



(12)发明专利申请

(10)申请公布号 CN 110753083 A
(43)申请公布日 2020.02.04

(21)申请号 201910840773.3

(22)申请日 2019.09.06

(71)申请人 江苏中云科技有限公司
地址 212000 江苏省镇江市新区丁卯经十
二路468号双子楼A座05-101室

(72)发明人 戴元顺

(74)专利代理机构 南京众联专利代理有限公司
32206
代理人 薛雨妍

(51)Int.Cl.
H04L 29/08(2006.01)

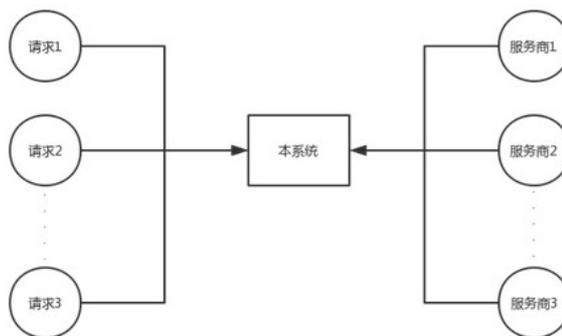
权利要求书1页 说明书3页 附图4页

(54)发明名称

一种面向多服务商的云服务资源统一分配系统

(57)摘要

本发明公开了一种面向多服务商的云服务资源统一分配系统,包括下列步骤:步骤1:收集用户请求阶段,该阶段会收到大量的用户请求;步骤2:调度阶段,该阶段已获得了最适合的服务商,将用户请求所需的服务通过服务分发模块分配给该服务商。步骤3:请求资源阶段,该阶段服务商收到请求,利用服务管理模块和数据设计模块,准备分配服务。步骤4:服务分配阶段,利用服务分发模块将该阶段服务商将服务提供给用户。本发明解决了服务商提供服务获得较大利益和保障用户服务质量问题,实现了二者的平衡,具备保障用户服务质量的前提下使得服务商的利益最大化特点。



1. 一种面向多服务商的云服务资源统一分配系统,其特征在于:包括下列步骤:
 - 步骤1:收集用户请求阶段,该阶段会收到大量的用户请求;
 - 步骤2:调度阶段,该阶段已获得了最适合的服务商,将用户请求所需的服务通过服务分发模块分配给该服务商;
 - 步骤3:请求资源阶段,该阶段服务商收到请求,利用服务管理模块和数据设计模块,准备分配服务;
 - 步骤4:服务分配阶段,利用服务分发模块将该阶段服务商将服务提供给用户。
2. 如权利要求1所述一种面向多服务商的云服务资源统一分配系统,其特征在于:所述步骤1包括以下子步骤:
 - 子步骤1-1:通过排队网络对用户请求进行划分,分到某个服务台;
 - 子步骤1-2:服务台根据用户请求对用户所能接受的最差服务质量进行量化,得到用户的服务质量保障的限制条件一;
 - 子步骤1-3:服务台根据服务商提供服务所需的成本和响应时间数据,得到一些高于用户最差服务质量的服务商群;
 - 子步骤:1-4:服务台获取服务商的资源利用率和资源过载限制;
 - 子步骤:1-5:根据上述子步骤1-2、子步骤1-3和子步骤1-4,使用多目标优化算法得到最适合的服务商。
3. 如权利要求1所述一种面向多服务商的云服务资源统一分配系统,其特征在于:监控模块同时实时监控请求收集、评估模块和请求处理模块。

一种面向多服务商的云服务资源统一分配系统

[0001]

技术领域

[0002] 本发明云平台下资源分配问题中为了保障用户的服务质量同时兼顾服务商收益问题而提出的一种资源分配系统。

背景技术

[0003] 目前各个云平台的服务商只为各自的用户提供服务,有时用户所需的服务对该服务商来说收益较小而为用户提供该服务的时候又需要占用较多资源,同时也会影响用户的服务体验。面对当前的大数据时代,云平台服务商每单位时间内分配出去的资源和使用云平台资源的客户每单位时间内提出的请求都是数量庞大的,面对这样的供需情况,需要统一调度和分配资源和请求。使得服务商和客户达到共赢。

