

# AN INTRODUCTION TO CLOUD COMPUTING AND AMAZON WEB SERVICES

---

Luigi Libero Lucio Starace

luigi.starace@gmail.com

March 27, 2019

University of Naples, Federico II

# CLOUD COMPUTING

---

Cloud computing is the on-demand delivery of computing resources through a cloud services platform via the internet with pay-as-you-go pricing.

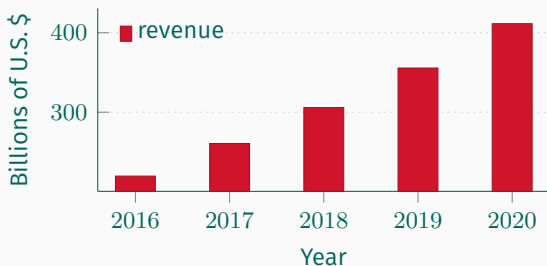
Cloud computing is the **on-demand delivery** of computing resources through a cloud services platform via the internet with pay-as-you-go pricing.

Cloud computing is the **on-demand delivery** of computing resources through a cloud services platform via the internet with **pay-as-you-go** pricing.

## SOME STATS

Worldwide Public Cloud Services Revenue Forecast (Billions of U.S. Dollars) [Gar17]

2016	2017	2018	2019	2020
219,6	260,6	305,8	355,6	411,4



## WHY MIGHT ONE USE CLOUD COMPUTING?

Suppose you have a great business idea: the **DATA** system (Data Analytics for Transport Agencies).

What steps do you take to start making money?

1. Estimate supply and demand;




1. Estimate supply and demand;
2. Estimate infrastructural needs;


1. Estimate supply and demand;
2. Estimate infrastructural needs;
3. Purchase and deploy infrastructure;


1. Estimate supply and demand;
2. Estimate infrastructural needs;
3. Purchase and deploy infrastructure;
4. Install and test your system;

1. Estimate supply and demand;
2. Estimate infrastructural needs;
3. Purchase and deploy infrastructure;
4. Install and test your system;
5. Offer your services to clients.


- Infrastructure is **very expensive**:


- Infrastructure is **very expensive**:
  - Hardware costs; 


- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;


- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;
  - Cooling system; Power bill;



- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;
  - Cooling system; Power bill;
  - A few systems engineers;

- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;
  - Cooling system; Power bill;
  - A few systems engineers;
  - Redundancy for high reliability.

- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;
  - Cooling system; Power bill;
  - A few systems engineers;
  - Redundancy for high reliability.
- It takes time and deploy the infrastructure;

- Infrastructure is **very expensive**:
  - Hardware costs;  Dell
  - Real estate costs;
  - Cooling system; Power bill;
  - A few systems engineers;
  - Redundancy for high reliability.
- It takes time and deploy the infrastructure;
- What if the estimations were wrong?

1. Choose one or more cloud services providers;

1. Choose one or more cloud services providers;
2. Deploy your systems on the cloud;

1. Choose one or more cloud services providers;
2. Deploy your systems on the cloud;
3. Offer your services to clients;

1. Choose one or more cloud services providers;
2. Deploy your systems on the cloud;
3. Offer your services to clients;
4. Pay for what you use.



## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;

### With Cloud Computing

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

× High investment risk;

### With Cloud Computing

✓ Reduced risk;

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;
- × Long time-to-market;

### With Cloud Computing

- ✓ Reduced risk;

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;
- × Long time-to-market;

### With Cloud Computing

- ✓ Reduced risk;
- ✓ Shorter time-to-market;

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;
- × Long time-to-market;
- ✓ Manages own data;

### With Cloud Computing

- ✓ Reduced risk;
- ✓ Shorter time-to-market;

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;
- × Long time-to-market;
- ✓ Manages own data;

### With Cloud Computing

- ✓ Reduced risk;
- ✓ Shorter time-to-market;
- × Trust the vendor?

## WHY MIGHT ONE USE CLOUD COMPUTING?

### Traditional process

- × High investment risk;
- × Long time-to-market;
- ✓ Manages own data;
- ✓ Completely in control;

### With Cloud Computing

- ✓ Reduced risk;
- ✓ Shorter time-to-market;
- × Trust the vendor?

