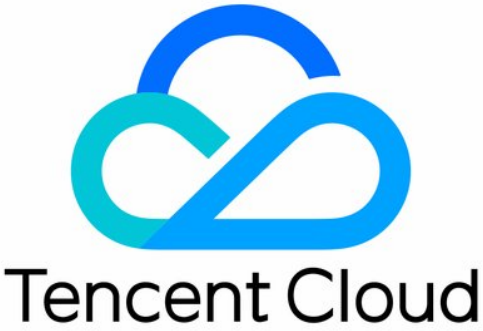
Cloud Load Balancer

제품 소개

제품 문서



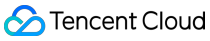
©2013-2022 Tencent Cloud. All rights reserved.

Copyright Notice

©2013-2024 Tencent Cloud. All rights reserved.

Copyright in this document is exclusively owned by Tencent Cloud. You must not reproduce, modify, copy or distribute in any way, in whole or in part, the contents of this document without Tencent Cloud's the prior written consent.

Trademark Notice



All trademarks associated with Tencent Cloud and its services are owned by Tencent Cloud Computing (Beijing) Company Limited and its affiliated companies. Trademarks of third parties referred to in this document are owned by their respective proprietors.

Service Statement

This document is intended to provide users with general information about Tencent Cloud's products and services only and does not form part of Tencent Cloud's terms and conditions. Tencent Cloud's products or services are subject to change. Specific products and services and the standards applicable to them are exclusively provided for in Tencent Cloud's applicable terms and conditions.

목록:

제품 소개
- 제품 개요
- 제품 장점
- 시나리오
- 기술 원리
- 사용 제한

©2013-2022 Tencent Cloud. All rights reserved.

제품 소개
제품 개요

최종 업데이트 날짜: : 2024-01-04 19:01:53

Cloud Load Balancer(CLB)란 무엇인가요?

CLB(Cloud Load Balancer)는 여러 리얼 서버에 트래픽을 분산시키는 서비스입니다. CLB는 트래픽 분산을 통해 응용 시스템의 외부 서비스 기능을 확장하고 단일 지점 장애를 제거하여 응용 시스템의 가용성을 향상시킵니다.
CLB 서비스는 버츄얼 서비스 주소(VIP)를 설정하여 **같은 리전**의 여러 리얼 서버를 성능과 가용성이 우수한 애플리케이션 서비스 풀로 가상화하는 것입니다. 애플리케이션에서 지정한 방식에 따라 클라이언트의 네트워크 요청을 서버 풀로 배포할 수 있습니다.
CLB 서비스는 서버 풀의 리얼 서버 인스턴스의 상태를 확인하고 비정상 상태의 인스턴스를 자동으로 격리하여 리얼 서버의 단일 실패 지점 문제를 해결하고 애플리케이션의 전반적인 서비스 성능을 향상시킵니다.
셀프 서비스 관리, 셀프 서비스 장애 복구 및 네트워크 공격 방어 기능을 갖춘 Tencent CLB는 기업, 커뮤니티, 전자 상거래 및 게임과 같은 애플리케이션 시나리오에 적용할 수 있습니다.

구성 요소

작업 CLB 그룹은 일반적으로 다음으로 구성됩니다.
Cloud Load Balancer: 트래픽 분산에 사용되는 CLB 인스턴스입니다.
VIP(virtual IP): CLB가 클라이언트에게 서비스를 제공하기 위해 사용하는 IP 주소입니다.
Backend/Real Server: 요청 처리에 사용되는 백엔드 리얼 서버 인스턴스입니다.
VPC/기본 네트워크: 전체 네트워크 환경입니다.
CLB 외부의 액세스 요청은 CLB 인스턴스를 통해 처리하기 위한 정책 및 전달 규칙을 기반으로 리얼 서버로 분산됩니다.

용어 사전

용어	전체 이름	설명
CLB	Cloud Load Balancer	Tencent Cloud에서 제공하는 네트워크 부하 분산 서비스로, 리얼 서버와 결합하여 사용자에게 TCP/UDP 및 HTTP 기반 부하 분산 서비스를 제공할 수 있습니다.
CLB	Load	수신 포트, 로드 밸런싱 정책 및 상태 확인 구성을 포함하며 각각 백엔드의 애플리케이션

©2013-2022 Tencent Cloud. All rights reserved.

