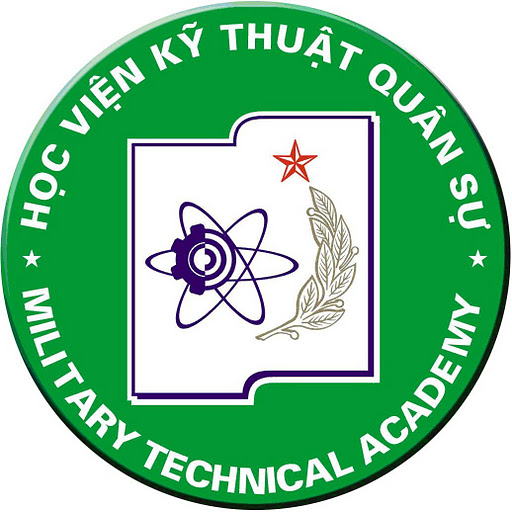
**BỘ GIÁO DỤC VÀ ĐÀO TẠO BỘ QUỐC PHÒNG**

**HỌC VIỆN KỸ THUẬT QUÂN SỰ**

****

**BÁO CÁO**

**Hệ quản trị cơ sở dữ liệu trên Cloud sử dụng Cassandra**

**Học phần:** Cloud Computing

**Giáo viên hướng dẫn:** Nguyễn Văn Giang

**Sinh viên thực hiện:**

Nguyễn Tuấn Anh - 15150152

Phạm Huy Tùng - 15150153

Hoàng Hùng Lâm - 15150138

Phan Văn Hải

**Lớp:** Công Nghệ Mạng 14

MỤC LỤC

[I. Giới thiệu chung 4](#_Toc20298517)

[II. Kiến trúc 5](#_Toc20298518)

[1. Mô hình Cloud: 5](#_Toc20298519)

[2. Mô hình cấp phát máy ảo: 6](#_Toc20298520)

[3. Mô hình chợ Cloud: 6](#_Toc20298521)

[III. Các đối tượng cơ bản trong CloudSim 7](#_Toc20298522)

[IV. Cài đặt, sử dụng CloudSim 9](#_Toc20298523)

[1. Phiên bản sử dụng: 9](#_Toc20298524)

[2. Kịch bản: 9](#_Toc20298525)

[2.1. Các API cơ bản của CloudSim 10](#_Toc20298526)

[2.2. Các tham số: 11](#_Toc20298527)

[2.3. Các bước tạo mô phỏng một đám mây trong CloudSim 13](#_Toc20298528)

[2.4. Kết nối Cloudsim với cấu hình mạngBRITE. 14](#_Toc20298529)

[2.5. Kịch bản 1: 18](#_Toc20298530)

[2.6. Kịch bản 2: 23](#_Toc20298531)

[3. Kết luận: 26](#_Toc20298532)

# Giới thiệu Google Cloud Platform:

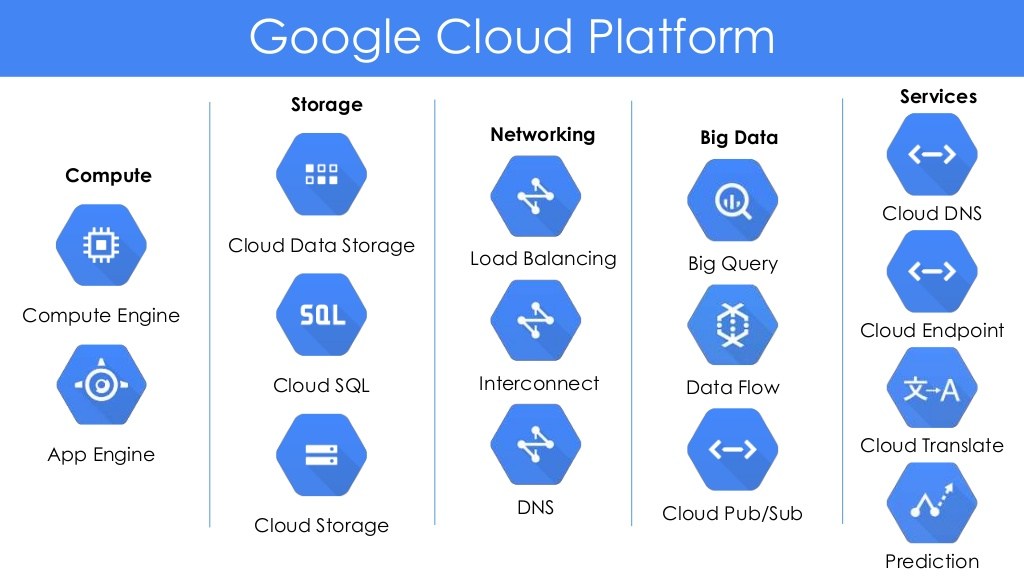
**Google Cloud Platform (GCP)** là một hệ thống các dịch vụ điện toán đám mây của Google. Google Cloud Platform gồm các dịch vụ lưu trữ phục vụ cho việc tính toán, lưu trữ, phát triển ứng dụng. Dịch vụ này được lưu trữ và chạy trên phần cứng của Google. Dịch vụ nền tảng Google Cloud Platform phục vụ cho các nhà phát triển phần mềm, quản trị viên điện toán đám mây, các chuyên gia công nghệ thông tin của doanh nghiệp, các nhà phát triển website thông qua mạng công cộng cũng như những mạng dành riêng.

