

CLOUD BASED ADDRESS RECOGNITION SYSTEM

Using AWS, Node.js, VS2015

Xin DU 926753861

Zhixuan WU 219831020

WHAT IS CLOUD?

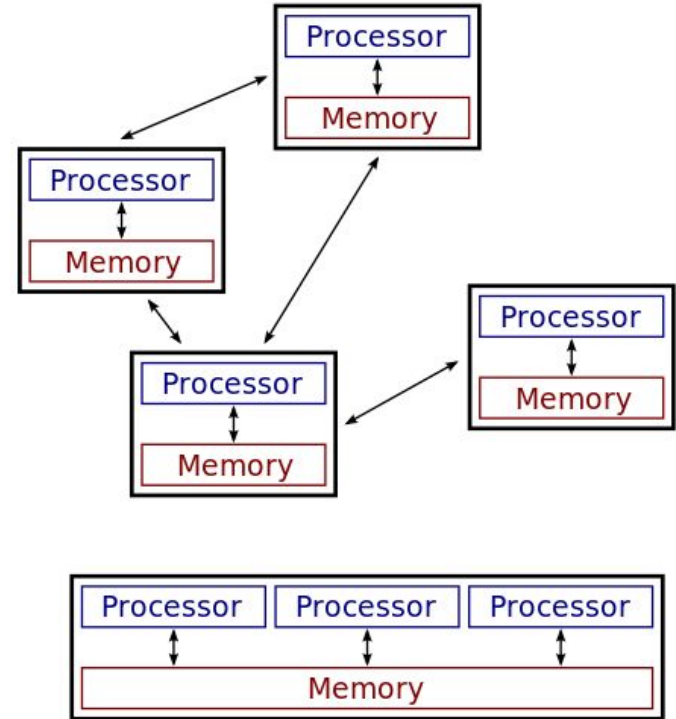
Cloud computing is a type of [Internet](#)-based computing that provides shared computer processing resources and data to computers and other devices on demand.

Distributed Computing is a field of [computer science](#) that studies distributed systems. A *distributed system* is a model in which components located on [networked computers](#) communicate and coordinate their actions by [passing messages](#).

DISTRIBUTED COMPUTING AND PARALLEL COMPUTING

In **distributed computing**, each processor has its own private memory (**distributed memory**). Information is exchanged by passing messages between the processors.

In **parallel computing**, all processors may have access to a shared memory to exchange information between processors.



COMPUTER CLUSTER AND GRID COMPUTERS

Cluster computing consists of a set of loosely or tightly connected **computers** that work together so that, in many respects, they can be viewed as a single system. Unlike **grid computers**, computer clusters have each node set to perform the same task, controlled and scheduled by software.

Grid computing is the collection of computer resources from multiple locations to reach a common goal. Grid computing is distinguished from conventional high performance computing systems such as **cluster** computing in that grid computers have each node set to perform a different task/application.

CLOUD SERVICE MODES

Software-as-a-Service(SaaS)

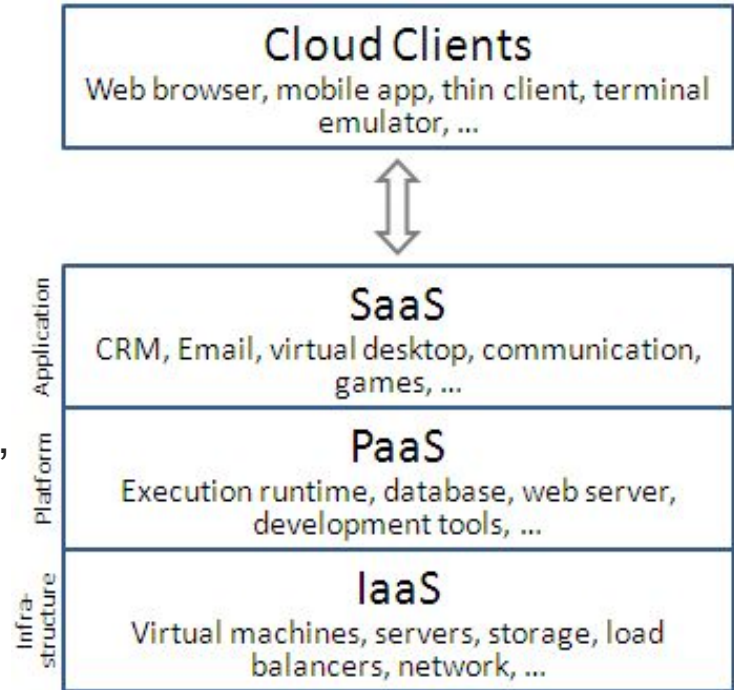
Users gain access to application software and databases.

