



基于AWS的HPC解决方案

代闻

wendai@amazon.com

AWS解决方案架构师





太阳能电池材料 - 有机物 vs. 硅基

Organic



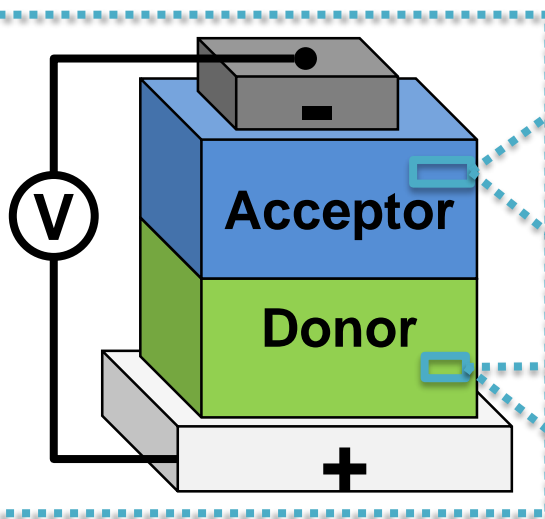
Inexpensive

Silicon

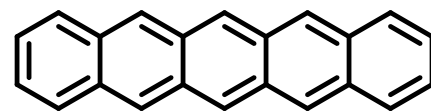
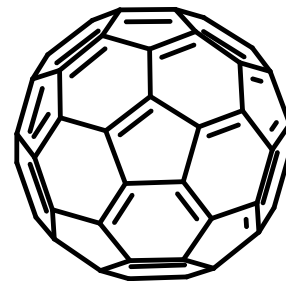


Efficient

有机物太阳能电池的原理

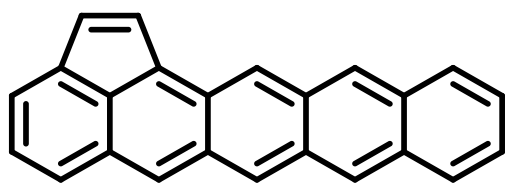


Bucky ball

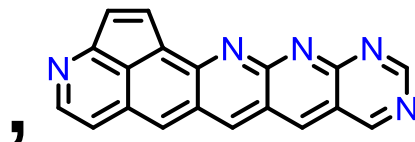
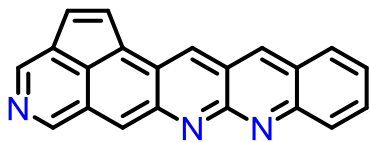
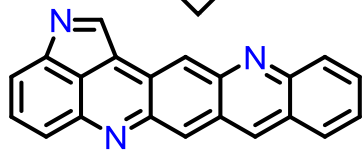


Pentacene

生成目标库



1

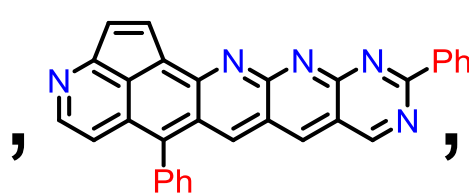
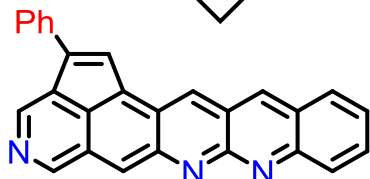
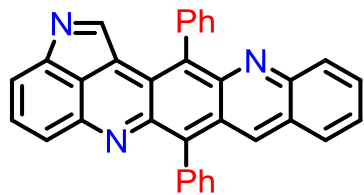


