版本号：V0.1

文档标识：YY-GYY-XQGG-WLYJK

易云公有云网络服务V1.3

需求规格说明书

修订记录

|  |  |  |  |
| --- | --- | --- | --- |
| 版本 | 时间 | 备注 | 编辑者 |
| 0.1 | 2016-9-28 | 初稿 | 张亚琪 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**[1.文档概述](#_Toc13656)** [2](#_Toc13656)

**[1.1. 概述](#_Toc20271)** [2](#_Toc20271)

**[1.2. 背景及目的](#_Toc1888)** [2](#_Toc1888)

**[2. 功能概述](#_Toc27035)** [2](#_Toc27035)

**[2.1. 统一术语](#_Toc3763)** [2](#_Toc3763)

**[2.2. 定义](#_Toc14211)** [2](#_Toc14211)

**[2.3. 功能列表](#_Toc17020)** [3](#_Toc17020)

**[3. 限制云主机最大带宽](#_Toc26693)** [4](#_Toc26693)

**[4. 负载均衡会话保持](#_Toc24058)** [4](#_Toc24058)

**[5. 监控](#_Toc22042)** [5](#_Toc22042)

**[5.1. 流量监控](#_Toc30386)** [5](#_Toc30386)

**[5.2. 流速监控](#_Toc31731)** [6](#_Toc31731)

**[5.3. 负载均衡监控](#_Toc4993)** [7](#_Toc4993)

**[5.3. 取值说明](#_Toc12779)** [8](#_Toc12779)

**[6. 性能要求](#_Toc25654)** [8](#_Toc25654)

# **1.文档概述**

## **1.1. 概述**

本文描述了公有云网络 QOS，网络监控，负载均衡监控以及负载均衡的会话保持的功能需求。本文将重点描述 stack部分网络以及负载均衡的功能及各项功能的详细描述。

**1.2. 背景及目的**

目前，网络1.2 版本已经实现了路由对外端口的速度限制，用户可以为每个私有网络设置最大带宽。为了应对用户提出的限制各个云主机上行流量的需求以及通过对大部分云厂商（阿里云，腾讯云，华为云，青云，云英）调研结果的分析，决定于网络1.3版本新增允许用户为云主机设置上行最大带宽且云主机与云主机之间独享最大带宽的需求。即用户可以在申请云主机的时候，自行设置出网最大带宽（最大带宽：在网络空闲的时候，可以充分的使用空闲的带宽，但是不能超过设置的最大带宽），目前带宽的可选范围是 [1, 100]（单位：Mbps）。同时，该功能还可以让用户更加明确云主机的计费情况。

网络1.2版本中已经实现了负载均衡功能，其作用就把不同的客户端的请求通过负载均衡策略分配到不同的云主机上去。但在某些需要保持登录状态的情况下，会要求客户端与云主机之间保持一个会话（session）以记录客户端的各种信息。例如，用户在云主机上部署了电商相关的Web应用，首先客户端在登录的时候，负载均衡服务会根据用户所选择的策略将该请求发送到云主机A上，然后在很短的时间内客户端又发送了一个请求，在负载均衡没有会话保持功能的时候，第二次的请求很有可能会分配到云主机B上，而客户端并没有在云主机B上登录过，所以客户端只能重新登录，可见购物者（客户端）的用户体验是非常不友好的。除了登录时会遇到此种问题，购物者购物的过程也会受到影响，例如刚刚放到购物车中的物品不见了等等一系列问题。那么如何得知这两次不同的请求其实是源于同一个客户端呢？这便是会话保持的价值所在了，会话保持可以保证同一会话的请求不会交付给不同的后端云主机。

为了能够满足用户想要通过全面掌控系统运行情况和资源使用状况进而方便定位分析问题的需求，遂在网络1.3版本中新增网络流速、网络流量、负载均衡请求数、连接数、平均响应时间、不同状态响应数、各状态码响应总数以及客户端访问负载均衡所消耗的流量等监控指标（作为监控的变量，数据点代表该变量随时间变化的值）。用户可以通过观察图表的方式获取到以上各项指标的监控数据，并能够根据自身情况自定义高警阈值。