Platform-as-a-Service (PaaS)

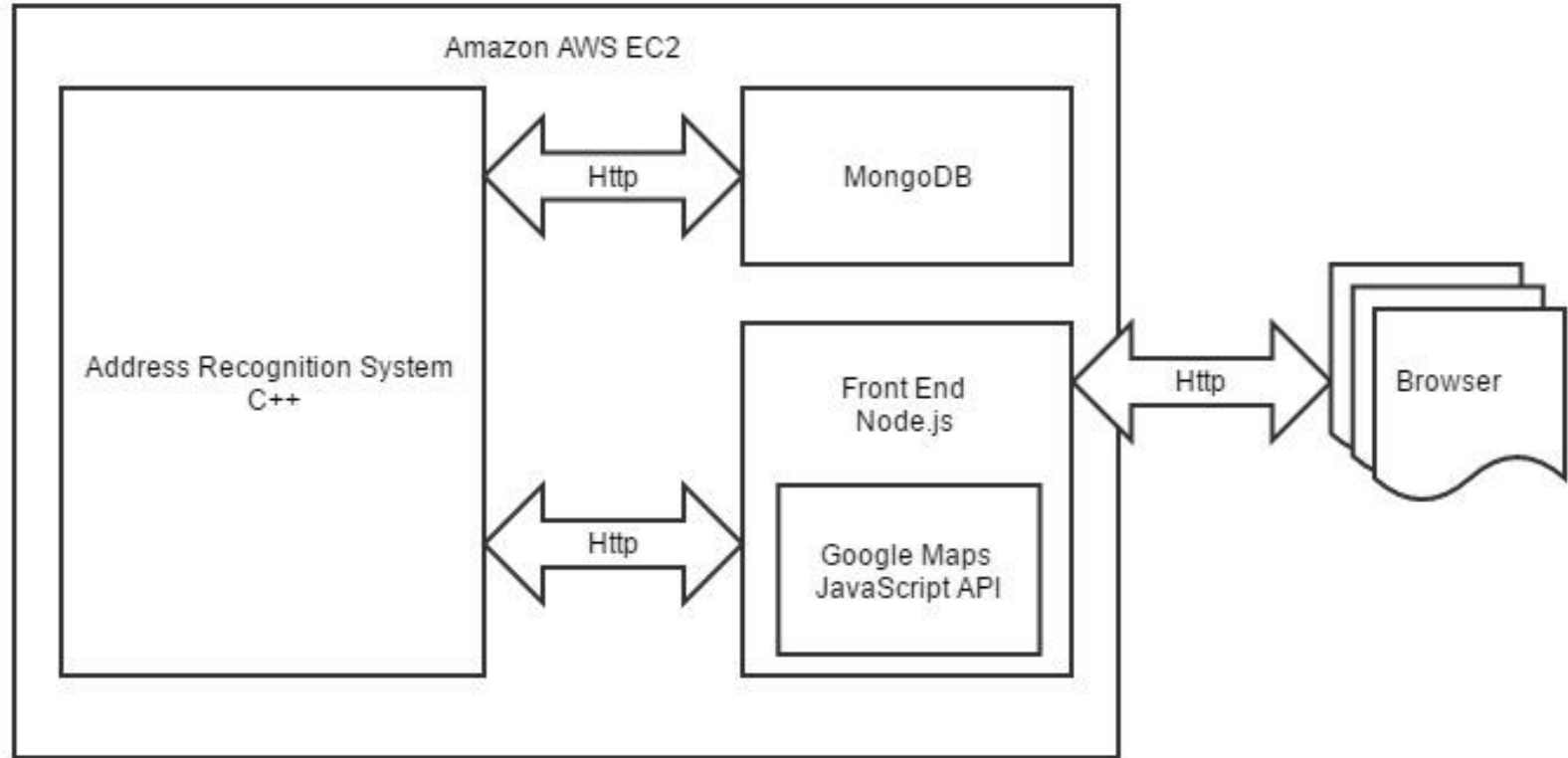
Users gain access to a **computing platform**, typically including operating system, database, and web server.