리스너	Balance Listener	서비스에 해당합니다.
리얼 서버	Real Server	사용자가 설정한 규칙에 따라 CLB에서 전달한 요청을 수신하는 리얼 서버 인스턴스입니다.
VIP	Virtual IP	시스템에서 할당한 서비스 주소(현재 IP 주소)입니다. 사용자는 서비스 주소 공개 여부를 선택하여 공중망/사설망 기반 로드 밸런싱 서비스를 생성할 수 있습니다. **공중망 VIP** 일반 IP: 일반 BGP IP로 네트워크 품질과 비용의 균형을 원하는 사용자에게 적합합니다. 가속화된 IP: Anycast를 사용하여 가속화하여 공중망 액세스를 보다 안정적이고 신뢰할 수 있으며 지연 시간을 단축합니다. 정적 단일 회선 IP: 단일 ISP를 통해 공중망에 액세스하므로 비용이 저렴하고 자체 스케줄링이 편리합니다. **사설망 VIP** VPC: VPC 내의 IP. 클래식 네트워크: 클래식 네트워크의 사설 IP.

작동 원리

기본 작동 원리

CLB 인스턴스는 클라이언트로부터 인바운드 트래픽을 수신하고 하나 이상의 가용존에 있는 리얼 서버 인스턴스로 요청을 라우팅하여 처리합니다.

CLB 서비스는 주로 로드 밸런싱 리스너에서 작동합니다. 리스너는 CLB 인스턴스에 대한 요청을 모니터링하고 정책을 실행하여 이를 리얼 서버에 배포합니다. CLB는 **클라이언트 - CLB** 및 **CLB - 리얼 서버**의 두 가지 차원에서 포워딩 프로토콜 및 프로토콜 포트를 구성하여 요청을 리얼 서버에 직접 포워딩할 수 있습니다.

CLB의 여러 가용존에 걸쳐 리얼 서버 인스턴스를 구성하는 것이 좋습니다. 이를 통해 CLB는 한 가용존을 사용할 수 없을 때 다른 가용존에서 정상적으로 실행되는 인스턴스로 트래픽을 라우팅할 수 있으므로 가용존 오류로 인한 서비스 중단을 피할 수 있습니다.

라우팅 요청

클라이언트는 도메인 이름을 통해 서비스에 액세스하도록 요청합니다. 요청이 로드 밸런서로 전송되기 전에 DNS 서버는 CLB 도메인 이름을 리졸브하고 요청된 CLB IP를 클라이언트에 반환합니다. CLB 리스너는 요청을 수신하면 다른 로드 밸런싱 알고리즘을 사용하여 요청을 리얼 서버에 배포합니다. Tencent Cloud는 현재 가중 라운드 로빈 및 ip_hash 가중 최소 연결을 포함한 여러 밸런싱 알고리즘을 지원합니다.

백엔드 서버 상태 모니터링

©2013-2022 Tencent Cloud. All rights reserved.

로드 밸런서는 백엔드 인스턴스의 실행 상태를 모니터링하여 트래픽이 정상적으로 실행되는 인스턴스로만 라우팅되도록 할 수도 있습니다. 로드 밸런서는 트래픽을 비정상적인 인스턴스로 라우팅하는 것을 중지하고 정상적으로 다시 실행될 때 트래픽을 다시 라우팅합니다.

관련 서비스

CLB는 다음 서비스와 협력하여 응용 프로그램 가용성 및 확장성을 개선합니다.
CVM 인스턴스: 애플리케이션이 클라우드에서 실행되는 가상 서버입니다.
Auto Scaling: 인스턴스 수량을 기본적으로 제어합니다. Auto Scaling에서 CLB가 활성화되면 확장된 인스턴스는 자동으로 CLB 그룹에 추가되고 종료된 인스턴스는 CLB 그룹에서 자동으로 삭제됩니다.
TCOP: CLB 및 모든 백엔드 인스턴스의 실행 상태를 모니터링하고 필요한 작업을 수행하는 데 도움이 됩니다.
도메인 이름 등록 및 DNS 레졸루션 DNSPod: 사용자 지정 도메인 이름(예: `www.example.com`)을 네트워크 통신에 사용되는 IP 주소(예: `192.0.2.1`)로 변환하여 요청을 CLB 인스턴스로 쉽고 빠르게 라우팅합니다.

©2013-2022 Tencent Cloud. All rights reserved.

제품 장점

최종 업데이트 날짜: : 2023-04-24 14:34:24

본문은 CLB의 장점을 설명합니다.

고성능

하나의 단일 CLB 클러스터(하나의 인스턴스가 아님)는 수억 개의 동시 접속을 지원하고 초당 수백만 개의 데이터 패킷을 처리할 수 있습니다. 이를 통해 매일 천만 회 이상의 페이지 뷰로 전자상거래 웹사이트, 소셜 네트워킹 플랫폼 및 게임 비즈니스를 쉽게 유지할 수 있습니다.

고가용성

CLB는 클러스터 기반 배포 모드를 채택하여 최대 99.95%의 가용성을 제공합니다. 하나의 CLB 물리적 서버만 사용할 수 있는 극단적인 상황에서도 수천만 개의 동시 접속을 지원할 수 있습니다. 클러스터 시스템은 결함이 있는 인스턴스를 제 때에 제거하고 정상적인 인스턴스를 유지하여 리얼 서버가 계속해서 제대로 작동하도록 합니다.