# **2. 功能概述**

**2.1.** **统一术语**

|  |  |
| --- | --- |
| **术语** | **解释** |
| 外网 | 非 Neutron 管理的网络 |
| 内网 | Neutron 管理的网络 |
| 路由器外网接口 | 通向外部网络的接口 |
| IDC上行流量 | 云主机向客户机传输信息的多少叫IDC上行流量 |
| IDC下行流量 | 云主机从客户机收到信息的多少流量叫IDC上行流量 |
| 公网网卡 | 连接着受管子网的实例网卡 |
| 内网网卡 | 连接着自管子网的实例网卡 |

**2.2. 定义**

1. **Qo****S：**是一种控制机制，它提供了针对不同用户或者不同数据流采用相应不同的优先级，或者是根据应用程序的要求，保证数据流的性能达到一定的水准。QoS的保证对于容量有限的网络来说是十分重要的。
2. **负载均衡：**是对多台云主机进行流量分发的服务。负载均衡可以通过流量分发扩展应用系统对外的服务能力，通过消除单点故障提升应用系统的可用性。
3. **会话保持：**会话保持是负载均衡最常见的问题之一，也是一个相对比较复杂的问题。会话保持有时候又叫做粘滞会话(Sticky Sessions)。会话保持是指在负载均衡器上的一种机制，可以识别客户端与服务器之间交互过程的关连性，在作负载均衡的同时还保证一系列相关连的访问请求会保持分配到一台服务器上｡
4. **Cookie：**总是保存在客户端中，按在客户端中的存储位置，可分为内存Cookie和硬盘Cookie，这里我们主要用到的是内存 Cookie。内存Cookie由浏览器维护，保存在内存中，浏览器关闭后就消失了，其存在时间是短暂的。也可以看作一个用来标识每一次请求的状态位。
5. **Connection：**客户端到服务器端建立的一条物理路径。HTTP持久连接是使用同一个TCP连接来发送和接收多个HTTP请求/响应，而不是为每一个新的请求/响应打开新的连接的方法。在 HTTP 1.1 中 所有的连接默认都是持续连接，除非特殊声明不支持。HTTP 持久连接不使用独立的 keepalive 信息，而是仅仅允许多个请求使用单个连接。然而， Apache 2.0 httpd 的默认连接过期时间是仅仅15秒，对于 Apache 2.2 只有5秒。短的过期时间的优点是能够快速的传输多个web页组件，而不会绑定多个服务器进程或线程太长时间。
6. **Session：**会话（session）是一种持久网络协议，在用户（或用户代理）端和服务器端之间创建关联，从而起到交换数据包的作用机制，session在网络协议（例如telnet或FTP）中是非常重要的部分。在不包含會話层（例如UDP）或者是无法长时间驻留会话层（例如HTTP）的传输协议中，會話的维持需要依靠在传输数据中的高级别程序。例如，在浏览器和远程主机之间的HTTP传输中，HTTP cookie就会被用来包含一些相关的信息，例如session ID，参数和权限信息等。

