批量计算

产品简介

腾讯云

版权所有：腾讯云计算（北京）有限责任公司

【版权声明】

©2013-2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

【商标声明】

腾讯云

及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【服务声明】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【联系我们】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

文档目录

产品简介
- 什么是批量计算?
- 应用场景
- 产品优势
- 名词解释
- 配额限制
- 数据安全

版权所有：腾讯云计算（北京）有限责任公司

产品简介
什么是批量计算?

最近更新时间：2024-09-23 14:47:51

您可以通过以下视频了解什么是批量计算：

观看视频

批量计算（BatchCompute，简称 Batch）是应对海量批处理作业的分布式计算平台，目的是给有大数据计算业务的企业、科研单位等提供高性价比且易用的计算服务。Batch 会智能的对任意规模批处理作业进行资源分配和作业执行管理，您可以将精力集中在如何分析和处理数据上。Batch 非常适合基因测序、医疗分析、科学研究、视频渲染、金融大数据、精准广告等领域。

作为一项完全安全的托管服务，Batch 使开发人员、科学家和工程师能够运行任意规模的批处理计算作业。利用 Batch，您仅需要提供计算执行的环境、命令以及对应输入输出数据存放地址，Batch 会根据工作负载去获取腾讯云的弹性资源，并自动调度作业执行流程以完成成千上万的批处理作业。Batch 可以降低运营的复杂性、节省时间和减少成本，从而使开发人员、科学家和工程师能够在腾讯云中轻松运行其批处理作业，也使企业、科研机构等不再需要预支巨额的费用去建设批处理数据中心。

