

可扩展的存储选项

张波
AWS资深技术讲师

日程

- 📦 AWS存储选项
- 📦 Amazon S3 & Glacier
- 📦 Amazon EBS存储

为应用选一款贴心的存储，你需要知道。。。。

📦 数据格式

- 结构化的数据 → 关系型数据库
- 松散定义的模式 → JSON/XML, NoSQL database
- 非结构化的数据 → 文本, 图片, 视频

📦 数据集和数据IO的大小(KBs, MBs, GBs, TBs, PBs)

📦 IOPS, MB/s, ms

📦 数据可用性和持久性的要求



一种存储无法满足所有需求，所以。。

典型负载

- Instance 存储
- Amazon Elastic Block Store
- Amazon Simple Storage Service (S3)
- Amazon Glacier

特定应用存储

- Amazon ElastiCache
- Amazon DynamoDB
- Amazon RDS
- Amazon CloudFront

数据迁移工具

- Amazon Direct Connect
- Amazon Export/Import
- Amazon Storage Gateway
- Tsunami, BitTorrent, 3rd party solutions



问题就来了，我应该选择哪种存储？

	ElasticCache	DynamoDB	Amazon RDS	EBS	Amazon S3	Amazon Glacier
平均延迟	ms	ms	ms,sec	ms	ms,sec,min (~ size)	hrs
数据量	GB	GB-TBs (no limit)	GB-TB (3 TB Max)	GB-TB	GB-PB (no limit)	GB-PB (no limit)
数据大小	KB (64 KB max)	KB (64 KB max)	KB (~rowsize)	KB	KB-GB (5 TB max)	GB (40 TB max)
并发访问量	Very High	Very High	High	High	Low-Very High (no limit)	Very Low (no limit)
成本/GB per month	¢¢	¢¢	¢¢	¢	¢	¢
持久性	Very High	Very High	High	High	Very High	Very High

热数据



冷数据

A person's silhouette is shown holding a smartphone horizontally. The phone's screen displays a scene of a concert with a crowd and bright stage lights. The background of the entire image is a blurred concert scene with green and blue lighting and bright spotlights. The text is overlaid on a semi-transparent blue horizontal bar.

每时每刻，我们不断地产生新的数据

用户？应用？机器？

大数据的时代需要海量的存储



- 📦 海量存储需求？
- 📦 高并发下的高性能？
- 📦 大规模下的可靠性？

Amazon Simple Storage Service (S3)