고탄성

CLB 클러스터는 업무 부하에 따라 애플리케이션 시스템의 서비스 기능을 탄력적으로 확장하고, Auto Scaling의 동적 스케일링 그룹을 통해 CVM 인스턴스를 자동으로 생성 및 해제합니다. 이러한 기능을 동적 모니터링 시스템 및 초 단위까지 정확한 청구 시스템과 함께 사용하면 리소스 요구 사항에 수동으로 개입하거나 예측할 필요가 없으므로 컴퓨팅 리소스를 합리적으로 할당하고 리소스 낭비를 방지할 수 있습니다.

보안 및 안정성

CLB는 Anti DDoS 분산식 방어 시스템을 적용하여 대부분의 네트워크 공격(예시: DDoS, Web 침입)을 방어합니다. 초 단위로 트래픽 공격을 방어하며, IP 차단, 대역폭 포화 등 경우를 원활히 방지합니다. CLB의 자체 synproxy 공격 방어 메커니즘을 통해, Anti DDoS 시스템을 활성화하기전에 백엔드 CVM이 공격으로 무너지는 것을 방지하며, 데이터의

©2013-2022 Tencent Cloud. All rights reserved.

안정과 보안을 지켜줍니다.
CLB는 각 테넌트의 트래픽을 엄격하게 격리하고 DDoS 공격에 대한 적극적인 보호를 제공합니다. 공중망 CLB는 기본적으로 Anti-DDoS Basic을 지원합니다. 또한 CLB는 Anti-DDoS Pro, Anti-DDoS Advanced, WAF 및 기타 보안 제품을 추가로 지원하여 비즈니스를 보호합니다.
설명：
더 높은 보호 요구 사항이 있는 경우 Anti-DDoS Pro를 구입할 수 있습니다. 최소 30Gbps 이상의 DDoS 방어 능력을 제공하며, 최대 방어 능력은 각 리전의 실제 네트워크 상황에 따라 동적으로 조정됩니다.
애플리케이션 레이어에 대한 보안이 필요한 경우 Web Application Firewall을 별도 구매하여 애플리케이션 레이어 Web 보호를 제공받을 수 있습니다. Web 취약점 공격, 악성 크롤러 및 CC 공격으로부터 애플리케이션 레이어의 Web 보안을 보호합니다.

저비용

CLB를 사용하면 지루한 OPS 작업에 할애하는 시간, 추가 로드 밸런싱 하드웨어에 투자할 필요가 없으므로 하드웨어 및 인건비의 최대 99%까지 절감할 수 있습니다. 필요에 따라 여러 과금 방식을 선택할 수 있습니다.

©2013-2022 Tencent Cloud. All rights reserved.

