# 交换网络环境的故障诊断

美国福禄克网络公司

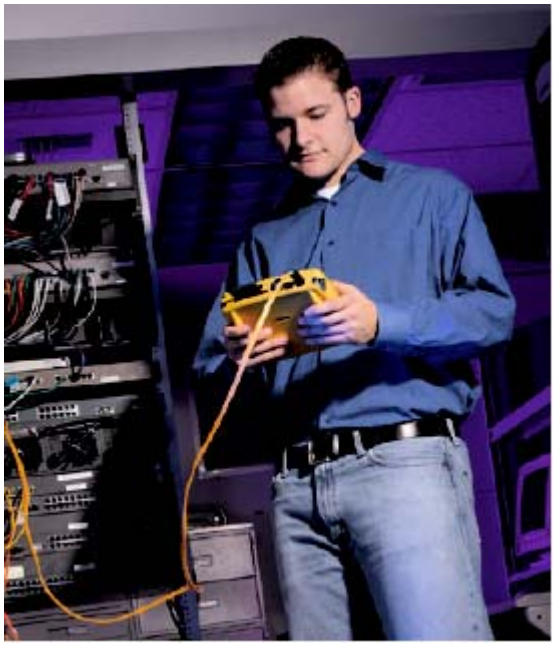
2004年11月

*十年前，网络相对简单。网络设备包括集线器、网桥和路由器等，每个设备都是一个独立部分，相互之间可以识别。故障诊断也相对简单。如果连接的是一台集线器，故障诊断则采取对冲突域进行故障诊断的原则。在冲突域连接至一个网桥的地方所有问题都消失了。故障诊断这时协议分析仪是故障诊断的最佳选择之一，特别是当用户了解了网络的基础和使用的协议后会更为有效。*

*在这之后，出现了交换机。*

*交换网络环境出现的问题与前面提到的共享介质环境的问题有些类似。例如：发生了什么问题，问题是谁带来的，问题有多么严重？最主要的区别是问题需要与交换机的一个特定端口相关联。*

- 在交换网络环境中应该考虑的问题包括：每个端口的忙碌状况
- 如何识别和跟踪错误源？
- 广播风暴的源头是什么？
- 交换转发表是否运行正常?
- 哪个站点连接在这个端口上？
- 交换机对协议或端口是否有速率限制？
- 这个端口在 VLAN 中吗？如果在是同服务器或服务在同一个 VLAN 中吗？



在一个交换网络里，您如何确定从哪里开始动手查找问题？想深入“透视”一个交换网络是非常困难的。首先，在2层交换的时候还是桥接转发方式，但到了3层交换却有了更高级的特性和转发规则，例如VLAN。到了4层交换，就更加复杂了，出现了更高级的转发和负载均衡技术，故障诊断故障诊断和解决就需要更多的交换机配置知识。

在安装完一台交换机后，每个交换机的半双工端口就构成了一个冲突域。如果该端口连接了一个集线器，集线器下面连接若干站点，那么冲突域会扩大。但随着交换产品的价格下跌，现在大多数新建的网络每个交换端口都只连接一个站点。因此，在半双工连接情况下，冲突域仅针对一个单独的电缆链路。

交换机通常是一个独立广播域的一部分，包括串连或者并连的任意数目的其他交换机。如果使用了OSI模型3层的功能，就可以创建多广播域，广播域的数目与VLAN数目相等。最极限的情况，如果交换机功能允许，每个端口可以配置为一个独立的广播域。可以把这种情况描述为路由到桌面。为每个端口创建一个独立的广播域后，故障诊断就会严格受限。但是如果我们把每个端口设置为一个单独的广播域，交换机在转发流量的时候，每个端口都需要路由服务，这会占用交换机CPU的有限资源。在网络环境中，对每个单独的端口进行路由请求和应答是非常困难的，我们应该避免这样的配置。不幸的是，这种情况在实际情况中非常常见，网络中经常发现服务器全部在一个子网或者广播域中，所有的客户在另外的子网或者广播域中。在这种情况下，

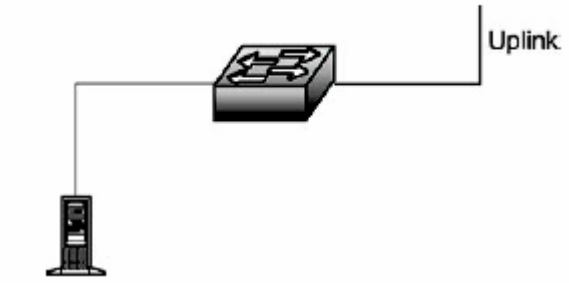
所有的请求都必须路由。如果维护行为限制在一个单独的服务器群里，那么考虑把服务器放进单独的VLAN里。然后把使用这台服务器的用户放到同一个VLAN。这样就可以使用2层交换的桥接方式来交换流量，只有很少的请求需要路由。如果服务器支撑多于一个用户区，可以在服务器上多装一块网卡来实现到用户的2层交换连接。

## 对交换机进行故障诊断的5种技术

可以采取5种基本方式来透视交换机。每一种方法都不同，都有积极或者消极的一面。类似在网络中遇到的其他问题一样,没有一个最好的答案。最合适的方案往往取决于您手中可以利用到的资源(什么工具可以使用或者以前安装过什么工具)，而且使用这些技术有可能造成服务中断。

即使把这些方式组合起来，也不能监测到所连接的网络，在交换的环境里面，也不像集线器那样方便监测。我们几乎不可能看到通过一个交换机的全部流量。大多数的故障诊断会假设流量会在站点和所连接的服务器之间或经过故障诊断交换机uplink口通过。而实际上如果2台主机直接传输信息的话，就不会使用交换机的uplink口或者任何其他的端口来交换流量。除非你知道具体用到哪个端口，否则是监测不到的。