, ... 3,473

2

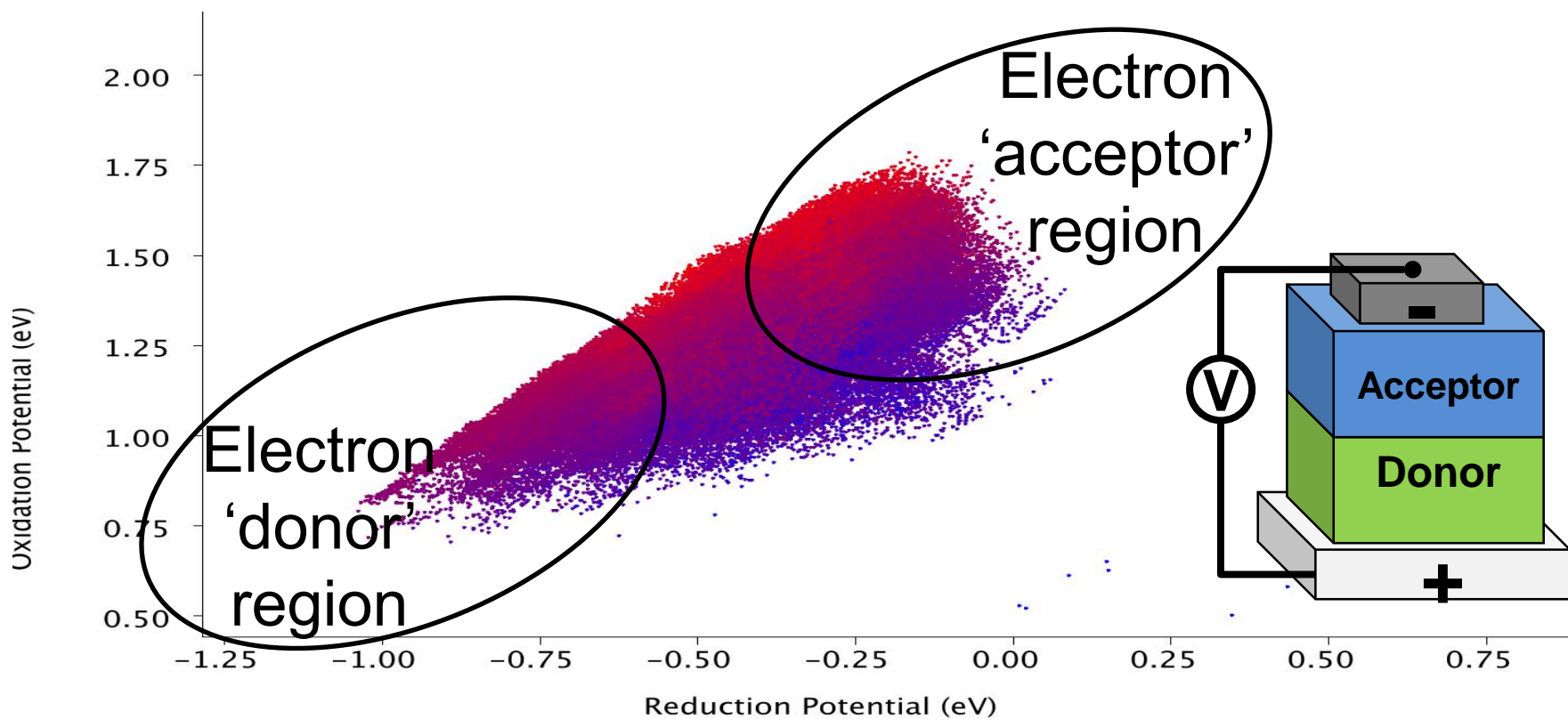
2

2

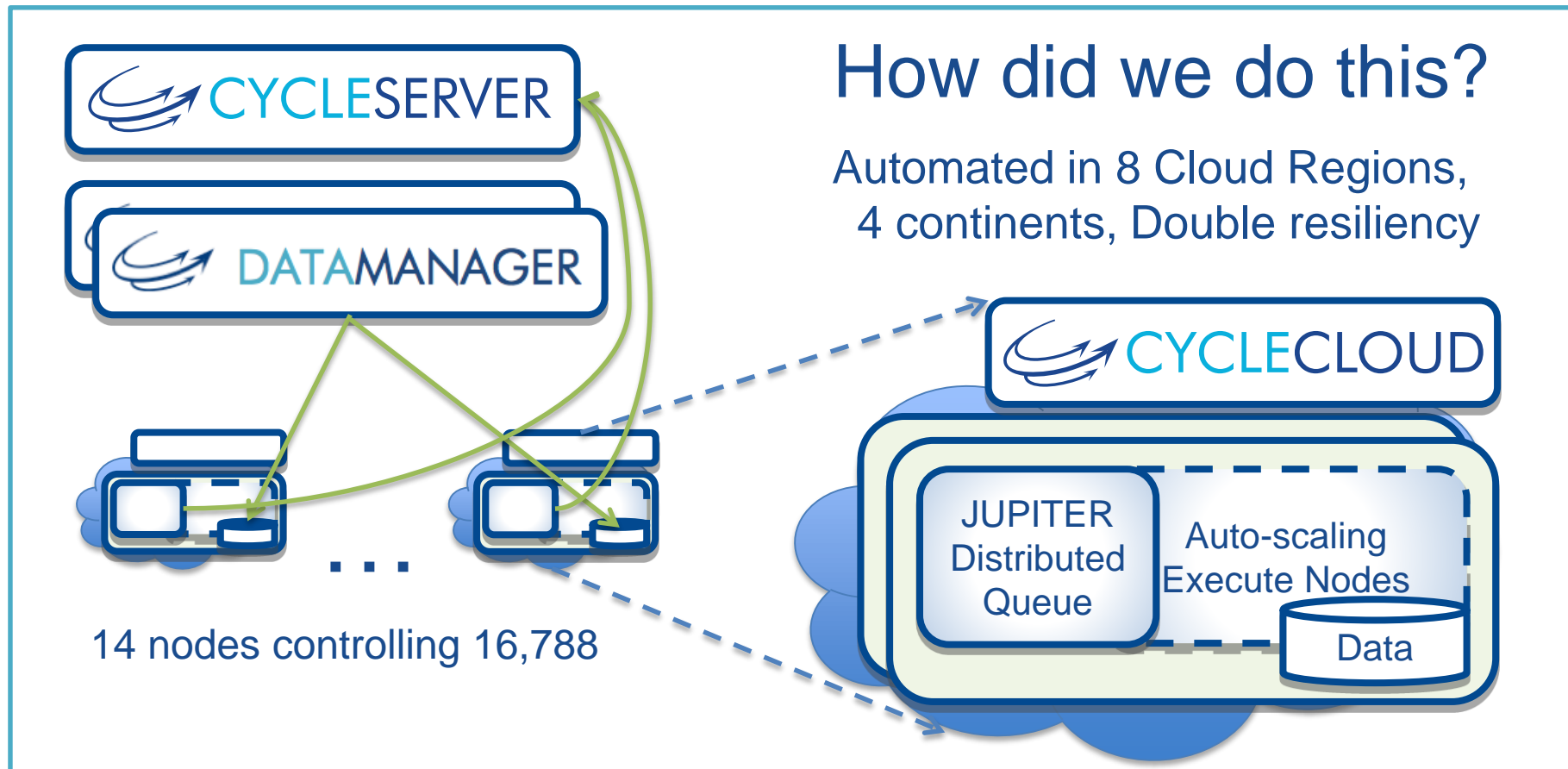


, ... 553,855

计算结果



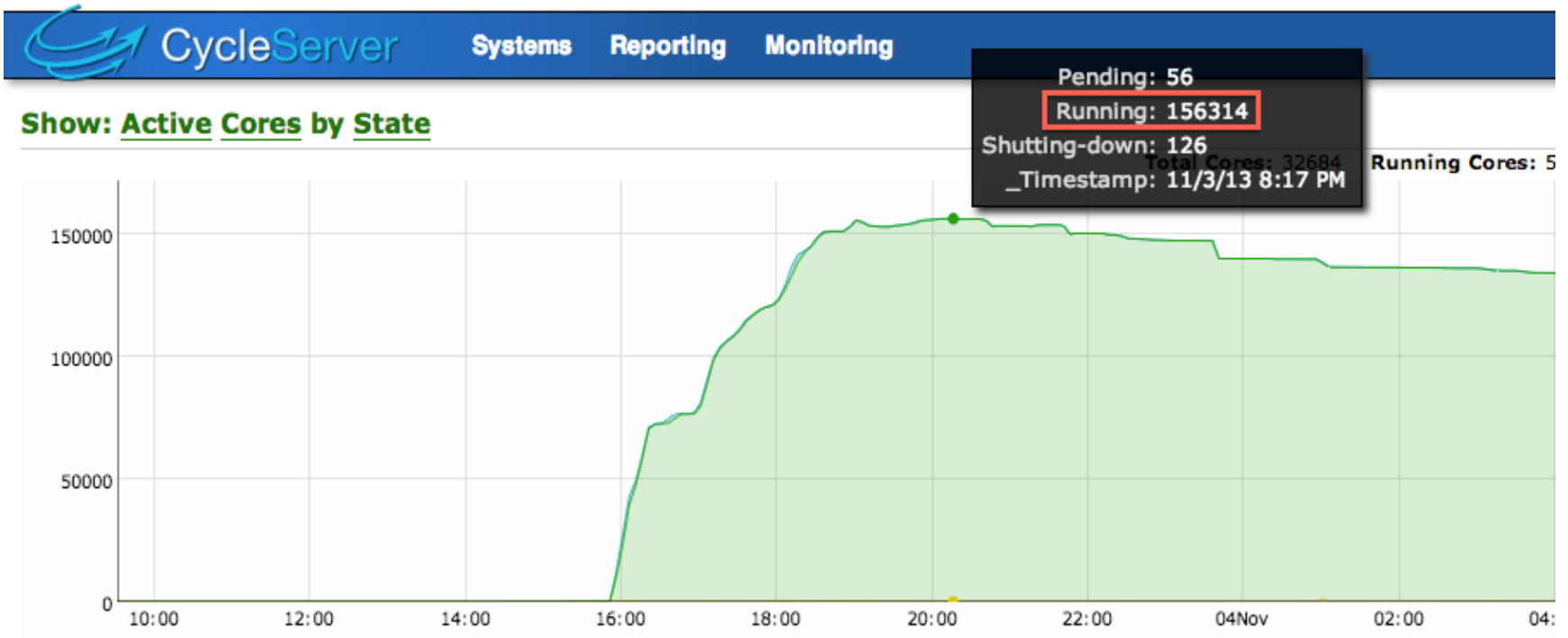
HPC 方案架构



How did we do this?

Automated in 8 Cloud Regions,
4 continents, Double resiliency

1.21 PetaFLOPS, 156,314 cores



PetaFIOP: 每秒千万亿次浮点运算

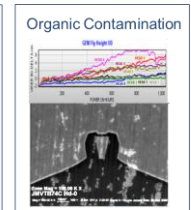
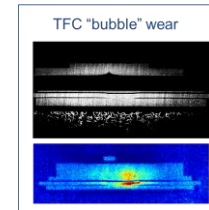
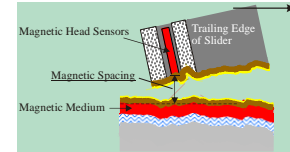
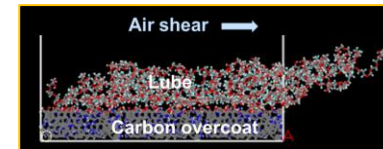
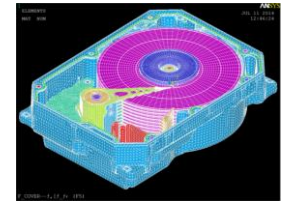
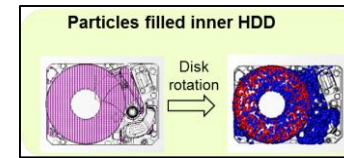
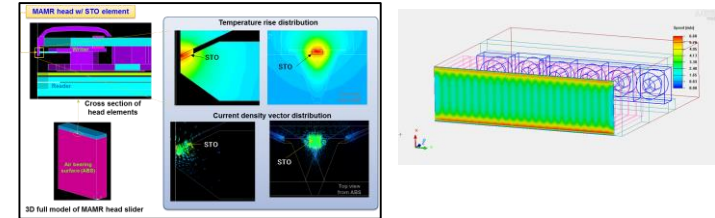
The logo features the letters 'HGST' in a bold, blue, sans-serif font. To the left of the 'H', there is a graphic element consisting of several thin, blue lines radiating outwards from a central point, resembling a stylized sunburst or a fan of data lines.

