật thường thì tốt hay tốt hơn các hệ thống truyền thống, một phần bởi các nhà cung cấp có thể dành nhiều nguồn lực cho việc giải quyết các vấn đề bảo mật mà nhiều khách hàng không có đủ chi phí để thực hiện. Các nhà cung cấp sẽ ghi nhớ (log) các truy cập, nhưng việc truy cập vào chính bản thân các audit log có thể khó khăn hay không thể.

Khả năng chịu đựng xảy ra thông qua việc tận dụng tài nguyên đã được cải thiện, các hệ thống hiệu quả hơn. Tuy nhiên, các máy tính và cơ sở hạ tầng kết hợp là những thứ tiêu thụ năng lượng chủ yếu.

VI. Xu hướng phát triển

Thuật ngữ “cloud computing” ra đời từ giữa năm 2007, cho đến nay đã không ngừng phát triển mạnh mẽ và được hiện thực bởi nhiều công ty lớn trên thế giới như IBM, Sun, Amazon, Google, Microsoft, Yahoo, Salesforce, ...



Hình 5: Một số nhà cung cấp dịch vụ cloud computing

Với mục tiêu giải quyết các bài toán về dữ liệu, tính toán, dịch vụ, ... cho khách hàng, cloud computing đã và đang mang lại lợi nhuận lớn, đem đến một sân chơi, một thị trường rộng lớn cho các nhà cung cấp dịch vụ, nên sự phát triển nhanh chóng của nó có thể được tính bằng từng ngày. Trong khi đó, thuật ngữ ban đầu của “grid computing” không mang tính kinh tế, lợi nhuận cao nên sự phát triển của nó đang ngày càng giảm sút, và chỉ đang được áp dụng vào lĩnh vực nghiên cứu khoa học.



Hình 6: Xu hướng phát triển

Theo báo VNUnet (28/04/2009), trong một cuộc khảo sát các khách hàng sử dụng các dịch vụ của cloud computing thì có tới 27,7% quan tâm về vấn đề lưu trữ dữ liệu (data storage); 17% về các ứng dụng liên quan tài chính; 12,8% về e-mail.

VII. Các khó khăn, thách thức

Trong quá trình hiện thực cloud computing, người ta nhận thấy một số khó khăn, thách thức sau:

Bảo mật

- Sở hữu trí tuệ (Intellectual property)
- Tính riêng tư (Privacy)
- Độ tin cậy (Trust)

Khả năng không kiểm soát dữ liệu

Độ trễ dữ liệu

Tính sẵn sàng của dịch vụ, dữ liệu

Các dịch vụ kèm theo

Các quy định pháp luật cho các dịch vụ, giữa khách hàng và nhà cung cấp

Chương 2. Hiện thực của Cloud Computing

I. Hiện thực cloud computing:

Các thông tin được tham khảo trên Wikipedia <http://en.wikipedia.org/wiki>

Cloud computing cung cấp hạ tầng, nền tảng và phần mềm như là dịch vụ, mà có thể được cung ứng như là một dịch vụ cho thuê trong cách dùng bao nhiêu trả bấy nhiêu đối với người dùng.

Cloud computing được hiện thực theo 3 kiểu:

u: u: u:u:u: u:u: iu: ụ ạ à iac iac á u: uờ ạ à u:
u: uờ u: uớ u: u:ứ ảiac u u: ô iu: u

Những kiến trúc ảo xếp chồng là một ví dụ của xu hướng mọi thứ là dịch vụ và có cùng những điểm chung. Hơn hẳn một máy chủ cho thuê, không gian lưu trữ tập trung hay thiết bị mạng, máy trạm thay vì đầu tư mua những nguyên thì có thể thuê đầy đủ dịch vụ bên ngoài. Những dịch vụ này thông thường được tính chi phí trên cơ sở tính toán chức năng và lượng tài nguyên sử dụng (và từ đó ra chi phí) sẽ phản ánh được mức độ của hoạt động. Đây là một sự phát triển của những giải pháp lưu trữ web và máy chủ cá nhân ảo.

Tên ban đầu được sử dụng là dịch vụ phần cứng (HaaS) và được tạo ra bởi một nhà kinh tế học Nicholas Carr vào tháng 3 năm 2006, nhưng điều này cần thiết. Nhưng từ này đã dần bị thay thế bởi khái niệm dịch vụ hạ tầng vào khoảng cuối năm 2006.

Những đặc trưng tiêu biểu:

Cung cấp tài nguyên như là dịch vụ: bao gồm cả máy chủ, thiết bị mạng, bộ nhớ, CPU, không gian đĩa cứng, trang thiết bị trung tâm dữ liệu.

Khả năng mở rộng linh hoạt

Chi phí thay đổi tùy theo thực tế

Nhiều người thuê có thể cùng dùng chung trên một tài nguyên

Cấp độ doanh nghiệp: đem lại lợi ích cho công ty bởi một nguồn tài nguyên tích
toán tổng hợp

Các ví dụ: Amazon EC2/S3, Elastra (Beta 2.0 2/2009), Nirvanix, AppNexus

pppp pp i u ề ả ia OD á OD ể ứ
 ụ ia ộ ết ắt t ừ tự

Cung cấp nền tảng tính toán và một tập các giải pháp nhiều lớp. Nó hỗ trợ việc triển khai ứng dụng mà không quan tâm đến chi phí hay sự phức tạp của việc trang bị và quản lý các lớp phần cứng và phần mềm bên dưới, cung cấp tất cả các tính năng cần thiết để hỗ trợ chu trình sống đầy đủ của việc xây dựng và cung cấp một ứng dụng và dịch vụ web sẵn sàng trên Internet mà không cần bất kỳ thao tác tải hay cài đặt phần mềm cho những người phát triển, quản lý tin học, hay người dùng cuối. Nó còn được biết đến với một tên khác là cloudware.

Cung cấp dịch vụ nền tảng (PaaS) bao gồm những điều kiện cho qui trình thiết kế ứng dụng, phát triển, kiểm thử, triển khai và lưu trữ ứng dụng có giá trị như là dịch vụ ứng dụng như cộng tác nhóm, sắp xếp và tích hợp dịch vụ web, tích hợp cơ sở dữ liệu, bảo mật, khả năng mở rộng, quản lý trạng thái, phiên bản ứng dụng, các lợi ích cho cộng đồng phát triển và nghiên cứu ứng dụng. Những dịch vụ này được chuẩn bị như là một giải pháp tính hợp trên nền web.

Những đặc trưng tiêu biểu:

Phục vụ cho việc phát triển, kiểm thử, triển khai và vận hành ứng dụng giống như là môi trường phát triển tích hợp
 Các công cụ khởi tạo với giao diện trên nền web.
 Kiến trúc đồng nhất
 Tích hợp dịch vụ web và cơ sở dữ liệu
 Hỗ trợ cộng tác nhóm phát triển
 Công cụ hỗ trợ tiện tích

Các yếu tố:

Thuận lợi:

- Dịch vụ nền tảng (PaaS) đang ở thời kì đầu và được ưa chuộng ở những tính năng vốn được ưa thích bởi dịch vụ phần mềm (đề cập sau), bên cạnh đó có tích hợp các yếu tố về nền tảng hệ thống.
- Ưu điểm trong những dự án tập hợp những công việc nhóm có sự phân tán về địa lý.
- Khả năng tích hợp nhiều nguồn của dịch vụ web
- Giảm chi phí ngoài lề khi tích hợp các dịch vụ về bảo mật, khả năng mở rộng, kiểm soát lỗi...
- Giảm chi phí khi trừu tượng hóa công việc lập trình ở mức cao để tạo dựng vụ, giao diện người dùng và các yếu tố ứng dụng khác.