GCP giúp giải quyết tất cả các vấn đề thiết yếu bao gồm Big Data, Storage, Compute Engine, Networking, Management, Developer Tools, Mobile, …. Do đó, doanh nghiệp của bạn chỉ cần tập trung vào những công việc thật sự quan trọng. Để phát triển doanh nghiệp hơn nữa, mà không cần phải quan tâm đến hệ thống bên dưới.

Ngoài cung cấp các dịch vụ quen thuộc, Google Cloud Platform mang đến sự khác biệt so với tất cả những nền tảng dịch vụ Cloud khác. Đó chính là những dịch vụ này được đặt trực tiếp tại Google. Nơi có một hệ thống Datacenter với mức độ bảo mật và an toàn dữ liệu cao nhất. Đáp ứng các tiêu chuẩn khắt khe nhất đối với một hệ thống điện toán đám mây.

Các sản phẩm chính mà Google Cloud Platform cung cấp bao gồm:

* **Compute** – App Engine, Compute Engine, Container Engine
* **Storage** – Cloud Storage, Cloud Datastore, Cloud SQL, Cloud Bigtable
* **Big Data** – BigQuery, Cloud Dataflow, Cloud Dataproc, Cloud Pub/Sub
* **Services** – Cloud Endpoints, Translate API, Prediction API



Compute Engine là tính năng sử dụng để tạo máy chủ của riêng mình, được gọi là VM (Virtual Machine), hay VM Instance (Đối với người AWS, đó là EC2).

Google Compute Engine ( GCE ) là thành phần Cơ sở hạ tầng dưới dạng Dịch vụ (IaaS) của Google Cloud Platform , được xây dựng trên cơ sở hạ tầng toàn cầu chạy công cụ tìm kiếm của Google, Gmail , YouTube và các dịch vụ khác. Google Compute Engine cho phép người dùng khởi chạy các máy ảo (VM) theo yêu cầu. Máy ảo có thể được khởi chạy từ image tiêu chuẩn hoặc image tùy chỉnh được tạo bởi người dùng. Người dùng GCE phải xác thực dựa trên OAuth 2.0 trước khi khởi chạy VM. Google Compute Engine có thể được truy cập thông qua Bảng điều khiển dành cho nhà phát triển, API RESTful hoặc giao diện dòng lệnh (CLI).

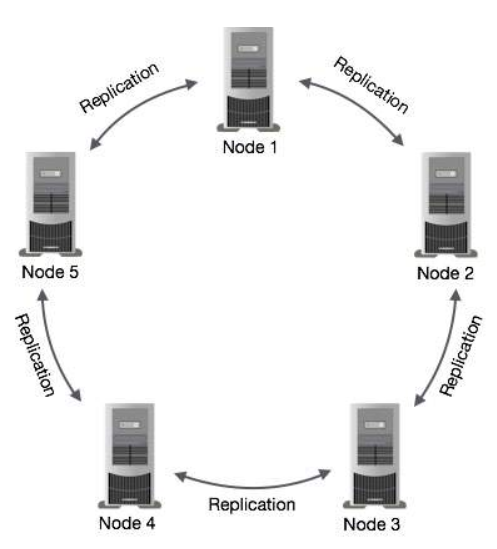
1. Cassandra
2. Cassandra là gì?

Là một quản trị hệ cơ sở dữ liệu phân tán mã nguồn mở được thiết kế để xử lý một khối lượng lớn dữ liệu giàn trải trên nhiều node mà vẫn đảm bảo tính sẵn sàng cao (Highly Availability), khả năng mở rộng hay thu giảm số node linh hoạt (Elastic Scalability) và chấp nhận một số lỗi (Fault Tolerant). Nó được phát triển bởi Facebook và vẫn còn tiếp tục phát triển và sử dụng cho mạng xã hội lớn nhất thới giới này. Năm 2008, Facebook chuyển nó cho cộng đồng mã nguồn mở và được Apache tiếp tục phát triển đến ngày hôm nay. Cassandra được coi là sự kết hợp của Amazon’s Dynamo và Google’s BigTable.

1. Kiến trúc Cassandra

Mục tiêu thiết kế của Cassandra là xử lý khối lượng công việc dữ liệu lớn trên nhiều node mà không có bất kỳ điểm thất bại nào. Cassandra có hệ thống phân tán ngang hàng trên các node của nó và dữ liệu được phân phối giữa tất cả các node trong một cụm.