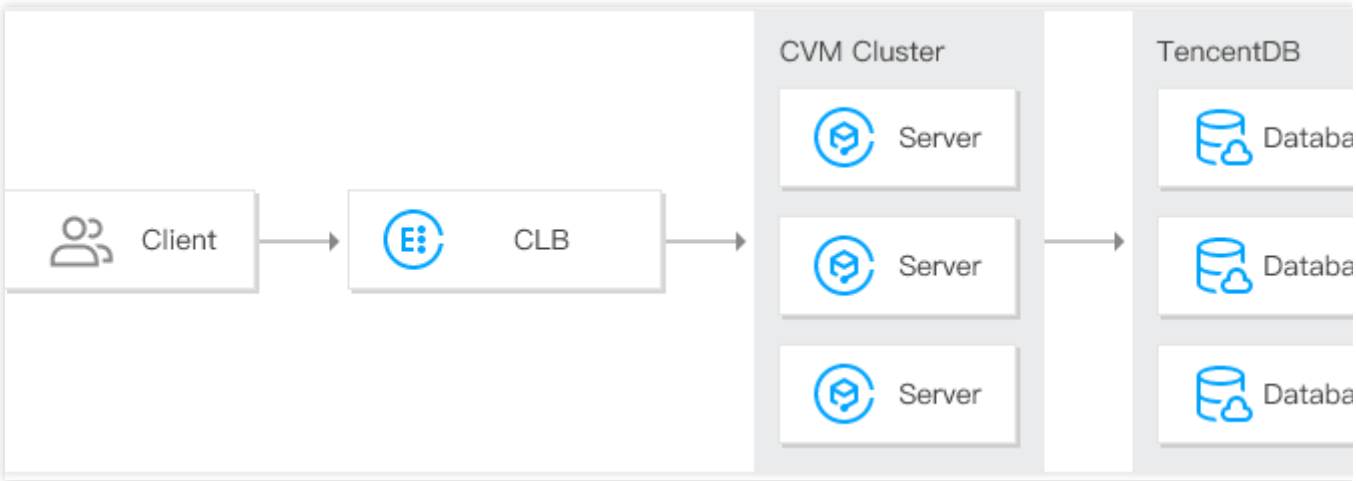
시나리오

최종 업데이트 날짜: : 2023-05-05 18:00:39

CLB는 주로 다음 시나리오에 적합합니다.
트래픽 분포. CLB는 다수의 액세스 요청이 있는 비즈니스 트래픽을 여러 리얼 서버에 분산할 수 있습니다.
단일 실패 지점 제거. 일부 리얼 서버를 사용할 수 없게 되면 CLB가 자동으로 차단하여 애플리케이션 시스템의 정상적인 작동을 보장할 수 있습니다.
수평적 확장성. 필요에 따라 애플리케이션 시스템의 서비스 기능을 확장할 수 있으며 이는 Web Server 및 App Server에 적합합니다.
글로벌 로드 밸런싱은 DNS 레졸루션 DNSPod를 통합하여 원격 재해 복구를 보장하는 글로벌 다중 리전 로드 밸런싱을 지원합니다.

단일 장애 지점의 트래픽 분산 및 제거

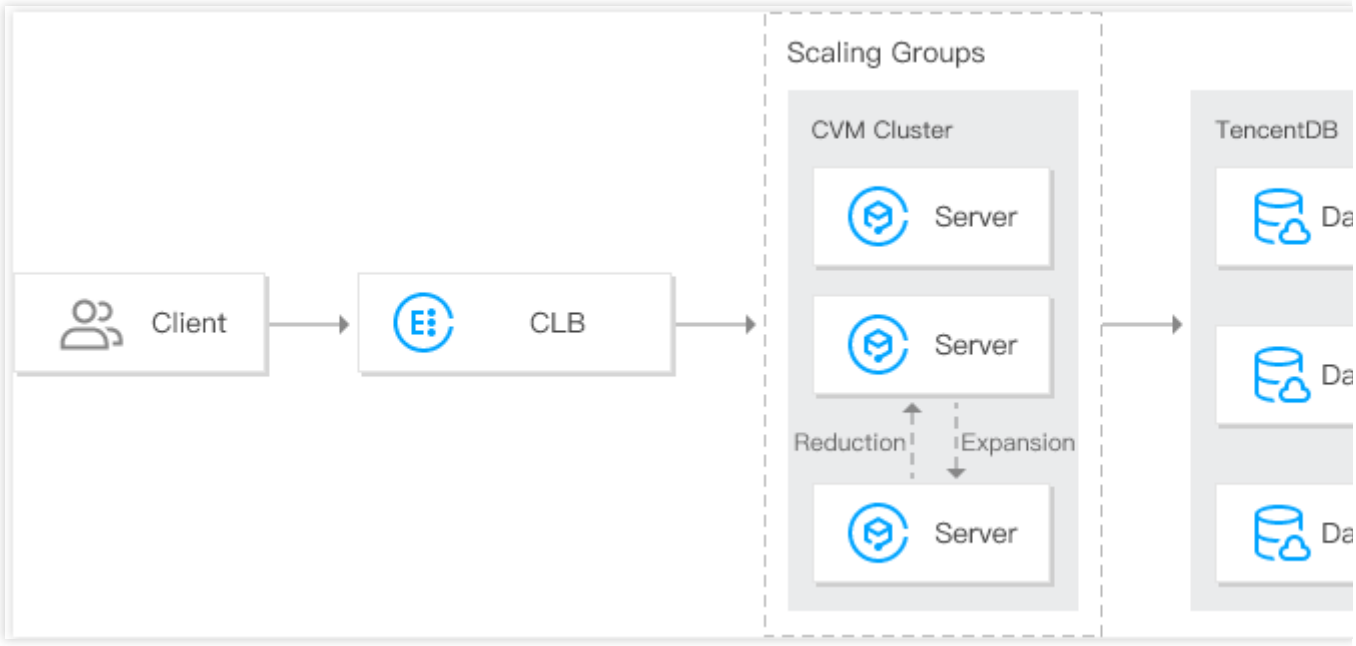
CLB를 사용하여 비즈니스 트래픽을 여러 리얼 서버에 배포할 수 있습니다.
비즈니스 클라이언트가 CLB에 액세스합니다.
여러 리얼 서버는 CLB가 비즈니스 트래픽을 포워딩하는 고성능 및 고가용성 서비스 풀을 형성합니다.
하나 이상의 리얼 서버를 사용할 수 없게 되면 CLB가 자동으로 해당 인스턴스를 차단하고 요청을 일반 CVM 인스턴스에 배포하여 애플리케이션 시스템의 작동을 보장할 수 있습니다.
비즈니스가 여러 가용존에 배포된 경우 실제 서버 레이어에서 다중 가용존 재해 복구를 보장하기 위해 CLB 인스턴스를 여러 가용존의 CVM 인스턴스에 동시에 바인딩하는 것이 좋습니다.
세션 지속성 기능은 동일한 클라이언트의 요청을 동일한 리얼 서버로 포워딩할 수 있어 액세스 효율성이 향상됩니다.