# WHY MIGHT ONE USE CLOUD COMPUTING?

## Traditional process

- × High investment risk;
- × Long time-to-market;
- ✓ Manages own data;
- ✓ Completely in control;

## With Cloud Computing

- ✓ Reduced risk;
- ✓ Shorter time-to-market;
- × Trust the vendor?
- × Dependant from a specific vendor?



- **Software as a Service (SaaS)**

- **Software as a Service (SaaS)**

The service vendor provides the user with a completed product that is run and managed by the service provider.

- **Software as a Service (SaaS)**

The service vendor provides the user with a completed product that is run and managed by the service provider.

- **Platform as a Service (PaaS)**

- **Software as a Service (SaaS)**

The service vendor provides the user with a completed product that is run and managed by the service provider.

- **Platform as a Service (PaaS)**

The service vendor provides the user with a set of API which can be used to build, test and deploy applications.

- **Software as a Service (SaaS)**

The service vendor provides the user with a completed product that is run and managed by the service provider.

- **Platform as a Service (PaaS)**

The service vendor provides the user with a set of API which can be used to build, test and deploy applications.

- **Infrastructure as a Service (IaaS)**

- **Software as a Service (SaaS)**

The service vendor provides the user with a completed product that is run and managed by the service provider.

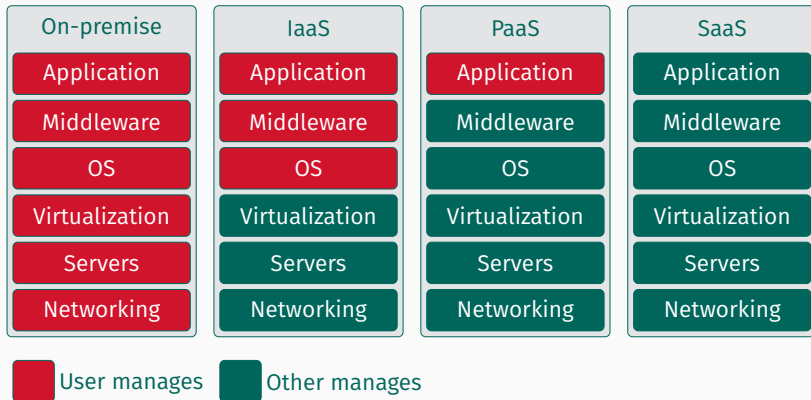
- **Platform as a Service (PaaS)**

The service vendor provides the user with a set of API which can be used to build, test and deploy applications.

- **Infrastructure as a Service (IaaS)**

The service vendor provides users access to computing resources such as servers, storage and networking.

# SERVICE MODELS: A VISUAL COMPARISON



**Figure 1:** A service models comparison

- Google



Google Cloud



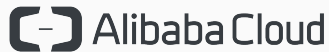
- Google
- IBM



- Google
- IBM
- Microsoft



- Google
- IBM
- Microsoft
- Alibaba



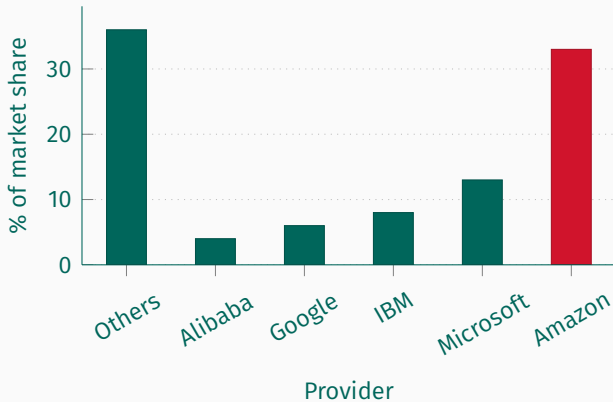
- Google
- IBM
- Microsoft
- Alibaba
- Oracle



- Google
- IBM
- Microsoft
- Alibaba
- Oracle
- Amazon



# MARKET SHARE



**Figure 2:** Market share in Q4 2017 (IaaS, PaaS, Hosted Private Cloud)  
[Syn18]

# AMAZON WEB SERVICES (AWS)





Amazon Web Services is a collection of cloud-based services.





