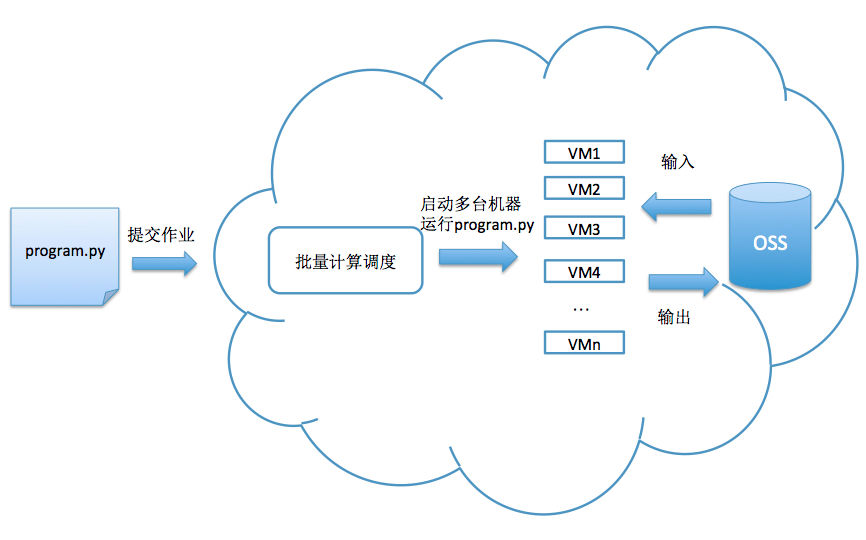
**什么是批量计算**

批量计算（BatchCompute）是一种适用于大规模并行批处理作业的分布式云服务。BatchCompute 可支持海量作业并发规模，系统自动完成资源管理、作业调度和数据加载，并按实际使用量计费。

通俗的讲，就是您可以提交一个任意的计算机程序，让它在阿里云的多个 VM 实例上同时运行，然后把结果写入到指定的持久化存储位置（如阿里云对象存储 OSS 或者文件存储 NAS），然后您可以去指定的位置查看结果。

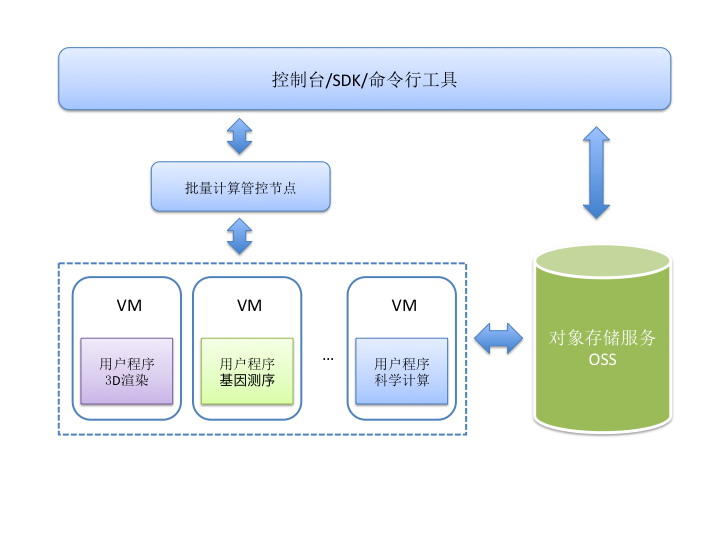
[](http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/assets/pic/27994/cn_zh/1524795335543/10_14_18__04_27_2018.jpg)

# 产品架构

更新时间：2018-10-26 19:53:10

BatchCompute 整体架构，包括：

* BatchCompute 入口
  + 批量计算服务提供基于 RESTful 风格的 API。在 API 之上，用户可以通过 SDK，命令行工具，控制台等方式使用批量计算。
  + 您可以使用这些工具向 BatchCompute 提交作业，查询作业运行状态，并管理作业生命周期（比如停止已提交的作业，释放已完成的作业等）。