수평적 확장성

©2013-2022 Tencent Cloud. All rights reserved.

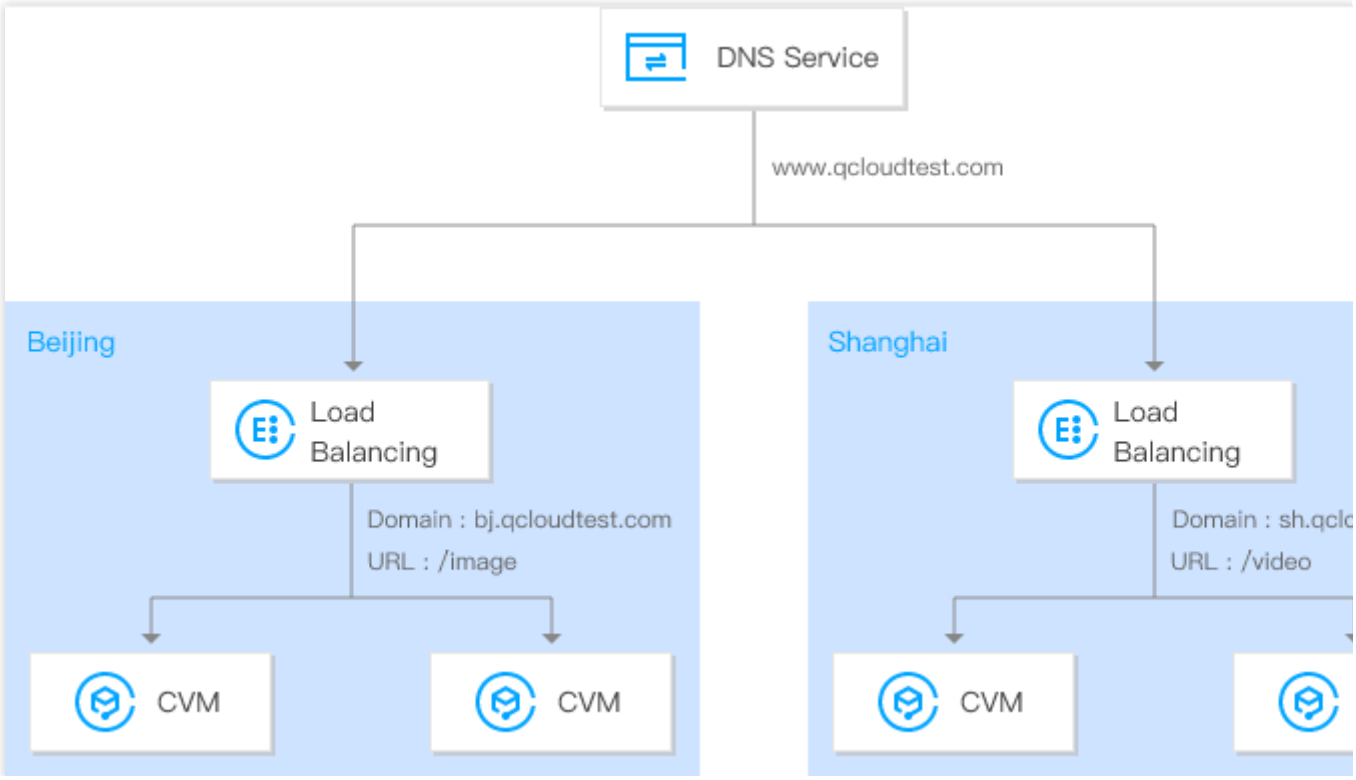
Auto Scaling을 사용하면 CLB가 비즈니스 요구 사항에 따라 CVM 인스턴스를 자동으로 생성하고 릴리스할 수 있습니다.
Auto Scaling 정책을 구성하여 CVM 인스턴스 수를 관리하고, 인스턴스 환경을 배포하고, 비즈니스 운영을 보장할 수 있습니다. CLB는 수요가 최고조에 달할 때 자동으로 CVM 인스턴스를 추가하여 고성능을 유지하고, 수요가 감소하면 비용을 절감하기 위해 CVM 인스턴스를 제거할 수 있습니다.
예를 들어 '블랙 프라이데이'와 같은 전자상거래의 주요 판매 캠페인 중에 Web 트래픽이 갑자기 10배 증가할 수 있으며 몇 시간 동안만 지속됩니다. 이 경우 CLB와 AS를 사용하여 IT 비용 절감을 극대화할 수 있습니다.