HGST

a Western Digital company

HGST的创新引擎：HPC 建模与仿真

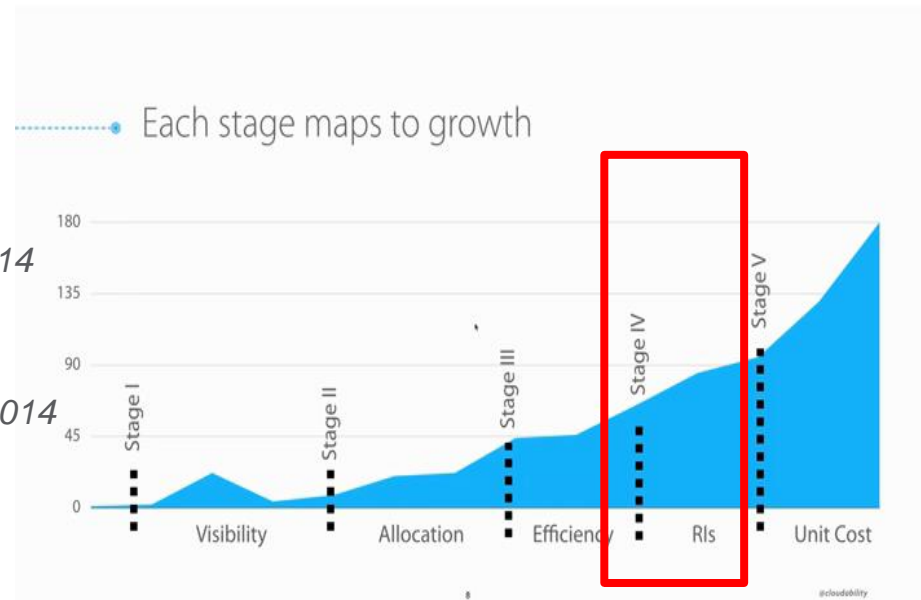
- 提高机械创新
 - Internal/External Mechanical Structural Analysis of HDD
 - Critical Lubricant Attributes and Physics
 - Airflow / He inside HDD
 - Optimal combination of HDD head and media compositions, spindle design, lubricants
 - Storage Array: HDD location, airflow investigations
- 提高单位面积存储密度(Aerial Density)
 - Micro magnetic analysis for Heat Assisted Magnetic Recording (HAMR)
 - Head-Medium Spacing (HMS)



HPC Doing The “Physics Work” Driving HGST Innovation

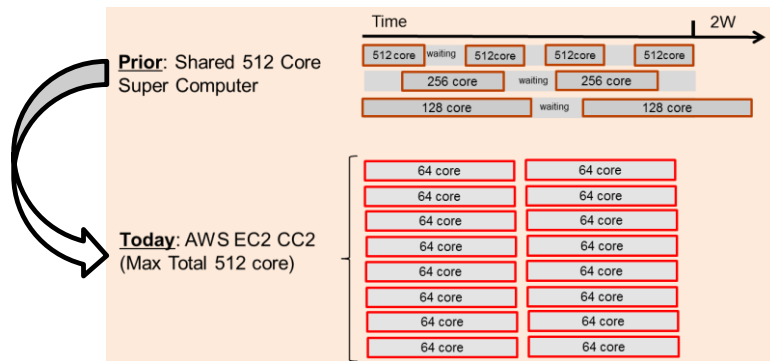
HGST基于AWS的HPC的五个阶段

- 阶段1：可行性验证测试（PoC）
 - *First HPC PoC: Sept 2013*
- 阶段2：小范围使用（Small Start）
 - *1st HPC Production Cluster: Nov 2013*
- 阶段3：优化工作负载和灵活性（Optimize Workloads And Flexibility）
 - *AWS C3 Deployments Jan 2014*
 - *4th HPC Production Cluster: June 2014*
- 阶段4：降低成本（Lower Cost）
 - *Use Spot And Reserved Instances: Oct / Nov / Dec 2014*
- 阶段5：结合业务指标（Business metrics）
 - *Utilization and cost reports to HGST engineers : Dec 2014*

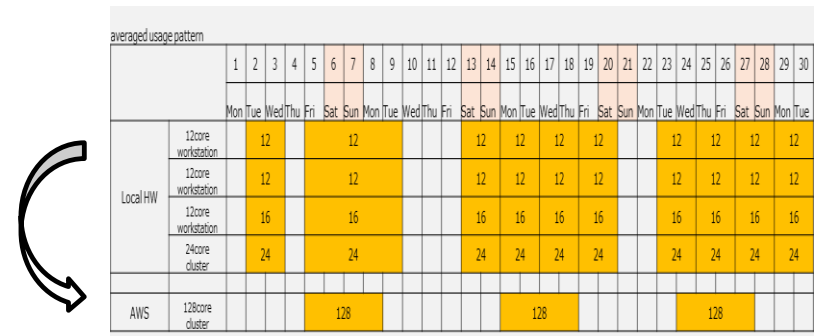


Stage 1 + 2 + 3: 定义并扩展计算规模

分子动力学 (Molecular Dynamics) 1.67x Initial Overall Throughput Gain



流体动力学 (Fluid Dynamics) 1.4x Overall Throughput Gain



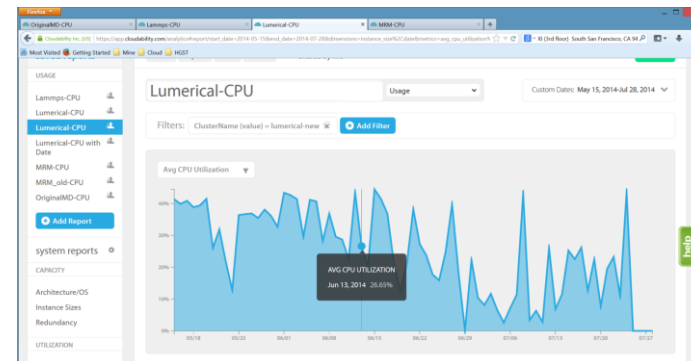
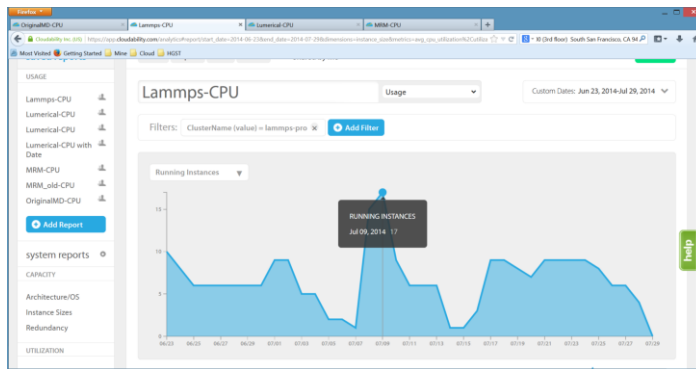
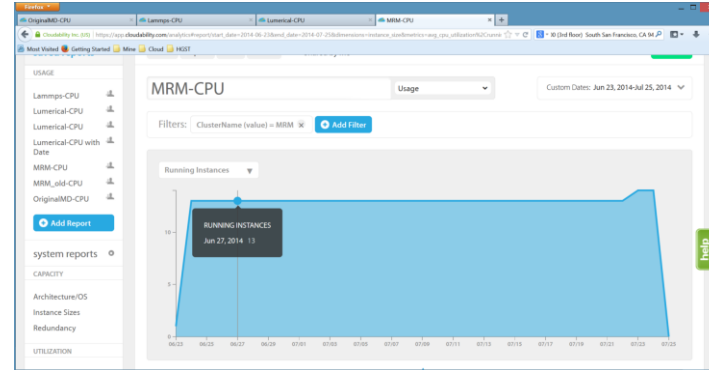
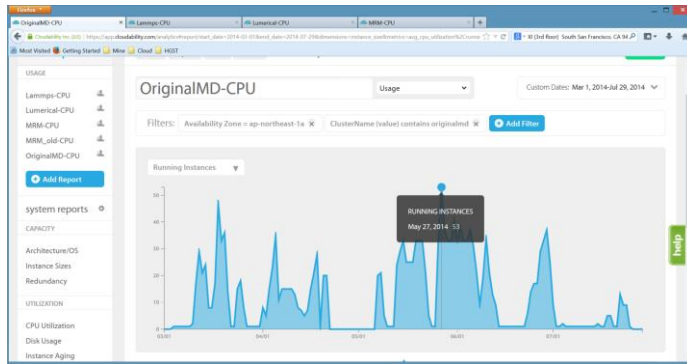
分子动力学 (Molecular Dynamics)

Simulation Type	Throughput Increase
Head Drive Interface Vacuum Gaps	1.99x
Vacuum Gap "collection"	4.00x
Media Grains for HAMR (FePt/C)	2.03x
4 Carbon Molecule Clusters	5.67x

微磁学 (MicroMagnetics)

