

艺术，该艺术专注于机架彼此连接和与边界路由器相连。近年来，数据中心网络的设计已经成为计算机网络研究的重要分支 [Al-Fares 2008; Greenberg 2009a; Greenberg 2009b; Mydotr 2009; Guo 2009; Chen 2010; Abu-Lib-deh 2010; Alizadeh 2010; Wang 2010; Farrington 2010; Halperin2011; Wilson 2011; Mudigonda 2011; Ballani 2011; Curtis 2011; Raiciu 2011]。

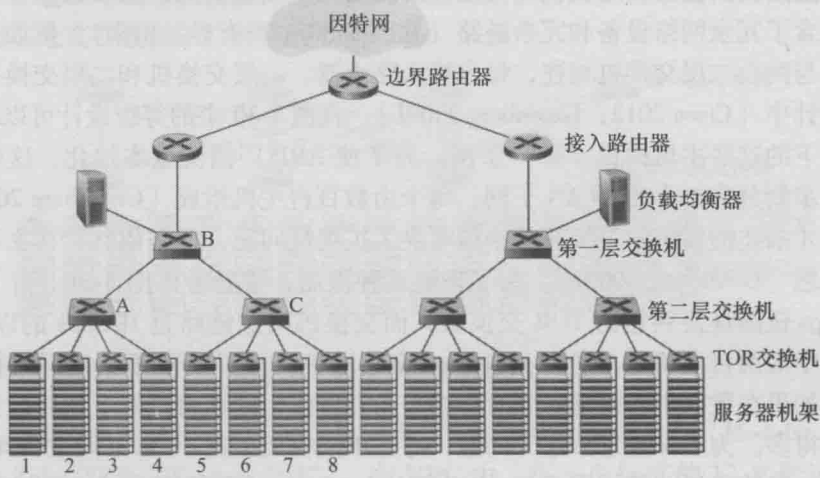


图 5-30 具有等级拓扑的数据中心网络

5.6.1 负载均衡

一个云数据中心，如一个谷歌或者微软的数据中心，能够同时提供诸如搜索、电子邮件和视频应用等许多应用。为了支持来自外部客户的请求，每一个应用都与一个公开可见的 IP 地址关联，外部用户向该地址发送其请求并从该地址接收响应。在数据中心内部，外部请求首先被定向到一个**负载均衡器**（load balancer）。负载均衡器的任务是向主机分发请求，以主机当前的负载作为函数来在主机之间均衡负载。一个大型的数据中心通常会有几台负载均衡器，每台服务于一组特定的云应用。由于负载均衡器基于分组的端口号（第4层）以及目的 IP 地址做决策，因此它们常被称为“第4层交换机”。一旦接收到一个对于特定应用程序的请求，负载均衡器将该请求分发到处理该应用的某一台主机上（该主机可能再调用其他主机的服务来协助处理该请求）。当主机处理完该请求后，向负载均衡器回送响应，再由负载均衡器将其中继发回给外部客户。负载均衡器不仅平衡主机间的工作负载，而且还提供类似 NAT 的功能，将外部 IP 地址转换为内部适当主机的 IP 地址，然后将反方向流向客户的分组按照相反的处理进行处理。这防止客户直接接触主机，从而具有隐藏网络内部结构和防止客户直接与主机交互等安全性益处。

5.6.2 等级体系结构

对于仅有数千台主机的小型数据中心，一个简单的网络也许就足够了。这种简单网络由一台边界路由器、一台负载均衡器和几十个机架组成，这些机架由单一以太网交换机进行互联。但是当主机规模扩展到几万至几十万的时候，数据中心通常应用**路由器和交换机等级结构**（hierarchy of router and switch），图 5-30 显示了这样的拓扑。在该等级结构的顶端，边界路由器与接入路由器相连（在图 5-30 中仅仅显示了两台，但是能够有更多）。在每台接入路