글로벌 로드 밸런싱

DNS 레졸루션 DNSPod와 결합된 CLB를 사용하면 비즈니스 트래픽을 다양한 리전의 글로벌 로드 밸런싱으로 해결할 수 있어 원격 멀티 사이트 활성-활성 및 재해 복구를 보장합니다.
다른 리전에 CLB 인스턴스를 배포하고 해당 리전의 리얼 서버에 바인딩할 수 있습니다.
DNSPod를 사용하여 다양한 리전의 CLB VIP에 대한 도메인 이름을 확인할 수 있습니다.
비즈니스 트래픽은 DNS 및 CLB를 통해 여러 리전의 여러 리얼 서버로 포워딩되어 글로벌 로드 밸런싱을 구현합니다.
리전을 사용할 수 없게 되면 해당 리전에서 CLB VIP의 레졸루션을 일시 중단하여 비즈니스에 영향을 미치지 않도록 할 수 있습니다.

©2013-2022 Tencent Cloud. All rights reserved.



©2013-2022 Tencent Cloud. All rights reserved.

기술 원리

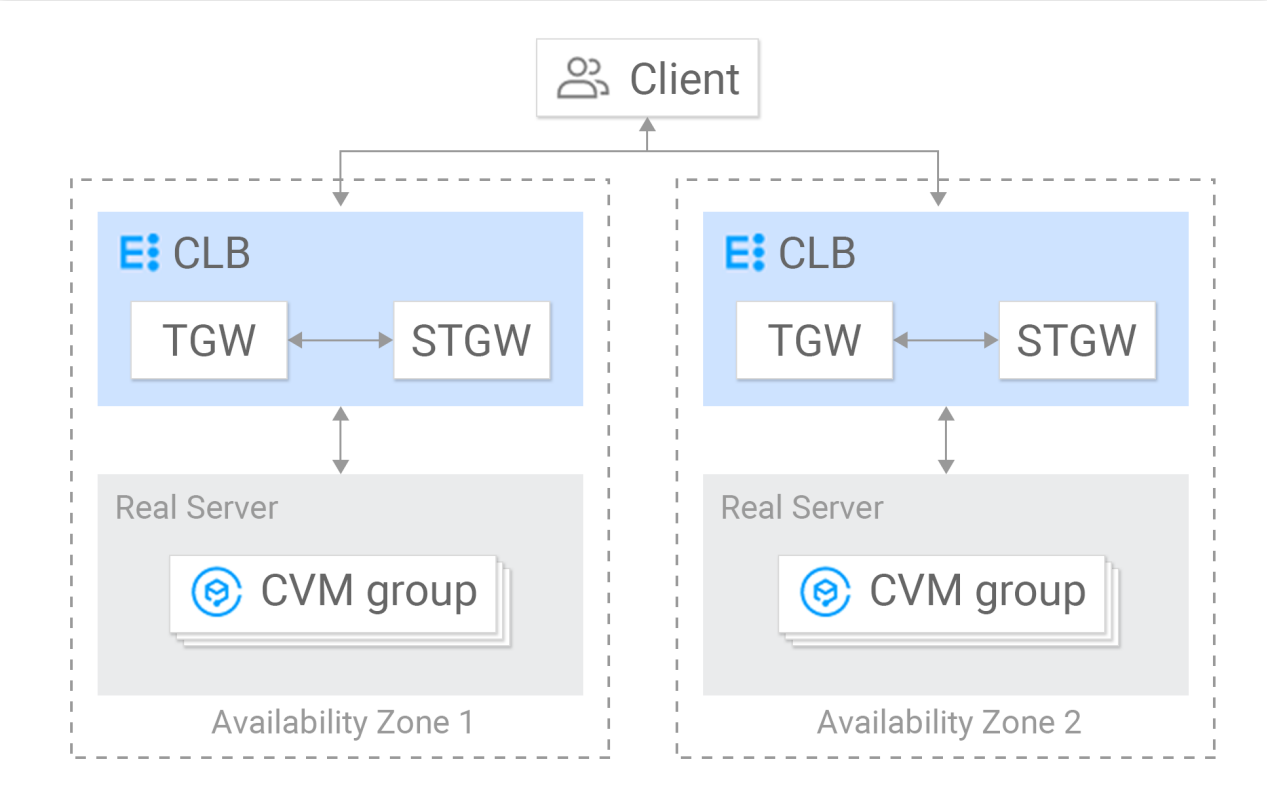
최종 업데이트 날짜: : 2023-03-23 12:01:14

CLB(Cloud Load Balancer)는 레이어 4(TCP, UDP 및 TCP SSL 프로토콜) 및 레이어 7(HTTP 및 HTTPS 프로토콜) 로드 밸런싱을 제공합니다. CLB를 사용하여 비즈니스 트래픽을 여러 실제 서버로 분산하여 단일 실패 지점을 제거하고 비즈니스 가용성을 보장할 수 있습니다. CLB는 클러스터 배치를 채택하여 세션 동기화를 달성하여 서버의 단일 장애 지점을 제거하고 시스템 이중화를 개선하여 서비스 안정성을 보장합니다. CLB는 도시 내 재해 복구를 구현하기 위해 동일한 리전의 여러 데이터 센터에 배포할 수 있습니다.