Parameter Sweeps	Throughput Gain
Model 1	1.23x - 1.78x
Model 2	1.01x - 1.67x
Model 3	1.23x - 1.69x
Model 4	Up to 2.7x

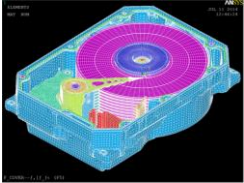
Stage 3: 针对不同的工作负载进行优化



Not all workloads and work require same compute resources 24 x 7 x 365

Stage 4: 竞价实例在研究中的应用

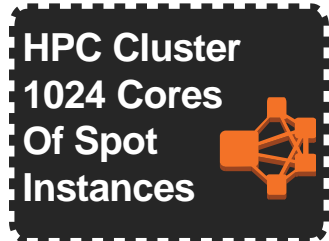
New Drive
Head
Design
Workloads



Submit jobs,
orchestrate HPC
clusters over VPN



Encrypt, route data to
AWS, return results



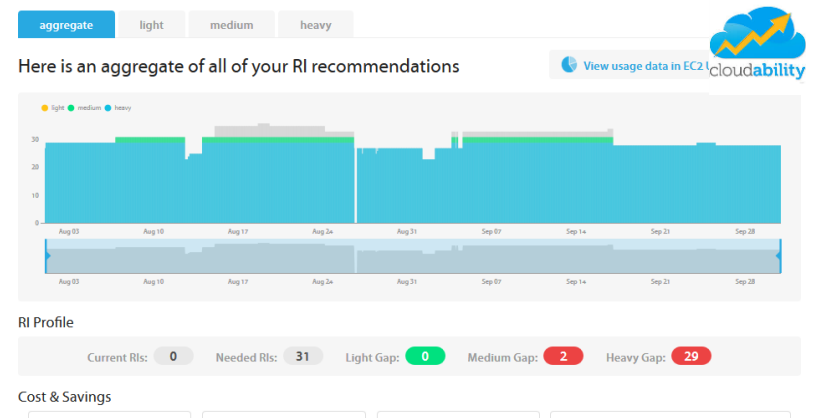
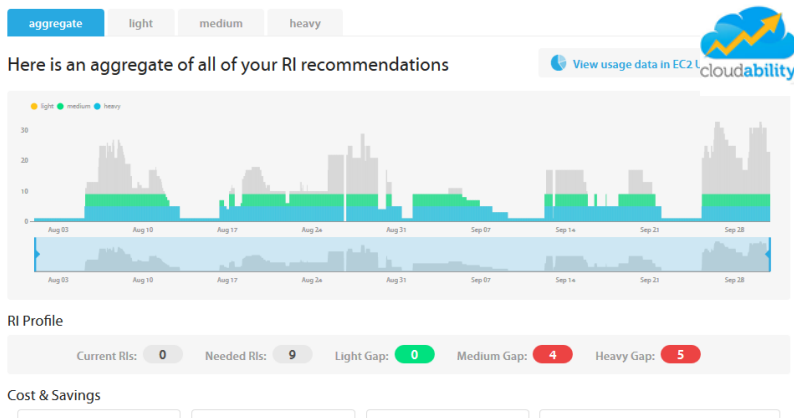
Simulated 22 advanced
head designs across 3
materials possibilities
= 15 compute years

Used AWS c3 instances

6x faster run-time:
Ran in 5 days, not 30!

Total cost:
\$4,026.02

Stage 4 + 5: 优化整体成本



Great Solutions Are Available To Ease Optimization Effort

Pfizer使用AWS处理HPC计算
高峰和激增需求。

“

AWS enables Pfizer's Worldwide Research and Development to explore specific difficult or deep scientific questions in a timely, scalable manner and helps Pfizer make better decisions more quickly.

Dr. Michael Miller
Head of HPC for
R&D, Pfizer



”

HARVARD医学院通过使用AWS，将HPC集群的成本降低了50%

“

The combination of our approach to biomedical computing and AWS allowed us to focus our time and energy on simulation development, rather than technology, to get results quickly. Without the benefits of AWS, we certainly would not be as far along as we are.

Dr. Peter Tonellato,
LPM, Center for
Biomedical Informatics,
Harvard Medical School



”

Unilever通过使用AWS，基因组序列的计算集群的运算速度提升了20倍。

“

The key advantage that AWS has over running this workflow on Unilever's existing cluster is the ability to scale up to a much larger number of parallel compute nodes on demand.

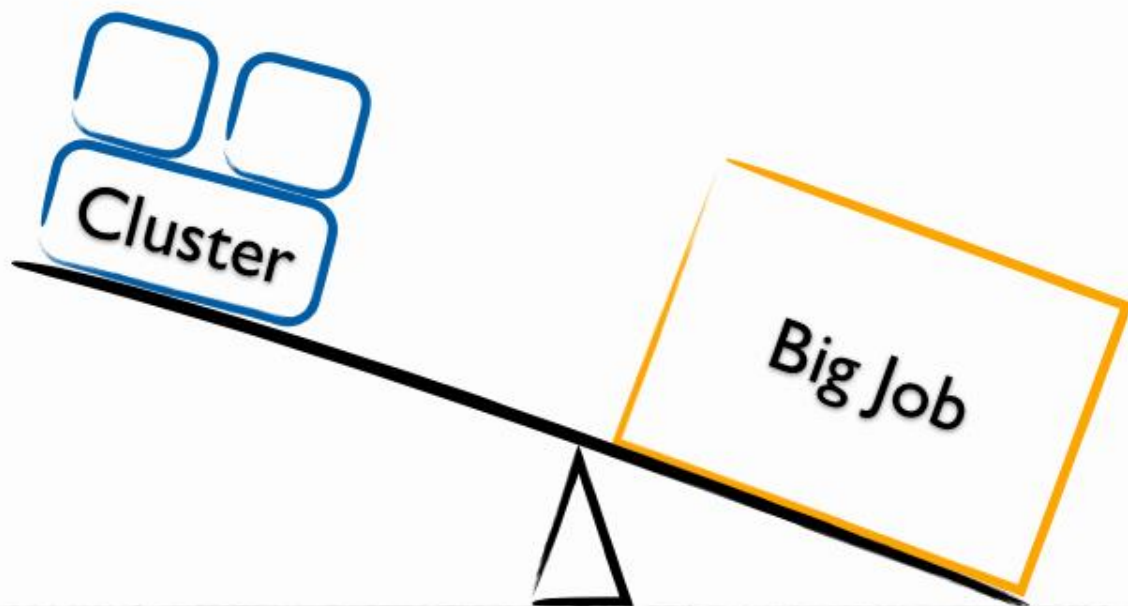
Pete Keeley
Unilever Researchs eScience
IT Lead for Cloud Solutions