面向开发者的简单存储

互联网存储

REST API

HTTP PUT
HTTP GET
HTTP POST
HTTP DELETE

海量的扩展性



X万亿的数据对象

- 无限的存储容量
- 高并发的吞吐量

峰值每秒钟300万请求

大规模下的高可靠

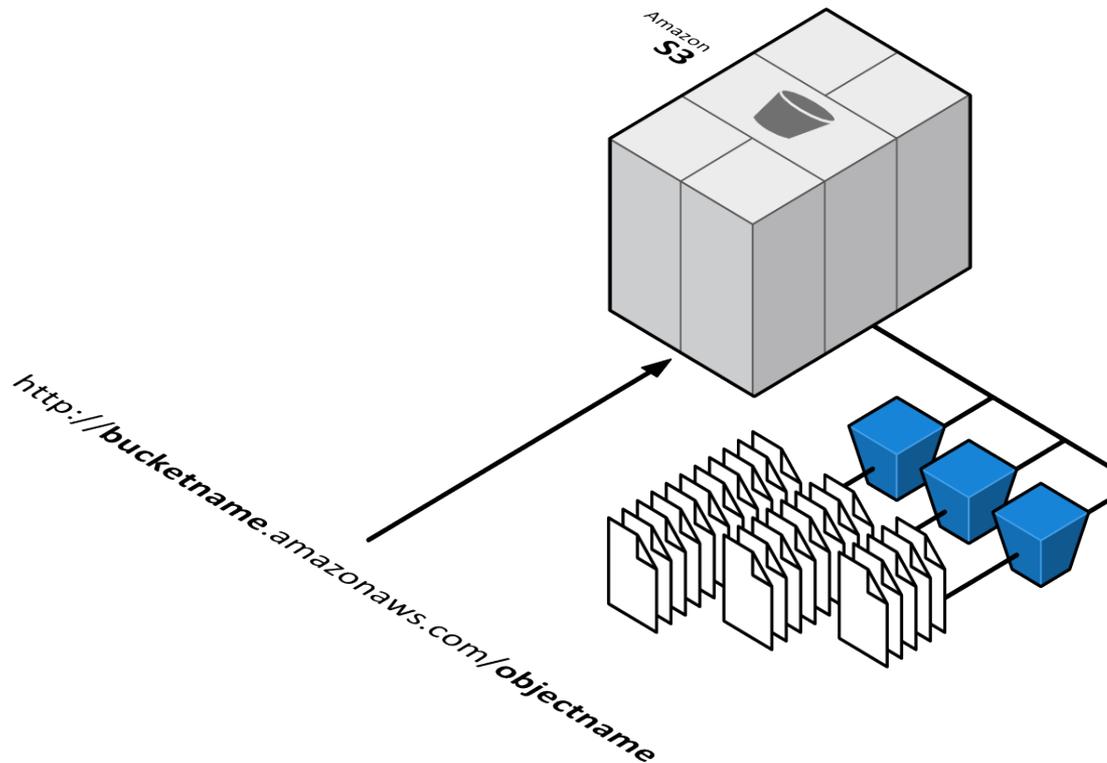


一个区域中跨可用区的多个副本

11个9的可持久性

版本功能和服务器端加密

Buckets + 无限数量的对象





```
public static void bucketCreate(Regions regions, String bucketName) {  
    AmazonS3 s3 = initClient();  
    Region region = Region.getRegion(regions);  
    s3.setRegion(region);  
    System.out.print("Creating bucket: " + bucketName);  
    s3.createBucket(bucketName);  
    System.out.println(" ... Done.");  
    bucketList(s3);  
}
```

Create Bucket

Actions ▾

All Buckets

	Name
	aws-training-share
	awsu-china
	cf-templates-15st5ou9fn52g-ap-northeast-1
	cf-templates-15st5ou9fn52g-ap-southeast-2

还有哪些应用场景？



📦 写一次，读多次 (WORM)