**2.3. 功能列表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 序号 | 功能模块 | 功能点 | | 描述 | 优先级 | 备注 |
| 1 | VPC | 限速 | | 为云主机设置上行最大带宽 | P0 |  |
| 2 | 为负载均衡器设置上行最大带宽 | P0 |
| 3 | 监控 | 流量 | 外网 → 公网网卡 | P0 |  |
| 4 | 公网网卡 → 外网 | P0 |
| 5 | 内网 → 内网网卡 | P1 |
| 6 | 内网网卡 → 内网 | P1 |
| 7 | 内网 → 公网网卡 | P2 |
| 8 | 公网网卡 → 内网 | P2 |  |
| 9 | 公网 → 路由器外网接口 | P0 |  |
| 10 | 路由器外网接口 → 公网 | P0 |  |
| 11 | 公网 → 负载均衡 | P0 |  |
| 12 | 负载均衡 → 公网 | P0 |  |
| 13 | 流速 | 外网 → 公网网卡 | P0 |  |
| 14 | 公网网卡 → 外网 | P0 |  |
| 15 | 内网 → 内网网卡 | P1 |  |
| 16 | 内网网卡 → 内网 | P1 |  |
| 17 | 内网 → 公网网卡 | P2 |  |
| 18 | 公网网卡 → 内网 | P2 |  |
| 19 | 负载均衡 | 会话保持 | | 保证用户相同关联的一组请求可以发送到相同负载均衡后端服务器中。 | P0 |  |
| 20 | 监控 | HTTP | 请求数 | P0 |  |
| 21 | 没开启会话保持：用于 HTTP 请求的 TCP 连接数（connction） | P0 |  |
| 22 | 开启会话保持：session 数（同一个浏览器进程发出的所有连接被看做一个session） | P0 |  |
| 23 | 后端服务器返回给负载均衡器响应码（1xx,2xx,3xx,4xx,5xx）的数量 | P1 |  |
| 24 | 负载均衡器返回个客户端的响应码（4xx,5xx）的数量 | P1 |  |
| 25 | member平均响应延迟时间 | P0 |  |
| 26 | TCP | 没开启会话保持：连接数（connction） | P0 |  |
| 27 | 开启会话保持：session 数（同一个IP发出的所有连接都看成是一个session） | P0 |  |

# **3. 限制云主机最大带宽**

|  |  |  |  |
| --- | --- | --- | --- |
| **序号** | **功能点** | **描述** | **备注** |
| 1 | 设置从云主机或负载均衡上行最大带宽 | 1. 允许设置的最大带宽范围为：[A, B]（单位：Mbps)，下面的两个条件必须同时成立。 2. A < B 3. A 和 B 都是正整数 4. 仅对从云主机流向外网这个方向进行带宽限制，不对网卡进行限速，因为如果云主机的某一个网卡分别有流向内网和外网的流量，注意不对流向内网的方向进行限速。 5. 不设置下行最大带宽。 6. 并不针对云主机的某一服务进行带宽限制。 |  |

# **4. 负载均衡会话保持**

| **序号** | **功能点** | **描述** | **备注** |
| --- | --- | --- | --- |
| 1 | 基于IP的会话保持  （四层会话保持） | 四层TCP方式的负载均衡是基于客户IP进行会话保持的，它会把同一IP地址的请求持续发往一台服务器，如果客户端的网络环境发生变化，比如从有线切成无线，4G网络切成无线，会因为IP地址的变化造成会话没有保持。 |  |
| 2 | 基于cookie的会话保持（七层会话保持） | 1. 第一次请求，客户端 http request 进入负载均衡器。 2. 根据负载均衡策略选择一台后端云主机，然后将请求发送给该云主机，云主机生成响应后，需要在响应头部加上 cookie 信息（添加 Set-Cookie），并在 cookie 中设置它在客户端中的存活时间（expire，用户可以自己选择存活时间的值：A，一天为上限，即 A 的范围是 [1,86400]（单位： s））。这样一来客户端浏览器设置中就包含服务器信息的特定 cookie了。 3. 客户端收到响应后，会将 cookie 信息写入本地。 4. 再次请求的时候，携带着特定 cookie 信息的 http request发到指定的服务器。 5. 后端服务器从请求头中获取到 cookie 信息，分析 cookie，如果对应的 cookie 是正确的就生成响应，并将响应发回给客户端。 |  |