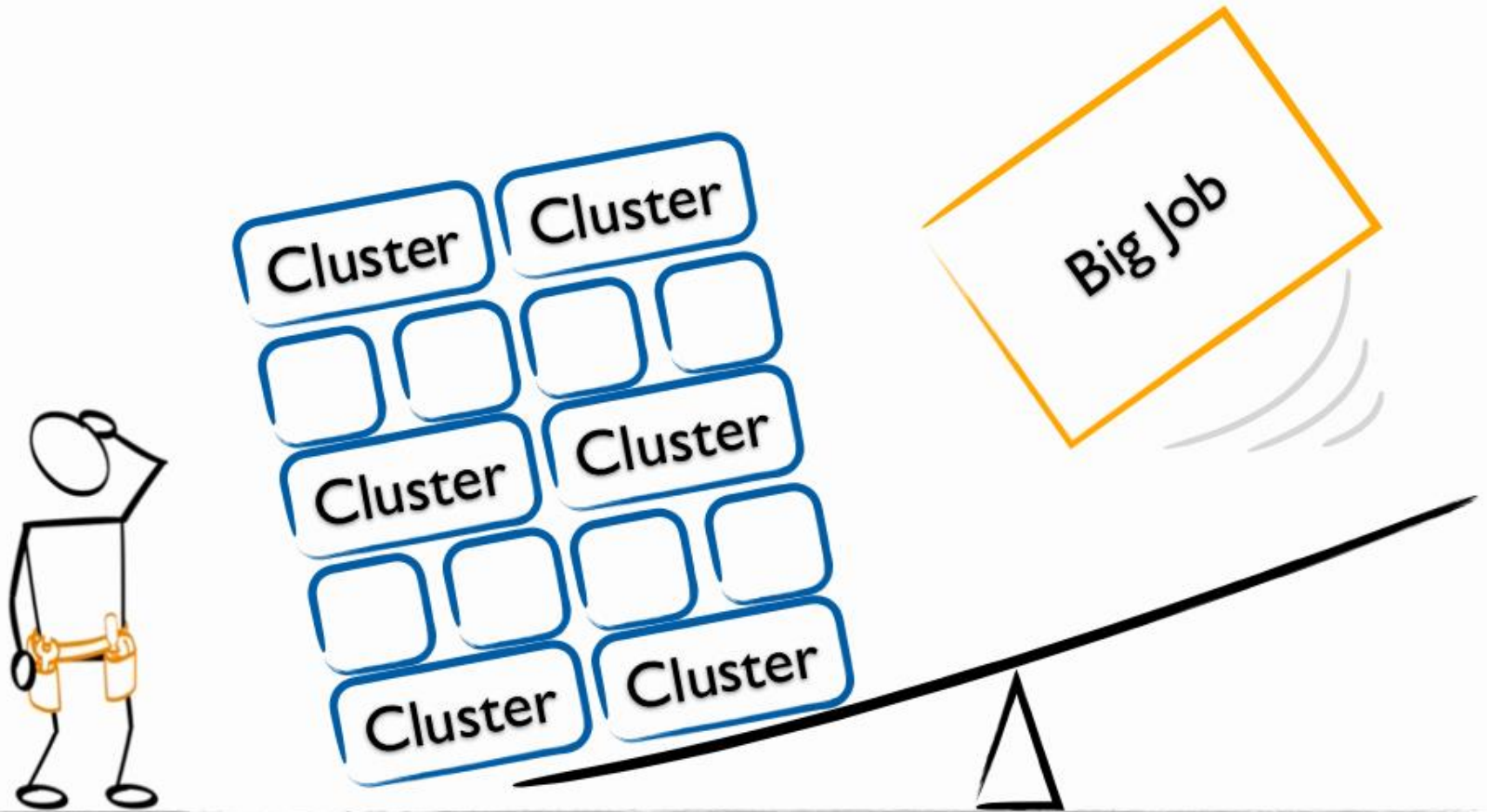
”

为什么在云上运行HPC?

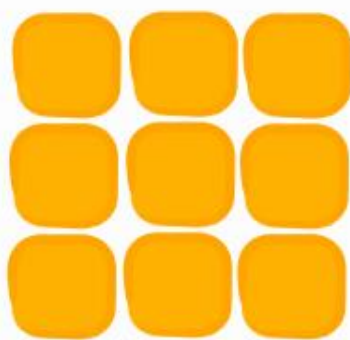
我们需要一个足够大的集群



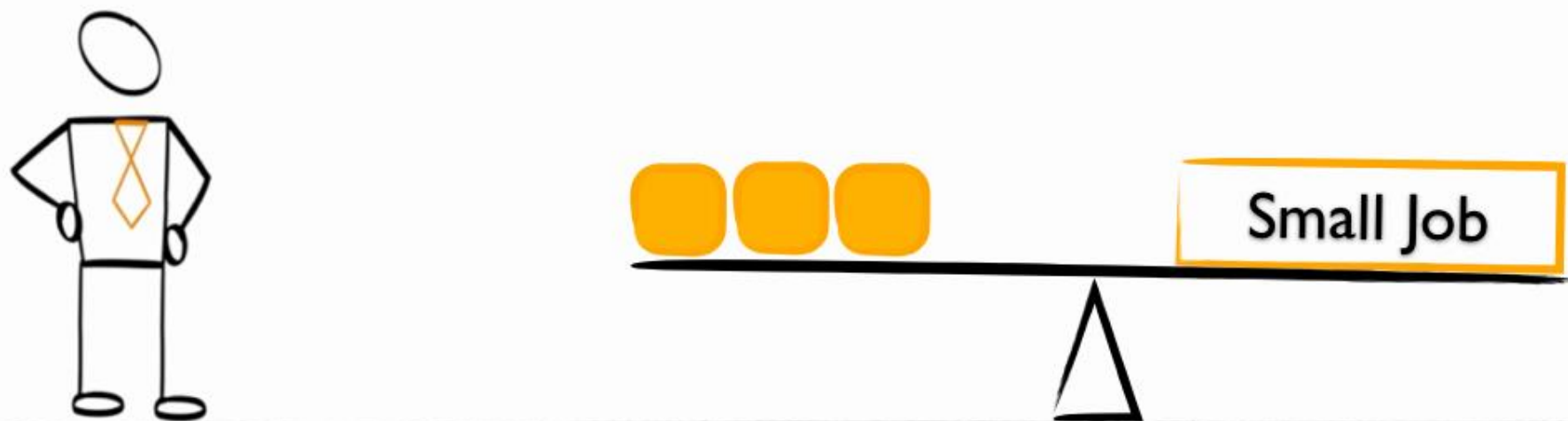
但太大的集群又会是浪费...



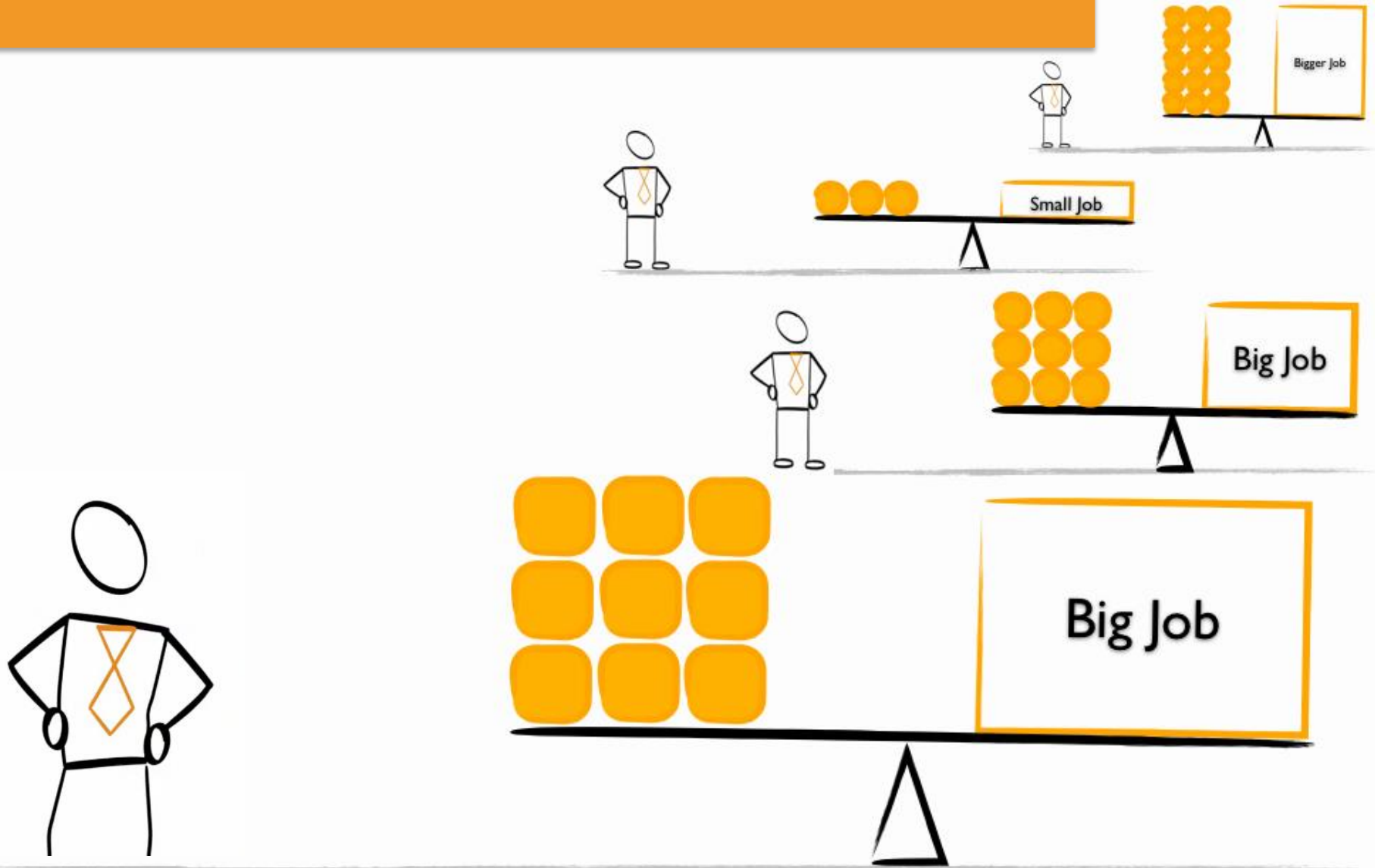
如果有一个大小刚好的集群...