- Mong đợi ở người dùng có kiến thức có thể tiếp tục hoàn thiện và hỗ trợ tương tác với nhiều người để giúp xác định mức độ khó khăn của vấn đề chúng ta gặp phải.
- Hướng việc sử dụng công nghệ để đạt được mục đích tạo điều kiện dễ dàng hơn cho việc phát triển ứng dụng đa người dùng cho những người không chỉ trong nhóm lập trình mà có thể kết hợp nhiều nhóm cùng làm việc

Khó khăn:

- Ràng buộc bởi nhà cung cấp: do giới hạn phụ thuộc và dịch vụ của nhà cung cấp
- Giới hạn phát triển: độ phức tạp khiến nó không phù hợp với yêu cầu phát triển nhanh vì những tính năng phức tạp khi hiện thực trên nền tảng web.

zoso (2/2008), Google App Engine, Salesforce, Heroku, Engine Yard

e e e e e e ie ụ à ã n khô á ie ụ e à êe e ư e
ia ừ

Dịch vụ phần mềm (SaaS) là một mô hình triển khai ứng dụng mà ở đó người cung cấp cho phép người dùng sử dụng dịch vụ theo yêu cầu. Những nhà cung cấp SaaS có thể lưu trữ ứng dụng trên máy chủ của họ hoặc tải ứng dụng xuống thiết bị khách hàng, vô hiệu hóa nó sau khi kết thúc thời hạn. Các chức năng theo yêu cầu có thể được kiểm soát bên trong để chia sẻ bản quyền của một nhà cung cấp ứng dụng thứ ba.

Những đặc trưng tiêu biểu

Phần mềm sẵn có đòi hỏi việc truy xuất, quản lý qua mạng.

Quản lý các hoạt động từ một vị trí tập trung hơn là tại mỗi nơi của khách hàng, cho phép khách hàng truy xuất từ xa thông qua web.

Cung cấp ứng dụng thông thường gần gũi với mô hình ánh xạ từ một đến nhiều hơn là mô hình 1:1 bao gồm cả các đặc trưng kiến trúc, giá cả và quản lý.

Những tính năng tập trung nâng cấp, giải phóng người dùng khỏi việc tải các bản vá lỗi và cập nhật.

Thường xuyên tích hợp những phần mềm giao tiếp trên mạng diện rộng

Các ví dụ: 3Tera (2/2006), Salesforce

II. Một số ví dụ về cách tính chi phí trong Cloud

Định giá cố định: nhà cung cấp sẽ xác định rõ đặc tả về khả năng tính toán cố định (dung lượng bộ nhớ được cấp phát, loại CPU và tốc độ .v.v...)

Định giá theo đơn vị: được áp dụng phổ biến cho lượng dữ liệu truyền tải, dung lượng bộ nhớ được cấp phát và sử dụng,... cách này uyển chuyển hơn cách trên

Định giá theo thuê bao: ứng dụng phần lớn trong mô hình dịch vụ phần mềm (SaaS) người dùng sẽ tiên đoán trước định mức sử dụng ứng dụng cloud (cách tính này thường khó đạt được độ chính xác cao)

III. Market oriented Cloud computing:

được cập nhật từ R. Buyya đề nghị trong bài báo

“Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities”

Với những tiến bộ của xã hội con người hiện đại, những dịch vụ thiết yếu được cung cấp rộng rãi để mọi người đều có thể tiếp cận như điện, nước, gas và điện thoại đã đóng vai trò quan trọng trong cuộc sống hằng ngày. Những dịch vụ tiện ích này có thể được sử dụng thương mại cần được sẵn sàng ở mọi nơi mà người dùng mong muốn vào mọi lúc. Khác hàng sau đó có thể trả cho nhà cung cấp dịch vụ dựa trên lượng sử dụng các dịch vụ tiện ích đó. Tương tự như các ví dụ trên, giữa nhà cung cấp dịch vụ và người sử dụng cần có những thỏa thuận cụ thể được nêu trong SLA (Service Level Agreement) trong đó xác định về yêu cầu chất lượng dịch vụ QoS (Quality of Service)

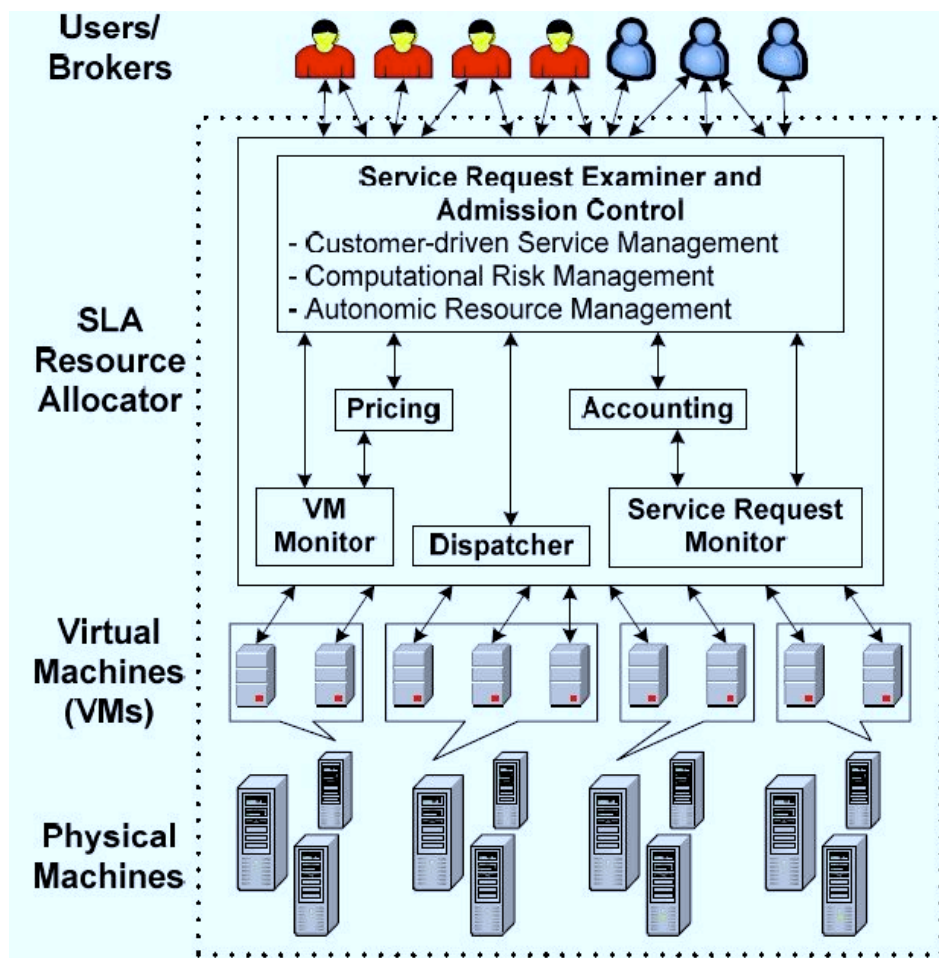
Đặc điểm của mô hình Cloud computing

Khách hàng phụ thuộc vào nhà cung cấp dịch vụ Cloud cung cấp tất cả năng lực tính toán họ cần, họ yêu cầu về chất lượng dịch vụ QoS phải được duy trì bởi nhà công cấp để phù hợp với mục tiêu và đảm bảo hoạt động của họ. Nhà cung cấp Cloud sẽ cần xem xét để đáp ứng các yêu cầu về dịch vụ của mỗi khách hàng để thương lượng với SLA cụ thể. Để đạt được điều này, nhà cung cấp dịch vụ Cloud không thể tiếp tục triển khai trên mô hình quản lý tài nguyên tập trung mà ở đó không cung cấp cho họ khả năng chia sẻ tài nguyên mà vẫn đáp ứng được tất cả các yêu cầu về dịch vụ. Thay vào đó, mô hình quản lý market-oriented được đề cập để đạt được sự cân đối giữa nhà cung cấp và khách hàng.