기초 아키텍처

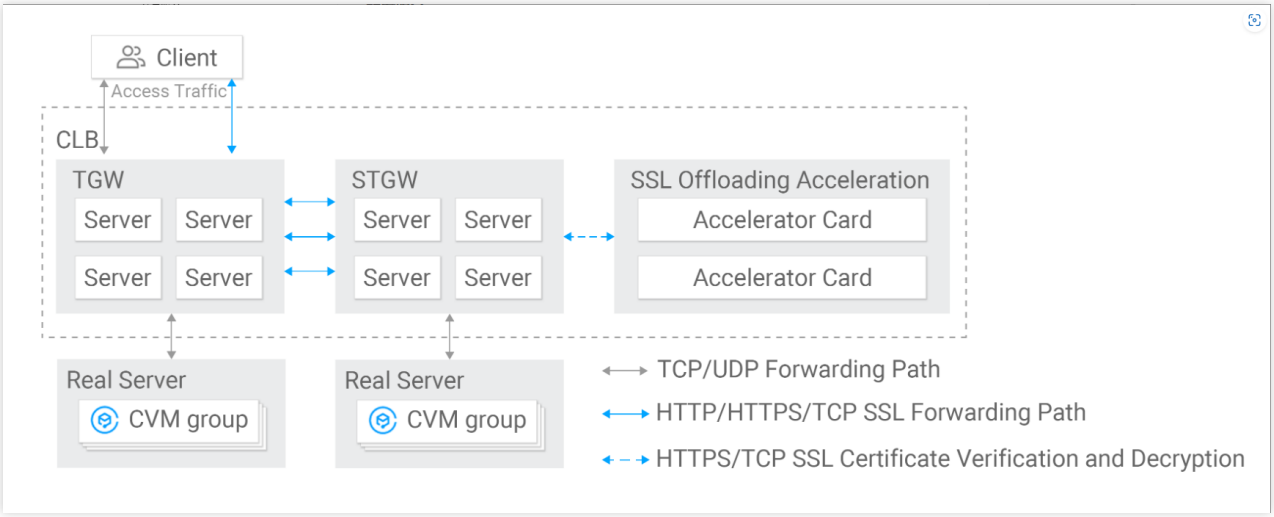
현재 Tencent Cloud CLB는 레이어 4 및 레이어 7 로드 밸런싱 서비스를 제공합니다.
레이어 4에서는 통합 TGW(Tencent Gateway)를 기반으로 로드 밸런싱이 구현됩니다. TGW는 고가용성, 고확장성, 고성능 및 강력한 공격 방지 기능과 같은 기능을 가지고 있습니다. DPDK(Data Plane Development Kit) 기반의 고성능 포워딩을 지원합니다. TGW를 사용하면 단일 클러스터가 수억 개의 동시 요청과 수천만 개의 PPS(초당 패킷)를 지원할 수 있습니다. Tencent 게임, Tencent 비디오, WeChat 및 QQ와 같은 많은 Tencent 비즈니스는 서비스 액세스를 위해 TGW를 사용합니다.
레이어 7에서는 STGW(Secure Tencent Gateway)를 기반으로 로드 밸런싱이 구현됩니다. Tencent가 Nginx 기반으로 개발한 로드 밸런싱 서비스로 대규모 동시성을 지원합니다. Tencent 뉴스, Licaitong, Tencent 게임 및 WeChat과 같은 Tencent의 레이어 7 비즈니스 트래픽을 대량으로 전달합니다.

©2013-2022 Tencent Cloud. All rights reserved.



포워딩 경로

CLB는 비즈니스 트래픽을 포워딩하고 실제 서버는 비즈니스 요청을 처리합니다. CLB는 Tencent Cloud 사설망을 통해 실제 서버와 통신합니다. TGW와 STGW는 모두 여러 서버에 배포되며 클러스터를 통해 로드 밸런싱 서비스를 제공합니다. CLB의 포워딩 경로는 다음과 같습니다.



©2013-2022 Tencent Cloud. All rights reserved.