而且是按需的、弹性的...



而且同时多个任务并行计算...



为什么在AWS上运行HPC?

Figure 1. Magic Quadrant for Cloud Infrastructure as a Service, Worldwide



Source: Gartner (May 2015)

AWS has ten times more cloud capacity in use than the other 14 providers combined.
-- Gartner (2015)

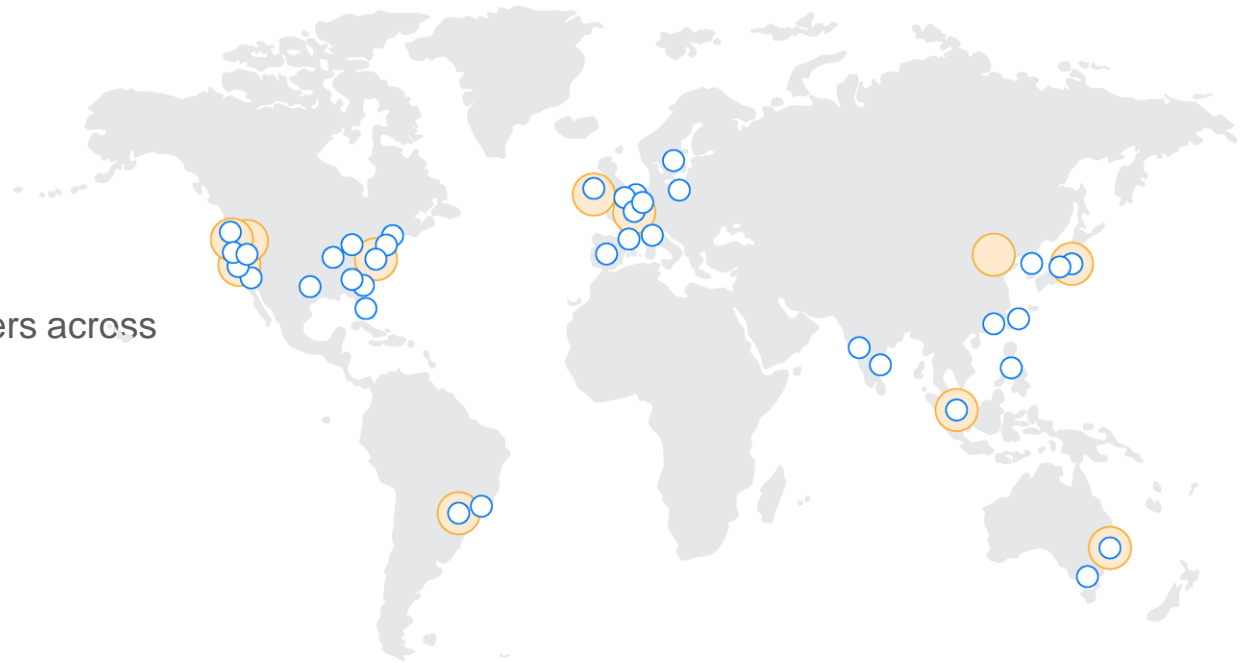
AWS全球基础架构

11 regions

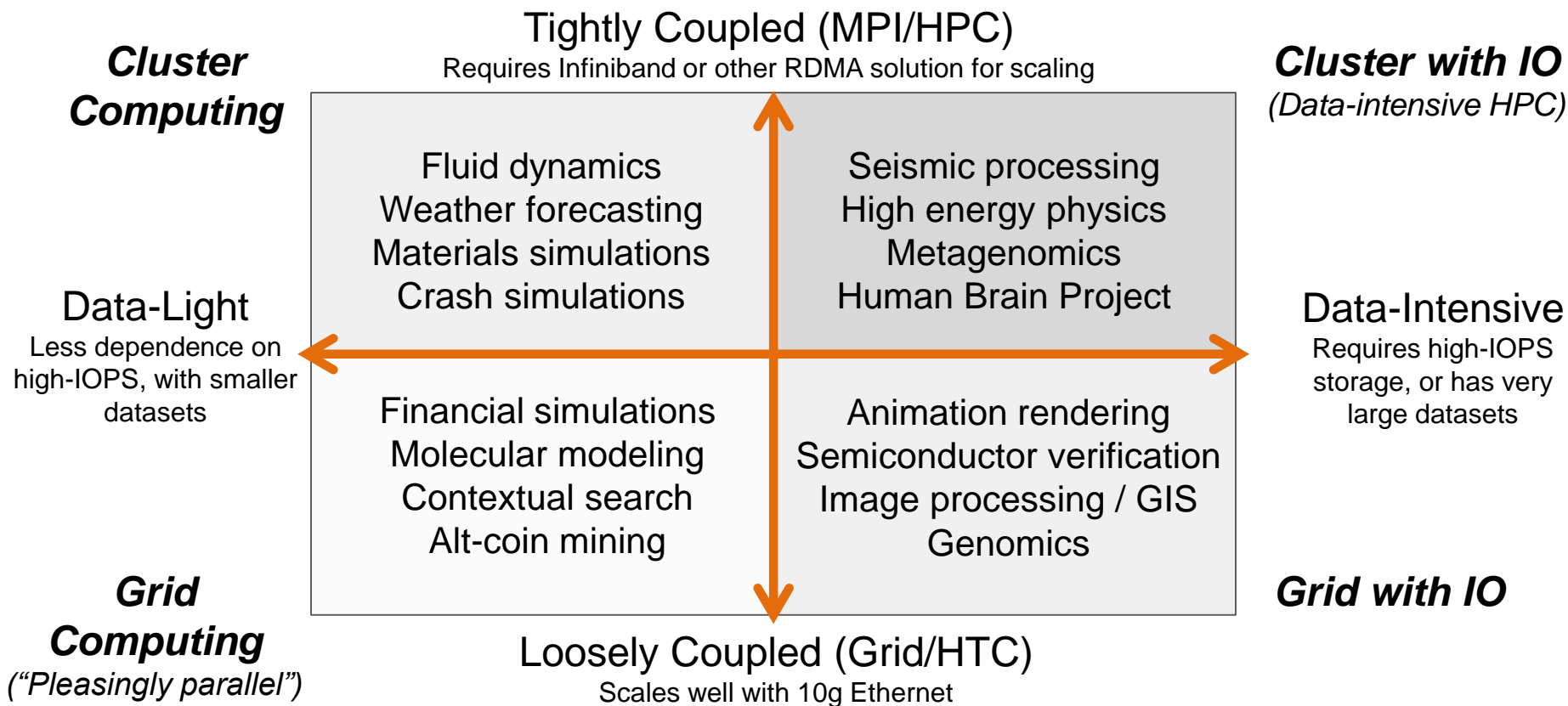
30 availability zones

53 edge locations

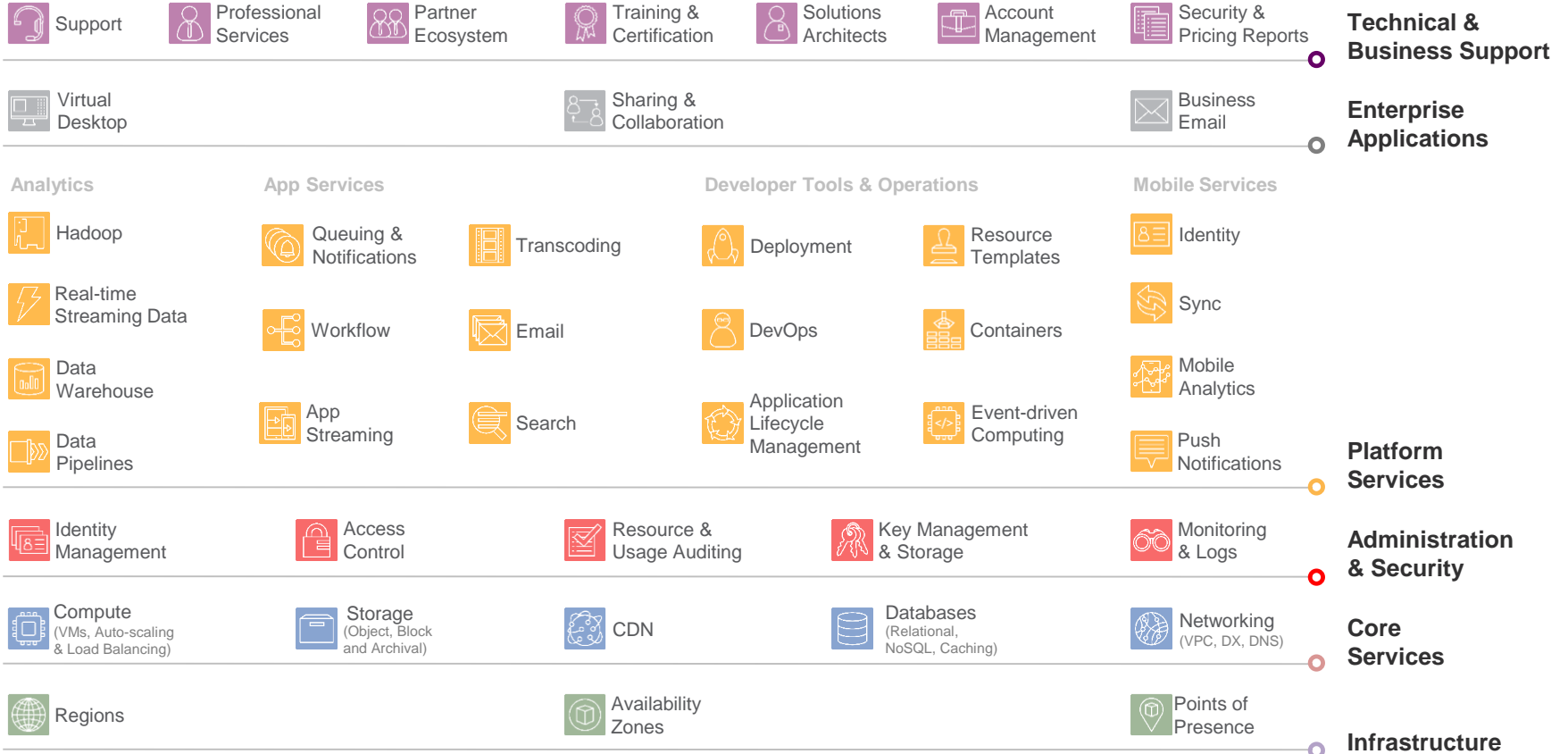
Over 1 million **active** customers across
190 countries



AWS全面支持HPC的各类应用



AWS服务分类






计算和网络资源

计算优化以及集群计算实例 (Cluster Compute)


Implement HVM process execution

Intel® Xeon® processors


10 Gigabit Ethernet – c3 has Enhanced networking, SR-IOV



32 vCPUs
2.8 GHz Intel Xeon
E5-2680v2 Ivy Bridge




60GB RAM




2 x 320 GB
Local SSD


c3.8xlarge



32 vCPUs
2.6 GHz Intel Xeon
E5-2670 Sandy Bridge



60.5 GB RAM



2 x 320 GB
Local SSD

cc2.8xlarge

GPU运算实例

CG1 instances

Intel® Xeon® X5570 processors