版权所有：腾讯云计算（北京）有限责任公司

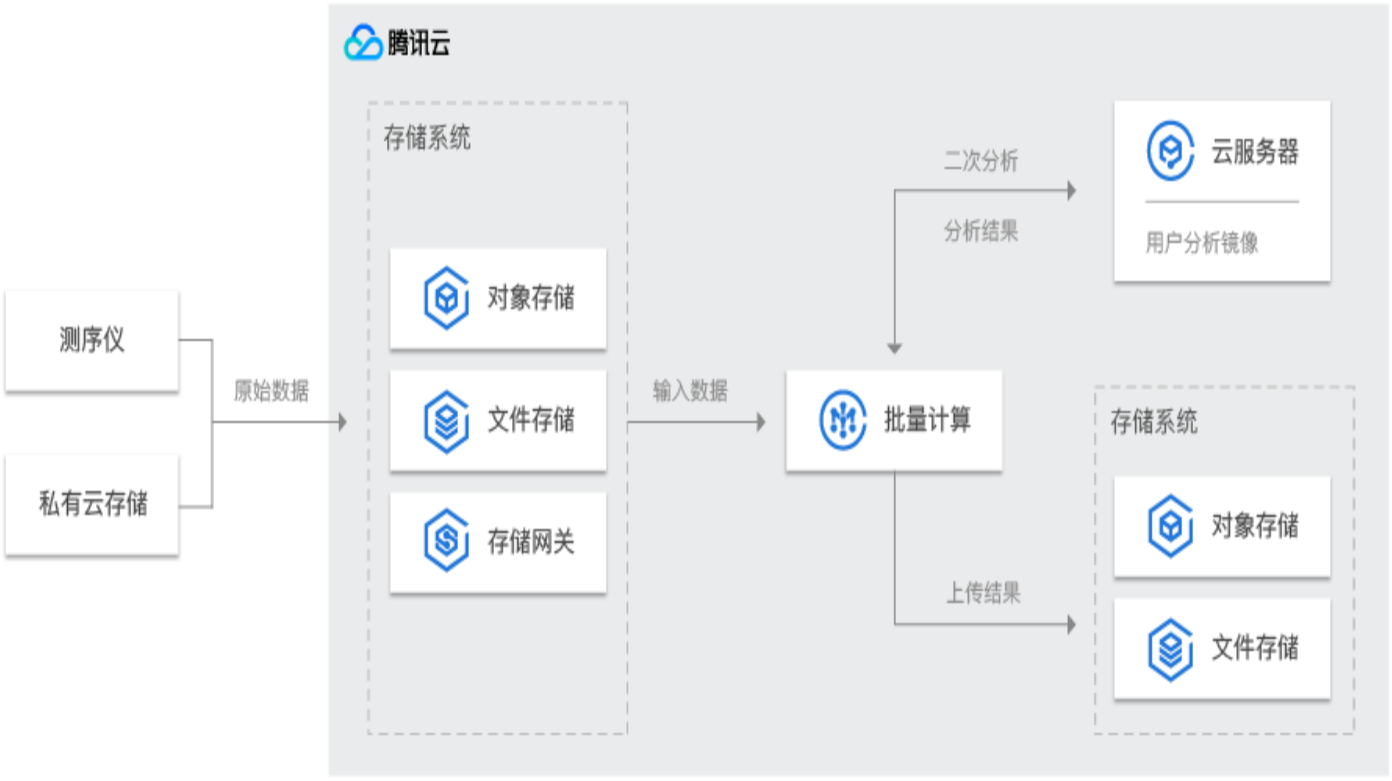
应用场景

最近更新时间：2024-09-23 14:47:51

以下视频介绍了批量计算的典型应用场景：
观看视频

基因测序

生物信息公司或实验室利用测序仪获取基因组序列的原始文件，在完成基因组序列的初步分析后将信息上传到云上存储系统，例如对象存储 COS，然后通过腾讯云批量计算 Batch 对信息进行更进一步分析。