ể hia s ò hia s ù ể

User/Broker (Người dùng/ Nhà môi giới): người dùng hay nhà phân phối sử dụng quyền ủy thác để gửi yêu cầu dịch vụ từ bất kỳ đâu trên thế giới tới Data center (trung tâm dữ liệu) hay Cloud để được xử lý

SLA Resource Allocator (Bộ phân phối tài nguyên SLA): đóng vai trò như một trung gian giữa các nhà cung cấp Data center và Cloud với người dùng/ nhà môi giới bên ngoài.



Kiến trúc Market-Oriented Cloud

Service Request Examiner and Admission Control: khi một yêu cầu dịch vụ được gửi lên lần đầu sẽ được phiên dịch thành các yêu cầu về chất lượng dịch vụ QoS trước khi xác định xem nó sẽ được chấp nhận hay từ chối. Do vậy, điều đó đảm bảo rằng không có tình trạng quá tải dịch vụ khi mà các yêu cầu dịch vụ không thể được đáp ứng đầy đủ vì giới hạn tài nguyên hệ thống sẵn

sàng. Nó cung cấp thông tin về trạng thái cuối cùng về tình trạng sẵn sàng của tài nguyên (từ cơ chế VM Monitor) và khả năng xử lý tải (từ cơ chế Service Request Monitor) theo thứ tự để quyết định việc phân phối tài nguyên một cách hiệu quả. Sau đó nó sẽ phân yện cầu cho các máy ảo VM và xác định đặc tả tài nguyên cho máy ảo được phân.

Pricing: cung cấp cơ chế quyết định cách các yêu cầu service được tính tiền. Ví dụ như yêu cầu có thể được tính tiền dựa theo thời gian các nhiệm vụ, tỷ lệ giá cả (cố định/thay đổi) hay tính sẵn sàng của tài nguyên (sẵn có/yêu cầu). Cơ chế định giá phục vụ như nền tảng cho cung cấp và yêu cầu tài nguyên tính toán trong Data center và các trang thiết bị trong việc cấp phát tài nguyên hiệu quả.

Accounting: cung cấp cơ chế để thao tác trên lưu lượng dùng tài nguyên được yêu cầu do đó chi phí cuối cùng có thể được tính toán và tính phí cho người dùng. Thêm vào đó, lịch sử sử dụng có thể được dùng để tối ưu bởi Service Request Examiner and Admission Control

VM Monitor: cung cấp cơ chế lưu vết những máy ảo sẵn sàng và các thông tin về tài nguyên của chúng.

Dispatcher: cung cấp cơ chế bắt đầu thực thi việc cấp phát máy ảo VM của những yêu cầu dịch vụ đã được chấp nhận.

Service Request Monitor: cung cấp cơ chế lưu vết tiến trình của yêu cầu dịch vụ

VMs (các máy ảo): nhiều máy ảo có thể được mở và tắt động trên một máy vật lý để phù hợp với yêu cầu dịch vụ, do đó việc chuẩn bị tối đa để có thể chia nhỏ tài nguyên để có thể đáp ứng các yêu cầu cụ thể của yêu cầu dịch vụ. Thêm vào đó, nhiều máy ảo VM có thể chạy động thời ứng dụng trên những môi trường hệ điều hành khác nhau trên một máy vật lý duy nhất do các máy ảo VM tách biệt hoàn toàn với các máy khác trên cùng máy vật lý

Physical Machines (các máy vật lý): Những trung tâm dữ liệu bao gồm nhiều máy chủ có thể cung cấp tài nguyên phù hợp với yêu cầu.

IV. Mô phỏng Cloud:

xem xét một ví dụ như CloudSim trong bài viết của R. N. Calheiro và các đồng nghiệp

“CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services”

ế on Control

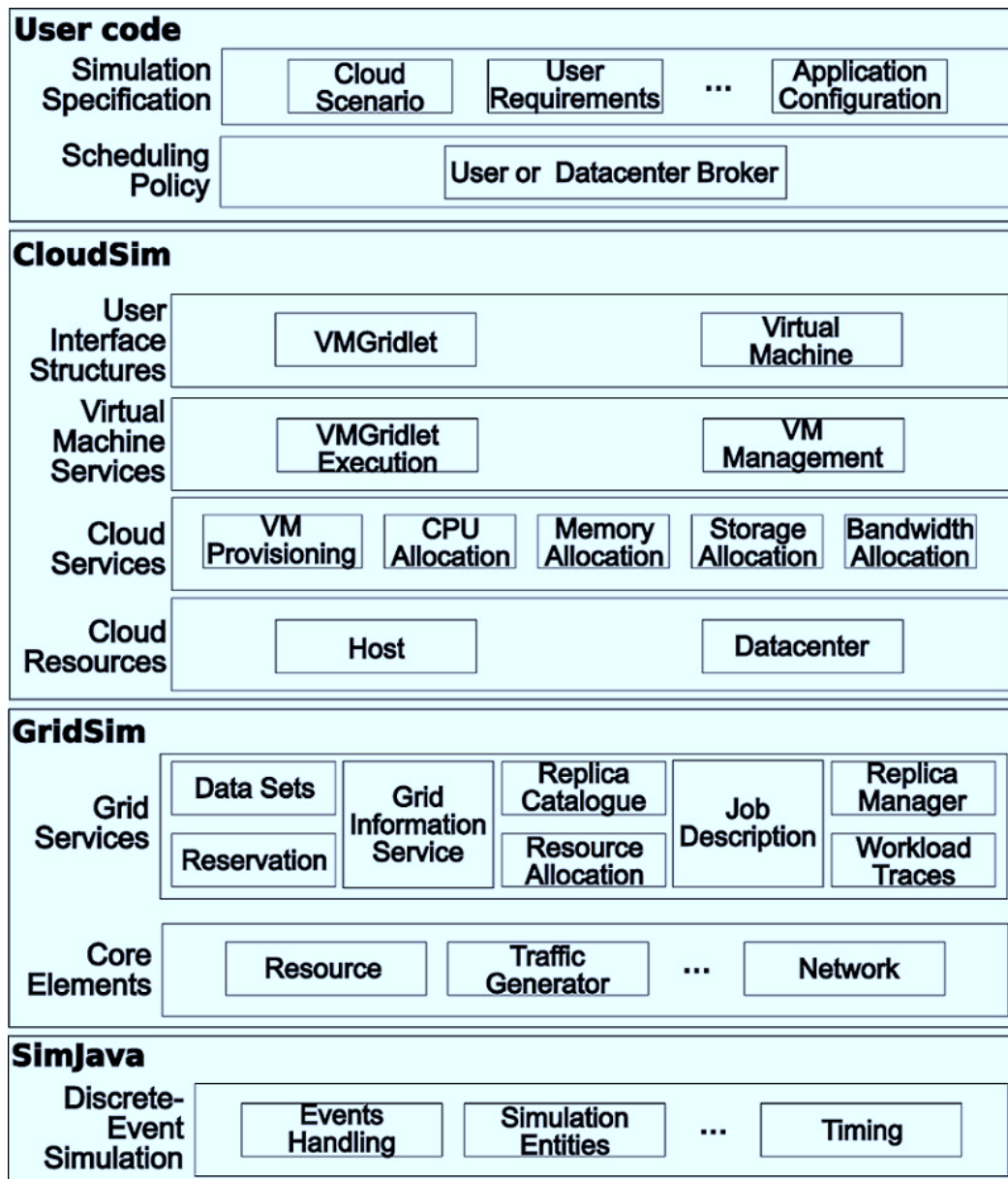
Một cách tổng thể bao gồm 4 lớp:

SimJava: mức thấp nhất trong kiến trúc bao gồm những công cụ mô phỏng sự kiện dùng để hiện thực những chức năng cốt lõi cần thiết cho việc mô phỏng ở lớp cao hơn như sắp xếp và xử lý sự kiện, khởi tạo các thành phần, các thành phần giao tiếp và quản lý mô phỏng đồng hồ .