2 x NVIDIA Tesla “Fermi” M2050 GPUs


I/O Performance: Very High (10 Gigabit Ethernet)

G2 instances


Intel® Intel Xeon E5-2670

1 NVIDIA Kepler GK104 GPU


I/O Performance: Very High (10 Gigabit Ethernet)



33.5 EC2 Compute
Units




20GB RAM




2x NVIDIA GPU
448 Cores
3GB Mem


cg1.8xlarge



26 EC2 Compute
Units



16GB RAM

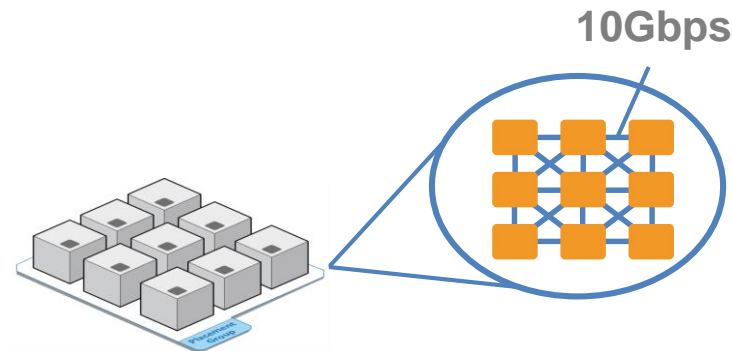


1x NVIDIA GPU
1536 Cores
4GB Mem

g2.2xlarge

万兆低延迟网络 (Network placement groups)

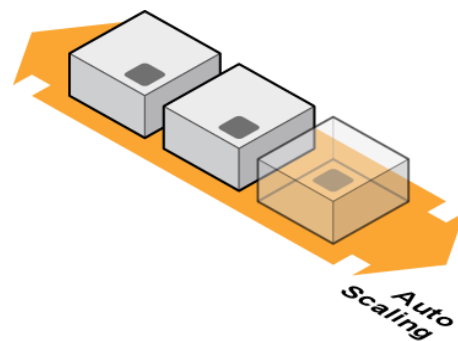
集群的计算实例部署在一个“Placement Group”中，通过低延迟的万兆网络互联。



自动化控制

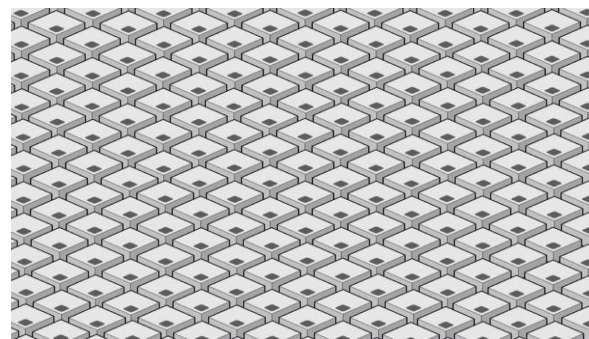
自动扩展 (Auto Scaling Group)

根据计算实例的工作状态 (如CPU、网络、存储IO等) 实现自动扩展和收缩，实现资源弹性和并行计算。

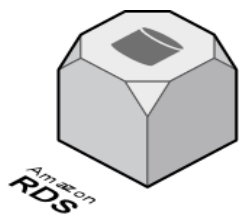


自动化控制

支持通过图形界面、命令行、SDK和API的方式控制资源，最大程度实现自动化，以及平台整合。

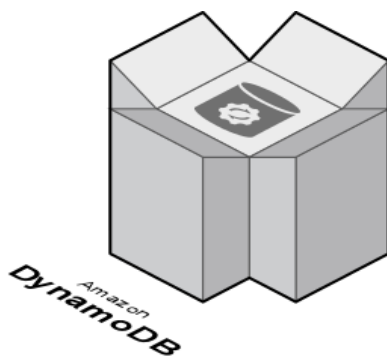


数据管理



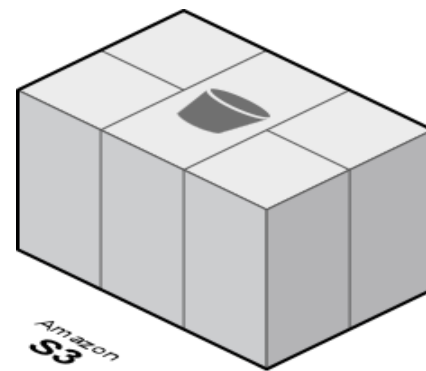
关系型数据库服务 (RDS)

(MySQL, Oracle, MSSQL, PostgreSQL, Aurora)



NoSQL数据库 (DynamoDB)