[](http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/assets/pic/42382/cn_zh/1468318746090/framework.png)

* 运行环境

BatchCompute 允许用户通过自定义虚拟机镜像或者 Docker 的方式对运行环境进行高度定制，可以支持 Windows 和 Linux 操作系统。用户程序运行在隔离的虚拟化环境中，确保用户环境和用户数据的安全性。

* 持久化存储
  + BatchCompute可以使用对象存储 OSS 或者文件存储 NAS 作为输入输出数据的持久化存储。
  + 用户的程序、自定义 Docker 镜像、作业的运行日志存储在 OSS 中。

# 功能特性

更新时间：2018-10-26 19:53:10

[**本页目录**](javascript:void(0))

* [1. 提交作业](https://help.aliyun.com/document_detail/27996.html?spm=a2c4g.11186623.6.544.52e2452faEYa5q#h2-1-1)
* [2. 管理我的作业](https://help.aliyun.com/document_detail/27996.html?spm=a2c4g.11186623.6.544.52e2452faEYa5q#h2-2-2)
* [3. 使用集群](https://help.aliyun.com/document_detail/27996.html?spm=a2c4g.11186623.6.544.52e2452faEYa5q#h2-3-3)
* [4. 自定义镜像](https://help.aliyun.com/document_detail/27996.html?spm=a2c4g.11186623.6.544.52e2452faEYa5q#h2-4-4)

## 1. 提交作业

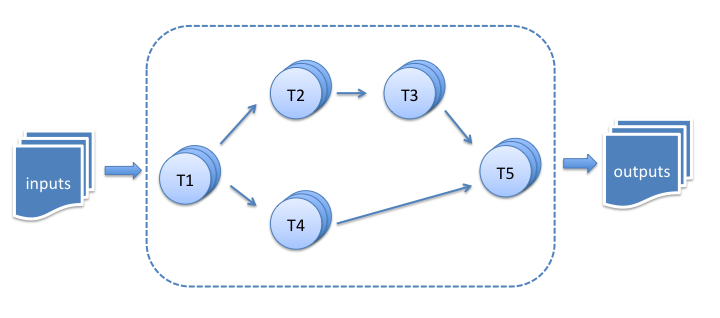
用户使用工具（如 SDK，命令行工具等）向 BatchCompute 提交作业，BatchCompute 使用用户指定的镜像（如：ubuntu）启动虚拟机（VM）, 在虚拟机中运行用户程序, 运行完成后释放虚拟机（VM）。

* BatchCompute 中使用 [OSS](https://www.aliyun.com/product/oss/) 作为持久化存储。**您可以在程序运行完成时将结果数据保存到 OSS 中。在批量计算中，也可以通过文件接口的方式访问 OSS 上的数据，请参阅**[**OSS 挂载功能**](https://help.aliyun.com/document_detail/28016.html)**。**
* BatchCompute 程序默认运行在 VM 中，也支持 Docker容器。 **也就是说，您可以**[**自定义 ECS**](https://help.aliyun.com/document_detail/28017.html)**镜像或者使用**[**Docker**](https://help.aliyun.com/document_detail/28022.html)**，在镜像中安装自己需要的任何软件，用来运行您的任何程序。**