[0004] 传统方法即单个服务商针对用户所需要的服务进行资源调度,但不足的是对于单个服务商来说,使用户得到较好的服务和缩减成本投入获取利益永远是矛盾的。

[0005] 中国联合网络通信集团有限公司,《一种资源服务系统及其资源分配方法》于2014-08-18日向中国国家知识产权局申请专利并获得批准,于2014-11-26日公开,公开号:104168318A。上述发明只在单个服务商内考虑资源分配,对用户请求划分优先级,不能给所有用户提供基本的服务质量。

[0006] 北京搜狐新媒体信息技术有限公司,王帅发明的《一种基于IaaS云平台的资源调度方法和装置》于2013-09-16日向中国国家知识产权局申请专利并获得批准,于2013-12-18日公开,公开号:103458052A,该发明只是单个服务提供商下的资源分配方案。并且可能导致该服务商的利益不能最大化。

[0007] 本发明要解决的技术问题在于:

1、针对相关方案一采用对用户请求划分优先级而使得部分低优先级用户服务质量变差问题,以及不能提供最低用户服务质量保障问题;

2、针对相关方案二采用尽力分配每个时间段内的资源而不考虑分配资源的成本投入和利润获得问题,以及不能使服务商获得最大利益问题。

发明内容

[0008] 针对现有技术的不足,本发明提供一种面向多服务商的云服务资源统一分配系统,包括下列步骤:

步骤1:收集用户请求阶段,该阶段会收到大量的用户请求;

步骤2:调度阶段,该阶段已获得了最适合的服务商,将用户请求所需的服务通过服务分发模块分配给该服务商。

[0009] 步骤3:请求资源阶段,该阶段服务商收到请求,利用服务管理模块和数据设计模块,准备分配服务。

[0010] 步骤4:服务分配阶段,利用服务分发模块将该阶段服务商将服务提供给用户。

[0011] 进一步改进在于:所述步骤1包括以下子步骤:

子步骤1-1:通过排队网络对用户请求进行划分,分到某个服务台。

[0012] 子步骤1-2:服务台根据用户请求对用户所能接受的最差服务质量进行量化,得到用户的服务质量保障的限制条件一。

[0013] 子步骤1-3:服务台根据服务商提供服务所需的成本和响应时间数据,得到一些高于用户最差服务质量的服务商群。

[0014] 子步骤:1-4:服务台获取服务商的资源利用率和资源过载限制;

子步骤:1-5:根据上述子步骤1-2、子步骤1-3和子步骤1-4,使用多目标优化算法得到最适合的服务商。

[0015] 进一步改进在于:监控模块同时实时监控请求收集、评估模块和请求处理模块。

[0016] 本发明是多用户和多服务商之间的中间平台,用户端方面接收大量的请求,服务商方面登记多个服务商能够提供的服务,每单位时间内都有大量的用户请求,本系统每单位时间内都接收这些请求,为了降低用户等待的响应时间,使用排队网络处理这些请求,排队网络中的服务台使用多目标优化算法计算当前请求在保证用户QoS的情况下需要的最差服务标准,接下来在已登记的资源簇里根据服务商的资源利用率计算最优的服务商,最后将当前请求转发给服务商,由服务商直接为用户提供服务。从单个服务商的角度来看,每个单位时间内提供的资源都一定能得到利润,相比较之前的单个服务商为所有用户请求提供资源,少了那些需要付出高成本低利润或者“入不敷出”的请求,多了能有更高利润的请求或将原本闲置的资源出售出去,这些请求无论是前者还是后者,都为单个服务商增加了利润。因此,本系统对多个服务商和用户来说是双赢的。

附图说明

[0017] 图1是本发明的结构图;

图2是使用本发明方法的模块图;

图3是本发明的服务提供流程图;

图4是本发明处理用户请求的排队网络流程图;

图5是用本发明后从单个用户请求角度来看的流程图。

具体实施方式

[0018] 下面将结合附图就本发明的发明目的、技术方案、发明优点作进一步详细说明。

[0019] 如图1-5,本实施例提供一种面向多服务商的云服务资源统一分配系统,包括下列步骤:

步骤1:收集用户请求阶段,该阶段会收到大量的用户请求;