Amazon Web Services is a collection of cloud-based services.  
**A very big one.**



Amazon Web Services is a collection of cloud-based services.  
**A VERY big one.**

# AMAZON WEB SERVICES (AWS)

---

COMPUTING

- (Virtual) Servers on demand



---

Azure: Virtual Machines [web](#)

Google Cloud: Compute Engine [web](#)

# AMAZON ELASTIC COMPUTE CLOUD (EC2)

- (Virtual) Servers on demand
- Different types of instances to suit computing needs



---

Azure: Virtual Machines 

Google Cloud: Compute Engine 

# AMAZON ELASTIC COMPUTE CLOUD (EC2)

- (Virtual) Servers on demand
- Different types of instances to suit computing needs
- Per-second (or per-hour) billing



---

Azure: Virtual Machines [web](#)

Google Cloud: Compute Engine [web](#)

# AMAZON ELASTIC COMPUTE CLOUD (EC2)

- (Virtual) Servers on demand
- Different types of instances to suit computing needs
- Per-second (or per-hour) billing
- Data transfer **not** included!



---

Azure: Virtual Machines [web](#)

Google Cloud: Compute Engine [web](#)

# AMAZON ELASTIC COMPUTE CLOUD (EC2)

- (Virtual) Servers on demand
- Different types of instances to suit computing needs
- Per-second (or per-hour) billing
- Data transfer **not** included!
- Persistent storage **not** included!



---

Azure: Virtual Machines [web](#)

Google Cloud: Compute Engine [web](#)



# AMAZON ELASTIC COMPUTE CLOUD (EC2)

- (Virtual) Servers on demand
- Different types of instances to suit computing needs
- Per-second (or per-hour) billing
- Data transfer **not** included!
- Persistent storage **not** included!
- Scaling **not** included!

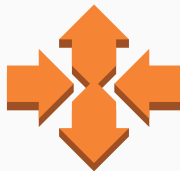


---

Azure: Virtual Machines [web](#)

Google Cloud: Compute Engine [web](#)

- *Scaling is the ability to increase or decrease the compute capacity of your application*



---

Azure: Virtual Machine Scale Sets [web](#)

Google Cloud: Load Balancing [web](#)

- *Scaling is the ability to increase or decrease the compute capacity of your application*
- Scale your application manually, on a scheduled basis or on demand

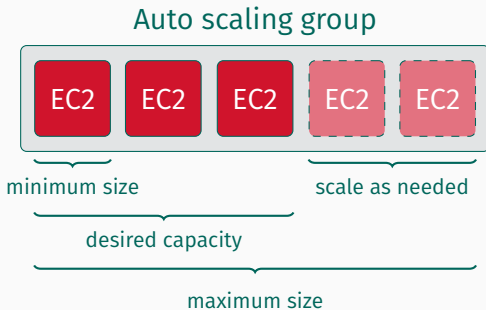


---

Azure: Virtual Machine Scale Sets 

Google Cloud: Load Balancing 

# AMAZON EC2 AUTO SCALING: DETAILS



- Distributes incoming traffic across multiple EC2 instances



# AMAZON ELASTIC LOAD BALANCING (ELB)

- Distributes incoming traffic across multiple EC2 instances
- Pay-per-use billing



- Distributes incoming traffic across multiple EC2 instances
- Pay-per-use billing
  - Execution time

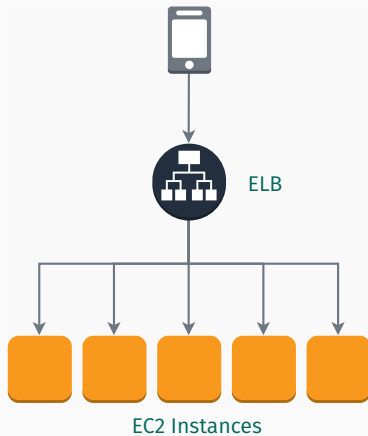


- Distributes incoming traffic across multiple EC2 instances
- Pay-per-use billing
  - Execution time
  - Number of requests / traffic