GridSim: bộ công cụ hỗ trợ các thành phần phần mềm cấp cao hơn để mô hình hóa nhiều nền tảng lưới, bao gồm cả hệ thống mạng và liên kết đồng bộ những thành phần cơ bản của lưới như tài nguyên, tập dữ liệu, lưu vết tải và dịch vụ thông tin.

CloudSim: là phần hiện thực ở mức thấp theo bởi việc mở rộng tự động các tính năng cơ bản được cung cấp bởi lớp GridSim. CloudSim cung cấp hỗ trợ lý thuyết cho việc mô hình và mô phỏng hóa những môi trường tập trung dữ liệu dựa trên nền tảng Cloud. Lớp CloudSim quản lý việc khởi tạo và thực thi các thực thể cốt lõi (máy ảo, máy chủ, trung tâm dữ liệu, ứng dụng) trong suốt quá trình mô phỏng. Lớp này có khả năng khởi tạo đồng thời và quản lý mở rộng trong suốt với những nền tảng Cloud bao gồm hàng nghìn thành phần hệ thống. Những vấn đề cơ bản như chuẩn bị triển khai máy ảo VM dựa trên yêu cầu người dùng, quản lý quá trình thực thi ứng dụng và theo dõi động đều được quản lý bởi lớp này. Người cung cấp Cloud có thể hiện thực những chiến lược của họ để tự động mở rộng lõi của chức năng chuẩn bị triển khai máy ảo VM.

User-code: đây là lớp trên cùng của hệ thống mô phỏng cho phép cấu hình những chức năng liên quan cho các máy chủ (số lượng, đặc tả cho chúng và những việc khác), ứng dụng (số lượng các tác vụ và yêu cầu của chúng), các máy ảo VM, số lượng người dùng và các loại ứng dụng và chính sách định thời môi giới. Một người phát triển ứng dụng Cloud có thể tạo ra nhiều sự phân tán yêu cầu người dùng, cấu hình ứng dụng, và những ngữ cảnh ở lớp này.



Kiến trúc lớp của CloudSim

im im aa

Kiến trúc dịch vụ nền tảng liên quan đến Cloud được mô hình hóa trong chương trình mô phỏng bởi thành phần Datacenter cho việc quản lý yêu cầu dịch vụ. Datacenter được tạo bởi các tập hợp các Host, có trách nhiệm để quản lý các máy ảo VM trong chu kỳ sống của chúng. Các Host là những thành phần biểu thị cho các nút trong Cloud: nó được cài đặt sẵn khả năng xử lý (biểu diễn qua đơn vị MIPS = milion of instruction per second), bộ nhớ, khả năng lưu trữ và chính sách định thời để xử lý việc cấp phân phát lỗi

tính toán cho các máy ảo. Các thành phần máy Host thực giao diện hỗ trợ mô hình hóa và mô phỏng cho cả những nút 1 nhân và đa nhân.

Việc phân phối máy ảo phục vụ cho ứng dụng cụ thể nào đó đến các thành phần Host là trách nhiệm của thành phần Virtual Machine Provisioner. Thành phần này cung cấp một tập các phương thức cho người nghiên cứu, với những hỗ trợ trong việc hiện thực chính sách mới dựa trên những mục tiêu về tối ưu. Những chính sách mặc định hiện có sẵn rất giản đơn theo hướng ai đến trước sẽ được phục vụ trước.

Với mỗi thành phần Host, sự cấp phát các lõi tính toán tới các máy ảo được thực hiện dựa trên việc cấp phát của host. Những chính sách tìm xem trong tài khoản có bao nhiêu lõi tính toán được giao cho mỗi máy ảo và bao nhiêu khả năng còn sẵn. Do vậy, có thể cấp phát lõi CPU cho máy ảo cụ thể hay là phân tán động giữa các máy ảo (chia sẻ theo thời gian) và gán lõi cho máy ảo theo yêu cầu hoặc là theo một chính sách cụ thể nào đó.

Mỗi thành phần Host khởi tạo một bộ định thời máy ảo hiện thực việc cấp phát theo không gian hay thời gian. Những người nghiên cứu hay phát triển Cloud có thể mở rộng bộ định thời máy ảo VM để thử nghiệm với một vài chính sách cấp phát. Kết đến chi tiết liên quan đến chính sách chia sẻ theo thời gian và không gian sẽ được đề cập.

p. p. á p. áp.

Một trong những ý tưởng khiến Cloud computing khác biệt với Grid computing là việc triển khai tối đa công nghệ ảo hóa và các công cụ.

Mô hình ứng dụng cổ điển ánh xạ cá chính xác các nốt tính toán và các ứng dụng riêng rẽ không còn phù hợp trong sự trừu tượng hóa tính toán trong môi trường Cloud. Ví dụ như hai Datacenter chạy trên một Host có 1 nhân. Mặc dù có vẻ như có sự tách biệt nhưng tổng lượng tài nguyên sẵn có bị ràng buộc bởi năng lực của host. Những yếu tố quan trọng này cần được xem xét trong tiến trình xem xét quyết định cấp phát tài nguyên.

Để cho phép giả lập những chính sách khác nhau, CloudSim hỗ trợ việc định thời máy ảo VM ở hai mức: trước tiên tại mức Host và sau đó là mức máy ảo VM. Ở mức đầu tiên, có thể xác định rõ tổng năng lực xử lý của mỗi nhân trong host sẽ được gán cho mỗi máy ảo. Tại mức tiếp theo, các máy ảo VM sẽ được phân rõ tổng năng lực xử lý cụ thể cho mỗi tác vụ được thực thi. Tại mỗi mức, CloudSim hiện thực chính sách cấp phát tài nguyên theo thời gian và không gian.

n. n. n. ợ p ph

Hỗ trợ dịch vụ đóng vai trò như là người môi giới giữa nhà cung cấp dịch vụ Cloud và khách hàng thông qua những dịch vụ so trùng là điểm nhấn của Cloud computing. Hơn

thể, những dịch vụ này cần cơ chế để xác định chi phí dịch vụ và các chính sách về giá. Mô hình chính sách chi phí và giá cả là một ý tưởng cần được xem xét khi thiết kế chương trình mô phỏng Cloud, bốn thuộc tính được xem xét đến cho một Datacenter:

- Chi phí mỗi bộ xử lý
- Chi phí mỗi đơn vị bộ nhớ
- Chi phí mỗi đơn vị lưu trữ
- Chi phí mỗi đơn vị băng thông sử dụng

Chi phí mỗi đơn vị bộ nhớ và lưu trữ được kèm theo trong quá trình khởi tạo máy ảo. Chi phí mỗi đơn vị băng thông sử dụng có trong quá trình truyền dữ liệu. Bên cạnh đó, các chi phí sử dụng bộ nhớ, lưu trữ, băng thông và các chi phí liên quan có mối liên hệ với việc sử dụng tài nguyên tính toán. Do vậy, nếu máy ảo VM được tạo mà không có tác vụ nào thực thi trên chúng, thì chỉ có chi phí về bộ nhớ và lưu trữ. Những vấn đề này có thể được thay đổi bởi người dùng.