* Tất cả các node trong một cụm đóng vai trò như nhau. Mỗi nút là độc lập và đồng thời liên kết với các nút khác.
* Mỗi node trong một cụm có thể chấp nhận các yêu cầu đọc và ghi, bất kể dữ liệu thực sự nằm ở đâu trong cụm.
* Khi một node bị hỏng, các yêu cầu đọc/ghi có thể được phục vụ từ các node khác trong mạng.



1. Mô hình dữ liệu Cassandra
2. Cluster

Cơ sở dữ liệu Cassandra được phân phối trên một số máy hoạt động cùng nhau. Container ngoài cùng được gọi là Cluster (cụm). Để xử lý lỗi, mỗi node chứa một bản sao và trong trường hợp một node bị lỗi thì node đó sẽ sao chép bản sao từ node bất kỳ trong Cluster. Cassandra sắp xếp các node trong một Cluster, theo định dạng vòng và gán dữ liệu cho chúng.

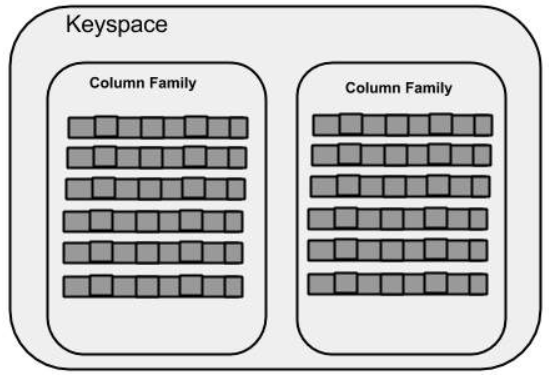
1. Keyspace

Keyspace là nơi chứa dữ liệu ngoài cùng trong Cassandra. Các thuộc tính cơ bản của Keyspace trong Cassandra là:

* **Replication factor**: Đó là số lượng máy trong cụm sẽ nhận được các bản sao của cùng một dữ liệu.
* **Replica placement strategy:** Không có gì ngoài cách đặt bản sao vào vòng. Chúng tôi có các cách như **simple strategy** (chỉ định một yếu tố sao chép đơn giản cho cụm), và **network topology strategy** (sử dụng tùy chọn này, bạn có thể đặt hệ số sao chép cho từng trung tâm dữ liệu một cách độc lập)
* **Column families**- Keyspace là một container chứa cho một danh sách một hoặc nhiều Column Family. Một Column Family, lần lượt là một container chứa một tập hợp các hàng. Mỗi hàng chứa các cột đã sắp xếp. Column Family đại diện cho cấu trúc dữ liệu của bạn. Mỗi không gian khóa có ít nhất một và thường có nhiều Column Family.

Cú pháp tạo Keyspace như sau:



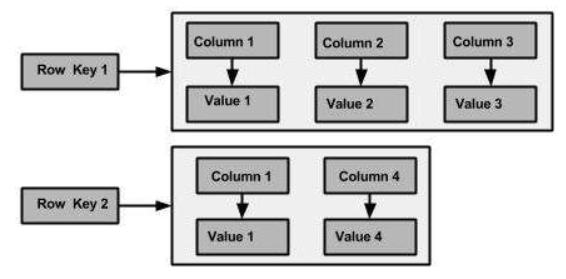


1. Column Family

Một Column Family là một container cho một tập các hàng. Mỗi hàng, lần lượt là một tập các cột.

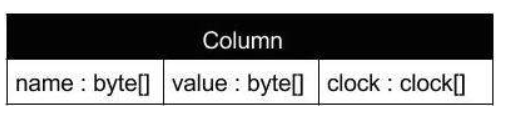
Một Column Family trong Cassandra có các thuộc tính sau:

* **keys\_cached** - Nó đại diện cho số lượng vị trí cần lưu trong bộ nhớ cache cho mỗi SSTable.
* **rows\_cached** - Nó đại diện cho số lượng hàng có toàn bộ nội dung sẽ được lưu trong bộ nhớ.
* **preload\_row\_cache** - Nó chỉ định xem bạn có muốn điền trước bộ đệm hàng hay không.