1. TCP/UDP 프로토콜:
TCP/UDP 프로토콜의 포워딩 로직은 TGW 클러스터에서 처리합니다.
TGW는 비즈니스 트래픽을 수신한 후 Tencent Cloud의 사설망을 통해 실제 서버로 포워딩합니다. 실제 서버의 반환 패킷도 TGW를 통해 클라이언트로 반환됩니다.
2. TCP SSL 프로토콜
TCP SSL 프로토콜이 처리되면 비즈니스 트래픽은 TGW 클러스터를 거쳐 STGW 클러스터를 거쳐 실제 서버로 트래픽을 포워딩합니다.
새 세션이 설정되기 전에 인증서 확인, 암호화, 암호 해독 및 기타 작업을 위해 가속기 카드 클러스터를 통과해야 합니다.
비즈니스 트래픽이 도착하면 Tencent Cloud의 사설망을 통해 TGW, STGW, 실제 서버를 차례로 통과합니다. 반환 패킷은 역순으로 클라이언트에 전송됩니다.
3. HTTP/HTTPS 프로토콜
HTTP 또는 HTTPS 프로토콜이 처리될 때 비즈니스 트래픽은 TGW 클러스터를 통과한 다음 HTTP 프로토콜을 식별하고 트래픽을 실제 서버로 포워딩하는 STGW 클러스터를 통과합니다.
새 HTTPS 세션이 설정되기 전에 인증서 확인, 암호화, 암호 해독 및 기타 작업을 위해 가속기 카드 클러스터를 통과해야 합니다. HTTPS는 HTTP 프로토콜로 변환된 다음 실제 서버로 포워딩됩니다.
비즈니스 트래픽이 도착하면 Tencent Cloud의 사설망을 통해 TGW, STGW, 실제 서버를 차례로 통과합니다. 반환 패킷은 역순으로 클라이언트에 전송됩니다.

©2013-2022 Tencent Cloud. All rights reserved.

사용 제한

최종 업데이트 날짜: : 2022-10-21 17:38:35

본문은 CLB의 사용 제한에 대해 설명합니다.

일반 제한 사항

Tencent Cloud CLB의 사용에는 특정 제한 사항이 있으며 다양한 유형의 CLB 인스턴스에는 고유한 사용 제한 사항이 있습니다. CLB 인스턴스 유형에 대한 자세한 내용은 Instance Types를 참고하십시오.

인스턴스 유형	리소스	기본 제한
모든 인스턴스에 대한 일반 제한 사항	단일 리전에서 생성할 수 있는 공중망 CLB 인스턴스 수	일반 IP 기반 인스턴스 100개. 개인 계정당 3개의 고정 IP 기반 인스턴스 또는 기업 계정당 15개의 고정 IP 기반 인스턴스.
	단일 리전에서 하나의 계정으로 생성할 수 있는 사설망 인스턴스 수	100
	인스턴스에 추가할 수 있는 리스너 수	50
	인스턴스의 리스너가 선택할 수 있는 포트	1 - 65535 사이의 정수
CLB(이전 '애플리케이션 CLB')	CLB 인스턴스의 HTTP/HTTPS 리스너에 대해 구성할 수 있는 도메인 이름 및 URL 포워딩 규칙의 수	50
	CLB 인스턴스에서 포워딩 규칙에 바인딩할 수 있는 서버 수	100
	CLB 인스턴스의 프런트엔드 포트에 해당할 수 있는 백엔드 포트 수	멀티 포트

CLB 인스턴스는 CVM 인스턴스에서 **바인딩을 자동으로 해제하지 않습니다**. CVM 인스턴스가 격리된 후(종량제 CVM 인스턴스가 2시간 이상 연체된 경우) CLB 인스턴스에서도 **바인딩이 해제되지 않습니다**.

대역폭 피크

최대 대역폭의 의미는 아래와 같이 네트워크 과금 방식에 따라 다릅니다.

©2013-2022 Tencent Cloud. All rights reserved.

과금 방식	차이점	설명
트래픽 과금	피크 대역폭은최대 피크 대역폭으로만 간주되며 커밋된 대역폭으로 간주되지 않습니다. 대역폭 리소스가 경합하는 경우 피크 대역폭이 제한될 수 있습니다.	실행 중인 모든 트래픽 과금 인스턴스의 피크 대역폭 합계는 단일 리전에서 5Gbps를 초과할 수 없습니다. 애플리케이션에 보장되거나 더 높은 대역폭이 필요한 경우 대역폭별 과금을 선택합니다.
BWP		공유 대역폭 패키지로 과금되는 모든 실행 중 인스턴스의 피크 대역폭 합계는 단일 리전에서 50Gbps를 초과할 수 없습니다. 더 높은 대역폭이 필요한 경우 영업 담당자에게 문의하십시오.

©2013-2022 Tencent Cloud. All rights reserved.