# **5. 监控**

**5.1. 流量监控**

| **序号** | **功能点** | **描述** | **API必需参数** | **备注**  **(监控粒度 & 公式)** |
| --- | --- | --- | --- | --- |
| 1 | 统计从外网流入某个云主机的流量，用于监控 | 统计从外网流向某一云主机所有公网网卡（目前最多只有一块公网网卡）流量的总和，用于查看下行流量，即云主机下载流量的总和。 | 1. Instance 2. 监控粒度 3. 流出端 4. 流入端   注意：流入端和流出端需要具体到云主机的某块网卡或者是路由器的某个接口 | 监控粒度分别为：1min、2min、15min  [流量计算公式](#_2.2.4. 一些说明) |
| 2 | 统计从云主机公网网卡流向外网的流量，用于监控、计费 | 统计从云主机每个公网网卡（目前一个云主机最多只能有一块公网网卡）流向外网的流量，为选择按需付费使用云主机的用户提供精确流量消耗的明细。 |
| 3 | 统计从内网流向某云主机内网网卡的流量，用于监控 | 统计从内网流向云主机每块内网网卡（目前每个云主机最多只能有一块内网网卡）的流量，便于用户监控东西向流量。 |
| 4 | 统计从某云主机内网网卡内流向内网的流量，用于监控 | 统计从云主机每个内网网卡（目前每个云主机最多只能有一块内网网卡）的流向内网的流量，便于用户监控东西向流量。 |
| 5 | 统计从内网流向某云主机公网网卡的流量，用于监控 | 统计从内网流向某一云主机每块公网网卡（目前云主机最多只能有一块公网网卡）的流量，便于用户监控东西向流量。 |
| 6 | 统计从某云主机公网网卡到内网的流量，用于监控 | 统计从某一云主机的每块公网网卡（目前云主机最多只能有一块公网网卡）流向内网的流量，便于用户监控东西向流量。 |
| 7 | 统计从某路由器外网接口流入内网的流量，用于监控 | 统计外网流入内网实例的流量的外加流入负载均衡器流量的总和，便于运维人员查看用户各私有网络的下行流量。 |
| 8 | 统计从内网流向路由器外网接口的流量，用于监控、计费 | 统计由实例和负载均衡器流向路由器外网接口的流量的总和，便于运维人员查看用户各私有网络的上行流量。 |

**5.2. 流速监控**

| **序号** | **功能点** | **描述** | **API必需参数** | **备注**  **(监控粒度 & 公式)** |
| --- | --- | --- | --- | --- |
| 1 | 统计由外网流向云主机公网网卡的平均流速，用于监控 | 统计从外网流入每个云主机的每块公网网卡（目前云主机最多只能有一块公网网卡）的平均流速，便于用户查看某一时间段内每个云主机的下行平均流速。 | 1. Instance 2. 监控粒度 3. 流出端 4. 流入端   注意：流出端流入端需要具体到云主机的某块网卡或者是路由器的某个接口 | [平均流速计算公式](#_2.2.4. 一些说明" \o "平均流速) |
| 2 | 统计云主机公网网卡流向外网的平均流速，用于监控 | 统计从每个云主机每块公网网卡（目前云主机最多只能有一块公网网卡）流向外网的平均流速，便于用户查看某一时间段内每个云主机的东西方向流速。 |
| 3 | 统计从内网流向云主机内网网卡的平均流速，用于监控 | 统计由内网流向每个云主机每块内网网卡（目前云主机最多只能有一块内网网卡）的平均流速，便于用户查看某一时间段内每个云主机东西向的平均流速。 |
| 4 | 统计从云主机的内网网卡流向内网的平均流速，用于监控 | 统计从每个云主机每块内网网卡（目前云主机最多只能有一块内网网卡）流向内网的平均流速，便于用户查看某一时间段内每个云主机东西向的平均流速。 |
| 5 | 统计从内网流向云主机公网网卡的平均流速，用于监控 | 统计由内网流向每个云主机每块公网网卡（目前云主机最多只能有一块公网网卡）的平均流速，便于用户查看某一时间段内每个云主机东西向的平均流速。 |
| 6 | 统计从云主机的公网网卡流向内网的平均流速，用于监控 | 统计从每个云主机每块公网网卡（目前云主机最多只能有一块公网网卡）流向内网的平均流速，便于用户查看某一时间段内每个云主机东西向的平均流速。 |