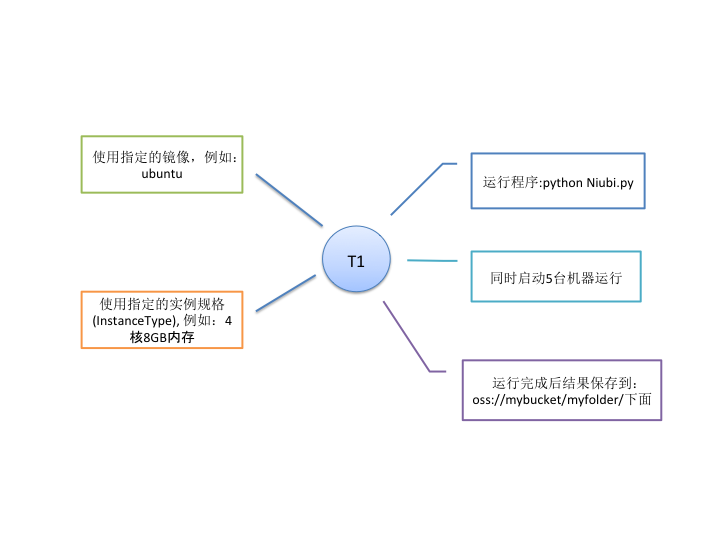
#### 作业描述

用户需要提交一个作业（Job）描述 JSON 文件到批量计算服务，该 JSON 文件中详细描述了需要执行哪些程序（支持多个程序），运行哪些程序需要启动多少台机器，机器的规格（内存和CPU等），运行日志打印到哪里，完成后结果输出到哪里等。

一个作业（Job）包含多个任务（Task）, 按照您指定 DAG 描述的顺序执行。

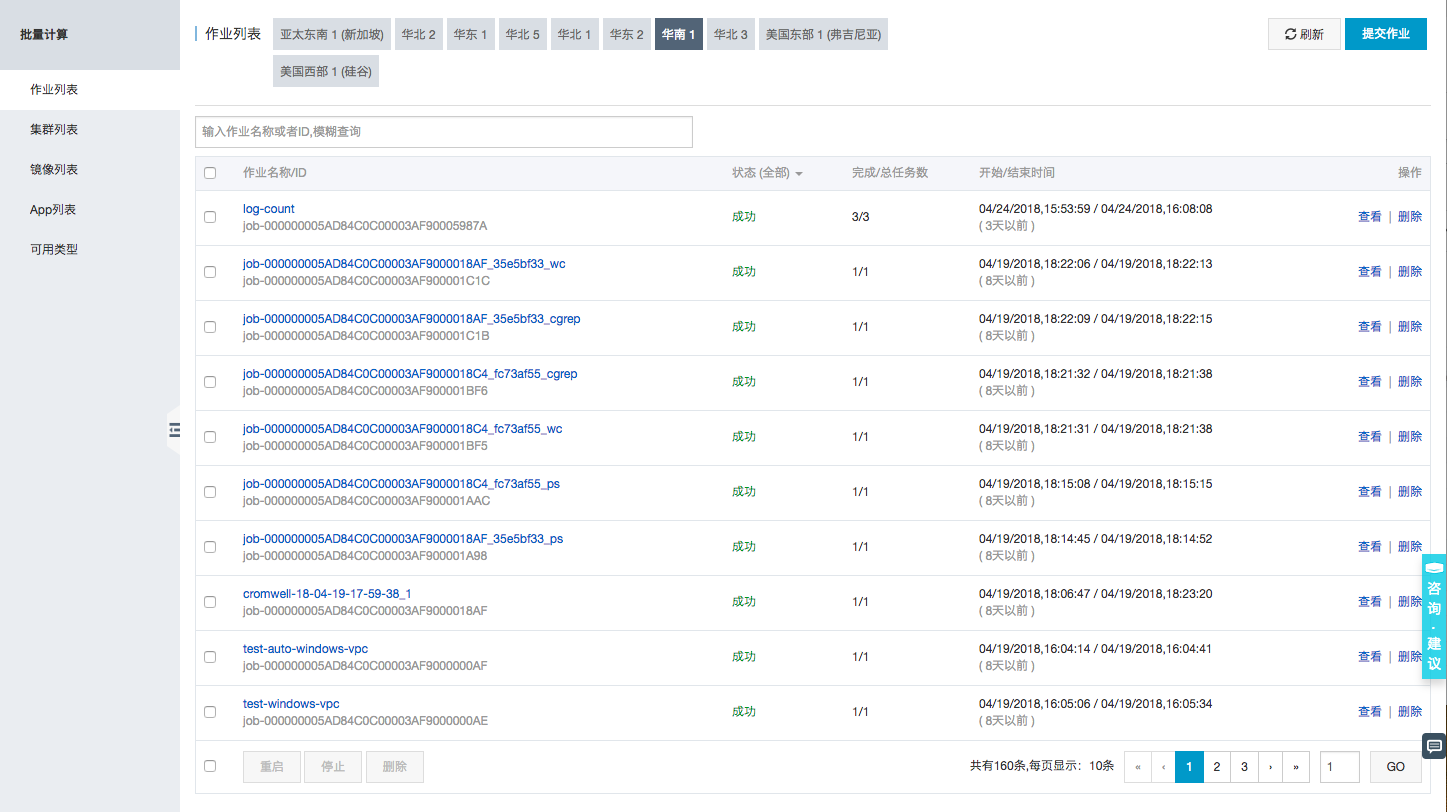
[](http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/assets/pic/42382/cn_zh/1468318787552/dag-tasks.png)

