大话传送网

传送网是什么

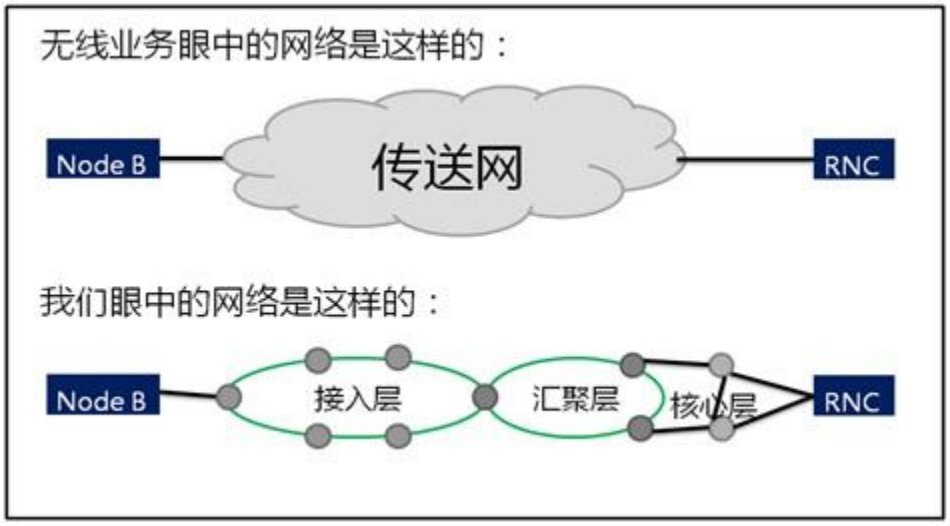
传送网是什么，这个问题不同的人会有不同的答案，可能有人会直观的理解传送网=传输设备+线路，或者是很多环和链等等，这个问题并没有标准答案，但作为一个刚刚进入这个领域的人，脑子里需要有个相对靠谱的理解。

如果把信息比作货物，传送网就是一张物流网。物流网承载的是各个企业、个人之间的业务往来，传送网承载的是各个业务网的信息往来。

固话、移动、宽带、数据、软交换、大客户等等都是靠传送网实现网元间的信息交互的，也就是说，我们之所以可以远距离的打电话、发短信、互联网上交流、看 IP 电视等，都是基于这张庞大而又复杂的传送网实现的。

传送网将遍布全球的业务层面的孤岛联成了固定电话网、移动通信网、宽带互联网，套用一句熟悉的广告词，我们不生产信息，我们只是通信系统的搬运工。

在我看来，传送网就是远距离传送信息的可靠的网络。



为什么说远距离呢，你要给你的办公室的同事或者邻居一个东西，就没必要叫快递公司。同样，信息的传递也不是处处都需要传送网，一般机房内的各种互联就可以直接对接，网线和 2M 同轴电缆可以传 100 米左右，有的设备配置单模光模块也可以传个十几甚至几十公里，但是几百公里甚至几千公里呢？远和近没有绝对的界限，只要是业务网鞭长莫及的，就要交给传送网了。

另外一方面就是容量，业务侧通过光纤直连在一定距离内固然可以实现，但是这么多专业都光纤直连，势必要消耗大量的光缆纤芯，付出的建设成本会很高。这就好比大家都不通过快递公司，而是自己开车、坐火车或飞机去送货，那肯定不是十几块钱能搞定的。传送网可以达到一对纤芯承载 8T 甚至更高的业务，传送效率越高就意味着单比特的传送成本越低，正所谓“因为专注，所以专业”。

再者就是安全，你货物交给快递肯定不希望弄丢了，传送网也必须要保证信息传递的安全性和准确性，需要提供各种容错机制、保护倒换作为安全性的保障。

其实传送网各种技术发展了几十年至今，无非就是这几个关键点：容量、安全、长距离。

1.1 支路到线路的复用

我们要通过快递寄东西，要先找来快递员填单子，将东西交给快递公司。业务网通过传送网承载业务，业务网和传送网设备之间也需要一个接口。快递寄东西，信封和包装箱有相应的

尺寸规格，业务网和传送网的接口也需要有一个标准，这个标准包含了接口的形状尺寸、电平值、速率、帧结构等。

如果尺寸不一致接头根本都塞不进去就更谈不上传送;电平值定义一致是为了接收端知道你发的电平值是代表 0 还是 1，就像古代的摔杯为号，都是事先商量好的，旁人根本傻傻搞不清楚；而速率一致才能保证一字不漏的接收信息；帧结构是规定了这一长串序列的哪几个比特是表示什么信息，就像用标点符号来断句一样。总之，想要通过传送网传递信号，就要遵循这个标准，否则就是驴唇不对马嘴。

还记得数年前，大家手机没电了要借充电器都是这样问：谁手机是诺基亚的，充电器借我用用，即使同一品牌，接口也不尽相同。现在基本安卓系统的手机就不存在这个问题，因为大家接口形状大小、充电电压都相同，这就是标准。标准统一可以实现多厂家互通，形成良性的市场竞争，避免垄断局面。

我们知道，传送网传递的是业务侧的 0 和 1 组成的码流，那么收发两端就需要这些码流以双方约定好的规则发送。传送网的发展从 PDH 到 MSTP 二十多年来，说起业务侧接口提起最多的就是 E1，所以我们先来了解一下 E1。

E1 是 PCM（脉冲编码调制）标准的一部分(日本、北美采用 T1，速率 1.544M)，那么 E1 到底是什么呢？早期的固定电话网的语音信号每路是 64K，E1 就是为传送 64K 语音信号而生的的接口，一个 E1 可以容纳 32 路 64K，那么 E1 的速率就是 32*64k=2.048Mbit/s，就是我们常说的 2M。1 路 E1 里的 32 路 64k 时隙中包含了 30 路语音信号、1 路同步信号和 1 路信令。

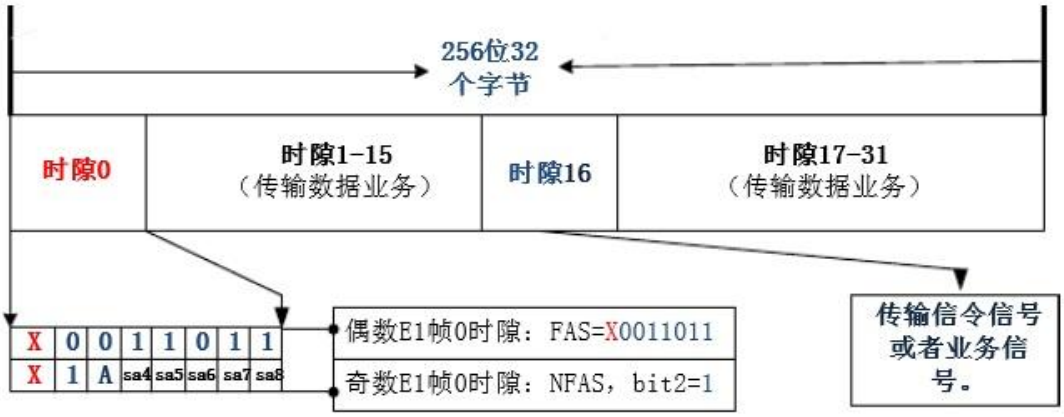
语音信号的 64K 是如何得来的，在大学通信原理中都讲过。根据奈奎斯特定律对语音信号进行每秒 8000 次的抽样就可以清晰的还原出语音信号,每次抽样的电平值用 1 个字节(8bit)表示，每路语音信号的速率就是 8K*8bit/s=64kbit/s。

E1 有 3 种用法:

一种是成复帧，用于时隙 16 传送随路信令的情况，需要将 16 帧的第 16 时隙组合起来才能传送完整的信令，所以要 16 帧捆绑起来用。

一种是信道化的 E1,就是时隙 16 不传送信令,除时隙 0 之外其余 31 个时隙用来传送信息。

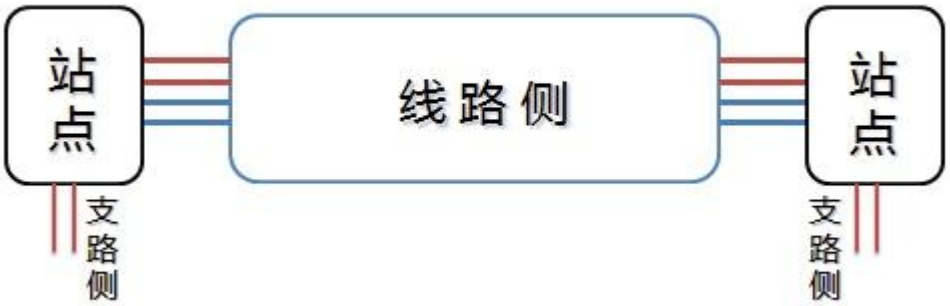
一种是非信道化的 E1，就是整个 E1 用来封装数据（如以太网），不区分 32 个时隙。



最初 E1 是因固定语音业务需求而生，后来这个 E1 也就成了传送网的接口标准之一。现如今，所有需要通过传送网传送的低速率业务，就需要遵循这个标准，如 GSM、3G 语音采用其他编码方式，速率也不是每路 64k，但接口都是沿用 E1，其他非语音的低速信号也统统沿用这个接口。这就像我们很熟悉的 5 号电池，直径 14mm，高度 49mm，我们不需要知道为什么是这个尺寸，是谁规定了这个尺寸，我们只知道生产厂家不按照这个尺寸生产，就一定卖不出去，这就是标准。

有了 E1 接口，语音业务可以接入到传送网中，可接下来业务怎么传递到目的地呢？快递公

司每收一个货物，会不会装上车就直接开往目的地？当然不会，那样和我们自己开车去送没什么区别，传送网就失去了他的意义。快递公司一定会把货物集中到一起，按照目的地分别装到大的货车中传送，这样高昂的运费分摊到每一个小包裹上就很少，成本就降下来了。
传送网需要在站点间建立一个可以传送多路业务信号的大的通道，这个通道一定比业务信号的带宽要大很多。对于传送网来说，业务接入（收发快递）叫做支路侧，站点间传送通道（物流运输）叫做线路侧，有了线路侧把站点之间连接起来，才能称之为网络。
把很多货物放到一个车厢里运输在传送网里有个专业的词，叫做复用，复用就是若干路信号合并到一起传送的过程。
下图就是一个最简单的传送网示意图，在两个站点之间建立一个 8M 的线路侧通道，可以容纳 4 个 E1 业务，站点间的 2 个 E1 业务通过线路侧的通道传送，其余两个 E1 作为冗余，可以计算出这个 8M 的带宽利用率为 50%。

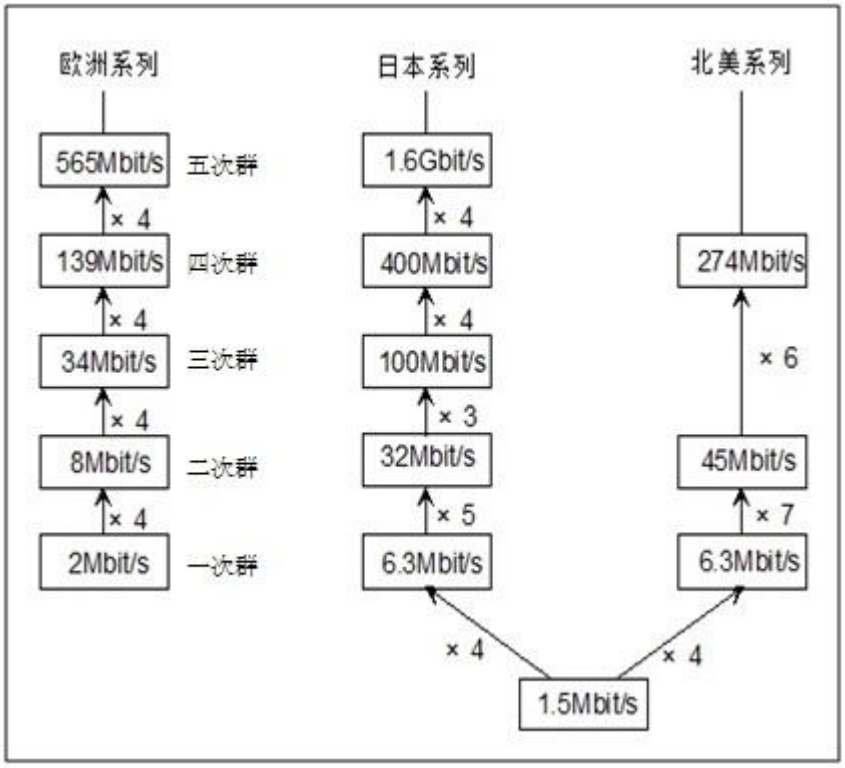


上图中，支路信号通过“时分复用”的方式装载到线路通道当中，这里有必要介绍一下各种复用方式：空分复用、时分复用、频分复用、码分复用。
我们打个比方，在一个房间里有四个人，两两成对的同时一对一交流，他们互相之间有会干扰，为了解决这个问题提高交流的效率，目前有以下几种办法：
空分复用：这个简单，让四个人分到两个房间里去对话，空间分离了，自然干扰就消除了，你走你的阳关道，我过我的独木桥。对于传送网来说，新建一个传输系统来提高容量就是空分复用。
时分复用：就是两组快速轮流说话。原本每组说一句话用 1 秒钟，现在改为每个组说 0.5 秒后换另一个组说，这样两组说话的时间互相不重叠，就像把时间切成一片片的分给大家使用，达到了快速传递信息消除干扰的目的。时分复用的等级越高，就需要说话的速度越快，就像中国好声音主持人华少那样。传送网的速率升级就是提高时分复用的等级，从 2M 到 8M，信号传送的时间不变，只是每个 bit 信号占有的时间窗口缩短到原来的 1/4。
频分复用：让两组分两个声部去说，就像女高音和男低音一同演唱那样，两组各自锁定收听各自的声部，由于声音之间差别较大易于分辨，也能达到消除干扰的效果。频分复用在生活中最常见的就是收音机，不同调频的节目都在空气中传播，我们通过调整收音机接收的频率去切换频道，只要频率保持一定的间隔，就不会收到其他频道的节目。传送网的波分复用就是让信号调制成不同的波长在一根光纤中传送，我们物理课都学过波长和频率是成反比的，波长不同就是频率不同，这实际上就是光纤内的频分复用。
码分复用：大家都有这样的经验，我们在聊天的时候，如果旁边有其他人说汉语，我们一定会觉得受打扰，但是如果旁边的人在说法语，而我们又不懂法语的话，旁边人说话对我们的干扰一定小很多，就当是背景噪声了。码分复用就是利用这个原理，让两组人分别用汉语和法语说话。码分复用在无线专业中听到的比较多，在传送网专业也有对应的 OCDMA（光码分

多址）的研究，但目前尚无应用。
了解了支路到线路的复用，那么接下来的问题是，支路侧和线路侧采用什么速率接口，支路侧接信号如何复用到线路接口中传送，我们又去怎样监控系统的工作状态等等，解决这些问题的方法需要一个完整的技术体系，比如我们接下来要说的 PDH 和 SDH。

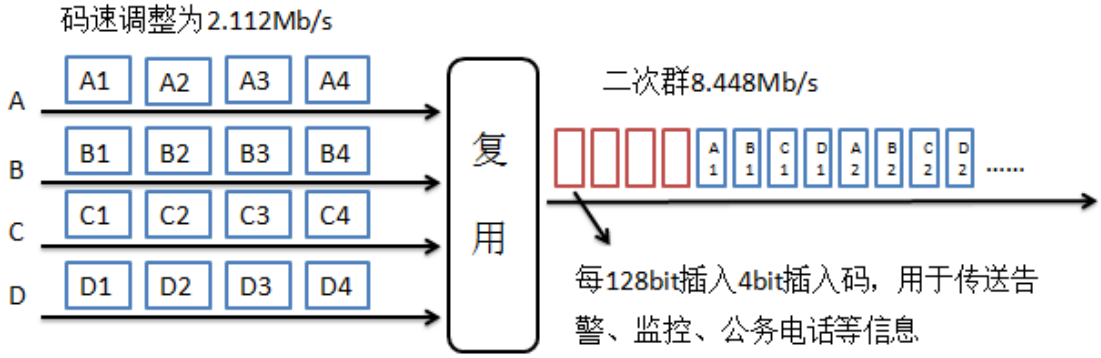
1.2 PDH 准同步数字体系

拿我们平时常用的交通工具来说，小汽车、中巴、大巴等等，这些交通工具提供的载客量是不尽相同的，但是也还是要有一定的规律可循，比如 7 座以下的小型车，一般就是 2 座（跑车）、5 座（小型车）和 7 座（商务车）。关于这个问题在行业内一定是有一定的规定或共识，遵循这个规定的基础上可以适当发挥，但是不可能乱来，而这些规定就是一个体系。
传送网也一样，把 E1 当做一个乘客来看，那么采用什么规格的线路侧接口（车厢）的容量就需要一个共识，如果 A 厂家的接口支持 10、40、160 路 E1，B 厂家支持 16、64、256 路，C 厂家又……这里面帧结构的定义就更是百花齐放，光模块种类也是五花八门，互通就不要想了，估计搞网络建设就一个头三个大，这里的必要性就不多说了。
言归正传，PDH 作为第一代光通信的标准，规定了一系列的速率等级，和等级间复用的方法，PDH 在全世界范围有两大体系三个标准，本文仅针对我国采用的欧洲系列简要介绍

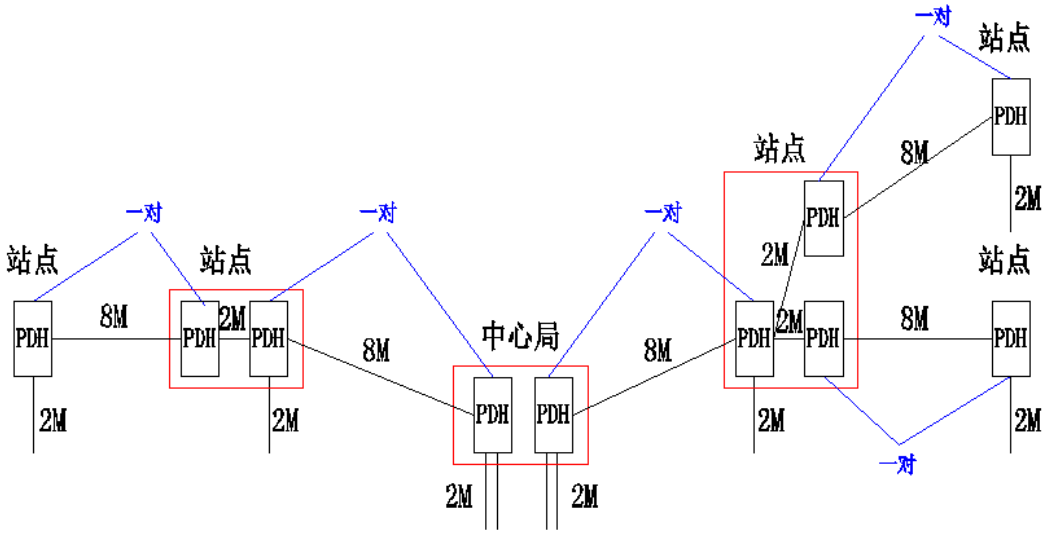


在 PDH 里各速率等级称为一次群、二次群……，我国采用的体系中，高次群和低次群容纳 E1 数量是 4 倍的关系，通俗点说，4 个 E1（一次群）被装到 8M（二次群）里，4 个 8M 被装到 34M（三次群）里，依次类推……
可是为什么各次群速率不是严格的 4 倍关系呢？因为低次群要复用成高次群之前，首先要经过码速调整。由于货物大小略有偏差，箱子的尺寸就要足够大，大于所有货物，那么当货物尺寸小于箱子的时候就要塞一些泡沫填充物。码速调整就是让各路准同步的信号变成完全同步，就是将标称速率 2.048Mb/s 但是有一定速率偏差的信号调整到 2.112Mb/s。
同步就是网元之间采用完全相同的速率，步调严格一致，你发第 1 比特的时候我根据时钟信

号就知道收第 1 比特，就像大家 7 点准时收看新闻联播，7 点半看天气预报一样，风雨无阻，年复一年。而准同步就是基本上同步，跟严格同步相比允许偏差那么一点点。
我们打电话就要求网络是同步的，否则我这边说话等了几秒钟你那边才听到，那通话就没办法进行下去，而互联网的业务就不要求同步，我们发了邮件可能由于网络这时比较繁忙，经过一定的延时对方才收到是可以被理解和忍受的。
同步之后，低次群采用按位复用的方法形成高次群。什么叫按位复用，比方需要对 4 路信号进行复用，4 路中中每路取 1 个 bit 组成第一个 4bit 的序列，然后取每路第 2 个 bit，如此循环下去。如果是每路信号取 8bit 就是按字节复用，每路信号取完整的一帧就是按帧复用。下图简单展示 4 路一次群按位复用成二次群的过程。



这里需要解释一下，线路侧和支路侧接口的区分是根据接口功能来区分的，和速率没有必然的对应关系，比如线路侧采用二次群，支路侧一次群，也可以线路侧四次群，支路侧三次群，总之线路侧的速率要大于支路侧。
遵循 PDH 规定这些速率等级，我们就可以搭建一张 PDH 传送网了。比如可以采用二次群（8M）作为线路侧，采用一次群（2M）作为支路侧，组成一个系统容量为 8M 的传送网，如下图所示：



从图中我们可以看出，PDH 设备是成对组网，一对设备组成一个点到点的链型系统。在需要同时连接多个方向的站点时，同机房需要放置多端设备（红色方框内是同一站点设备），设备间的电路转接需要靠支路口之间通过线缆互联实现。

为什么 PDH 是成对的呢？这一点有必要详细的解释一下，因为后面的密集波分复用系统（DWDM）也是成对的点到点组网，理解了 PDH 的特点，后面 PDH 和 SDH 的区别以及 DWDM 和 OTN 的区别都不难掌握。

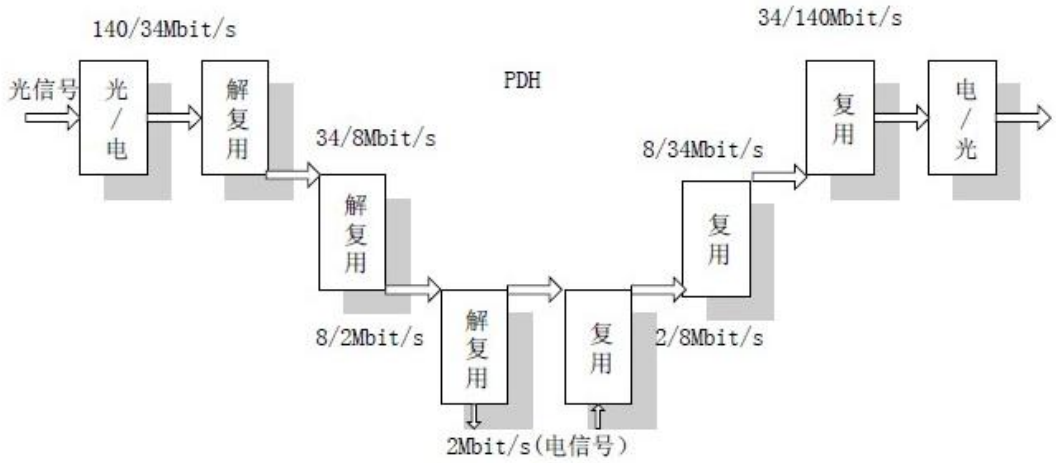
PDH 设备实现的就是一个简单的复用和解复用的功能，只能将 4 路 2M 复用成 1 路 8M，将 1 路 8M 解复用成 2M。如果需要往下一个站点传送，就需要另一个设备将信号再复用。在传送网这样的设备叫做 TM（终端复用器），TM 只有一个线路侧的接口，所以一个站点有多少个光方向，就需要有多少个 PDH 设备。

与 TM 相对应的是 ADM（分插复用器），ADM 可以支持多个光方向，所以一个 SDH 的站点无论有几个光方向都只需要一端 SDH 设备。多个方向的电路解复用之后，可以通过交叉矩阵互相调度。交叉是一个节点技术，就相当于物流调度站有一个智能的自动调度系统，将货物搬来搬去，而没有这个交叉功能就需要把所有的车货物全部卸下来，堆在地上，一些工人去进行手动的分拣，效率自然很低。支持交叉功能的站点省去了业务跳接复杂的物理连线，物理连线要靠人工完成，而交叉矩阵的电路调度可以在网管上操作，省去了人力成本。

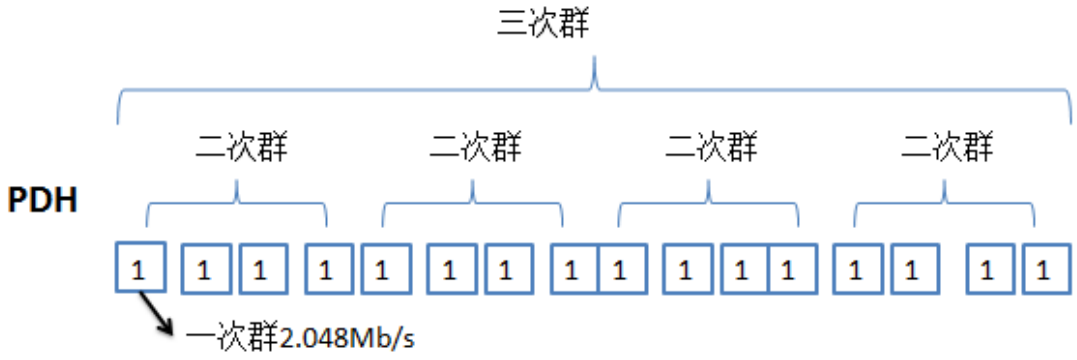
其实对于语音侧接口已经逐步 IP 化的今天来说，PDH 各次群之间如何复用我们可以不必过多详细的去了解，我们只要关注一下 PDH 的几个特点，从而明白为什么 PDH 会被 SDH 替代就可以了:

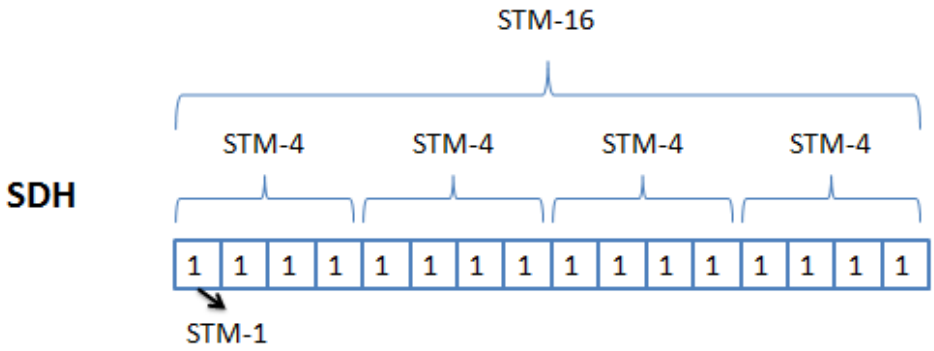
1，　　准同步就是基本差不多同步，为什么要说“准”呢，就是各个网元的时钟没有严格的统一，虽然标称的速率一致，但是总会有小范围的误差，进行高次群复用之前要通过码速调整使信号完全同步，再进行同步复用;

PDH 低次群需要逐级的复用成高次群，就是说支路信号是一次群时线路只能复用成二次群，无法直接复用成三次或四次群，如下图所示，想要复用成更高的群，需要多个背靠背的复用解复用设备，那场面是相当的壮观。



原因是这些低次群和高次群之间不是严格的 N 倍关系，中间插入的码元需要逐级的取出还原信号。就好像大箱子里面凌乱的摆放着小箱子，需要一层一层打开才知道具体的位置:





2，　　PDH 只支持点到点的网络，无法形成环路保护；

3，　　PDH 中的开销字节非常少，很多复杂的监控、管理功能都无法实现。

4，　　PDH 没有全世界统一的标准，系统间互通困难；

5，　　PDH 实际应用到四次群，系统速率偏低。

可能有人会说，为什么不一开始就制定一个很强大全面的同步系统呢，这就像我们不可能一毕业就买个奔驰宝马海景大别墅一样，技术发展也是一步一步来的，有了前面的技术和经验的积累，才能使技术向前发展。

1.3　　SDH 同步数字体系

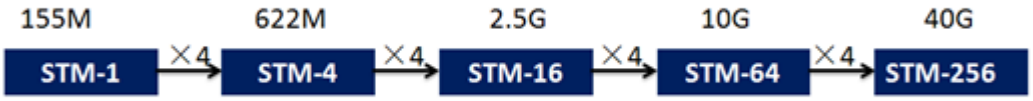
还记得上学的时候老师讲，想当年啊，别的国家的 PDH 发展到了四次群，只有我们国家研究到五次群！俺们心中顿时油然而生一种民族自豪感，紧接着老师说——因为其他国家已经开始发展 SDH 了。哎，时间过得真快，现如今，SDH 这个纵横江湖 20 来年的神一样的技术终于也要淡出历史舞台了。

下面我们开始介绍 SDH，首先照搬一个标准的定义。

SDH（同步数字体系）定义：根据 ITU-T 的建议定义，是不同速度的数位信号的传输提供相应等级的信息结构，包括复用方法和映射方法，以及相关的同步方法组成的一个技术体制。SDH 采用的信息结构等级称为同步传送模块 STM－N 。

这里提到一个词”映射”，　”映射”　就像把你要邮寄的货物装到快递公司提供的箱子里，“相应等级”就是看你的货物大小决定给你什么规格的箱子。对于 SDH 来说“映射”就是将业务信号装载到各种规格容器中，SDH 提供多种“C 容器”来装载各种速率的信号，例如将 E1 装到 C12 中，将 140M PDH 信号装到 C4 中。

SDH 的信息等级 STM-N 就和 PDH 中的各次群的概念一样，SDH 的 STM-N 的速率对应如下图所示，STM-1、STM-4、STM-16、STM-64、STM-256 这些速率等级是严格的 4 倍的关系。



SDH 与 PDH 相比较，有以下几个优点:

1、 严格同步的系统，在 STM-N 中可以直接上下低速信号，节省大量背靠背设备。

2、 统一的标准方便互通，不同国家、运营商之间可以直接对接，不同厂家设备之间也可以混合组网。

3、 单光口系统速率高达 10Gb/s，容量比 PDH 有了大幅提高。

4、 丰富的开销字节能够实现强大的网管能力，可以对 STM-N、STM-1、VC12 等不同等级的颗粒实现全面的监控。

5、 具有环网自愈保护能力，在具备两条不同路由的光缆的前提下，可以在发生故障时业务自动倒换到备用路由，保证业务不会中断，网络安全性高。

SDH 分层结构

不管是 SDH 还是 OSI 七层协议模型，业务都是从上层逐渐打包层层封装交到下层处理，直到最底层后在物理链路上传送。分层的处理信息可以将各部分的功能模块化，每一个模块需要扩展、更改的时候不至于牵一发而动全身，比如我们的电脑，硬盘和内存不够都可以单独扩容，而不会因为硬盘不够大了去更换一台电脑。

对于 SDH 而言，从小的通道逐层封装到大的 STM-N，依次经历了通道层、复用段层、再生段层。这个概念可能比较抽象，那我们来

举个例子:

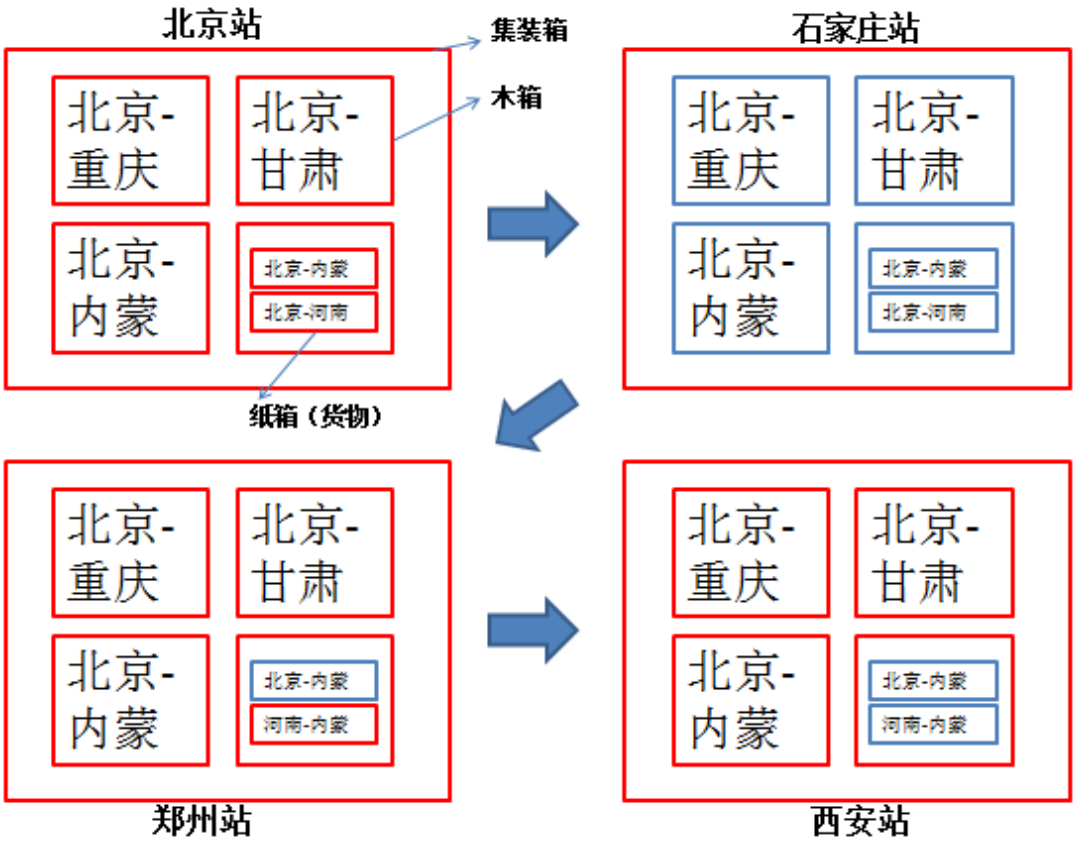
从北京途经石家庄发往郑州、西安、兰州、重庆、呼和浩特的一个货物集装箱，装箱的过程是这样的: 每个货物被分别装在了一个个的纸箱里。若干个纸箱被贴上标签装到木箱里，若干个木箱贴上标签后再装到集装箱里,如此这个层层包装的过程就分为“纸箱层”、“木箱层”、“集装箱层”。那么货物在每个中转站、调度站时是怎样打开箱子、检查标签、分发货物的呢?

石家庄中转站: 货车途经石家庄中转站的时候，石家庄只负责检查一下“集装箱层“的标签以及集装箱有无破损，确认后继续上路发往郑州，石家庄就不处理“木箱层”和“纸箱层”。

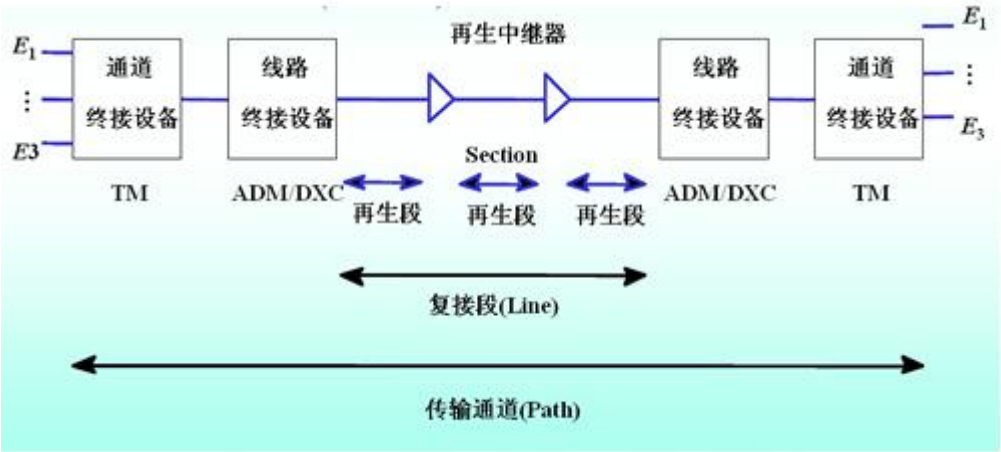
郑州调度站: 郑州方面由于有接收的货物，需要将集装箱打开，查看“木箱层”标签后，取出郑州站对应的木箱，打开木箱子查看“纸箱层”标签，取出目的地是郑州的纸箱，然后再将郑州发出的纸箱贴标签，和木箱中原来其他纸箱一起再装回木箱，再装回集装箱，发往下一站——西安。

西安调度站: 西安是一个大的货物调度中心，但是不收发货物，所以只要打开集装箱查看“木箱层”标签，将木箱子按照目的地分别装到发往重庆、兰州、呼和浩特的车上即可。但如果货物不是整木箱的发往一个目的地，西安站就需要对木箱内的货物进行整合，那么西安也需要查看纸箱的标签，进行取出重新装木箱。

以上的例子中，所有站点都处理了“集装箱层”标签，西安处理了“集装箱层”和“木箱层”标签，郑州站处理了全部三个层的标签。从下图可以看出，在每个站点红色的方框是被打开检查、读写的相应层面的标签（开销）。



现在我们可以很容易的理解，SDH 从下到上分为再生段层、复用段层、通道层，分别对应的颗粒为 STM-N、STM-1、E1 通道。再生段层就是对信号进行整形放大，不进行其他处理；复用段层就是需要打开 STM-N，对里面的 STM-1 进行重新组合。通道层就是要把业务通道终结落地。再生段对应于每两个站点之间，复用段对应于有业务上下的站点之间，通道对应于一条业务的两个端点。



1.4 SDH 帧结构和开销

要了解 SDH 的工作原理，知其所以然，就有必要大致了解一下 SDH STM-N 的帧结构，首先有必要介绍帧结构是一个什么东西:

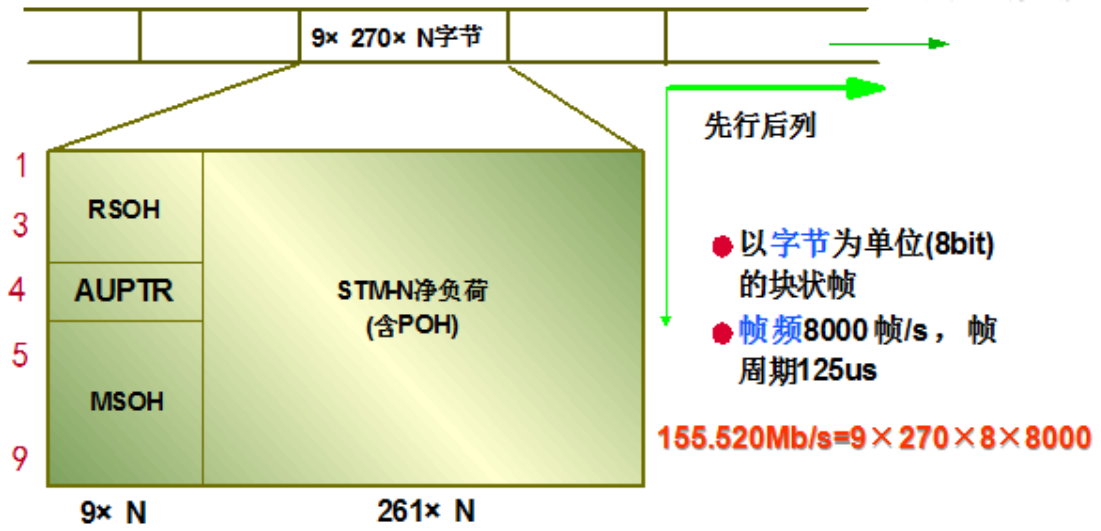
设备间在发送数据的时候，双方会按某种协议约定一个发送的顺序，每一部分有约定好的意义。对于 SDH 而言，哪些字节告诉你这一帧的起点在哪里（定帧字节），哪些字节告诉你这

整个一帧的工作状态或者里面某一通道的工作状态（开销），哪些字节告诉你这里面有没有比特串位（指针），哪些字节是真正要发送给你的数据（净荷），这几个就是 SDH 帧的组成部分。

对于接收端来说，也要按照这个标准去接收分析，才能将这一个长长的序列拆分开，能看懂里面的每一部分内容。就像我们的手机号码，+86 1XX XXXX XXXX,前面+86 代表国家，后面 3 位代表运营商，再后面 4 位代表地区，最后 4 位是卡的编号，利用这个规则，通过归属地查询软件就能够告诉我们这个号码是北京移动的或者上海联通的。

再比如以太网帧，要先发几个比特代表我这一帧开始（帧头），然后告诉你我是谁、我要找谁（源宿地址），后面是长度，告诉你我这一帧有多长，你就知道收多少个 bit 结束，然后就是数据净荷，最后是提供一个序列让你计算是否有误码（校验）。每一种协议有自己的工作方法，就有了自己相应的帧结构的组成部分。

下面我们看一下 SDH 的块状帧结构，之所以是块状帧，只是为了理解和分析方便，实际上传送的时候也是一个长长的序列，按照先行后列的顺序一行一行的传送。



SDH 帧包含了 RSOH、MSOH、POH、AUPTR、净荷几部分，每部分的作用如下:

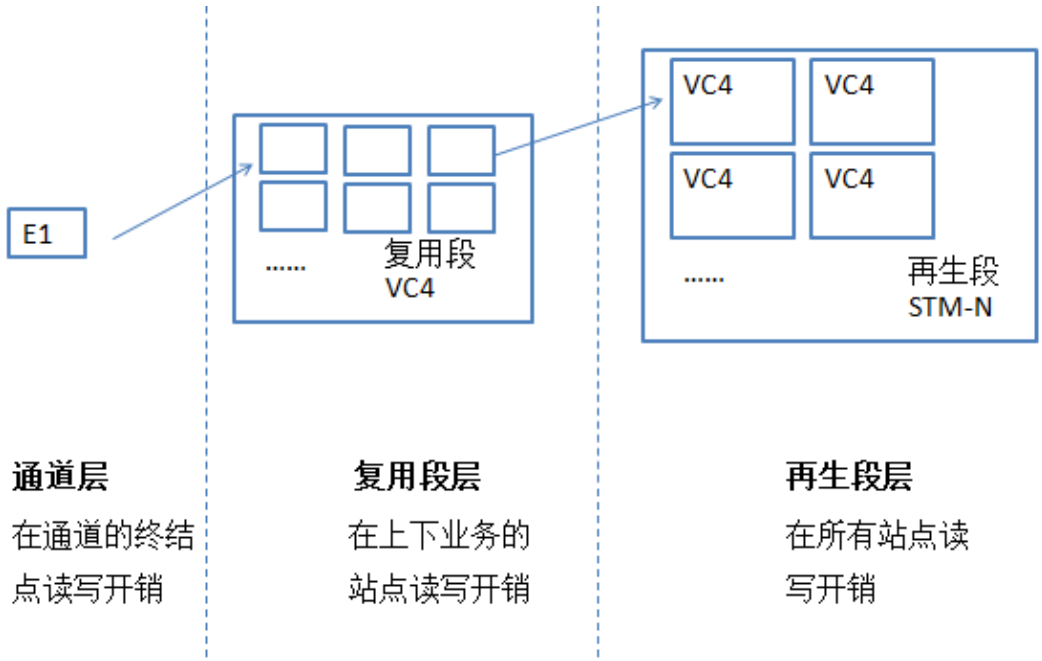
再生段开销(RSOH)—对 STM-N 整体信号进行监控;

复用段开销(MSOH)—对 STM-N 中的每一个 STM-1 信号进行监控;

指针(AUPTR)—对帧的微量偏移进行校正;

POH（通道开销）—对 STM-1 中的 VC12 等通道进行监控。

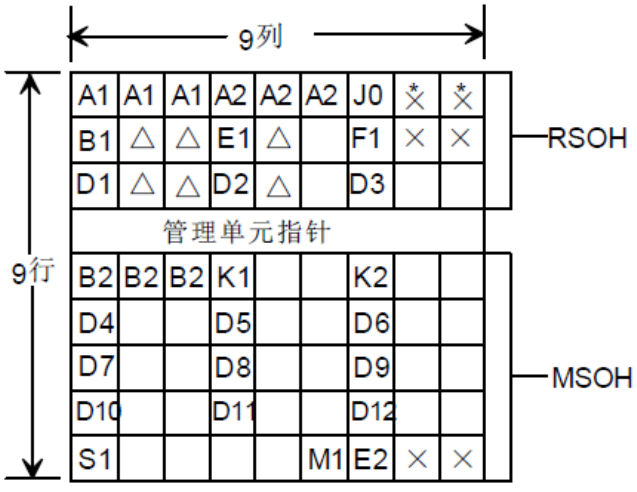
对于中继站点，仅对于 STM-N 的整体进行放大再生，所以只需要读写再生段开销（RSOH），没有必要打开帧查看每一个 STM-1。而在业务有上下的站点需要根据上下业务的颗粒将 STM-N 拆成 STM-1 或者 E1，查看读写 STM-1 工作状态对应复用段开销(MSOH)或 E1 对应的通道开销(POH)进行业务的交叉，但是对于其他站点的业务(通道)没有必要查看。通道开销(POH)仅在该业务需要调度的站点和源宿两端进行读写。



SDH 凭借各种开销组成层层细化的监控体制，能够实现对每一层的信号工作状态进行监控和管理，出现问题也能够迅速的定位到哪个 VC12。试想一下，如果没有这些开销，就像物流公司运输过程发现一个大集装箱里丢失了一个包裹，发到目的地对你说:“反正货丢了一部分，你自己看看丢了什么吧”，让人情何以堪。同时 SDH 缺点是带宽利用率也较低（开销占近 20%），不过本人觉得这不算什么缺点，为了提高安全性付出的代价是正常的。

记得以前我们寄包裹寄信，只能一遍一遍的打电话问东西到没到，现在网上就可以查到每一单货的踪迹，货到哪了由谁在派送都能够一清二楚，这个就是物流体系的发展进步，但另外一方面我们也能想到，这个查询服务系统的背后一定有一个庞大的团队和管理体系，势必也会增加了不少成本，物流的这个飞跃就同从 PDH 到 SDH 的发展颇为相似。

前面我们介绍了帧结构的组成部分，各种开销的监控范围。那么开销到底是如何工作的呢，下面再将帧结构中的开销部分放大来看一看。



△ 为与传输媒质有关的特征字节（暂用）；
× 为国内使用保留字节；
※ 为不扰码字节；
所有未标记字节待将来国际标准确定（与媒质有关的应用，附加国内使用和其他用途）

这张图是 SDH 的段开销的结构图，各字节的名称和和作用都有着详细的定义，我们大致来了解一下:

A1、A2: 定帧字节， A1 固定是 11110110，A2 固定是 00101000， 当接收端收到连续 3*N 个正确的 A1 和 A2 帧时，便知道，新的 STM-N 帧已经到来。

J0: 再生段踪迹字节,代表收发两端在再生段这一层是保持连接的。

D1-D12: 数据通信通路，是用来传送网管信息的，包括网管的操作命令、管理维护信息等。其中 D1-D3 对应再生段，D4-D12 对应复用段，SDH 就是靠这 768k（12*64k）的通道实现强大的网管功能。

E1、E2: 公务联络字节。SDH 设备上都配有一个公务电话，用于在开通设备的时候上下游站点间方便联络，公务联络字节就是传送公务电话的语音信号的。其中 E1（此 E1 非彼 E1）对应再生段层，E2 对应复用段层。如果用 E2 字节的话，就无法和中继站互通公务电话了。