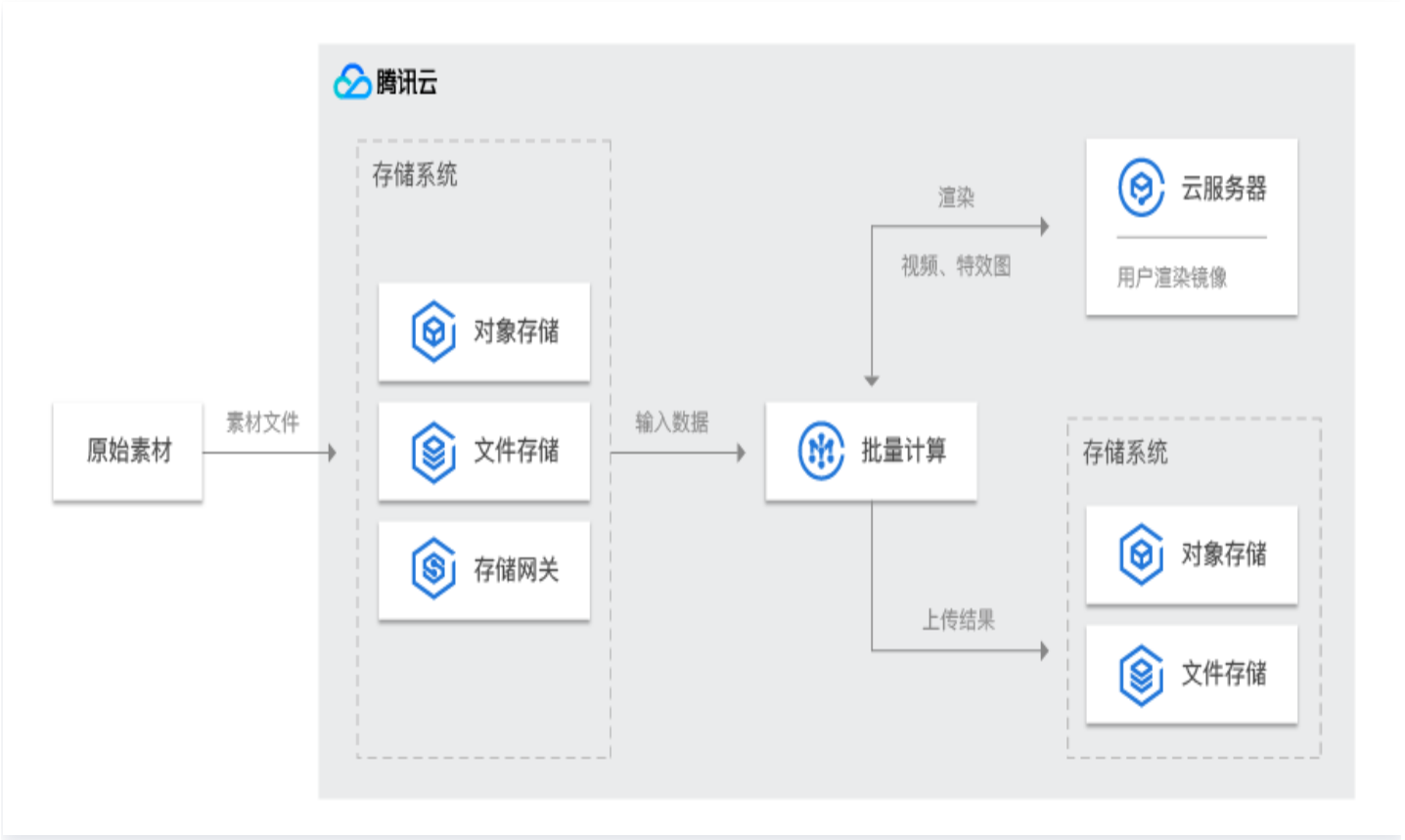


场景常见步骤

1. 生物信息专家从测序仪或私有云上的存储获取原始信息，并上传到腾讯云存储服务，例如 对象存储 COS 、文件存储 CFS 。
2. 用户定义用来分析信息的 Batch 作业，然后提交作业，作业的存储配置与前一步上传的原始信息关联。
3. Batch 自动调度资源，并将用户上传的自定义分析镜像部署到被调度的 CVM 里，同时自动调度作业开始分析原始信息。
4. CVM 上的计算完毕后，Batch 会自动将分析结果上传到用户指定位置。
5. 将待分析的原始信息上传到 COS。

影视和效果图渲染

在影视、广告、建筑规划等视觉创作行业中，内容制作者和后期制作公司需要使用到海量机器来完成影视特效、三维动画、特效图等相关渲染工作，腾讯云批量处理 Batch 为用户提供了自动化内容渲染工作流水线的能力，用户可以构建自己的渲染依赖流程，同时利用 Batch 的海量资源和作业调度能力来高效的完成视觉创作工作。



场景常见步骤

1. 用户准备好渲染需要的原始素材文件，并上传到腾讯云存储服务，例如 COS、CFS。
2. 用户定义用来分析信息的 Batch 作业，然后提交作业，作业的存储配置与前一步上传的原始信息关联。
3. Batch 自动调度资源，并将用户上传的自定义渲染镜像部署到被调度的 CVM 里，同时自动调度作业开始渲染视频或效果图。
4. Batch 在 CVM 上的渲染程序完成后，将生成的视频或效果图上传到用户指定位置。

版权所有：腾讯云计算（北京）有限责任公司

产品优势

最近更新时间：2024-09-24 17:26:21

以下视频比较了腾讯云批量计算 Batch 和自建计算集群，进一步突显了 Batch 的优势：

观看视频

优势	腾讯云批量计算 Batch	自建计算集群
低成本	**零预投入成本** 用户可以在需要时免费调用海量的云服务器资源，为 Batch 服务所用，仅需为您使用到的云服务器付费。同时，用户无需担心建设和维护基础设施，节约成本及时间。	**建设费用高昂** 自建基础设施，无论是购买还是长期租用，都有着高昂的成本，一个大规模计算集群建设成本最少也是以千万计。 同时计算集群的建设也是一个繁琐冗长的过程，其中的未知开销和风险将让您难以聚焦于业务本身。
弹性资源	**按需伸缩资源规模** 批量计算 Batch 完全按需使用，您可以在有计算作业时，通过 Batch 来获取您需要的资源。	**算力浪费或者算力不足** 无论您如何去精细规划自建资源集群的大小，在不可预测的业务需求面前，浪费或者不足最终是难以避免。资源闲置带来的成本审核风险，抑或是资源不足而影响到业务发展，都是自建集群时常要面对的问题。
智能调度	**完全托管** 批量计算 Batch 将批处理计算中最麻烦的资源调度和流程调度完全接管，用户只需要定义和提交计算作业，集中精力在分析结果和解决问题上。	**手动支持** 手动或者自行开发调度软件来处理资源分配和流程调度是非常繁琐和易错的事情，特别面对海量资源和计算作业时，很难预估将要浪费多少人力和资源。
强大易用	**丰富功能，简单使用** 批量计算 Batch 提供了大量辅助功能，帮助用户定义和提交计算作业，例如多种执行方法定义、计算流程编排、状态监控等，批处理任务里所有常用功能用户可以在 Batch 上找到。 同时还深度结合其他相关云服务，例如对象存储 COS。Batch 从计算材料获取、计算作业提交和进行、计算结果上传三个大环节上提供了一站式闭环服务。	**自行开发或第三方工具** 自行开发海量批处理系统工作量大且与用户主营业务无关，第三方工具大多缺乏及时的技术支持，需要用户自行摸索使用方法，这两种方法都要分散掉用户大量精力，而无法聚焦主营业务。