子步骤1-1:通过排队网络对用户请求进行划分,分到某个服务台。

[0020] 子步骤1-2:服务台根据用户请求对用户所能接受的最差服务质量进行量化,得到用户的服务质量保障的限制条件一。

[0021] 子步骤1-3:服务台根据服务商提供服务所需的成本和响应时间数据,得到一些高于用户最差服务质量的服务商群。

[0022] 子步骤:1-4:服务台获取服务商的资源利用率和资源过载限制;

子步骤:1-5:根据上述子步骤1-2、子步骤1-3和子步骤1-4,使用多目标优化算法得到最适合的服务商。

[0023] 步骤2:调度阶段,该阶段已获得了最适合的服务商,将用户请求所需的服务通过服务分发模块分配给该服务商。

[0024] 步骤3:请求资源阶段,该阶段服务商收到请求,利用服务管理模块和数据设计模块,准备分配服务。

[0025] 步骤4:服务分配阶段,利用服务分发模块将该阶段服务商将服务提供给用户。

[0026] 其中监控模块同时实时监控请求收集、评估模块和请求处理模块。

[0027] 本实施例解决了服务商提供服务获得较大利益和保障用户服务质量问题,实现了二者的平衡,具备保障用户服务质量的前提下使得服务商的利益最大化特点。

[0028] 以上仅是本发明的优选实施方式,应当指出,以上实施列对本发明不构成限定,相关工作人员在不偏离本发明技术思想的范围内,所进行的多样变化和修改,均落在本发明的保护范围内。

