

CDN&负载均衡

CDN

(Content Delivery Network, 内容分发网络)

核心概念

CDN 的主要作用是**让用户更快地获取静态资源**（图片、视频、JS、CSS 等），并减轻主服务器的压力。

怎么理解？

想象你开了一家网店：

- 你的仓库在广州。
- 深圳、北京、上海的顾客都来下单，如果每次都让货从广州寄过去，速度慢，快递压力大。
- 于是，你在深圳、北京、上海各放一个**小仓库**，存一些热销商品。顾客下单时，最近的仓库直接发货，这样就快很多。

这里：

- 主仓库 = **原始服务器**
- 各地小仓库 = **CDN 节点**
- 顾客拿到商品 = **用户访问资源**

特点：

- 靠近用户，提高访问速度。
- 减少服务器压力。
- 高峰期也能稳定访问。

前端开发相关点

- 对于图片、JS、CSS、视频等静态资源，通常会配置 CDN。
- 访问 URL 类似： `https://cdn.example.com/logo.png` ，浏览器会自动去最近的节点拉取。

实现

负载均衡

核心概念

负载均衡是把用户请求分配给多台服务器，保证系统稳定、高效运行。

怎么理解？

继续用网店类比：

- 你开了 10 家实体店。
- 顾客来了，不可能都排队等第一家店。
- 有一个“分流人员”把顾客引导到人少的店，这样每家店都不会太拥挤。

这里：

- 顾客请求 = 用户访问请求
- 分流人员 = 负载均衡器 (LB)
- 各家店 = 多台服务器

特点：

- 防止某台服务器过载。
- 提高访问效率。
- 系统可用性更高（某台服务器挂了，流量自动转到其他服务器）。

常用策略

1. 轮询 (Round Robin) ：每次请求轮流分配给服务器。
2. 最少连接 (Least Connections) ：分配给当前连接数最少的服务器。
3. IP 哈希 (IP Hash) ：根据用户 IP 决定分配到哪台服务器，保证同一用户请求总是去同一台服务器（有状态需求时用）。

对比

3 CDN vs 负载均衡

特点	CDN	负载均衡
作用	静态资源就近分发	请求分流到多台服务器
关注点	加速访问、节约带宽	系统稳定、高可用
使用对象	图片、JS、CSS、视频等	业务服务器、API 服务
工作方式	缓存节点缓存资源	分配请求到服务器

💡 简单理解：

- **CDN** = 内容“搬到用户附近”，加速访问。
- **负载均衡** = 流量“合理分配”，防止服务器挤爆。