举个例子，如**图1**，一台服务器接入一台交换机。在反映有问题的用户中，一部分是直接与这台交换机相连，另外的一部分用户是由这台交换机的uplink口从其他路由器或者交换机连接上来的。故障报告是访问服务器“慢”，这样的故障报告对技术支持工程师来说基本上没有任何价值。




*图一、一个最基本的交换机环境*

## 方法1：通过TELNET或者串行口接入服务器

高级的网络技术支持工程师或其他知道交换机密码的人在进行故障诊断时可以选择通过TELENET或者交换机的串口登陆，来检查交换机的配置。**（如图2）**

（）

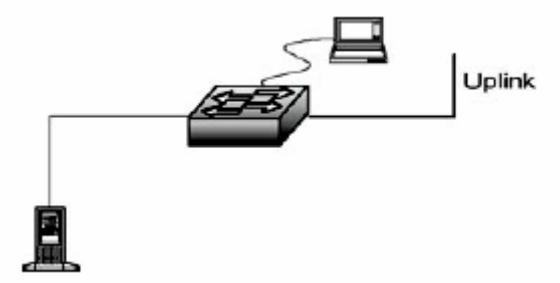



*图 2、使用RS-232 控制端口*

交换机配置可以通过上面提到的2种方法查看，虽然问题不一定是配置引起的。不管问题是操作系统有BUG还是配置不完善，都不能从配置列表中轻易的查看出。配置信息在定位交换机是否像预期的那样运行上比较有用，但针对故障诊断就不是

了。为了验证交换机的配置，往往需要使用多种的交换机故障诊断方法配合。

很多交换机都带有实时的故障诊断工具，因为交换机生产厂家和型号的不同，这些故障解决工具的特征也各不相同。但是要使用好这些工具，必须依靠一定的理论知识和实际经验。

## 方法2：连接到一个空闲端口

最简单的故障诊断方法是在交换机的空闲端口接入一个监测工具，例如协议分析仪。

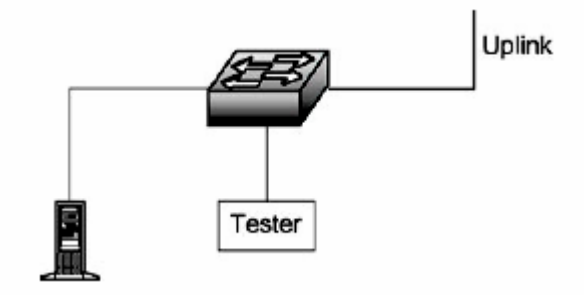



*图3、从任意端口监测*

把监测工具接入交换机的一个空闲端口，不用中断服务就可以查看所属广播域。该监测工具与广播域里的其他站点一样有相同的权限。

不幸的是，交换机（做为一个多端口的桥接设备）几乎不转发流量到监测端口。因为桥接设备就是这样设计的，流量直转发到所属的目的端口，不会去其他的端口。协议分析仪因此几乎监测不到流量。

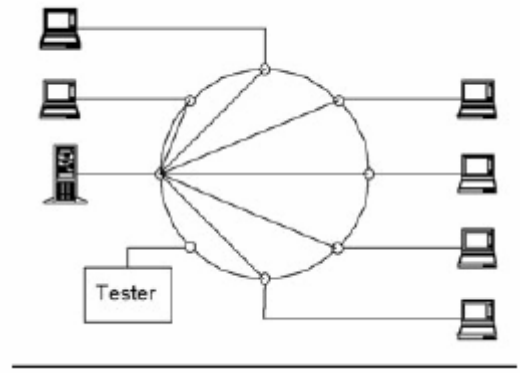



*图4、交换机在源端口和目的端口之间转发流量。非常少的流量会转到其他端口。站点和服务器之间可能每秒钟会转发几千个帧，但是监测端口每分钟只能看到几个帧*

转发到监测端口的流量几乎全部都是广播，包含一些零星的目的地址不明的帧。这些零星的帧是由于路由转发表老化的结果，经常是目的端口不明的帧。一些经验不够的技术人员看到这么高的广播（接近100%），却没有注意到端口利用率很低，就误判网络出现了广播风暴，其实不是。

这样查看交换网络几乎没有用，因为监测工具必须获取流量。获得的流量或者对广播域的查询对网络搜索和发现其他类型问题是有很有帮助的，但对解决用户连接慢的问题并没有多大的帮助。

对大多数交换机来说，都有一个更好的选择，可以把需要监测的端口流量备份到一个专门的空闲口。（见**图5**）这种技术通常称为端口镜像。

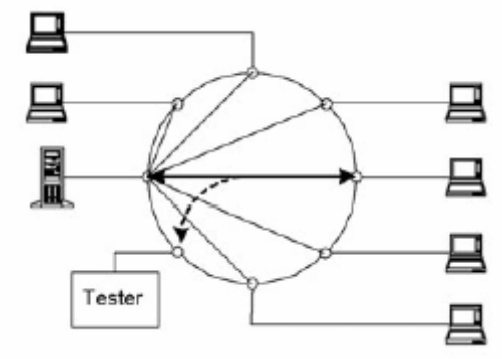



*图5、配置镜像口的逻辑效果*

大多数交换机厂家都提供备份或镜像流量的功能，可以把监测工具接入交换机一个专门配置过的端口。老的交换机必须指定一个专门的监测口做为镜像口，但现在大多数新的交换机可以指定任何一个端口做为镜像口。

虽然交换机厂家实现镜像的方式各不相同，但是有一些基本相同的监测选项。值得注意的是，几乎在所有的情况下，交换机在转发流量到镜像口的时候，同时把错误都过滤掉了。对于故障诊断来说，这意味着同时过滤掉了有用的信息。

此外，实际操作当中需要我们通过控制口（交换机的RS232端口），或者Telnet进程来配置镜像。这意味着除了监测工具之外，我们通常还需要带一台电脑或者终端来对交换机进行配置。