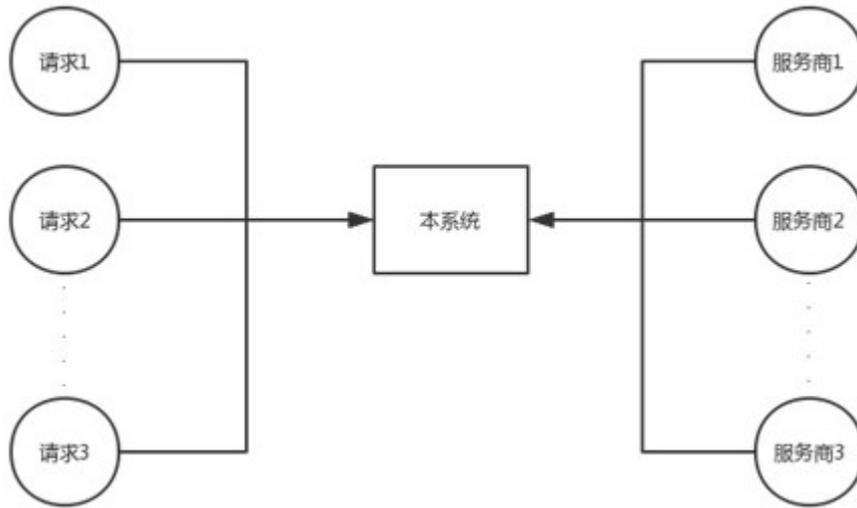


图1

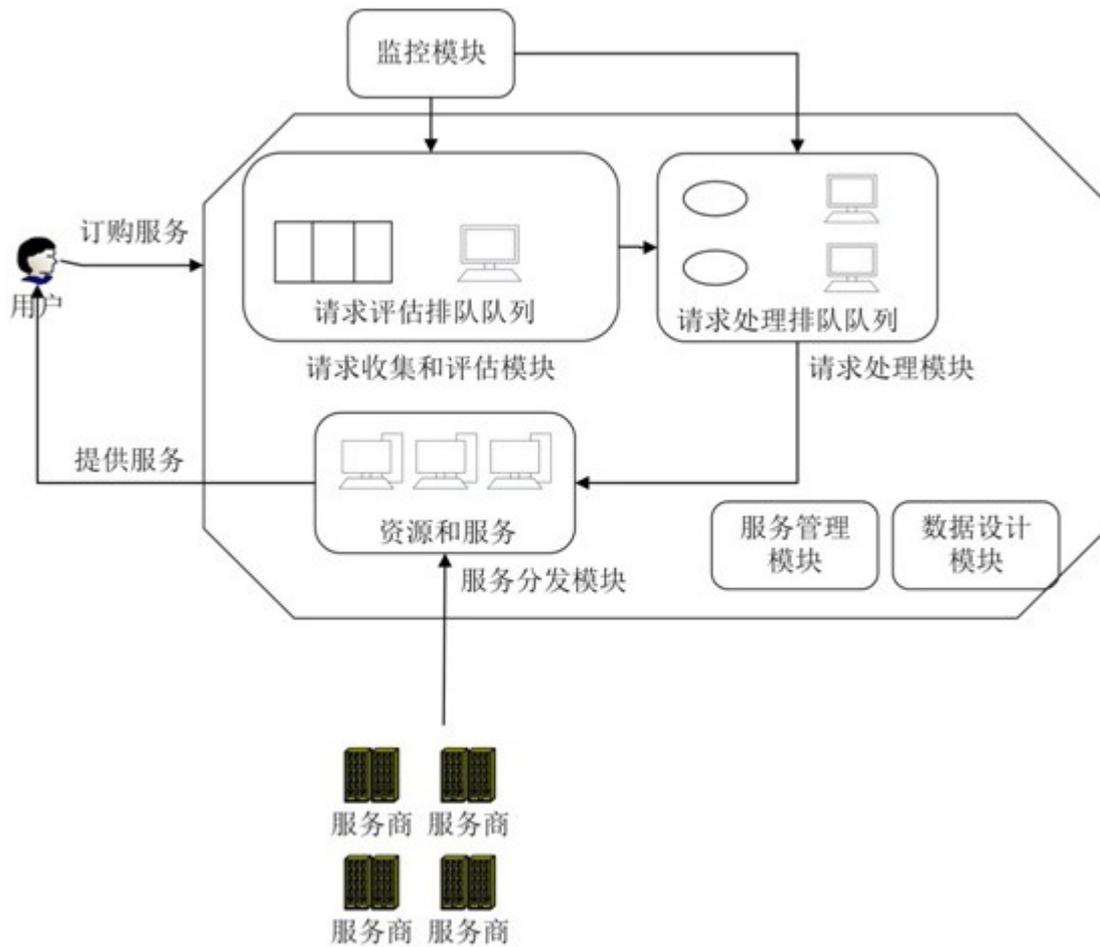