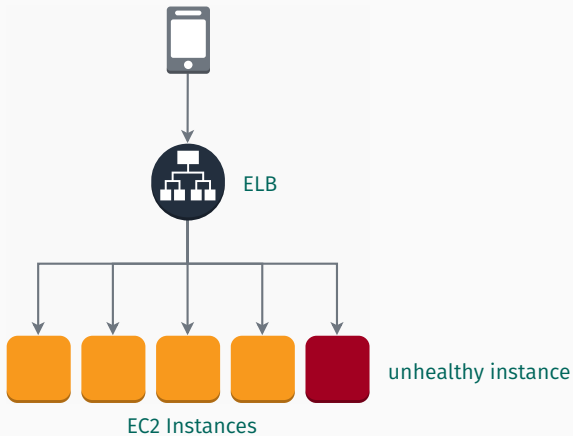




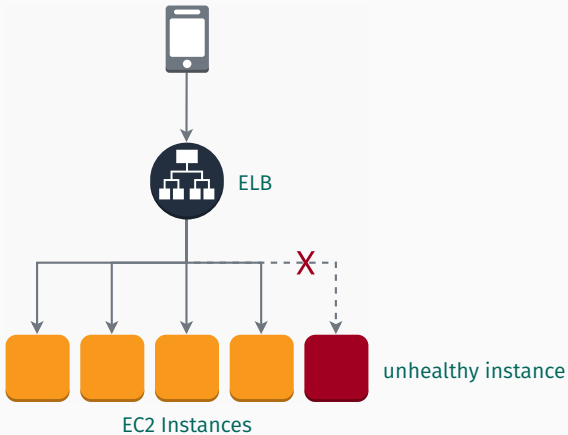
# AMAZON ELASTIC LOAD BALANCING (ELB)



# AMAZON ELASTIC LOAD BALANCING (ELB)



# AMAZON ELASTIC LOAD BALANCING (ELB)



# AMAZON WEB SERVICES (AWS)

---

## CLOUD STORAGE

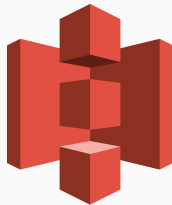
- Elastic Block Storage (EBS)
  - Persistent local storage for EC2 instances.



- Elastic Block Storage (EBS)
  - Persistent local storage for EC2 instances.
- Elastic File System (EFS)
  - File system interface to share data between EC2 instances.



- Elastic Block Storage (EBS)
  - Persistent local storage for EC2 instances.
- Elastic File System (EFS)
  - File system interface to share data between EC2 instances.
- Simple Storage Service (S3)



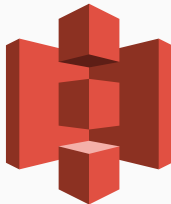
- Elastic Block Storage (EBS)
  - Persistent local storage for EC2 instances.
- Elastic File System (EFS)
  - File system interface to share data between EC2 instances.
- Simple Storage Service (S3)
- Glacier
  - Durable and cheap long-term storage.





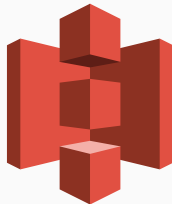
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*



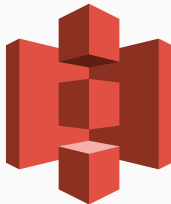
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)



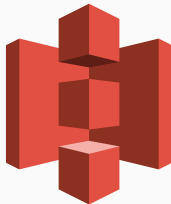
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones



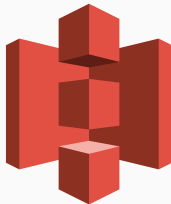
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*



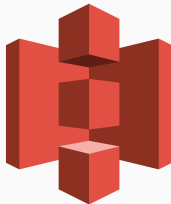
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*
- Multiple storage classes



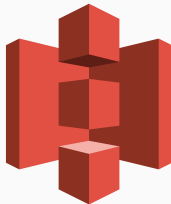
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*
- Multiple storage classes
  - Standard



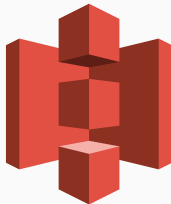
# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*
- Multiple storage classes
  - Standard
  - Infrequent Access



# AMAZON SIMPLE STORAGE SERVICE (S3)

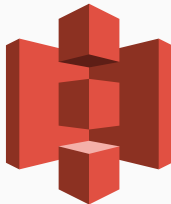
- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*
- Multiple storage classes
  - Standard
  - Infrequent Access
  - One zone-Infrequent Access