镜像端口经常只是一个“监听”端口，不过很多交换机厂家允许把该端口配置成全双工的。配置了镜像口，监测工具就可以查看报告连接慢的主机和服务器之间的实际流量的备份。镜像口可以只监测交换机的任意一个端口，甚至可以是Uplink口，也可以同时监测交换机的多个端口。但是同时监测的端口很多的话，过高的流量就有可能会超过镜像口的接收能力。

监测端口的输出能力是一个很重要的问题。镜像口可以收，也可以发。在配置的时候，经常关掉了镜像口发的功能。但不管有没有关掉镜像口发的功能（不管镜像口是全双工或者不是），镜像口的接收能力都是有限制的。如果被监测的全双工端口的速率和镜像口是一样的话，交换机在转发流量的时候很容易就会丢包，但是交换机不会通知您。

假设您在监测一个以100M全双工速率连接到交换机的服务器的话，那么服务器在全双工工作的时候，服务器的收发速率都是100M，那么总共就有了200M。然而交换机的100M镜像口最多只能接收100M的流量。所以任何交换机的端口（全双工的）利用率超过50%的时候，镜像口接收到的包就会有丢失。

如果把多个端口镜像到一个端口，丢包的问题就会更加的严重。因为大多数交换机都工作在低容量，这个问题并不会被立刻注意到。大多数用户连接的平均利用率都很低。只是偶尔会有流量的突发。

如果选择一个高速的镜像口，就可以减少丢包的问题。例如把**图6**中的100M镜像口换成1000M，那么就可以很容易的接收200M的监测流量了。

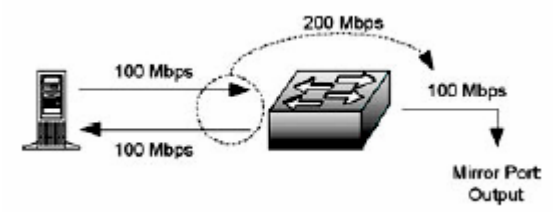



*图6、监测口限制了输出流量*

## 方法3：在链路上接入集线器

使用集线器很具有战略意义。对很多网络来说，大多数发送和接收的流量都来源于文件服务器之类的共享设备。在交换机端口和文件服务器中间接入一个集线器，再把分析仪接入集线器，实际上就把分析仪和文件服务器接入了同一个广播域。如**图7**所示。使用这种方法，技术支持人员就可以看到文件服务器所有进出的流量，帮助技术支持人员解决一系列的问题，包括用户登陆失败、性能低效、连接丢失等。

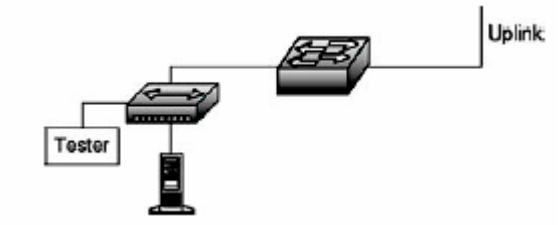



*图7、使用集线器监测交换机端口*

接入集线器的方法很多时候都不实用，特别是在需要监测多个服务器的时候。在哪里接入集线器合适？所有的服务器都要连接吗？如果是用一个集线器，换来换去连接的话，您一定不希望您的网络这样频繁地被干扰。连接集线器所带来的时延，经常会带来连接的丢失。另外，很多时候监测工具并不支持服务器所采用的技术或者连接速率。

使用共享集线器监测一条链路上的所有流量和错误仍然是一个有效的方法。这几乎是唯一一种可以在交换网络环境中实际查看和分析MAC层错误的方法。使用SNMP来发现这些错误也可以。但是，为了更好地进行错误分析，还是用监测工具直接查看最直接。

接入集线器的方法有2种主要缺陷。服务器链路有可能不是全双工的，或者和集线器的端口双工状态不匹配，这会给监测带来更多的不愿意看到的错误结果。而且使用这种方法时，手头必须要有一个共享集线器。现在很多新型的集线器都类似于交换机，而不是共享的转发设备。接入这种新型的集线器，相当于接入了一个新的交换机，您会看不到想要查看的流量，对监测起不到什么作用。如果接入的是双速率的集线器，例如10M/100M双速率的，可能每个速率都提供了一个广播域，两个速率之间再进行转发。在这种情况下，需要确认被监测链路和监测工具运行在相同速率，才能够使用这种双速率集线器。还有一些集线器提供在所有端口之间转发的功能，更因此把自己标榜为价格便宜的交换机，给人造成误解。他们都不能用在这种监测方法上。

## 方法4：使用一个TAP（监测接口盒）或者分流器

这种方法类似于加了一个共享集线器，不同点是TAP链路只是接收流量，不允许监测工具发出流量。

TAP和分流器这2个词有时候可以互换，虽然分流器通常应用于光纤链路。在光纤链路上，分流器会把光在初始路径和监测路径上进行分光。典型的分光比率包括80：20、70：30以及50：50。以80：20为例，80％的光通过分光器继续传送到原始路径，20％的光转发到监测路径。如果光纤本来就有问题，或者传输的距离很长的话，光分流器带来的20％光丢失，很容易造成链路出现问题。分流器在光纤链路上很容易就会带来3 dB的衰减。有些分流器要耐用一些，因此即使在链路的一端安装分流器造成链路中断，还可以将其换到另一端去安装，让链路正常工作。光分流器不需要电源就可以工作。需要注意的是，分流器是带内（Inbound）监测设备，所以分流器的线缆正确连接就非常重要。