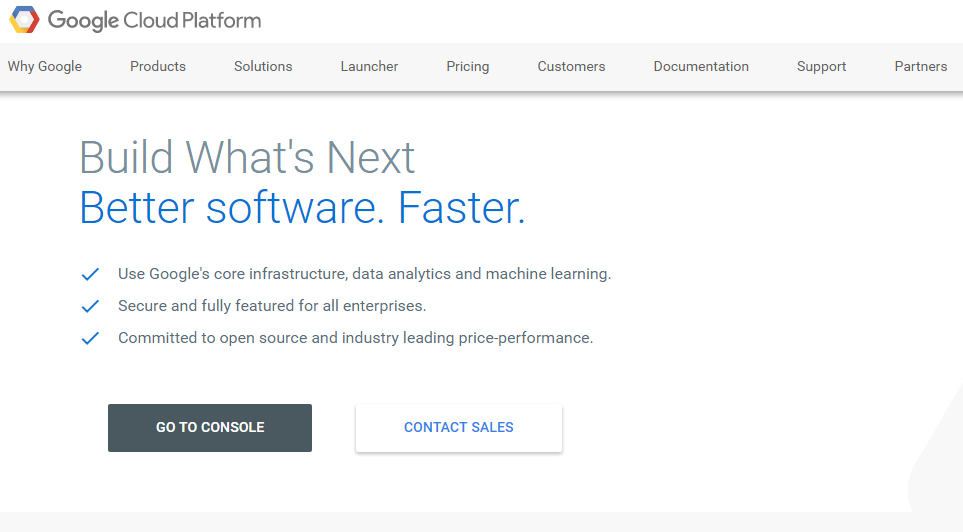
1. Column

Cột là cấu trúc dữ liệu cơ bản của Cassandra với ba giá trị, đó là tên khóa hoặc tên cột, giá trị và dấu thời gian. Dưới đây là cấu trúc của một cột.

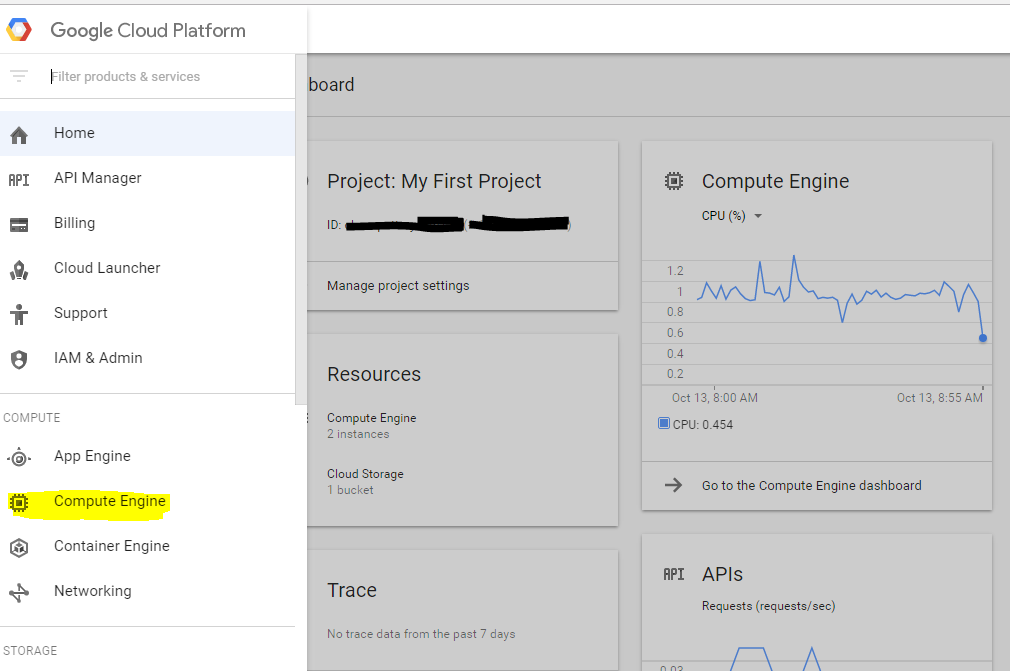


1. Cài đặt và Cấu hình Cassandra trên Google Cloud

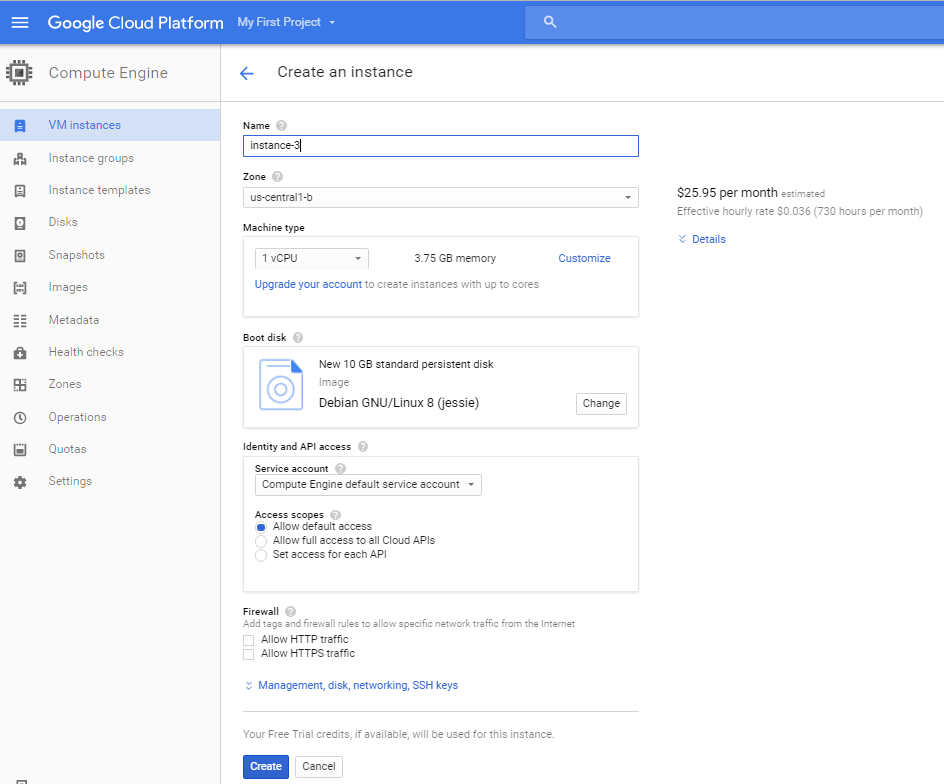
* Step 1: Tạo tài khoản Google cloud dùng thử trong 60 ngày:



* Bây giờ ta vào màn hình Console của Google cloud và xem được tất cả dịch vụ mà Google cloud cung cấp.

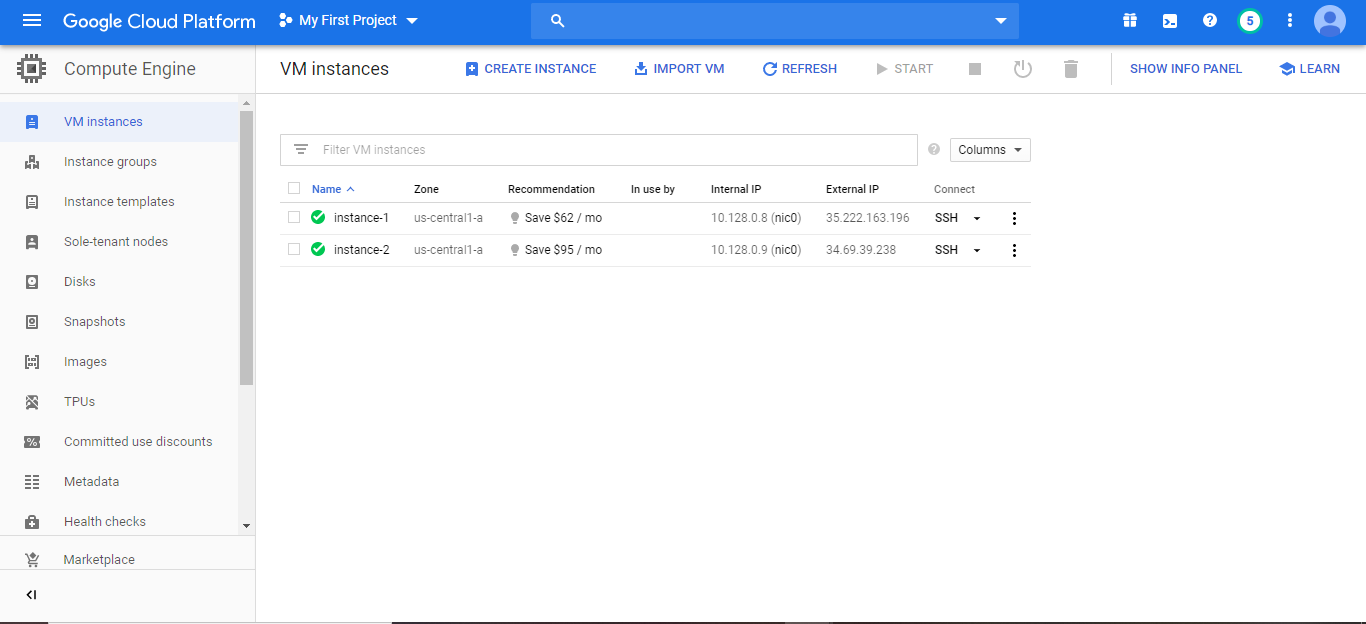


* Step 2: Chọn Compute Engine và tạo server của chính bạn, ở tab VM Instance. Rồi tạo Instance và chọn từng options mà bạn muốn:

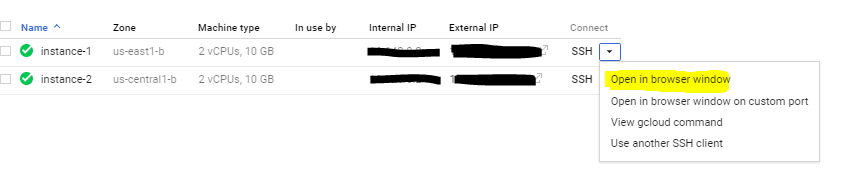


* Step 3: Chọn theo config sau:
* Name: tùy bạn muốn đặt
* Zone: us-central1-a
* Machine type: n1-highmem-4 (4 vCPUs, 26 GB memory)
* Boot Disk (OS you want in your server) — select ubuntu18.04