F1: 使用者通路字节，也是提供一个 64k 的通道，可以传送语音和数据，可以理解为给运营商备用。

B1、B2: 比特间插奇偶校验码，B1 对应再生段，B2 对应复用段。什么是校验呢，这个我们生活中也经常用到，比如我打电话告诉你我的银行卡号，卡号读完之后告诉你一共是 16 位数字，你数一下发现不是 16 位就会告诉我再重说一次，这“共 16 位数字”就相当于一个校验码，帮助你验证是否有漏记的（比特丢失），又或者将卡号再读一遍，第二遍读卡号也相当于一个校验码，可以验证卡号记录是否有错误（收到误码）。奇偶校验顾名思义，就是利用发送的所有比特中 1 的数量是奇数还是偶数的原理，如果是奇数就在后面加校验码 1，是偶数就在后面加校验码 0，这样到了接收端，无论信息净荷是奇数还是偶数，加上校验码就一定是偶数，如果中间有单个比特出现了误码，0 变成了 1 或者 1 变成了 0，收端校验计算得出的是奇数，则判定误码。如果有 2 个比特同时误码则是无法判定的，但是同一帧中有 2 比特误码的概率是极低的。

K1、K2（b1-b5）: 自动保护倒换（APS）字节，用于实现复用段倒换保护，这个保护方式后面会有详细介绍。

K2（b6-b8））: 复用段远端失效指示，就是在使用复用段保护时，这 3 个比特告知前方有故障，信号传不过去，需要触发倒换保护。

S1:同步状态字节，值越小代表时钟等级越高，用于判定是否进行时钟切换。

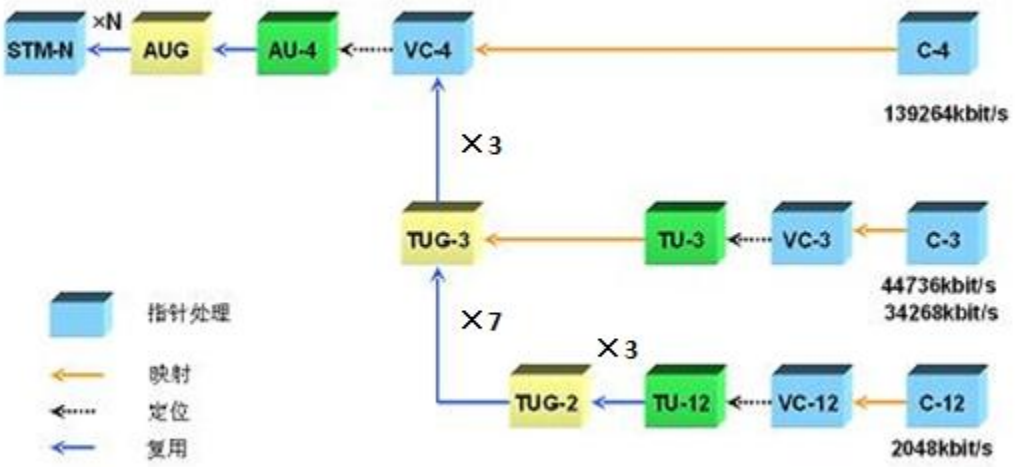
M1: 复用段远端误码块指示，用于接收端告诉发送端，我接收到了误码。

保留字节: 没有明确规定用途，厂家可以自己定义，从而实现厂家的特有功能，各种专利技术。

这里介绍这些开销目的是为了让大家大致了解开销的工作原理,至于实现的细节方面不过多解释，通道开销基本上也是这个思路，本文不逐一介绍，需要了解请参照相关资料。

1.5 SDH 复用和交叉

前文多次提到了映射（装箱子）、复用（小箱子装入大箱子），这个过程是如何实现的呢？SDH 规定了一系列的映射复用的方式，下图是我国使用的方式的示意图。



图中的各个单元对应如下:
C: 容器，VC: 虚荣器，TU: 支路单元，TUG: 支路单元组，AU: 管理单元，AUG: 管理单元组。
在这里我们对于E1封装到STM-1的过程进行简单的说明,目的是大概了解这个过程的原理，毕竟我们不是搞开发的人员。
还是把E1当做一个货物来看待,首先SDH提供一个叫做C12的箱子,这个箱子尺寸(速率)略大于E1，E1装入C12时要塞一些泡沫固定（码速调整），C12贴上标签（通道开销）之后形成了带标签的箱子(VC12)。VC12被绳子绑在了固定的位置(指针定位)之后形成TU12。3个TU12组合在一起（复用）形成了TUG2，7个TUG-2组合在一起（复用）形成了TUG3，3个TUG3又组合（复用）在一起装在了一个叫做C4的大的箱子里，C4贴上标签(通道开销)后形成VC4,VC4又被绳子固定（指针定位）后形成AU4，AU4加上车头（SOH）后最终形成了货车（STM-1）。N辆货车（STM-1）组成了长长的车队（STM-N）。
其他速率的信号复用的过程也大致类似，无非就是装箱（码速调整）、贴标签（通道开销）、绑定位置（指针定位）、组合（复用）、加车头（段开销）几个过程，从复用的路线图中都容易去理解。
SDH可以提供多种容器，包括C12、C3、C4，支路侧可以支持E1、34M、45M、140M的PDH信号，同时STM-N也可以作为支路业务，如果线路侧速率是STM-M，支路侧速率是STM-N，只要M大于N就可以。
这里SDH的复用是采用字节间插的方式，和PDH的按bit间插有所区别。字节间插可以一定程度保证信号的完整性，但需要的缓存要大一些。
特别需要说明的是，在这些容器中，VC（虚容器）是我们工作中最耳熟能详的一个，因为VC是作为一个独立的单元被调度（交叉）的，从我们举的装车的例子也容易理解，被贴上了标签的箱子作为调度、运输的基本单元被搬来搬去，而没有必要带上绳子和车头。

SDH的交叉

将STM-N打开，对里面VC4、VC12等颗粒进行读写、重新排列位置的过程称为交叉。支持了交叉功能的系统就像我们国家现在的高速公路一样，哪怕我们从西藏开车去东北，这一路经

过很多条高速公路，但是高速之间可以通过互通立交自由的切换，而不用频繁的下高速、上高速。

交叉分为高阶和低阶交叉，高阶交叉对应的颗粒是 VC4（大箱子），低阶交叉对应的颗粒是 VC12（小箱子）。交叉是靠交叉矩阵实现的，交叉矩阵将各个方向来的各种级别的信号进行调度。

交叉能力是 SDH 设备的一个重要指标，高阶交叉能力是一个设备层次的定位，比如核心层一般对应的是 300G 以上，汇聚层大概 100 多个 G，接入层一般就是几十 G。高阶交叉能力一般有两种表示方法，一种是 N*VC4，一种是多少个 G，两者之间可以换算，比如 128*VC4 换算过来就是 128*155M，大约就是 20G。

低阶交叉对应的是设备对小颗粒业务的处理能力，靠低阶交叉模块来实现，低阶交叉只有在打开 VC4 处理 E1 的时候才会用到，对于不上下 E1 业务的汇聚点来说，只对 VC4 级别进行调度不需要低阶交叉，高阶交叉能力强的设备低阶交叉能力不一定强。低阶交叉能力也有两种表示方法，一种是 N*VC12，一种是多少个 G，和高阶交叉能力一样可以换算，2016*VC12=32*155M=2*2.5G=5G。

高阶交叉

低阶交叉

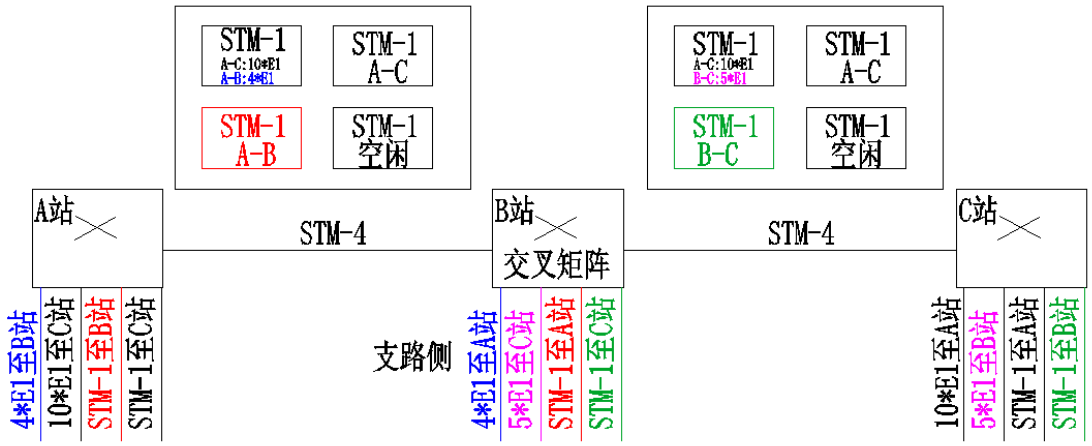
我们从下面示例来了解一下高阶交叉和低阶交叉的工作，假设有 A、B、C 三个站点组成一个链型系统，B 站为 ADM（光分叉复用节点，包含 2 个以上线路接口），A、C 站为 TM（光终端复用节点，只有 1 个线路接口），站点间业务需求如下:

A 站-B 站业务需求: 1*STM-1，4*E1

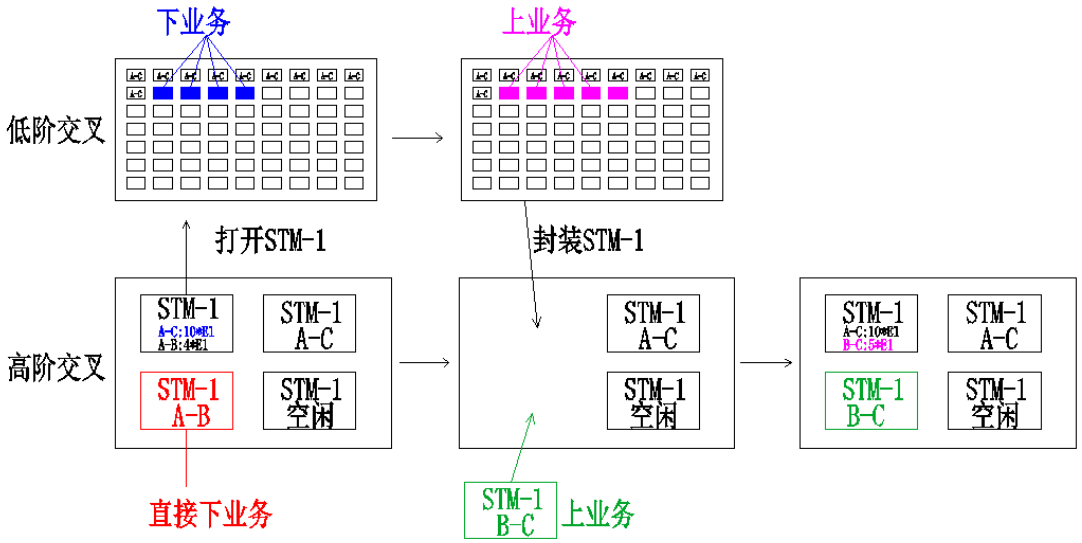
A 站-C 站业务需求: 1*STM-1，10*E1

B 站-C 站业务需求: 1*STM-1，5*E1

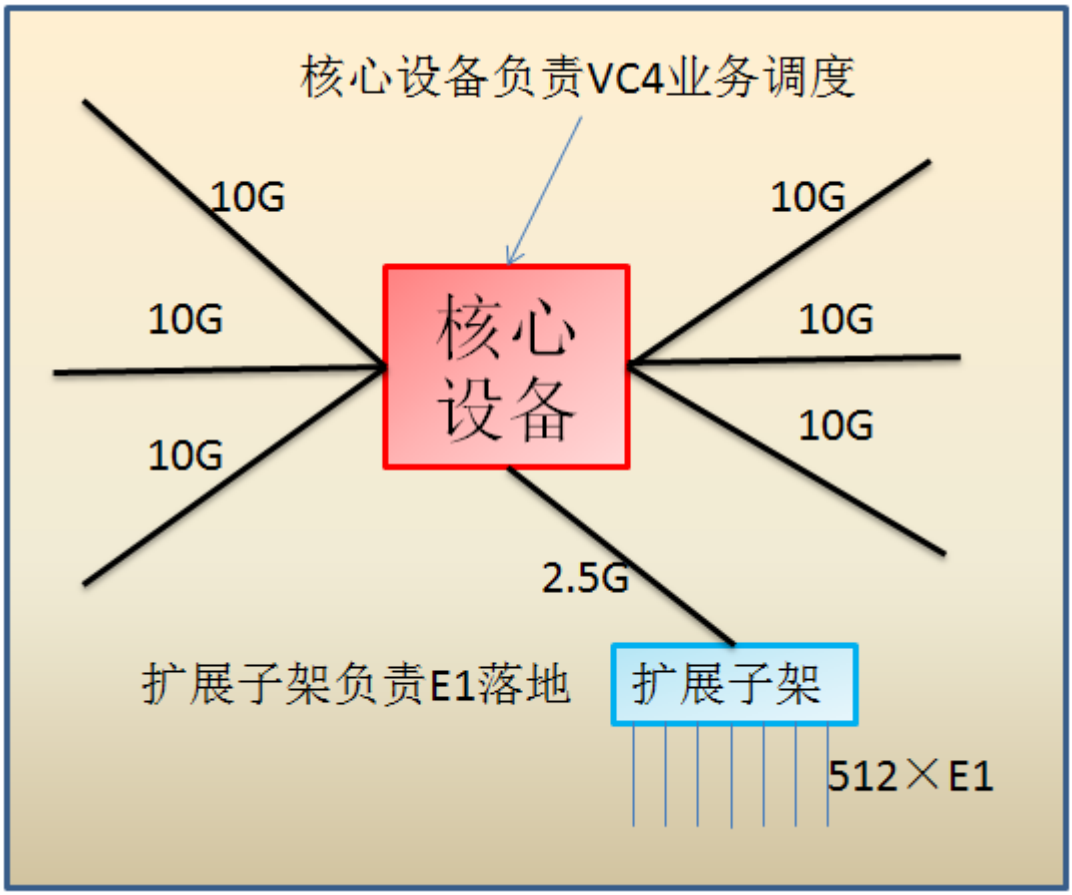
下图可以看出从 A 站发往 B 站的 STM-4 帧在经过了 B 站的交叉矩阵后,被封装成了新的 STM-4 帧发往 C 站。图中在 B 站经过交叉矩阵调度的业务和对应的支路接口用了相同的颜色表示，可以清楚的看出业务和接口的对应关系。



下图可以进一步看出高阶交叉和低阶交叉的区别和工作原理，其中 STM-1 的业务经过高阶交叉直接上下业务，而左上角的 STM-1 帧中封装了不同去向的 E1 业务，在 B 站需要将 STM-1 打开进行调度。在这个示例中，B 站点占用了 8*VC4 的高阶交叉容量和 126*VC12 的低阶交叉容量。



一般来说，厂家提供的 SDH 设备中，定位于纯核心层的大容量交叉设备一般不提供低阶交叉能力，核心层设备一般都会下挂扩展子架，进行低阶业务的处理。这是因为核心层设备都比较贵，核心设备的槽位也是“寸土寸金“，占用一个大容量的槽位去接入 E1 这样小的业务非常浪费。就像公司的总经理日理万机，一些端茶倒水的小事让他去做大材小用，所以给总经理配助理或秘书。

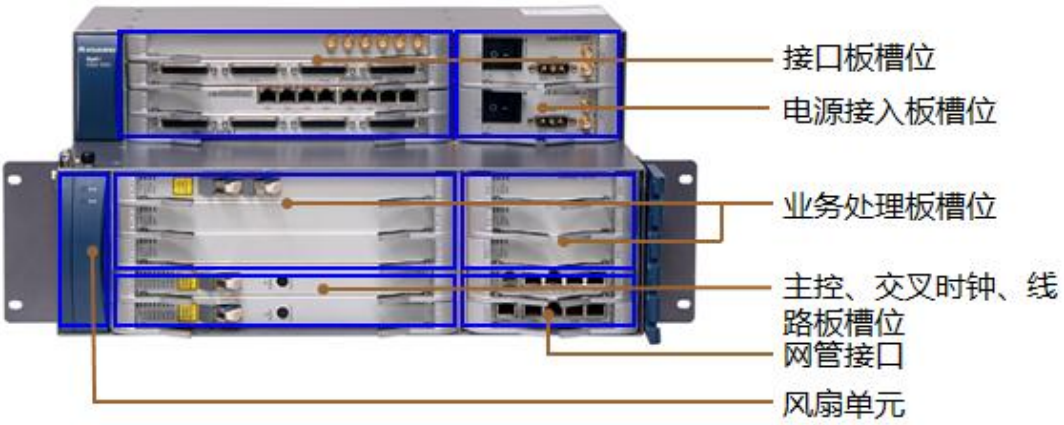


运营商一般在集采的时候会对各个层面设备的交叉能力、业务接入能力等设定一个下限要求，各厂家需满足条件才有资格参与投标。

1.6 SDH 设备组成及参数

前面部分为大家简单介绍了 SDH 体系的原理部分，下面将对这些抽象的概念具体化，通过设备实物逐步带大家接触这个真实的网络。我们以一端 SDH 设备为例，介绍一下设备的主要组成部分和参数，在我们需要的时候，可以查询厂家的设备资料进一步了解。

下图就是一端 SDH 设备的照片。

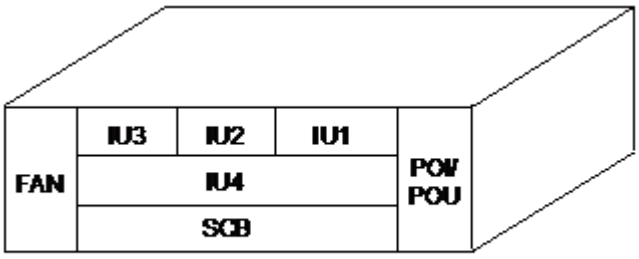


SDH 设备按照配置来说由子架、公共单板、业务单板组成。

子架是 SDH 设备的骨架，就像一个空壳子，子架后面是设备背板，背板提供一定数量槽位，可以插入各种类型的单板。
SDH 设备的单板可以分为公共单板和业务单板两类:
公共单板一般包括电源、主控、交叉、时钟、风扇等，公共单板是一个设备能够正常运转的必配的板件，是负责设备的供电、散热、时钟提供、交叉矩阵等必备的功能单元。
业务单板是我们打交道最多的单板，是负责业务的接入和处理的单板，业务单板的配置是根据业务需求来选配的，一个业务单板有这几个参数: 接口数量（几路）、接口速率（STM-N、E1 等）、接口类型（光/电、传输距离等），比如我们通常说的 8 口短距 155M 光板、2 口长距 10G 光板、16 路 E1 电接口板等等。
一个 SDH 设备的主要参数有尺寸、重量、电源端子需求、功耗、交叉能力、最高速率、最大接入能力、业务槽位数、单业务的最大接入数量。
尺寸、重量、电源端子、功耗可以用来判断一个机房是否具备安装条件，机柜内是否有空间安装，机房承重是否满足要求，电源端子、整流模块、蓄电池是否满足设备供电要求。
交叉能力是一个设备能力的表达，就像我们电脑的处理器是 I5 还是 I7 一样，能够反映一个设备的层次级别。
最高速率是指设备可以提供的最高速率端口是多大，就是我们常说的这个设备是 10G 设备还是 2.5G 设备，可以组成多大速率的系统。
最大接入能力是指设备插满最大速率的单板之后可以接入多少业务量，比如共有 12 个槽位，每个槽位最大可插 2 路 10G 的单板，设备最大接入能力是 240G。
业务槽位数是指设备可用于插业务单板的槽位数量，槽位数越多，设备可插单板就越多，设备接入能力就越大，配置越灵活。
单业务最大接入数量是指对于单一业务而言，设备最大可以提供多少路接入，比如设备所有槽位都用来插 622M 单板，可以插 12 块 4 路 622M 单板，622M 的最大接入能力就是 48 路。

下面以华为 Metro 1000 设备为例，来了解一款具体的设备。

OptiX 155/622H 的主机外形尺寸为: 436mm（宽）×293mm（深）×86mm（高），按照 19 英寸标准设计，满配置重量不超过 10kg，最大功耗不超过 100W，有了这些参数，我们就知道在一个机房安装此设备，需要机柜有 2U 的空间，需要增加机房负载约 2A，需要占用 2 个 6A 的电源端子。
OptiX 155/622H 交叉容量为 21.25G 高阶全交叉和 5G 低阶全交叉，设备支持最大速率 2.5G，设备共有 4 个业务槽位，槽位图如下:



Metro1000 设备各种接口的最大接入数量：16×STM-1（光）、6×STM-1（电）、8×STM-4（光）、2×STM-16（光）、112×E1、9×E3/T3、24×FE（电）、8×FE（光）、3×GE（光）。

有了初步的了解之后可以知道这款设备大致定位于网络的末端接入层，我们还可以进一步了解设备哪些槽位可以对应插哪些类型的单板，在工作中可以根据组网的端口实际需求，去选择适合的设备和确定设备的单板配置。

1.7 SDH 保护和组网

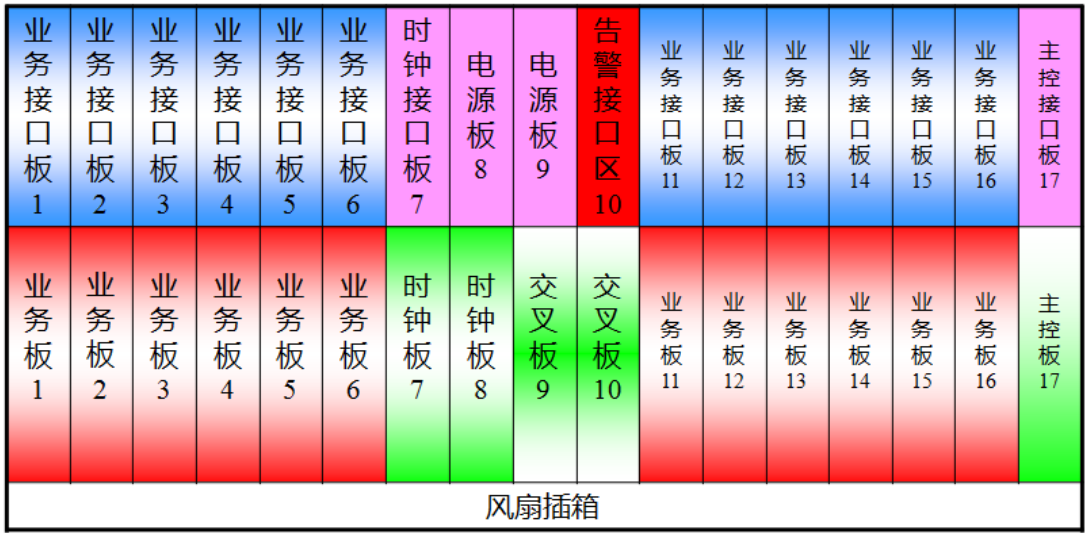
1.7.1 SDH 保护方式

SDH 的丰富、快速的保护机制使网络的安全性得到了很高的保障，我们听到过“类 SDH 保护”这样的词，这说明 SDH 保护机制在网络演进的过程中是经过市场考验的，甚至作为衡量其他技术安全性的标杆。

从保护的层面来说，SDH 保护分为单板级保护和网络级保护。

单板保护是指通过单板的冗余配置，在一块板故障的时候，另一块可以继续工作，不影响业务处理。一般情况 SDH 设备的电源板、交叉主控板都配置 2 块互为备份，这种保护一般称之为 1+1 备份。

下图可以看出，SDH 设备的电源、交叉、时钟板都分别有 2 个槽位。



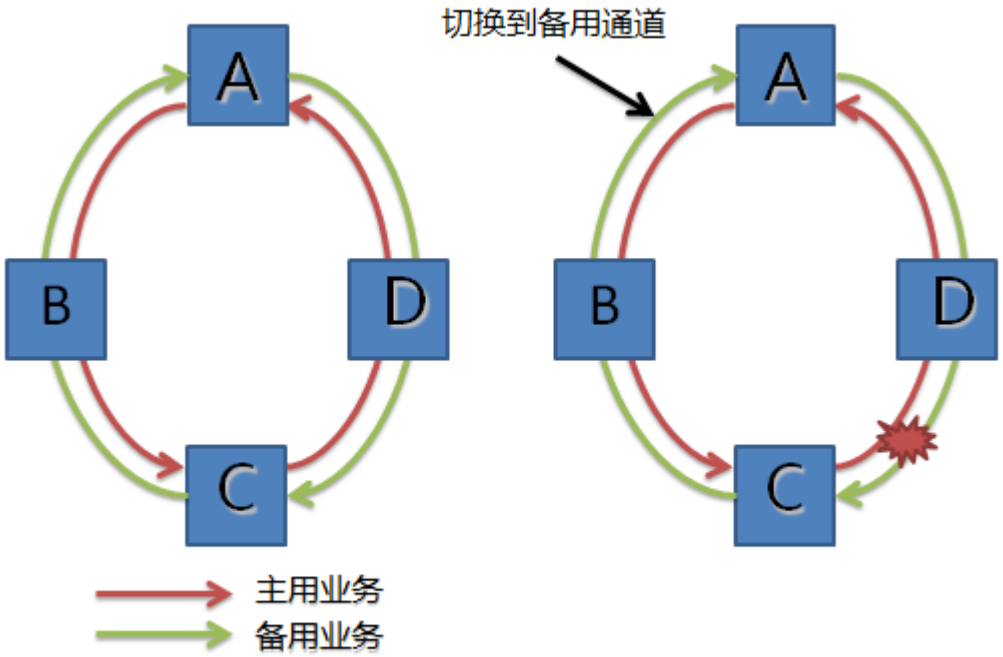
汇聚层以上的 SDH 设备业务单板一般支持 1 块备用单板为 N 块主用单板提供备份的功能，N 块业务单板中的 1 块损坏时，业务可以自动切换到备用板件上处理，这种保护一般称之为 1:N 保护。

网络保护是指当两点间的链路或节点设备故障时，业务可以通过其他路径倒换传送，保证业务不会中断。任何网络保护的前提是要有至少 2 条光缆路径可以到达目的地，也就是我们通常说的物理成环。

SDH 网络保护方式主要有两大类：通道保护和复用段保护。

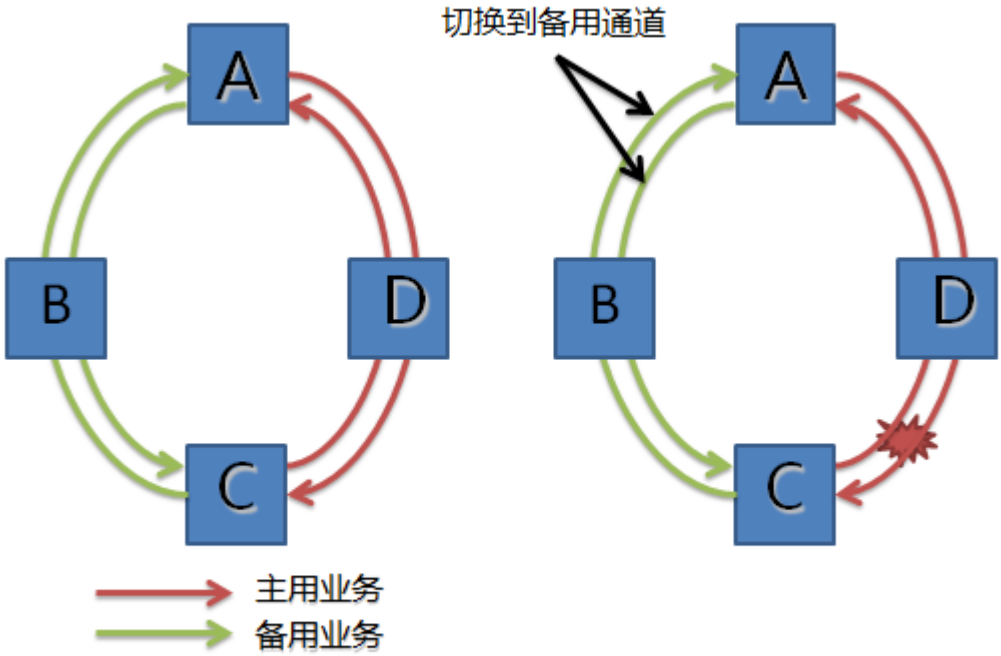
通道保护是最简单快速的保护方式，通道保护的原理概括为“并发优收”四个字，保护的颗粒是以通道（VC）为基础，就是说某个 E1 出现了问题，可以这个 E1 为单位单独进行倒换，不影响其他业务。

通道保护的具体过程是这样的：发送端将主用和备用信号从东西向同时发送，接收端接收主用信号，主用信号中断或劣化时，收端根据信号质量决定是否切换。如下图所示，A 和 C 之间的业务都通过 B、D 两个方向同时发送，正常情况接收端接收红色线条的主用业务，当红色线条路由中断时，接收端自动切换到绿色的备用通道上去。



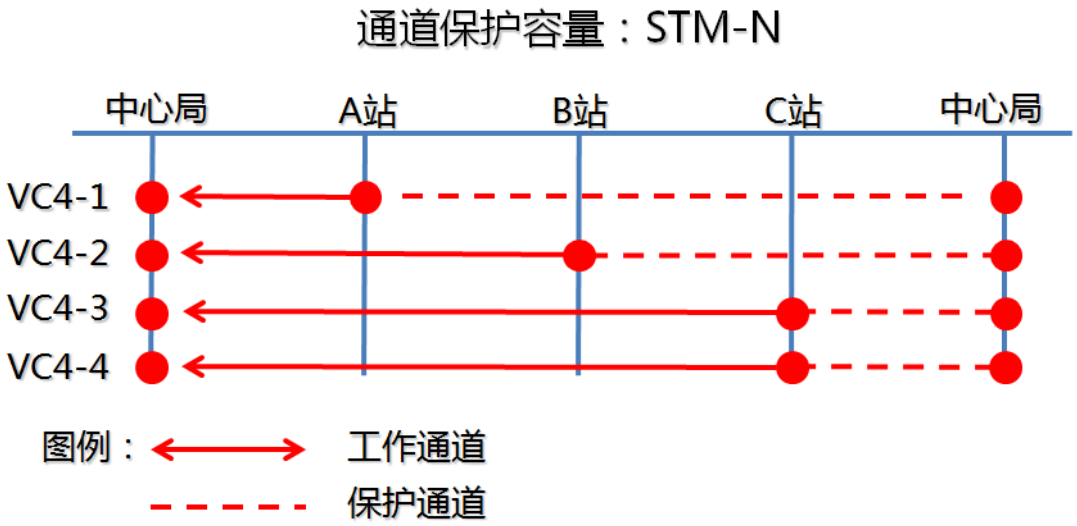
二纤单向通道保护环

通道保护分为单向和双向，A-C 和 C-A 的主用业务分别走不同的路径称为单向，A-C 和 C-A 的主用业务走同一个路径（都经过 D 点）的是双向。双向通道保护不常用，其实原理上单向和双向并没有大的区别，下面附上双向通道保护的图，大家感受一下。



二纤双向通道保护环

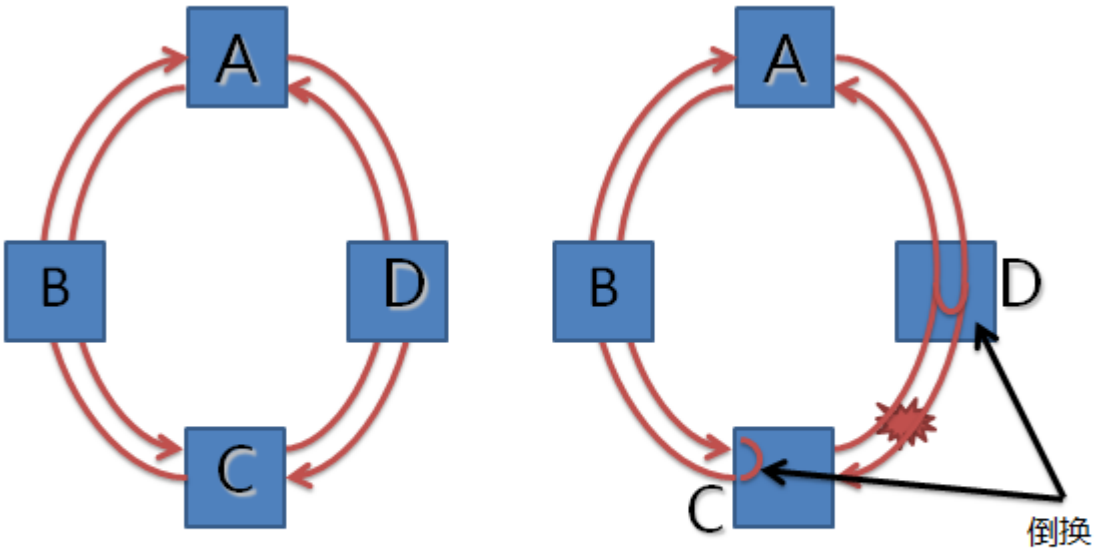
对于一个 STM-N 的环路来说，使用通道保护，系统的容量就是 STM-N，因为对于每一个通道，除主用业务占用的路径外，其余的路径全部用于业务保护倒换，比如中心局到 A 站配置一个 VC4 的业务，那这个 VC4 通道剩下的 A-B-C-中心局的段落全部用作保护通道，其他业务不能占用，如下图所示：



复用段保护是利用段开销的 K1 和 K2（b1-b5）字节实现的保护方式，复用段保护是以 VC4 为单位进行倒换，只能用于 STM-4 以上的网络。保护原理复杂一些，需要进行双端倒换和启用 APS 保护协议，所以保护时间要比通道保护稍微慢一些（≤50 毫秒）。

复用段保护分为二纤单向、二纤双向、四纤双向，下面重点对常用的二纤双向复用段保护进行介绍。

首先将 STM-N 的一半预留为备用通道，以 STM-16 环路为例，容量共 16 个 VC4，则将 1-8#VC4 用作传送业务，其余 9-16#VC4 留作备用。当线路发生中断时，设备检测到故障触发 APS 协议，在故障点两端的设备内部进行倒换，将中断的业务倒换到反向的 9-16#VC4 中传送。



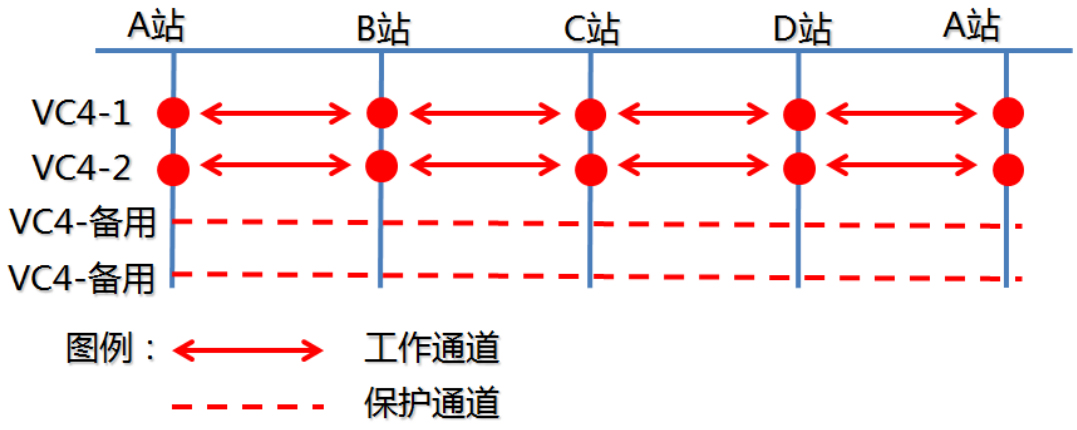
二纤双向复用段保护环

复用段保护的优势在于理论容量大于通道保护，二纤双向复用段保护的理论容量=M/2*STM-N，其中 M 为节点数。这个容量怎么计算来的呢？

首先复用段环由于要预留一半通道，所以可用的通道数量为 1/2*STM-N。

然后剩余的 1/2 的工作通道可以传送任意两点之间的业务，在极端情况下任意相邻节点之间均有业务需求，这种情况下 M 个站点组成的环路的一个通道就可以传送 M 条业务，因此 STM-N 环路就可以达到 M/2*STM-N 的容量。如下图，一个 4 节点的 STM-4 复用段环最大可以传送 2*STM-4 的业务（8 条 VC4）。

复用段保护容量M/2*STM-N



实际情况容量都不会达到这么理想，所以这个容量称为理论容量。在环路业务为集中型业务的时候，复用段环的容量与通道环容量是相同的，在分散型业务的情况下复用段环的容量大于通道环。

什么是分散型和集中型业务？分散型业务就像公交车，乘客在每一个站上上下下，也就是 A 到 B、B 到 C、C 到 D、D 到 A 都有业务；集中型业务就像机场大巴，大家从不同站点上车，但目的地只有一个—机场，也就是 A 到 B、A 到 C、A 到 D 有业务，B、C、D 三点间没有业务需求。

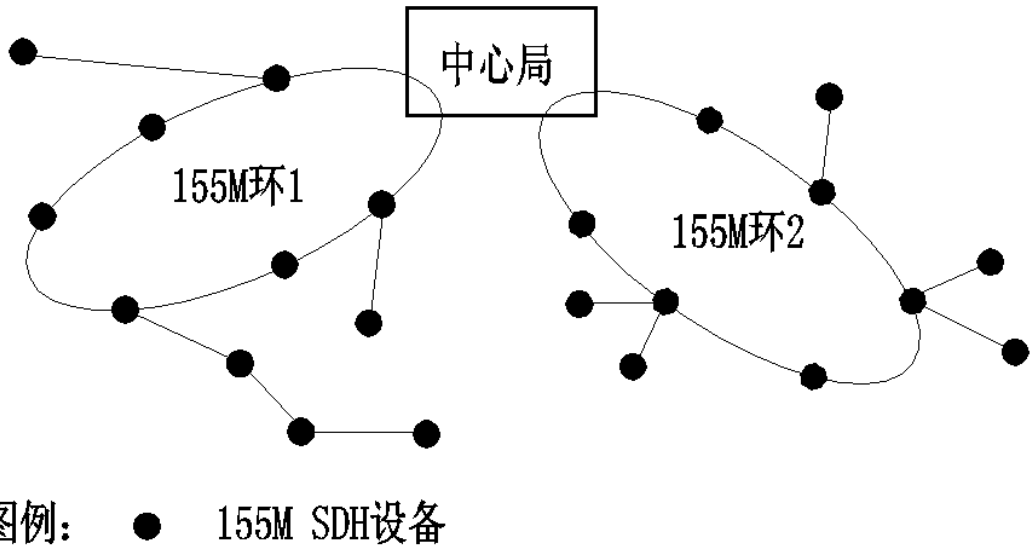
一般来讲，接入层的业务通常为集中型业务，业务往往都集中于核心/汇聚节点，这种情况下两种保护容量相同，通道保护的倒换时间较短，所以在接入层一般选用通道保护。

核心汇聚层的业务通常为分散型业务，各节点都可能承担某个业务网的核心节点功能，业务流向比接入层复杂一些，这种情况下复用段保护的容量优势比较突出，所以核心汇聚层一般选用复用段保护。

1.7.2 SDH 组网

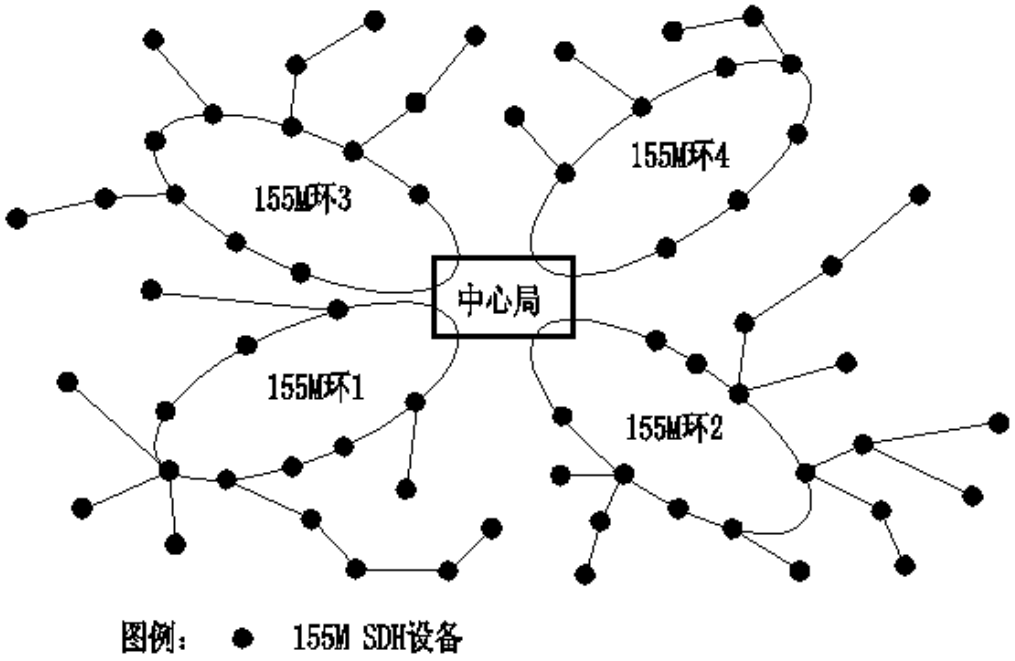
传送网组网结构主要有 3 种：环状、链状、网状，实际上传送网从 20 世纪 80 年代发展至今近 30 年的时间，组网结构上没有太大变化，无论 SDH、DWDM、分组传送网组网结构一直以环状或环带链结构为主，而网状结构由于对光缆线路要求较高，且需要使用控制层面进行业务调度，应用较少。

SDH 大约从 1996 年左右开始在我国规模商用，当时由于无线网络的站点较少，传送网的规模也相应较小，SDH 发展初期网络结构是这样的：



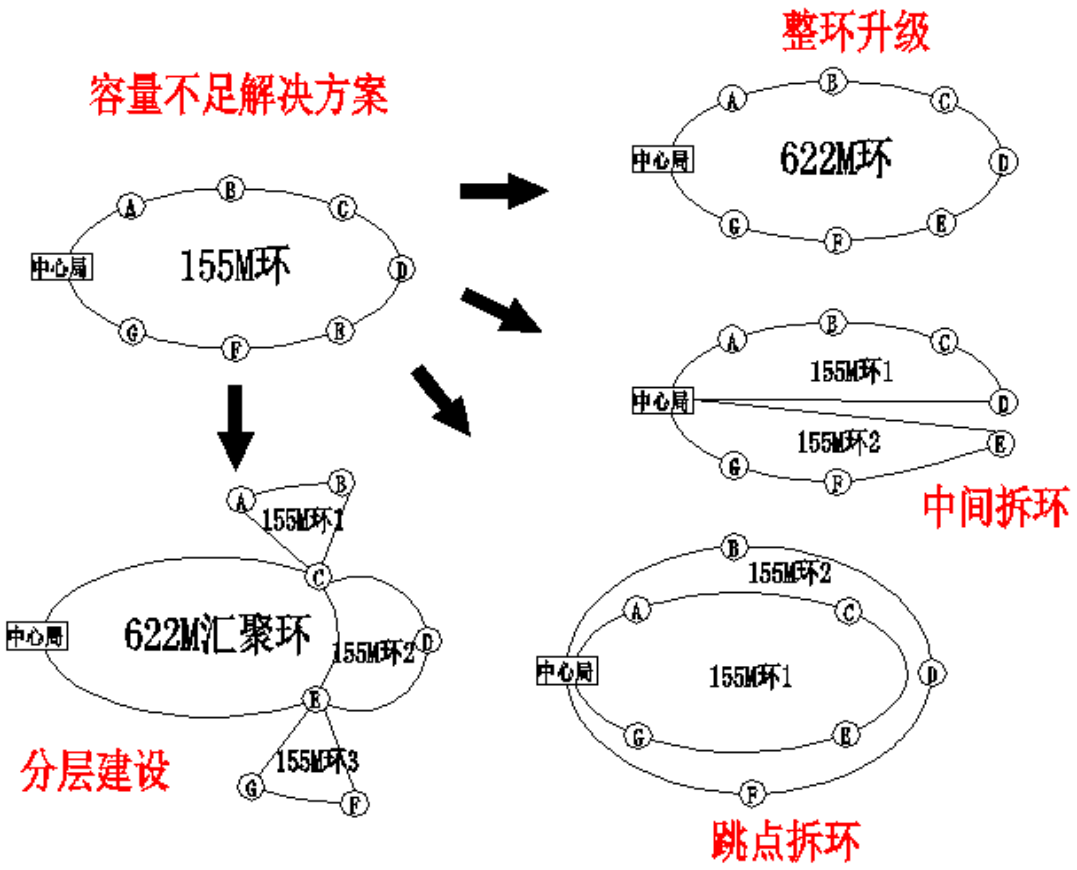
这个时期的无线网络还在 2G 阶段，每个基站的带宽需求为 1 个 E1，一个 155M 的环路按照 30%的带宽预留，只考虑无线业务大约可以带 40 个基站，足够满足业务需求，所以早期的网络以 155M 为主。

随着无线网络的发展，无线基站逐步增多，SDH 网络也渐渐壮大起来。



随着无线 GSM 1800M 基站的建设和大客户专线业务的接入，承载在 SDH 网络上的业务也逐渐多元化，单站的带宽需求逐步从 1 个 E1 发展到 2-4 个 E1 甚至更多。

此时，SDH 网络面临了两个问题，一是环路容量不足的问题，二是环路越来越多导致中心局的光口和入局光缆越来越多，中心局的压力很大。面临这两个问题如何去解决呢？首先，问题一的解决方法有如下几种:



方法一：环路升级。将容量不足的环路升级为 622M，容量提升为原来的 4 倍，且不额外占用光纤，但是环上每个节点都要扩容 622M 光板，投资较大。

方法二：拆环。将环路拆分为 2 个环路，容量增加 1 倍。按照具体实施方法可以分为跳点拆环和中间拆环，但是跳点拆环需要额外占用一对纤芯，中间拆环也需要具备光缆路由。拆环后各站点仍是 2 个光方向，所以只需要在中心局增加光板，投资较低。

方法三：增加汇聚层，也就是我们经常说的分层建设。

在站点中选择业务量较大且相互位置较分散的站点升级为汇聚点，设备升级为 622M，汇聚点组成 622M 汇聚环路，其余站点根据光缆路由下挂在汇聚点下作为接入环。这种方法需要额外占用纤芯，需要给中心局和汇聚点新增光板，投资较低，系统容量是原来的 4 倍。

分层建设有如下几个优点：

1、 能够解决上面说的问题 2，能够缓解中心局压力，中心局只与汇聚节点成环，接入点均下挂至汇聚点之下，和中心局不直接相连；

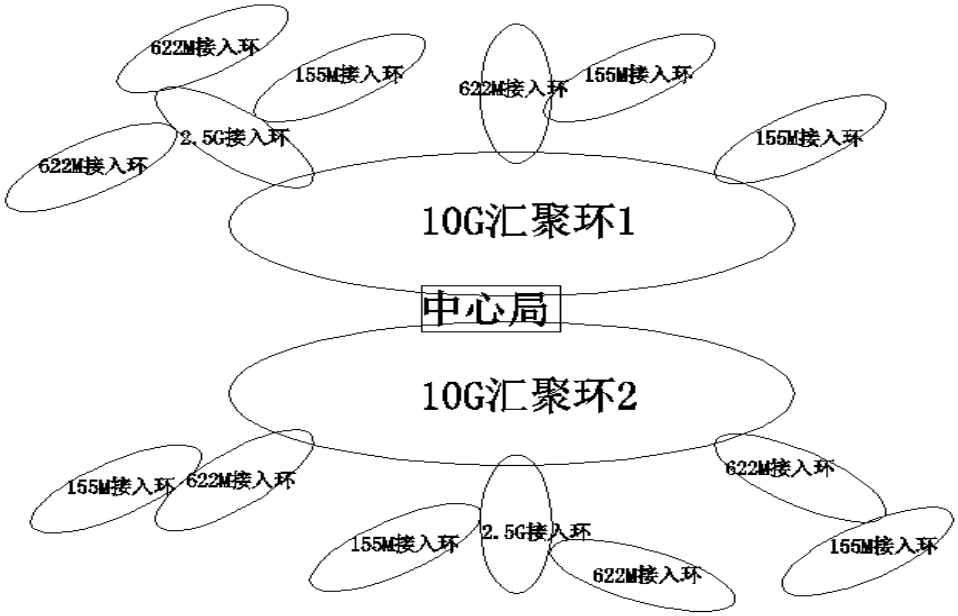
2、 投资较小，网络容量提升大（4 倍）；

3、 按照汇聚点的分布，对接入点进行分区汇聚，避免接入层环路过大导致大量迂回路由，可减少光缆资源消耗；

4、 后期扩容灵活，后期可针对容量不足的接入环单独升级改造，不需整环升级，影响面较小。

实际上传送网的分层建设的思路在生活中比比皆是，上到国家小到公司均采用这种层层管理的方式，国家的行政区域划分为省、市、县、乡镇、行政村的分层结构，道路按照层面也分为国道、省道、县道几级，以道路为例，如果不分层建设道路，省、市、县、乡镇、村都修一条路到北京，北京的交通肯定不堪重负，省、市、县的车流量都在同一条道路上跑，如果道路拥堵就要整条道路拓宽改造或新建，那也是巨大的工程。分层建设的话，北京只通过高