# AMAZON SIMPLE STORAGE SERVICE (S3)

- *store and retrieve any amount of data from anywhere*
- 99.999999999% durability (nine nines!)
- Data is distributed across a *minimum* of three availability zones
- A logical unit of storage is a *bucket*
- Multiple storage classes
  - Standard
  - Infrequent Access
  - One zone-Infrequent Access
  - Amazon Glacier

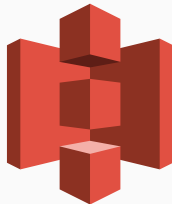


### Multiple storage classes

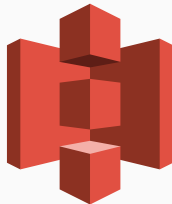
Storage class	Storage	Retrieval (per 1K req.)
Standard	\$0.022 per GB	\$0.0004
Infrequent access	\$0.0125 per GB	\$0.001
IA single zone	\$0.01 per GB	\$0.001

**Table 1:** S3 pricing (Ireland)

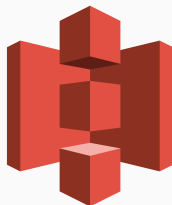
- Well-integrated with other services



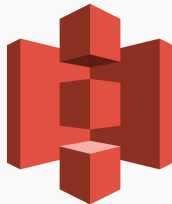
- Well-integrated with other services
  - Machine Learning



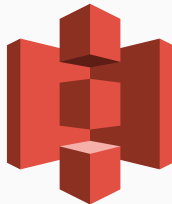
- Well-integrated with other services
  - Machine Learning
  - Big Data Analysis



- Well-integrated with other services
  - Machine Learning
  - Big Data Analysis
- REST API



- Well-integrated with other services
  - Machine Learning
  - Big Data Analysis
- REST API
- Can be used to host static websites



# AMAZON WEB SERVICES (AWS)

---

## DATABASE SERVICES



- Set up, operate a relational database in the cloud.



- Set up, operate a relational database in the cloud.
- Takes care of backups, patching.



- Set up, operate a relational database in the cloud.
- Takes care of backups, patching.
- Supports:



- Set up, operate a relational database in the cloud.
- Takes care of backups, patching.
- Supports:
  - MySQL, PostgreSQL, MariaDB



- Set up, operate a relational database in the cloud.
- Takes care of backups, patching.
- Supports:
  - MySQL, PostgreSQL, MariaDB
  - Oracle, MS SQL Server



- Set up, operate a relational database in the cloud.
- Takes care of backups, patching.
- Supports:
  - MySQL, PostgreSQL, MariaDB
  - Oracle, MS SQL Server
  - Amazon Aurora



- DynamoDB
  - *Fast and flexible NoSQL database service for any scale.*



- **DynamoDB**
  - *Fast and flexible NoSQL database service for any scale.*
- **ElastiCache**
  - In memory data store.
  - Supports memcached, Redis





- **DynamoDB**
  - *Fast and flexible NoSQL database service for any scale.*
- **ElastiCache**
  - In memory data store.
  - Supports memcached, Redis
- **Neptune**
  - Graph database service
  - Supports RDF, SPARQL, ...



# AMAZON WEB SERVICES (AWS)

---

## DATA ANALYSIS

- Run SQL-like queries on S3-stored data in seconds;



- Run SQL-like queries on S3-stored data in seconds;
- Completely managed;



- Run SQL-like queries on S3-stored data in seconds;
- Completely managed;
- You are charged for the number of bytes scanned per query, rounded up to the nearest megabyte, with a 10MB minimum per query. Scanning 1TB costs 5\$.



- Easily Run and Scale Big Data Frameworks such as Apache Spark and Hadoop;



- Easily Run and Scale Big Data Frameworks such as Apache Spark and Hadoop;
- You pay a per-instance rate for every minute used;



- Business Intelligence service that makes it easy to deliver insights to everyone;



---

[Amazon QuickSight](#)

[QuickSight overview \(Youtube\)](#)



- Business Intelligence service that makes it easy to deliver insights to everyone;
- Load data from anywhere;



---

[Amazon QuickSight](#)

[QuickSight overview \(Youtube\)](#)

- Business Intelligence service that makes it easy to deliver insights to everyone;
- Load data from anywhere;
- You pay a per-session rate (from 0,30\$ per session up to 5\$ a month);



---

[Amazon QuickSight](#)

[QuickSight overview \(Youtube\)](#)

# AMAZON WEB SERVICES (AWS)

---

## MACHINE LEARNING

- Amazon SageMaker



- Amazon SageMaker
  - Preconfigured for Tensorflow, MXNet...