Đợi cho đến khi Cloud tạo xong sẽ hiện ra giao diện như sau:



* Step 4: Connect tới server mà bạn vừa tạo thông qua SSH:



* Có thể SSH thông qua máy local của mình bằng External IP mà Google cloud cung cấp sẵn thông qua lệnh:

SSH [IP External]

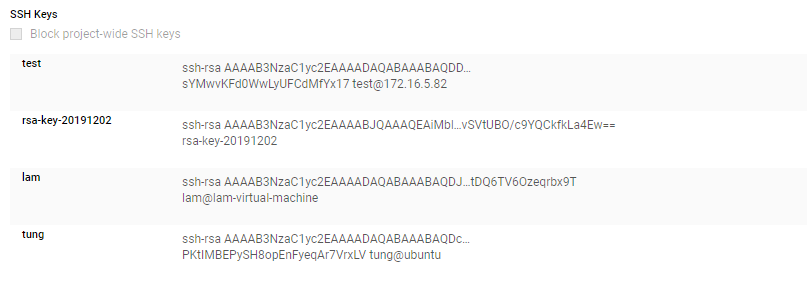
Nhưng trước đó máy local của bạn cần phải dùng lệnh sau để gen ra key SSH:

ssh-keygen -t rsa

Sau đó sử dụng lệnh Cat để xem được key đã gen:

cat ~/.ssh/id\_rsa.pub

Sau đó vào Google cloud và copy key vào phần:



Và Save lại. Sau đó Refresh là đc



* Step 5: Sau khi SSH xong ta sẽ tải Cassandra bằng lệnh:

Wget <http://mirrors.viethosting.com/apache/cassandra/3.0.19/apache-cassandra-3.0.19-bin.tar.gz>

* Step 6: Unzip Cassandra:

tar -xzvf apache-cassandra-3.0.19-bin.tar.gz

* Step 7: Cấu hình trên máy tên Instance-1:

cd apache-cassandra-3.0.19  
cd conf

* Mở file Cassandra.yaml bằng lệnh nano để sửa các thuộc tính sau:

seeds: "10.128.0.8,10.128.0.9"(thay localhost)  
listen\_address: 10.128.0.8  
start\_rpc: true  
rpc\_address: 10.128.0.8

* Lưu lại và Start Cassandra tại máy tên Instance-1:

cd ~/apache-cassandra-3.0.19/bin  
./cassandra

* Step 8: Tại máy tên Instance-2 ta sẽ config như sau:

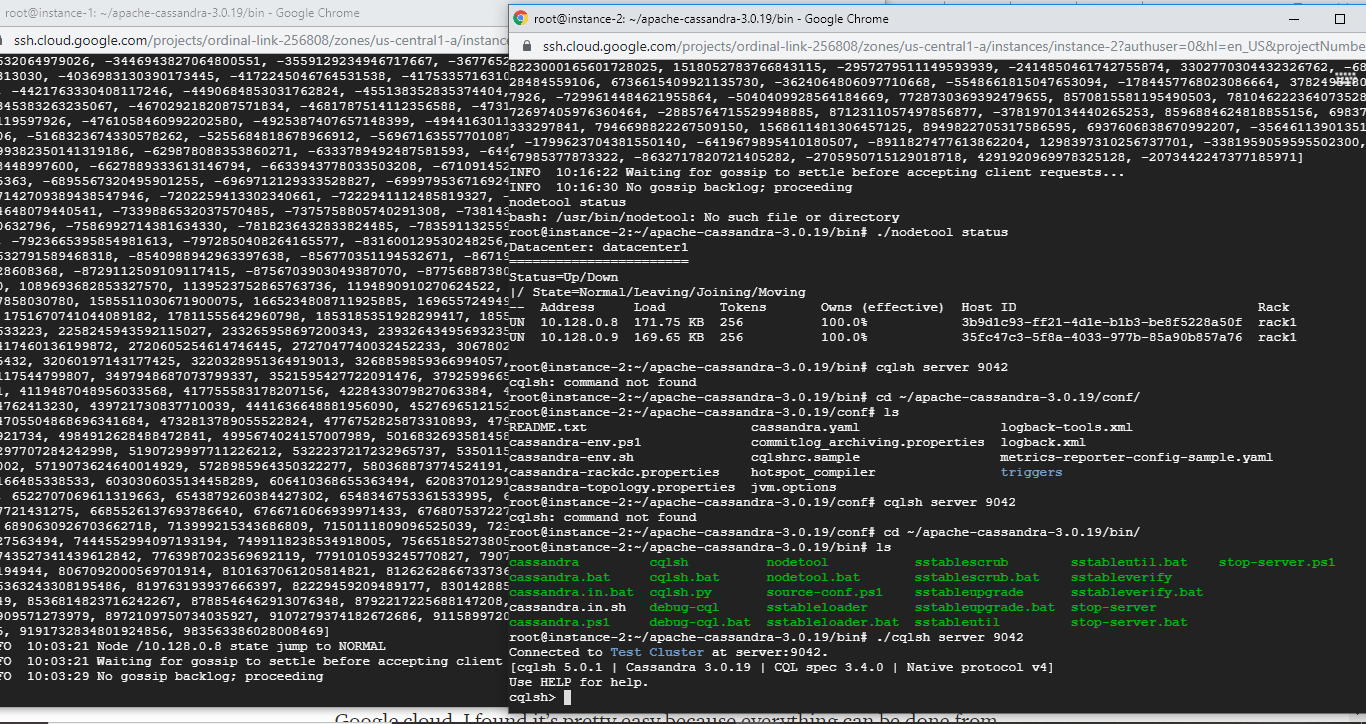
seeds: "10.128.0.8"