每个任务定义了使用哪个镜像，使用什么实例规格，运行哪个程序，需要多少台机器运行，还有结果存储在哪里等。

[](http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/assets/pic/42382/cn_zh/1469498741945/3.png)

## 2. 管理我的作业

您可以使用工具（控制台，命令行工具等），查看我提交的作业，可以停止，重启，或删除作业。 查看各个任务的情况，各个实例（VM 实例）的情况和日志。下图是控制台的作业管理界面:

[](http://docs-aliyun.cn-hangzhou.oss.aliyun-inc.com/assets/pic/27996/intl_zh/1524795921082/%E5%B1%8F%E5%B9%95%E5%BF%AB%E7%85%A7%202018-04-27%20%E4%B8%8A%E5%8D%8810.23.39.png)

## 3. 使用集群

由于每次运行程序前需要启动虚拟机，会占用一定的时间（一般几分钟左右），遇到忙时有可能申请不到资源（虚拟机被其他客户使用了），您提交作业后可能需要等待一段时间才能运行。

如果您想要提高运行效率，可以先创建好集群，指定需要的虚拟机数量（比如：5台）和镜像ID，BatchCompute 会为您分配好机器并启动，这些机器会一直处于运行状态，一旦您提交作业上来，就可以直接运行，效率较高。

## 4. 自定义镜像

您提交作业或者创建集群时，可以使用批量计算官方提供的镜像，也可以使用自定义镜像。自定义镜像的好处是，可以自己安装需要的软件。详情可以看 [这里](https://help.aliyun.com/document_detail/44571.html)

# 应用场景

更新时间：2018-10-26 19:53:10

[**本页目录**](javascript:void(0))