无缝变更, 弹性扩展的NoSQL数据库服务



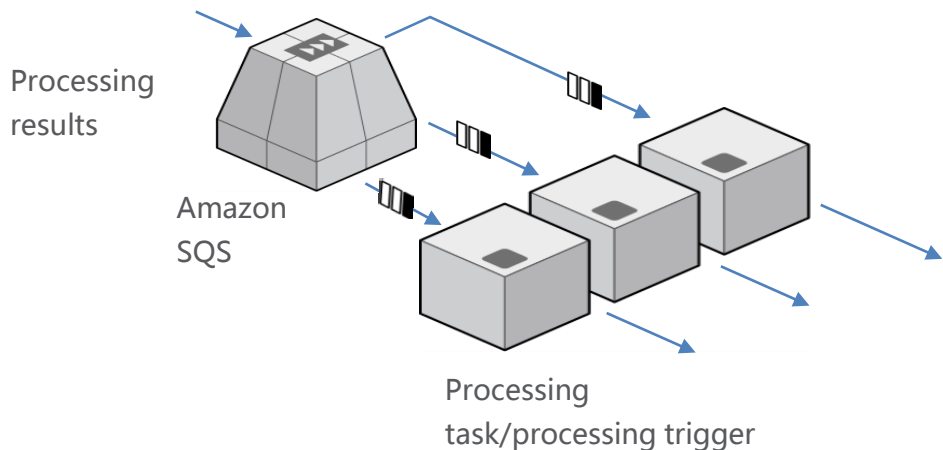
对象存储 (S3)

单个文件最大支持 5TB，面向 Internet，并具有灵活的访问控制。

负载分配与 workflow

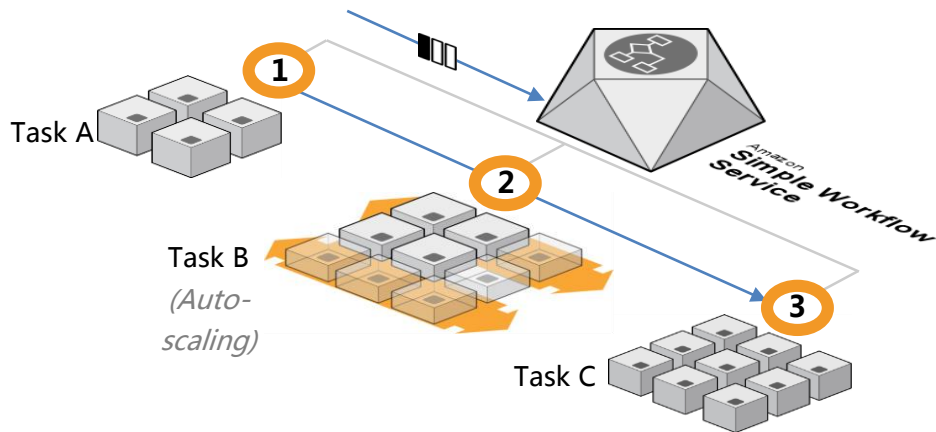
消息队列服务 (SQS) 实现负载分配

通过高可用的消息队列 (Simple Queue Service) 来分发任务负载。



workflow 服务协调负载和任务集群

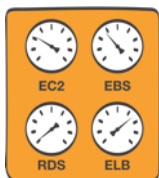
处理流程较长的任务，使用 workflow 服务 (Simple Workflow Service) 按照步骤调用各个节点，并定义不同的策略和工作逻辑。



自动化部署和管理服务

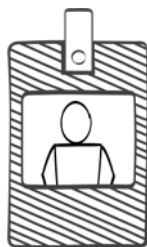
Amazon CloudWatch

Monitor resources



AWS IAM (Identity & Access Mgmt)

Manage **users, groups & permissions**



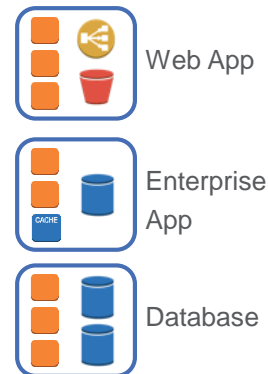
AWS OpsWorks

Dev-Ops framework for application lifecycle management



AWS CloudFormation

Templates to deploy & manage



AWS Elastic Beanstalk

Automate resource management



成本优化：多种计价模型

免费资源

第一年有一定量的免费资源，仅限海外。

主要用于PoC或对AWS环境的熟悉过程



按需实例

按小时付费

临时使用，应对突发的负载。



预留实例

按年承诺使用，支持无首付、部分首付和全付。与按需实例相比，年平均成本大大降低。

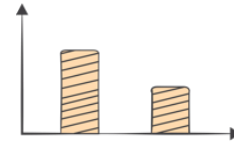
适用于必需的计算资源。



竞标实例

对AWS空闲计算资源进行竞标，有机会以极低价格获得计算实例。

用于补充需要花费大量时间的负载的计算能力，或提供临时的计算能力。



专用实例

开启计算实例时要求运行在给一个用户专用的硬件设备上

适用于异常敏感或对合规有要求的计算负载。



AWS定价策略

- 自2006年成立以来，AWS宣布了49次降价
- 新功能往往会更强大，并且更低价



在AWS上运行HPC的客户还有很多

Lawrence Livermore
National Laboratory



HITACHI

amazon.com



BROOKHAVEN
NATIONAL LABORATORY



nimbic

LIONSGATE

illumina®



Berkeley
UNIVERSITY OF CALIFORNIA



Numerate

Autodesk®

NYU School of Medicine
NYU LANGONE MEDICAL CENTER



UniCredit



THE CLIMATE
CORPORATION



CYCLECOMPUTING
LEADER IN CONDO GRID COMPUTING SOLUTIONS



ERICSSON



bankinter.



HGST
a Western Digital company

NASDAQ®

Capital IQ
A Standard & Poor's Business

DNAexus

Mentor
Graphics®

MIT
Massachusetts
Institute of
Technology

Pathwork
Diagnostics

同时，客户使用AWS处理越来越多的HPC负载

石油与天然气

Seismic Data Processing

Reservoir Simulations, Modeling

Geospatial applications

Predictive Maintenance

制造工程

Computational Fluid Dynamics (CFD)

Finite Element Analysis (FEA)

Wind Simulation

生命科学

Genome Analysis

Molecular Modeling

Protein Docking

媒体和娱乐

Transcoding and Encoding

DRM, Encryption

Rendering

科学计算

Computational Chemistry

High Energy Physics

Stochastic Modeling

Quantum Analysis

Climate Models

金融

Monte Carlo Simulations

Wealth Management Simulations

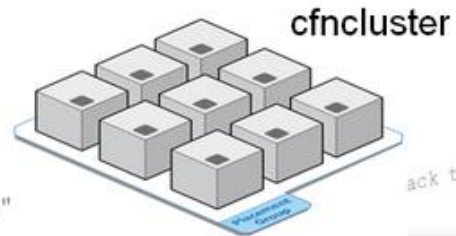
Portfolio, Credit Risk Analytics

High Frequency Trading Analytics

AWS的HPC合作伙伴



AWS HPC Demo资源



cfnccluster ("CloudFormation cluster")

Command Line Interface Tool

Deploy and demo an HPC cluster

For more info:

<http://aws.amazon.com/hpc/resources>

```
"AWSTemplateFormatVersion" : "2010-09-01",
"Description" : "This template creates your live event.",
"Parameters" : {
  "InstanceType" : {
    "Type" : "String",
    "Description" : "EC2 Instance Type",
    "Default" : "m3.xlarge",
    "AllowedValues" : [ "m3.xlarge", "m3.2xlarge", "m3.4xlarge" ],
    "ConstraintDescription" : "Must be one of the allowed values."
  }
}
```

```
aws: cfncluster --help
usage: cfncluster [-h] [--config CONFIG_FILE] [--region REGION] [--noexit]
                 [cluster, stop, create, list, update, instances, subnets, delete]
cfncluster is the tool to launch and manage cluster.

positional arguments:
  cluster, stop, create, list, update, instances, subnets, delete
  cluster                creates a cluster
  update                 updates a running cluster
  stop                  stops a cluster
  delete                 deletes a cluster
  status                pulls the current status of the cluster
  list                  displays a list of stacks associated with cfncluster
  instances              displays a list of all instances in a cluster
  subnets               sets to Master Instance

optional arguments:
  -h, --help            show this help message and exit
  --config CONFIG_FILE, -c CONFIG_FILE
                        specify an alternative config file
  --region REGION, -r REGION
                        specify a specific region to connect to
  --noexit, --no       do not exit for stack events, after executing stack
                        command
```

ack that uses Amazon Clo

Val

第十七届全国科学计算与信息化会议 AWS 反馈问卷



您的反馈，很重要！

谢谢大家填写反馈建议，
您将得到50美金AWS服务抵扣券。

感谢您的宝贵时间，欢迎提问！



China SMB and Startups on AWS



CAMERA360



China Enterprises on AWS



Amazon Partners Network in China



Neusoft



ThoughtWorks®