Infrastructure-as-a-Service (IaaS)

Users gain access to resources from cloud provider's large pools of equipment.



PROJECT STRUCTURE



AMAZON WEB SERVICES (AWS)

Amazon Web Services (AWS), a subsidiary of [Amazon.com](https://www.amazon.com), offers a suite of [cloud-computing](#) services that make up an [on-demand computing platform](#).

[Amazon Elastic Compute Cloud \(EC2\)](#)

[Amazon Virtual Private Cloud \(VPC\)](#)

[Amazon Simple Storage Service \(S3\)](#)

[Amazon Relational Database Service \(RDS\)](#)

LAUNCH AN INSTANCE OF AWS

EC2 Dashboard

Events

Tags

Reports

Limits

INSTANCES

Instances

Spot Requests

Reserved Instances

IMAGES

AMIs

Bundle Tasks

ELASTIC BLOCK STORE

Volumes

Snapshots

Resources

You are using the following Amazon EC2 resources in the US East (N. Virginia) region:

0 Running Instances

0 Volumes

0 Key Pairs

0 Placement Groups

0 Elastic

0 Snapsh

0 Load B

1 Securit

...

 Easily deploy and operate applications - use Chef recipes, manage SSH users, and more. [Try OpsWorks](#)

Create Instance

To start using Amazon EC2 you will want to launch a virtual server, known as an Amazon EC2 instance.

Launch Instance

Note: Your instances will launch in the US East (N. Virginia) region

Step 1: Choose an Amazon Machine Image (AMI)

[Cancel and Exit](#)

An AMI is a template that contains the software configuration (operating system, application server, and applications) required to launch your instance. You can select an AMI provided by AWS, our user community, or the AWS Marketplace; or you can select one of your own AMIs.

Quick Start

1 to 31 of 31 AMIs

My AMIs

AWS Marketplace

Community AMIs

☐ Free tier only ⓘ



Amazon Linux
Free tier eligible

Amazon Linux AMI 2016.09.1 (HVM), SSD Volume Type - ami-0b33d91d

The Amazon Linux AMI is an EBS-backed, AWS-supported image. The default image includes AWS command line tools, Python, Ruby, Perl, and Java. The repositories include Docker, PHP, MySQL, PostgreSQL, and other packages.

Root device type: ebs Virtualization type: hvm

Select

64-bit



Red Hat
Free tier eligible

Red Hat Enterprise Linux 7.3 (HVM), SSD Volume Type - ami-b63769a1

Red Hat Enterprise Linux version 7.3 (HVM), EBS General Purpose (SSD) Volume Type

Root device type: ebs Virtualization type: hvm

Select

64-bit



SUSE Linux
Free tier eligible

SUSE Linux Enterprise Server 12 SP2 (HVM), SSD Volume Type - ami-fde4ebea

SUSE Linux Enterprise Server 12 Service Pack 2 (HVM), EBS General Purpose (SSD) Volume Type. Public Cloud, Advanced Systems Management, Web and Scripting, and Legacy modules enabled.

Root device type: ebs Virtualization type: hvm

Select

64-bit

Step 2: Choose an Instance Type

Amazon EC2 provides a wide selection of instance types optimized to fit different use cases. Instances are virtual servers that can run applications. They have varying combinations of CPU, memory, storage, and networking capacity, and give you the flexibility to choose the appropriate mix of resources for your applications. [Learn more](#) about instance types and how they can meet your computing needs.