📦 场景：

- 内容存储和分发
- 数据分析的存储
- 备份、归档和灾难恢复
- 静态网站托管

Glacier – 面向冷数据的归档存储



- ❏ 低成本
- ❏ 11个9的高持久性
- ❏ 读取时间3-5小时
- ❏ 通过Glacier的API
 - 或者S3的API

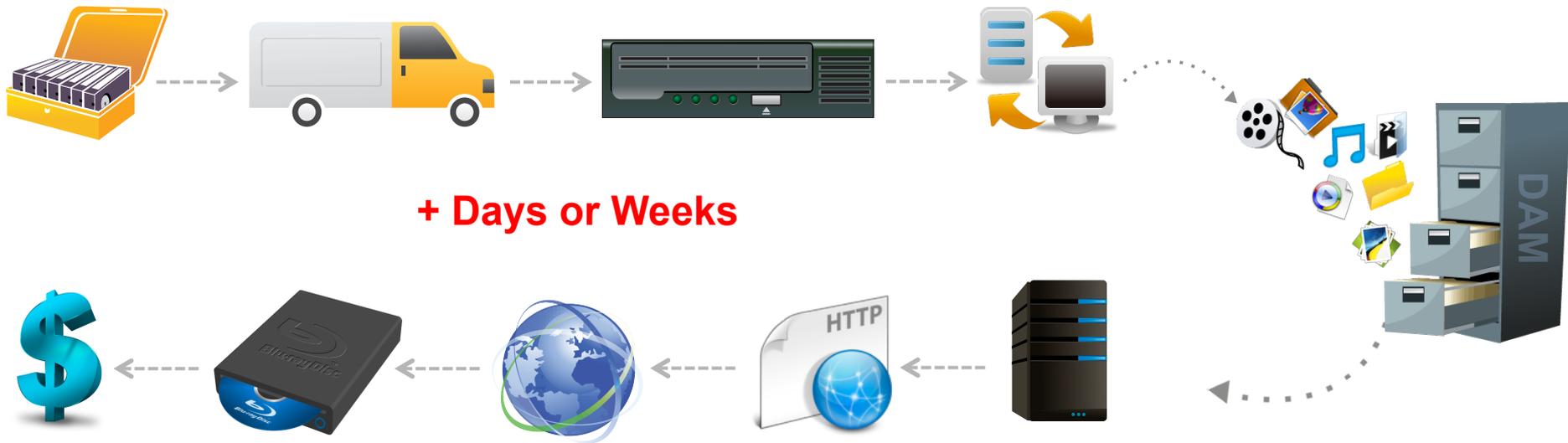


基于策略的生命周期管理

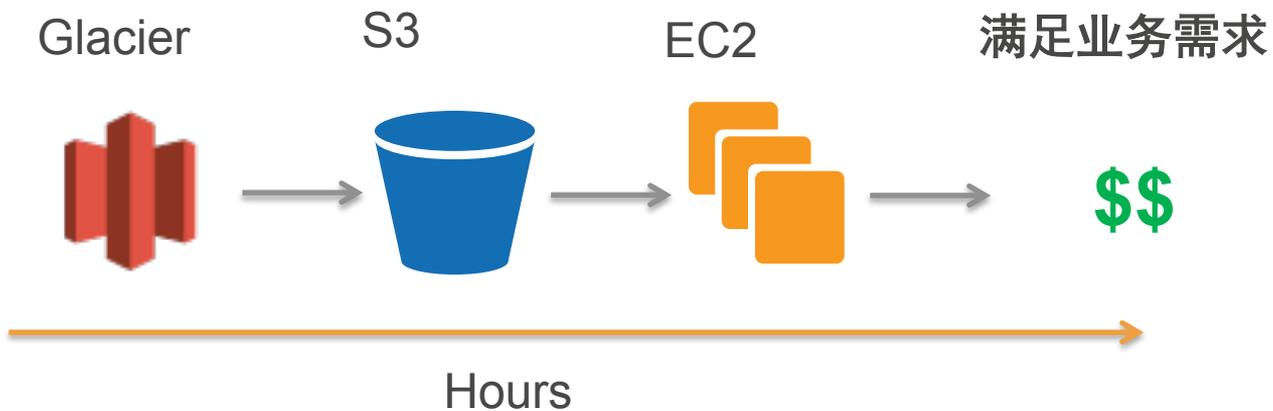
- Key prefix "logs/"
- Transition objects to Glacier 30 days after creation
- Delete 365 days after creation date.

```
<LifecycleConfiguration>
  <Rule>
    <ID>archive-in-30-days</ID>
    <Prefix>logs/</Prefix>
    <Status>Enabled</Status>
    <Transition>
      <Days>30</Days>
      <StorageClass>GLACIER</StorageClass>
    </Transition>
    <Expiration>
      <Days>365</Days>
    </Expiration>
  </Rule>
</LifecycleConfiguration>
```