版权所有：腾讯云计算（北京）有限责任公司

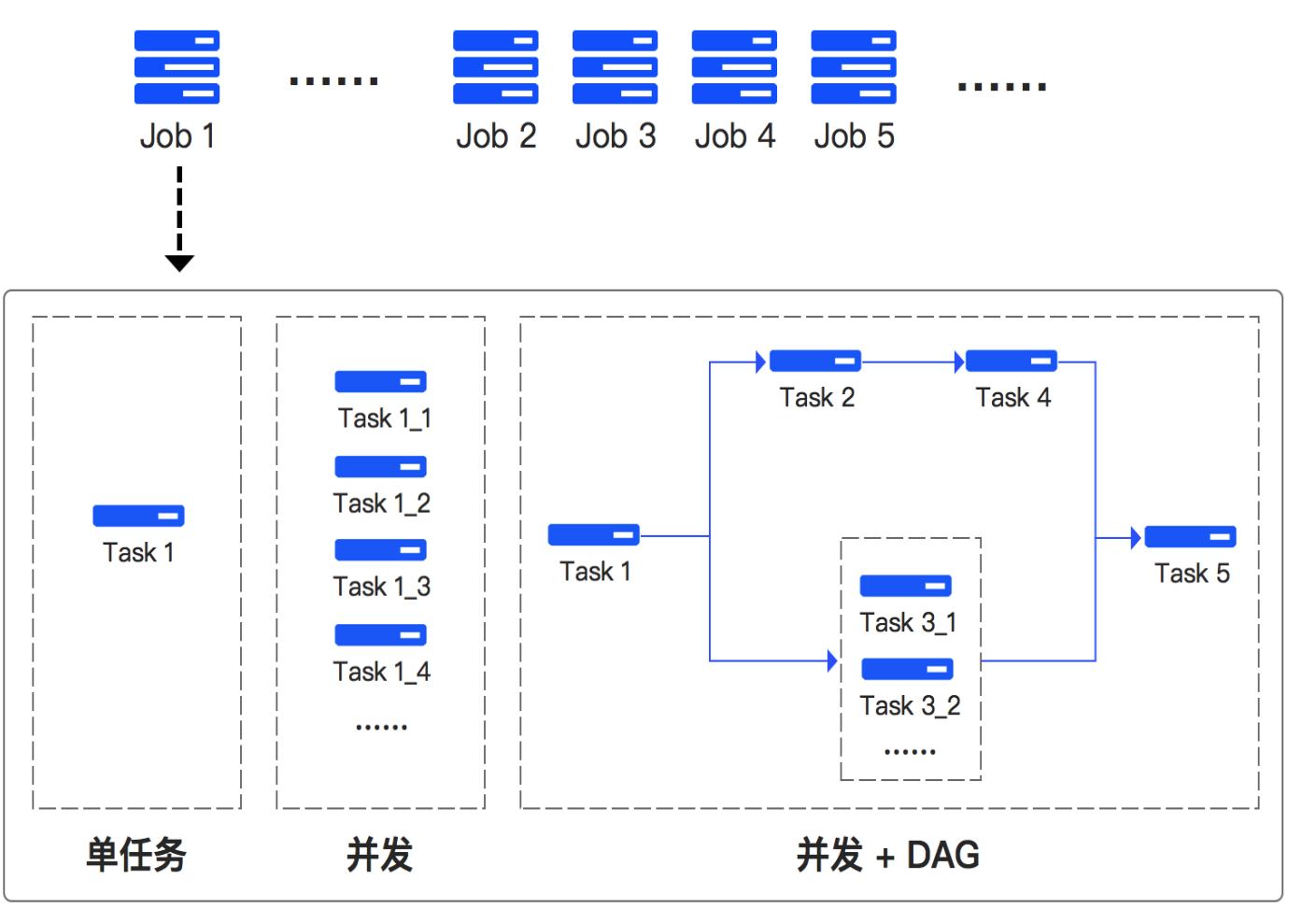
名词解释

最近更新时间：2024-09-24 17:26:21

您可通过以下视频了解批量计算的相关名词及简要工作流程：
观看视频

作业

作业（Job）是用户提交批处理工作的最小单位，它由单个或多个有前后依赖关系的任务组成。可以通过非常易用的 DAG 语法，来给多个批处理任务设置依赖关系，共同组成一个作业，然后依次执行各个任务，直到所有任务完成，即作业完成。任务间的依赖关系只能在作业提交时指定，提交完成后不能进行修改。