Chương 3. So sánh giữa Cloud Computing và Grid Computing

I. Tổng quan

Cloud Computing là một gợi ý cho tương lai, là thời điểm chúng ta không tính toán trên các máy tính cục bộ mà thực hiện tính toán trên các tiện ích tập trung được điều hành bởi thành phần thứ ba (third party). Từ giữa những thập niên 90, thuật ngữ Grid đã được xem xét để mô tả các công nghệ cho phép người tiêu thụ lấy về sức mạnh tính toán theo yêu cầu. Ian Foster và các cộng sự thừa nhận rằng “Bằng việc chuẩn hóa các giao thức được sử dụng để yêu cầu sức mạnh tính toán họ có thể thúc đẩy quá trình tạo ra Grid Computing tương tự như lưới điện”. Kết quả của ý tưởng này là các nhà nghiên cứu đã cho ra đời các hệ thống liên kết khả mở rộng theo nhiều khuynh hướng khác nhau như: TeraGrid, Open Science Grid, caBIG, EGEE, Earth System Grid, chúng không chỉ cung cấp sức mạnh tính toán mà còn cả dữ liệu và phần mềm theo nhu cầu. “Cloud Computing” phải chăng là tên gọi mới của Grid? Câu trả lời là về mặt chi tiết chúng khác nhau, nhưng cả hai cộng đồng đều phải đối mặt với nhiều thách thức giống nhau.

3.1.1 Định nghĩa Cloud Computing

Có nhiều định nghĩa về Cloud, chúng ta xem xét một định nghĩa tiêu biểu trước khi đi vào phân tích các đặc điểm của nó.

Cloud Computing là một mẫu tính toán phân bố quy mô lớn được quan tâm bởi khả năng mở rộng mang tính kinh tế, trong đó một khối các dịch vụ, nền, bộ nhớ và sức mạnh tính toán được quản lý linh động, ảo hóa, trừu tượng và được phân phối theo nhu cầu đến người sử dụng bên ngoài trên toàn Internet.

Một số đặc điểm chính cần xem xét trong định nghĩa này: Cloud Computing là một mẫu tính toán phân bố đặc biệt. Nó khác với hệ phân bố truyền thống ở chỗ: khả năng mở rộng vô cùng lớn, có thể được gói gọn như là một thực thể trừu tượng để phân phối các cấp độ dịch vụ khác nhau đến người sử dụng bên ngoài, nó có khả năng mở rộng mang tính kinh tế, các dịch vụ có thể được cấu hình động và phân phối theo nhu cầu.

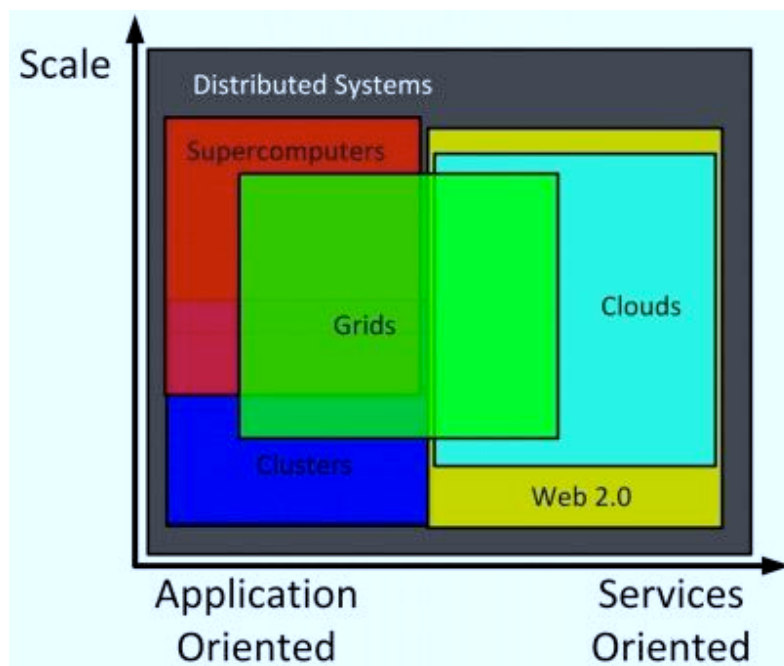
Các tổ chức, các viện nghiên cứu và các ngành công nghiệp hàng đầu đang nhanh chóng tiếp cận Cloud Computing để giải quyết bài toán nhu cầu tính toán và

lưu trữ ngày một tăng cao trong kỷ nguyên Internet. Có 3 yếu tố chính đẩy nên làn sóng Cloud Computing:

- Chi phí phần cứng giảm đi nhanh chóng, khả năng tính toán và lưu trữ ngày một tăng, sự xuất hiện của kiến trúc đa nhân (multi-core) và các siêu máy tính lên đến hàng trăm ngàn nhân.*
- Dữ liệu trong nghiên cứu khoa học tăng lên theo lũy thừa và internet đã trở nên quá thông dụng.*
- Sự chấp nhận rộng rãi trong tính toán dịch vụ và các ứng dụng Web 2.0*

3.1.2 Clouds, Grids và Distributed Systems

Khi xem xét các định nghĩa về Clouds, Grids, Distributed Systems (DS), chúng ta dễ dàng thấy rằng định nghĩa của Clouds có điểm trùng lặp với các định nghĩa của Grids và DS. Điều này không phải là một vấn đề đáng ngạc nhiên bởi vì Clouds không ra đời một cách riêng lẻ hay độc lập mà nó dựa trên nền tảng của các công nghệ trước đó. Hình 1 minh họa mối liên hệ giữa Clouds và các công nghệ khác:



Hình 1 Tổng quan về Grids và Clouds

Clouds và các miền khác trùng lặp lên nhau. Web 2.0 bao phủ hầu hết các không gian của ứng dụng hướng dịch vụ tồn tại trong Clouds. Grids Computing

chồng lấp lên tất cả các lĩnh vực này nhưng khả năng mở rộng được xem là ít hơn supercomputer và Clouds.

Trong ba đặc điểm được đề ra bởi Ian Foster để nhận diện Grids thì chỉ có đặc điểm thứ 3 mới thỏa Clouds, hai đặc điểm còn lại không phù hợp nữa. Phần sau trình bày một so sánh chi tiết từng khía cạnh giữa Clouds và Grids.

II. So sánh giữa Grids và Clouds

Phần này tập trung so sánh giữa Grids và Clouds ngang qua nhiều khía cạnh từ kiến trúc, mô hình bảo mật, mô hình thương mại, mô hình lập trình, ảo hóa, mô hình dữ liệu, mô hình tính toán đến “nguồn gốc (provenance)” và ứng dụng. Cả hai mô hình tính toán về mặt tổng quát thì giống nhau nhưng trong cụ thể thì có một số điểm khác biệt.

3.2.1 Mô hình thương mại (Business Model)

Mô hình thương mại truyền thống trong các phần mềm trước đây là các hình thức tính phí theo một máy tính. Trong Clouds mô hình thanh toán phí linh hoạt hơn nhiều, người khách hàng chỉ cần trả theo nhu cầu sử dụng như các loại phí sinh hoạt hàng ngày mà họ phải trả: điện, nước, gas ..v.v. Ngoài ra Clouds còn hỗ trợ khả năng mở rộng hệ thống mang tính kinh tế, tức là người sử dụng có thể triển khai với hệ thống ngày một lớn hơn với chi phí phù hợp. Các tiềm năng hiện của có là các trung tâm dữ liệu lớn của các tập đoàn Amazon, IBM, Sun, Google ..v.v. Người dùng trong tương lai chỉ cần một thẻ tín dụng đã có thể truy cập theo nhu cầu đến hàng 100 000 bộ xử lý ngang qua hàng chục trung tâm dữ liệu trải khắp thế giới.