旧世界：磁带的归档恢复



新世界：Amazon Glacier上的归档数据恢复

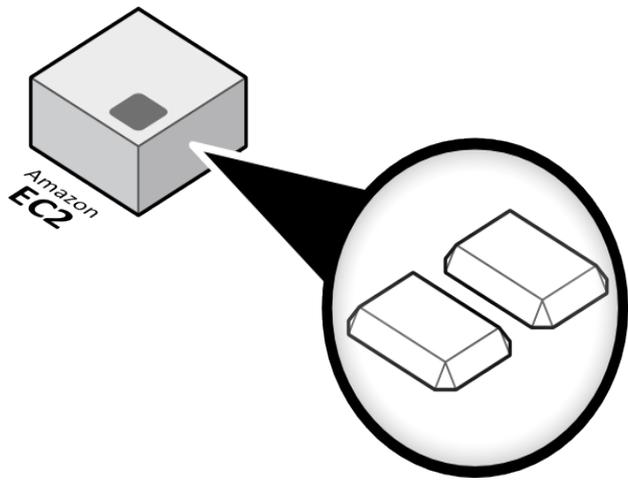


数据库应用需要什么样的存储？

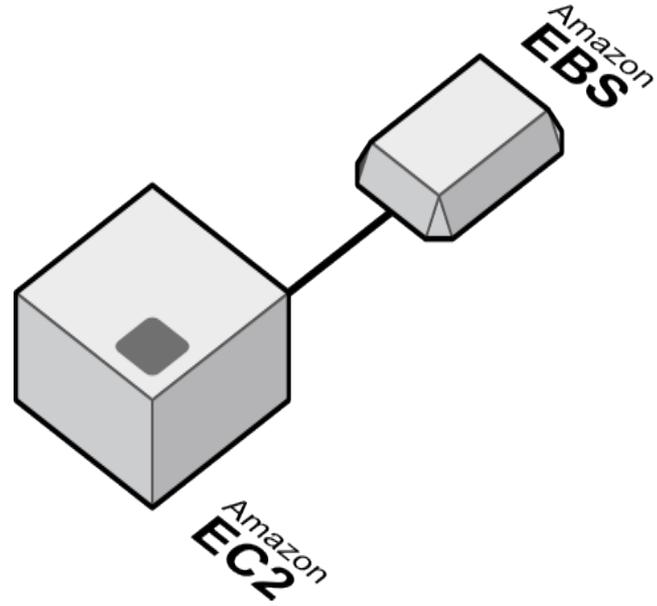
数据库应用对存储的需求

- ❏ 数据块访问
 - ❏ 低延迟
 - ❏ 高并发
 - ❏ 高可用
 - ❏ 高可靠

实例存储 vs EBS 卷



持久性?
性能?
价格?



EBS 性能&容量

10 GB -> 1TB

100 -> 4000 IOPS

Create Volume ✕

Type	<input type="text" value="Provisioned IOPS (SSD)"/>
Size (GiB)	<input type="text" value="500"/> (Min: 10GiB, Max: 1024GiB)
IOPS	<input type="text" value="3000"/> Min: 100 IOPS, Max: 4000 IOPS

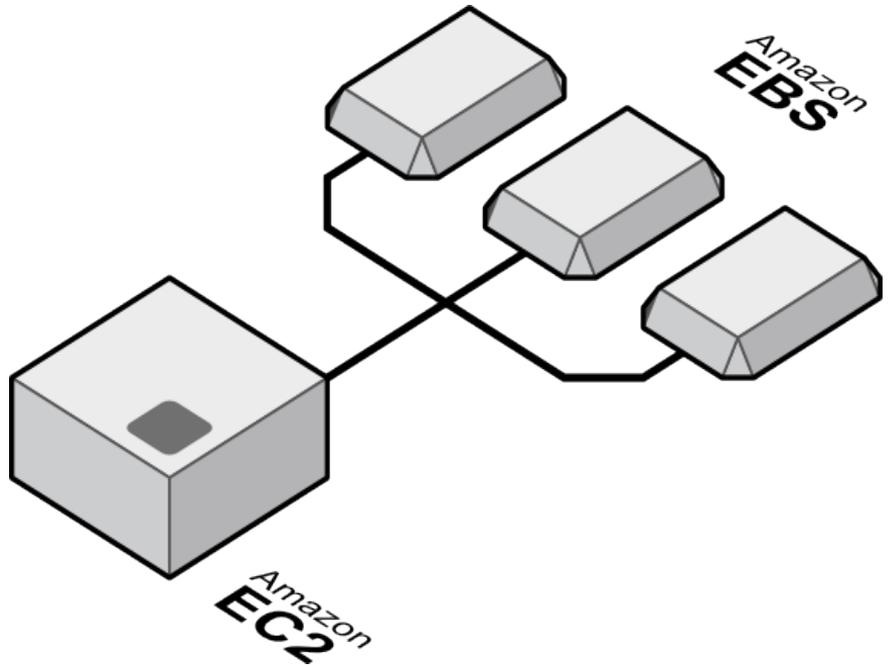
Availability Zone

Snapshot ID	<input type="text" value="Search (case-insensitive)"/>
Encryption	<input type="checkbox"/> Encrypt this volume

EBS Optimized

[Cancel](#) [Create](#)