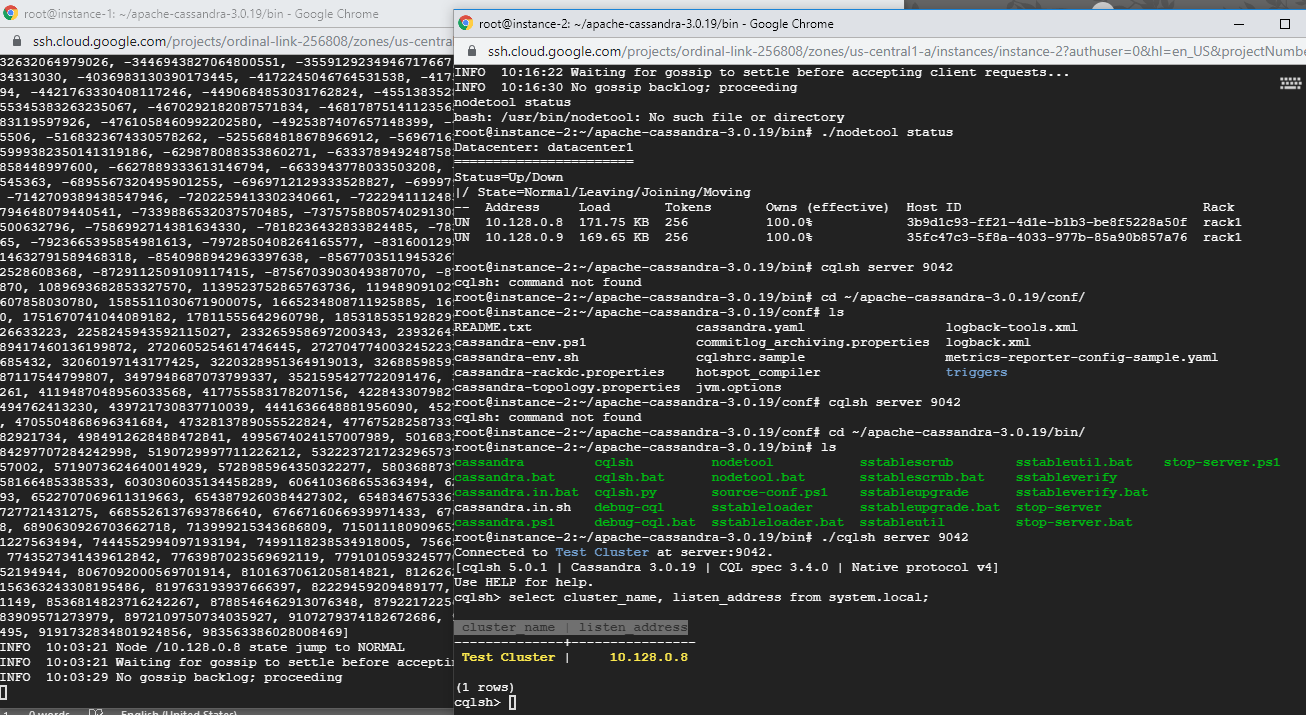
listen\_address: 10.128.0.9

start\_rpc: true

rpc\_address: 10.128.0.9

* Lưu lại và chạy Cassandra như ở máy Instance-1 để 2 máy bắt tay vào 1 Cluster:

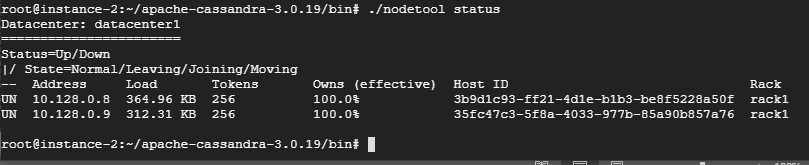




* Để chắc chắn ta sẽ kiểm tra trên máy Instance-1 và Instance-2 bằng lệnh:

cd apache-cassandra-3.0.19/bin/

./nodetool status



Bao giờ check được status như trên là bạn đã thành công.