任务

任务（Task）是作业的基本组成单位，包含了实际在一台云服务器上执行的应用程序的相关信息。Batch 调度系统根据用户提交的配置，自动创建 CVM，安装镜像，执行程序。任务不可以直接提交，必须放入一个作业后才可以提交执行，作业支持放置单个任务，也可以放置多个任务。
任务最核心的可配置属性如下：

- `CVM 实例配置`：任务是执行在 CVM 上的，您需要根据您计算任务的特性，去配置 CVM 实例的类型和配置，例如选择计算型实例(C2)还是高IO型实例（IO2）、内存和磁盘大小、实例所处于的 VPC 网络等。
- `执行环境`：包含镜像和命令配置，镜像一般是自定义镜像，里面包含您的应用程序和它所依赖的环境，而命令则指定如何启动这些应用程序来进行计算。
- `远程存储映射`：支持将远程存储地址映射到本地文件系统地址，目前支持 COS，详细见后面单独名词解释。
- `标准输出`：您可以配置标准输出的映射地址（COS），应用程序里输出到 stdout 和 stderr 的信息将会在任务完成后被上传到对应地址，这样可以在任务完成后回溯计算过程。

任务模板

可以将常用的任务制作成任务模板，基于任务模板定制不同的任务，实现作业的快速提交。

远程存储映射

存储映射即将远程存储地址（对象存储 COS 或文件存储 CFS）映射到 CVM 的本地文件系统，这样可以操作本地文件系统的方式来读写远程存储。

任务实例

任务实例是 Batch 调度和执行的最小单元。一个任务包含单个或多个任务实例。每一个任务实例运行在一台 CVM 实例上，执行相应的计算任务。您可以在任务配置里设置需要并发执行的实例数目，Batch 会在任务被启动时以同样的配置调度指定数目的实例，您可以在执行程序里通过环境变量来区分这些实例。
多实例的典型运用场景是输入数据可切割然后并行处理的场景，这样可以充分的运用云上弹性资源的优势进行并行计算，提高工作效率。

镜像

镜像（Image）配置在任务里，创建实例时使用，它必须是一个标准的或者自定义的 CVM 镜像，您需要预先将计算环境和应用程序准备好并制作成自定义镜像，您可以在 镜像控制台 查看您的自定义镜像，如何制作自定义镜像请参见 创建自定义镜像 。

计算环境

计算环境（ComputeEnv）是由单台或多台 CVM 实例组成的计算集群，可根据业务需求扩容或缩容。当任务配置指定了计算环境，任务实例会调度到指定计算环境的节点上执行；当没有指定计算环境时，Batch 会创建 CVM 实例用于执行任务实例，并默认在任务实例完成后销毁 CVM 实例。

版权所有：腾讯云计算（北京）有限责任公司

配额限制

最近更新时间：2024-09-22 20:24:01

配额限制类型	描述
云服务器配额限制	Batch 在执行作业过程中可能会用到按量计费和竞价实例模式的云服务器（CVM），云服务器会对用户在不同可用区下可同一时间持有的**云服务器数量**做限制，详情请参见 购买按量计费云服务器实例限制 。
作业配额限制	单用户作业数量最大为1000。
任务模板配额限制	单用户任务模板数量最大为50。
计算环境配额限制	• 单用户计算环境数量最大为10。 • 单计算环境 CVM 计算节点数量最大为2000。 • 单计算环境黑石计算节点数量最大为500。

版权所有：腾讯云计算（北京）有限责任公司

数据安全

最近更新时间：2024-09-23 20:24:01

数据私密性

批量计算服务对计算环境、作业和任务模板中涉及到的敏感元数据进行加密，密文持久化存储在云数据库服务中。

数据知情权

使用批量计算服务过程中，经过计算产生的数据存放在云硬盘、消息队列、对象存储和文件存储等云存储产品中。批量计算服务提供算力和任务调度，不保留计算数据。

版权所有：腾讯云计算（北京）有限责任公司