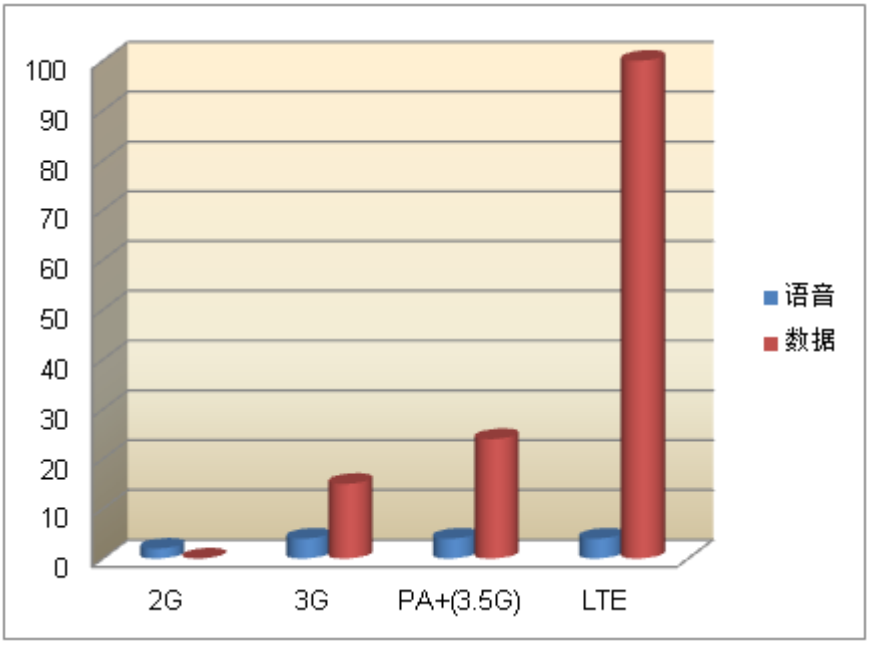
速、国道与一些省相连，省内各市通过省道相连，这样某个省道路拥堵就不会影响其他省的交通状况。
传送网形成了分层的结构之后，网络架构就相对稳定了，接下来传送网的建设只是根据无线业务分布区域调整或新建汇聚环路，根据无线新建站点配套建设接入层环路，对容量不足的汇聚和接入环路进行升级，发展到今基本就是下面图这个样子（本图为示意图，末端支链未体现）：



关于 SDH 的网络优化这部分内容可以重点了解一下，因为这些方法带有普遍性，无论传送网技术发展到什么程度，这些方法都是适用的。

1.8 MSTP 多业务传送平台

2009 年各运营商开始大力建设 3G，3G 与 2G 的最大不同就是数据业务的暴增。虽然 2G 时代也可以低速上网浏览一些网页 QQ 之类，但实现技术上说（GPRS 和 EDGE）都是无线侧利用语音通道实现的，对于传送网来说传送的还是 E1 业务。
3G 时代，手机上网速度明显快了，看视频听音乐下载电影都可以实现了，从单站带宽上说，3G 单站带宽达到 20M 左右，到了 HSPA+时代单站带宽达到 50M 左右，LTE 阶段更是达到了 100M 以上。这几十倍于 2G 时代的带宽使传送网悄然面临一场脱胎换骨的变革。要说起这场变革，还是要从语音业务和数据业务说起。



前面书说过，语音 TDM 业务的帧无论是 64K 的话路还是 E1 电路，都是固定的每秒 8000 帧的帧频，这是语音业务的特点，因为无论你说不说话电话都在每秒 8000 次的抽样，大家拨通电话时就占用了一个时隙，这个时隙一直为你保持，直到你挂机通道才撤销给别人使用。

数据业务和语音业务的特点截然不同，我们都有这方面的经验，无论是宽带还是手机上网，产生的数据流量和你的上网的行为是有关的，在你看视频下载电影的时候数据流量高，但是在浏览网页、聊 QQ 的时候数据流量很低，在你不上网的时候，虽然你的宽带或 3G 数据保持连接，但是基本并没有流量产生。数据业务发送的是数据包，数据包的大小和发送时间是不固定的，在没有数据的时候只发送一些信令，例如设备间打打招呼之类。

这里把语音业务比作 A 公司，每隔一小时就需要往火车站发一个固定大小的货物，A 公司可以和地铁公司签订协议，买断地铁的一个空间单位（时隙），这个空间在 A 公司和火车站之间是属于 A 公司所有的，其他业务不得占用，这样的运输的机制就完全适合 A 公司的需求；如果把数据业务比作 B 公司，发货时间、货物大小都不一定，那么 A 公司的运输机制一定不适合 B 公司，那个预留的空间可能多数时候都是空闲的，但是在 B 公司业务多的时候可能又装不下，那 B 公司应该选择出租车，货物多的时候可以打多个车。

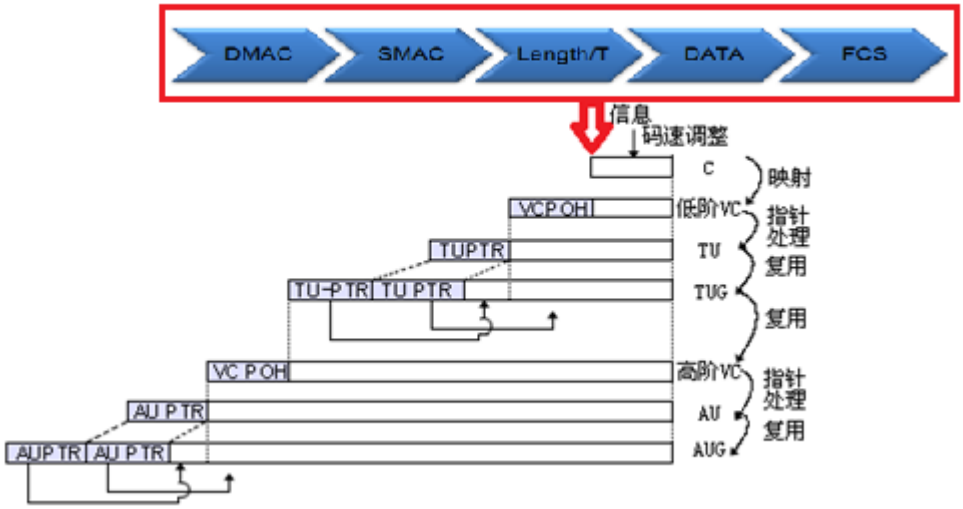
上面的例子可以看出，不同的传送机制针对不同的业务的传送效果是不同的，

A 公司的运输机制就是 SDH，B 公司就是分组交换。那么 B 公司是否可以采用 A 公司的机制呢？当然是可以的，虽然 B 公司闲的时候没有业务，忙的时候假设最多需要有 100 个空间单位，那就把这 100 个单位全部买断下来，肯定够用了，只是浪费比较多。

既然这么浪费，为什么还使用这种方案呢？因为 B 公司之前已经买断了地铁 50 个空间单位，并与地铁公司签订了长期合作发展的协议，容量不够的情况下再购买 50 个空间的投资不大，也许比租车的费用还要节省一些。运营商在考虑采用什么技术承载的时候，承载效率只是一方面，另外一方面就是保护建设投资，将两种方案的投资和建设效果对比一下就知道那种方案更优。

毕竟 SDH 网络建设十多年来，投入了大量的建设成本，任何事物发展都有个过程，无线数据业务的发展也是逐步增长的，在 3G 网络建设初期，单站数据带宽大概 8-15M 左右，1 个 622M 的接入环也可以带 30 个基站左右，这个阶段 SDH 技术还不至于直接淘汰，还应该让这张网继续发出它的光和热，从技术上可以对 SDH 设备做一些改进，使其支持以太网业务接口，实现对数据业务的承载。

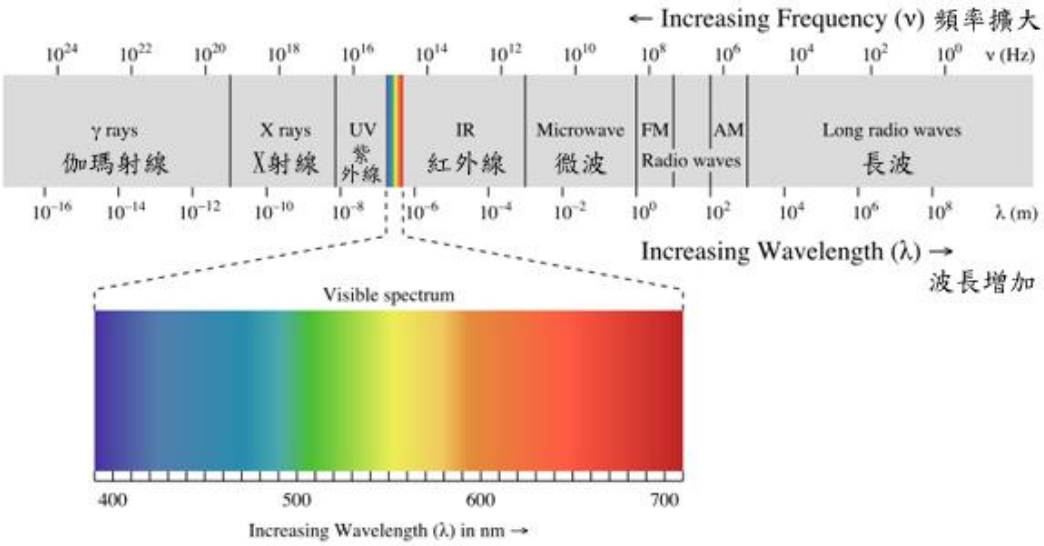
前面说过，业务的接入要有一个标准的接口，对于数据业务来说这个接口叫做以太网接口，3G 时代使用最大速率 100M 的 FE 接口（快速以太网），LTE 时代因为单站带宽超过 100M 所以使用 GE 接口（千兆以太网）。可是，SDH 提供的接口中并没有对以太网接口的支持，要在 SDH 中承载数据业务，需要对 SDH 进行升级改造，于是就有了这个新的名词——MSTP。
MSTP（Multi-Service Transfer Platform 多业务传送平台）是指基于 SDH 平台同时实现 TDM、以太网等业务的接入、处理和传送，提供统一网管的多业务节点。
那么 MSTP 是如何在 SDH 基础上支持以太网业务的接入呢？ SDH 通过 FE 或 GE 接口收到以太网数据，将完整的一帧缓存后，通过一种叫做通用成帧规程（GFP）的协议，对应封装到 SDH 的 C12、C4 等容器里，剩下的步骤就是 SDH 体系的那一套，层层的封装成 STM-N 在线路上传送，到了收端再将以太网帧数据从容器 C 中取出，发送给业务侧。



FE 接口的最高带宽是 100M，但是从以太网帧结构可以看出帧的数据内容长度是可变的，帧长度范围是 64-1518 字节，而且以太网帧没有固定的帧频，FE 接口的实际数据流量从 0 到 100M 皆有可能，那 SDH 给这个业务预留多少带宽可以取决于这个接口的最大带宽，一个基站的数据流量有一个峰值带宽比如是 10M，SDH 就事先为这个基站预留 5 个 E1 的通道，在实际数据流量小于 10M 的时候就插入空闲帧。实际应用中，为了保证带宽的使用效率，也可以按照以太网业务的平均流量来配置带宽，这样在高峰期流量超过均值的时候，就会发生拥塞。

1.9 传输距离计算

我们上学时都接触过下面这张光谱图，我们肉眼可见光部分波长范围是 390¯760nm，大于 760nm 部分是红外线，小于 390nm 部分是紫外线。光谱中不同波长的光有着各种不同的特性（穿透力、能量、折射率、杀伤力等），使其可以应用到各个领域中，提起各种射线，总让人不自主的想起医院里那些冰冷神秘的设备，一种莫名的不适感涌遍全身。

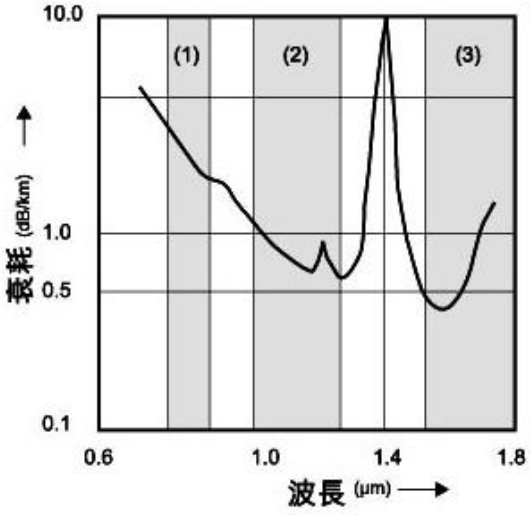


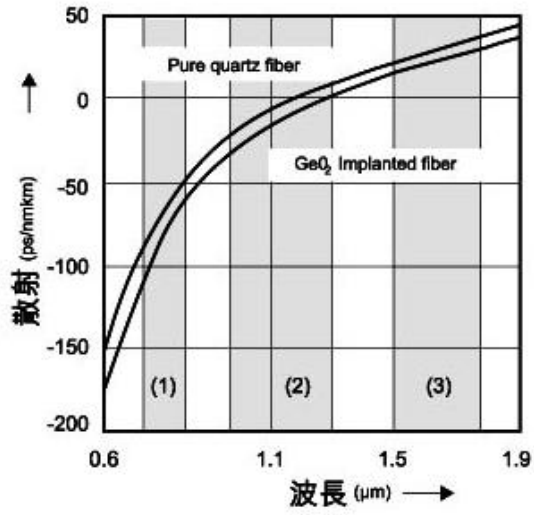
我们都知道，光纤通信是利用光的全反射的原理，如果入射光大于一定角度，入射光在纤芯和包层的分界面会全部被反射，在光纤中不停的“碰壁—反弹”直至到达另一端。

如果站在光通信的角度，我们关注的是各种波长的光在光纤中传送的特性：衰耗、色散。衰耗是指发送的光信号经过光纤传送之后功率的衰落，也就是信号由强变弱，当衰落到接收端无法正确识别便产生误码或中断；色散是指不同波长的光信号在光纤中的传播效应不同导致的信号脉冲变形，产生误码。从脉冲信号的形状上来看，衰耗是脉冲由高变矮了，色散是脉冲由瘦变胖了。

在光谱中不是所有的频段都适合光传输的，要考虑不同频率光信号的色散和衰减性能，选用合适的频段才能使信号传的更远。不同波长的光在光纤中的衰耗和色散曲线不同，下图中，阴影部分（1）（2）（3）是光通信应用的波长窗口，可以看出，衰耗较小的波长主要集中在 1310nm 和 1550nm 左右两个窗口上（850nm 衰耗较高，一般用于多模光纤），而色散是随着波长的增加而增加的。

1310nm 信号的衰耗较大但色散较小，因此一般称为零色散窗口；1550nm 附近信号的衰耗较小但色散较大，称为低损耗窗口。1310 窗口光信号的零色散特性可以通过某种神奇的技术（利用光纤材料中的石英材料色散与纤芯结构色散的合成抵消特性￥*%*@￥#…！））位移到 1550 窗口附近，因此长距离的传输一般使用 1550nm。





下面我们分别对于衰耗和色散，来说明一下光信号的最大传输距离的计算

衰耗受限传输距离计算

衰耗对传输距离的限制由发光功率、接收灵敏度、线路损耗决定，这个比较容易理解，就像你隔一定距离向我喊，声音能够传多远有三个因素，第一是你那边声音的大小，第二是我这边耳朵灵敏程度，第三是传播途径的风和噪音大小，也就是源、宿、传播介质三要素。
衰耗（增益）的单位是 db，db=10logX，其中 X 是衰耗的倍数。为什么用 db 来表示呢？因为信号的衰减和增益可不是增减少多少瓦（W）这么小家子气的，动辄就是 10 的 N 次方倍的数量级，如果用绝对值 W 来表示会非常不直观，就需要数小数点之前或之后 0 的个数才能知道衰耗（增益）了多少倍。用对数来表示就简单很多，比如增益（衰耗）10 倍就是 10db（-10db），100 倍就是 2 个 10 倍就是 20db，增益 1000000000000000 倍就是增益 150db，是不是用 db 表示清爽了许多？这里强调一下，db 之间是相加减的关系，不能乘除。
下面我们看一下衰耗受限的传输距离计算公式:

$$L=(Ps-Pr-Pp-C-Mc)/(af+as)$$

$$L=(Ps-Pr-Pp-C)/(af+as+\triangle Mc)$$

这两个公式差不多，唯一区别就是 Mc 的计取方式不同。公式中:
L — 再生段最大距离(km)
Ps — S 点寿命终了（EOL）最小平均发送功率(dBm)
Pr — R 点寿命终了（EOL）最差灵敏度(dBm)
Pp — 光通道代价，它包括反射、码间干扰等产生的总色散功率代价。一般在 1310nm 波长时取 1dB，在 1550nm 波长时根据传输距离的长短分别取 1dB 或 2dB。
C — 所有活动连接器衰减之和，每个连接器衰减取 0.5dB，共两个连接器。
Mc — 光缆富裕度，光纤长短不同取值不同，最大取值为 3dB。△Mc 单位为 dB/km，一般为 0.02～0.03 dB/km，Mc 和△Mc 的区别是将富裕度一次性计取，还是分摊的单公里线路中。
af — 光纤衰减系数（dB/km），与工作波长密切相关。在一定工作波长上，光纤的衰减为一定值，不随传输信号速率的高低而变化。在 1310nm 窗口一般在 0.35 dB/km 左右，1550 窗口一般在 0.25 dB/km 左右，具体值与光纤的质量有关，应以实际测试的为准。

as — 光纤熔接接头每公里衰减系数（dB/km），与光缆质量，熔接机性能，操作水平有关。工程中取 0.01～0.02dB/km。

这个公式看似有些复杂，其实很简单，就是用发光功率减去接收灵敏度，得到线路上可以供衰耗的容限，这个容限再减去接头损耗、富裕度等固定要发生或预留的值，剩下的就是可以在光纤中容忍的衰耗值，用这个值除以每公里光纤的衰耗，得出的就是最大传输距离。

当衰耗受限距离无法满足传送要求时，需要使用 EDFA（光纤掺铒放大器），进行衰减补偿。光放大板分为三种功率放大板 OBA、前置放大板 OPA 和光线路放大板 OLA。在 SDH 工程中，一般只使用 OBA 和 OPA，OBA 的作用是提高发送端的光功率，也就是增大公式中的 Ps；OPA 的作用是提高接收端的灵敏度，也就是降低公式中的 Pr。

我们工程中配置的 SDH 光接口有各种各样的型号，例如 S1.1、S4.1、L16.2、L64.2 等等，其中 S 和 L 代表光口类型是长距还是短距，决定了发光功率 Ps 和接收灵敏度 Pr，这个参数有相关的标准，可以在设备的技术资料中查到；后面的数字 1、4、16、64 就代表 STM-N 中的 N，代表了光口的速率。小数点后面的 1 和 2 代表了工作波长，1 表示 1310nm，2 表示 1550nm，这个数字决定了光通道代价 Pp 和光纤衰减系数 af。线路的衰耗值一般可以经过实际测试得出，也可以根据光纤衰减系数来估算。

而分组设备的光模块一般按照 10km、40km、80km 去标注，看起来直观了一些，实际上给出的公里数也是一个参考值，如果要准确的计算传输距离，使用的方法和上述介绍的相同。

色散受限传输距离计算

色散其实说起原理较为复杂，但是计算色散受限传输距离要容易的多，色散受限的传输距离计算公式：

Ld= ε / Dm（这个公式应该不用解释）

其中：Ld—传输距离

ε —光源色散容限

Dm —每公里色散值

色散值的单位 ps/(nm*km)，含义是单位波长间隔内各波长成分通过单位长度光纤所产生的时延，色散容限值单位是 ps/nm，这个值由光源决定。

举个例子，ε=1600 ps/nm，Dm=20 ps/(nm*km)，那么 Ld 就等于 1600 ps/nm /2020 ps/(nm*km)=80 公里，就是说如果考虑光纤的色散效应，信号最多可以传 80 公里。超过 80 公里就需要色散补偿，色散补偿就相当于补偿 ε，一般色散补偿模块的规格直接用公里数表示。

设备的最大传输距离必需同时满足上述两个主要受限因素，传输距离值遵循木桶理论，决定于两个因素受限距离的最小值，传输距离不满足时，哪个因素受限就需要做相应的补偿。

OTN 的传输距离计算相比 SDH 和分组网要复杂的多，本人未参与过相关的工作，OTN 是在新建系统时将每一段的参数调整好并留有一定余量，基本是一劳永逸，工作中一般也涉及不到 OTN 的传输距离计算，需要了解请查阅相关资料。

第二章　　波分复用（WDM）

2.1　　波分复用基本概念

随着业务的发展传送网经常面临着容量不足的问题，问题的解决方法前面提到过一些，可以升级系统速率，可以对环路进行拆分优化，也可以新建一个传输系统来解决。

首先说升级速率，155M 可以升级 622M，2.5G 可以升级到 10G，可是速率升级总有个上限，目前来说无论 SDH 还是分组传送网，大规模商用的最高速率基本还是 10G。

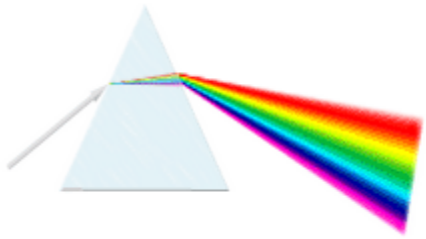
一个技术从提出到讨论到标准制定成熟，再到经过测试、试商用、商用，最终规模的应用在各大运营商需要几年至十几年很长的周期，虽然 40G、100G 技术都已经成熟，但是离大规模的应用还有一定的距离。光器件本身的物理特性也决定了传输速率不可能无限的增长，到了一定的极限再去突破的难度会成指数级提高，在无法解决一些技术壁垒的时候，就只能想别的办法。

环路拆分优化和新建，等于依靠增加系统的数量来提高容量，貌似可以无限的复制，但是有一个无法回避的现实的问题，就是对光缆纤芯的使用成倍的增加，在光缆已经没有纤芯资源可利用的时候，需要新建光缆会遇到两个问题，一是建设的可行性，二是建设的成本，可能会有各种问题或矛盾导致建设不被允许，即便可以新建，建设的成本可能由于光缆的长度、建设的难度等因素而非常高昂，说白了就是有关部门不让建或者太贵了建不起的问题。

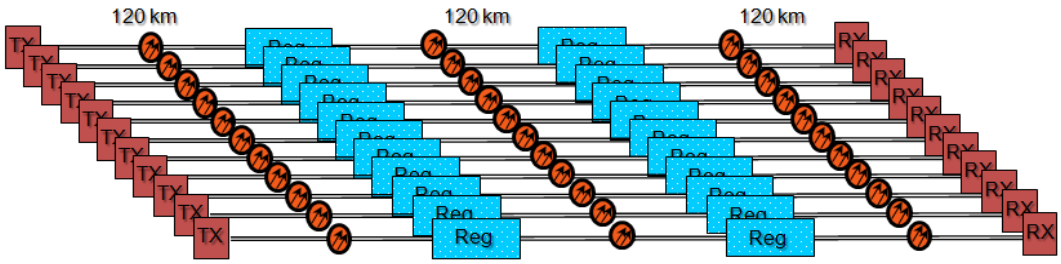
在上述的方法都行不通的时候，利用波分复用技术可以迎刃而解。波分复用可以大幅的提高单根光纤中的传输速率，当然设备价格也是非常高昂的，不过运营商会算这笔账，只要比新建光缆便宜，或者可以解决其他手段解决不了的问题，就是值得的。

----波分复用你这么神奇，你麻麻造么？

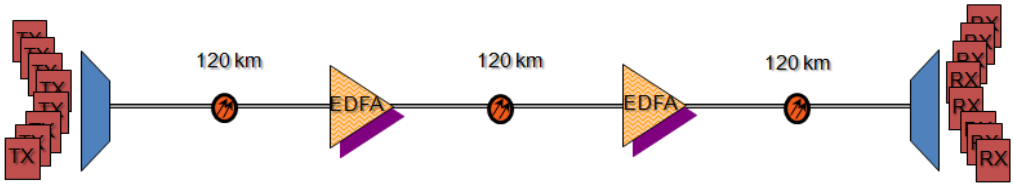
----我麻麻不但神奇，还很漂亮呢，我麻麻名字叫彩虹！



我们大家都知道，利用三棱镜可以把一束白光分成红橙黄绿蓝靛紫七色光，相反的七色光按照一定的角度入射，通过三棱镜也可以合成一束白光。如果七色光各自代表着不同的信号源，那通过类似三棱镜的器件（合波器）就可以将多路信号合并到一起，这样就可以做到在一根光纤中传送多路信号，如果单路信号的最大速率是 10G，合波器可以合路 80 个波长，那么一根光纤传送的速率就提高了 80 倍，达到 80*10G=800G，波分复用的原理大致就是如此。

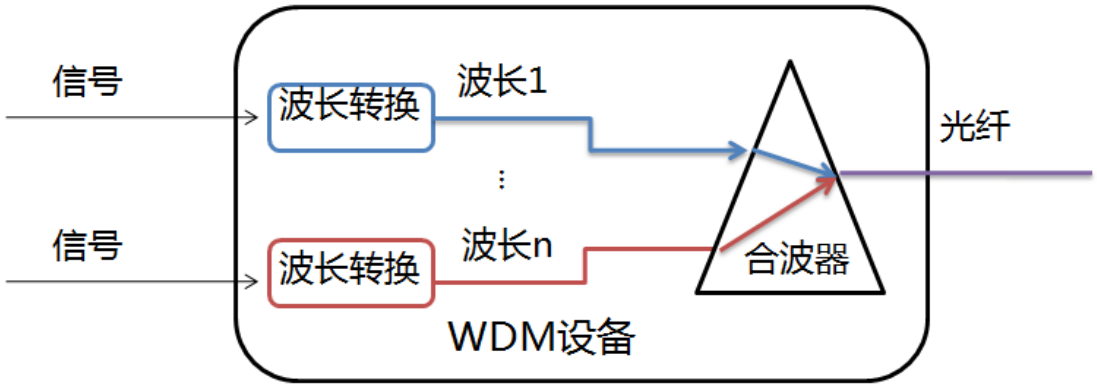


SDH系统 每系统占用2芯光缆



DWDM系统 80系统只占用2芯光缆

波分复用（ Wavelength Division Multiplexing ）技术是在一根光纤中同时传输多个波长光信号的一项技术。传送网的 MSTP 和分组传送网技术，都是在一根光纤中传送单个的波长，速率的提高是靠缩短单 bit 信号的时间来实现，越高的速率意味着 bit 信号之间就越拥挤。而波分复用在一根光纤中传送的多个波长之间是互不相干的，就像我们的广播电台在空中传送多个频道，频道多传递信息量就大。这样我们不用费劲心思在 SDH 的车厢里尽量多塞一些人，只要引进双层（多层）巴士，容量解决了，大家也不用挤的那么辛苦。
波分复用和频分复用基本差不多，只是传输的介质不同，下面这张图直观的表示了波分复用的原理。



下图是用于光通信的波段和波长的对应关系，可以看出能够适用于长距离大容量传输的波段资源是非常有限的，波道之间的间隔就决定了对于资源的使用效率，频率间隔越小则可以复用的波长就越多。根据波长间隔的疏密，波分复用可以分为 DWDM（密集波分复用）和 CWDM（稀疏波分复用，简称粗波分）。

DWDM 的波长间隔为 0.4nm 或 0.8nm 左右，使用 C 波段和 L 波段。一般的 40 波系统就是采用 C 波段波道间隔 0.8nm，而 80 波系统一般使用 C 波段波道间隔 0.4nm，波道间隔缩小一半，可容纳波道数量就增加一倍。下表是 ITU-T 建议的 C 波段波道 80 波的频率图，将 80 波中的偶数波去掉剩下的就是波道间隔 0.8nm 的 40 波系统，我们将 40 个奇数波道和 40 个偶数波道分别称为 C 和 C+波段。

波道编号	中心频率（THz）	波道编号	中心频率（THz）	波道编号	中心频率（THz）	波道编号	中心频率（THz）
1	192.1	21	193.1	41	194.1	61	195.1
2	192.15	22	193.15	42	194.15	62	195.15
3	192.2	23	193.2	43	194.2	63	195.2
4	192.25	24	193.25	44	194.25	64	195.25
5	192.3	25	193.3	45	194.3	65	195.3
6	192.35	26	193.35	46	194.35	66	195.35
7	192.4	27	193.4	47	194.4	67	195.4
8	192.45	28	193.45	48	194.45	68	195.45
9	192.5	29	193.5	49	194.5	69	195.5
10	192.55	30	193.55	50	194.55	70	195.55
11	192.6	31	193.6	51	194.6	71	195.6
12	192.65	32	193.65	52	194.65	72	195.65
13	192.7	33	193.7	53	194.7	73	195.7
14	192.75	34	193.75	54	194.75	74	195.75
15	192.8	35	193.8	55	194.8	75	195.8
16	192.85	36	193.85	56	194.85	76	195.85
17	192.9	37	193.9	57	194.9	77	195.9
18	192.95	38	193.95	58	194.95	78	195.95
19	193	39	194	59	195	79	196
20	193.05	40	194.05	60	195.05	80	196.05

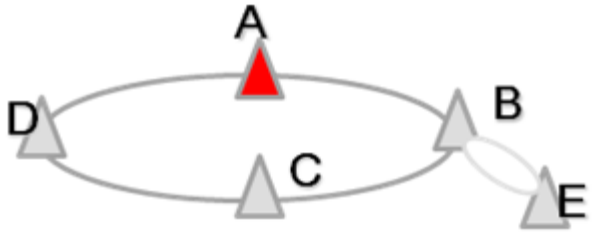
在 80 波不能满足容量要求的时候可以使用 L 波段，原理同 C 波段相同。

CWDM 的频率间隔为 20nm，使用 O、E、S、C、L 五个波段，波道数支持最大 16 个。

由于密集波分的波长间隔小（ITU-T G.692 建议 WDM 中心波长的偏差不大于信道间隔±20%），而温度对于波长的稳定性影响较大，所以密集波分需要使用冷却激光器和温度控制功能；而对于粗波分来说，温度变化导致的波长漂移仍然在容许范围内，激光器无需温度控制机制，另外密集波分的合分波器的工艺要求也高于粗波分，因此密集波分的成本要大于粗波分。

粗波分由于波道容量较小，一般应用于传送网接入层，而密集波分目前大规模应用于各大运营商的省际干线（一干）、省内干线（二干）和本地网核心汇聚层。

DWDM 系统常用组网结构为环形或环+链结构，不具备光缆双路由的边远县城、乡镇节点一般以链型方式接入到 OTN 环路中。

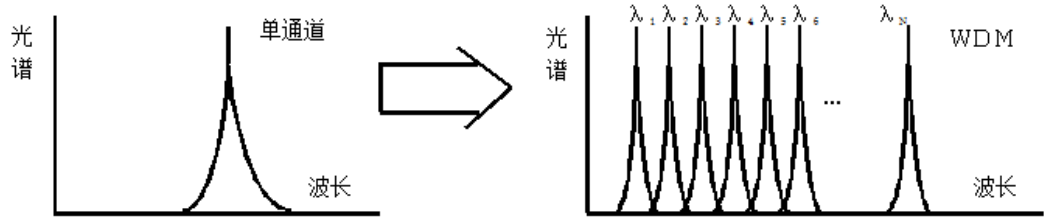


2.2 WDM 系统组成

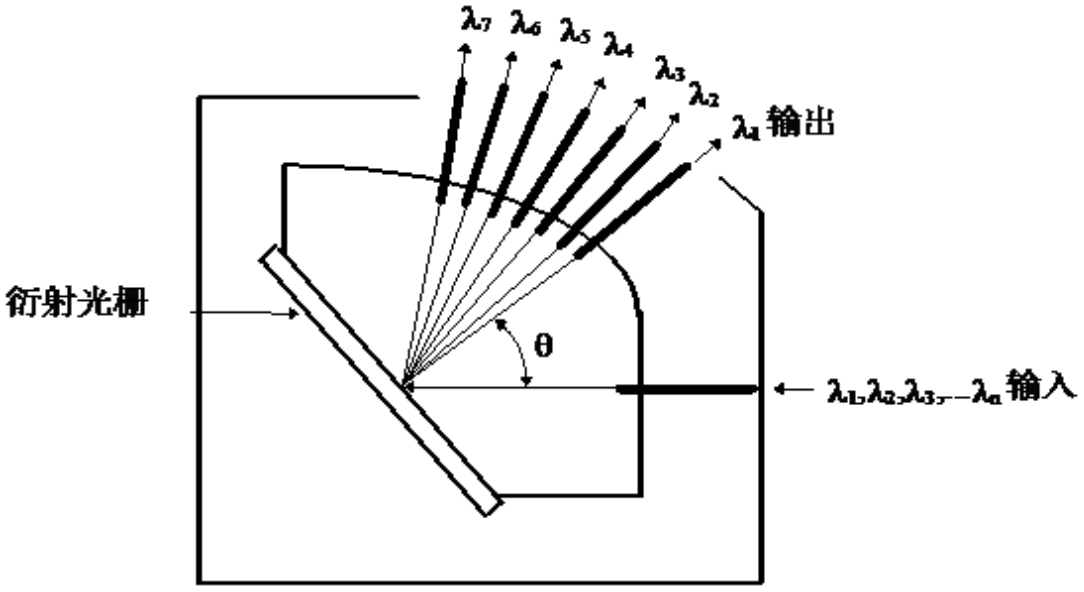
波分系统如何将各路信号合并到一起传送的?这一节我们介绍一下 WDM 系统每一部分负责的功能。

首先，我们业务侧的信号都是频谱较宽的单波长信号，我们通常称之为白光。合并之前需要将信号转换成系统规定的精准波长的信号（彩光），所以首先需要经过具有波长转换功能的光转发单元（OTU），OTU 将客户侧信号转换为电信号之后，再通过固定或可调波长的光模块转换为规定的波长，实现了实现任意波长光信号（如 G.957）到满足 G.692 要求的波长转换的功能。

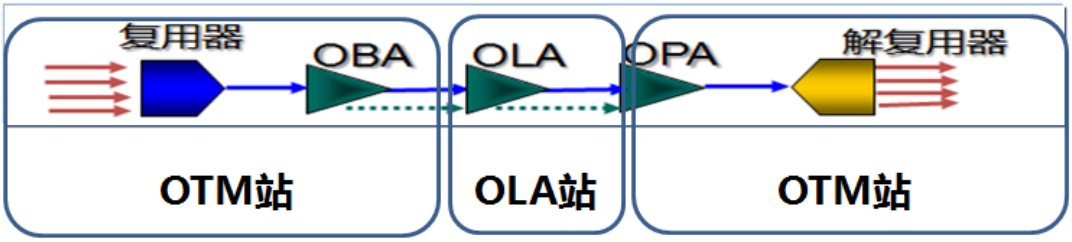
就像原本你一个人很霸气的睡在一个大床上（单波信号），现在要在这个床上挤 N 个人，那就得拜托你收敛一下，不能再占那么大地方，才能将这些人安排到固定而拥挤的位置上（波分复用），而这个拥挤是指频率上的，和时分复用是不同的。



经过 OTU 的各路光信号进入合波器合并到一起，到了对端再通过分波器将线路中的信号分离出来。合分波器是无源光器件，可分为衍射光栅型、棱镜型、波导型等几种类型。



由于信号在合波的过程中有一定的衰耗，为了满足信号的长距离传输，合路的信号在送入光纤之前需要经过光功率放大器（OBA）进行放大，在到达接收端时信号已经衰耗的很厉害，为了能够被接收端识别需要经过光前置放大器（OPA）进行放大，在发端和收端距离较远时需要经过一个或若干个光线路放大器（OLA）进行放大。



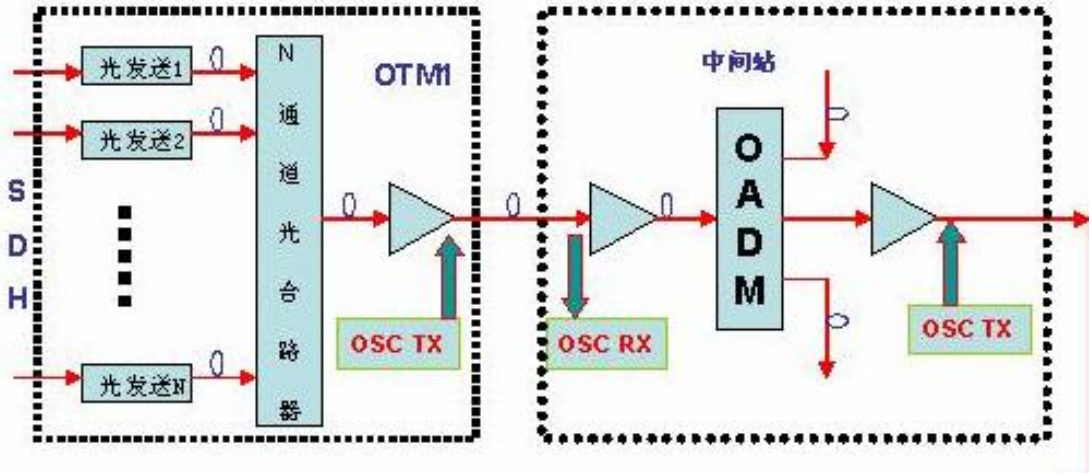
举个例子，如果一个车子要跑长途，那么 OBA 就像在出发前给车子先加满油，OLA 就像中途没油了进入加油站加油，而 OPA 就相当于到了终点再加一次油。如果一箱油就可以跑到终点，就不用中间进加油站（OLA）。

OLA 是作为一个独立的站点存在的，而 OBA 和 OPA 是插在发端和收端站点波分设备机框内的光放大板。不同放大器由于处在网络中的位置不同，输入光功率和增益性能也不同。OBA 是处于发送端，这时信号主要是由于经过合波器造成相对较小的衰耗，所以 OBA 输入光功率比较高增益比较小；而在收端和 OLA 站点，信号经过了“长途跋涉”已经几乎衰减到了临界点，所以 OPA 和 OLA 的输入光功率很低；OLA 放大后还需要把信号再次送上线路传送，而 OPA 只需要进行一些放大让收端能够识别就可以，所以 OLA 增益比 OPA 大。

WDM 系统按照站型分可以分为光终端复用站（OTM）、光分插复用站（OADM）、光线路放大站（OLA），在 OTM 站所有波长都全部上下，OLA 站点只放大信号不上下波长，OADM 站点可以上下 N 个波长，其他波道业务直通过去。实际上 OADM 的波长穿通一般也是物理连纤实现的，就是从西向分波器下来，再用光纤连到东向合波器上去，本质上与 OTM 没有区别，所以我们可以理解为，WDM 网络都是由 OTM 和 OLA 站组成。

为了使管理与监控信息不依赖于传输的业务，波分系统使用一个单独的信道来管理 WDM 设备，这个信道就是光监控信道 OSC。

G.692 规范的带外 OSC 使用的标称波长为 1510nm，数据速率取为 2Mbit/s，OSC 得不到 EDFA 的放大，靠低速率下高的接收灵敏度（优于-50dBm）仍能正常工作。但必须在 EDFA 之前下光路，而在 EDFA 之后上光路，就是说你坐着另外一辆小车去监控这辆大车，而不是身在此山中，不然光放坏了，监控信号也就跟着挂了。



2.3 光传送网 OTN

DWDM 技术的应用给传送网带来了质的飞越，优点主要有以下几方面:

1、 单光纤的可传送带宽得到了极大的提升，在业务流量较大的干线、本地网核心汇聚层节省了大量的光纤资源。

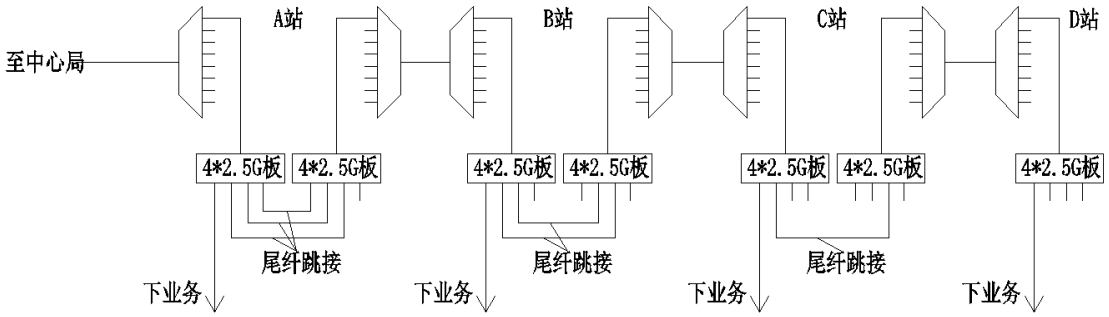
2、 扩容便捷性大大提高，传送网系统的扩容只需要增加一些波分板件就可以实现，相比较敷设光缆来说大大缩短了建设周期

3、 波分系统的无电中继距离能够达到几千公里，大大提高了传送网的传输距离。无电中继距离就是信号在途中不经过光电转换，只靠对光信号的放大、补偿整形能够达到的传输距离。如果使用电中继理论上的传输距离是无限远。

可是最初的 DWDM（相对于 OTN 我们一般称之为传统波分）只是简单的将各路信号变换波长后合路传送，到接收方解复用，于是在传统波分的应用过程中就逐渐暴露了一些问题:

1、 业务调度不灵活。DWDM 系统只对信号的波长进行转换，不对信号的帧结构进行处理，也就是说每个 10G 波道的内部帧结构 DWDM 系统是看不到的，所以传统的 10G 波分系统就支持 10G 的客户信号，虽然可以通过 TMUX 单板将 4 路 2.5G 信号合并为 1 路 10G 信号，但本质上相当于将 SDH 的复用功能内置到了 DWDM 设备中，实际对 DWDM 来说还是透传 10G 信号。即便如此传统 10G DWDM 系统也只能支持 10G 和 2.5G 业务颗粒，而 GE 等一些小颗粒业务则无法直接接入 DWDM 系统，需要经过 SDH 设备进行低速信号的交叉复用。

2、 仅支持点-点组网结构，传统 DWDM 所谓的环实际上是多个点-点系统组成的，同一站点的不同方向的业务调度只能靠尾纤跳接来实现， DWDM 和 PDH 设备一样，也是一对对背靠背的 TM 组成，组网结构如下图所示:



3、 网络运行维护、管理不灵活，DWDM 系统的监控通道仅为 2M，前面说过 SDH 的 STM-1 的帧结构中各种开销带宽就达到了 20 多 M，而 DWDM 对整个 40 或 80 路 10G 信号的监控仅仅只有 2M，所以可想而知，DWDM 如此低的带宽只能对整个光通道的一些非常重要的指标和性能进行监控。

4、 DWDM 系统保护方式仅支持对光缆线路和单个波道进行保护，因为 DWDM 的最小业务单元就是波道，而对于波道以下的低速信号 DWDM 并不关心，所以自然也无法提供类似 SDH 通道级别的保护，保护方式不够灵活。

我们不难发现，DWDM 的这些问题恰好 SDH 都曾经很好的解决过，SDH 通过体系规定的映射复用方式可以接入并监控各种低速信号，SDH 通过交叉单元实现了一台设备上多个光方向之间的业务调度，SDH 也提供强大的维护管理功能，支持任意级别的通道的快速倒换保护。

那我们不免会想能否有一个新的技术，可以结合 DWDM 大容量的优势和 SDH 的组网灵活、保护完善、管理功能强大的特性，使两者的优点结合起来呢？事实证明鱼与熊掌是可以兼得的，OTN（光传送网）系统就是这样的一个新的技术体制。

OTN 做了几件事:

1、 定义了一系列速率等级和帧结构: OTUk、ODUk、OPUk。OTN 和 DWDM 的最大区别也是在于此，OTN 有了自己的帧结构，基于不同等级的 ODUk 颗粒，就可以实现类似 SDH 的电交叉功能，使小颗粒的信号可以合并在大的通道中传送，OTN 的一个波道中也像 SDH 那样有了大大小小的容器，所以对于 2.5G 以下的低速信号，OTN 从体制上就具备了接入和处理的能力，提高带宽利用效率。而传统 DWDM 只是简单粗暴的将波长合并和分离。

2、 通过 WSS（波长选择开关）等技术实现了波长之间的灵活调度，支持以光波长信号为基础的灵活调度，提升业务调度的灵活性。也就是以波道为颗粒进行交叉，有点像 SDH 里高阶交叉的概念，但是交叉的单元是光信号，所以叫光交叉。

3、 OTN 帧结构中引入了类似 SDH 的丰富的开销机制，强大了网管能力。既然 OTN 有了自己的帧结构，那么顺便规定一些字节用于管理，这也是从 SDH 那的舶来之品。

2.4 OTN 电交叉

OTN 的业务处理分为光层和电层，电层的处理的是 ODUk 的颗粒，而光层的基本单元是单个波道。

电层则将单个波道中包含的不同等级的 ODUk 数据帧进行映射、交叉、复用，光层负责将波道合并、分离，将波长信号在各站点上下、调度。基于 ODUk 和基于波道的调度分别是 OTN 的两大功能: 电交叉和光交叉。光交叉是 OTN 特有的概念，因为 SDH 每个光方向都是单一波长，而波分每个方向上都有多个波长信号，OTN 光交叉可以让这些波长信号不经过光电转换，而在各个方向之间自由的“穿行”。

OTN 电交叉:

首先明确一个问题，对于电交叉这部分功能，既然我们用 DWDM+SDH 的方式也可以实现，为什么要制定 OTN 的帧结构呢?

