

核心路由器的交换矩阵设计 PDF转换可能丢失图片或格式，  
建议阅读原文

[https://www.100test.com/kao\\_ti2020/251/2021\\_2022\\_\\_E6\\_A0\\_B8\\_E5\\_BF\\_83\\_E8\\_B7\\_AF\\_E7\\_c101\\_251149.htm](https://www.100test.com/kao_ti2020/251/2021_2022__E6_A0_B8_E5_BF_83_E8_B7_AF_E7_c101_251149.htm) 交换矩阵是一个路由器的核心。 综述 在设计一个路由器时，第一步就是要决定交换矩阵的工作方式以及交换矩阵与板卡（或物理端口）之间的界面，而不是板卡的种类或是转发芯片的功能。只有在交换矩阵的结构设计完成后，板卡的设计才能启动。而随着需求的变化，板卡可以被重新设计，交换矩阵的界面却不可以，否则旧的板卡就不能和新的板卡兼容。因此，交换矩阵是用来区分路由器最好的界标，交换矩阵的特性，决定于其提速比（Speedup Ratio），对缓存的管理，延展性的好坏，制造的成本，冗余的能力，复制地能力，以及总容量。提速比 - 交换矩阵的输出速率与输入速率的比例。通常说，提速比越高，交换矩阵内的堵塞率（Blocking Ratio）就越小，但是，这样的矩阵系统的造价也将越高。因此，一味地追求的低堵塞率并不是一个现实的做法，只要能将堵塞率降至可用网络设计的方式解决，就可以合理的成本制造出接近完美的交换矩阵。 缓存的管理 - 交换矩阵可以被分成无缓存型及有缓存型，但是，不管交换矩阵是否有缓存，她都必须管理缓存，特别是进入交换矩阵前的缓存。缓存的管理不善，很容易因为缓存系统的尾端掉包，造成服务品质政策的扭曲，因此，好的交换矩阵必须有好的缓存管理。 延展性 - 任何系统的延展性都与其所用的零部件有关，规避零部件的物理限制是增加延展性的不二法则，举例说，动态随机存取记忆体（DRAM）的存取时间在20年间（1985 - 2004）仅仅缩小

的50%，而半导体的速度却增加了2720倍，因此，动态随机存取记忆体就成为一个提升系统延展性的阻碍。设计交换矩阵系统时，也必须考虑所采用的零部件是否有很好的成长空间，这样一来，就可以在不变更其工作原理及界面的条件下，使系统的容量能有数十倍，甚至数百倍的提升。制造的成本 - 新的半导体制造技术，无论是在密度及缩小耗电上都有长足的进步，但是她的价格却也上升许多。一个交换矩阵的系统，在需要增大原设计的总容量4倍，16倍，64倍，256倍，甚至数千倍时，可能会因为原系统中所采用的半导体零部件的物理限制，而必须使用新的半导体制造技术，使系统的成本大幅度的提高。因此，交换矩阵的设计者，总是在成本及性能间作平衡。一方面，系统本身必须达到升级的需求（缩小耗电，提高总容量），另一方面，又必须注意制造的成本，祈求系统能达到最好的性价比。冗余的能力 - 在交换矩阵系统的设计初期，因考虑规避半导体零部件在未来升级中，所可能遇见的瓶颈，交换矩阵系统的设计多采用多平面的结构，使核心部件在可预见的未来，不会成为系统升级的瓶颈。但多平面的结构可能会因为转发顺序的维护而造成不尽理想的冗余特性。如何能使交换矩阵在能保证转发的顺序的同时，又能有良好的冗余能力，就变成设计交换矩阵的一大挑战。点对多点的复制功能 - 交换矩阵的复制功能对其总容量有绝对的影响，一个640G的交换矩阵如不能自行复制点对多点（组播，Multicast）的流量，则其总容量，在大量的多点广播的业务环境中，会出现大幅度缩减，因此，一个优秀的交换矩阵必须要能自然地（天生）支持点对多点的复制功能。总容量 - 交换矩阵的总容量决定系统本身的吞吐能力及支

持的板卡种类。在考虑前述的几个因素下，我们将对常见的五种交换矩阵做检视（examine）。下列是五种常见的交换矩阵系统：

- \* 共享式内存 Shared Memory
- \* 全连接 Fully Meshed Connections
- \* 多方存取的数据总线 Multi-access BUS
- \* 交叉矩阵交换 CrossBar Switch
- \* 空间域多级交换矩阵 Space-domain Staged Switching Fabric

共享式内存 共享式内存作为交换矩阵的基本动作原理是利用共享的内存作为帧的交换点。对于交换矩阵而言，输入板卡将帧存入共享的内存中，并同时将在帧内的必要讯息，如MPLS的SHIM HEADER，IP的原地址及终点地址等，连同帧在共享内存的位置存入原始排序，交换晶片利用这些存于原始排序的资料找寻出输出板卡及二层封装的讯息，连同帧在共享内存的位置存入输出板卡的排序中，输出板卡利用排序中的资料将帧从共享式的内存中读出，并将帧加上适当的二层封装后，由物理端口送出。共享式的内存是用记忆体所组成，不像是查找路由，需要高速的定址能力，共享式的内存需要有的是高速的吞吐能力。同步静态记忆体（SSRAM）的工作周期短，又不需要更新（refreshing），但是其造价昂贵，且容量低，因此，同步静态记忆体更适合存放路由表，而并不适用于组成交换矩阵。同步动态记忆体（SDRAM）。单价低，容量大。但需要被周期性的更新，其吞吐能力也因为其较长的工作周期而不突出。解决这两个问题可以用多个交互更新的记忆体组 - 交互更新使得其中的一部分的记忆体组能交互地产出资料，而像是流水线的资料读取则可使吞吐能力大幅的提升。但是，使用多个记忆体组会使系统的数据接脚（data pins）快速的增加，造成系统的可服务性大幅下降。 100Test 下载频道开通，各类

考试题目直接下载。详细请访问 [www.100test.com](http://www.100test.com)