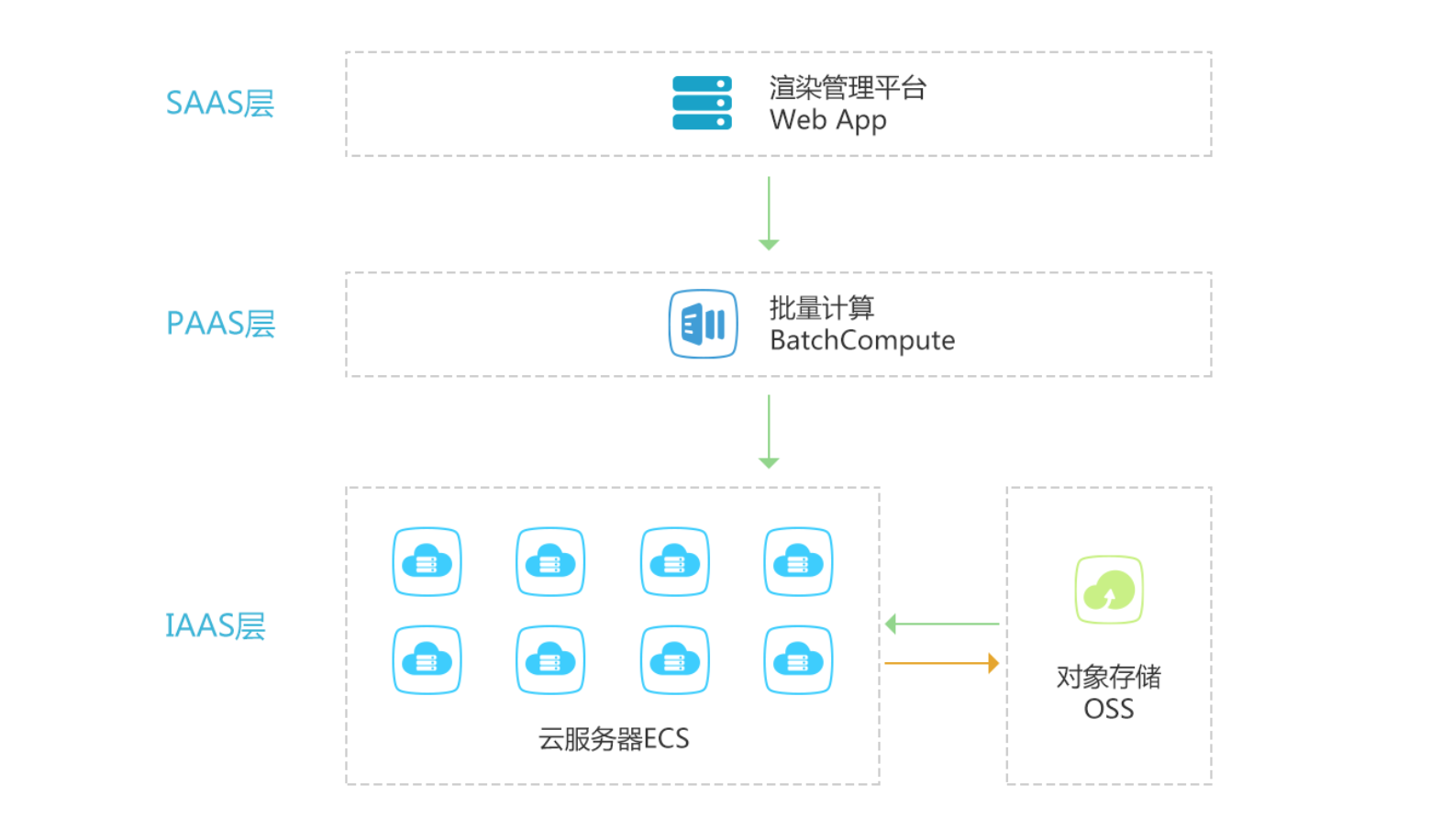
* [典型案例](https://help.aliyun.com/document_detail/27999.html?spm=a2c4g.11186623.6.547.4d657668kO2YR3#h2-u5178u578Bu6848u4F8B1)