## **5.3. 负载均衡监控**

| **序号** | **功能点** | **描述** | **(监控粒度 & 取值说明)** |
| --- | --- | --- | --- |
| 1 | 流入流量 | 统计从外网流向负载均衡器所消耗的流量 | [流量取值说明](#_5.3. 取值说明) |
| 2 | 流出流量 | 表示从负载均衡器流向外网所消耗的流量 |
| 3 | 请求数 | 在一个时间粒度内，统计负载均衡（HTTP 协议）接收到的来自客户端的请求总数。（累计值） | [请求数取值说明](#_2.2.4. 一些说明) |
| 4 | 用于 HTTP 请求的 TCP 连接总数 | 未开启会话保持的情况下，统计客户端经由负载均衡器向云主机发起请求并建立 HTTP 连接的总数。注意这里的连接指的是 Connection，即负载均衡处理的最小单位。 | [连接数取值说明](#_5.3. 取值说明) |
| 5 | 经过负载均衡（HTTP 协议） session 总数 | 开启会话保持的情况下，统计某时间段内 session 总数量。开启会话保持的情况下，一个浏览器进程发出的所有连接被认为是一个 session。 | [session数取值说明](#_5.3. 取值说明) |
| 6 | 各 member（ip + port） 平均响应延迟时间 | member响应时间指负载均衡（HTTP 协议）向后端云主机发出请求到云主机返回响应给负载均衡服务器所用的时间 |  |
| 7 | 后端云主机返回给负载均衡服务的1xx/2xx/3xx/4xx/5xx 状态码的响应数 | 统计一段时间内 HTTP 协议的监听器下所有后端返回的对应状态码的响应数。  时刻 B 的某状态码响应数含义：从时刻 A 到时刻 B 的某状态码响应数 | [响应数取值说明](#_5.3. 取值说明) |
| 8 | 由负载均衡直接返回给客户端的 4xx/5xx 的状态码的响应数 | 统计一段时间内由负载均衡直接返回给客户端的对应的状态码的响应数。 |
| 9 | TCP 的连接总数 | 未开启会话保持的情况下，统计客户端经由 TCP 协议监听器向云主机发起请求并建立 TCP 连接的总数。注意这里的连接指的是 Connection，即负载均衡处理的最小单位。 | [连接数取值说明](#_5.3. 取值说明) |
| 10 | 经过负载均衡（TCP 协 议）session 总数 | 开启会话保持的情况下，由同一个 IP 地址发出的所有连接为一个 session，源 IP 会话保持过期后，该 session 被删除掉。 | [session数取值说明](#_5.3. 取值说明) |

## **5.3. 取值说明**

|  |  |  |
| --- | --- | --- |
| **监控项取值** | **说明** | **备注** |
| 平均流速 | 时刻n的流速(单位：Bps) = (时刻n的累积流量 - 时刻n-g的累积流量)(单位：Byte)/ (时刻n - 时刻n-g)(单位：秒) | g 表示监控粒度(monitor granularity),目前出了负载均衡中的请求数、连接数、session数的监控粒度为：10s、50s  其他的流量和流速的粒度分别为：1min、2min、15min  注意：如果时刻n（n=0）,即云主机创建的时间，则取值 0 |
| 流量 | 时刻n的流量(单位：Byte) = 时刻n的累积流量 - 时刻n-g的累积流量(单位：Byte) |
| 请求数 | 时刻 n 的请求数(单位：次) = 时刻n的累积请求数 - 时刻n-g的累积请求数(单位：次) |
| 连接数 | 时刻 n 的连接数（单位：个）= 时刻n的累积连接数 - 时刻n-g的累积连接数（单位：个） |
| session 数 | 时刻 n 的会话数（单位：个）= 时刻n的累积连接数 - 时刻 n-g 的累积连接数（单位：个） |
| 响应数 | 状态码A的响应数（单位：个）= 时刻n的累积响应数 - 时刻 n-g 的累积响应数（单位：个） |

1. **性能要求**

**该版本目前没有对性能方面的特殊要求。**