Filter by: All instance types Current generation [Show/Hide Columns](#)

Currently selected: t2.micro (Variable ECUs, 1 vCPUs, 2.5 GHz, Intel Xeon Family, 1 GiB memory, EBS only)

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance	IPv6 Support
☐	General purpose	t2.nano	1	0.5	EBS only	-	Low to Moderate	Yes
☒	General purpose	t2.micro Free tier eligible	1	1	EBS only	-	Low to Moderate	Yes
☐	General purpose	t2.small	1	2	EBS only	-	Low to Moderate	Yes
☐	General purpose	t2.medium	2	4	EBS only	-	Low to Moderate	Yes
☐	General purpose	t2.large	2	8	EBS only	-	Low to Moderate	Yes

Step 6: Configure Security Group

A security group is a set of firewall rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group: ☒ Create a **new** security group
☐ Select an **existing** security group

Security group name:

Description:

Type i	Protocol i	Port Range i	Source i	
RDP ▾	TCP	3389	Custom ▾ 0.0.0.0/0	✕

Add Rule



Warning

Rules with source of 0.0.0.0/0 allow all IP addresses to access your instance. We recommend setting security group rules to allow access from known IP addresses only.

1. Choose AMI
2. Choose Instance Type
3. Configure Instance
4. Add Storage
5. Add Tags
6. Configure Security Group
7. Review

Step 7: Review Instance Launch

Please review your instance launch details. You can



Improve your instances' security

Your instances may be accessible from the Internet. You can also open additional ports in your security groups.

AMI Details



Microsoft Windows Server 2016 Datacenter

Free tier
eligible

Microsoft Windows 2016 Datacenter

Root Device Type: ebs Virtualization

If you plan to use this AMI for an application that

Instance Type

Instance Type	ECUs
t2.micro	Variable

Select an existing key pair or create a new key pair



A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

Choose an existing key pair

Create a new key pair

Proceed without a key pair

☐ I acknowledge that I have access to the selected private key file (aaa.pem), and that without this file, I won't be able to log into my instance.

Cancel

Launch Instances

Cancel

Previous

Launch



Feedback



English

© 2008 - 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.

[Privacy Policy](#)

[Terms of Use](#)

MONGODB

MongoDB (from humongous) is a free and open-source cross-platform document-oriented database program.

Classified as a **NoSQL** database program, MongoDB uses **JSON**-like documents with **schemas**.

NODE.JS

Node.js is an [open-source](#), [cross-platform JavaScript runtime environment](#) for developing a diverse variety of [server](#) tools and applications.

Allows you to build scalable network applications using JavaScript on the server-side.

Express.js, AngularJS...

ADDRESS RECOGNITION SYSTEM

What it is?

Recognition system for
automated letter/package sorting

Enables postal and shipping
agencies to process higher
volumes of parcel addresses more
effectively

sto 申通快递 express		3303962602507	
收件信息		周凤玲 地址：安徽省滁州市来安县新城镇南大街 130号欧韩潮流馆	
托运物品名称		数量：1件 tb5295999_00 一：1件 TX BT 圣诞老人亮片毛衣 翻领 均码	
已验视 内件包装		收件签收：	
收件信息		收件人：周凤玲 电话：18075295999 地址：安徽省滁州市来安县新城镇南大街 130号欧韩潮流馆	
其 申通快递 express		ID: tb5295999_00	
目的地寄件人		安徽滁州 来安县	
小C家 13588064734		3303962602507	

ADDRESS RECOGNITION SYSTEM

Why bother?

Fast developing express delivery business

Widely used printed tag

Ancient Sorting System (EMS)

Existing English based recognition products

ADDRESS RECOGNITION SYSTEM

Structure:

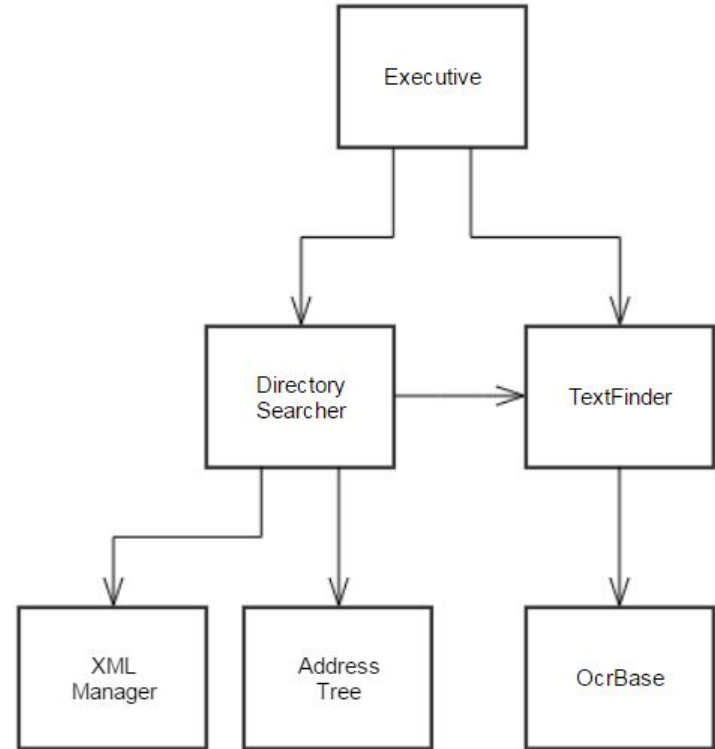
word recognition system(OCR)