首先是组网复杂，DWDM+SDH 两套设备自然要占用更多的机房空间、功耗，还要多出许多复杂的设备间的连纤。就像我们现在的智能手机一样，一部手机就可以集成了移动电话+MP3+照相机+游戏机等等功能，谁又愿意带着这么多东西在身上呢，高集成度多功能化，这是科技发展的趋势。

另外，SDH 的交叉颗粒是从 2M-10G，而 OTN 的业务颗粒是 GE 以上到 100G，颗粒度远远大于 SDH，SDH 能够解决的也仅仅是 GE 和 2.5G 颗粒的交叉，对于 10G 以上的颗粒无法支持。即便是 SDH 能够实现的 GE 和 2.5G 的交叉，SDH 上实现的成本也要高于 OTN，就像运输整箱整车的大件货物的话，火车的成本要低于小汽车的成本。

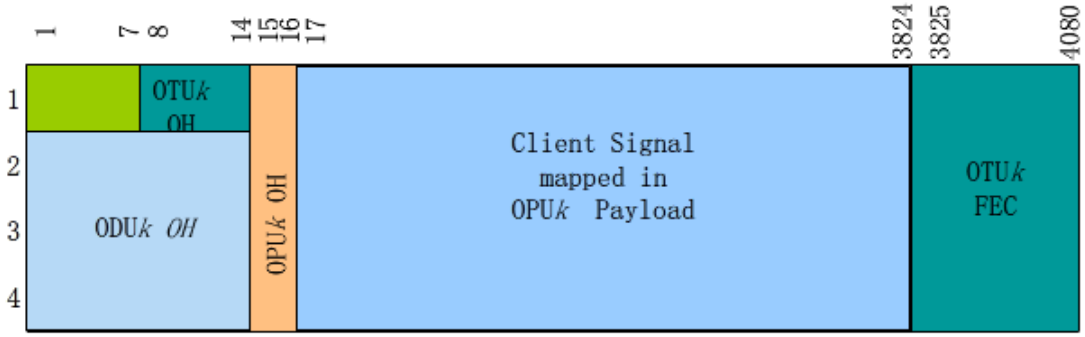
下面言归正传，OTN 在电层规定了一系列的速率等级和容器: OTUk、ODUk、 OPUk。这个 OTUk 就和 SDH 的 STM-N 是类似的概念，区别是 OTUk 的容量和 STM-N 不在一个级别上，ODUk 就相当于 SDH 中的虚容器 VC，OPUk 就相当于 SDH 中的容器 C。

ODUk 是 OTN 电交叉的基本单元，对应的速率和业务类型如下表:

ODUk 等级	ODUk 速率(kbit/s)	对应业务
ODU0	1,244,160	GE
ODU1	2,498,775	2.5G
ODU2	10,037,273	10G
ODU2e	10,399,525	10GE
ODU3	40,319,218	40GE
ODU4	104,794,446	100GE

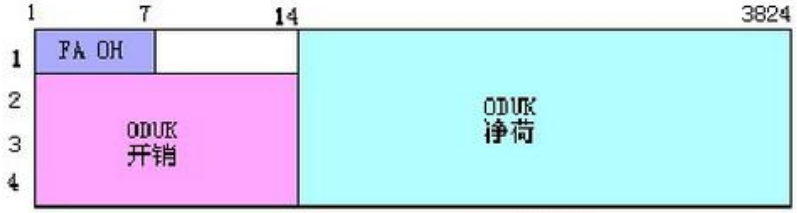
除此之外，为了更灵活适应不同速率的业务颗粒，OTN 还支持 ODU flex，ODU flex 是速率灵活可变的容器，可支持 2.5G 以上的任何速率（1.25G 以下映射到 ODU0，1.25G-2.5G 映射到 ODU1），系统会根据业务速率自动指配相应的 ODUk 组合，速率间隔是 1.25G（因为 OTN 的最小颗粒就是 ODU0--1.25G）。比如客户侧信号是 6G，系统自动分配 1*ODU0+2*ODU1=6.25G 来封装。

OTUk、ODUk、 OPUk 的帧结构如下图所示：



OPUk 中装载着客户的信息净荷，加上 OPUk 开销后成为 OPUk 帧（ODUk 净荷），而 OPUk 帧加上 ODUk 开销和 FA 帧（帧定位字节）后成为 ODUk 帧（OTUk 净荷），ODUk 帧加上 FEC（向前纠错码）后成为 OTUk 帧。

OPUK 帧结构

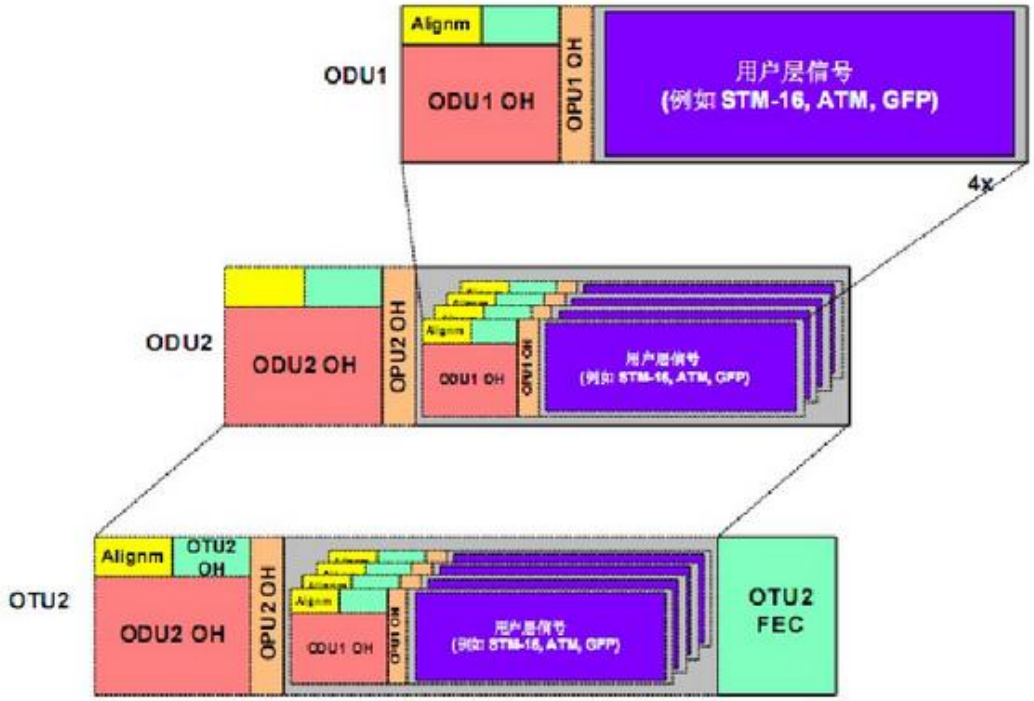


ODUK帧结构

OTUK帧结构

需要注意的是，OTN 的帧结构与 SDH 有一点最大的不同，SDH 的 STM-N 帧结构中 N 不同对应的帧结构不同，STM-4 的列数是 STM-1 的 4 倍，但是帧频都是 8000 帧/秒。而 OTN 的 ODUk 帧格式不随着 k 的改变而改变，都是 4*4080 字节块状帧，但不同 ODUk 等级对应的帧频不同。

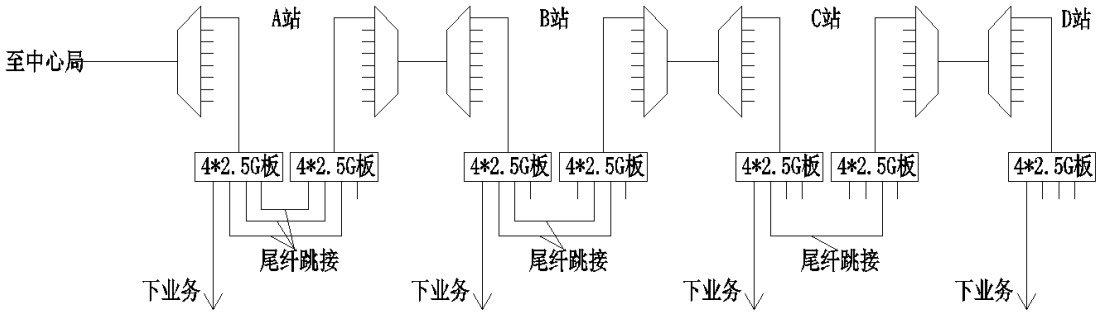
低等级速率的 ODUk 按照一定的规则映射到高等级的 ODUk 中，下图以 4 路 ODU1 映射到 ODU2 中为例，从图中可以大致了解映射的过程。



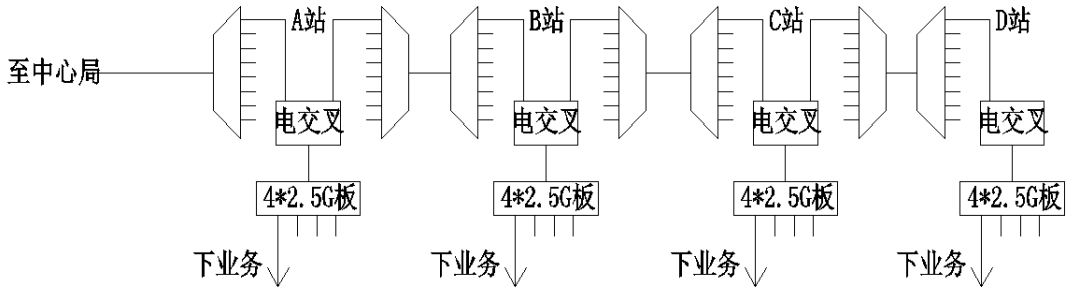
有了这些业务等级之后，OTN 就可以引入类似 SDH 的交叉，就可以调度任意方向的 ODUk 业务（子波长业务），OTN 的 ODUk 交叉和 SDH 的 VC 颗粒的交叉都是对电信号的处理，基于电交叉矩阵实现的，所以称之为电交叉。

我们从下面这两个图中可以对比出，有交叉和无交叉矩阵的站点业务转接的方式的不同，假设图中中心局至 ABCD 四个站点分别有 1 个 2.5G 的业务需求:

无交叉功能的站点业务调度

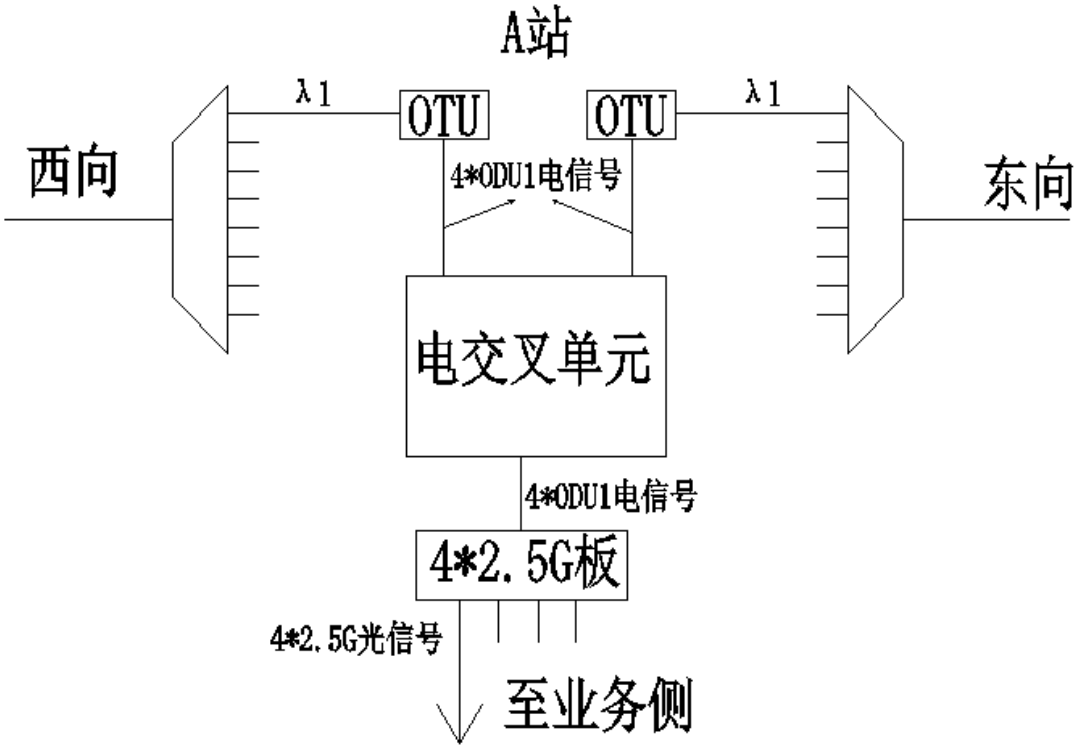


有交叉功能的 OTN 系统业务调度



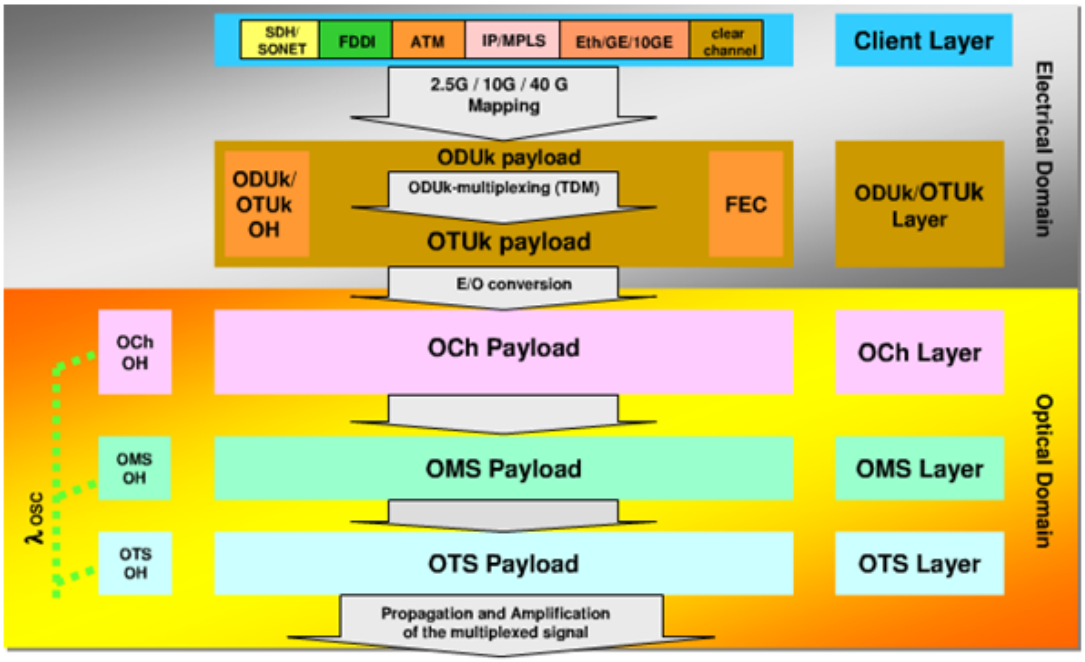
如下图 A 站点，从西向过来的光信号经过分波分成 40 波光信号，通过线路板将第 1 波解成电信号 4*ODU1，接到电交叉单元上，东向与西向相同也是 4*ODU1 接交叉单元；支路侧的 4*2.5G 板通过 4*ODU1 背板接口与交叉单元相连。支路板的背板接口要查看厂家单板的相关数据，系统侧接口的容量不一定等于业务侧接口容量，比如业务接口是 8 口 0-2.5G，有可能系统侧只有 4*ODU1。

与交叉单元相连的都是通过背板总线连接，不需要人工连线，而支路板的 4*2.5G 接口与业务设备相连是需要我们人工连接的。就像我们的电脑 CPU 和内存、硬盘各个部件相连是系统内部连线，而网线、耳机等这些线是需要我们去连接的。



电交叉单元负责做什么呢，这三个方向来的 12*ODU1 信号，就像 12 个箱子，电交叉单元可以将箱子打开解成 ODU0，也可以不打开；可以将线路侧的箱子取出放到支路侧，将支路侧的箱子放入线路侧；或者还可以将西向的 1 号和 2 号变个位置，变成 2、1、3、4 的顺序放到东向 4 个箱子里，当然也可以什么都不做。而不管怎样调度，占用的交叉容量都是这 12*ODU1=30G，因为你将箱子放在了调度中心，就占用了调度中心的一块地方，你做很复杂的交叉和不交叉都一样。

电层的工作完成了，最终两个方向的 OTU2 都层层打包完毕，接下来 OTU2 经过电-光转换就成为光通道层的单个 10G 波道信号，以此类推，40G 系统对应 OTU3，100G 系统对应 OTU4。OTN 的光层也像 SDH 一样分为光通道层、光复用段层和光传送层，电层和光层的完整体系结构如下图:



OPUk：光通路净荷单元　　ODUk：光通路数据单元　　OTUk：光通路传送单元

FEC：前向纠错编码　　OCh：光通道

OMS：光复用段层　　OTS：光传送层

在这一部分我们不详细展开介绍 OTN 的帧结构、开销定义、复用映射流程，本人工作中涉及不到，多数人也用不到，有兴趣可以参照专业的技术资料相关的内容，大家可以对照 SDH 的相关内容去看去类比理解。

2.5　　OTN 光交叉

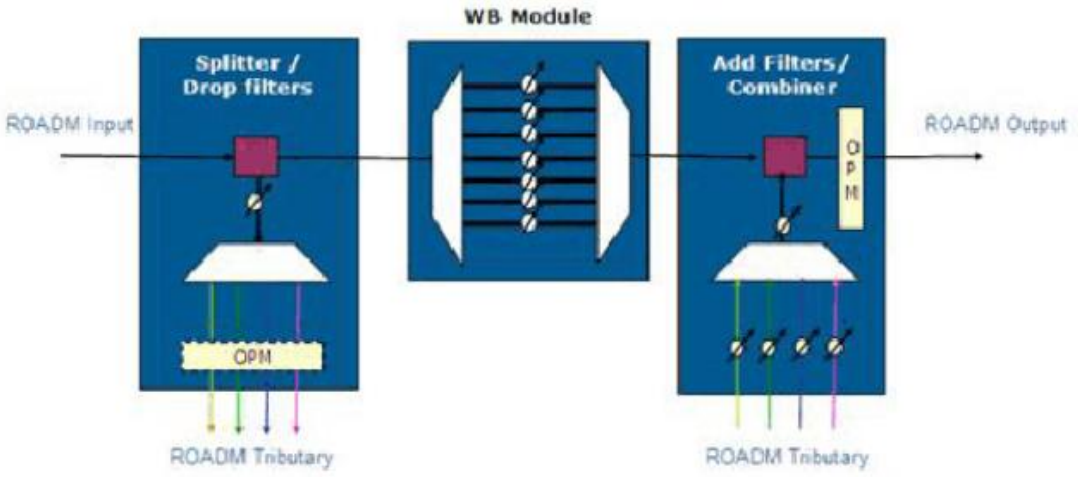
一个 OTN 站点有 N 个（2 个或以上）的光方向，每个方向传送过来的都是 40 个波长合路的光信号，那么这 N*40 个波长信号在站点中都是何去何从的，是我们这一部分要关注的问题，我们来从简单的二维（2 个光方向）来说起。

假设 A 站点从西向过来的合路光信号中，第 1-8 波需要在本站下业务，其余 9-40 波在该站点直通过去，这直通的 32 波需要人为的在东西向的合分波板之间跳纤，这种站点称之为 FOADM（固定光分插复用器），其中的固定是相对 ROADM 来讲的，哪些波长上下或直通可以通过人为的操作去调整。

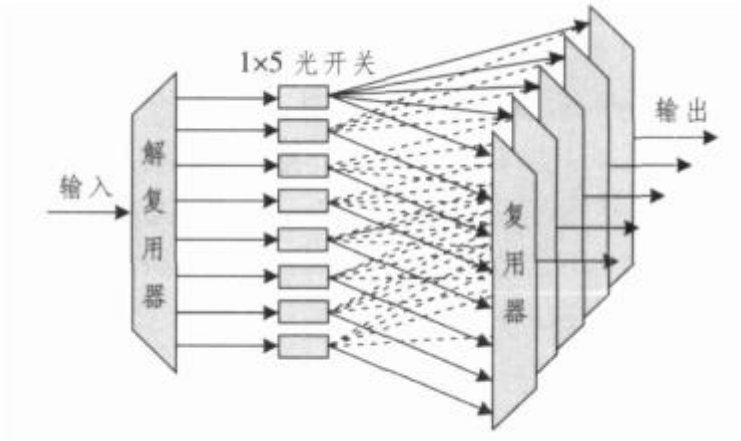
如果 A 站点业务增加，原有 8 波无法满足需求，需要使用 9、10 波，就需要工作人员到 A 站现场，将第 9、10 波的跳纤拆除，通过 OTU 板上下业务。而如果第 9-16 波原本是给 B 站预留的，在 A 站做了穿通但是 B 站没有，那就需要工作人员再到 B 站将 9、10 波跳纤直通过去。

这种方式通常在网络建设的时候就将波道资源给各站点划分预留好，需要变更的时候就需要上述的繁琐的人为手动操作过程，而 ROADM 则可以动态的在网管上配置波长，远程指配每个波长的透传或阻断。ROADM（可重构的光分插复用），顾名思义，是波分系统中的一种具备在波长层面远程控制光信号分插复用状态能力的设备形态，采用可配置的光器件，实现 OTN

节点任意波长的上下和直通配置。 二维的 ROADM 可以通过 WB（波长阻断器）和 PLC（平面光波导）技术来实现，而多维的 ROADM 通过 WSS（波长选择开关）来实现。
波长阻断器的原理是通过使用功分器把全部波长的信号分为两束，一束经过 WB 模块，传输至下一个站点，另一束则传到下行支路，WB 模块的作用是将需要下行的波长阻断。WB 模块最常见的结构是使用解复用器-可变光衰减器-复用器结构，即解复用后每个波长都接一个可程控的可变光衰减器，根据需要将已下行的波长衰减掉，剩余的波长在经波分复用器复用后传输到下一个网元。



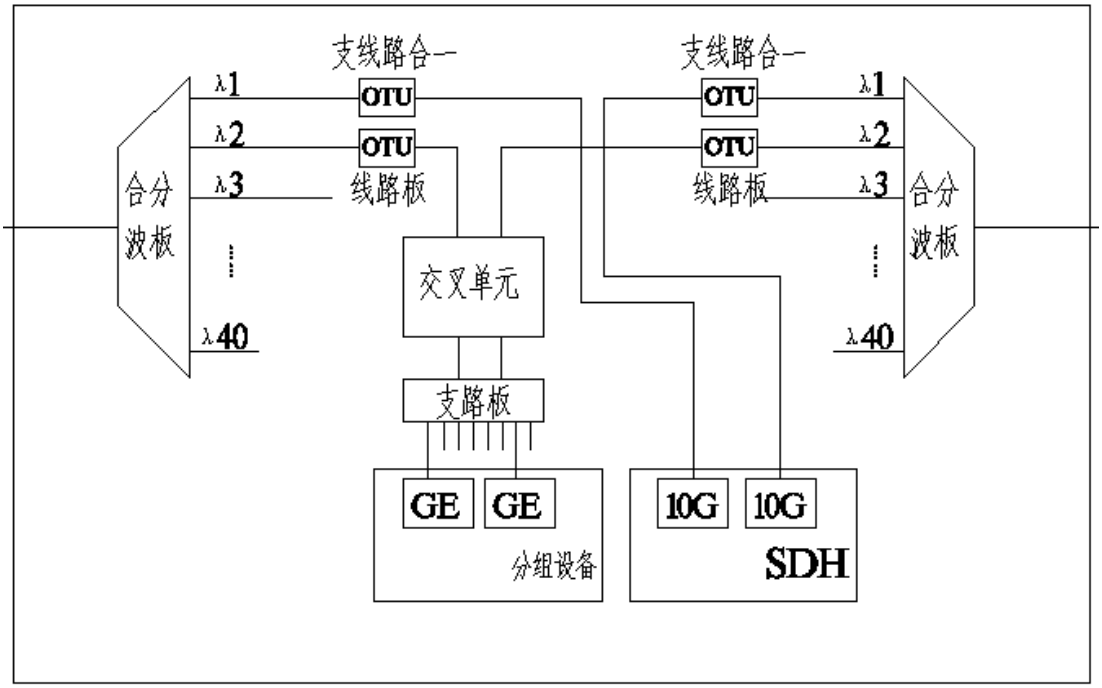
WB 只是能够控制哪些波长的穿通，完整的动态下波长业务还需要借助其他器件来实现，PLC 实现的功能和波长阻断器相同，只是将上下波长的功能和穿通部分集成到了一个芯片上，集成度较高，损耗较小。
在上面的例子里面，40 波的合波光信号被分成功率不同的两路，两路信号都包含 40 个波长，下业务的一路经过分波器之后波长 1-8 下路，而穿通的一路信号经过 WB 之后 1-8 波经过较大的损耗而被过滤掉，而 9-40 波损耗较小可以通过，传到下一站点。将 1-8 还是 1-10 波阻断可以通过网管配置来实现，不需要人工操作，大大提高了波长配置的灵活性。
如果该站点有 2 个以上光方向的话，我们需要在这多个方向之间调度波长信号，只有两个维度的 WB 和 PLC 无法实现，可以通过 WSS（波长选择开关）可来实现。
WSS 是一个多端口的模块，包括一个公共端口和 N 个与之对应的光口，在公共端口的任意波长可以远程指配到 N 个光端口中的任意一个，原理如下图所示:



这样我们可以将一个方向来的任意一个波长,通过网管配置到任意一个光方向中的任意波长去（需要使用可调波长的 OTU 板），业务的配置灵活性又得到了更大的提高。
从需求的角度讲，由于运营商的各大骨干、汇聚节点之间的业务需求相对稳定，目前没有大量的光波长级别的灵活调度需求，所以目前光交叉在我国应用较少，仅在国干层面有少量应用。

2.6 OTN 系统单板

按照惯例，OTN 系统板件分为公共单板和业务单板。
公共单板:
公共单板属于基本配置，特点是一次性配置终生使用，一般情况不需更换和扩容，至少不频繁，这是和业务板的最大不同。所以业务类单板我们打交道比较多，而公共单板基本都是设备新购的时候的标配打包，都是成套的。这就像我们吃火锅，业务单板就像是菜品，根据需求来选择，想吃什么吃什么，而公共单板就像是锅底，想吃你就必须得来一个。
首先和所有设备一样，设备要运转，都需要子架、电源、主控、风扇等。
交叉单元: OTN 有交叉单元，而传统波分没有这部分功能。交叉单元的指标是交叉容量，一般用多少个 G 来表示，比如 320G，是设备对电信号的处理能力的表现。
合分波板: 分为合波和分波两部分，发送端使用合波，接收端使用分波，因为是成对出现所以一般称为合分波板。40 波的系统每个站点的一个光方向需要使用一对，一般两个光方向就是东西向各一对，如果是 80 波系统，每个方向使用两对，然后用梳状滤波器再将两个 40 波信号合并成一路 80 波，反之将 80 波信号分成两路 40 波。
光放板: OA 板的主要参数包括工作窗口、增益、额定输出光功率、接口类型。比如 OA 板有如下的型号标识: OBA2520（C,25dB,20dBm,LC），其中 C 是指该板工作在 C 波段，25dB 是指对信号最大的增益值，代表最大能将输入信号放大的倍数，20dbm 是信号的最大输出光功率，LC 是接头类型，就是我们常说的小方头。拿汽车加油来说，最大增益就是最大加油量，而最大输出光功率就是油箱的容量，油箱满了多加也加不进去。如果是输入功率 15dbm，那输出光功率最大也是 20dbm，此时增益就是 5db。
前面提到过 db，db 是表示增益和衰耗的倍数，而这里的 dbm 是用来表示功率，dbm=10 lg（P/1mW），比如 0dbm=1mw，3dbm≈2mw，10dbm=10mw，就是一个简单的对数计算。dbm 之间同样只有加减没有乘除，dbm 相减实际就是功率之间相除，得出的是 db，是增益或衰耗的倍数。
业务类单板:
业务单板主要分为两种: 线路板、支路板。线路板和支路板功能合到一起，就是支线路合一板。
我们说过，传送网的线路侧是指站点之间的接口，而支路侧是本站上下业务的接口。对于波分来讲，线路接口指的就是波道，就是 40 波其中的任何一波，我们要使用某一波，就要配置线路板，也就是波长转换（OTU）板。支路板提供业务接口，可以接入 SDH、分组网、大客户、IP 城域网等等业务需求。
系统中的一个站点的内部结构图如下图所示，可以看出各类业务单板在系统中的位置:



线路板:

从接口数量来分分为单路、双路、四路，速率上来说分为 10G、40G、100G，按照波长是否可变可以分为可调波长和固定波长。

可调波长的线路板可以根据需要选择波长，提高波道利用的灵活性，同时也可作为备件，但价格要高于固定波长线路板。固定波长线路板需要在购买时告知厂家是第几波，波长无法更改。

线路板的配置就是根据站点要使用波道的情况，比如 A 站点要使用第 1 波，东向西向各一个线路侧接口，就需要两块单口或一块双口的线路板。一个站点在某一波配了线路板就可以上下业务，没有配置线路板就是该波直通，该站点在这一波就是中继的角色。

通常我们说的 40 波系统，这 40 波是指最大波道数量，其中配置了线路板的叫做已配置波道，只有配了线路板的波道才能够传送业务，其余没有配置线路板的叫做空波道，配置波道数/波道总数=波道配置率。已配置波道是一种形成能力,未配置波道只能说是未开发的资源,就像我们买了一块地可以盖 40 栋房子，只要盖好的房子才有住人的能力，没盖房子的只是空地皮。

支路板:

支路板是负责客户业务的接入，将信号输送到电交叉矩阵。型号一般就是 N（端口数）*XX（接口类型）接口板，接口数一般为 4、8、16，对于 40*10G 波分系统，支路接口速率包括 GE、2.5G、0-2.5G 自适应，自适应的接口是可以接 2.5G 以下任何速率的“智能接口”。

支路板的配置很简单,就看站点要接入业务的数量和类型去选择配置,接口不够了需要扩容。

支线路合一板:

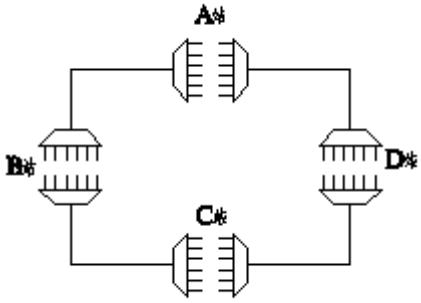
对与波道速率相同的业务（单波道容量是 10G，支路业务也是 10G），接入后不需要经过交叉单元，信号接入后直接送给光转发单元，支路板和线路板的功能集成到一起，叫做支线路合一板。支线路合一板按照接口数量分为单路、双路等。

传统的 DWDM 系统都是采用支线路合一板，OTN 支持电交叉后才将支线路分离开，在支路业务为子波道业务时（支路速率小于线路侧，例如 40*10G 系统中的 2.5G 和 GE 业务）使用分离的支路板，这样线路侧和支路侧都可以根据业务需求灵活配置。

波分系统配置流程:

首先要收集业务需求，将各类业务需求整理成表格，然后根据需求表和现状波道图画出扩容波道图，最后根据波道图确定扩容线路板数量，根据业务接口和现有支路板使用情况，确定扩容支路板数量。

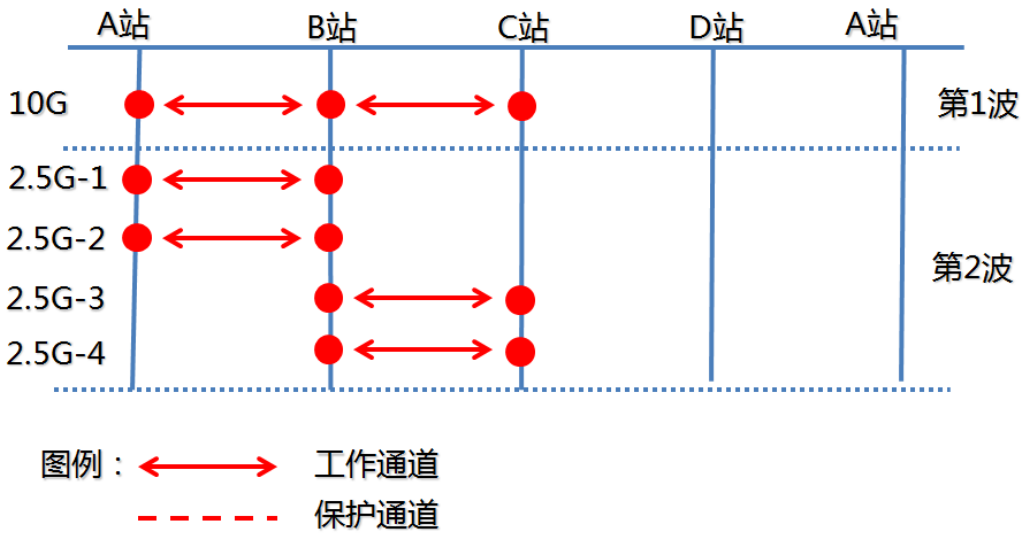
下面举个例子来具体了解一下 OTN 设备的配置：图中的 A、B、C、D 组成一个 OTN 环路。



如果不考虑业务保护，假设 A 站到 B 站、B 站到 C 站分别有 1*10G+2*2.5G 的业务需求，需求表如下：

起点	终点	速率	数量
A 站	B 站	10G	1
A 站	B 站	2.5G	2
B 站	C 站	10G	1
B 站	C 站	2.5G	2

根据需求表画出对应波道图，如下图所示：



B 站点需要配置 2 块 10G 支线路合一板，2 块 10G 线路板和 4 个 2.5G 的支路接口，A 和 C 站点分别需要配置 1 块 10G 支线路合一板、1 块 10G 线路板和 2 个 2.5G 的支路接口，支路接口考虑本次的需求和适当的预留，可以选择 4 口或者 8 口 2.5G 支路板。

2.7 OTN 系统的保护

OTN 的保护分为单板级保护和网络级保护。

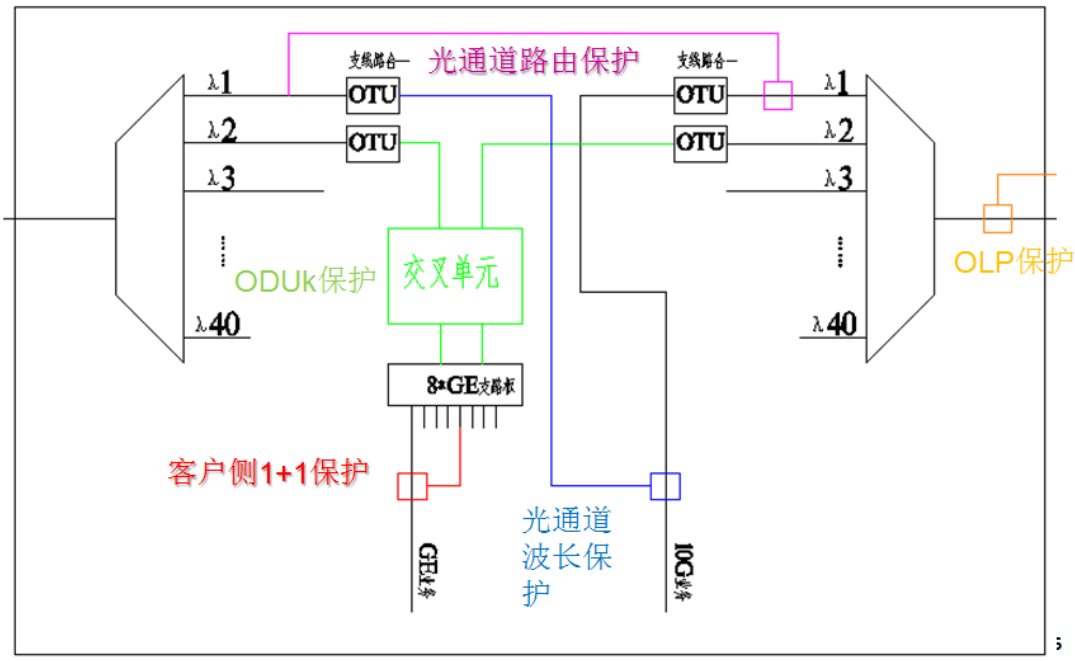
单板级保护：

单板级的 1+1 保护原理容易理解，前面 SDH 部分已经介绍过，就是 2 块（1+1 备份）或者多块（1: N 备份）相同板件的互为备份，道理和我们的手机配两块电池一样，一块没电了换另外一块，照样打电话。

网络级保护:

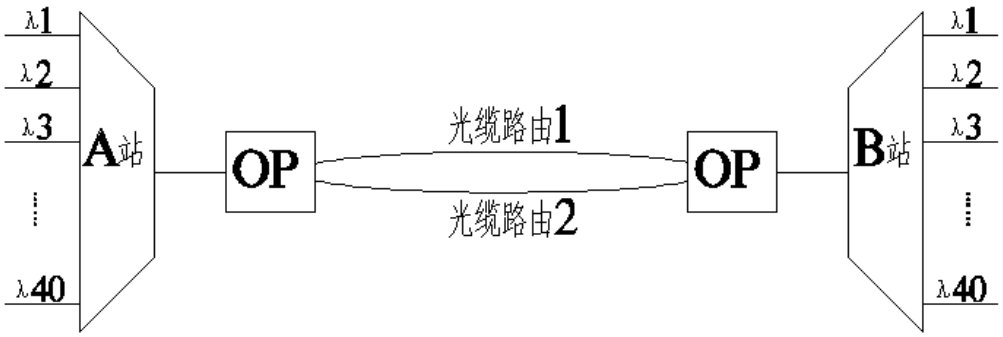
OTN 的网络级保护按照保护对象可以分为光线路保护（OLP）、光通道保护（OCP）、ODUk 保护和客户侧保护。其中光线路保护、光通道保护、客户侧 1+1 保护属于传统 DWDM 保护方式，ODUk 保护为 OTN 相对 DWDM 特有的保护方式，因为 DWDM 就没有 ODUk 的概念。

下面这张图中可以看出不同保护方式 OP 板所处的位置和保护的对象的不同:



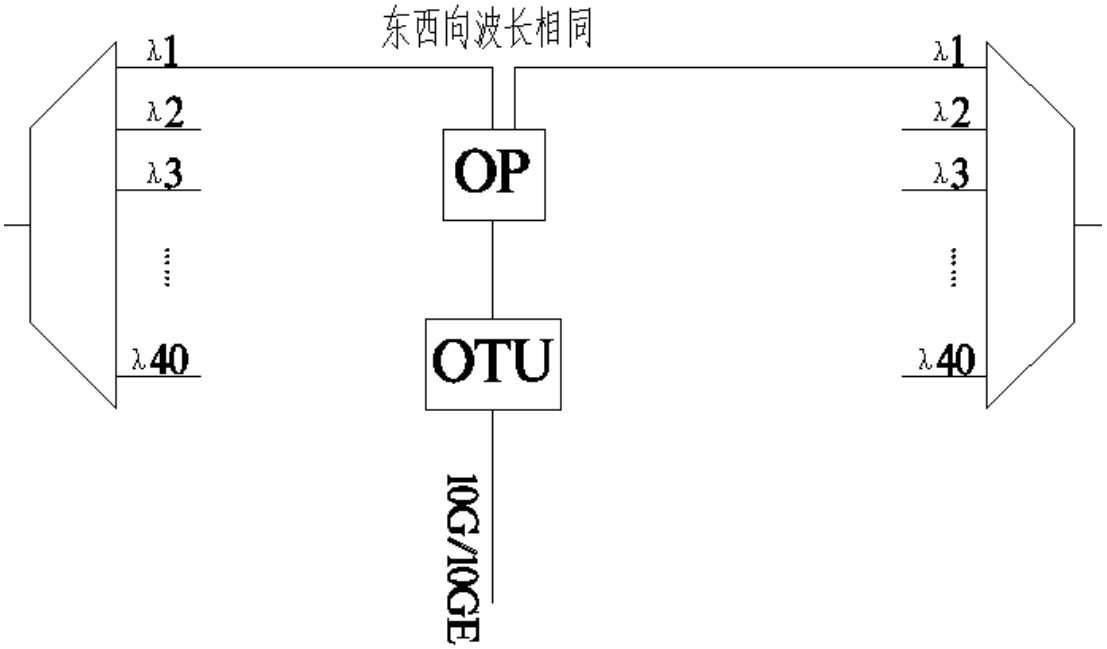
光线路保护:

光线路保护（OLP）是对两个站点之间的光缆线路进行保护，简单说就是 AB 两点间的光信号走两条不同路由的光缆，来防止某条光缆故障导致的业务中断。OLP 保护对象是整个合波后（40 波）的线路信号，原理就是将合波放大后的信号通过 OP 板复制，主备信号分别走两条路由，接收端在主用业务信号低于门限值时切换到备用业务，与 SDH 的通道保护的机制一“并发优收”相同。OLP 保护需要具备不同路由的光缆线路才能实现，一般多用于链型系统，用于光缆线路经常中断的段落。



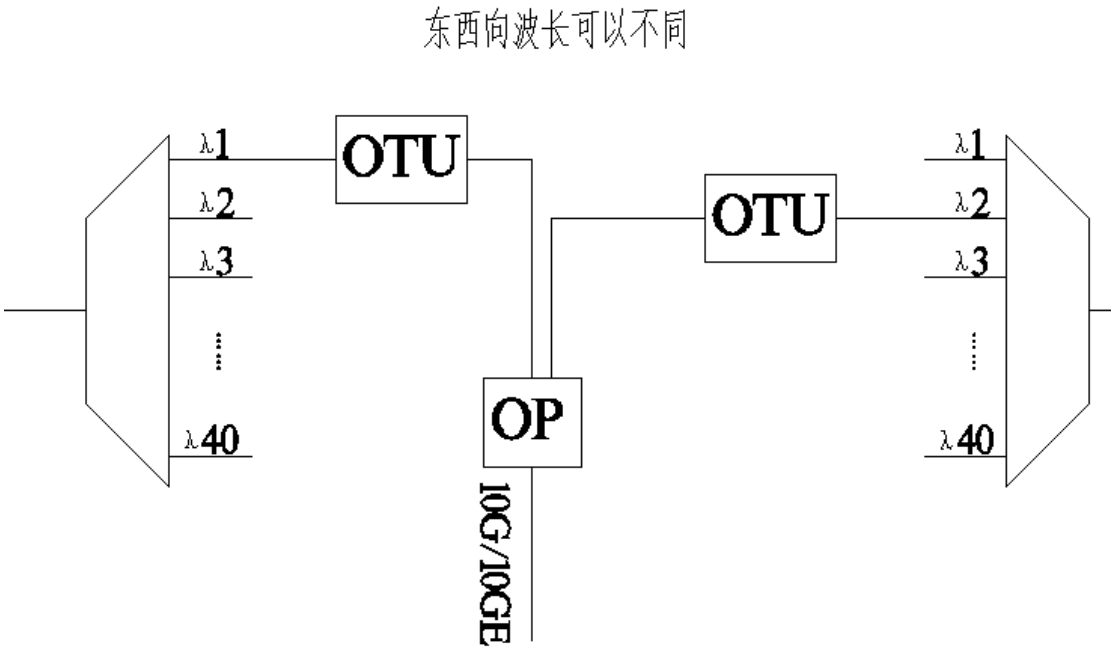
光通道保护:

光通道保护的对象是单个的波道，光通道保护有不同的实现方式，一种是对已经转换波长的信号复制后分别送到东西向的合分波器，如果东边路由有问题可以从西向传送，保护对象是OTU后端，能够实现对路由的保护，称为光通道路由保护。光通道路由保护也称OTU板内1+1保护，板内的意思就是只有一块OTU板:



另一种是将客户信号复制以后，分别送给不同的OTU板，可以将信号调制为2个不同的波长（当然也可以是相同波长），称之为光通道波长保护。光通道路由保护和波长保护的区别就是信号的一分为二是在OTU之前还是之后。在OTU之前分路信号的，需要两块OTU板来调制波长，所以保护的对象除了东西向路由，也能够对OT U板的故障和单个波长的信号失效都能有效的保护。

光通道波长保护也称OTU板间1+1保护:



对于一个波道保护来说，光通道波长保护需要 2 块 OTU 板，建设成本也高了一倍，要多花几万块钱，但是可以有效的保护单板、线路。而光通道路由保护仅需要增加 OP 板即可，但只能对线路进行保护，成本很低，所以实际应用的多一些。

光波长共享保护:

光通道保护还有一种方式叫做光波长共享保护（OWSP），光波长共享保护和光通道波长保护的区别类似于 SDH 的复用段保护和通道保护的区别，SDH 复用段保护是以 VC4 为单位，而 OTN 的 OWSP 保护是以波道为单位 ，原理都是双端倒换，需要 MSP 协议，只能应用于环形系统。

ODUk 的保护分为 ODUk SNCP 和 ODUk Spring 两种，保护对象是 ODUk 颗粒。SNCP 和 Spring 就相当于 SDH 的通道保护和复用段保护，只是颗粒不同，可以参见 SDH 部分的图文介绍，这里就不详细解释了。

而客户侧 1+1 保护，是将客户侧的业务直接复制为主备两路，分别经过不同的支路接口、OTN 板，经过不同方向的合分波器，对全程都能有效的保护，但是成本也要高一些，一般应用较少。

各种保护方式对比:

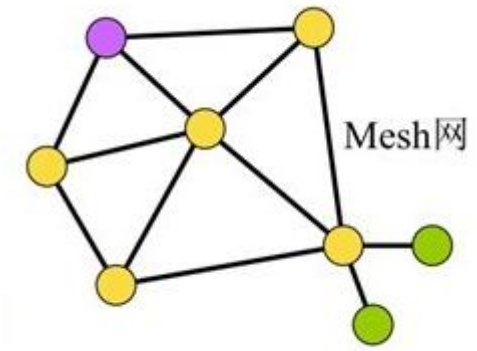
保护方式	客户侧1+1保护	ODUk SNCP	ODUk Spring	光通道路由保护	光通道波长保护	光波长共享保护OWSP	线路1+1保护
保护颗粒	客户侧业务	ODUk	ODUk	波长	波长	波长	线路
保护范围	全部	单板、线路	单板、线路	线路	单板、线路	单板、线路	线路
是否需要APS协议	否	否	是	否	否	是	否
倒换时间	小于30ms	小于30ms	小于50ms	小于30ms	小于30ms	小于50ms	小于30ms
是否有节点数限制	否	否	是	否	否	是	否
适用网络拓扑	全部	全部	环型	环型	全部	环型	链型

从适用业务类型来看，ODUkspring 和 OWSP 这类需要 APS 协议的保护方式适用于分散型业务，ODUk SNCP、光通道这类双发选收的保护方式适用于集中型业务，而线路 1+1 保护和客户侧 1+1 保护则适用于所有业务。

关于保护再说两句:

如果说所有的保护都不需要成本，那么一定是能上的全上。但是事实是，所有的保护都有一个特点，越高的安全级别就需要付出越高的建设成本，所以这里有一个度的问题。

我们要保护光缆路由，就需要建设两条不同路由的光缆，成本就翻一倍左右；如果想要很多条光缆都断了都不影响业务，就要建成 MESH 网，网络还要加载控制层面，支持业务的自动配置，搞不好设备都要换掉，建设成本我们这里不详细计算，反正是很高。至于保护的效果，谁也不能说没有效果，就像我们给汽车买全险一样，买一个心安。



那么究竟保护到一个什么程度是合适的呢？不同的运营商不一样，不同的业务不一样，不同的网络层次也不一样，所以这个问题答案肯定不是一概而论的，这里就是谈谈而已，不下结论。

首先说说网络保护，我们不谈保护方式，就谈一般习惯性的做法。

不同层面节点的保护:

我们知道，传送网中“环上“的节点是有网络级保护的，而”链上“的节点则没有，因为网络保护的前提就是双路由，一般是指不同路由（物理成环）。同缆环是不得已为之的一种手段，同缆异芯的组环，这种方式只能保护光口不能保护光缆，设备光口是 2 个，但光缆只有一条，农民伯伯一锄头下去所有的纤芯全部断掉，信号就有去无回。

一般核心汇聚节点绝大多数都在环上，因为它们位置太重要。核心节点负责整个行政区域内的所有业务，汇聚站点也像封疆大吏一样，掌管一方的命运，所以态度上这些节点不能成也要成环，资源就必须向这些节点倾斜，所以问题不大。

而接入层，我们都知道有个指标叫做成环率，如果说核心汇聚层是必须成环，接入层就成了尽力而为，两者之间差距是很大的。成环率=环上节点数/节点总数，这总数*（1-成环率）的这部分就是链上的节点。

接入层有三个特点:

1、业务相对核心汇聚层来说不够重要，一般链上就是三五个节点，也就是三五个站的业务，和上层的成百上千比，数量上有差距。

2、节点数量巨大，这一点和上面一点反过来了，这就是我们的二八理论，少数站点掌握着大量的资源。如果说每一个站点都要成环，汇聚层的成本要低的多，因为问题点少。就像某公司年终给每个骨干发一辆宝马彰显实力，可是要是给所有员工每人一辆 QQ 也未必吃得消。