1. **Truy vấn CQL**
2. **Bảng Sản Phẩm**

**Tạo bảng:**

CREATE TABLE sanpham ( Id\_sp int, Id\_km int, Id\_dm int, Ten\_sp text, Anh\_sp text, Gia\_sp int, So\_luong int, Kich\_thuoc text, Trong\_luong text, Mau\_sac text, Am\_thanh text, Bo\_nho text, He\_dieu\_hanh text, The\_nho text, Camera text, Pin text, Bao\_hanh text, Ket\_noi text, Gia\_km int, Batdau\_km date, Ketthuc\_km date, PRIMARY KEY (Id\_sp));

**Thêm dữ liệu:**

INSERT INTO dienmaypurple.sanpham (Id\_sp,Id\_km, Id\_dm, Ten\_sp, Gia\_sp, So\_luong, Mau\_sac) VALUES (1,1, 1, 'SAMSUNG S5', 100000, 5, 'Red');

1. **Bảng Danh Mục Sản Phẩm**

**Tạo bảng:**

CREATE TABLE tinh\_trang ( Id\_tinh\_trang int, Tinh\_trang text, PRIMARY KEY (Id\_tinh\_trang));

**Thêm dữ liệu:**

INSERT INTO tinh\_trang (id\_tinh\_trang , tinh\_trang ) VALUES ( 1, 'Dang giao');

1. **Bảng Tình Trạng**

**Tạo bảng:**

CREATE TABLE tinh\_trang ( Id\_tinh\_trang int, Tinh\_trang text, PRIMARY KEY (Id\_tinh\_trang));

**Thêm dữ liệu:**

INSERT INTO tinh\_trang (id\_tinh\_trang , tinh\_trang ) VALUES ( 1, 'Dang giao');

1. **Bảng Khách Hàng**

**Tạo bảng:**

CREATE TABLE khachhang (Id\_kd int primary key,Ten\_kh text,Sdt text,Mail text);

**Thêm dữ liệu:**

INSERT INTO khachhang(Id\_kd,Ten\_kh,Sdt,Mail) values (1,'Phan Van Hai','0123456789','phanvanhai1997@gmail.com');

1. **Bảng Người Dùng**

**Tạo bảng:**

CREATE TABLE nguoidung(Id\_nd int primary key,Ten text,Taikhoan text,Matkhau text);

**Thêm dữ liệu:**

INSERT INTO nguoidung(Id\_nd,Ten,Taikhoan,Matkhau) values (1,'Hoang Hung Lam','lam','1');

1. **Bảng Đơn Đặt Hàng**

**Tạo bảng:**

CREATE TABLE don\_dh(Id\_hd int primary key,Id\_kh int,Id\_tinh\_trang int,Id\_nvgh int, Ngay\_lap date, Tong\_gia int, Noi\_nhan text, Ghi\_chu text);

**Thêm dữ liệu:**

INSERT INTO don\_dh (Id\_hd,Id\_kh,Id\_tinh\_trang,Id\_nvgh,Ngay\_lap,Tong\_gia,Noi\_nhan) values (1,1,1,1,todate(now()),5,'Quan Hai Ba Trung, Tp.Ha Noi');

1. **Bảng Chi Tiết Đơn Đặt Hàng**

**Tạo bảng:**

CREATE TABLE ct\_ddh (Id\_ct\_hd int primary key,Id\_hd int,Id\_sp int,So\_luong\_mua int, Don\_gia int);

**Thêm dữ liệu:**

INSERT INTO ct\_ddh (Id\_ct,Id\_hd,Id\_sp,So\_luong\_mua,Don\_gia) values (1,1,1,5,1);

1. **Bảng Bình Luận**

Tạo bảng:

CREATE TABLE binhluan(Id\_bl int,Id\_sp int,Ho\_ten text,Ngay\_gio date,Noi\_dung text,Dien\_thoai varchar,primary KEY (Id\_bl));

Thêm dữ liệu:

INSERT INTO binhluan (id\_bl,dien\_thoai,ho\_ten,id\_sp,ngay\_gio,noi\_dung) VALUES (1,'0912456789','Le Tuan Anh',1,'2019-03-29','San pham rat tot, toi rat hai long!');

1. **Bảng Nhân Viên Giao Hàng**

Tạo bảng:

CREATE TABLE nv\_gh(Id\_nvgh int,Ten\_nvgh varchar,Sdt\_1 int,Sdt\_2 int,Primary Key(Id\_nvgh));

Thêm dữ liệu:

INSERT INTO nv\_gh (id\_nvgh, ten\_nvgh , sdt\_1, sdt\_2 ) VALUES ( 5, 'Pham Huy Tung', '0912345767', '0945906198');

1. **Bảng Sản Phẩm Bán Ra**

Tạo bảng:

CREATE TABLE sp\_ban(Id\_sp\_ban int,Id\_sp int,So\_luong\_ban int,Primary Key(Id\_sp\_ban));

Thêm dữ liệu:

INSERT INTO sp\_ban (id\_sp\_ban, id\_sp , so\_luong\_ban) VALUES ( 1, 1, 15);