更高的性能？ 更大的容量？

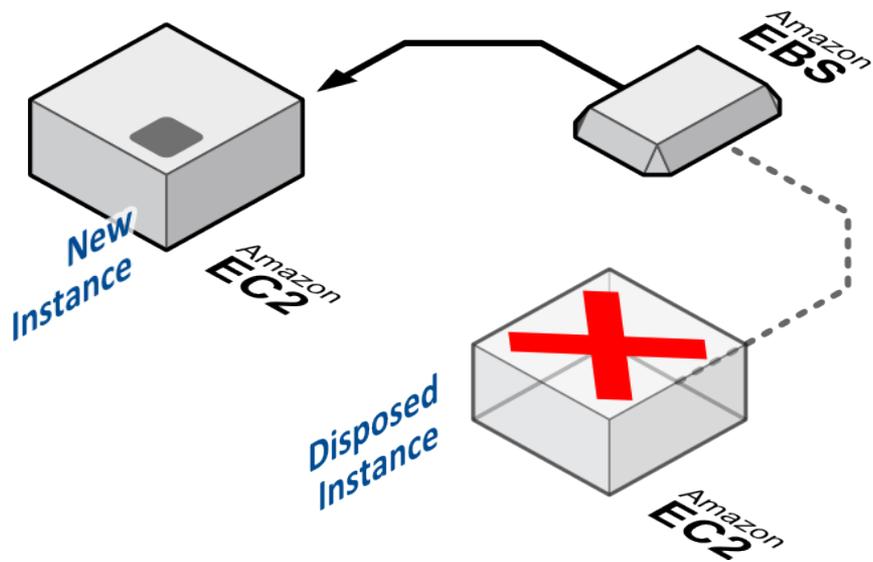


单卷更大容量！更高性能！

Provisioned
IOPS

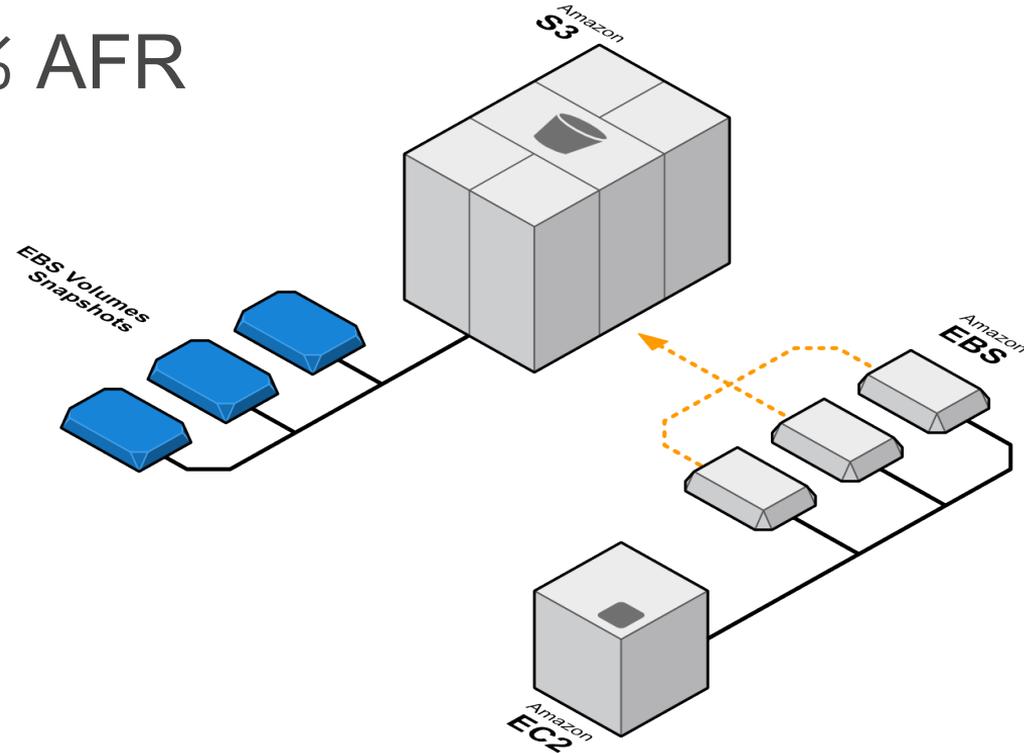
20,000
16TB

高可用的保护？



更大的数据持久性?