3、建设难度大，这个问题比上一个问题更加突出，因为上一点说的就是成本，运营商有差钱的也有不差钱的，可是建设困难面前大家都一筹莫展。接入层站点分布范围广，延伸到了地图板块的每一个乡镇、村落，这些角落里有很多天然的屏障让我们无法建设两条不同路的光缆，有的连一条都建不成，还要用微波。

这三点是为了解释为什么是尽力而为，但尽力而为不是无为，我们规定了成环率、长链（3 节点以内）和超长链（5 节点）的比例，这些问题还是要持续关注和改善的，虽然只是一些多年不变的老话题、硬骨头，至少邻居省市之间要互相比较一下，不要搞到最后一名。

业务侧保护和传输侧保护:

这里就说一点，我们的保护一般就是一倍的关系，也就是对于 1 个业务我们配置 2 倍的资源去实现 1V1 的保护，但是这个业务侧和传输侧的概念是有些人容易模糊的。

比如数据网提出 A 站点到中心局 2 条 10GE 的需求，那么这两条业务之间本身就是互为保护的关系，我们可以东西向配置在一个波道里。

这种情况是业务侧已有了保护，传输侧不在提供额外的保护。
我们举个例子，张三交给你 1 把钥匙让你转交给李四，这时候传送网的做法就是将钥匙配 1 把备用的，1 把丢了还有另外 1 把，这是传输侧的保护。
但是，如果张三交给你 2 把相同的钥匙，说明他已经配过了，我们就不需要再去将 2 把配成 4 把，这是业务侧的保护。
同样，我们 OTN 系统承载的 MSTP 网的 10G 环路，实际波分侧是没有保护的，我们在一个波道里承载的是 A-B、B-C、C-D、D-A 的四条业务需求。如果 A-B 的光缆断了，对于波分来讲，A-B 这条业务就中断了，但是 MSTP 网并没有中断，这时的倒换环回是 MSTP 通过 B-C-D-A 实现的，是业务层面的保护。
概括一下，就是业务无传输有，业务有传输无的关系。
当然，像前面提到的 MESH 的例子，显然不是 1 倍的保护，只要业务足够重要足够值得，也是可以有的，这个帐很好算。

支路侧保护:

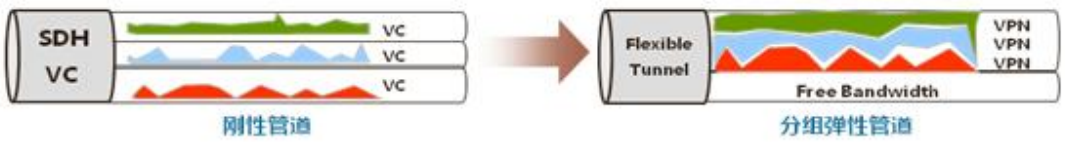
很多的设备是支持支路侧保护的，从设备来讲就是一个支路接口桥接的问题，能不能做是厂家的问题，做与不做是建设者和设计者要面临的问题。
一般来讲可以一刀切，核心层做汇聚层不做，或者核心汇聚层做接入层不做，这里的问题还是那个二八原理，做会花更多的钱，但安全性肯定更上一层楼，主要的分水岭，到底在哪，不能一概而论。

总结:

关于保护，这就是一个花多少钱办多少事的问题，说这么多，主要想让大家去了解明了，结论，还要自己去下。

第三章　　分组传送网 3.1　　MSTP 的末路

前面介绍 MSTP 的部分提到，MSTP 是针对语音业务而生的一种传送机制，MSTP 是刚性通道，而分组传送网则是弹性通道，业务可以统计复用，更适应数据业务的承载。
什么是统计复用，打个比方，有 100 个人的公司每天中午订餐，每个人吃 1-2 个馒头，吃 1 个还是吃 2 个是不确定的，那么对于快餐公司，如何确定给这 100 个人配送多少个馒头?刚性通道的方式就像是固定每天送 200 个，每人 2 个，吃不了就扔掉，而统计复用就是根据经验统计出，每天总共馒头消耗数量不会超过 150 个，那么就送 150 个馒头，每个人能吃多少吃多少，这节省的 50 个馒头就是统计复用的效果，如果这个波动范围更大，比如 0-100 之间，统计复用的效果就会更加明显，所以传送网面临从 MSTP 到分组传送网的转型，这是一个必然的趋势，在 3G 后期 HSPA+、DC 到 LTE 时代这个趋势可以用迫在眉睫来形容。
在平均 100M 单站带宽情况下复用效率可以达到很多倍，具体数字不好估计，至少应该在 4 倍以上，这个倍数叫做收敛比， 比如 40 个 100M 需求的基站带宽需求量是 4G，实际配置只需要 1G，收敛比就是 4：1.
可见，统计复用可以提高利用率减少资源浪费。从下图能够看出，带宽不能复用的刚性通道和统计复用的弹性通道的承载效率:



无线网从 3G 发展到 HSPA+到 LTE，数据业务的带宽在剧烈增长，LTE 阶段单站带宽达到 100M 以上，而语音业务从始至终都是 1-2 个 E1。这个趋势下，数据业务量越大，技术上的转型就越迫切。

我们打个比方，就像一个中餐馆的主营业务是盖浇饭，同时也提供汉堡满足少量的顾客需求，但如果点汉堡的人越来越多，已经远远超过了主营业务盖浇饭，那就有必要改招牌换厨子，干脆改成肯德基算了，毕竟肯德基做汉堡的质量、速度、成本都要有优势。

除了带宽利用率的大幅提高之外，分组传送网带来的还有建设成本的下降，就像前面的例子，肯德基做一个汉堡的成本一定是比非专业的中餐馆要低的多，从实际工程采购经验来看，一端 622M 的 MSTP 设备需要 3 万左右，一端 2.5G 的 MSTP 设备需要 10-15 万左右，一端 10G 的 MSTP 设备需要 40-50 万左右，而一端 GE 的分组设备只有 1-2 万，10GE 的汇聚层设备也只有十几万，这么省钱的技术，运营商们表示非常喜欢。

从设备集成度来看，SDH 的常见数据接口板一般就是 2 口 GE 或者 8 口 FE，这还是指核心汇聚层设备，而分组的接入层设备一般就是 8 口 GE，汇聚层可以达到 24 口、48 口 GE，这意味着同样大小的 MSTP 和分组设备提供 GE、10GE 接口的能力是有很大差距的。

可见，承载数据业务的前提下，分组传送网基本是可以秒杀 MSTP 的节奏，不过术业有专攻，要是比 SDH 那一套 E1、STM-1 承载的话，形势就要反过来看了。

业务的全 IP 化、宽带化驱动了传送网从 MSTP 到分组传送网的转型，但变革总是伴随着很多问题和阻力的，这个演进过程也注定了无比的艰难，因为运营商对分组传送网的建设、维护经验都相当欠缺，原来对 MSTP 的那一套经验和分组传送网的要求有很大差异，即便是原来搞数据网专业的人员，恐怕也极少有面对几百台甚至几千台设备的大网的经验，但困难归困难，挑战归挑战，技术演进的车轮是不会原地等待的。

3.2 IP 网基础-分层结构

由于无线数据业务的快速发展，IP 技术以其简单开放的特性一统天下，使多年以前原本没什么交集的两张网—电话网和计算机网渐渐的融合，可谓殊途同归。分组传送网是从数据网设备二层交换、三层路由的那一套技术发展过来，数据网设备是用于解决不同地点的主机（计算机）互相通信的问题。其实数据网设备我们并不陌生，日常在我们的办公室和家里都可以见到交换机、路由器。在了解分组传送网之前，我们先走进数据网去了解认识一下这些对于 MSTP 来讲全新的世界。

首先，按照惯例，我们要先了解其分层的概念。

为什么又要分层，为什么要说“又”呢，因为层面很重要。传送网乃至通信网中，很多概念的掌握就基于这个“层”的理解。

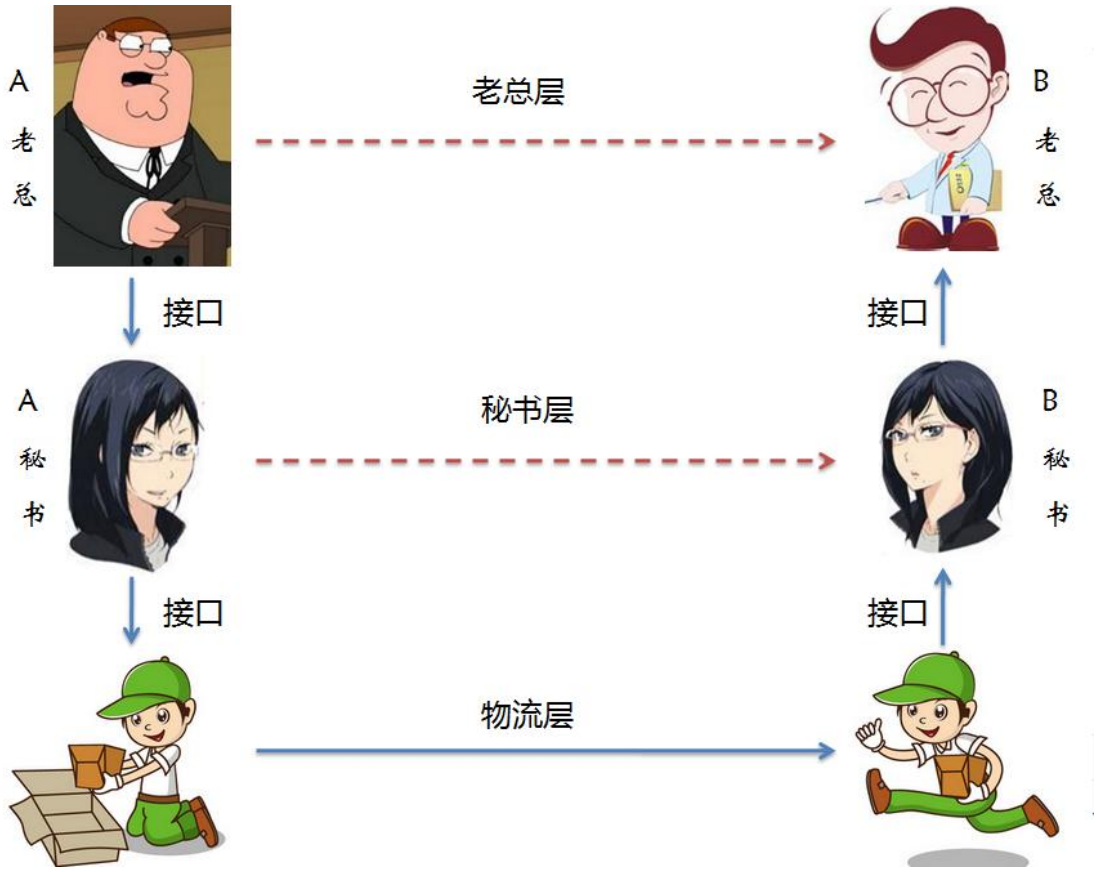
这里再举个例子:

A 公司和 B 公司是密切的贸易合作伙伴。某天，A 公司老总（简称 A 总）大笔一挥，签了个合同，要寄给 B 公司的老总（简称 B 总）。

这时，A 总将合同装在文件袋中，他叫来 A 公司秘书（简称 A 秘），将袋子交给秘书，吩咐其将袋子寄给 B 公司。

A 秘叫来快递公司，填上 B 公司的电话地址，将快递寄出。

快递公司将文件袋包装之后，包裹通过物流系统 N 次中转，最终送达 B 公司。B 公司秘书（简称 B 秘），收到以后，将包裹拆开之后，将文件袋转交给 B 总。



这是我们生活中容易接触到的实例，基于此例子，我们来类比一些概念。

逻辑连接和物理连接:

上图中，虚线表示逻辑连接，实线表示物理连接。

A 总并没有把合同直接交给 B 总手里，而是交给了秘书，通过层层转交，通过整个“网络”最终达到了 B 总那，这样效果和直接交给 B 总是一样的，B 总不可能说，A 总你没有给我合同，合同是我的秘书给我的，这是抬杠。

实际生活中，我们也更关注最终起到实质性交流的主角，说“A 总将合同给了 B 总“符合我们的语言习惯，因为其他的“A 秘””X 通快递”并不是重点。

A 总和 B 总之间发生了间接的联系，我们可以换种说法，A 总和 B 总之间建立一个逻辑连接。同理，站在网络的角度，我们往往更关心的是逻辑关系，否则任何一个专业讲方案、规划，可能都要将传送网的方案扯进来。

而图中 A 总-A 秘-物流、物流-B 秘-B 总之间是有实际接触的，在通信网里来讲，就是有物理接口的，所以叫做物理连接。一般 A 总和 A 秘是一个站点，B 总和 B 秘是一个站点，也就是上下层对接是在同一站点中进行的。对于通信网，上下层接口有可能是同一专业内部接口，也有可能是不同专业之间的接口。

层级概念

例子中 A 总和 B 总是一个层面，我们叫做老总层；A 秘和 B 秘是一个层面，我们叫做秘书层；而负责收和发快递的两个快递员是一个层面，我们叫做物流层；

A 总可以给 B 总打个电话，我给你发的合同你收到了没？

A 秘可以给 B 秘打个电话，我给你发的东东你收到了没？

只有一个层面的两个实体（网元）之间才可以对话，A 秘不会给 B 总打电话，因为这叫不在一个层面上，B 总根本就不认识 A 秘。

同一层面之间的网元组成这一层的网络，比如老总层还有 C 总、D 总等等，如果画一个各公司老总层的关系图，上面就都是老总没有秘书(有些秘可能比总大，这个我们不讨论)，他们之间沟通的也都是公司层面的事。

同样，A 秘的电话本里也有 B 秘、C 秘、D 秘，这些都是他们工作中的接口人，他们又是一个层面，叫做秘书层，就像二号首长里写的一样，领导在里面开大会，秘书在外面开小会，这两个会就是不同层面各自的会话。

不同层的信息单元

老总层的信息单元是合同，秘书层的信息单元是文件袋，物流层的信息单元是包裹，老总只面对合同，他发出、收到、处理的只能是合同，包裹和袋子会在老总发出合同之后加上，在收端在交给老总之前被拆掉。这就像通信中不同层的信息单元一样，有帧、包、比特流，一个层的网元只会收、发、处理这个层的信息单元。

A 秘书和 B 秘书并不知道文件袋里装的是合同，反正给我的时候啥样，送过去的时候还是啥样。在通信网里，这叫做上层信息对下层来说是透明的，或者说下层将上层将信息透传。

层间的关系

最后，老总层和秘书层是什么关系，我们说，秘书层是为老总层服务的。同理，我们也常说，传送网就是给业务网提供服务的。老总不管秘书怎么将合同送达，他只提要求，你要在明天下午两点之前送到。业务网也不关心传送网是怎么组网、怎么保护、怎么管理，业务网也只提要求，比如业务的开通时间要求、起止站点、电路带宽、最大延时等。

传送网和其他网的关系:

有些人问，业务网（无线网、宽带网等）和传送网是什么关系，根据上面例子的分析，就是上下层的关系，业务侧 A 点到 B 点要一个 100M，传送网分别在 A 点和 B 点和业务网对接，这 100M 就连上了，真和快递一样一样的。传输永远是在最下层，就是跑腿的，所以本文中总是拿物流做比喻，不但工作方式雷同，连命运都相似……

传送网内部， MSTP/分组和波分之间是什么关系？也是上下层的关系， MSTP/分组在上层，波分在下层。SDH 将客户侧业务送达，如果是通过光缆，光缆就是 SDH 的下层。如果通过波分承载，波分将 N 个 SDH/分组业务送达，波分就位于 SDH/分组和光缆之间。对于比如 E1 业务，SDH 承载了 E1，波分承载了 SDH 的 10G 或者 2.5G 大颗粒，而光缆承载了波分的合路信号。

最后说承载网和传送网的关系，我们可以简单概括说，也是上下层的关系，但是情况跟业务网 over 传送网比稍稍有点复杂。

过去，承载网是和 MSTP 网并行的，革命分工不同。MSTP 主要承载无线、大客户这类专网业务，以 TDM 业务为主。承载网主要面对公众用户，以 IP 业务为主。分组网的崛起之后，分组网和承载网实际上有了相同的功能，但是两者承载的业务种类有所区别，短时间还是各自为营，两者的融合也是必然的趋势，但是要有个过程。

对于数据网在不同的省份、市区、县城之间互通是通过波分承载的，波分属于传送网，所以说数据网是传送网的上层也没有错。简单的说，承载网和 MSTP/分组平级，是波分的上层。OK，言归正传，下面接着说 IP 网的分层。

IP 网的分层:

关于网络互联 ISO 定义了 OSI（开放系统互联）七层参考模型，其中应用、表示、会话上三层属于资源子网，物理、链路、网络下三层属于通信子网，传输层是资源子网和通信子网之间的桥梁。OSI 七层模型只是一个参考，在实际网络中不会严格按照七层来，比如 TCP/IP 将其简化为四层结构。

这个图看起来像两个高楼，生活中你从这个楼要去那个楼，一定要先下到一楼，对于通信也是一样，上面的连接都是逻辑的，最后都得下到物理层，变成比特流。

数据在网络中传送对应下面三层：网络层、链路层和物理层，而上三层是主机对主机之间的一些协议，总之数据到了用户端，能够还原成我们直观感受到的音视频等各种应用文件，我们在网络的角度可以不用去了解，笔者也确实是很不了解，总之，网络工作者们只关心数据如何送达。

我们常听到某个设备、网络或者业务是二层的或者三层的，这个二层和三层就是指数据链路层和网络层。网络层寻址使用 IP 地址，寻址的方式是三层路由，发送的是数据包；数据链路层寻址使用物理地址（MAC 地址），寻址方式是二层交换，发送的信息是数据帧；物理层是建立端到端的连接，发送的是比特流。

我们要发送的信息经过网络层，会加上网络层的帧头形成了三层的数据包；到了数据链路层会加上二层的帧头帧尾形成二层数据帧；数据帧到了物理层变成 0 和 1 组成的比特流，经过物理链路传送到下一站点。数据在到达节点或者终点，经过上述的相反过程。

还是拿物流网打个比方，这里澄清一下，我也不是物流公司的工作人员，最多就是个某宝剁手党，物流领域说的不对的，还请各位看官海涵。

每个快递包裹都会填写发件人和收件人，包括姓名、地址、电话。“姓名”就像是 MAC 地址，而“地址”就相当于网络中的 IP 地址，那分层处理在物流网中的体现就很好理解，快递在每个站点分拣调度的时候，只需要看地址，不需要看姓名；而货物到达了某个小区之后，快递员就只需要看姓名、电话；货物在路上运输的时候，司机才不管这些货物的起点终点是哪里，所有地址统统不看，更不会管发件人和收件人姓甚名谁，他只是不停的比如在北京和上海之间往返运输，将货物从北京送到上海就完成了他的任务。这里基于发件地址去分拣调度货物就对应网络层，而根据姓名电话将货物送达就对应了数据链路层，而在运输途中就对应物理层。

如果某个节点的设备是二层交换机，那数据在这个节点只打开和封装二层的帧格式，按照物理地址进行二层交换去寻址转发，而不必去修改网络层的数据，二层交换机也没有这个功能。寻址的概念是之前的 SDH 没有的，因为 SDH 的通道层和段都是面向连接的，起点和终点都是固定的唯一的， 就像屋子里只有我们两个人，那我说话肯定是给你说的。如果和 OSI 七层模型对应，SDH 就是工作在物理层。而数据网络是非面向连接的，一个数据包（帧）在网络里有很多个岔路口，要依靠地址才能找到一条通往目标的路，这个过程就是寻址。SDH 就像走一条条专用直达的高速路，只要上高速就只有目的地一个出口；而数据网就像是在城市里开车，要去哪里必须要知道地址需要使用导航，否则会迷路。

3.3 IP 网基础—二层交换

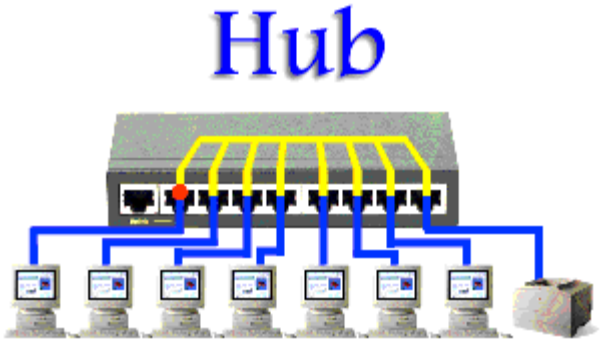
要了解 IP 网每一层的工作方式，我们由简单到复杂的组网情况说起。

假如 A 和 B 两台主机需要通信，只需要通过网线将两个网卡连接在一起就可以，这个连接是物理层的，中间不需经过交换机路由器等设备。物理层对接头、线缆相关的参数进行规定，比如 RJ45 的 8 芯线缆中一端的 1、2 芯用来发、收信号，而另一端 3 和 6 芯用来收和发，这样的规定使两台主机能够在物理上互相通信成为可能，就像我们在 A 和 B 之间修建了一条马路，有了马路之后 A 和 B 才能互相来往。

这个时候 A 和 B 的通信就不依赖于地址，因为整个世界只有 A 和 B，A 说 Hello，B 就一定会收到，这就是小两口的二人世界。

这个时候，又加入了 C 和 D，我们也可以将 ABCD 通过一根线连接到了一起，相当于四台主机以时分复用的方式共享一条线缆。这时候由于 ABCD 之间没有隔离，如果 A 说话 BCD 全部都能听到（广播域），A 在说话前会先侦听有没有其他的主机正在说话，如果有就等待一段时间再侦听，直到链路空闲时再发言；如果 A 和 B 同时说话就会发生撞车（冲突域），这时 A 和 B 都随机生成一个数字去等待，比如 A 等待 1ms，B 等待 3ms，1ms 之后 A 重新开始说话，而 B 的等待时间比 A 长，所以继续等待 A 说完话再侦听线路，等待时机。

这种工作机制就是集线器（Hub），在互联网在我国刚刚兴起时还比较常见，主要是因为价格比交换机便宜。集线器是工作在物理层的设备，没有二三层设备的那些存储转发学习的功能，也不会使用 MAC 地址或者 IP 地址，上面例子里 ABCD 之间的工作机制叫做 CSMA/CD（载波监听多路访问-冲突检测），其实根本上说就是靠广播，大家完全没有秘密可言。由于信号总是“撞车”，可想而知，集线器的带宽利用率很低，现在已经很少使用。集线器的各个端口同属于一个冲突域，也属于一个广播域，这个概念从字面上就容易理解。



如果把集线器换成二层交换机，ABCD 都连接到交换机的 1-4 号端口下，这时四台主机就组成了一个小的二层局域网，虽然交换机和集线器从外观上看几乎一模一样，但是主机间通信的机制完全不同，交换机高大上了许多。

二层交换是根据 MAC 地址进行转发的，这里介绍一下 MAC 地址：每一个主机和网络设备的 MAC 地址是在出厂的时候就被分配好的，MAC 地址就像我们的身份证号，从一出生就伴随你一辈子，一般情况下不会更改。MAC 地址共 48 个比特，为了读写方便通常用 12 个 16 进制数字来表示（如 00-E0-FC-00-00-06），其中前 24 位是由国际标准组织 IEEE 分配给设备厂商，后 24 位由厂商自己定义。

这个时候 A 要对 D 说话，交换机是如何转达给 D 的呢？

1) 交换机收到了 A 的数据之后，首先查看源 MAC 地址，知道了 1 端口的这个主机叫做 A，交换机拿出本子记下了：“主机 A—端口 1”，这时候形成了他的第一条地址表；

2) 这时交换机并不清楚 D 在哪里，于是他对除了 A 之外的 3 个端口都说，A 在呼叫 D，收到请回答；

3) 这时候 D 通过目的 MAC 地址判断是在呼叫自己，会对交换机回应数据包告知俺就是你要找的 D，而其他 B 和 C 两个主机看到数据包目的地址不是自己，则直接将数据包丢弃，就像什么都没有发生过；

4) A 又通过 D 返回的数据包中得知了，4 端口对应主机 D，建立了又一条地址表，主机 D—端口 4；

这样 A 到 D 就通过交换机完成了一次通信，交换机也学习到了 A 和 D 的 2 个 MAC 地址，下一次发往 A 和 D 的数据就不会再告知天下，而是直接发到所对应的端口。交换机根据源 MAC 地址建立地址表这个过程就叫做学习，经过了如此这般的若干个回合之后，交换机对于其各个端口和主机的对应关系应该熟记于心，形成了一个 MAC 地址和端口的对应关系，就是 MAC 地址表。

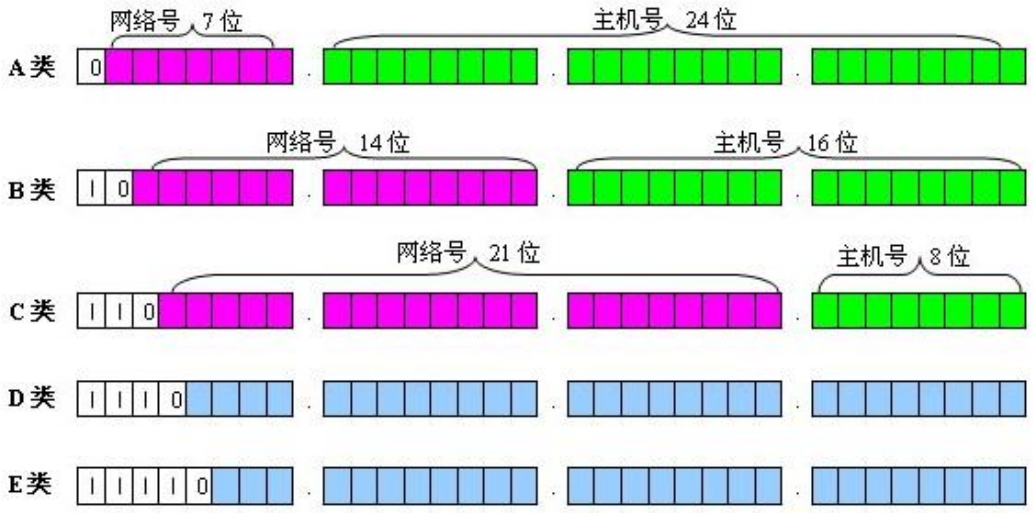
MAC 地址	所在端口
MAC A	1
MAC D	4
MAC B	2
MAC C	3

交换机会根据数据帧的目的 MAC 地址查找地址表，决定从哪个端口转发数据，如果查找不到就广播给除源端口之外的所有端口，所以交换机的各个端口还同属于一个广播域。
交换机有一定的缓存能力，能够同时存储 ABCD 四个主机发送的数据，这个时候 ABCD 之间便没有了冲突，即便链路繁忙，交换机会在链路空闲时将信息转发出去，所以交换机各个端口不属于同一个冲突域，实现了冲突域的隔离。
我们可以大概了解一下以太网帧结构及各部分定义:
DMAC:目的 MAC 地址； SMAC: 源 MAC 地址;
length/T:大于 1500 时为 Ethernet II，表示类型，小于等于 1500 时为 802.3，表示长度;
DATA: 数据内容; FCS: 帧校验。

以太网帧结构里除了前面说过的源宿 MAC 地址之外，还有长度字节，用来表示这一帧的大小，因为以太网帧的长度是可变的，而 SDH 是固定的大小，所以没有长度一项。对于 802.3 帧，长度的最大值 1500 能看出帧的最大长度是 1.5K 左右，最后的 FCS 帧校验用来确保数据在链路层的传输是可靠的。
二层交换的方式被形象的称作一通信基本靠吼。二层交换在规模较小的局域网中非常适合，转发通过硬件来实现，速度快效率高。但是如果网络中的主机非常多，这样的机制就有了很大的问题。交换机各个端口属于同一广播域，主机过多，广播就会频繁占用带宽资源，造成带宽的浪费。如果有一个人不停的在办公室里喊，谁是张三有人找，谁是李四来一趟，大家一定会觉得太扰民，而且公司大了，那面人员流动性也很大，人员的大量变动对交换机来说都要一一的去学习，交换机表示压力很大。

3.4 IP 网基础—三层路由

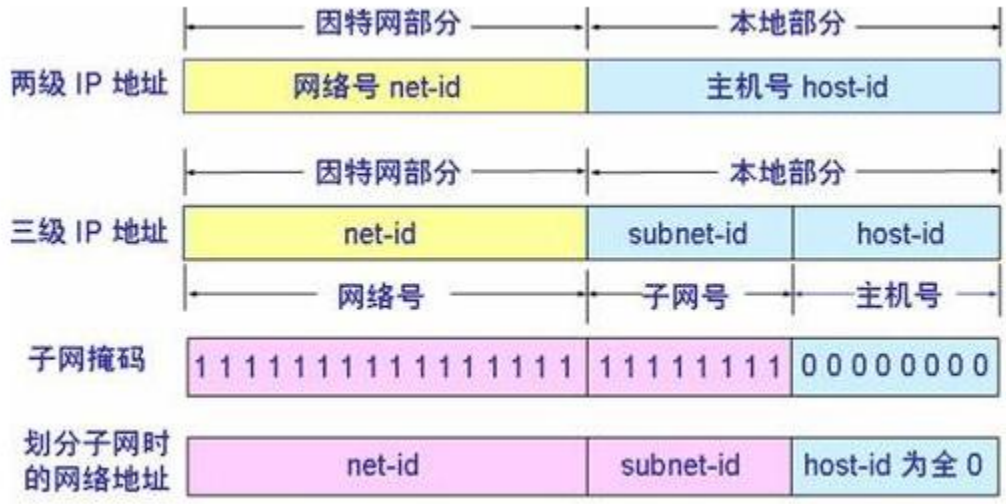
依靠二层交换机，我们可以组建一个小型的局域网，可以包含数十台主机。但如果要组建一个城市、国家甚至到全球的网络，二层交换技术由于其广播的工作方式，必然导致网络拥塞甚至瘫痪，必须将广播域隔离在一个小范围之内，将广域网划分为无数个小型的局域网。而局域网之间的互通，就需要三层路由来实现，实现路由功能的设备叫做路由器，路由器根据三层 IP 地址实现寻址功能。
了解三层路由之前，首先我们来了解一下路由使用的地址--IP 地址。
IP 地址是 IP 协议提供的一种统一的地址格式，它为互联网上的每一个网络和每一台主机分配一个“逻辑地址”（和物理地址对应），一台主机的 IP 地址是随着其所处的位置不同而更改的，就像我们的通信地址会随着我们所处城市的改变而变化。
IP 地址是一个 32 位的二进制数，通常被分割为 4 个“8 位二进制数”（也就是 4 个字节）。为了读写方便，IP 地址通常转换为“点分十进制”表示成(a.b.c.d)的形式，其中，a,b,c,d 都是 0~255（00000000-11111111，共 28=256 个）之间的十进制整数，例如我们熟悉的 192.168.0.1。
InternetNIC 在全球范围内统一分配 IP 地址，将 IP 地址空间划分为 A、B、C、D、E 五类，其中 A、B、C 是基本类，D、E 类作为多播和保留使用。
InternetNIC 规定一部分 IP 地址是用于局域网使用，也就是私网 IP，不在公网中使用，分别是： 10.0.0.0 ~ 10.255.255.255， 172.16.0.0 ~ 172.31.255.255， 192.168.0.0 ~ 192.168.255.255。这些私网 IP 大家应该都很熟悉，一般我们在公司里配置的 IP 地址就是这几个其中之一，私网 IP 就像我们在办公室里可以称呼老王、小张、大个、眼镜儿等等这些代号，在其他公司内部也有这些称呼，但是出了公司到外面就都要称呼大名，否则人家不知道你在叫谁。



从图可以看出，每类地址的网络号和主机号的位数不同，一个 A 类地址可以容纳 16777216（224-2）台主机，B 类容纳 65534（2 16-2）台主机，而 C 类可以容纳 254（28-2）台主机。减 2 是因为一个网络里主机号全 0 表示网络本身，全 1 表示广播，这两个地址不能占用。路由器之所以能够组建大网，实现全球内数以亿计的主机之间的通信，正是因为一个 IP 地址可以包含多台主机，例如一个 A 类地址 110.0.0.0 就包含了从 110.0.0.1 至 110.255.255.254 共 1600 多万台主机，在路由器的路由表中只要一行就可以表达。而一个主机必须对应一个 MAC 地址，假设有 1600 万台主机连在一个交换机下，交换机的 MAC 地址表就要有 1600 万行。

A、B、C 类公网地址中，主机数最少的 C 类地址也包含 254 台主机，而我们实际应用中一般将一个部门或者一个办公室作为一个局域网，网络规模都没有这么大，这无疑对 IP 地址资源造成一种浪费，所以又引入了子网的概念。子网就是将 IP 地址中的主机号部分再划分出一部分作为子网号，这样一个 IP 地址又可以划分成多个子网来使用，比如一个 256 的 C 类地址可以分成 4 个 64 台主机的子网，也可以分成 8 个 32 台主机的子网，这样划分 IP 地址的使用效率大大提高。

我们在接入一个网络的时候，需要先设置子网掩码，子网掩码就是用来表示 IP 地址中哪些位表示子网号，哪些为是主机号，从而将一个 IP 地址细分使用的。



将一个局域网内的主机划分为若干个小的局域网，还可以用划分 VLAN 的方式。VLAN 顾名思义就是虚拟局域网， VLAN 是通过在帧结构中插入一个 VLAN 标签（802.1Q），通过 VLAN 号来判定哪些主机属于同一个 VLAN，VLAN 可以通过 IP 地址来划分，也可以通过 MAC 地址来绑定或者通过端口来随意的划分局域网。

有了 IP 地址将所有的主机划分成一个个的网络（或者子网），这时就有了内网和外网的概念，同一网络中的主机处在一个内网里，而其他网络的主机对于本网的主机来说就都是外网--外面的世界。

内网的通信靠二层交换来实现，而要与外网的主机通信，外网的主机交换机是无法广播到的，就需要经过路由器，通过路由来到达。路由器的各个接口连接的是一个个的网络，路由器就像不同网络之间沟通信息的中转站。

大家知道，我们的计算机要上网都要事先设定网关的 IP 地址。网关，顾名思义，就是网络关口，也就是内网和外网之间的那道门，是二层网络和三层网络之间的桥节点，所有要出入本网络的数据均由网关来转发，就像海关一样，里面和外面的世界我不管，但是要进来和出去的，就必须经过我这一道关口。

交换机可以通过 IP 地址判定数据包是否是局域网内，如果目的 IP 地址不是本网络（或子网），就会将目的 MAC 地址改为网关的 MAC 地址，将该数据包转发给网关，再由网关转发出去。网关可能是路由器或者具有路由功能的三层交换机，也可能是一台主机上安装了两个网卡，一个对内一个对外。

假设你的名字叫小明，你住在一个大院子里（内网），你的邻居有很多小伙伴（内网主机），当你想跟院子里的某个小伙伴玩，只要你在院子里大喊一声他的名字，他听到了就会回应你，并且跑出来跟你玩（二层交换）。

但是你家长不允许你走出大门，你想与外界发生的一切联系，都必须由父母（网关）帮助你联系。假如你想找你的同学小静聊天，小静家住在很远的另外一个院子里，他家里也有父母（小静的网关）。但是你不知道小静家的地址，不过你的班主任老师有一份全体同学的名单和地址对照表，你的老师就是 DNS（域名解析系统，将域名翻译成 IP 地址）服务器。于是你在家里和父母有了下面的对话：

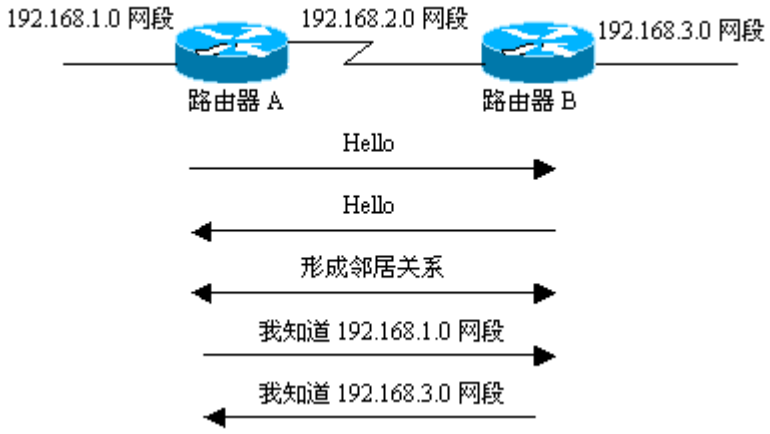
小明（内网主机）：妈妈我想和小静玩行吗？

家长（网关）：好，我打电话问一下老师，小静家住在哪里。

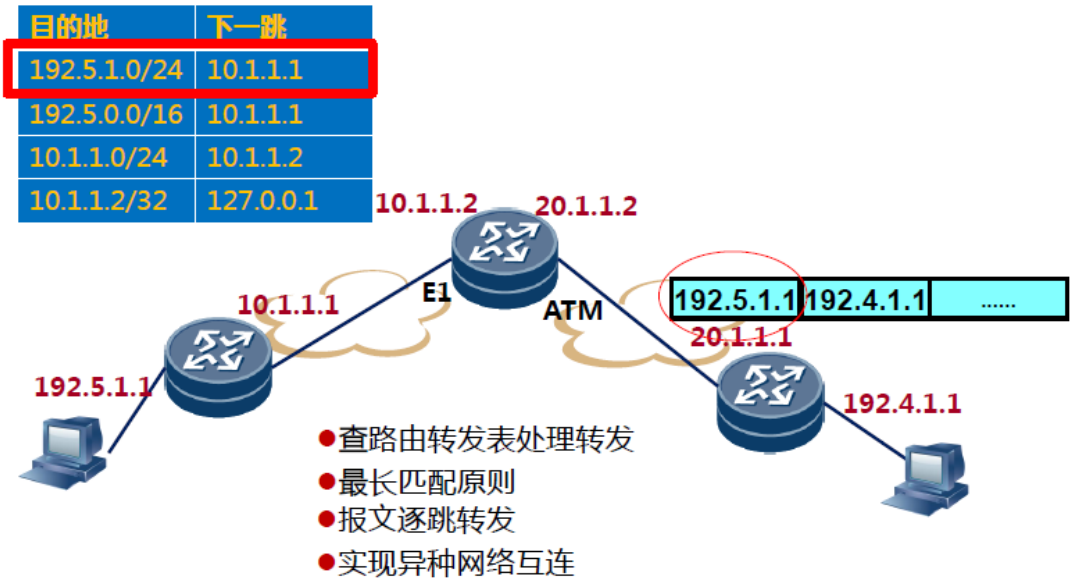
老师（DNS 服务器）：小静家的地址是……

小明：太好了！妈妈我去找小静玩了。

妈妈：可是孩子，你不知道去小静家的路怎么走啊，我带你一起去吧。
于是妈妈带领着小明，拿着地址去找小静，路上碰见了好多好心的警察叔叔（路由器），警察叔叔告诉小明小静家应该往哪个方向走，最终，小明找到了小静一起愉快的玩耍。
警察叔叔对城市的道路都了然于胸，可是路由器是如何知道下一步应该往哪个方向走呢？路由器在初始化时会启动某种动态路由协议，动态路由的成功依赖于路由器的两个基本功能：维护路由选择表，向其他路由器发布路由表。一个路由选择协议的内容包含了： 如何发送更新信息（发送的过程的约定）、更新信息里包含哪些内容、什么时侯发送这些信息（触发更新、定期更新）、这些信息发送给谁（广播、组播）等等。
路由器自动发现周围的邻居，多个路由器之间会交流各自收集的信息，这样，大家将一片片的局部地图拼凑起来，就形成了世界地图，也就是完整的路由表。比如说，我将“我有个邻居叫张三”的消息告诉另一个邻居李四，李四就知道了张三是他的邻居的邻居，张三到李四的距离是 2 跳。



路由器通过静态配置或动态路由协议生成路由表，路由表和 MAC 地址表类似，是 IP 地址和端口的对应关系，路由表形成以后的工作方式就和交换机差不多了，发过来的数据包和路由表进行匹配查找，在相应的端口将数据转发出去。



路由的过程就像我们在高速公路上行驶，我们在经过每一个岔路口（路由器）的时候，都会有一个指示牌（路由表）将我们指引到正确的方向，其实我们并不知道目的地具体在哪里，也不知道从起点到终点全程的路线怎么走，只要我们在每个岔路都知道选择正确的方向，就一定会一步一步的接近目的地，在最后将要到达的一个指路牌，才会看到目的地。

可能有人会问,路由器这么强大,为什么还要用二层交换呢,每一个主机都连接到路由器上,全部使用路由器进行三层组网，岂不是更简单？试想一下，办公室里有一个张三，站在同事李四的面前，左手拿着手机百度地图，右手拿着你的通信地址--中华人民共和国 XX 省 XX 市 XX 区 XX 路 XX 号 XX 大厦 XX 楼 XX 室李四，张三自言自语：“我要怎么才能找到他呢？先看看地图吧，中国在哪里？”……最后折腾了一气终于找到了李四，估计眼泪都掉下来了，世界上最遥远的距离就是，你就在我的面前，而我却用路由的方式寻找你。很显然，直接喊一声“谁是李四？”这个问题就解决了。

能简单粗暴解决的问题就没有必要复杂化，所以在局域网内，使用交换机会提高转发效率，而且交换机的成本也要远低于路由器。

还有一点，我们知道 IP 地址的资源是有限的，人家美国的 internet 组织划分 IP 地址的时候近水楼台先得月，43 亿 IP 地址北美占了大概 3/4，而我们 14 亿人口的泱泱大国一共才 2.5 亿个，是可忍孰不可忍，所以我们发扬了勤俭节约的优良传统，通过二层交换+三层路由的方式，一个局域网可以只使用一个公网地址，而在局域网内部可以使用私网地址，可以大量的节省 IP 地址资源。数据包在从内网到外网时，需要做一个私网到公网地址的转换。

路由器的路由表中有三种路由：静态路由、动态路由和缺省路由。

静态路由：管理员手工定义到一个目的地网络或者几个网络的路由，静态路由不能对线路不通、节点变化等路由变化作出反应。静态路由是路由器的私有路由，一般不向外广播，一般在到达某个网络的路径唯一的情况下，使用静态路由。

动态路由：路由器根据路由选择协议所定义的规则来交换路由信息，并且独立地选择最佳路径。动态路由协议是智能的、自动的，在网络节点发生变化或链路故障时，路由协议会自动计算和更新拓扑结构，重新计算可用的路由，而不需要人工维护。

缺省路由：当路由表中与包的目的地址之间无匹配的表项时路由器的选择，地址在路由表中查不到，就按这个路径转发。

一般,路由器查找路由的顺序为静态路由,动态路由,如果以上路由表中都没有合适的路由，则通过缺省路由将数据包转发出去。

动态路由选择协议的分类:

IGP 和 EGP: 根据自治域来划分，路由协议可以分为内部网关协议（IGP）和外部网关协议（EGP）。自治域内部采用的路由选择协议称为内部网关协议，常用的有 RIP、IGRP、EIGRP、OSPF; 外部网关协议主要用于多个自治域之间的路由选择，常用的是 BGP 和 BGP-4。

自治域: 自治域系统是指处在一个统一管理的域下的一组网络的集合，说白了就是我的底盘我做主，比如联通和电信就是不同的自治域，一个公司内部也是一个自治域，因为他们无法管理对方的网络设备。

有类/无类路由协议：根据是否支持子网掩码的传播，路由协议可以分为**有类路由协议**和**无类路由协议**。有类的路由协议包括 RIP v1、IGRP 等。这一类的路由协议不支持可变长度的子网掩码，不能从邻居那里学到子网，所有关于子网的路由在被学到时都会自动变成子网的主类网（按照标准的 IP 地址分类）。无类类的路由协议支持可变长度的子网掩码，能够从邻居那里学到子网，所有关于子网的路由在被学到时都以子网的形式直接进入路由表。

也就是说有类路由协议会无视你划分的子网，你将一个 C 类地址划分的多个网段，有类路由协议都直接还原成 C 类地址。

距离矢量路由协议和链路状态路由协议: 路由器会得到很多条可到达目标的路由信息，一条路由失效时可以通过其他路径把数据包传递到目的地，确保网络的畅通。在众多路由中，路由器要选择出一条或多条最佳的路由添加到路由表中,可是不同路由协议认为的最好的协议可能是不同的，因为衡量路由间的“好”与“坏”的标准是不同的，依据不同的标准给路由去打一个分数也是不同的，这个“分数”就是**度量值**。

所谓**度量值（value），**就是路由器根据自己的路由算法计算出来的一条路径的优先级。当有多条路径到达同一个目的地时,度量值最小的路径是最佳的路径,被路由器添加进路由表。路由器中最常用的度量值包括: 带宽（bandwidth）、 延迟（delay）、负载（load）、可靠性（reliability）、跳数（hop count）、开销（cost，一个自定义的任意值）等 。

比如 RIP（路由信息协议）是一种距离矢量协议，RIP 使用的度量值就是跳数，而不关注带宽、可靠性等指标，最大支持 16 跳，只能应用于小型网络，在距离超过 16 跳之后 RIP 就认为路由是不可达的。

而 OSPF 是一种链路状态协议，使用基于带宽的度量值计算出 cost，比如 10M 的链路 cost 值为 10，100M 链路 cost 值为 1。OSPF 路由器之间交换的并不是路由表，而是链路状态，比如接口上的 IP 地址，子网掩码，网络类型，Cost 值等等。OSPF 通过获得网络中所有的链路状态信息，从而计算出到达每个目标精确的网络路径。

二层交换+三层路由的方式我们已经大致了解，要进一步的了解 IP 的世界，里面的技术和概念还很多，本人也非数据专业，只能略表一二，更进一步的了解还需要大家查阅更深入的相关资料，本文的初衷还是让大家能够对传送网相关的各个模块有一个初步的认识，起到一个穿针引线的作用。

3.5 从 IP 到传送网

前面对 IP 网的工作方式有了一个大概的了解，从本节起，我们要逐渐回到我们的主题：传送网。

IP 网是一个成熟开放的强大的技术，而另一方面，我们的传送网的业务也已经逐渐 IP 化，业务的 IP 化推动着传送网向 IP 化的转型，那么一个新的问题产生了，如果用 IP 网络直接去承载我们的电信业务，是否可行呢?或者说，还存在哪些问题?

我们知道，IP 网是一张面向无连接的网络，是一张尽力而为的网络，而我们的传送网承载的是以基站回传为主的业务，是对安全性要求很高的号称“电信级”业务，这尽力而为的网络去承载电信级业务，就是最大的问题所在。

我们宽带上网，邮件发不出去、下载文件中断的情况都是时有发生的，就算整个网断了，最多也是打个电话报修或者投诉一下，维护人员来修就是了。宽带网从用户到 OLT（光线路终端，宽带接入网设备）之间都是没有保护的，线路断了业务就中断，就只能抢修，毕竟我们就几十块一个月的的包月费，很多时候还是免费蹭来的 WIFI，我们又能说什么呢?