- Amazon SageMaker
  - Preconfigured for Tensorflow, MXNet...
  - Build, Train and Deploy phases



- Amazon SageMaker
  - Preconfigured for Tensorflow, MXNet...
  - Build, Train and Deploy phases
  - Pay based on build time, train time and hosting time



- Comprehend (for NLP) [Dashboard](#)





- Comprehend (for NLP) [Dashboard](#)
- Rekognition (Visual Analysis) [Dashboard](#)



- Comprehend (for NLP) [Dashboard](#)
- Rekognition (Visual Analysis) [Dashboard](#)
- Translate

- Comprehend (for NLP) [Dashboard](#)
- Rekognition (Visual Analysis) [Dashboard](#)
- Translate
- Polly (text-to-speech)

- Comprehend (for NLP) [Dashboard](#)
- Rekognition (Visual Analysis) [Dashboard](#)
- Translate
- Polly (text-to-speech)
- Transcribe (speech-to-text)

# AMAZON WEB SERVICES (AWS)

---

MISCELLANEA

## ■ Cognito



- Cognito
  - Sign-up and authentication

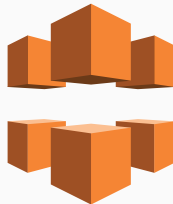


- Cognito
  - Sign-up and authentication
  - Federated identities





- Cognito
  - Sign-up and authentication
  - Federated identities
- CloudFront



- Cognito
  - Sign-up and authentication
  - Federated identities
- CloudFront
  - Content Delivery Network



- Cognito
  - Sign-up and authentication
  - Federated identities
- CloudFront
  - Content Delivery Network
  - 116 Points of Presence in 56 cities across 24 countries



**PRACTICE TIME!**




Your **DATA** (Data Analytics for Transport Agencies) service is ready for deployment on the Cloud (using AWS), and your first client is eager to start using it!

The system you want to deploy consists in the following components:

- internet-connected boxes, equipped with the necessary sensors, which are installed in every vehicle the client wants to monitor and periodically send data to your application;
- a Relational Database (PostgreSQL);
- a web service which receives data from the vehicles, performs ETL operations, archives raw data and inserts new records in the Relational Database;
- a web application which allows your client to perform analytical queries on the relational database;

Your first client wants to manage a fleet of 100 vehicles. Each vehicle is active for about 10 hours a day and, when active, sends 1KB of data every minute to your application.

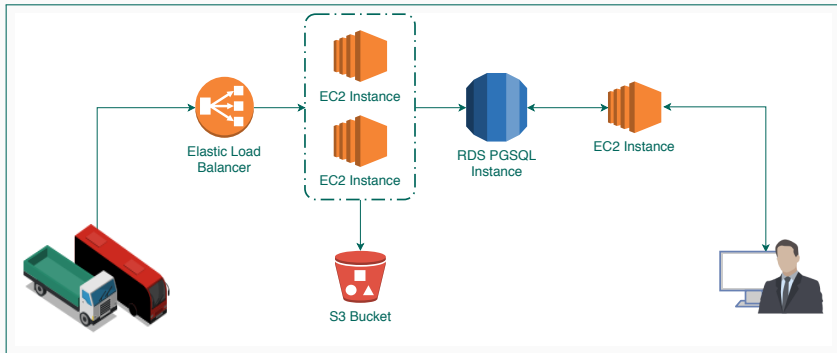
- Task 1** Think of a Cloud infrastructure for **DATA** using AWS. Which services would you use? Is there any weak spot in your architecture?
- Task 2** Use AWS Simple Monthly Calculator  to estimate the monthly bill for your application.