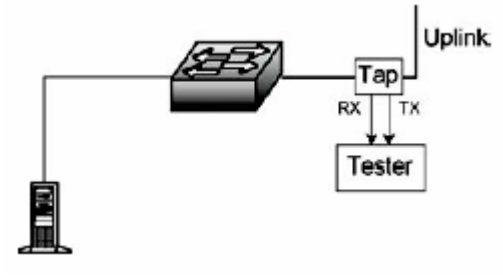



*图8、使用TAP或者分流器*

电口的TAP也会带来信号丢失的问题，因为TAP需要信号来识别通过的流量。对电缆来说，这相当于增加了衰减，如果链路本身已经有一定问题或者链路很长的话，TAP的引入有可能会造成连接中断。电口的TAP工作需要电源，信号被恢复并重传到监测端口。如果设计的好，在TAP掉电的时候，链路应该也不会中断。

对链路使用TAP进行监测的方式是一个很好的查看链路流量的方法。一旦安装成功，TAP对被监测的设备来说就是透明的，可以随时使用，而且不会带来更多干扰。不幸的是，在接入TAP的时候，链路必须暂时中断。此外，TAP或者分流器会按照2个独立的方向提供流量。也就是说，发送和接收是分开的。

为了同时监测通过TAP链路的请求及响应，需要一个带两个输入口的监测工具。双端口的监测工具可以分别监测每个方向，也可以把两个方向的链路集中在一起分析。您也可以选择每次只监测一个方向的流量，但这样分析起来会比较困难。对TAP来说，监测全双工链路和半双工链路，操作上没有什么区别，都可以监测。您可以选择一个单端口的监测工具，监测单一的方向，或者选择一个双端口的监测工具，同时的监测两个方向。

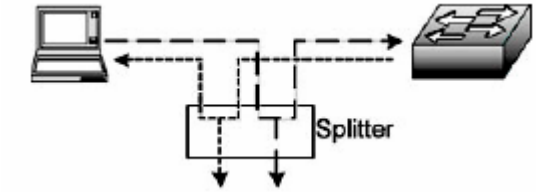



*图9、TAP工作原理的逻辑图*

## 方法5：用SNMP查询交换机

对一个交换网络进行故障诊断的最有效办法，应该是通过直接询问交换机来查看网络的状况。这可以通过SNMP或者连接到交换机的控制口实现。显然，直接连接到交换机的控制口不是理想的办法，因为这就需要对网络中的每台交换机都有物理上的连接。稍微理想一点的替代方法是搭建连接到交换机控制口的终端服务器。SNMP是一个更好的选择，它可以在交换网络带内的任何地方进行查询，不需要附加的硬件。如果您部署了网管系统，还可以配置当利用率、错误、或者其他参数超过门限的时候，交换机主动发出SNMP陷阱。然后利用网管或者监测工具，研究是什么原因造成了门限超出。

事实上几乎所有的交换机都提供SNMP功能，哪怕是最便宜的交换机。它们之间主要的区别就是提供的信息多少。一些价格便宜的交换机只提供简单的SNMP信息，且是针对整个交换机的；而那些价格贵一些的交换机，还可以提供交换机每个端口的详细信息。

SNMP可能是监测交换网络最常用和干扰最少的办法。SNMP控制台不需要非常靠近被监测的设备，只要求有路由可达就可以了，同时交换机的安全配置允许控制台与交换机的代理进行通信。

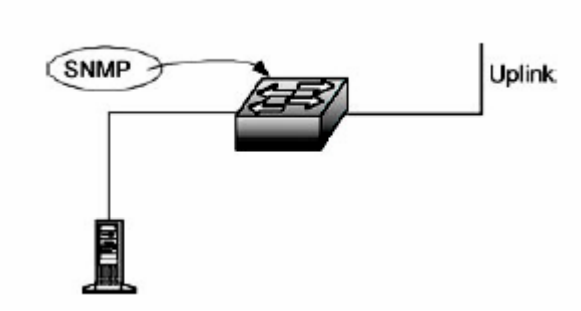



*图10、使用SNMP监测交换机*

虽然交换机可以识别到错误，但交换机本身并不定时地报告错误，所以使用SNMP查询或许是最好的办法。

支持SNMP的交换机有不同的MIB库（管理信息库）。每一种MIB都不同。除了某些对自己的交换机提供支持的私有MIB库，标准的MIB库对交换网络的监测也非常有用。下面是对故障诊断非常有用的一些MIB库。

RFC 1213 - MIB II
RFC 1643 - Ethernet-Like Interface MIB
RFC 2819 - RMON Ethernet
RFC 2021 - RMON 2
RFC 2613 - SMON

很多RFC生成之后就不断地在更新和增强。因此我们要检查最近更新的RFC。例如RFC1213，至少更新和增强了五次，生成了5个新的RFC（2011，2012，2013，2358和2665）。除了定义利用率和错误的RFC之外，有关桥接的MIB（RFC1493）也是非常有用的。

使用SNMP监测网络的时候，必须注意安全性。如果SNMP代理没有限制，那么潜在的任何地方的任何人都可以监测到您的网络动态或修改交换机配置。交换机售出的时候默认打开了SNMP，并且使用的是一个非常通用的密码。SNMP密码叫做通信字符串，使用明文传播，这带来了潜在的危险。SNMP V3提供对通信字符串的加密，减少了这种危险，但是SNMP V3还没有广泛使用。最常用的通信字符串是public。现在，使用public，很多Internet上的SNMP代理都可以被接入。

我们应该立即修改通信字符串。SNMP代理应该为不同的字符串配置不同的接入级别，不同的IP地址、不同的子网也有不同的接入级别。或者根据其它的配置来限制接入的级别。通过路由器接入SNMP代理可能会对SNMP的限制带来一些影响。防火墙也有可能完全阻止SNMP。即使您能够通过SNMP接入代理，也要求代理支持您所要查询的MIB库。大部分厂家完全支持标准的MIB库。然而，也有一些厂家不支持。有时候为了支持期望的MIB，还需要先对交换机的操作系统进行升级。这种方法还有一个问题，如果SNMP代理执行的MIB不正确的话，那么响应就完全是错误的了。虽然这并不是经常发生的，但有时候程序设计的错误，会带来错误的响应。