可是无线语音业务不同,如果基站掉站,附近的用户都打不成电话,如果通信中断几个小时，在运营商来讲就是重大事故，举个例子，我们来看一个新闻报导:

新华网快讯：12 日四川汶川县发生 7.8 级强烈地震已经 6 个多小时，目前通信仍处于中断状况，记者用固话、手机进行联系，均没有成功。

又一则快讯：中国电信汶川分公司安排精干人员，在持续的余震中，在已经成为危房内的机房里,用了半个小时修通了设备,县城内的固定电话及小灵通恢复正常使用。由于电力中断，他们启用了油机(柴油发电机)发电。

这些报导说明了什么，通讯和电力一样，是关系到民生的大事，怎么没有报道宽带中断几个小时？所以为什么我们一直强调 “电信级业务”这个概念。

从另一个角度说，打电话每分钟都是钱，无线 3G、4G 的数据业务也是每 Mbit 要几毛钱，从运营商的收益上来说也是产生巨大经济效益的业务。
说了这么多其实就是一句话，IP 网安全性不够，我们要的不是尽力而为，而是必须保证。另外一点，我们知道，传送网是有着很强大的网络管理系统的，能够对网络的故障进行精准的定位，并且有着强大的保护倒换作为保障。可是对于 IP 这个无连接的网络来讲，这些都是无法做到的，每秒钟在网络中传递的天文数字级别的数据包，在网络中的行踪都是不确定的，连数据包去哪了都不知道，管理又从何说起呢。这就像我们可以对列车、地铁严格的管控，每隔几分钟一班车，车辆的运行情况都是掌握的准确无误的，可是对一个城市所有的机动车来讲，要实现相同级别的管理就是痴人说梦了。
综上所述，首先要解决的就是这个有连接和无连接的问题。我们需要一个办法，让无连接的 IP 网变成有连接，以便实现维护管理、保护倒换等传送网所要求的一系列功能，提高安全性，实现这一步质变的技术基础就是 MPLS。

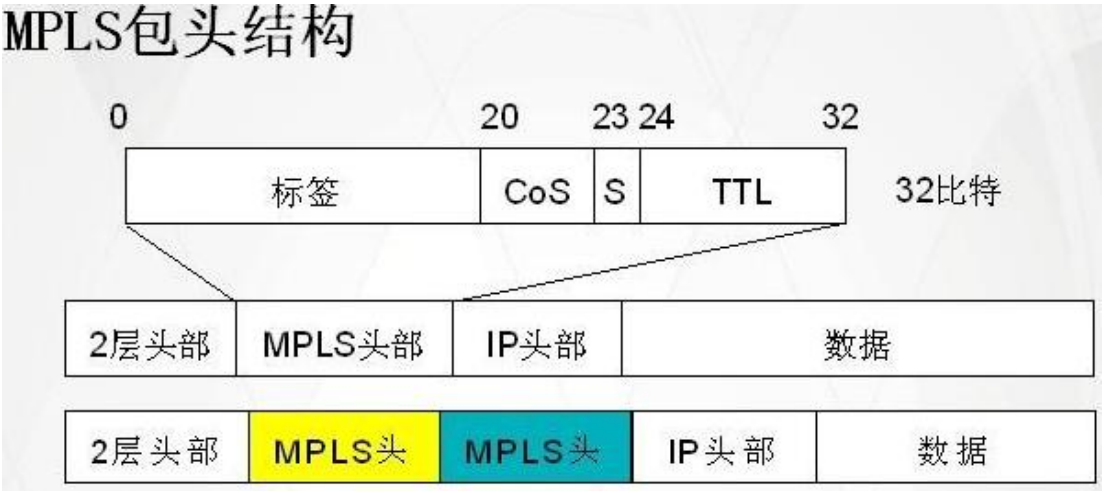
3.6 MPLS 多协议标签交换

MPLS—(Multi-Protocol Label Switching)多协议标签交换，是一种用标签交换代替路由，实现数据包快速转发的技术体系，它的价值在于能够在一个无连接的网络中引入连接模式的特性；其主要优点是在提供 IP 业务时能确保 QoS 和安全性，具有流量工程能力。
举个例子，我们在某网站上下载一部几个 G 的电影，文件会分成至少几百万个数据包从服务器传送到你的电脑，这几百万个数据包的源地址和目的地址都是相同的。按照传统的路由方式，对于每一个数据包路由器都需要查找一次路由表，决定从哪个接口转发，一共要进行几百万次的路由，这样的方式显然是很浪费路由器的能力资源的。
这就像一百个人走到一个地方问你路，他们的目的地是同一个地方，你就没有必要给每个人都说一遍怎么走，只要给第一个人指路之后，告诉后面的人：全部跟着第一个人走就好了。对于数据包的路由，我们同样可以只对第一个数据包进行路由选择，后面的相同目的地的数据包都贴上相同的标签，按照第一次路由的结果进行转发。
MPLS 技术最初就是为了提高路由器的转发效率而问世的，从另一方面我们也不难发现，对于 MPLS 网络来说，到相同的目的的数据包被分发相同的标签，走相同的路径，这样就在 IP 网中打通出一条虚拟的“路（LSP）”出来，也就是将 IP 网从无连接变成了有连接的网络，那真是，世界上本没有路，走的人多了，便变成了路，有了“路”之后在此基础上对数据转发路径的管理也就可以实现了。
总结一下：标签交换的效果有 2 个：提高了转发效率，无连接变为有连接。

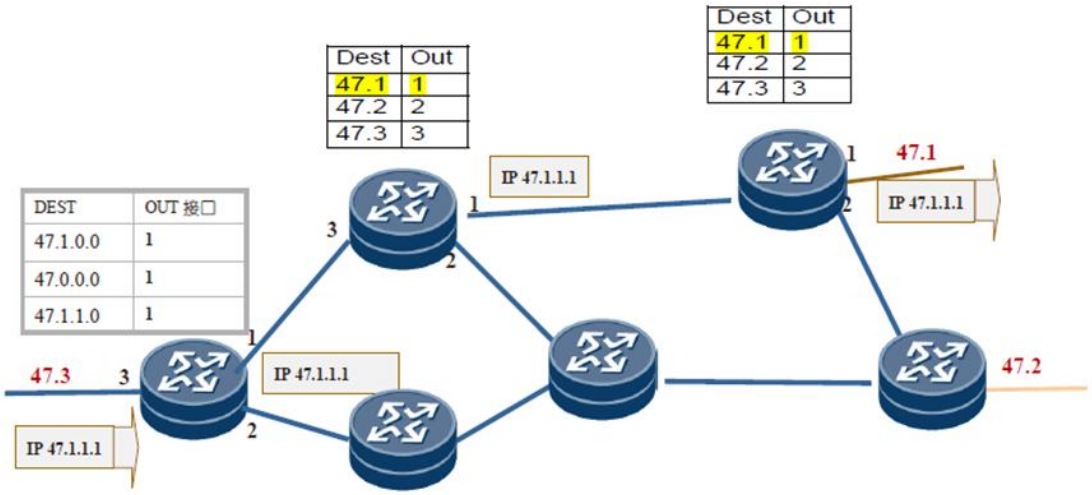
MPLS 相关名词、术语的解释:

多协议： MPLS 不但可以支持多种网络层协议，还可以兼容第二层的多种链路层技术。
标签交换: 它提供了一种方式，在 MPLS 边缘的路由器（LER）将 IP 地址映射为简单的具有固定长度的标签，将标签添加到二层和三层帧格式之间，在 MPLS 网络的内部标签交换路由器（LSR）使用标签快速交换代替路由。
QOS: 服务质量(Quality of Service)，是一种网络安全机制。当网络过载或拥塞时，QoS 能通过业务的优先级划分，确保重要业务量不受延迟或丢弃。克拉玛依火灾中“让领导先走”就是 QOS 在现实中的生动的例子。
流量工程: 合理分配流量，保证网络资源得到充分的利用，避免网络过度闲置和拥塞。与流量工程对应的是网络工程，如果将网络工程比作道路新建、改扩建的话，流量工程就是缓堵保畅。
FEC: 转发等价类，如果入口路由器收到分组都到达同一子网的，则这些分组就属于同一类，叫做转发等价类，同一 FEC 的分组都会转发给同样的下一跳。前面问路的例子中，那 100 个人就属于同一个转发等价类。

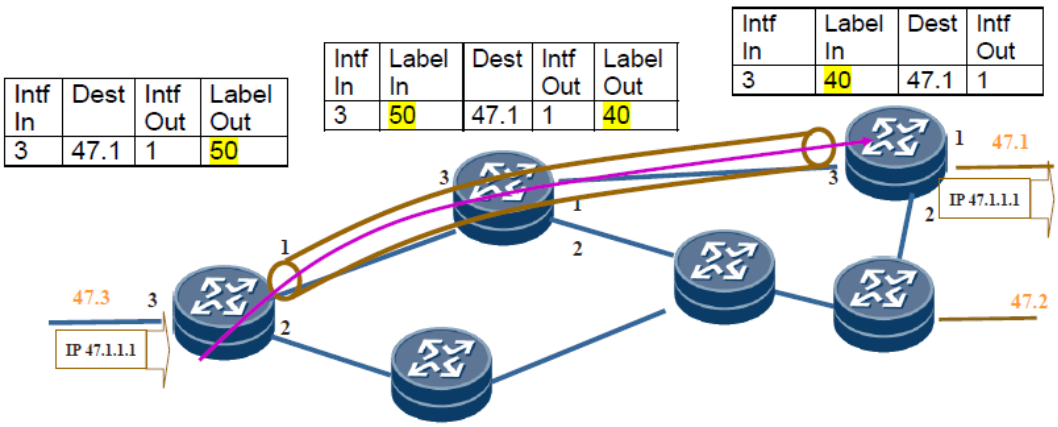
LSP: 标签交换路径（Label Switched Path），一个转发等价类在 MPLS 网络中经过的路径。
LER: 标签边缘路由器（Label Edge Router），位于 MPLS 域边缘连接其它用户网络的路由器，主要完成连接 MPLS 域和非 MPLS 域以及连接不同 MPLS 域，实现 FEC 划分，分发标签，剥去标签。
LSR: 标签交换路由器（Label switching Router），MPLS 区域内部的路由器，负责标签交换和标签分发。
MPLS 的原理:
MPLS 的原理概括下来就是：一次路由，多次转发（交换），具体过程如下:
MPLS 根据数据包的流向的路径，按照某种方式为每个路由器分配标签，同一 FEC 可以对应一个或者多个标签（为了负荷分担），但是不同 FEC 一定对应不同的标签。
标签分发方式有两种，下游按需标签分发（DOD）和下游自主标签分发（DU），DOD 就是收到上游标签分配申请才分配标签，DU 是下游自主分配标签，无需上游申请。
标签分发之后 LSR 需要将 FEC 与标签映射关系向其他 LSR 通告，通告分为两种方式：独立 LSP 控制和有序 LSP 控制方式。独立控制是指 LSR 分发标签后即通告给上游 LSR，有序控制是指 LSR 需要收到下游 LSR 对此 LSP 通告后向上游发送映射通告。
LSR 对于收到的 FEC 映射标签，可以完全保留，也可以只选择 LSP 的映射标签保持，前者叫做自有保持方式，后者叫做保守保持方式。
标签的分发、通告、保持方式，涵盖了 LSR 对于标签处理的不同方面，假设一条 LSP 从上游至下游依次经过 A、B、C、D、E 多个 LSR，分发方式 DOD 就是只有路径上这些路由器才会为此 FEC 分发标签，而 DU 方式是指其他非 LSP 上的节点也会为此 FEC 分配标签；独立控制方式是指 ABCDE 同时向上游通告，而有序通告则需要 E 通告给 D 之后，D 才会给 C 通告，以此类推。自由保持方式是指收到的映射全部保留，保守保持是指 A 只保留 B 通告的映射，将其他的删除。
其实这些方式都是协议内部的事，我们只需要知道，大家通过一系列的流程，获得了对于一个 FEC 的标签分配方式。



标签全部分发通告完成之后，LSR 就可以根据标签进行交换：进入 MPLS 网络的数据包，在 LER 处根据 IP 头部判定 FEC 并查找标签、出接口，将标签插入到二层头和三层头部中间，从出接口转发出去。
带有标签的 MPLS 包在 LSR 路由器处，只查找入标签对应的出标签和出口，进行标签交换（用入标签替换出标签），在出接口发送分组，标签在离开 MPLS 域的 LER 处被剥离。由于标签交换的原理类似于二层交换，都是固定字符的匹配查找，所以数据转发的效率会大大提高。
传统路由方式



MPLS 标签转发



MPLS 协议栈

同路由的分类一样，标签交换路径 LSP 分为静态 LSP 和动态 LSP 两种。静态 LSP 由管理员手工配置，动态 LSP 则利用路由协议和标签发布协议动态产生。

动态的 LSP 功能实现模块由两部分组成：控制单元和转发单元。

控制单元负责标签的分配、路由的选择、标签转发表的建立、标签交换路径的建立、拆除等工作，转发单元负责按照 LIB（标签信息表）替换标签并转发数据包，类似于二层交换机的功能。

控制单元协议包括路由协议和信令协议。路由协议就是前面介绍的 OSPF、ISIS、BGP 等，路由协议通过路由器之间的信息交互形成路由表、网络拓扑，收集链路状态信息。

信令协议（标签分发协议）是 MPLS 的控制协议，负责 FEC 划分、标签的分配以及 LSP 的建立和维护等一系列操作，MPLS 可以使用多种标签发布协议，包括专为标签发布而制定的协议，例如：LDP（标签分发协议）、CR-LDP（限制路由的标签分发协议），也包括现有协议扩展后支持标签发布的，例如：RSVP-TE（扩展资源预留协议），MP-BGP（多协议扩展 BGP）。

通过路由协议扩展，可以在每台路由器上维护网络的链路属性和拓朴属性，包括最大链路带宽、最大可预留带宽、当前预留带宽等，形成流量工程数据库 TED，路由决策都是由 IGP 协议及其 TE 扩展和 CSPF 作出的。

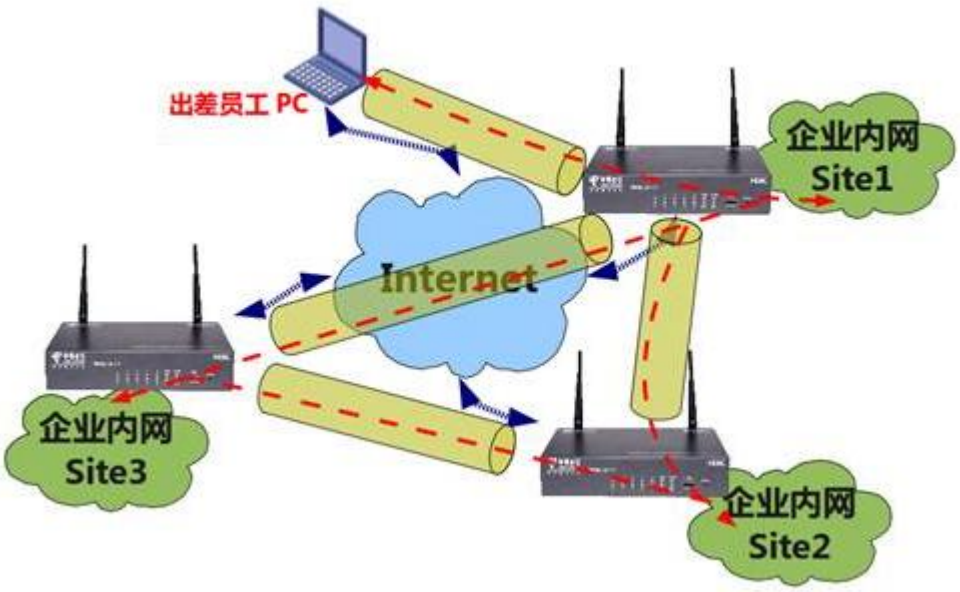
其中 LDP 仅负责标签分发，不去关心链路状态，也无法指定经过的路由，仅能够实现快速转发的功能，如果需要支持显式路由、流量工程和 QoS 等业务时，就必须使用其他标记分发协议。CR-LDP 和 RSVP-TE 在功能上比较相似，但在协议具体实现上不同。

MPLS 利用标签在 IP 网中打通了有连接的隧道，但是我们传送网要承载无线、宽带、大客户专线等业务，可能在一个运营商站点下面同时下挂了某大客户 CE 路由器、移动基站 NodeB、宽带接入网 OLT 等设备，从业务安全的角度，MPLS 网中这些业务彼此之间是需要相互隔离的，我们需要为数据包多加一层标签---私网标签，私网标签用来区分不同的业务，比如 A 公司和 B 公司之间通过私网标签形成两个相互隔离的 VPN。

举个例子，A 公司在某酒店举行年终总结大会，在某酒店定了 30 个房间，那这个会议的所有成员就相当于一个 VPN，在酒店前台就会有针对 A 公司会议的 30 个房间号对应的人员的清单，这个清单就相当于 VRF 表，到达酒店的客人只要报出 A 公司名称（相当于 VPN 号），就能且仅能与会议成员取得联系。虽然酒店有 100 个房间，但是在会议来看，可以看成一个 30 个房间的专用酒店。

这样用户数据包虽然在 MPLS 网络中在同一隧道中传送，但是在运营商网络和用户 CE 连接处，PE 为每个 VPN 都维护一个单独的 VRF（虚拟路由表），表中只有本 VPN 对应的站点，这样就达到了隔离的效果。

下面一节，我们来了解一下 VPN。



3.7 VPN（虚拟专用网）

专网的概念我们都有所了解，我们国家的军队、铁路、电力等单位都有自己的专网，这些单位对业务的安全性要求非常高，这类的专网是单位自己花钱建设的一张独立的网络，物理上与公用网络是隔离的，独立建设、管理、维护，所以安全性和带宽都有很高保障，但是付出的代价无疑是最高的，不过这些土豪不在乎，有钱就是这么任性。

另外一些企业，比如各大银行系统，采取了租用运营商 SDH 电路等方式来搭建专网，虽然物理网络不是专用的，但是 SDH 的 E1 通道是用户独享的，也使业务的带宽和安全性能够得到保障，效果与土豪的专网也差不多。

租用 SDH 电路的价格也不菲，对于更多公司来讲还是难以接受，但从业务的需求来讲，公司各个分支机构之间需要一个网络来连接，需要承载视频会议系统、邮件系统、各种办公系统

等需求，对网络的服务质量还有着较高的要求，这类的客户需要一个经济实惠又能够达到专网效果的办法—VPN。

VPN（虚拟专用网）是一种逻辑上的专用网络，但本身却不是一个独立的物理网络。VPN 就是在利用公共网络建立虚拟私有网，就是用某种技术在公网上面建立一条条的虚拟连接，将公司的各个分部连接起来， VPN 用户与其他用户互相视而不见，也就是逻辑上的隔离。

用道路举例来说:

城市交通道路就是公网，大家的车都可以在上面跑;

铁路、地铁就是专网，你再有钱也不能买个列车上人家轨道上去开;

而公交专用道就相当于 VPN。公交专道是在城市道路中划出一条车道专门走公交车，在高峰期拥堵的时候其他禁止车辆行驶，利用这种专道专用的规则达到了专线的效果，从成本上来说很明显比单独修一条路要小得多。

MPLS VPN

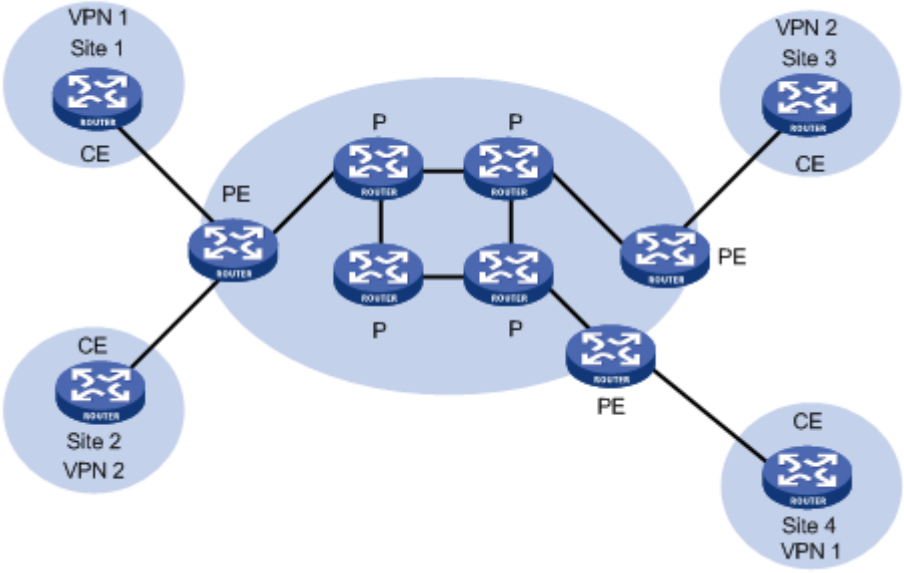
MPLS VPN 是 VPN 的一种实现方式，MPLS VPN 通过对不同 VPN 用户分配两层标签，即公网标签和私网标签,其中公网标签处于外层，私网标签处于内层，私网标签用来实现不同 VPN 用户的隔离。公网标签用于在 PE 设备之间形成数据传输的隧道，而私网标签则用于 PE 对不同 VPN 用户数据的区分。

二层头部	公网标签	私网标签	IP 头部	数据

要了解 MPLS VPN 的工作过程，先要了解几个 VPN 的概念:

首先，运营商网络内部的设备分为两种: P(运营商核心路由器)和 PE(运营商边缘路由器)。这里说的 P 和 PE 是 VPN 对不同位置和功能定位的设备的叫法，P 和 MPLS 里的 LSR 对应，PE 和 MPLS 里的 LER 对应。就像是一个公司的老大，公司内部管理来说管他叫董事长，从法律来讲叫他法人代表，其实是一个人，强调的点不同。

PE 就是和用户设备直接相连的运营商设备，而 P 是运营商网络内部的设备，不和用户设备相连，PE 就相当于一个公司的业务员，而 P 就相当于内勤人员。CE 是用户边缘设备，也就是用户自己的设备中和运营商相连的那个。一个 PE 下的一个 VPN 用户可能有一个或多个 CE，这些 CE 就称之为一个 site（站点）;



比如我开一个公司，北京有 2 台路由器，上海有 3 台路由器，想通过运营商去开通一个 VPN，将这 5 个路由器连接起来。这 5 台路由器就都是 CE，北京的 2 台连接到运营商的一台路由器上，运营商的这个设备就是 PE，上海同样也有一个 PE；我北京这 2 台 CE 叫做一个 site，上海那 3 台 CE 也是一个 sete。而运营商在北京和上海之间除了这 2 个 PE 的其他设备就是 P 了。

当属于某一 VPN 的用户数据进入 MPLS 主干网时，在 CE 路由器与 PE 路由器连接的接口上可以识别出该 CE 路由器属于那一个 VPN，进而到该 VPN 对应的 VRF 中去读取下一跳的标签，并将标签作为内部标签加入标签协议栈。PE 路由器继续查找自己的全局路由表获得下一跳的接口和标签后，将该标签作为外部标签加入标签协议栈并将加入两层标签的数据包从相应的接口发给 P 路由器。

我公司的数据到了 PE 之后，首先 PE 分配一个内网标签，也就相当于贴上公司的名字。然后我的数据包要从北京到上海，PE 再根据北京到上海的道路情况，给我分配了一个外层标签，这个标签对应一个隧道，也叫隧道标签，就是指引我的数据到达上海的“通行证“。

同样从北京到上海的还有其他公司的数据，大家都穿着同样的外衣—外层标签，P 路由器根据外层标签转发数据包直到出口 PE 路由器，不同公司的数据走着同样的一条路到达了上海。上海的 PE 发现自己已经是最后一站了，将外层标签去掉，就露出了内网标签—VPN 标签，PE 根据这个标签，将不同公司的数据作为一般的 IP 包发给不同公司的 CE，整个通信的过程就完成了。

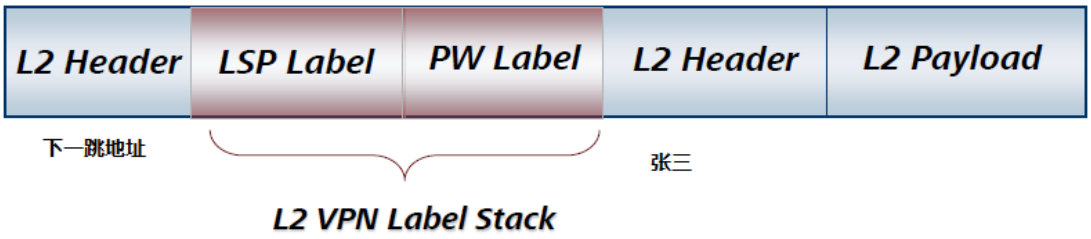
数据传输过程中，内层标签只由 PE 设备进行处理，P 设备并不理会他的存在，也就是说 P 设备并不知晓和关心数据包属于哪个 VPN，或者说，私网标签对 P 来说是透明的。

3.8 PWE3 和 L3 VPN

根据用户业务类型的不同，MPLS VPN 可以采用基于二层或者三层的解决方案，分别叫做 PWE3（L2VPN）和 L3 VPN。两者的区别就是 MPLS 标签在数据包的位置不同，当然工作原理也有很大区别，下面分别说明：

PWE3(端到端伪线仿真)

先放一段“专业”的介绍，PWE3(Pseudo-Wire Emulation Edge to Edge 端到端伪线仿真)是指在分组交换网络中尽可能真实地模仿 ATM、帧中继、以太网、低速 TDM 电路和 SDH 等业务的基本行为和特征的一种二层业务承载技术。



从帧结构图中可以看出，PWE3 是在二层头部之前加上了隧道和 PW 标签之后，又加上了一个二层头部，这两个二层头是不一样的。

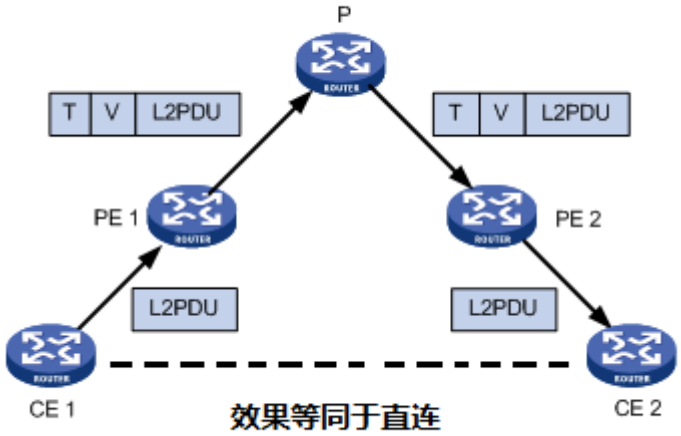
我们先从最常用的点到点的业务模型讲起，前面讲过二层地址就相当于一个人的姓名，假设我要去找一个叫张三的人，在要进入 VPN 网络之前，我的二层目的地址就是张三，这个地址在整个过程中是不变的。张三在哪里我并不知道，我只知道这个名字，而 VPN 网络一定预先为我铺好了一条通往张三的路，否则我和网络都不知道要去哪还怎么通信呢。

数据包从 CE 到达 VPN 的 PE，PE 查找预先配置好的“业务映射表”，就是这个端口上来的人要去找张三，对应的 VPN 通行证和隧道都是配置好的，PE 根据业务映射表为数据包打上两层标签：隧道标签和 PW 标签之后，又加上一个二层头，这个头是在网络中每一跳的 MAC 地址，数据的每一次转发都会将这个地址更改为下一跳地址。

在到达目的地之后，PE 将前面的二层头、两层标签都去掉之后，露出了“张三”这个地址，这时，我的面前一定站着唯一一个人，他叫“张三”。

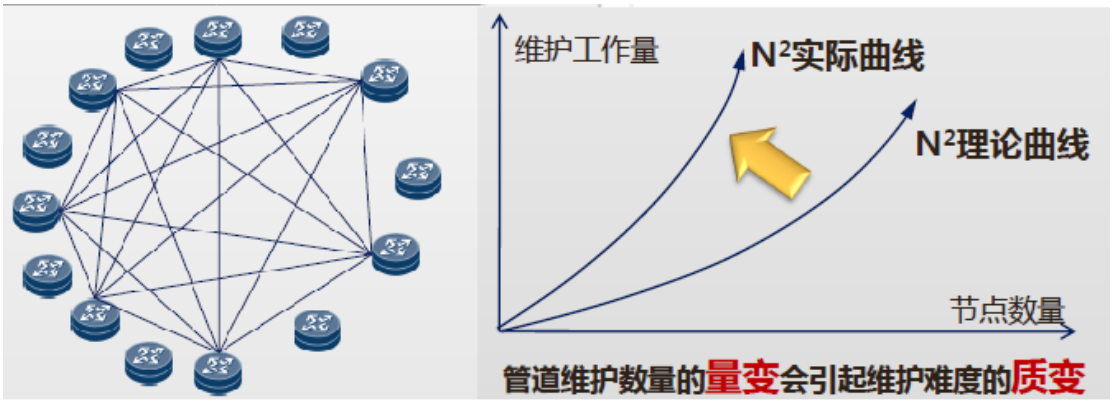
PWE3 就是在分组网络上透明传送用户的二层数据，从用户角度来看，该分组网络就是一个二层交换网络，用户 CE 之间就像通过网线或者交换机实现互联一样。PWE3 应用上更贴近于传送网的功能，公网设备相对于私网设备来说相当于是私网的下层，是透明的，就像我们玩网络游戏的传送门一样，嗖的一声不知怎么的就到了。

如下图所示，CE1 与 CE2 通过 L2VPN 互联，效果同 CE1 与 CE2 用网线直连相同。



对于每一个 VPN 的 site，与其相连的 PE 都有一个其对应的业务映射表，就是一个端口和 PW、隧道的对应关系，根据这个表格，PE 才知道每个端口发来的数据包要去哪里、要发到哪个隧道。PE 只需要配置用户端口和 VPN 的映射关系。

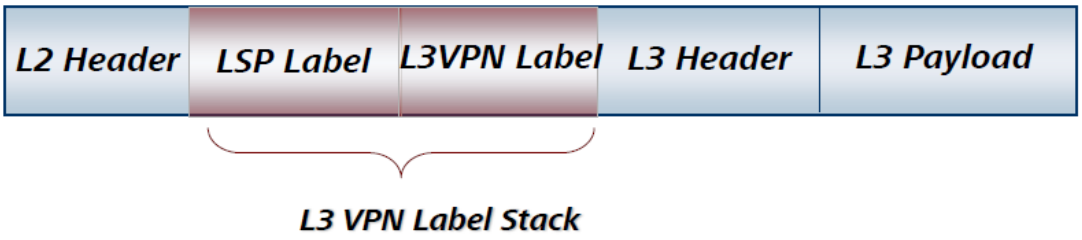
PE 不需要维护 VPN 内部的路由信息，数据包到达 PE 后，PE 根据 PW 标签转发给对应的端口，就完成了它的任务，至于数据包到达用户侧 CE 之后如何送达目的地是用户自己的事，。我要找的“张三”实际上是一个接口人，就像我们下了飞机有一个举着大牌子的接站人员一样，至于张三怎么带着我找到我的最终目的地，那是我们公司内部的事，与 VPN 网络无关。
我要找张三这个例子是一个点对点的业务模型，MPLS 网络只需要建立一条 PW 业务就实现了连接（E-line），这是我们实际应用中最常见的情况。
实际上还有一种点对多点的业务模型，比如我要去找张三、李四……情况稍微复杂一点点，PE 只要相对应的建立 N 条 PW 可以实现（E-LAN），PE 根据目的 MAC 地址去查找对应的 PW 和隧道。
可是还有一种多点对多点的模型，例如 LTE 的 X2 接口业务，N 基站之间都要互相通信，那我们就要建立 N*(N-1)/2 条 PW，如果 N 等于 100，那么就要建立 4950 条 PW，这个工作量就是巨大的，N 要是几千呢？那就不可想象。



所以 PWE3 比较适合点对点和点对多点的业务，而对于多点对多点业务，我们通过 L3 VPN 去实现更适合

MPLS L3 VPN

从帧结构图中看出，L3VPN 是在 IP 头前面加上了隧道和 VPN 标签，前面加上的二层头和 PWE3 里的作用相同。



与 PWE3 的业务映射方式不同，在 L3VPN 中每台 pe 路由器为其直接相连的每个 site 维护一个 vrf（虚拟路由转发表），ce 路由器把站点的本地路由广播到 pe 路由器，并从 pe 路由器上学习远程 vpn 路由，也就是说，你一个公司本地的内部的路由怎么走，远端的每个分支机构都有哪些 IP 地址，对应的要走哪条路，PE 全都要知道，很明显，L3VPN 的 PE 更高大上了，也更忙了。
PE 收到 CE 发来的数据包之后，根据对应的端口识别出属于哪个 VPN，根据目的 IP 地址在 VRF 表中查找路由，然后打上对应的 VPN 和隧道标签转发出去。对端的 PE 接收到数据包之后，根据 VPN 标签知道去查找对应的 VRF 表。
还是上面的例子，L3VPN 来讲我就不是找张三了，我要从北京出发找到我们上海分公司办事处的 X 号楼 X 层 XXX 室。首先我到达 PE 之后，PE 根据端口知道了我是哪个公司的，拿出对

应的 VRF 表一查，上海分公司在哪怎么走，给我贴上了公司名称(VPN 标签)和通行证（隧道标签）之后，一路我来到了上海的 PE。上海的 PE 撕掉我的通行证（已经到达就没用了）之后一看我是某公司的，拿出一个某公司的 VRF 表，对应我要找的房间号，告诉我，你去找这个 CE，PE 参与了一次路由之后也完成了它的使命，最后 CE 根据我的目的地址带领我到达了我的目的地。

L3 VPN 中，公网设备和私网设备在网络层次上相当于处于同一平面，整个提供 VPN 公网有点像是一台路由器。L3VPN 与 PWE3 对比来看，PWE3 的 PE 只知道与之相连的 CE 的端口，只要是发往这个 VPN 的数据包，PE 统统转发给这个端口；而 L3 VPN 的 PE 设备清楚 VPN 内部的 IP 地址，PE 会根据 VRF 表去决定转发的下一跳。

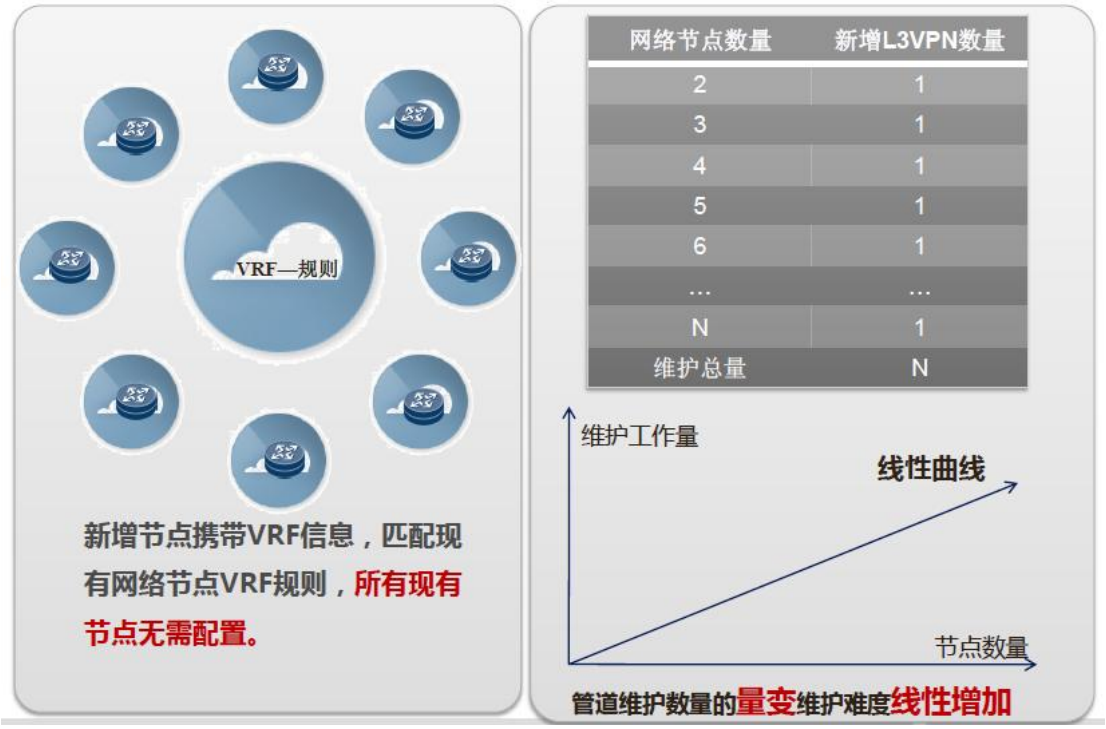
MPLS L2 VPN 与 L3VPN 的对比

从透传的数据单元来看：L3 VPN 透传的是三层 IP 数据，故也只支持 IP 数据透传了；PWE3 透传的是二层数据单元，故能支持很多种业务类型的透传，比如以太网帧，帧中继，ATM，TDM 等；

从适用业务来看：PWE3 主要适用于大客户专线、传统 TDM 业务；L3VPN 适用于例如 LTE NodeB 承载、有 L3 需求的大客户。

从组网应用上来看： PWE3 一般用于点对点、点对多点的简单组网形式，L3 VPN 可以用于多点对多点的复杂组网形式。

比如 A 公司有北京、上海、深圳、南京等分部，如果通过 L3 VPN 组网，每个 site 有多个 CE 与 PE 相连， 这样 A 公司不同分部多个 CE 之间的互通都由 MPLS 网络的自动路由来实现，整个公司所在同一个 VPN，对于新增站点，只要携带与该 VPN 匹配的 VRF 信息，就会自动加入网络，这样对于网络的开通和维护的工作量都是线性增长的，与上面 PWE3 的平方级增长相比小很多。



最后，为了再次说明 PWE3 和 L3VPN 的区别，我们再举个不恰当的例子，A 市政府经常有领导需要前往火车站乘坐高铁，如果是直接开车前往，这就相当于 IP 模式，可是如果碰到晚高峰期碰见个堵车把行程耽误了，领导很生气，后果很严重。

如果 A 市政府给交警部门打个招呼，在公司到火车站这一路上留出一条车道，其他路段自由放行，这就相当于 PWE3 模式，如果经常还有领导要去机场，那就再由公司到机场再预留一条车道，还是通过 PWE3 实现。

可是问题是，如果政府在市区有很多个分支机构，那么留那么多条车道是不现实的，就不如所有政府的车辆都挂上政府的牌子，只要一上路大家就会自动让出车道，这种方式就是 L3VPN。

3.9 分组传送网

经过了上面诸多章节一步一步的介绍到这里，我们的主题—分组传送网已经离我们越来越近了，通过 MPLS、VPN 等技术对于 IP 技术的改造、强化基本大功告成，使其已经能够满足我们“电信级业务”的承载要求，上一节介绍的 PWE3、L3 VPN 实际上就作为分组传送网的核心技术，并以此为基础针对传送网的需求进行了修改，增加或去掉某些功能，制定了适应电信级业务承载的技术标准。

再回头看看我们对分组传送网的要求有哪些，分组传送网又如何去实现。这里插一句，很多的概念我们不去展开一节一节的介绍了，有些时候我们难以理解一个问题不是我们的信息量太少，相反是相关的术语、概念、协议、原理太多，多的我们别说一一去搞懂，就连分清这些概念的关系、层次都很困难，我们顺着本文的这个主线去了解分组网的本质就达到了本文的初衷，其他的概念一笔带过。

1 带宽统计复用

统计复用是 IP 的天然支持的特性，当然一如既往的支持。

1 端到端的管理和保护

基于外层隧道（Tunnel）和内层 VPN 两层标签，分组传送网实际上也形成了类似 SDH 的分层结构:

分组传送通道层（PTC,Packet Transport Channel），该层对应封装每一条客户侧业务的 PW 或 L3VPN,相当于 SDH 通道层中的低阶通道（E1）; 分组传送通路层（PTP, Packet Transport path），该层对应包含多条业务的隧道，相当于 SDH 通道层中的高阶通道（STM-N）; 分组传送段层（PTS, Packet Transport Section），对应于两个设备的互联接口之间，相当于 SDH 的复用段层; 物理媒介层，和所有网络的物理层一样。

有了这么多分层和对应的数据单元，分组网对于每一条业务、每一条隧道和每两个站点之间的物理、逻辑链路都能够尽在掌握，在每一层都可以传送用于管理、检测、倒换的数据包，再通过冗余的资源配置就实现了保护倒换，比如对于两点之间配置 2 条隧道就可以实现隧道 1+1 保护，对于一条业务配置 2 条 PW 就实现了业务 1+1 等保护功能。

1 QOS 服务质量

分组网不像 SDH 那样的带宽固定的大家井水不犯河水的刚性通道，分组网的多条业务在带宽资源不足时会去“争抢”同一条链路或者隧道的带宽资源，那么对于多条业务我们不能袖手旁观的去让他们通过“自由竞争”决定谁能胜出，必须要加以干预，因为业务和业务的重要性不同，在分组网中传送的不同业务对于时延、安全的要求不同，你那发个微博不差这几秒钟，我这打电话呢，几秒钟传不过去就掉线了，所以需要对不同业务加以区别对待。

当网络过载或拥塞时，需要确保重要业务量不受延迟或丢弃，换句话说，需要通过某种方法让网络知道业务的级别，级别越高越重要，分出个高低贵贱。比如语音业务、大客户业务就是必须保障的，而宽带上网的业务就是级别比较低的。

那么怎么让网络能知道这个级别呢？给每个业务编个号就行了，通过 MPLS 标签中的 EXP 字段（3bit）可以区分 8 个服务等级，PE 通过业务映射表，可以知道不同业务的端口、IP 地址等信息和服务等级的对应关系，为其打上相应的 EXP 的标签。业务通过的网络节点只要打开标签查看 EXP，就知道了这条业务是 VIP 还是平民百姓。

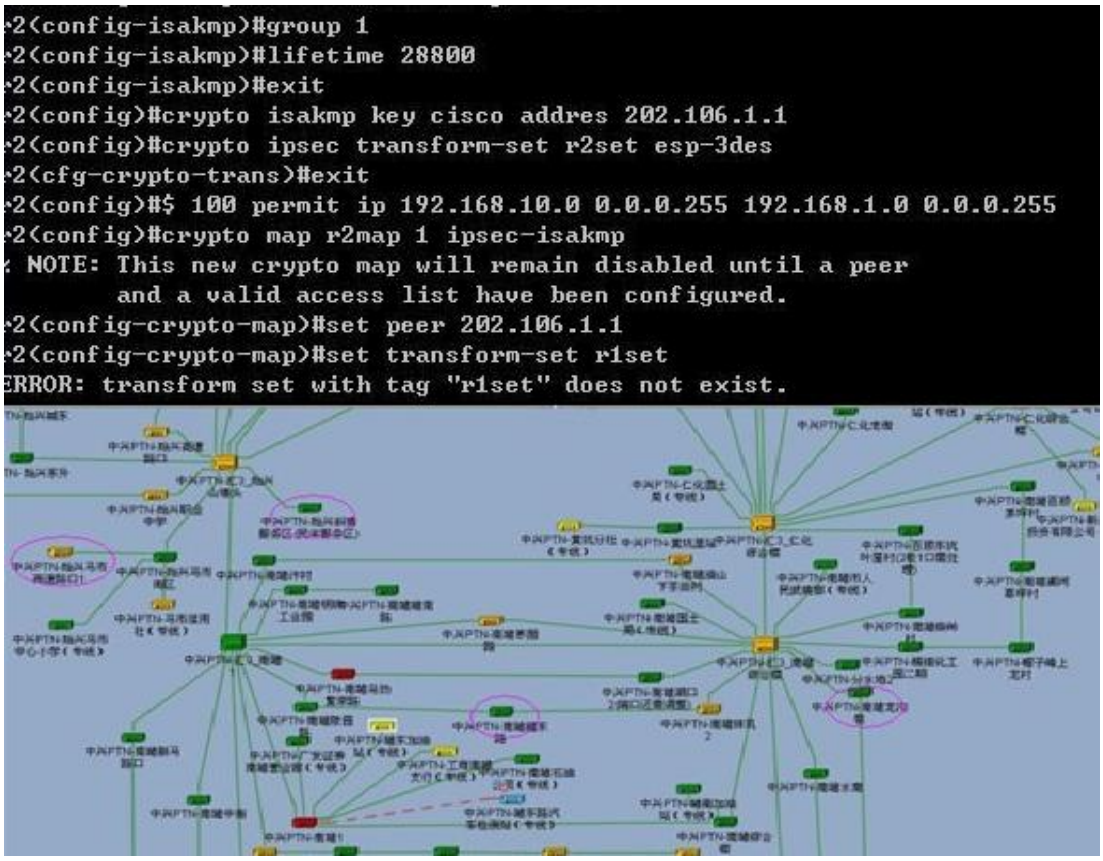
1 **流量工程**

我们的道路在高峰期总会出现拥堵，同时可能有些路段车流量并不大，就需要交通台路况通报这样的系统让大家去避开拥堵路段。网络也是一样，需要避免带宽使用的不均衡。分组网可以通过 RSVP-TE 等动态协议或者网管的静态分配去合理的建立业务路径，保证网络的高效运行。当然如果所有的路都堵死了，那就不是流量工程能解决的了，得考虑网络新建、扩容和优化。

1 **时钟同步和时间同步传递**

分组传送网通过同步以太技术，实现了物理层的频率同步，在此基础上实现了对 1588v2 时间同步协议的支持。

1 **网管界面图形化**



这张图我们能感受到命令行式的和图形化的网管的差异，恐怕我们搞传输的看见那些命令行都头疼不已。数据网使用这种方式去配置业务、管理设备、诊断故障，有两点原因，一是相关的从业人员确实技术很牛，想干这个必须得有 CCNP、CCIE 等认证证书，技术门槛很高；另一方面，数据网的设备数量都比较少，如果是几千台设备区通过这种方式去配置，光是对设备拓扑图的空间想象力都是不可思议的，谁要是能做到，必须推荐他去最强大脑代表中国战队，能在脑力界混了肯定也就不玩通信了。

所以，必须改，怎么改？网管把所有网元的信息都收集上来之后，剩下的就是个网管界面开发的事了，这个我们不关心，只要和原来 SDH 网管一模一样用的顺手就 OK。

分组传送网的功能平面

每一条业务在分组网中的承载都经过了很多的处理，我们按照大的方向将所有的功能分为三个平面去说，再去回顾一下前面讲的内容，去看一下分组网设备的工作流程。

转发平面：

每一个数据包在分组网中经过的每一个节点处都要被转发,转发平面干的活实际上是个完全不烧脑的纯体力活。数据包来了发到哪呢？咱不是有 MPLS 表么，看看标签，在表里查找对应的下一跳标签和出口，啪一下贴上标签，走你！转发平面就介绍完了。

网管平面:

网管平面是人和网络交互的界面，通过图形化的操作界面，实现整个网络的故障、性能、安全、业务配置的监控和管理。

网络哪里断了，在网管界面会有告警显示；网络的丢包率等性能指标也可以在网管去查看；通过网管可以为业务配置保护方式；通过网管去创建一条条的业务，让 PE 知道每一条业务的类型（VPN）、起点和终点、服务等级。

控制平面:

控制平面负责标签的分发，我们之所以能够实现流量工程、在网络中为每一条业务留出一条路（Tunnel）来，都是基于控制平面实现的，控制平面是一个决策者，通过路由协议收集网络拓扑和资源之后，根据这些情报去决策业务的转发路径，而转发层面只是控制平面做出决策后的执行者。

一条业务的历程:

A 站点是一个 GSM 站点，它要去向 BSC 传递本基站语音业务。首先，我们在管理平面告诉和 A 站点连接的 PE1，你的“1 号 E1 接口”连接的 A 站点是 2G 基站，VPN 号是 1，它要去找 PE2 下面的 BSC，对应 PE2 的是“1 号 E1 接口”。

控制平面根据网管平面的这一指令，根据当前网络的带宽资源，为其找到了一条带宽足够的“路”，沿途的每个 P 站点都分配了外层标签，标签中包含了表示这条 E1 业务的最高优先级的 EXP。