0.1%-0.5% AFR



传统存储，你需要。。。

性能

- 每个应用有特定的性能需求
- SSD, SAN, NAS, DAS。。。

持久性和可用性

- RAID, B2D, B2T

成本

- 昂贵的采购成本

弹性和扩展

- 通常需要超配

云上的存储，你可以

性能

- 预配置的IOPS, SSD, EBS 优化的实例

持久性和可用性

- 简单的S3快照；可用区内 / 跨可用区 / 跨区域的复制

成本

- 用量付费

弹性和扩展

- 从小开始，弹性扩展，没有限制

什么情况下用EBS？什么情况用实例存储？

- ❏ 长期持久化的存储
- ❏ 数据频繁更改
- ❏ PIOS 用于数据库
- ❏ 文件系统：NTFS, ExtFS, RAID, LVM...
- ❏ 裸设备访问

更多资源

- ❏ AWS 架构中心

<http://aws.amazon.com/architecture/>

- ❏ AWS 上的存储选项

http://media.amazonwebservices.com/AWS_Storage_Options.pdf

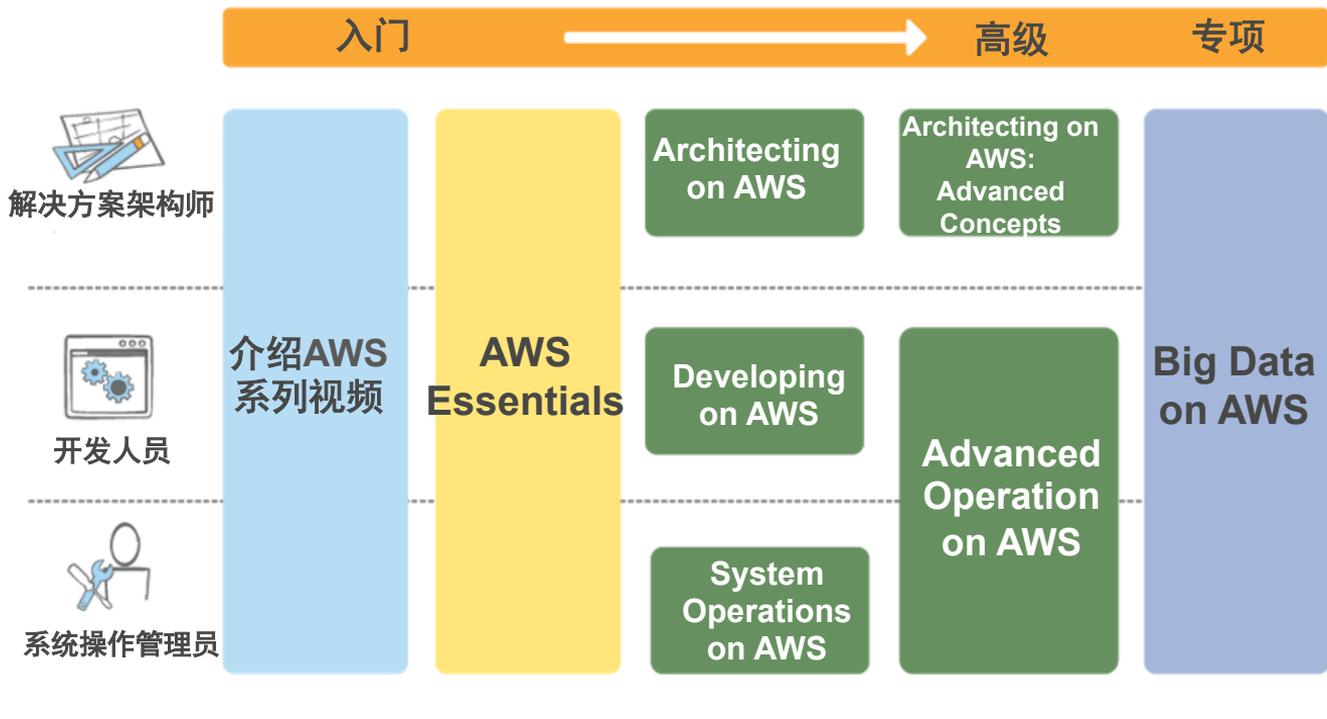