图2

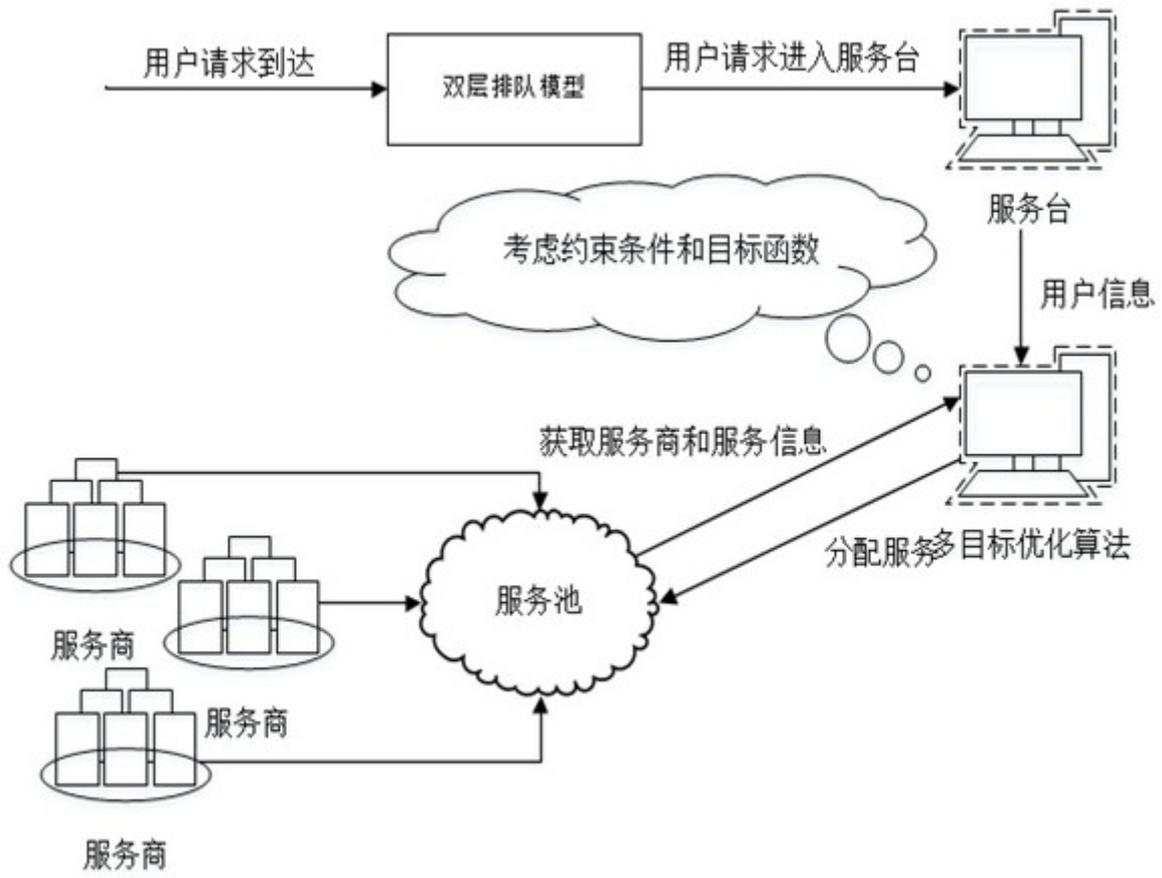


图3