交换机不响应SNMP的查询有很多原因。一旦这些问题都解决了，SNMP就能够提供非常有效的监测和趋势分析。

## 结论

故障诊断的一个普遍方法是等待用户的投诉。这个方法虽然简单，但是非常有效。用户能够感知到网络正常的性能是怎样的。一旦有性能下降，网络支持中心就会很快收到客户的投诉。有了用户投诉，您就应该从他的接入点开始做故障诊断了。这种方法的缺点是完全是被动的，不具有前瞻性的。

理想的方法是使用前瞻性地监测。包括定期地查询每个交换机、监测每个交换端口的流量、流量的趋势，同时检测其他的相关网段。把问题解决从故障诊断方式变成故障预防方式。

## EtherScope™ ES 网络通，通过对交换网络的超级透视能力，快速地定位问题，解决了交换网络故障诊断的难题

将ES网络通接入网络，您立即就可以透视交换机。Trace SwitchRoute功能显示了交换网络中任意2个设备之间通信的路径。问题无处可逃。

**利用ES网络通，您将会看到：**

- 交换机端口利用率、错误和配置
- 端口连接的主机
- 交换机的VLAN划分和交换机端口所属VLAN
- 交换机扫描，交换机端口的流量概览

让我们看看ES网络通是如何解决交换网络中诸多棘手问题的。

***站点可以访问网络和关键服务吗？***

这是故障诊断或者网络验证的起点。我们必须首先确认网络连接的状态，站点是否能连接到关键服务。很多网络故障解决专家都提到必须先从物理层开始进行验证。

对于网络连接的丢失，ES网络通首先诊断电缆的连接和链路的信号。首先是电缆连接的LINK LED链路灯点亮，表示以太网连接的脉冲存在。接着诊断连接的速率和全双工状态，下一步再进行网络、VLAN和设备的查找。

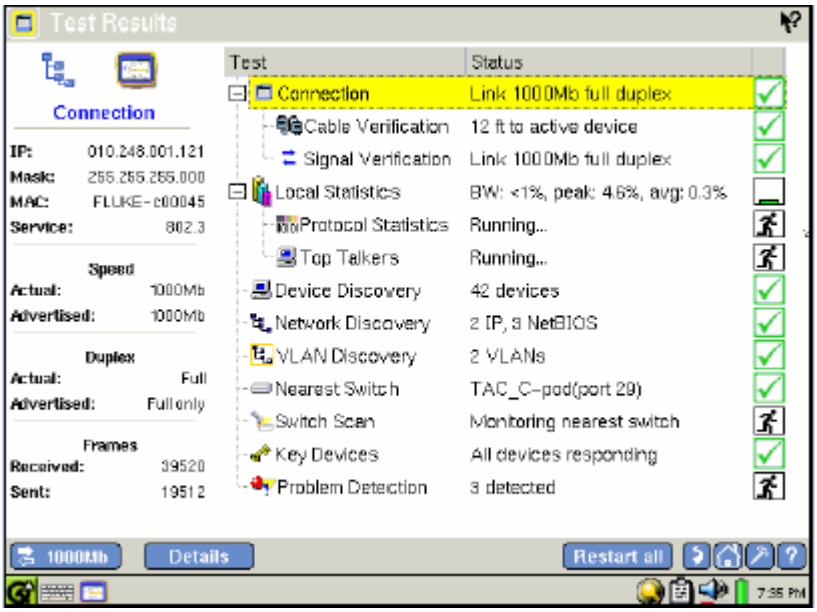



*图1、正在进行的自动测试。LED灯显示和快速诊断验证网络连接*

通过查看电缆认证的详细信息，可以诊断电缆连接的任何错误。显示的结果包括电缆的状态、是否存在线缆错误（开路、

短路、串绕等）。

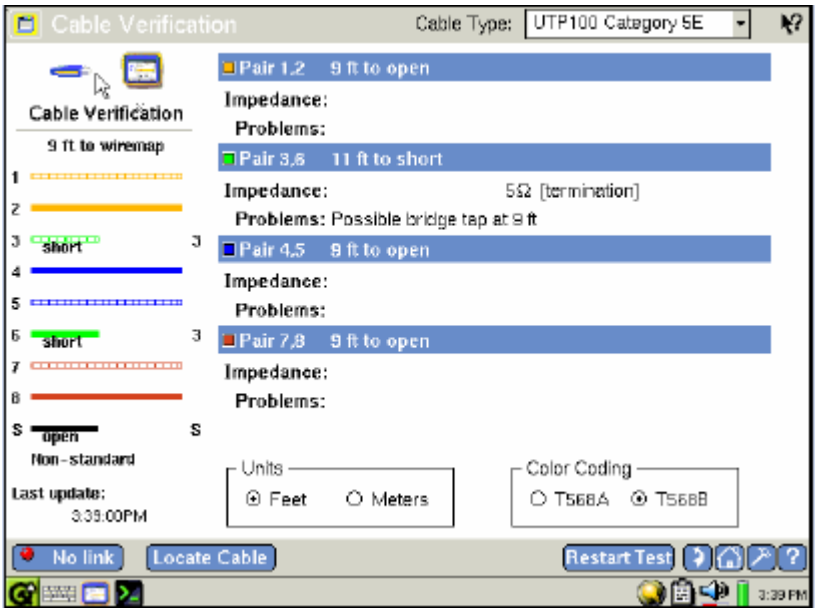



*图2、电缆认证结果显示电缆状况良好*

尽管以太网设备都承诺互操作性，但是自动协商依然是困扰很多网管人员的问题。以前遇到这样的问题，很难确定具体原因。现在有了ES网络通的信号验证诊断功能，就可以透视设备自协商的过程。在这个测试进程里，ES网络通监测有故障链路的电信号，扫描直流电压，扫描POE（Power Over Ethernet）、链路脉冲和数据信号、实际的FLP（快速以太网脉冲）信号、链路两端设备的速率和双工设置状态，从而解决链路的自协商问题。