- ❏ AWS 安全白皮书

http://media.amazonwebservices.com/pdf/AWS_Security_Whitepaper.pdf

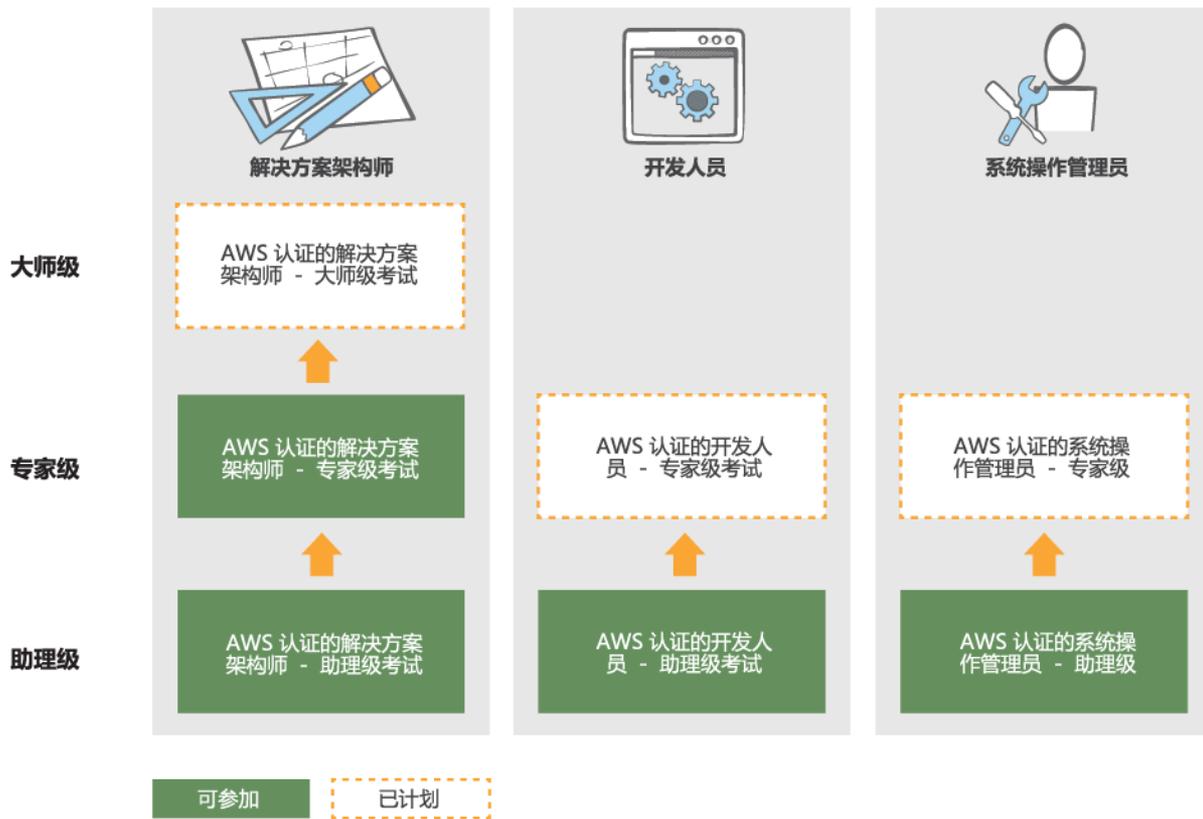
- ❏ AWS Test Drive Labs – Try your application in the Cloud

<http://www.appsassociates.com/awslabs/>

基于IT角色的讲师指导课程



AWS 认证路线图



AWS培训与认证

免费教学视频

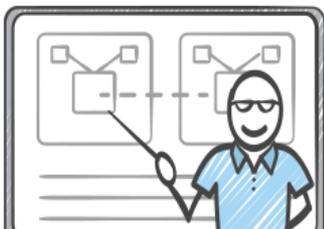


亚马逊EC2介绍

免费的教学视频让您在30分钟之内快速了解 AWS 云服务

aws.amazon.com/bokecc.com

讲师指导课程



利用AWS设计，部署和操作可扩展的，高效应用的专业能力

aws.amazon.com/cn/training

认证计划



彰显你在AWS平台上的专业技术能力

aws.amazon.com/cn/certification

问答环节