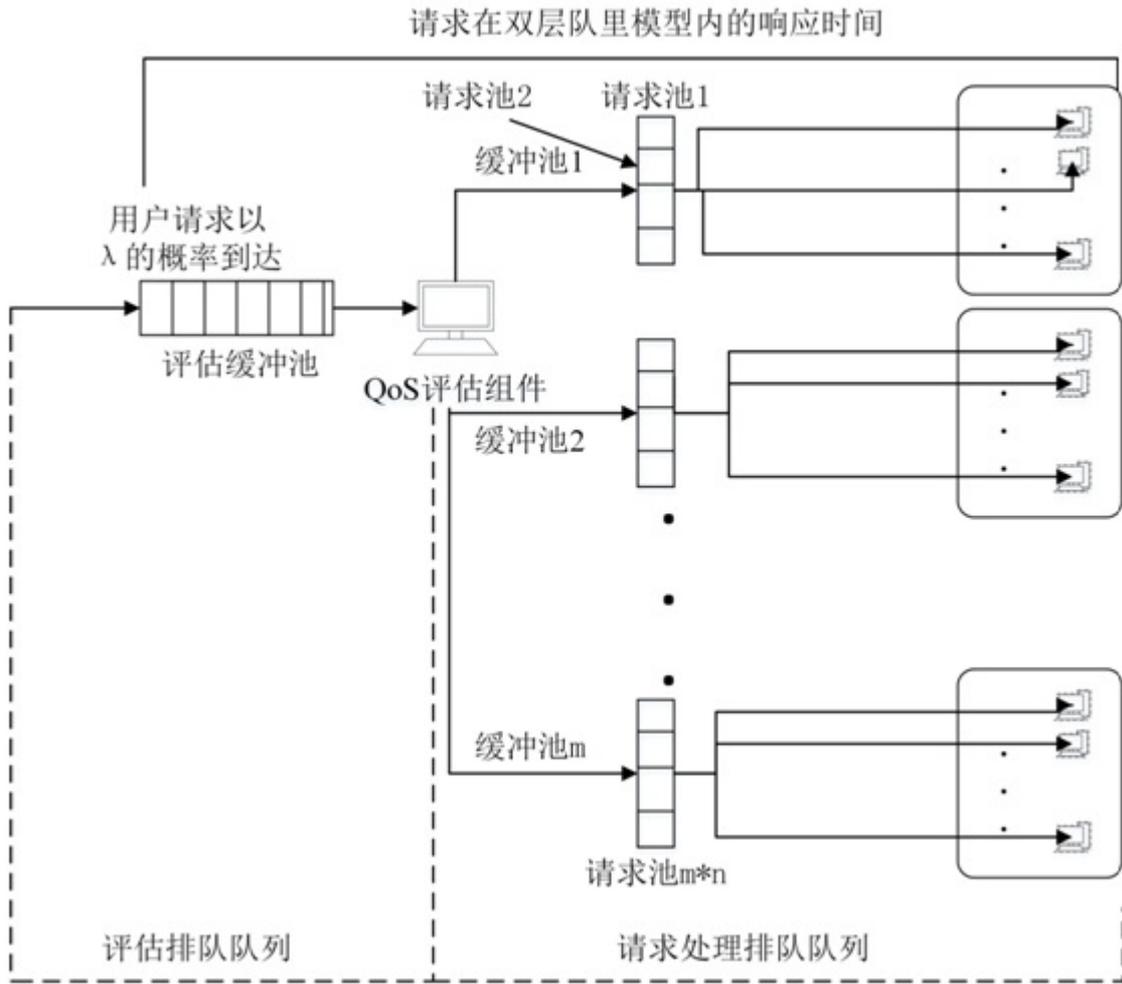


图4

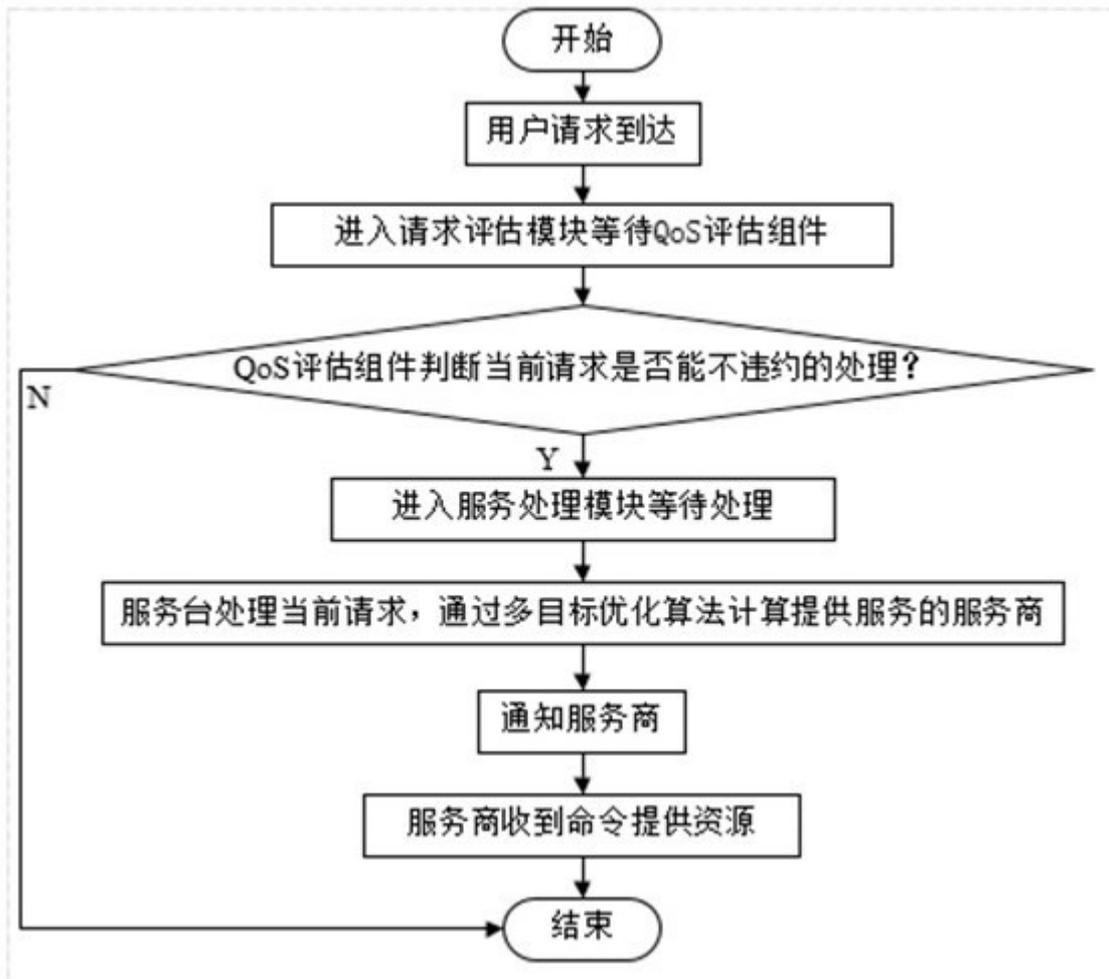


图5