BatchCompute 广泛应用于电影动画渲染、生物数据分析、多媒体转码、金融保险分析等领域。

## 典型案例

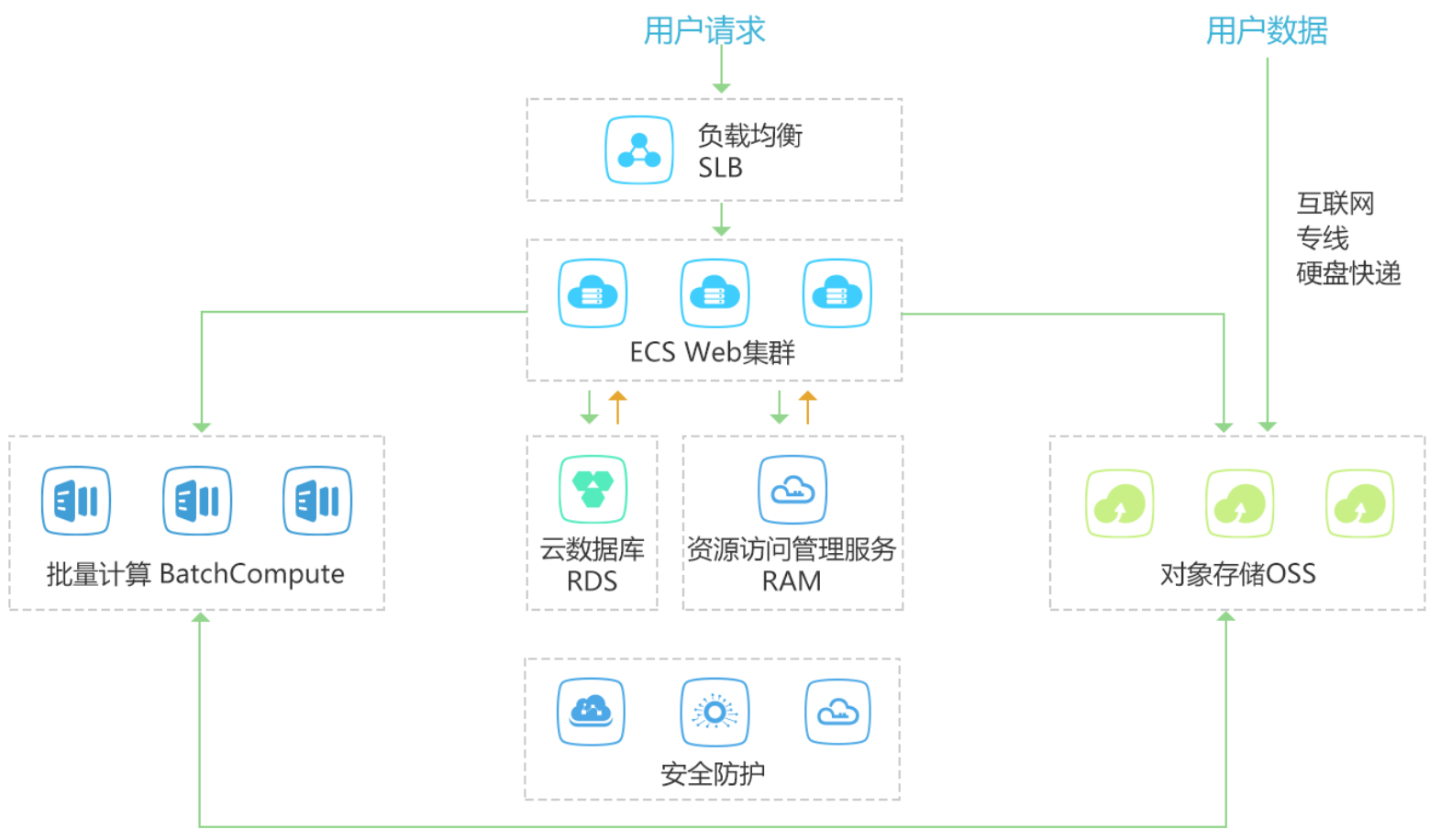
### 电影动画渲染

阿里云渲染云解决方案基于 BatchCompute 搭建。国产动画片如小门神、昆塔等，利用BatchCompute 调用阿里云日夜不停进行影片视效渲染，极大提高了渲染效率。

[](https://gtms01.alicdn.com/tps/i1/TB1HsFhLFXXXXcyXXXXfmp_MpXX-1572-916.png)

### 生物数据分析

生物基因企业利用 BatchCompute 完成大规模基因组测序分析。

[](https://gtms04.alicdn.com/tps/i4/TB1gPBgLFXXXXXnXpXXHKuULXXX-1570-916.png)