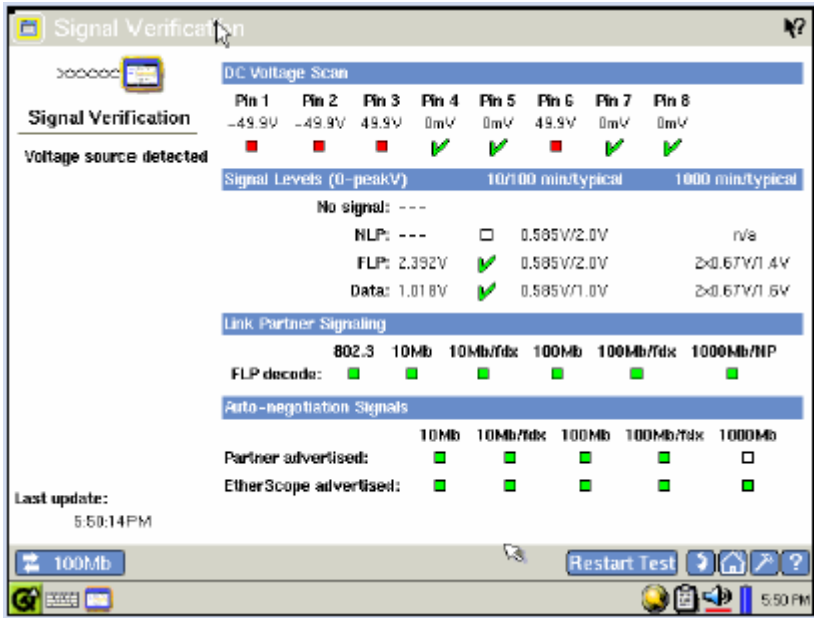



*图3、信号验证指出存在PoE*

**工作站连接在那里？**

使用 ES 网络通的“最近的交换测试”功能可以发现临近交换机的插槽和端口号，在 ES 网络通的自动测试结果中会显示交换机名和所连接的插槽/端口号。突出加亮这一测试结果，预览面板将显示进出端口流量、VLAN 配置和所连接的工作站的数量。

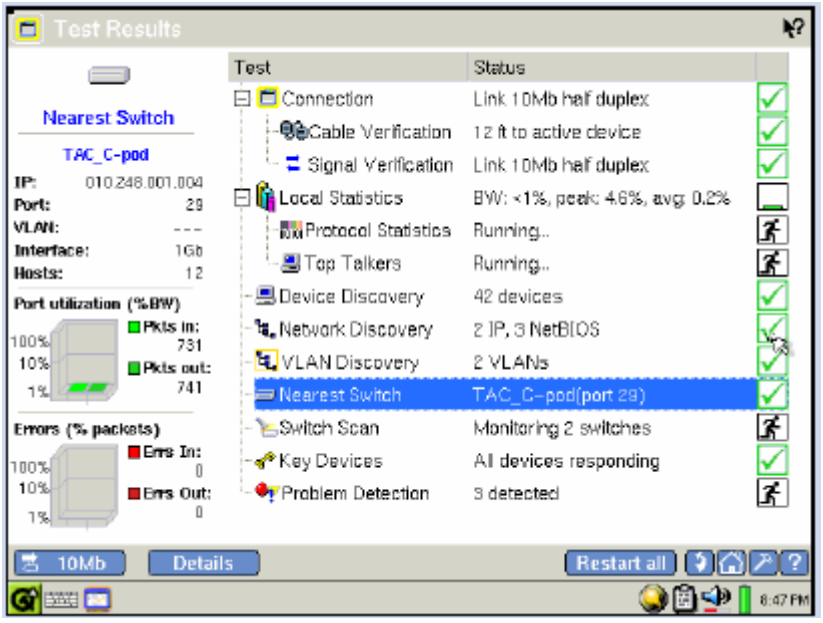



*图4、最近的交换机*

ES 网络通还没有停止测试，它利用网络主动搜索功能记录在广播域内的每台设备。通过对交换机桥转发表条目上的 MAC 地址的比较，ES 网络通就可以知道工作站连接在网络的位置。任一台被搜索到的交换机的“设备详细信息”都可以提供给您丰富的配置信息，包括最近的交换机和连接端口号。