PE1 根据网管的指令，为其打上号码为“1”的内层标签，根据控制平面的路径分配指令，为其打上了隧道标签之后，从某个接口将数据转发出去。

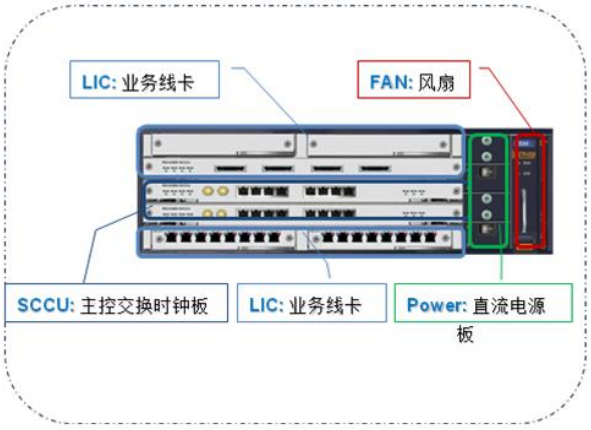
数据在路上只经过了转发平面的处理，一路到达了 PE2。PE2 看到自己 MPLS 表中此业务没有下一跳标签，说明到达了终点站，将隧道标签弹出，露出了“1”这个内层标签，它知道这是个 2G 基站，管理平面早就给他打过招呼，这个业务发送给“1 号 E1 接口”这个 BSC，整个通信就完成了。

分组网设备介绍

无论 SDH 还是分组，从原理讲到设备的时候，技术含量都从大学一下子降到了小学。分组网设备从外观上来看，和 SDH、交换机、路由器等没有什么大的区别，和 SDH 明显的区别是以太网口的集成度明显变高了。

我们去学习设备的方法和 SDH 也是一样,关心一些主要的参数,主要是槽位数、单板的类型，在配置的时候可以知道这台设备能够最大支持的接口数，哪些槽位可以插哪些类型的单板。和 SDH 比分组网有两个特别的参数，“交换容量”就和 SDH 的高阶交叉容量差不多，是对于数据带宽的最大处理能力,也都是用多少个 G 去表示;而“包转发率”是分组网特有的概念，是指每秒能够转发包的数量，单位是 Mpps，这个数量我们不好去精确的计算是否满足，总之是层次越高的设备包转发率也越大，而边缘层设备我们也不会让它“太忙”，所以一定也是够用的。SDH 为什么没有包转发率，SDH 没有包的概念，每个接口都是 8000 帧/秒，这个处理能力是一定满足的。

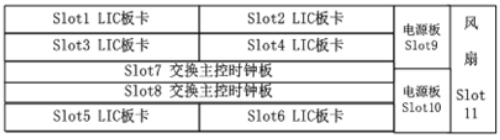
这里就以中兴 CTN 6220 为例简单罗列一些指标。



尺寸	482.6*130.5*240 mm（宽×高×深）
总槽位数	8
业务卡槽位	6
背板容量	88Gbps
交换容量	88Gbps
线速接入容量	44Gbps（单向）
包转发率	65.47Mpps
典型功耗	150W
产品定位	接入/汇聚层设备
FIB表	32K

分组网常用的接口 10GE、GE、FE、STM-1、E1，其中线路侧接口可以选择 10GE 和 GE，需要配置哪些单板就看我们的业务需求，需要几个配几个。比如一个站点要组 1 个 10GE 环，下面带 2 个 GE 的链，要接 5 个 LTE 基站需要 5 个 GE 光口，1 个 3G 基站需求 1 个 FE 口，共计需要 2 口 10GE、7 口 GE、1 口 FE，我们就配置 2 块单口 10GE、1 块 8 口 GE、1 块 4 口 FE，一共占用 4 个槽位空余 2 个，设备配置工作就是个数字统计的过程。

接口类型	板卡描述	单板端口密度	整机集成度
10GE	62XX 10GE（增强型）光接口板	1	2
GE	62XX 8GE（增强型）光接口板/ 62XX 8GE（增强型）电接口板/ 62XX 4GE（增强型）光电混合接口板	4/8	24
FE	62XX 8FE（增强型）光接口板/ 62XX 8FE（增强型）电接口板	8	48
STM-4	62XX STM-4通道化单板/ 62XX STM-4网关单板/ 62XX STM-4 POS板	1	6
STM-1	62XX STM-1通道化单板/ 62XX STM-1网关单板/ 62XX STM-1 POS板	4	24
E1	62XX 16路E1接口板/ 62XX 63路E1接口板	16/63	111



✓6220接口种类丰富，提供各种宽窄带业务接口。

✓6220所有端口都可达到线速

✓6220提供2路10GE接口，多达48个线速FE接口。

这里有一点和 SDH 不一样的，分组网的光模块是灵活按需配置的，比如 8 端口的 GE 板，光模块我们可以配置 4 个，不够用了后期扩容，而 SDH 都是一次性配齐的。

3.10 PTN 和 IPRAN

提到分组传送网，曾经被大家议论最多的两个词就是 PTN 和 IPRAN，虽然两个技术标准都在大规模的使用，但是技术之争已经逐渐的淡化了。

从字面上去理解，PTN 应该是分组传送网的统称，也就是包含现在的 IPRAN 在内的所有解决方案的集合，但实际说到 PTN 是指基于 MPLS-TP 实现的分组传送网，这是由于 PTN 的概念推出时，就一直在 T-MPLS 和 MPLS-TP 的方向上不断演进，而基于 IP/MPLS 的方案直接被思科命名为 IPRAN，PTN 的一个子集另立门户，所以两者就分别作为两种标准的代名词被沿用下来了。

广义的 PTN 和狭义的 PTN 就是一个范围大小的区别，而 IPRAN 则是被人为扭曲了的一个定义。IPRAN 本意是指无线接入网（Radio Access Network）的 IP 化，即 Node B 至 RNC 回传的 IP 化，是无线侧的概念，靠谱一点的理解应该是 IPRAN 是 PTN 的业务驱动，而 PTN 是无线 IPRAN 的承载手段。思科能够这样去包装一个技术，使 IPRAN 成为其解决方案的代名词，也是基于它的江湖地位的，这让我想起了蒙牛的“随变”，不但代表了蒙牛的一款雪糕，很多时候有人去买雪糕，大家表示“不知道吃什么，随便买吧”，于是就买回来一兜“随变”回来，“随变”就成了“随便”的代名词，也是营销的经典案例，和 IPRAN 有着异曲同工之妙。

本文中的 PTN 是指基于 MPLS-TP 的分组网，而 IPRAN 是指基于 IP/MPLS 的分组网。

先说说 PTN（MPLS-TP），我们经常见到一个表达 PTN 技术特点的一个公式，MPLS-TP = MPLS - L3 复杂性 + OAM 。

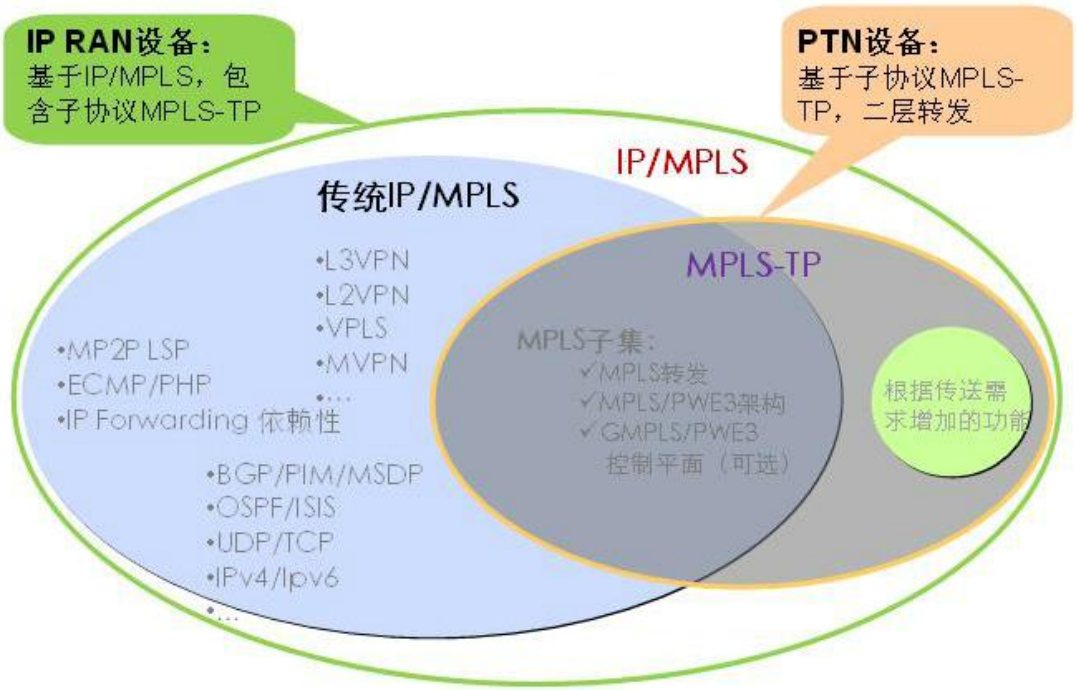
MPLS 本身具备了基于标签的转发和基于 IP 的转发两者的功能，MPLS-TP 是为传送网量身定做的标准，是需要面向连接的，所以 PTN 去掉了 MPLS 无连接的基于 IP 的转发，增加了 SDH 网络原本具有的端到端的 OAM 功能。 MPLS-TP 的电路连接的搭建采用 PWE3 的方式，而业务的保护和管理、维护等功能都参照 MSTP 的方式，可以理解为除了内核由刚性变为弹性之外，与 MSTP 的其他方面非常类似。

再说简单一点，PTN 就是按照原本 SDH 的思路，将 MPLS 的 L2VPN 概念实行“拿来主义”，保留了 MPLS 面向连接和 IP 的统计复用，其余的功能都尽量原版 COPY 原 SDH 的技术。

和最初的 PTN 相比， IPRAN 支持全面的三层转发及路由功能，支持 L3 VPN 功能和三层组播功能，并同 PTN 一样对网管界面做了图形化的改进，对业务实现端到端的精细化管理。所以有些人理解 PTN 和 IPRAN 的区别在于 PTN 是二层 VPN，而 IPRAN 是支持三层的，实际上由于 LTE 业务需求的推动，PTN 设备也通过升级支持了三层 VPN 的功能，所以二层还是三层并不是两者的本质区别。

实际上，PTN 和 IPRAN 两者的最大区别在于控制平面的实现不同，PTN 的控制平面是通过网管实现，相当于有一个站在所有设备之上的管理者，去根据全网的路由、带宽信息去统筹分配路径、带宽、分发标签，实际上 PTN 就不存在控制平面，因为控制平面和管理平面合二为一了，PTN 管理平面集成了 IPRAN 管理和控制两个平面的功能，之所以将 PTN 控制平面也单独出来说，也是为了突出两者的区别。

而 IPRAN 的控制平面是在设备上实现的，设备之间通过各种路由协议、标签分发协议，相互沟通、商量着实现了路径选择、资源预留等功能，从下图中也能看出，IPRAN 包含的协议要比 PTN 多很多，IPRAN 的设备承担了控制平面这一重大功能。



很多时候，一个貌似很复杂的协议，其实实现的是一个很简单的功能，原因就在于，机器不是人，对于机器来讲一个 bit 没定义清楚它也听不懂，可能就导致一片混乱，所以就必须非常严格、规范的制定各种规章制度，确保机器能够准确无误的去完成人类想做的事情，而人不一样，很多事情可以反复的沟通，直至问题明确的被解决。我们都学过编程，编程和协议就是一个道理，你一个标点符号、命令格式用的不对，程序就无法正确的执行。

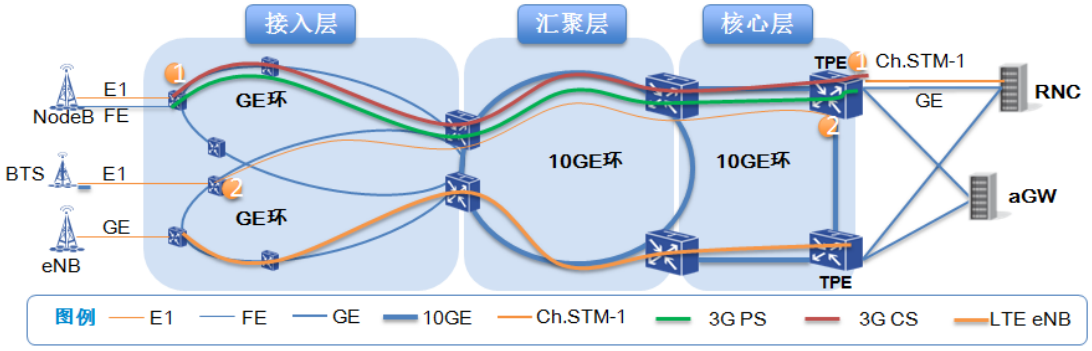
所以概括一下两者的区别，PTN 是中央集权，IPRAN 是民主共和。PTN 是网管位高权重、独揽大权，一人全部说了算，所有的设备上面不需要复杂的协议，只要能够听懂、接收命令就行了；而 IPRAN 由设备去计算路由，对于设备来讲相互交流拓扑结构、带宽资源、标签分配、VPN 路由等每一个信息都需要复杂的协议作为约定，实际上实现的功能和 PTN 是相同的。

在 2G 和 3G 时代，传送网提供的只是基站至中心局的管道，在 LTE 阶段，X2 接口以及 MME/SAE Pool 承载需求，基站之间需要路由型网络，随着技术的发展，PTN 也逐渐支持了三层 VPN 的功能，在各大运营商的承载技术尘埃落定之后，技术之间的竞争也渐渐偃旗息鼓，最终两种技术也将随着通信的大趋势逐步走向融合。

不管黑猫白猫，抓到耗子就是好猫，对于分组传送网的技术选择来说也是这样，不管是 PTN 还是 IPRAN，只要能够满足分组传送网的要求就是可用的技术。分组传送网的要求无非上一节说到几点：支持多种端到端业务（面向连接），高带宽利用率（统计复用），类 SDH 的保护和管理、支持 QoS、流量工程、时间同步等功能，两者在功能上没有多大差别。

3.11 分组网组网结构

分组传送网组网与 SDH的组网，单从组网图上去看区别不大，都是环状或环带链的结构，都是分为核心+汇聚+接入层三层，如果本地网规模较小，核心汇聚层节点数量不多时，不区分核心层和汇聚层，统称核心汇聚层。接入层、汇聚层一般采用环形结构，核心层可以组环状或网状结构。分组网每一层的功能层次定位也与 SDH 网络相同。



相同的地方就不介绍了，下面就说说分组网和 SDH 网络的主要区别：

1、双上联结构

分组网接入上联汇聚、汇聚上联核心尽量采用双上联的方式，也就是一个接入环上联至两个汇聚节点，汇聚环上联至两个核心节点。这是由分组网的保护机制决定的，分组网可以通过 VRRP 保护（下一节介绍）在两个上联点（网关）之间自动的切换，可以保证一个上联点失效之后，另一个上联点承担起业务收敛的重任，确保业务不中断。

而 SDH 网中双上联的应用较少，多数是采用单点挂环的方式，因为 SDH 需要做到两个上联节点之间的业务倒换，需要人工进行逐条业务的配置，实现难度较大，不易于全网部署。

2、多组环、组小环

组小环就是环上节点数量严格限制，比如不超过 8 个。其实 SDH 时代我们也同样倡导组小环，这个原则并没有改变过，可是实际上 SDH 环上节点超过 10 个的情况“大有环在”，我们总会碰见这样那样的困难和理由，而无法重视这个原则。

其实这里面就是一个很简单的道理，一个环的速率一定的情况下，环上节点越多，单站的可用带宽就越小。在分组网时代，我们面临的业务带宽需求已经今非昔比了，动不动就几十 M、上百 M 的需求，我们接入环是以 GE 环为主，按照 LTE 单站 100M 的带宽去计算，考虑 1:2 统计复用的效应和带宽的预留，一个环也就接入 8 个左右基站，对于 10GE 环也有比如环+链节点不超过 20 个的要求，运营商不同要求也不同，具体的原则我们不讨论，就说这个事。

所以组小环就是必须实现的目标。那些实在没办法的山沟沟里，新建光缆、拆环都无法实现的区域，一般也不会有太大的带宽需求，再怎么样最起码也应该是“尽力而为”吧。

另外一方面，LTE 要求端到端的时延小于 20ms，这就要求从中心局 EPC 到基站有一个最大跳数，一般不超过 30 跳，这也是我们限制环上节点和链上节点的另一个原因。

链上节点尽量不超过 3 个，链越长越不安全，这是和 SDH 一样的要求。

3、容量计算

SDH 容量的计算是个简单的统计数字的过程，比如一个 2.5G 环，就 1008 个 E1，各家各户用多少报上来，一统计不超过 70%就 OK，超过了就要扩容，这是个“硬”道理。

SDH 的带宽需求是“数字”的、“离散”的，而谁能说出分组网一个站的带宽占用是多少 M？分组网的容量是“软”容量，具体使用多少是要去实时监测的，不像 SDH 那样满了就一个站也接不进来，从带宽上说分组网理论上没有严格的最大接入站点数量，上面说的只是一个建议数量，就像我们的马路，到底最大能跑多少辆汽车谁也无法准确估计，只能在拥堵的时候对道路进行改造。

我们在做网络规划、设计的时候，按照运营商的指导意见，每个基站有一个“承诺带宽 CR”和”峰值带宽 PR”的值，峰值带宽我们一般不做考虑，按照“承诺带宽”也就是“必须保证”的带宽这个标准也可以像 SDH 一样去计算，按照计算结果去选择速率等级、规划网络，但这个计算结果是有多一厢情愿，小雨哥又忍不住说几句。

某次我们公司要做一个“分组网带宽统计及网络规划”的课题的时候，我们在各厂家的支持下，在对本地网各环路接口的带宽检测至少一周以上的情况下，拿到的数据很是意想不到，如果说平均带宽利用率不足 10%会不会觉得夸张？所以对于李克强总理三次督促“降费提速”,我相信总理此之前是做了实际调研的。

当然设计原则也没有错，这就像我们再盖一个住宅小区的时候，按照住户 1:1 的比例配置了停车位，这个算法是没有错的，可是住户究竟有多少能买得起汽车，我们计算过没有？

我们来算一笔很粗很粗的帐,按照相关数据统计,全国 3G\4G 移动用户 2014 年底达到 6 亿，其中中国移动 3.36 亿，中国联通 1.49 亿，中国电信 1.18 亿，每个用户的月均上网流量按照 300M 计算，具体的数据没有找到权威性的，反正在“降费提速”实行之前这个 300M 带宽应该是达不到的。这样一个月的总流量=总用户数*月均流量，这个单位是 MB，换算成 bit 再乘 8 得到 Mb，我们把这个流量分摊到每一秒，总流量÷30 天÷24 小时÷3600 秒之后得到的是每秒的总流量（Mb），再分摊到全国 332 个地级市本地网，得出的每秒的数据速率三家运营商在 300M-1Gbit/s。而再来看看我们的本地网一般的本地网的分组网汇聚层也至少有几个 10G 环，这里我们不考虑运营商的用户占比、本地网之间的差异，这个粗略的数字已经可以给我们一个答案了。

回归正题，分组网的系统速率的选择与 SDH 有很大不同，SDH 每个速率等级是 4 倍的关系，提供的速率等级也较分组网多（STM-1、4、16、64 四种），可以更灵活的选择。而分组网只有 GE 和 10GE 两种速率，如果 GE 不满足就只能选择高 10 倍的 10GE 给人的感觉有些浪费，所以在分组网建设中，接入层一般建议采用 GE 速率。

光口拉远接入基站

光口拉远，这不是分组网和 SDH 的区别，而是语音接口 IP 化的成果。在语音接口使用 E1 的时代，E1 的传输距离就是 100 米，基本没有哪两个基站之间距离是小于 100 米的，所以我们必须每个基站配置一台设备。

在语音接口 IP 化之后，语音信号通过以太网接口去传送，这意味着只要光口能传送到的区域，一台设备就可以接入方圆 XX 公里范围内的基站，对于光缆我们必然还要铺设到位，影响不大，但是设备的数量可以缩减到原来的几分之一。

当然,用光口拉远基站节省设备带来的问题就是组网结构变为星形，理论上的网络安全性的下降是不争的事实，这里就是一个取舍的问题，关于保护的度，参见 DWDM 保护中的一些言论，还是那句话，我们就说道理，不得出结论，多一份选择，毕竟不是坏事。

IPRAN 的组网限制

这个问题仅针对于 IPRAN，因为 IPRAN 设备上加载多种协议实现路径选择、标签分配，所以我们在建设网络的时候会有这样那样的不允许、不建议的组网场景，而 PTN 由于是网管实现控制功能，设备层只要路是畅通的就可以指配隧道，因此不存在此问题。

IPRAN 在运行 OSPF、ISIS 等协议时，为了能使路由计算更快同时减少设备的压力，会将每个汇聚环、每个接入环/链分成不同层次、区域或者进程，将一张分组网分为一个个小部分去独立的运算,这样就需要一个接入环的两个上联汇聚点在一个汇聚环上; 而且为了简化业务配置尽量是相邻节点，因为如果两个上联点不相邻，中间点要做一个 VLAN 的交换建立二层连接。即便是物理上跨域去连接其他环汇聚的汇聚点，和本接入环不同区域的那条连接也是无效的，会被路由域隔离，也达不到我们想要的多一条路由多一份保护的效果。

其他还有一些规则，比如一个接入环的两个上联点必须都是汇聚节点，汇聚环上联点是两个核心节点，这样的层次分明、等级森严，实际上所有的“限制”基本就是几个目的，为了业务的配置方便快捷，为了业务的管理更加清晰可控，为了给设备的运行计算减少压力。

3.12 分组网业务承载

前面已经讲过 PWE3（L2VPN）和 L3VPN 的区别，这里再简单的概括总结一下，对于一条 PW，两端的业务侧的设备就相当于直连，分组网络就是一个传送门；而 L3VPN，业务侧设备将数据发到了网关，业务侧“委托”分组网根据 IP 地质去查找业务的目的地，去计算、决定业务转发的路径。

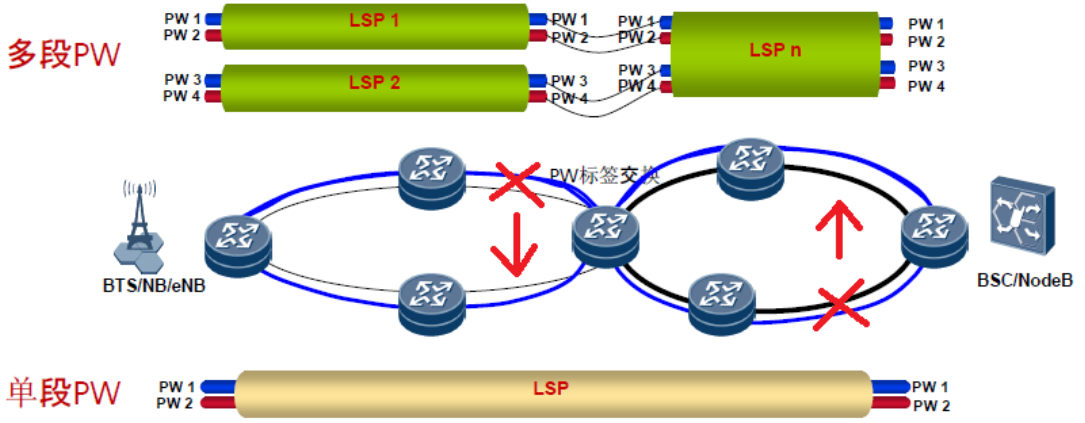
我们几乎每次接触分组网就会听到三层到核心、三层到汇聚、三层到接入这样的字眼，这些二层、三层的词语就像蜜蜂一样整天在我们耳朵跟前萦绕，其实说的就是对于每一种业务到底要用什么方式去承载更合理，这里，我们就分 2G/3G 和 LTE 两种情况去看一看不同承载方式有什么区别。

2G/3G

2G/3G 时代，业务侧的组网模型是星形结构，也就是所有基站对中心局的点对点连接，SDH 就是用一条条的 E1 承载了这些基站到局端的电路，到了分组网我们自然而然的就用 PW 去解决，因为业务的两端都是确定的，用动态的 L3VPN 去解决除了把问题复杂化了之外，结果和 PW 都一样，而 PW 几乎就是专门用来解决这种点到点业务的。那么如何用 PW 去解决，我们可以选择 2 种承载方式：端到端 PW 和分段 PW（MS-PW）。

端到端 PW，就是每个接入点配置一条 PW 直接到达核心层（下图左），而分段 PW 就是接入层各点通过 PW 到汇聚点，汇聚点在将接入层的 PW 打包用一条 PW 回传至中心局（下图右）。

端到端的 PW 有两个缺点，一是核心层设备的压力巨大，比如全网有 1000 个基站，核心层就要处理 1000 条 PW 的隧道信息。二是对故障的保护能力较弱，比如我们基站到核心配置端到端的 2 条 PW，分别为 PW1 和 PW2，如果 PW1 的接入段和 PW2 的汇聚段同时发生了中断，系统就会认为这 2 条 PW 全部中断，而导致整个业务中断。而分层 PW，会在接入段和汇聚端分别实现 Tunnel 的倒换。

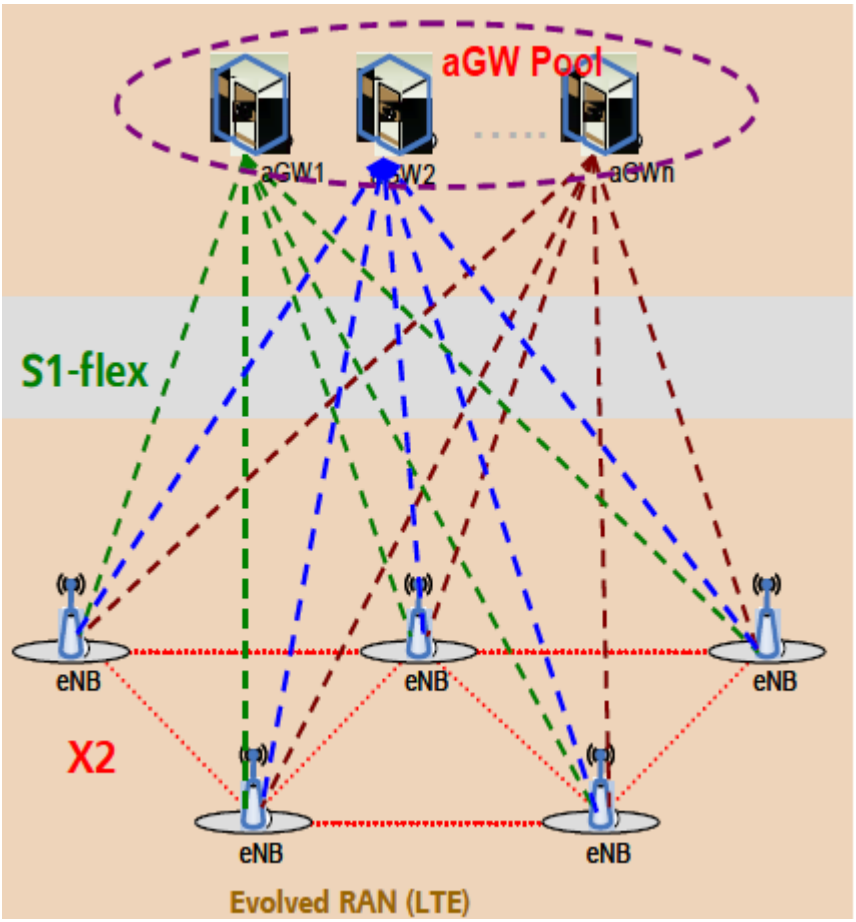


MS-PW 就像我们每一个企业的人员结构一样，从上到下分为若干等级分别有相应的部门领导，而各个部门领导面向总经理做本部门（区域内）的业务接口，这样总经理只需要管理几个部门经理就 OK，如果让总经理去管理每一个员工，他会 very 的忙。

MS-PW 的汇聚节点处，需要对接入层的 PWE3 终结之后，将多条 PW 全部贴上另一个 PW 标签（标签交换），进入汇聚环上的 tunnel 之中，将信息送上汇聚直达核心的“专用高速”。

LTE 业务

为什么把 LTE 单独拿出来说，LTE 和 2G/3G 的最大区别就是多了一个 X2 接口，也就是相邻基站之间的一个通信接口，2G/3G 时代的基站只管和中心局 BSC、RNC 的交互，而 LTE 将部分控制功能下移至基站，基站之间也需要互相打打招呼，所以基站名称也从 NodeB 升级为 ENodeB。对于 S1 接口和 2G 的 Abis 接口、3G 的 Iub 接口一样，采用 MS-PW 的方式承载，这里要说的也就是 X2 接口如何处理。



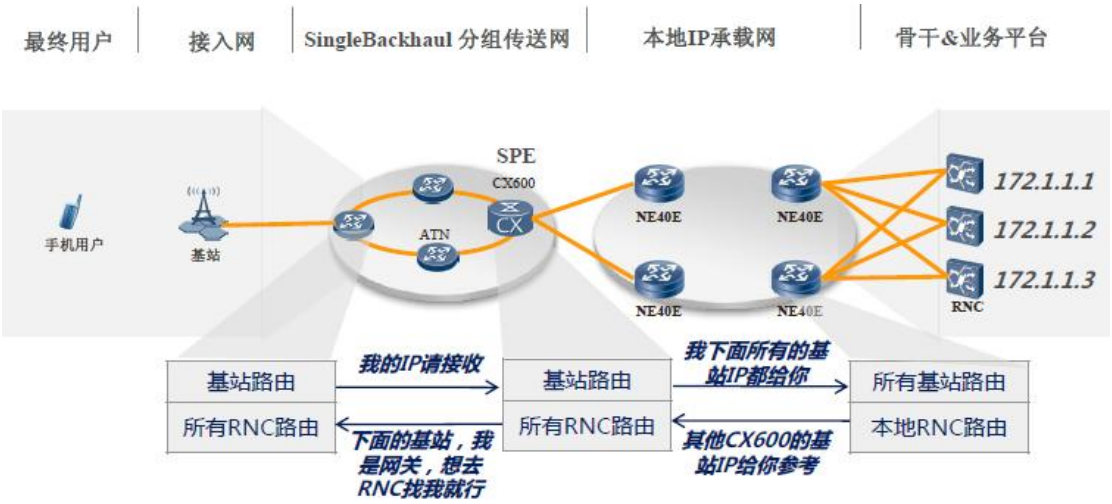
三层到核心，就是整个分组网除了核心设备具有 L3 功能之外，汇聚和接入设备只做 PWE3 的处理；三层到汇聚就是核心到汇聚采用 L3VPN，汇聚到接入采用 PWE3；而三层到接入，是指全网都采用 L3VPN 的方式接入业务。

L3 到核心，任意两个基站之间要通信，他们都靠 PW 连到中心局，局端的核心设备具有路由功能，会帮你找到你要找的那个 ENodeB，这和前面说的端到端 PW 一样的道理，核心设备就成了全网所有基站的“接线员”，而且所有的业务都要绕汇聚环路到达核心层，这就像大家下了班不管家离公司有多远，都上绕城高速走一圈？那绕城就一定堵的一塌糊涂。

L3 到接入，每个 ENodeB 要找另外的 ENodeB，每个 ENodeB 接入的设备就是他的网关，去实现 X2 接口都靠 L3 功能自动寻址，分组网是面向连接的，每一个 X2 接口的通信就会建立一条条的 Tunnel，这样 Tunnel 的数量也不说 N*(N-1)/2 吧，毕竟这个组合中很多基站之间是没有通信可能的，但至少也是基站数量的好几倍，再加上 abis、iub、S1 接口的业务，各种 L2、L3 大客户，接入层的盒式设备能否从容处理这么大的信息量是个问题，到底有没有必要让接入设备去这么“智能”又是个问题。

L3 到汇聚，也就是我们说的分层 VPN（HVPN），接入到汇聚采用 PW 的二层方式，汇聚层使用 L3VPN，这就是我们实际使用最多的一个解决方案。

接入设备对汇聚设备说“我的眼里只有你”“你就是我的唯一”，每个基站 X2 只要建立一条对汇聚点的 PW 就轻松搞定；汇聚设备说，环上几个弟兄，要找我小弟的给我说一声，跨部门协调就是我的本质工作，就这最多几十条路由问题不大；核心设备说，这些汇聚层的兄弟们跨环通信请找我，麻麻再也不用担心我的能力；从此以后大家都表示能担重任，皆大欢喜。



3.13 分组网保护方式

分组网的保护方式分为单板级保护和网络级保护，单板级保护前面 MSTP 和 DWDM 部分都有介绍，下面对网络层保护进行说明:

BFD（双向转发检测）:

SDH 的保护是基于站点基于 SDH 帧结构的检测，SDH 每秒 8000 帧，每一帧都包含了对于通道、复用段的工作情况的信息，因此能够很快发现故障。

对于 IP 网来说，路由器之间也有着类似的机制，叫做慢 HELLO 机制。为什么叫做“慢 hello”呢？因为它比较慢。本身 IP 数据网就是定位于承载公用业务的，基于“秒”级的故障检测、重路由机制是可以接受的，可是现在要承载 TDM 等“电信级业务”了，秒级就真的太慢了，要让它快起来。

BFD 的原理就是网络对等体之间通过互联接口每隔一定时间的互相握手，这个接口可以是物理的，比如是一个以太网口，也可以是逻辑的，比如一条 PW 或者 LSP 的两端，通过三次互相打招呼，确定这个路由是可用的，也就是”UP”状态，一旦“DOWN”了，就说明出现了问题，就需要触发倒换保护。

比如张三和李四是对等体，他们之间的接口目前属于“DOWN”状态，也就是不可用。张三给李四发一个“我这里显示你是断的”，李四回复张三“我这里显示你也是断的”，这是第一回合。张三接因为接收到李四的信息，又对李四说“我们好像通了”，李四也给张三回复“是的，我们好像是通了”。然后张三又给李四说“我确定我们通了”，李四也回复张三“是的，我们确实通了”。

这就是 BFD 的原理，很简单，可是作用非常重要。基于这个简单的协议，使 IP 网、MPLS 网能够承载“电信级业务”。BFD（双向转发检测）是一套用来实现快速检测的国际标准协议，提供轻负荷、持续时间短的检测。

BFD 不是一种保护，但是是所有保护的前提， BFD 能够在系统之间的任何类型通道上进行故障检测，这些通道包括直接的物理链路、虚电路、隧道、MPSL LSP 、多跳路由通道。在 IPRAN 网络中，不管是隧道层面、业务层面、物理链路层面，均可采用 BFD 进行快速的故障检测。BFD 是 IPRAN 中使用的协议，在 PTN 中使用的协议是 Y.1731。

隧道保护:

LSP 保护和 SDH 的通道保护原理大致相同。

无论 PWE3 还是 L3VPN，每一条业务都在网络中经过了一条隧道 (LSP)，我们在配置隧道的时候，像 SDH 那样配置一主一备两条隧道，并指定两者是互为备份的关系，就实现了隧道 1+1 和 1：1 的保护。

1+1 保护就是主备通道同时传送业务，接收端择优接受，无需启用 APS 协议。1：1 保护平时保护通道可以传送额外业务，当工作通道故障时，业务切换到保护通道，如果保护通道带宽不足，保护通道上原来传送的额外业务被丢弃，需要启用 APS 协议。实际应用中一般配置的是 1:1 保护。

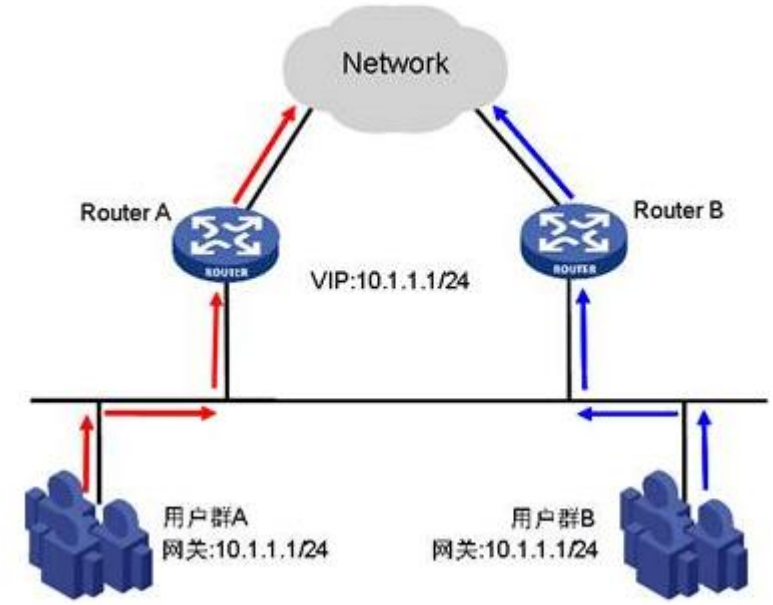
LSP1：1 保护是 IP RAN 网络中基本的保护方式，分为恢复式和非恢复式两种，恢复式是指倒换之后迅速进行原主用 LSP 的再次协议建链，若能建立成功，将在一段时间后，将流量再倒回原主用 LSP 上，这个动作为倒回。

LSP 保护能够对链路和途径节点的故障实现保护倒换，但对于业务的源宿节点的故障，无法实现保护。

网关保护:

分组网络采用 HVPN 的分层结构时，接入层双上联的两个汇聚节点作为 PW 的终结和二三层桥节点，是二层 VPN 网络到三层 VPN 网络的网关，在中心局侧，中心局内的 2 台设备也作为 BSC、RNC 等设备进入三层网络的网关，在网关设备故障或者网关至用户侧设备链路失效的时候，用户侧设备是不支持自动重路由寻找备用网关的，所以网关之间需要启用 VRRP 保护已防止这种情况下导致的业务中断。

VRRP(VirtualRouterRedundancyProtocol,虚拟路由冗余协议)是一种容错协议。 VRRP 将局域网的一组路由器（包括一个 Master 即活动路由器和若干个 Backup 即备份路由器）组织成一个虚拟路由器，称之为一个备份组。这个虚拟的路由器拥有自己的 IP 地址。当缺省路由器 down 掉(即端口关闭)之后，这时，虚拟路由将启用备份路由器，从而实现全网通信。



VRRP 保护在 PTN 中叫做双节点子网保护。

业务保护:

PW 保护

对于 PWE3 业务来说，配置两条一主一备的 PW 可以实现对 PW 的保护，两条 PW 是同源不同宿的，如果 PW 的终点是 HVPN 中的二三层桥节点，还需要配合 VRRP 保护。

VPN FRR 保护

对于 L3VPN 来说， CE（如 BSC、RNC）和 2 台 PE 相连， 当一台 PE 或者 CE 和 PE 之间的链路故障时，CE 可以通过 VRRP 切换至备用 PE 上，同时远端 CE 通过预先在远端 PE 中设置指向主用 PE 和被用 PE 的主备用转发项，启用 VPN FRR 保护自动切换备用路由至备用 PE，防止业务的中断。

PTN 环网保护

环网保护是 PTN 特有的保护机制，IPRAN 不支持此保护方式。环网保护即创建一个环形的保护隧道为工作隧道的保护通道。当网络上节点检测到网络失效时，故障相邻节点通过 APS 协议向相邻节点发出倒换请求，当某个节点检测到失效或接收到倒换请求时，转发至失效节点的业务将倒换至另一个方向。

环网保护分为 Wrapping 和 Steering 两种组网方式，这两种方式组网方式相同，也都需要启动 APS 协议，差别是倒换机制，Wrapping 模式下，业务在故障相邻的两侧节点进行换回，Steering 模式下，业务在源宿节点进行反向。

第四章　传送网设计应用

4.1　理论实践结合

巴拉巴拉写了有几万字，传送网的基础知识就大体介绍完了，呼——（长出一口气），写到这里，对于传送网也只能算介绍了一点皮毛，虽然不够深入，也已然是我的全部，这些也就差不多可以应付基本的工作了。

在工作中,我给大家做培训的时候,总有些人会有一个疑惑:学习这些东东有什么用？的确，相比较概预算培训、机房勘察流程培训等内容，这些纯理论的东西乍一接触感觉不是那么接地气，所以从这一部分起，我们把之前介绍的知识和具体的工作内容逐渐结合起来。

首先我想占用一点篇幅，来探讨一下理论和实践的关系，一个学习方法的问题，因为作为一个新人，能够迅速的提升自己，尽快脱颖而出承担重要的工作岗位，能够为个人的发展、晋升奠定一个重要的基础。提高技术水平无非就是两种渠道，理论学习和工作实践，也就是看技术资料和跟着老鸟实战。

首先说理论学习，现如今是信息高度发达时代，网上的资料从专业广度、技术深度、研究角度来说可谓无所不有、包罗万象，新手们面临这个资料库就像大海里的一只小鱼一样找不到方向，一个很现实的问题就是--看不太懂，这无疑是很影响积极性和上进心的。其实实际上这并没什么，很多资料是没有必要深究的，比如就算同样是传送网的资料，先不说具体哪个知识模块，单纯就作者角度来说也分面对产品开发、技术普及、网络建设、工程设计、运营维护等等太多不同，而我们只需要在资料的海洋里找到适合我们的，剩下相关度不高的虽然写的很好也可以做个记号以后慢慢看，所以对于资料的筛选就是一个关键的问题。

再说工作中的实践，工作中我们接触到的又是另外一个天地，不仅包含了技术、经验和技能被具体化了的一个个的问题，又包含了心理学、营销学等更多的方方面面，职场就像一个被社会化了的大经验池，这又是一个更漫长的学习历程。

如果我们不去实践，单凭看技术资料，经过一段时间之后很遗憾的告诉你，你的进步几乎是零，因为职场去定义你的水平是不管你看了多少本技术书籍，考核你的只有一条，你能干什么工作。可是我们如果不去研究理论而仅仅凭借实践，比如你刚毕业，迅速的掌握了如何绘制图纸、文本表格编辑，熟练掌握各种工作的流程，只能说做到迅速上手，但到了一定的阶段你会发现你再也难以突破自己。

工程实践是基础，有实战的机会一定积极参与，多多益善。在工作实践前后进行针对性的知识补充，就是上学时老师的经典语录: 课前预习，课后复习。

比如师傅要带你去机房看一下，去之前就找一些关于机房内设备、机房工艺、勘察流程等资料了解一下，到了机房就不至于大脑空白。在机房的时候多拍些照片多问些问题，这样师傅会觉得你孺子可教也，你就成了重点培养的对象。回来之后再看些资料，和见到的实物对照

一下，这一次小小的经历也会有不错的收获。如果不做准备，去了就只是看热闹，回来师傅问你看机房感觉怎么样，你总不能说：机房很好！很大！很凉快！
技术理论可以使你从熟练工提升到专家，就像唱歌一样，不懂音乐也能唱，那最多就是歌手，如果精通音乐理论就是音乐人。从实际工作中接触到的一些问题，通过理论学习，将接触到的东西梳理、总结，形成一个逐渐完整的知识结构，而不是一个个凌乱的知识点。
技术要学到什么深度？个人建议，至少要比工作中能够用到的稍微深入一点，要适当的拔高，这样我们就会轻松完成工作，在公司才会有更大的上升空间。这就像两个举重选手，张三最好抓举成绩 100 公斤，李四最好抓举成绩 110 公斤，那么两者抓举 100 公斤的感觉不同，后者一定更游刃有余、淡定从容。当然深度也是适可而止的，作为网络设计人员，了解设备参数、配置原则，了解组网原则是根本，了解设备工作原理是拔高，去深究协议的细节就没必要了，每个人时间是有限的，用在知识面的广度拓展上效果会更好。

4.2 走进机房看一看

本部分内容属于一些机房初级概念介绍，传输老鸟们可以自动跳过。

4.2.1 走进基站

作为一个传输新人，参加工作后的某一天，师傅过来跟你说，走，和我到站上看看去，具体任务可能是去抢修、勘察、巡检，无论是什么目的，这是不错的一次机会，让我们与通信设备终于有了第一次亲密接触。
进入机房，第一印象最深的一定是嗡嗡嗡的声音，伴随着机器烦闷的轰鸣，自己的头也开始渐渐的膨胀开来，一方面看到基站机房内的形形色色的机架、设备、交错的连线，一下子感觉信息量好大，另一方面也看过很多通信技术的资料，但是感觉不到这两者有什么关系，理论和实践像是不相干的两个世界，在大脑里一时间没有找到结合点。
没关系我们一步一步来，首先我们把握住要点，我们是传输人，我们要锁定那个屹立在一排机柜中的 Mr.right：传输机柜。打开及柜门能看到一根根细细的黄线（尾纤）从设备上接出来，还可能有一根根灰色的细线（2M 线）和网线，最关键的是设备的标牌上写着我们认识的一款传输设备型号，OK 就是它了。

接着我们蹲下来凑近去看，每一块板子上都有一个单板型号，我们拿出手机对照着百度一下设备资料，就知道这个站的传输设备的具体配置，比如一端 PTN 1900 设备业务单板配置为：2*10GE 光口、8*GE 光口、8*FE 电口、8*E1 电口。其中光口使用的线缆都是尾纤，E1 接口对应 E1 电缆，FE、GE 电接口对应网线，除了电源线我们能用到的线缆就这么多。尾纤比光缆轻便柔软可以灵活的弯曲，一般作为基站内的连线使用，与无线基站的 7/8 馈线作用一样。

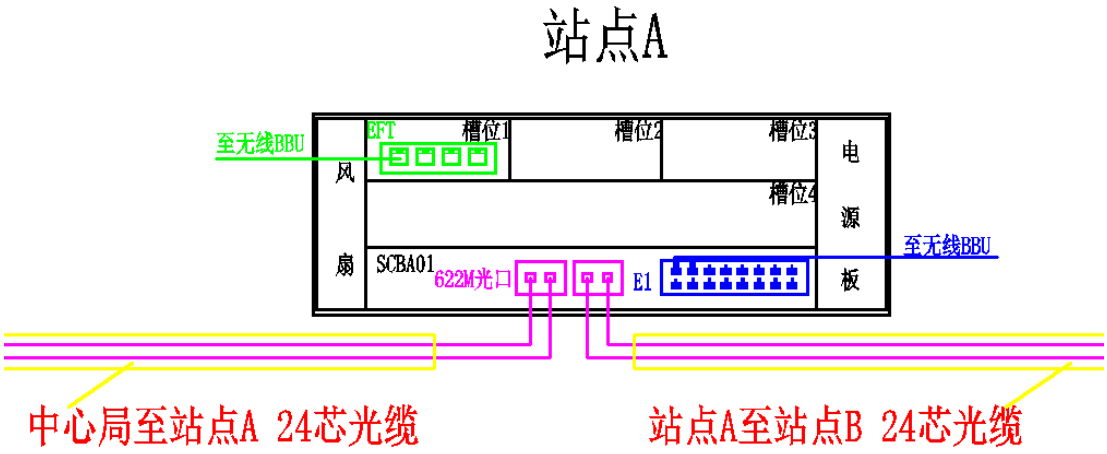
2M线

尾纤

网线

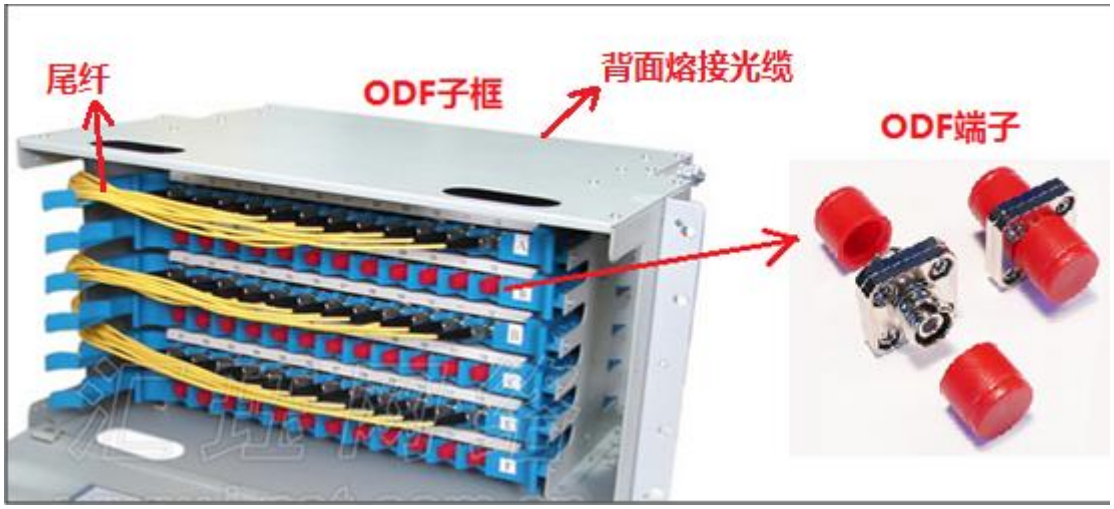
这些光接口中，主线路侧接口一定是速率最高的，比如有 10GE 和 GE，那么线路侧就是 10GE 速率，而低速接口 GE 也有可能是线路侧接口，用来接入下一层的环或者链，也有可能是支路接口。