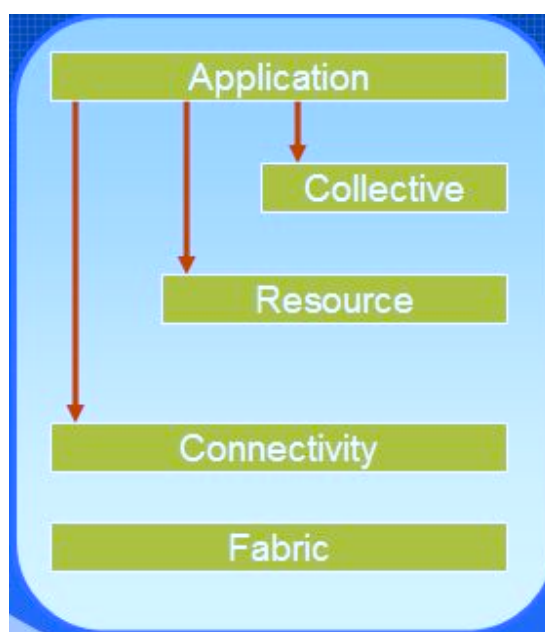
Mô hình thương mại của Grids là hướng đến các dự án nghiên cứu trong môi trường học thuật như các đại học, viện nghiên cứu, phòng thí nghiệm của chính phủ. TeraGrid làm một ví dụ điển hình cho hình thức thương mại của Grid, nó sử dụng hàng chục Grid khác nhau từ các viện nghiên cứu trong cùng một quốc gia. Khi một tổ chức nào đó gia nhập vào TeraGrid thì sẽ có thể truy cập đến các Grids khác và đồng thời có thể sử dụng tài nguyên, kết quả thí nghiệm ... trên hệ thống Grid này. Mô hình này đã được ứng dụng rộng rãi nhiều nơi trên thế giới. Có nhiều nỗ lực để xây dựng một Grid kinh tế cho cơ sở hạ tầng Grid toàn cầu, nó hỗ trợ thương mại, đàm phán, cung cấp dịch vụ theo nhu cầu sử dụng, khả năng rủi ro, chi phí và sở thích của người sử dụng. Nếu như các cố gắng này mang đến kết quả khả quan thì mô hình thương mại của Grids và Clouds thật khó phân biệt trong tương lai.

3.2.2 Kiến trúc (Architecture)

Phần này trình bày mô hình kiến trúc của Grids và Clouds để làm nổi bật sự khác biệt trong hướng tiếp cận của cả hai. Trong khi Grids tập trung trên việc tích hợp các tài nguyên sẵn có gồm cả phần cứng, hệ điều hành, cơ sở hạ tầng an ninh của các hệ thống thì Clouds hướng đến các cấp độ khác nhau của dịch vụ nhằm đáp ứng tối đa nhu cầu người sử dụng như SaaS, IaaS, PaaS.

Để hỗ trợ cho việc tạo ra các tổ chức ảo (Virtual Organizations) – một thực thể luận lý mà bên trong nó các tài nguyên phân bố có thể được khám phá và chia sẻ như thể trong cùng một tổ chức, Grids đã định nghĩa và cung cấp một tập các giao thức chuẩn, phần mềm cơ sở (middleware), bộ công cụ và các dịch vụ được xây dựng trên tập giao thức này. Khả năng hoạt động liên kết và tính an toàn là những vấn đề chính được quan tâm cho cơ sở hạ tầng Grids bởi vì các tài nguyên có thể đến từ các miền quản trị khác nhau, có cả chính sách sử dụng tài nguyên cục bộ và toàn cục khác nhau, các nền và cấu hình phần cứng và phần mềm cũng khác nhau về khả năng sử dụng và tính sẵn sàng của chúng.

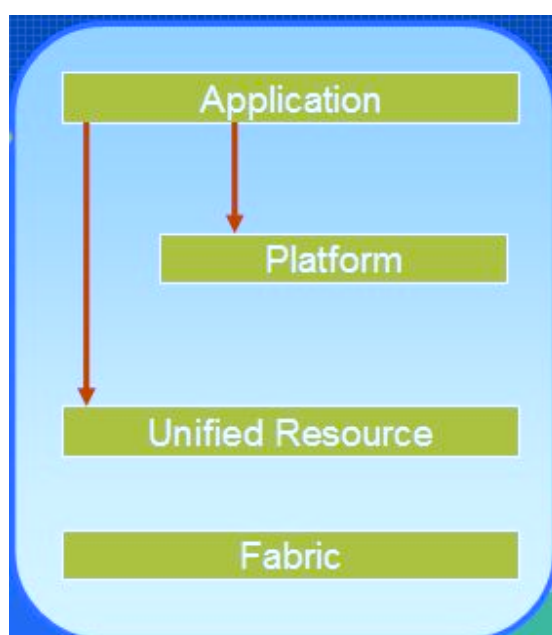
Grids cung cấp các giao thức và dịch vụ ở 5 lớp khác nhau như được minh họa bởi hình sau:



Hình 2 Kiến trúc Grid

Lớp connectivity định nghĩa các giao tiếp và chứng thực cốt lõi cho quá trình giao tác an toàn và dễ dàng qua mạng. Lớp resource định nghĩa các giao thức công bố, khám phá, đàm phán, giám sát, kế toán và thanh toán chi phí cho các hoạt động trên các tài nguyên riêng biệt. Lớp collective nắm giữ sự tương tác ngang qua các tập tài nguyên, dịch vụ thư mục cho phép giám sát và khám phá ra các tài nguyên tổ chức ảo. Cuối cùng là lớp application bao gồm các ứng dụng người dùng được xây dựng trên đầu của các giao thức.

Clouds được phát triển để giải quyết những bài toán tính toán mở rộng qua Internet trong đó một số giả thiết là khác biệt so với Grids. Clouds thường được xem như là một “hồ” tính toán và lưu trữ có thể được truy cập thông qua các giao thức chuẩn và qua một giao tiếp trừu tượng. Thật ra Clouds có thể được hiện thực trên nhưng công nghệ Grids đã tồn tại hàng thập niên nhưng hướng vào kiến trúc 4 lớp sau: Lớp Fabric gồm các tài nguyên phần cứng, lớp Unified Resource chứa các tài nguyên được trừu tượng hóa để có thể xuất hiện đồng nhất với người dùng, lớp Platform thêm vào một tập các công cụ, phần mềm và dịch vụ trên lớp Unified Resource, lớp Application là lớp chứa các ứng dụng chạy trên Clouds. Hình 3 minh họa một kiến trúc của Clouds.



Hình 3 Kiến trúc Clouds

Kiến trúc của Clouds nhằm hướng đến ba mô hình dịch vụ sau: Software as a Service – SaaS, Platform as a Service – PaaS, Infrastructure as a Service – IaaS.

Tuy nhiên người sử dụng có thể yêu cầu đan xen hay tổng hợp các loại hình dịch vụ này.

Infrastructure as a Service: Mô hình này cho phép cung cấp phần cứng, phần mềm và thiết bị với hình thức chi trả dựa trên tài nguyên sử dụng. Cơ sở hạ tầng (infrastructure) có thể mở rộng hay thu nhỏ một cách linh hoạt tùy theo nhu cầu. Các ví dụ tiêu biểu là Amazon EC2 (Elastic Cloud Computing), S3 (Simple Storage Service).

Platform as a Service: đưa ra môi trường tích hợp cấp cao để xây dựng, kiểm tra, và triển khai các ứng dụng tùy ý. Một cách tổng quát các nhà phát triển ứng dụng sẽ phải chấp nhận một số hạn chế trên các kiểu phần mềm mà họ có thể viết đổi lại tính mở rộng gắn liền với ứng dụng. Ví dụ điển hình là Google App Engine.