address block locator

directory searcher

Input: picture

Output: standard address



ADDRESS RECOGNITION SYSTEM

Challenge:

- Hard to recognize character

 - Chinese characters are logograms (大, 太, 犬)

 - Noise (logo, headline, barcode, jams)

- Hard to locate sentence

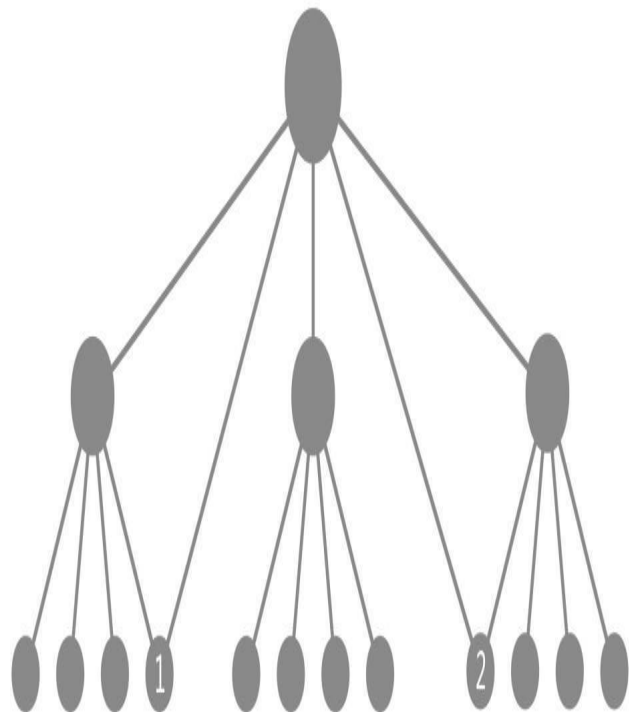
 - Unstructured formats (abbreviation, nickname)

- Need cross-validates address with database

```

<place>
  <place code="120000000000">
    <name full="天津市">
      <short>津</short>
      <simp>天津</simp>
    </name>
    <place code="120100000000">
      <name full="市辖区">
        <short></short>
        <simp>市辖区</simp>
      </name>
      <place code="120115000000">
        <name full="宝坻区">
          <short></short>
          <simp>宝坻</simp>
        </name>
        <place>
          <place code="120116000000">
            <name full="滨海新区">
              <short></short>
              <simp>滨海</simp>
            </name>
            <name full="塘沽区">
              <short></short>
              <simp>塘沽</simp>
            </name>
            <name full="大港区">
              <short></short>
              <simp>大港</simp>
            </name>
            <name full="汉沽区">
              <short></short>
              <simp>汉沽</simp>
            </name>
          </place>
        </place>
      </place>
    </place>
  </place>

```



```

<Line>
  <Sentences>
    <candidate>
      安徽滁州来安县
    </candidate>
    <candidate>
      安徽源删亲妄具
    </candidate>
    <candidate>
      寅傲涂川采寅罢
    </candidate>
  </Sentences>
  <SourceCuts>
    <cut>
      安徽滁州来安县
    </cut>
  </SourceCuts>
</Line>

```

ADDRESS RECOGNITION SYSTEM

Advantages on cloud:

- Easy to access

- Flexible capacity (Fast deploy)

- Database service

- Other Services & support

ADDRESS RECOGNITION SYSTEM

Further tasks:

Socket-based message-passing functionality

File transfer mechanism

To serve concurrent users (threading/BlockingQueue)

Wire with Google Map service

Increase efficiency (NoSQL/AWS API)

Xml generator with pluggable rules

THANK YOU