我们去看 10GE 光口尾纤上面的标签，如果没有标签那谁也无能为力，就当做是有吧。尾纤都是一对一对的使用，两条尾纤对应一个接口的收和发，也就是我们说的一个光方向，上面有着同样去向的标签。这 2 个 10GE 光口上就有 2 对 4 条尾纤代表本站的两个光方向，分别为上下游站点名称，于是我们便知道这个基站在网络上有两个邻居，比如“人民医院”、“市一中”，本站很可能是在一个 10GE 环上，也有可能是 10GE 链上的一个中间节点。对于 GE 光口也同样，标签上写着其他站点名字的，代表线路光接口，如果是写着“LTE”“无线”“3G”这样的标签就代表是本站的业务接口，而没有连线的接口是空余的，可以后期使用。



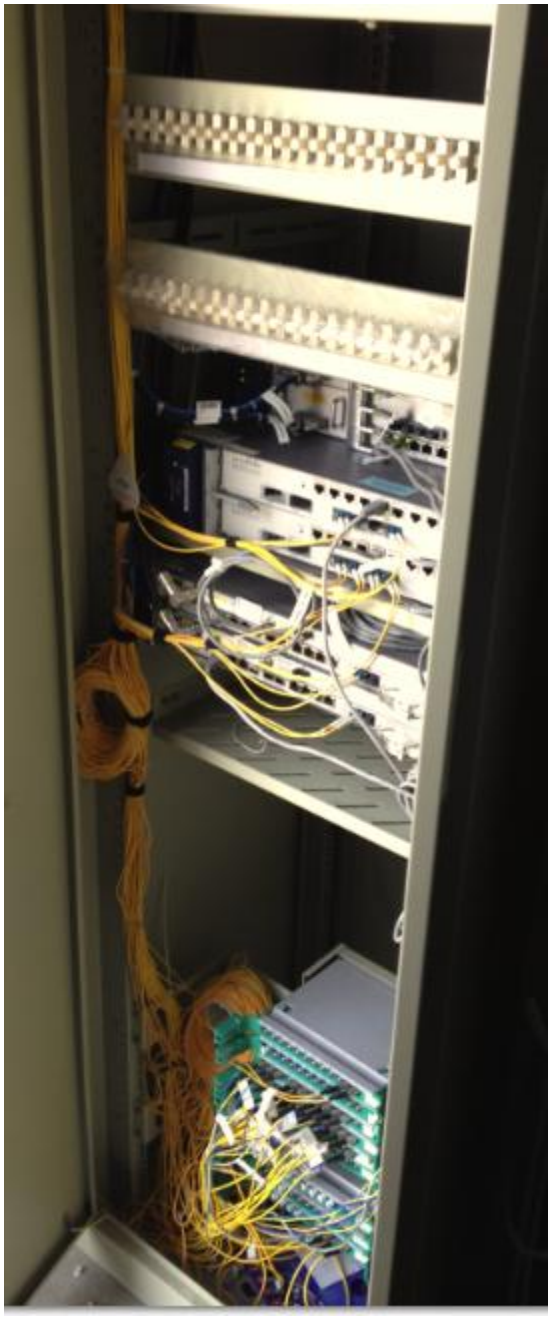
传输设备和光缆之间不会直接对接，中间经过一个 ODF（光纤配线架），ODF 原理和水管对接用的法兰一样，作为传输设备和线路专业的分工界面。ODF 一侧熔接光缆，另外一侧通过尾纤与传输设备的光口连接，在建设基站时进站光缆熔接到 ODF 上，该基站需要使用第几芯，就用尾纤连接到 ODF 对应的端子上，这样规范的安装方式便于后期维护工作。

ODF 每一排端子（熔纤盘）后面也都贴着标签，代表光缆的对端站点，这个标签就可以和设备上线路接口的标签对应上，如果设备上有“人民医院”、“市一中”，ODF 标签上也应该能找到通往对应站点的光缆。



FE 是支路侧接口，一般不经过配线架，直接用网线连接到业务侧设备上。
E1 接口同光口一样需要经过 DDF（数字配线架），作用和 ODF 相同，DDF 一侧连接传输设备，一侧连接业务侧设备，作为传输和其他专业的分工界面，新建基站时传输设备将所有 E1 全部连接到 DDF 的传输侧，业务设备需要使用 E1 就连接到 DDF 对应端子上，每个 DDF 端子上面也有标签，表示本端子是被什么业务占用。

ODF、DDF 在汇聚、核心节点都是独立的 ODF 架，而一般的基站的设备尺寸较小，ODF 和 DDF 配线的容量也较小，所以为了节省机房空间，很多都是传输设备和 ODF、DDF 安装在一个机柜中，叫做综合配线柜。



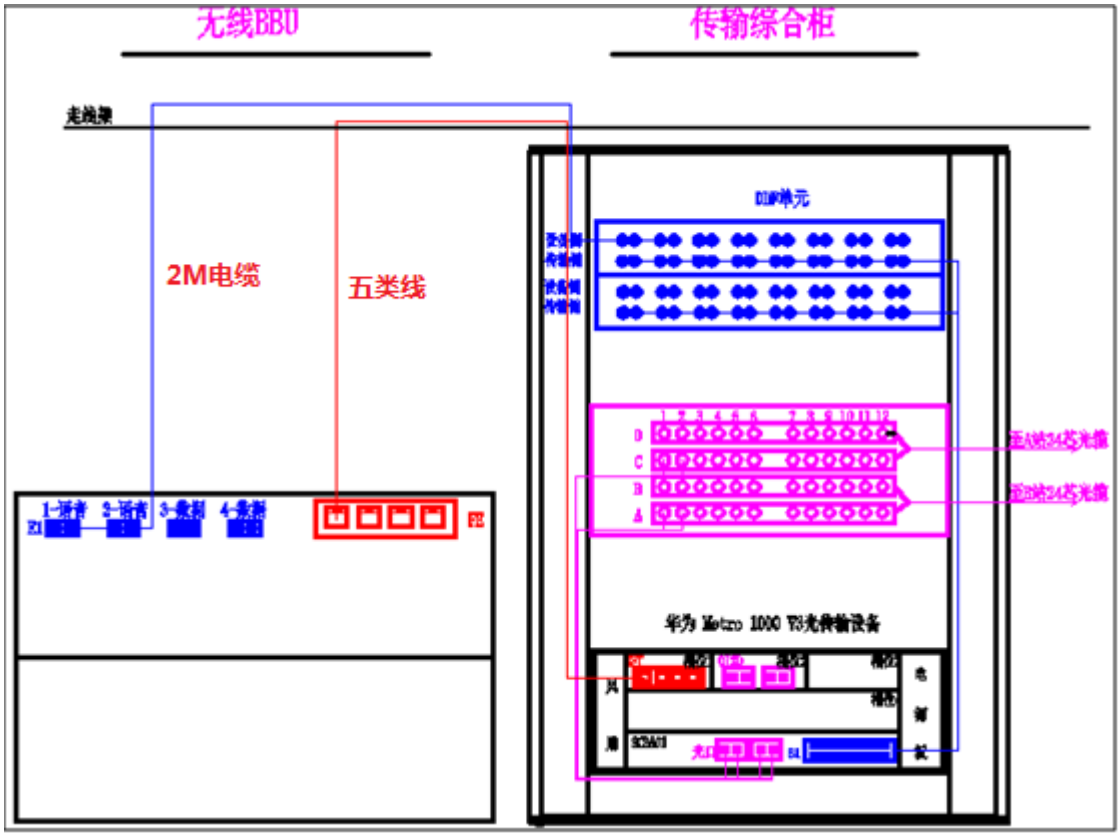
为了机房布线的整洁美观，不同机柜之间的线缆都是通过走线架布放的：

总结一下，通过一个基站的学习，我们了解到了什么内容:

1、认识了一款型号设备的外观和其部分单板，了解了单板型号和对应的接口;

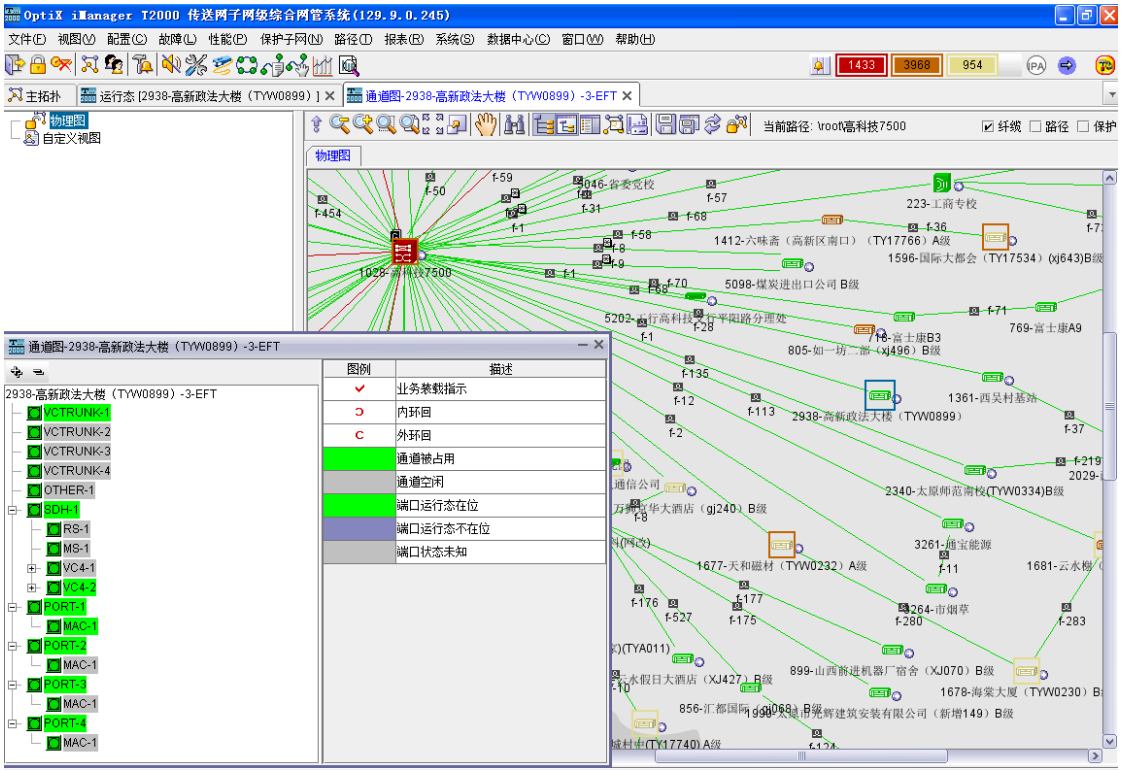
2、通过设备接口标签、ODF 和 DDF 标签了解了基站的线路、支路侧接口的去向，知道了基站的上下游站点和本站的业务侧接口;

3、通过对所见信息的整理，脑子里就有了一个整个基站的连线图，如下图所示:



我们再去到网管上打开我们去过的这个基站的传输设备界面，会看到这些单板的配置、使用情况，和我们在基站中实地考察的结果相同，所以今后如果是要了解这些信息，就不必跑下

去一个一个的看，在网管上就可以一目了然。但是如果要了解机房和机柜内空间使用情况，或者要了解光缆芯数的使用情况、电源端子占用情况等网管上看不到的信息，就需要逐站实地勘察。



看过了一个基站之后，回到办公室，我们发现这些设备资料和网络拓扑图已经不再是纯粹的概念了，再去看时也多了几分亲切，也知道了有些东西有必要记在心里，而不至于“书到用时方恨少”，临场手忙脚乱的百度了。其实这些东西和道理都不复杂，“世上无难事，只怕有心人”，稍微用点心，如果能在一个机房中能够掌握了这些知识，就已经很可以了，因为很多人最初都没有这么用心，至少我没有，干了几个月之后才恍然大悟。

4.2.2 走进中心局

又有一天，师傅过来说，走，我带你去中心局看一看，有了前面的经历，我们于是信心满满、兴高采烈的跟着去见识一番。到了地方套上鞋套踩在高大上的防静电地板上，看见一排排机柜的那一刻，是不是有了一种迷失的感觉？还是脑子又嗡的一声迅速大了好几圈，就像从一个小镇子来到首都一样，心里一直重复一句话“这是什么情况”？

没关系，从基站到中心局其实就是个量变，本质都是一样的，我们耐下心来一点一点的去寻找我们要的线索，而且，基本上作为一个通信人，能够接触见识到的阵容也就不过如此了，世界就那么大一点，我们得去看一看。

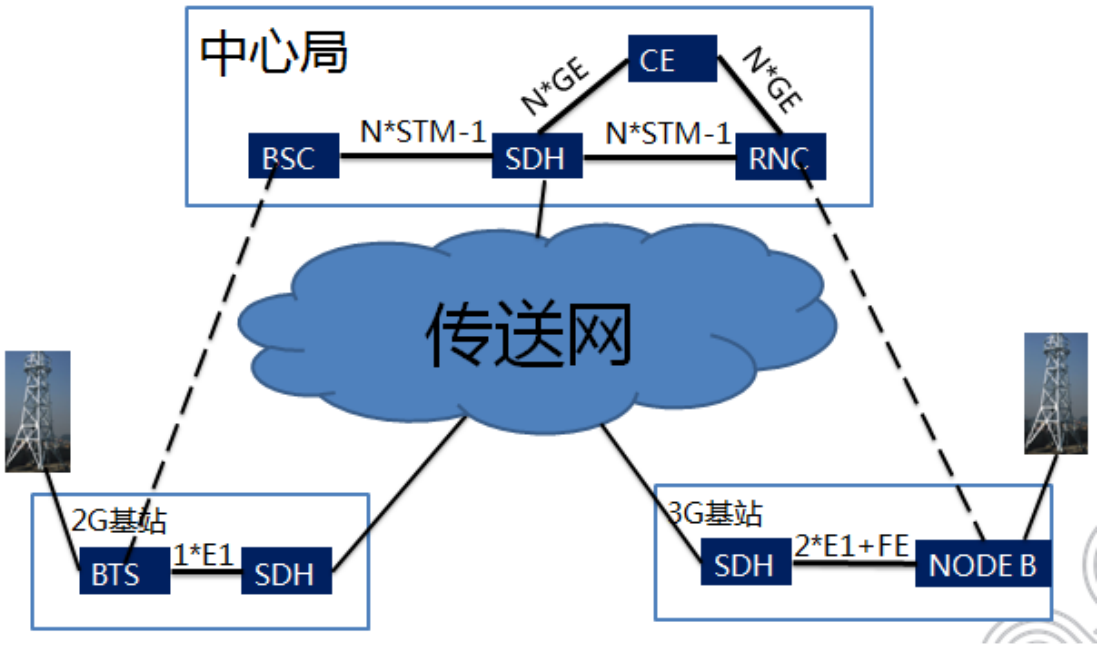
基站的业务侧设备一般就是一两个机柜，而中心局业务侧设备要将全网的所有业务汇集到局端进行交换，所以每个网的阵容足有几排机柜甚至是独立的“无线机房”、“固网机房”、“数据机房”；中心局的 DDF 架和 ODF 架也是一排排一列列，就像***广场阅兵一样整齐庄严；基站的传输设备一般就是一个柜子，而中心局的传输设备要下挂各个本地网 MSTP、分组网、波分环路之外，还有一干二干的环路，所以至少需要一两排柜子。

有了前面基站的经历，我们在去中心局之前应该对组网结构有所了解的，本地网市区环 1 环 2……郊县东南西北环……一干二干各种环，应该将组网图打印出来拿在手里。中心局传输设备位置摆放是分类的，本地网的柜子都是挨着的，干线的柜子可能在另外一头或者另起一排，这样对应着这些柜子上的标签和手里的图纸，耐下心来就能够一一的对号入座。

以前一进入中心局传输机房，中心局的传输机房几乎都是被一排排的 DDF 占据了大半面积，因为中心局无线语音业务落地使用 E1 接口，每个 E1 都是一个端子，所以柜子的数量巨大，现在逐步在进行光接口改造，将 E1 接口改为 155M 接口，或者 IP 化改造为 GE 接口，可以将很多个 DDF 系统缩减为 1 个 ODF 端子对接，逐渐 DDF 会用的越来越少直到消失。

对于 ODF，中心局除了进局光缆数量、芯数巨大，光缆成端的 ODF 架数量剧增之外，还多了一类的 ODF，叫做调度 ODF 架，也就是不同专业设备之间跳纤经过的中转站。普通基站的传输和业务侧也就那么几根尾纤，直接从走线架就“飞”过去了，可是现在面临的是几十上百条的尾纤，这样飞来飞去的可不好，需要在传输侧和业务侧分别使用调度 ODF，去将这些连线规范一下。

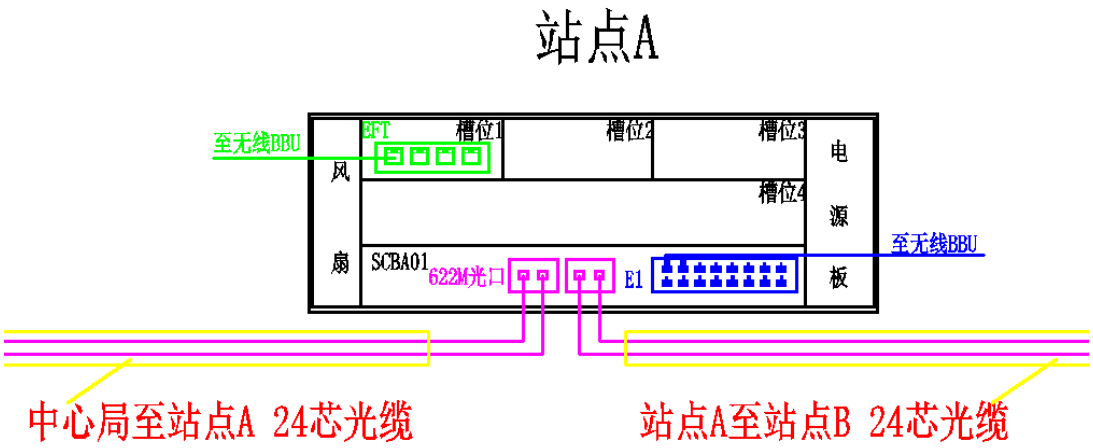
对于 2G 业务，分组网和 MSTP 都通过 E1 接口或者 STM-1 接口和 BSC 互联；对于 3G 业务，语音业务使用 STM-1 接口，IP 化改造之后使用 GE 接口，对于数据业务传输设备和无线局端设备 RNC 之间以太网业务通过 CE 相连，使用 GE 或 10GE 接口。CE 是一台路由器，将传送网发来的业务进行带宽汇聚、端口收敛之后发送给 RNC，分组网有的也通过 CE 转接，有的则是核心设备充当了 CE 的角色。下图中的实线是物理连接，虚线是逻辑连接。



接下来说波分设备，波分设备最显眼的就是合分波板了，尾纤密集不说，而且不是收发成对的，所有波道的收和发都分别连接到合波板和分波版上。波分细说起来突然又没有什么可说的，因为波分也没什么特别的，无非是波分的线路支路速率很高，波分的支路侧是我们 MSTP、分组网的线路侧，也就仅此而已，速率再高也就是一对尾纤，外观上看都没什么区别，其实还是看标签。老说标签标签，因为所有的这些连接关系，就要实地去看就要看标签，没有标

签难道让我们爬到走线架上、地板底下去顺藤摸瓜？可是实际中标签有和没有、规范和不规范都很常见，这是一个安装工艺的问题。

42-3.png (19.46 KB, 下载次数: 10)



4.3 接入工程，小试牛刀

日子刚刚平静下来，还没来得及去整理这些信息、资料，师傅又来了，甩给你一个站表，说“新批下来几百个基站，你给做一下立项报告”，你心里又说“what？什么情况？”。别急，我们来解读一下这个任务，这么多基站要接入进我们现有的网络里来，我们要买多少设备，

建多少光缆，扩容多少个光板以及相关的配套？这些就是建设规模，我们最终要的结果是需要花多少钱，就是工程造价或者总投资。我们做的这个事情就是一个基站传输接入工程，是我们传送网最基本的工程，满足新建基站的接入需求。传输接入类工程还有室分接入、WLAN接入等工程，模式都差不多。

做这个事之前，首先我们要收集一些基础资料，我们传送网现网的拓扑图，设备的配置情况，这些可以在网管上导出表格或者逐站去查看，还需要每个站点的具体位置（经纬度），还需要知道这些新建基站的经纬度、站型配置（不同站型对于不同带宽需求）。我们就拿一个新建基站举例子，将新建站和原有站的经纬度分别用不同颜色的图标导入到地图里。

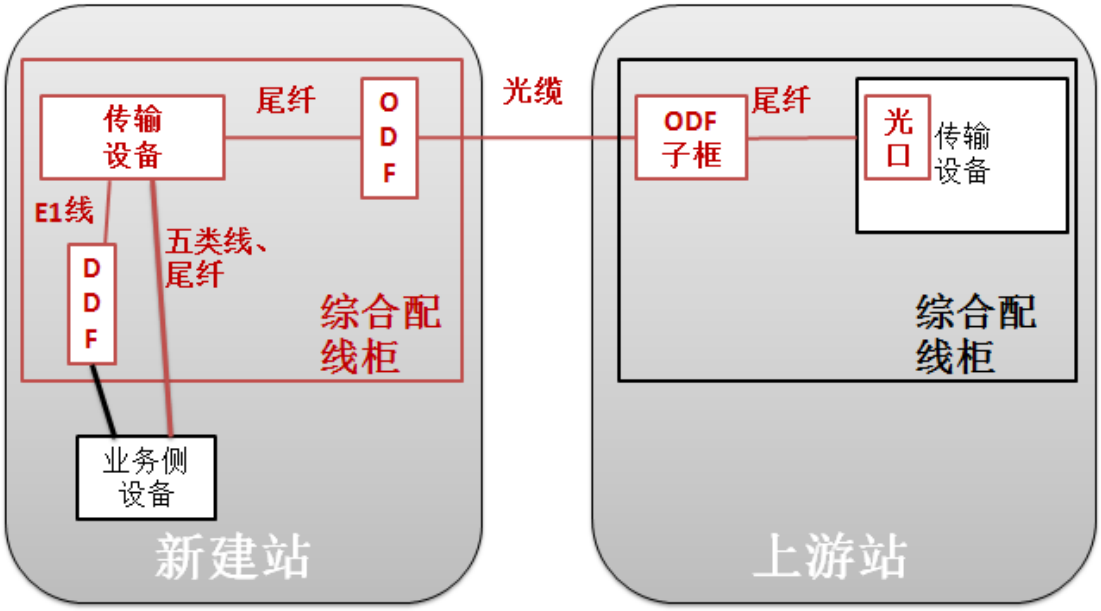
导进地图之后，我们发现这个新建基站周围这么多老大哥，小弟和大家打个招呼：”HI 大家好我是新来的“西大街”，很高兴即将加入你们这个大家族”。周边的邻居也非常友好的欢迎这个即将加入的新成员，不过首先一个问题，大家都是混“光纤通信界”的，要加入进来好歹你得有一条光缆和我们互通吧？

“西大街”这时应该和谁互通呢？如果没有特殊的原因（比如物业纠分），基本就两个原则，一是尽量选环上的节点，二是就近接入的原则，这个近不是指直线距离，而是光缆距离，我们光缆毕竟不能飞过去，得沿着街道去走管道或者杆路吧。西大街选了一个上家叫做“金苑酒店”的环上节点，说“哥，今后我就跟你混了”。新建光缆需要多长？我们在地图上用测量工具沿着街道量过去，300 米，给点余量 400 米光缆就差不多。如果新建基站要纳入环中，还要在环上选另一个和“金苑酒店”相邻的基站，要再建一条光缆，不管一条还是两条，光缆这事算是搞定了。

有了光缆就有了基础，大哥“金苑酒店”说，你是要什么接口啊，是纯光的还是要电的？要是光口，哥这给你直接拉过去，要是电口，哥管不了你那么远，你得自己有个设备啊。

我们就当是需要 FE 和 E1 电口吧，这时我们就需要新购一台设备，什么设备呢？把新建站的线路、支路接口统计一下，你是要接入环路还是单链，下面还要不要再挂“小弟”？环路就是 2 个 GE/10GE 接口，和原环路的速率一致，单链就是 1 个 GE；支路呢，你是 2G、3G、4G？分别需要什么接口，要几个，一统计就有个总的需求出来了，基本上末端基站的设备也就那么几种配置模型，选一个能满足需求的模型，设备选型也就完成了。接下来相关配套的综合配线柜、DDF 单元、ODF 子框和配套的各种电缆、尾纤这些也要考虑进来，一个也不能少。

新建站这头是完事了，可是你要 GE 或者 10GE 拉到上游站“金苑酒店”，人家有没有口给你接呢？有空余光口自然万事大吉，没有口是要扩光模块还是光板，你也得给人家考虑进去，万一没槽位光板都没得扩，就得换个大一点的设备，还有，上游站也需要熔接新建光缆的 ODF 模块，没有空余的话也要新增，这些东西都是咱得买单的。整个需要我们考虑的东西都在下面这张图里，用红颜色表示：



接入工程就是这样简单而繁琐，几百个站需要多少光缆、多少设备、多少板件，那就一个一个站的按部就班做下来，按照采购单价估算一下，总共要花多少钱也就出来了。最后计算出平均每站花了多少钱，用了多少光缆，和前期工程比较一下看一下指标是否合理，如果不合理，需要看看偏高或者偏低是有什么原因，还是计算过程有误。

至于这么多站接入进来之后，带来的环路流量的增长，接入环、汇聚环能否满足业务需求，这些还有另外的工程去考虑，也就是我们要讲的下一节—网络优化。

4.4 我给网络当医生

就像我们人会生病一样，传送网每隔一定的时间会出现各种问题，有问题就需要解决，没有问题也可以对网络进行调整，使网络更加健壮，这些工作统称为网络优化。如果把接入类项目比作吃饭穿衣，那么网络优化就是定期体检、看医生，有病就要治病调理，没病也要预防保健，而我们每一个网络建设管理、设计、维护人员就充当了网络的医生的角色。

医生看病开方需要病历、化验单，需要望闻问切，我们网络医生也需要收集网络的相关数据、基础资料作为参考，需要掌握网络的组网图、业务配置、单板配置这些情况，还要知道网络的哪部分出现了什么“症状”。接下来我们就 MSTP/分组网和 OTN 分别说明。

MSTP/分组网优化

首先我们给网络做一个“血常规化验”，也就是针对一些共性问题，包括容量问题和安全问题。

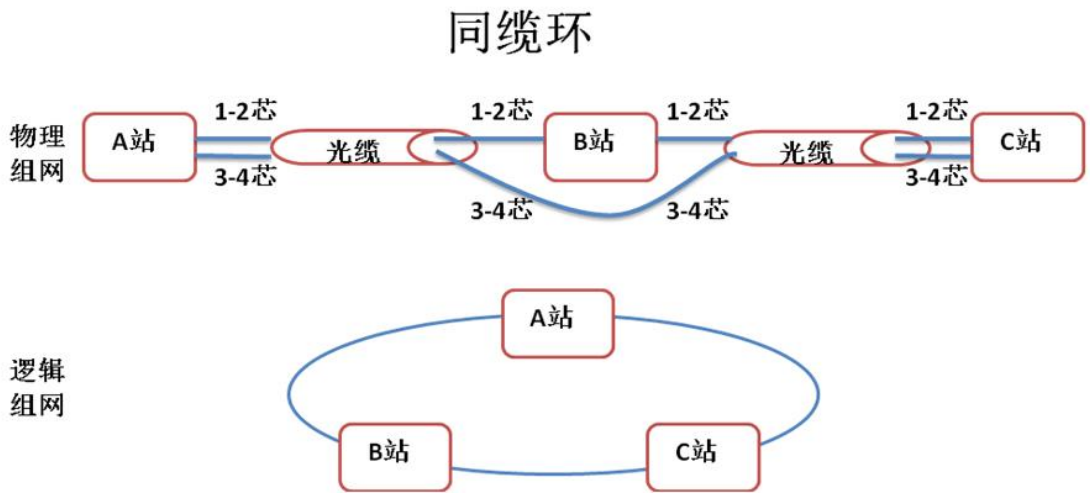
MSTP/分组网容量问题:

网络没有容量就接入不了业务，所以容量问题是硬伤，是必须解决的。我们对于网路容量给带宽利用率设定一个门限值，一般是 70%左右，对于 MSTP 在网管上进行时隙利用率统计，对于分组网需要一段时间的流量监测，就能得出一个当前的利用率，在此基础上还要考虑到

一段时间内的业务预留，一般是指本年度的新建基站和其他一些可预见的带宽占用，得出的最终利用率如果超过门限值就说明带宽不够用了，需要调整。
如何调整，在前面第一章 1.7 节--MSTP 保护和组网部分有介绍，包括新建环路、拆环、整环升级、部分升级等，这里不再重复。
上面说的只是理论上的方法，实际上目前对于 MSTP 容量不足的问题，我们从开始大力建设分组网之时，已经开始控制 MSTP 的建设投入，随着分组网的部署，MSTP 需要承载的业务已经少量增长甚至下滑，基本上也就是新增一些大客户的专线需求，对于原来在 MSTP 上承载的 3G 数据业务，可以逐步割接到分组网上承载。也就是说 MSTP 这张网我们不再作为重点去建设，只是尽可能的发挥它目前的能力和价值，出现问题尽量采用其他办法去解决。
另外，设备的能力不足、槽位不足也属于容量问题范畴之内，这类问题只能是更换更高级别的设备去解决。

MSTP/分组网安全问题:

安全问题包括超大环、同缆环、超长链等等。
超大环是指环上节点过多，比如超过 8 个或者 10 个，首先需要有一个标准去衡量，超大环既影响单站带宽又不安全，怎么改造？拆呗！
超长链，一般是指链上节点 5 个以上的，怎么改造，新建一条光缆把链改造成环。同缆环是特殊的一种环，我们常说逻辑成环，是一种不得已的组网手段，也就是在网管上看逻辑上组网是个环，实际上是个链，也就是光缆实际上是单路由的。
同缆环能够保护的仅仅是设备、单板层面，比如图中 B 站设备瘫了，A 和 C 还可以正常通信。但是同缆环保护不了光缆线路，因为光缆就一条，要是断也十有八九是全部纤芯都断了，比如 A 和 B 之间光缆中断，那么 B 和 C 的业务也就中断了。事实上多数的故障原因都来自于光缆线路，所以同缆环也就是形同虚设。



同缆环怎么改造呢？新建第二路由光缆，也就是把上例中的 3-4 芯光缆用另外一条光缆去承载，这里建设可行性和难度我们不讨论，就说纯理论，具体实施难度肯定是相当大的，否则当初为什么要去建同缆环呢？
很多建不了光缆的地方只能使用微波，微波的稳定性比较差，所以有条件了就要建光缆改造成光纤通信；还有一些在网运行多年的厂家已经停产的设备，设备性能较差而且厂家也无法提供正常的技术服务，有条件就要换成新的。
这些是常见的问题，对于其他的问题我们具体问题具体分析:
比如一些汇聚节点存在低阶交叉容量不足的问题，多数是因为一些历史原因，导致对低阶交叉的使用不合理造成的，还是举物流的例子，这么大一车的货物运到西安中心，需要每个箱

子都打开重新去分拣、装箱，这本可以在其他节点完成业务打包整理的，所以是不合理的，需要依靠业务调整的手段去解决。

还有个别的环路，由于光缆经常 2 处以上中断导致业务丢失，传输俗称开环，这是纯物理线路的问题，也就只能从光缆层面去解决，要么找个安全的路由，要么只能考虑拆环、纤芯租用、纤芯置换这些方式去解决。

波分系统优化

同样，我们分容量和安全问题去说。

波分系统安全问题:

波分系统也有单链/同缆环的问题，这里就不说超长链了，因为波分系统容量大业务重要，一个节点的单链对于波分就是迫不得已的，或者说不允许的，长和超长就更说不过去了，改造方式同上，建光缆，链改环。

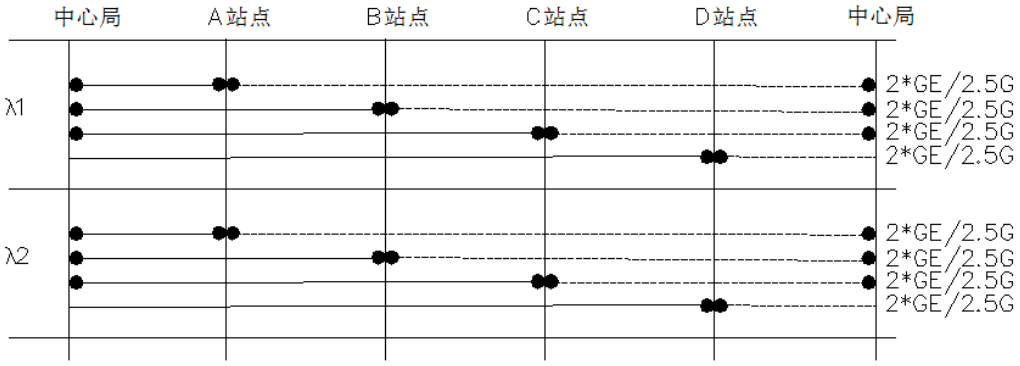
波分安全问题还有和上面一样的开环的问题，光缆线路不稳定的段落超过 1 段，对于传送网来说是致命的灾难。我们可以在问题段落使用 OLP 进行重点保护，但也需要两点间的不同路由的光缆。其实传送网的安全问题归根结底就是光缆的问题，所以解决问题也要从光缆上去落实，没有光缆路由，什么保护都是白说。

波分系统容量问题:

对于波分系统容量，多数时候我们面临的不是容量不足的问题，如果一个 80 波的 OTN 系统，80 波全部用的满满的无法满足后期业务需求，这个问题不需要讨论，只能是新建系统，在干线传送网上业务量巨大，建几个平面都是业务驱动所使然。

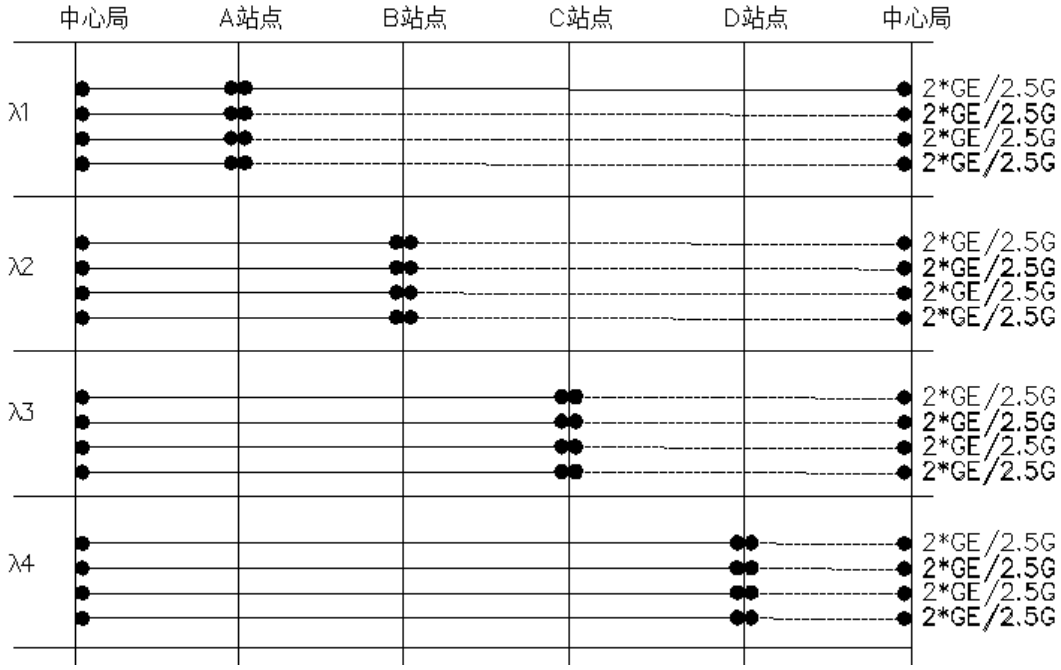
而对于一个本地网，一个 80 波的系统用了可能只有几波、十几二十波，看起来容量还有很大空间，貌似容量没问题，可我们要重点说的，是资源利用是否合理的问题，也就是波道优化、整合的问题。这个问题一般指的是子速率波道，也就是单波 10G 系统中的 GE、2.5G 业务。

在网络建设的早期，GE、2.5G 的需求占主要地位，各种业务需求的带宽比较小的时候，我们一个环上几个节点共享一个 10G 波道，这样去开通业务是无可厚非的，这种情形下，每个站点 2 块线路板一块支路板这是标配。后期业务量增大，ABCD 四个点一个子波道已经不够用，又面临扩容，我们又扩了第二波，和第一波的建设方式一样。



对于这个例子我们算一笔账，比如这一组板件（2 线路+1 支路）造价是 20 万，我们这 2 波就需要 5*2*20 万=200 万，实现的总带宽是 20G，也就是 10 万/G 的单位带宽造价，给每个站点带来的带宽是 5G。

而如果我们下图的方式，中心局到每个点开通一个 10G 波道，造价就是 40 万*4=160 万，实现总带宽是 40G，单位带宽造价为 4 万/G，每个站点独享 10G 带宽，波道配置如下图:



上下两张图一对比，后者单站带宽大一倍，而投资却小于共享波道的方式，单位带宽的造价后者也节省了 60%，这投资效果的差距是非常大的，这也是我们要讨论的问题的关键所在。而另一方面，前者只占用了 2 个波道资源，而后者占用了 4 个，不过这不是问题，就像原本一个车位卖 10 万，现在 10 万可以买两个车位，那多占了一个车位资源是问题吗？

波分的波道资源和无线的频谱资源是不同的，无线的频谱是运营商斥资多少个亿买来的，而且无线信号都是在空气中传播，互相是可能有干扰的，所以要加以合理规划、分配、使用。而波分的一个系统搭建起来的成本很低，只有占用的两芯光缆可能比较宝贵，波分的波道资源又是每个环路相互独立的可以重复使用，所以一般空波道不需要去过分的纠结（注意这是“一般”，不绝对，如果光线资源紧缺，波道配置率又较高，节省波道资源也是重点考虑），只有配置了线路板的波道才是有价值的，是可用的。而这些线路板、支路板却是花了昂贵的价钱购买来的，才是宝贵的资源。

前者的波道配置方式可能在建设初期是合理的，是有历史原因的，但是随着业务量的逐渐增大，合理的规划波道配置非常重要。有时我们面临的是这些即成的事实，钱已经花了省不下了，我们做这些优化工作的结果就是可以节省很多单板或者用这些单板去实现更多的价值，就需要去进行波道整合。

波道整合的思路就是上面讲的，去按照各节点的带宽需求、业务发展预留去综合考虑，去算账，然后制定合理的方案之后去调整波道。现实应用中的业务需求和波道配置五花八门不尽相同，这里举的例子是一个典型情况，可能有些站点需求可能 5G 就足够了，我们可以将 2 个点共享一波，投资效果也是优于多点共用一波的，要具体问题具体分析。

支路板是可以通用的，而线路板可能是可调的或者非可调的，如果是可调的，优化之后结余下来的单板可以灵活调配；但如果是非可调的，我们就拆下来很多多比如λ1-4 的单板，这些单板到了其他系统也是无用的，因为我们都是从λ1 开始建设网络的，你这富裕几块λ1-4 的板子，对不起我这也不缺，这些单板就只能用于新建系统中使用。有时原来 GE 需求的站点直接升级为 10GE，原 GE 波道的线路板、支路板也就空余出来了，也面临同样的问题。针对这个问题又引出了一个新的问题，可调的线路板的价格是高于非可调线路板的，我们不建议全网都购买可调的线路板。现网中多数 OTN 网络的波道配置都是从 1 开始的，如果在网络新建时，如果一个本地网有 N 个环路，各个环路分别从 1、11、21、31、41……号波道开

始配置，后期波道调整的时候，大家都可以互通有无，所以我们节省这些资源能否得到再利用，就取决于我们的一个建设思路和习惯，这个在后期是有很大区别的。

4.5 网络架构很重要

这一节我们不谈系统，我们要说说网络架构，要说的不多但是很重要。

MSTP、分组网、OTN 的组网结构都是水平分区、垂直分层。水平分区是将大的网络分成若干个小的相对独立的区域,每一个区域都有对应的汇聚节点、接入环路负责传送本区域的业务，可以使业务能够就近接入，就像我们国家分为 34 个省、自治区、直辖市、特别行政区；垂直分层是将网络分为核心、汇聚、接入层级结构，每一层的节点各司其职，就像我们国家管省、省管市、市管县、县管乡镇、乡镇管行政村，不这样分层国家也管不过来，分层组网的好处在 MSTP 组网部分也已经提到。

既然每个网都是分层分区的，都有着网络架构的概念，为什么还要把网络架构单独拿出来说呢？此架构和彼架构相比重点有两点，一是要长期稳定，二是要通盘考虑，尤其重要的是第一点，所以我们先说稳定的问题。

网络架构的稳定

汇聚节点就像我们一个单位的中层管理干部一样，干部的待遇、权限、资源都应该有相应的优势，去支撑他去行使权力尽其义务，同样汇聚节点的地理位置、机房产权、机房面积、后备电源、出入机房管道光缆等也应该是完善而优越的，毕竟汇聚点的设备大、光缆多、业务也丰富，既然要处理这么多的业务必然一些资源的需求和消耗大一些。

可实际上有些汇聚机房和普通基站机房并没有太大的差别,仅仅是安装了某个网络的汇聚设备，就被推上了汇聚节点的位置，比如某县城的某基站装了一台 10G MSTP 设备，那这个基站就是 MSTP 的汇聚节点，那这样有什么问题呢?--频繁的机房搬迁。首先就是产权问题，房子是租来的，人家要你搬你就不得不搬；其次就是面积，汇聚机房要装的设备和光缆都多，设备机架、ODF 架这些都得有地方摆，摆不下了怎么办？搬!

搬家的经历我们都有过吧，费钱、费力、费心，搞的你身心俱惫，可不是找个搬家公司那么简单，而且汇聚机房的出入管道、光缆这些也没有办法搬走，怎么办呢？再新建。对于网络建设来讲，机房搬迁带来的网络重新规划设计、业务的割接这些事情费心费力都认了，借用 TVB 的那句经典台词，呐，搬家这种事情呢大家都不想的，做人呢作重要的是开心。但是费钱可 hold 不住，那是真金白银的，一个汇聚机房搬迁怎么着不得个几十上百万呢？

所以，机房稳定的重要性就出来了，稳定最主要的就是产权、面积，房子是我自己买的，谁还让我搬家？面积足够我几十年用的就更没问题，至于其他的相对而言都是次要的。

各汇聚节点都能保持长期稳定，整个网络就稳定和谐，如果能够做到这一点，剩下的网络建设、优化这些工作都轻松不少。这里为什么没谈核心节点？核心节点就像公司老总，该配的应该都配了，一般情况不用我们去操心。

中层领导稳定了,那底层领导呢？就像 MSTP 里的 2.5G 节点或者分组网里的 10GE 接入节点？这个问题和汇聚节点稳定是一个问题，只是资源和政策的倾斜度不同，汇聚层稳定了，接入层也当然要尽量的稳定，那是最好了。

网络架构的一致性

MSTP 有汇聚点，分组网有汇聚点，OTN 有汇聚点，数据网和有些业务网也都有汇聚点，这么多汇聚点怎么考虑,各自为营互不干涉吗？都按照前面说的这个标准,买房子而且要大房子，机房装修配套整一遍，那不浪费钱吗？

网络架构一致性，就是指各个网络之间充分协调、沟通，选出合适的地理位置的机房，一经确定尽量不改变，大家把面积、动力、承重等各种需求汇到一起，按照大家共同的要求去建设网络架构，所以说网络架构不是传送网自己的事，而是整个网络的共同的一个梦想。

光缆网

上面机房的事都确定了之后，光缆也要分层去建设，目的也是要将光缆的职责分工明确。汇聚节点之间要互通要组网就需要光缆,叫做汇聚光缆。汇聚层之间的光缆纤芯需求是较小的，因为汇聚层的节点少、容量大，但为什么还要单独分层呢?
以往没有分层建设光缆时，汇聚和接入光缆共用，接入层对光缆纤芯的消耗量是巨大的，也是难以预测的，如果接入层把纤芯用完了，汇聚层要用哪怕 2 芯都找不到的时候怎么办？汇聚层纤芯需求小，但是业务重要，合理的规划建设就是为了避免这些不可控的情况，使重要的业务能够得到充分的保障。
如何保障？建设直达光缆，汇聚光缆只在汇聚节点处上下，中间不开口，这样就杜绝了接入层使用汇聚光缆的可能性，把接入层彻底给屏蔽掉。
汇聚光缆下面还有接入主干光缆，虽然层面不同但道理一样，这些二级汇聚、接入主干节点之间的光缆也只在节点、光交处上下，普通基站可以用但是要接到对应的主干节点、光交，这样方便纤芯的统一管理。

管道网

本文通篇基本上没有提到管道的内容，一是对于管道来说，网的概念相对很淡，二是本人也不太懂，也没人需要我去普及管道是用来穿放光缆的，所以也只能站在整网的角度少说上两句。
一个管道网，合格不合格、好与不好都没有很量化的标准，这不是一个能完全去按照策略、思路去按部就班实施的工程，不好说又不能不说，所以我们也有一些关于管道简单量化的指标，比如主干道路覆盖率，管孔/子孔的利用率。
覆盖率自然是越高越好，但是哪里能建哪里不能建又不是我们说了算，只能作为横向对比的一个参考，谁也不能说覆盖率低于多少就不行，建设难度谁负责谁知道。利用率呢，太高了说明资源很紧张，太低了又是资源没有充分的利用，貌似是不高不低最好。
但是作为我们自己对于管道的建设是否合理还是有一个大致标准的,我们打开一个本地网的市区管道分布图，管道布局疏密有致，业务密集区域则密，业务欠发达区域则疏，这应该是合情合理的，说明管道的建设没有跑偏；无论疏与密，管道都应该是呈网状的，至少隔几百米至一公里要和其他管道互通，而不是一条道跑到黑，和其他管道老死不相往来，这样的管道使用效果就不会好（指市区管道，高速、干线除外）；管道的建设应该是有倾向性的，核心、汇聚节点之间的路由、主干道路上的资源应该相对丰富，因为业务量大、业务等级较高。

4.6 贺岁大片-三年滚动规划

将上面 4.3-4.5 的工作内容，根据需求做连续 3 年的建设方案，将眼光放远，将需求做实，把握方向，通盘考虑，适当超前，得出每一年每一块的具体的方案和投资，然后汇流成河，得出三年的总投资，最终得到的结果就是我们每年一度的贺岁大片--三年滚动规划。
其实关于滚动规划原本想写很多，牢骚也不少，不过最后还是觉得此时无声胜有声的效果更好

所以标题也起的这么喜感，那就不啰嗦了，滚滚更健康，相信很多同仁也会有一些共鸣。

最后一句话总结：天空飘来五个字，那都不是事！
谢谢！
全文完。