You can download these slides from:

- <http://tiny.cc/cloud-aws-slides>
- <http://bit.do/cloud-aws-slides>

# PROPOSED ARCHITECTURE



## PROPOSED ARCHITECTURE: COSTS

One vehicle sends in a day  $1\text{KB} \times 60 \times 10 = 600\text{KB}$  of data, therefore 100 vehicles send about 60MB of data every day. In one month we collect about 1,8GB of data. In the proposed solution the web service runs on two `t3.micro` (2 vCPU, 1GB memory) EC2 instances, while the web applications runs in a single `t3.small` (2 vCPU, 2GB memory) instance. The database runs on a `db.m5.large` (2 vCPU, 8GB memory) RDS PostgreSQL instance with 240GB SSD storage. A Load Balancer distributes incoming data among the two web service instances, and the raw data is saved in S3 buckets.

The estimated monthly bill is

## PROPOSED ARCHITECTURE: COSTS

One vehicle sends in a day  $1\text{KB} \times 60 \times 10 = 600\text{KB}$  of data, therefore 100 vehicles send about 60MB of data every day. In one month we collect about 1,8GB of data. In the proposed solution the web service runs on two `t3.micro` (2 vCPU, 1GB memory) EC2 instances, while the web applications runs in a single `t3.small` (2 vCPU, 2GB memory) instance. The database runs on a `db.m5.large` (2 vCPU, 8GB memory) RDS PostgreSQL instance with 240GB SSD storage. A Load Balancer distributes incoming data among the two web service instances, and the raw data is saved in S3 buckets.

The estimated monthly bill is .

## PROPOSED ARCHITECTURE: COSTS

One vehicle sends in a day  $1\text{KB} \times 60 \times 10 = 600\text{KB}$  of data, therefore 100 vehicles send about 60MB of data every day. In one month we collect about 1,8GB of data. In the proposed solution the web service runs on two `t3.micro` (2 vCPU, 1GB memory) EC2 instances, while the web applications runs in a single `t3.small` (2 vCPU, 2GB memory) instance. The database runs on a `db.m5.large` (2 vCPU, 8GB memory) RDS PostgreSQL instance with 240GB SSD storage. A Load Balancer distributes incoming data among the two web service instances, and the raw data is saved in S3 buckets.

The estimated monthly bill is . .

## PROPOSED ARCHITECTURE: COSTS

One vehicle sends in a day  $1\text{KB} \times 60 \times 10 = 600\text{KB}$  of data, therefore 100 vehicles send about 60MB of data every day. In one month we collect about 1,8GB of data. In the proposed solution the web service runs on two `t3.micro` (2 vCPU, 1GB memory) EC2 instances, while the web applications runs in a single `t3.small` (2 vCPU, 2GB memory) instance. The database runs on a `db.m5.large` (2 vCPU, 8GB memory) RDS PostgreSQL instance with 240GB SSD storage. A Load Balancer distributes incoming data among the two web service instances, and the raw data is saved in S3 buckets.

The estimated monthly bill is . . .

## PROPOSED ARCHITECTURE: COSTS

One vehicle sends in a day  $1\text{KB} \times 60 \times 10 = 600\text{KB}$  of data, therefore 100 vehicles send about 60MB of data every day. In one month we collect about 1,8GB of data. In the proposed solution the web service runs on two `t3.micro` (2 vCPU, 1GB memory) EC2 instances, while the web applications runs in a single `t3.small` (2 vCPU, 2GB memory) instance. The database runs on a `db.m5.large` (2 vCPU, 8GB memory) RDS PostgreSQL instance with 240GB SSD storage. A Load Balancer distributes incoming data among the two web service instances, and the raw data is saved in S3 buckets.

The estimated monthly bill is 246,32 \$ 

## REFERENCES I

- [Gar17] Gartner. *Gartner Forecasts Worldwide Public Cloud Services Revenue to Reach \$260 Billion in 2017*. Oct. 12, 2017. URL: <https://www.gartner.com/newsroom/id/3815165> (visited on 03/30/2018).
- [Syn18] Synergy Research Group. *Cloud Growth Rate Increases; Amazon, Microsoft & Google all Gain Market Share*. Feb. 2, 2018. URL: <https://www.srgresearch.com/articles/cloud-growth-rate-increases-amazon-microsoft-google-all-gain-market-share> (visited on 03/30/2018).