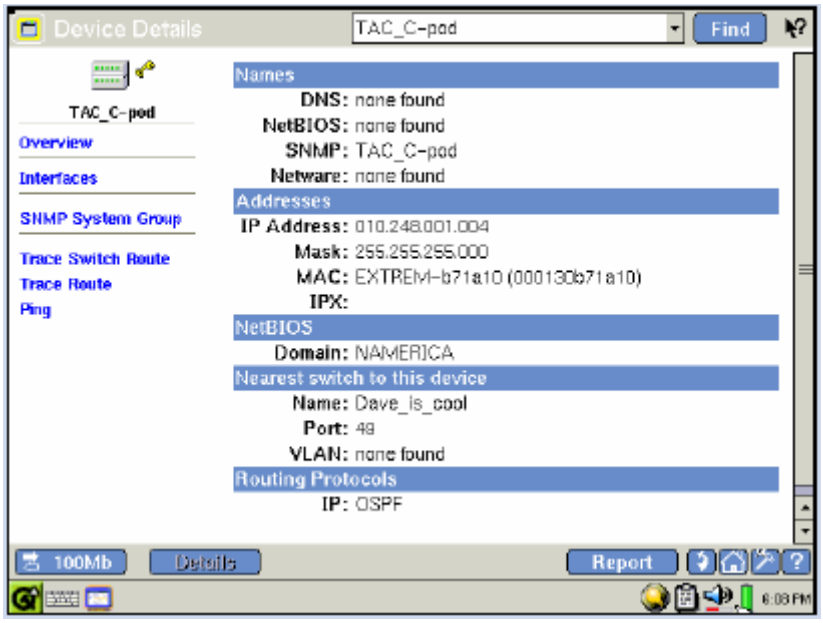



*图5、设备详细信息*

**每个端口的繁忙程度如何？**

通过 SNMP 协议，ES 网络通在交换机扫描对话中对最近的交换机和另一台进行询问，比如收/发端口利用率、速度设置和错误指示（如果存在）等将被显示出来。

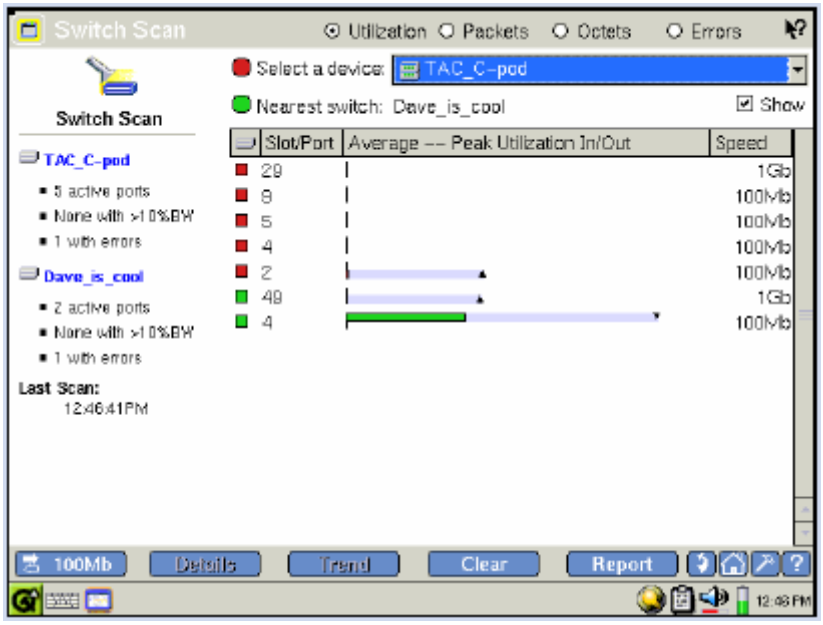



*图6、交换机扫描*

**当前端口连接哪些工作站？**

所选端口的“设备详细信息”显示连接在这个端口上的每台设备。利用方便的搜索功能，通过通过匹配主机全部或部分名称、IP 地址或 MAC 地址，可以快度地定位设备。

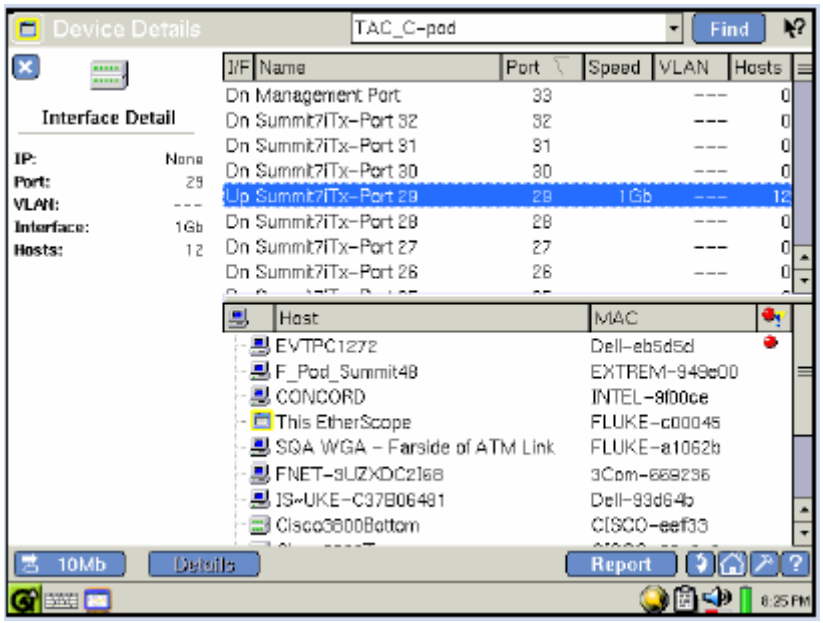



*图7、接口详细信息*

**广播风暴源是什么？怎样找到错误根源？**

“发送流量最多者”将发送包数量最多的站点进行列表，可以根据通过协议类型进行过滤，通过 MAC 地址、错误、广播和多播来进行排序，找到错误源或过量广播包非常简单，只需选择标准然后查看最顶部的站点即可。还可深入查找交互机、插槽和端口的详细信息。