Software as a Service: hướng tới việc phân phối phần mềm với yêu cầu cụ thể, trong mô hình này người sử dụng có thể truy cập từ xa thông qua Internet và chỉ trả theo mức độ sử dụng. Salesforce là một trong những nhà tiên phong cung cấp mô hình dịch vụ này. Ngoài ra còn có Live Mesh của Microsoft cũng cho phép chia sẻ tập tin, thư mục đồng thời qua nhiều thiết bị.

Từ kiến trúc của Clouds và Grids cho thấy do mục tiêu hướng đến khác nhau nên chúng có cách giải quyết vấn đề và tổ chức hệ thống khác nhau.

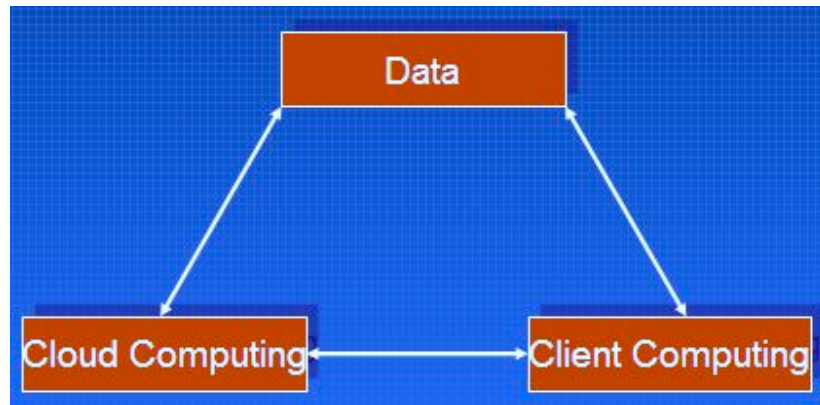
3.2.3 Quản lý tài nguyên (Resource Management)

Quản lý tài nguyên là vấn đề quyết định khả năng hoạt động của hệ thống và đương nhiên nó không thể thiếu trong các hệ thống lớn như Grids và Clouds. Trong phần này, đề cập đến những thách thức chính mà cả Grids và Clouds phải đối mặt để thấy được điểm tương đồng và khác biệt trong hai mô hình.

Mô hình tính toán (Compute Model): Hầu hết các Grids sử dụng mô hình tính toán bó (batch-scheduled compute model) cùng với một bộ quản lý tài nguyên cục bộ như PBS, SGE, Condor để quản lý tài nguyên tại các vị trí khác nhau. Với phương thức quản lý này. Grids không thể phục vụ tốt cho các chương trình yêu cầu nhiều bộ xử lý và thực thi trong thời gian dài. Chẳng hạn trong trường hợp một chương trình cần 100 bộ xử lý và thực thi trong 60 phút, nó phải đợi đến khi hệ thống có đủ 100 bộ xử lý và rảnh trong 60 phút. Mô hình tính toán trong Clouds hoàn toàn khác, các người dùng được sử dụng tài nguyên đồng thời mặc dù

phải đảm bảo được chất lượng dịch vụ cho người sử dụng. Đây cũng là một trong những thách thức của Clouds về khả năng mở rộng và khi số lượng người dùng lớn.

Mô hình dữ liệu (Data Model): Mô hình dữ liệu có khuynh hướng trong tương lai là mô hình tương tác tam giác như chỉ ra trong hình 4.



Hình 4 : Mô hình dữ liệu.

Tính toán Internet sẽ xoay quanh mô hình dữ liệu ở trên do các nguyên nhân khách quan từ thực tiễn. Người khách hàng không muốn đưa các dữ liệu nhạy cảm hoặc tối quan trọng lên xử lý và lưu trữ trên Clouds. Đồng thời, người dùng cũng muốn truy cập đến dữ liệu riêng tư của họ ngay khi các giao tiếp mạng chậm hay hư hỏng. Ngoài ra, công nghệ đa nhân (multicore - technology) cũng mang đến cho người dùng nhiều hệ thống con mạnh mẽ trong tương lai.

Đối với Grids, thì mô hình tính toán tập trung vào dữ liệu đã được quan tâm từ sớm, chẳng hạn như Data Grid đã được thiết kế dành riêng cho các ứng dụng tập trung lớn trên dữ liệu. Do đó, Grids không cần thiết đưa ra định hướng phát triển mô hình dữ liệu như Clouds thay vào đó nó đã có mô hình chuyên biệt để sử dụng.

Tính cục bộ của dữ liệu (Data Locality): Một thách thức chính trong vấn đề mở rộng ứng dụng một cách hiệu quả là vị trí dữ liệu liên quan đối với các tài nguyên đã sẵn sàng. Việc di dời dữ liệu ở xa về các bộ xử lý một cách liên tục là một hạn chế rõ rệt, hơn nữa quá trình nhập xuất (I/O) dữ liệu từ xa có khác biệt lớn với I/O trên dữ liệu cục bộ và ảnh hưởng trực tiếp đến hiệu suất của hệ thống. Đây là một định hướng được Clouds quan tâm giải quyết. Quá trình xử lý dữ liệu là kết hợp dữ liệu liên quan và tiến trình (process) xử lý nó trong cùng một nút (node), do đó vấn đề “data locality” sẽ được giải quyết một cách hiệu quả. Về phía

Grids,

dữ liệu được lưu trữ dựa trên các hệ thống file chia sẻ như NFS, GPFS, PVFS ..v.v. Nơi mà tính “data locality” không dễ dàng áp dụng, đây cũng là một điểm yếu của Grids so với Clouds.

Kết hợp tính toán và quản lý dữ liệu (Combining compute and data management): Tính cục bộ của dữ liệu chỉ hiệu quả khi mà số lượng người dùng và qui mô chưa thực sự lớn. Khi hệ thống ngày càng mở rộng thì vấn đề kết hợp tính toán và quản lý dữ liệu trở thành yếu tố quan trọng cho việc cải thiện hiệu suất. Đó là bài toán định thời công việc (job) gần với dữ liệu mà nó sử dụng trong suốt quá trình thực thi tránh tối đa hoạt động di dời dữ liệu qua mạng. Grids đã có các mô hình tập trung cho các ứng dụng dữ liệu khổng lồ còn Clouds thì sẽ phải đối mặt với vấn đề này trong tương lai.

Ảo hóa (Virtualization): Đây là công nghệ được khai thác mạnh mẽ trong hầu hết các Clouds. Không giống như mô hình tính toán của Grids, Clouds đòi hỏi chạy nhiều ứng dụng người dùng và tất cả các ứng dụng này phải được thực hiện một cách đồng thời đối với người sử dụng. Bên cạnh đó mỗi người dùng có cảm giác như là họ đại sở hữu toàn bộ tài nguyên sẵn có. Công nghệ ảo hóa cung cấp mức trừu tượng thiết yếu cho việc hợp nhất tài nguyên nhằm đạt được mục tiêu này. Đồng thời ảo hóa còn cho phép mỗi ứng dụng có thể được đóng gói

(encapsulate) để có thể cấu hình, triển khai, bắt đầu, di chuyển, tạm dừng, tiếp tục, dừng hẳn, ..v.v. và vì vậy cung cấp tính độc lập, khả năng quản lý, bảo mật tốt hơn. Grids không dựa vào công nghệ ảo hóa nhiều như Clouds, nhưng để đảm bảo được tính riêng tư cho các tổ chức ảo trên tài nguyên thì công nghệ ảo hóa cũng được sử dụng trong lưới chẳng hạn như Nimbus.