如果交换机扫描显示在某个端口上村杂错误，点击这个端口的“设备详细信息”，有问题的设备将作为主机被列出，CRC

错误和过量包冲突能指示自动协商方面的问题。

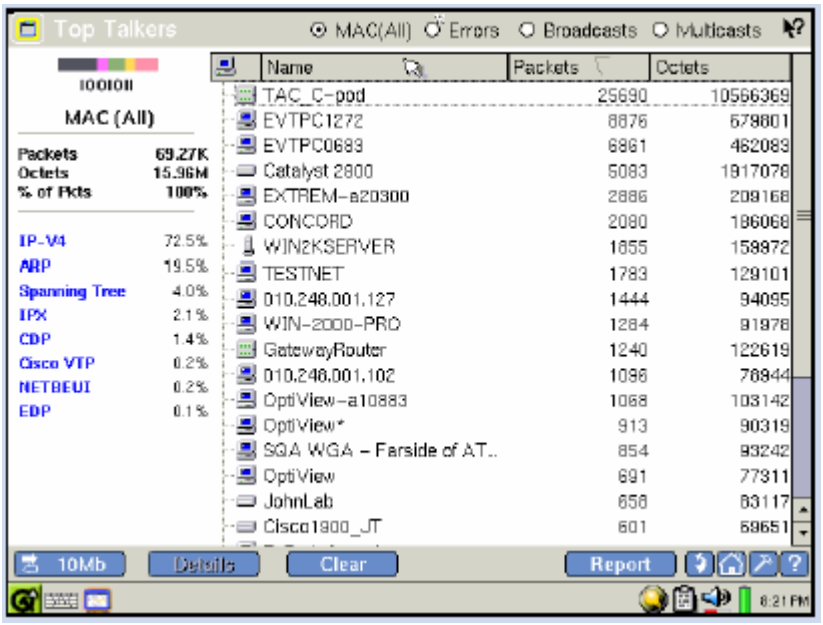



*图8、发送流量最多者*

**桥转发表运行是否正常？**

每个网络管理员都会依赖“Trace Route”命令去测试网络的连通性，了解设备之间的第三层路径，但却不能知道相应的物理端口和第二层的连接路径，这就使得网络故障解决人员消耗大量时间（或者依赖过期的文档）去寻找设备之间的连接关系。ES 网络通的“Trace SwitchRoute”功能可以收集交换机和端口上的每个连接以及网络中交换机之间的连接，显著的减少了缩小故障范围的时间。

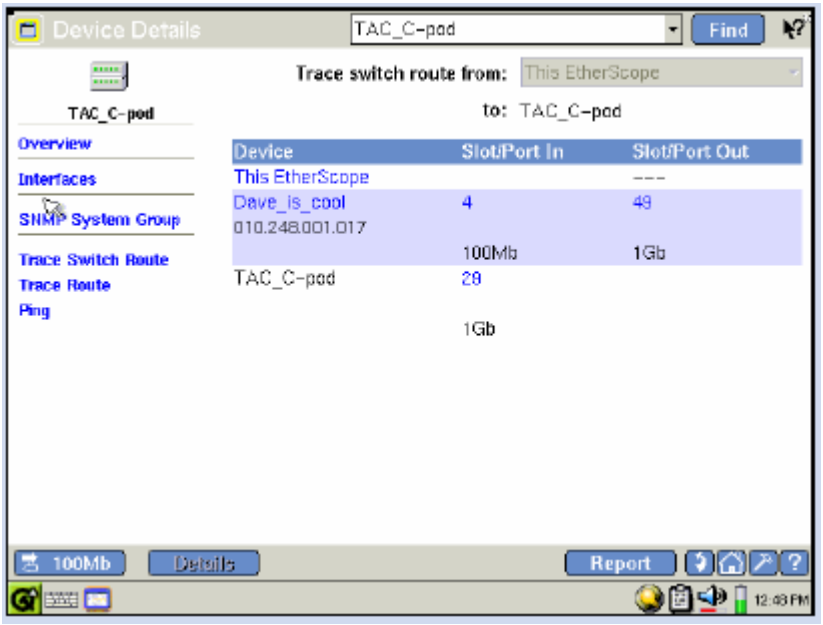



*图9、Trace SwitchRoute*

**当前端口是否属于某个 VLAN？如果是，它与服务器在同一个 VLAN 吗？**

作为网络自动搜索过程的一部份，ES网络通可以发现广播域内的VLAN配置信息。点击首页，然后点击“VLAN 搜索”，可以进行VLAN目录的预览。选择“详细信息”可以查看VLAN的列表，当VLAN被展开后，可以看到交换机端口与VLAN的对应关系。

点击一个端口，选择“详细信息”，就可以得到这个端口的状态、端口流量和设备配置等信息。ES网络通是一款主要应用于广播域的产品，可以自动搜索到被测广播域内所有的设备，包括所有交换机。为了帮助自动搜索过程，使用“设备添加”功能，可以将本地或非本广播域内的设备添加到“搜索”数据库中，这个功能位于用户界面的“设备搜索”和“关键设备”两个页面上。用户增加的设备将保存在仪器中，故障解决过程中将一直存在。如果仪器复位到出厂初始值，它们将丢失。

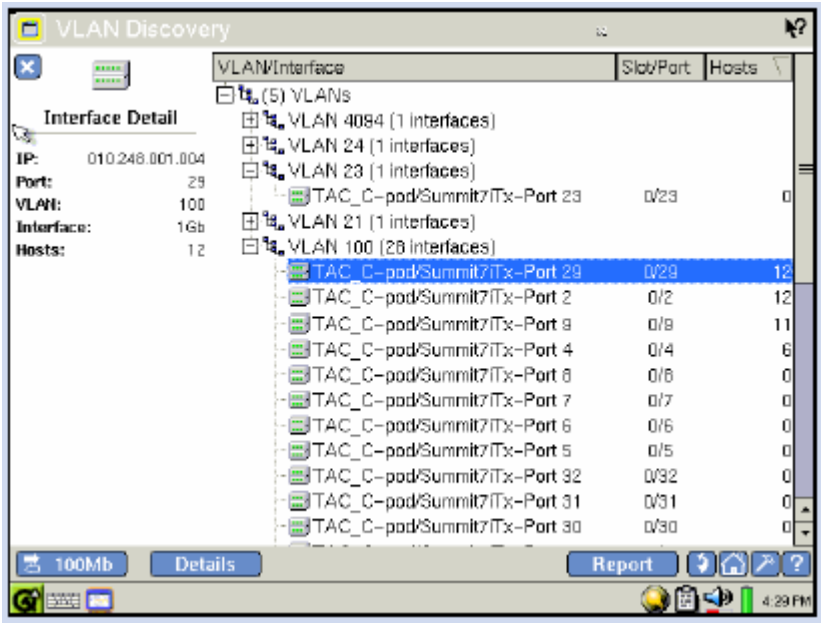



*图10、VLAN搜索*

访问 www.flukenetworks.com.cn/EtherScope 在线观看 ES 网络通的虚拟演示，或查看其它相关信息。