Giám sát (Monitoring): Một thách thức khác mà công nghệ ảo hóa mang lại cho Clouds đó là việc giám sát tài nguyên. Vấn đề giám sát tài nguyên trên Clouds chưa được tổ chức trực tiếp như Grids. Grids đề ra các mô hình tin cậy khác nhau để người dùng được ủy quyền có thể truy cập các tài nguyên khác nhau trên các Grid khác nhau và quá trình duyệt tài nguyên này được đảm bảo an toàn. Hơn nữa, Grids không có độ trừu tượng cao và tận dụng ảo hóa như Clouds, ví dụ công cụ giám sát Ganglia có thể theo dõi các cluster và Grids phân tán một cách khá hiệu quả và được chấp nhận rộng rãi trong cộng đồng Grids. Khả năng giám sát tài nguyên của Clouds gặp nhiều thách thức do mục tiêu đề ra của nó về việc cân bằng quá trình giám sát các ứng dụng thương mại, quản lý máy chủ xí nghiệp, giám sát máy ảo, bảo trì phần cứng..v.v. Do đó trong tương lai gần Clouds sẽ

hướng đến việc cung cấp khả năng tự bào trì, cấu hình và quản lý về phía người dùng.

Nguồn gốc dữ liệu (Provenance): Đây là quá trình hướng dẫn về nguồn gốc, lai lịch của các sản phẩm dữ liệu bao gồm tất cả các dữ liệu nguồn, sản phẩm dữ liệu trung gian, và các thủ tục được ứng dụng vào trong sản phẩm dữ liệu. Thông tin này là cần thiết để hiểu, khám phá, hợp lệ và chia sẻ các sản phẩm dữ liệu cũng như các ứng dụng, chương trình dẫn xuất từ nó. Trong Grids “provenance” đã được triển khai thành các hệ thống workflow và được sử dụng hiệu quả. Ngược lại, Clouds còn chưa khai thác lĩnh vực này và có nhiều thách thức trong việc theo vết dữ liệu ngang qua các nhà cung cấp dịch vụ khác nhau, ngang qua các phần mềm khác nhau và các lớp phần cứng trừu tượng thuộc một nhà cung cấp.

3.2.4 Mô hình lập trình (Programming Model)

Mặc dù mô hình lập trình trong môi trường Grids không có các khác biệt về cơ bản so với các môi trường lập trình song song và phân bố truyền thống. Nó có một số vấn đề phức tạp như nhiều miền quản trị, tài nguyên bất đồng bộ, hiệu suất và tính ổn định, điều khiển ngoại lệ trong môi trường có tính linh động cao..v.v. Mục đích chính của Grids là các ứng dụng khoa học đòi hỏi khả năng mở rộng cao, mở rộng đến khối lượng tài nguyên lớn và thực thi vừa nhanh vừa hiệu quả, do đó các chương trình cần phải hoàn tất một cách chính xác nên phải xem xét đến độ tin cậy và khả năng kháng lỗi trong môi trường Grids.

Ngoài các môi trường lập trình truyền thống như MPI, Globus Toolkit, còn có một mô hình lập trình được quan tâm gần đây là MapReduce. MapReduce thật ra chỉ là một mô hình lập trình song song khác nhằm vào các hệ thống với khối lượng dữ liệu lớn. Nó được xây dựng dựa trên hai từ khóa chính là “map” và “reduce” mượn từ ngôn ngữ tự nhiên. Chức năng “map” là cung cấp thao tác cụ thể lên một tập dữ liệu và sinh ra một tập dữ liệu mới. Thao tác “reduce” ngược lại sẽ thu gom các dữ liệu lại. Một hiện thực mã mở của mô hình này là Hadoop và đang được quan tâm trong Clouds.

Clouds về khái quát chấp nhận các Web Services APIs, nơi mà người dùng cấu hình và lập trình dịch vụ Cloud dựa trên các API được trình bày trong Web Services. Các giao thức HTTP và SOAP là những giao thức chuẩn được chọn trong các dịch vụ này. Mặc dù Clouds chấp nhận một số giao thức tổng quát như HTTP và SOAP, khả năng tích hợp và liên kết giữa các dịch vụ lẫn ứng dụng vẫn

là một thách thức lớn khi người sử dụng làm việc với các nhà cung cấp Clouds khác nhau.

3.2.5 Mô hình ứng dụng (Application Model)

Grids một cách tổng quát hỗ trợ nhiều loại ứng dụng khác nhau và hỗ trợ tốt cho các ứng dụng hiệu suất cao (high performance computing - HPC). HPC là các ứng dụng song song chạy trên một hệ thống cục bộ không phải qua internet. Mặt khác Clouds Computing cũng hỗ trợ các loại ứng dụng tương tự Grids nhưng không đạt được tính hiệu quả đối với các ứng dụng HPC. Bởi vì Clouds còn trong quá trình sơ khai, nó còn nhiều vấn đề cần giải quyết.

3.2.6 Mô hình bảo mật (Security Model)

Clouds hầu như bao gồm các trung tâm dữ liệu (data center) chuyên dụng thuộc về cùng một tổ chức, nó mang tính đồng nhất khác biệt hoàn toàn với môi trường Grids. Mô hình bảo mật của Clouds hiện tại đơn giản và ít an toàn hơn các mô hình sẵn có của Grids. Tiêu biểu là Clouds dựa trên Web tạo ra tài khoản người dùng, cho phép họ đặt lại mật khẩu và nhận mật khẩu mới thông qua email trong một môi trường không mã hóa và không an toàn. Chú ý rằng, người dùng có thể dễ dàng sử dụng Clouds ngay lập tức với chỉ một thẻ tín dụng hay địa chỉ email. Ngược lại hoàn toàn, Grids hạn chế hơn về tính bảo mật. Chẳng hạn, mặc dù thông qua Web để tạo tài khoản người dùng, thông tin nhạy cảm về tài khoản, mật khẩu được yêu cầu giao tiếp giữa người với người để xác nhận rõ ràng người dùng được ủy quyền.

Tính bảo mật là một vấn đề quan trọng được quan tâm trong Cloud Computing. Có 7 điểm chính mà một người dùng Clouds cần chú ý khi quyết định sử dụng:

1/ Truy cập bởi người dùng có quyền hạn: Việc dữ liệu nhạy cảm được xử lý bên ngoài xí nghiệp cần đảm bảo rằng nó chỉ truy cập và lan truyền bởi người sử dụng có đặc quyền.

2/ Đồng thuận về điều chỉnh : Người khách hàng cần phải biết được nhà cung cấp Clouds có thực hiện các thao tác ngoài luồng trên dữ liệu của họ hay không.

3/ Vị trí dữ liệu: Do người dùng không biết nơi lưu trữ dữ liệu của họ, nhà cung cấp Clouds phải đảm bảo lưu trữ và xử lý dữ liệu đúng theo thỏa thuận pháp lý ban đầu và phải tuân thủ yêu cầu về tính riêng tư của người khách hàng.

4/ Sự tách biệt về dữ liệu: Dữ liệu giữa các người sử dụng khác nhau phải được lưu trữ riêng biệt.

5/ Khôi phục: Nhà cung cấp Clouds phải có cơ chế nhân bản và khôi phục dữ liệu trong trường hợp có hỏng hóc xảy ra.

6/ Hỗ trợ kiểm tra: Các dịch vụ Clouds rất khó kiểm tra, nếu như điều đó là quan trọng với khách hàng thì nhà cung cấp dịch vụ cũng phải có cơ chế hỗ trợ.

7/ Tính lâu dài: Dữ liệu trên Clouds phải tồn tại lâu dài bền vững ngay khi nhà cung cấp Clouds bị thu mua bởi một tập đoàn khác.

Việc so sánh giữa Grids và Clouds trong thời điểm hiện tại chỉ phản ánh được phần nào các ưu và khuyết điểm hiện có của chúng. Nhưng cả hai cộng đồng này vẫn đang không ngừng phát triển để hoàn thiện.

