;login:

THE MAGAZINE OF USENIX & SAGE
April 2002 • Volume 27 • Number 2

inside:

SYSADMIN

ISPadmin

BY ROBERT HARRIS

USENIX & SAGE

The Advanced Computing Systems Association &
The System Administrators Guild

ISPadmin

Network Design and Operation

by Robert Haskins

Robert Haskins is currently employed by WorldNET Internet Services, an ISP based in Norwood, MA. After many years of saying he wouldn't work for a telephone company, he is now affiliated with one.

rhaskins@usenix.org

Introduction

This installment of ISPadmin examines how service providers large and small might set up their IP (and associated) networks to provide services to their customers. After covering some network basics, the article illustrates how a small dialup provider might set up its network and how a larger provider might. Issues surrounding the traditional small and large dialup ISP are examined. Finally, such topics as staff requirements, service level agreements, and network design considerations are pondered.

ISP Networking Background

This section contains basic networking concepts and terms and their meanings.

Paid Egress

Egress, synonymous with "exit," is how network engineers refer to the points where traffic leaves the provider's network and enters another entity's network. There are two types of egress: paid and peer. *Paid egress* is bandwidth that the provider buys from another provider to deliver traffic that is not destined for the provider's, or peer's, network. In the greater Boston area, it runs about $500 per megabit/sec (Mbps) per month without local loop charges.

Peer Egress

The second type of egress is *peer egress*, or *peering*, where little or no cost besides hardware is incurred. Peering is exchanging traffic destined for someone else's network directly with them, rather than using paid bandwidth. There are two types of peering arrangements, public and private. *Private peering agreements* are connections that take place in private, common facilities. GlobalNAPS (or GNAPS, a CLEC associated with my employer) allows no cost peering for its customers. If two customers co-located in GNAPS facilities would like to peer, and GNAPS doesn't incur any cost, the providers are allowed to peer without additional cost from GNAPS.

In order for most larger providers to peer, they require a considerable amount of traffic to be exchanged and that the traffic be "roughly balanced." For example, WorldCom requires 150 Mbps of traffic from the provider's network to WorldCom and 150 Mbps from WorldCom's network to the provider's network. (Even providers who qualify for WorldCom's free peering are required to pay for a connection into their facility.) *Public peering points* are facilities set up for the express purpose of enabling peering relationships (e.g., MAE EAST, WorldCom's widely known public peering facility; there are many such public peering points run by a wide variety of providers). In the case of public peering points, the host of the exchange point usually charges for connections into the facility, in order to cover its costs and make a profit.

Autonomous System Number (ASN)

When a provider has multiple egress points in its network, an ASN is used to identify what network the traffic originated from (in the case of outgoing packets) or is destined for (in the case of incoming packets). It usually consists of a unique four-digit number (e.g., "AS1234") which tells other devices on the Internet which network a particular packet belongs to, when a network is multi-homed. An ASN is assigned by the American Registry for Internet Numbers (ARIN) or other Internet numbering authority.

Here are common circuit acronyms and associated speeds for the United States (from the ISP Glossary listed in the references) and a few common service provider acronyms.

Dedicated Circuit Acronyms and Speeds

DS0 (Digital Service 0):	64 Kbps clear channel (normally provisioned by the telephone company as 56 Kbps)
T1 (DS1):	24 DS0s or 1.544 Mbps
PRI (Primary Rate Interface):	single ISDN channel normally provisioned on a T1, supports both ISDN and plain old telephone service (POTS) connections
ISDN (Integrated Services Digital Network):	64 Kbps
T3 (DS3):	672 DS0s or about 43 Mbps
OC3 (Optical Carrier):	155.52 Mbps
OC12:	622.08 Mbps
OC48:	2.488 Gbps

Some Common Service Provider Acronyms

ILEC:	Incumbent Local Exchange Carrier (e.g., Verizon, Qwest)
CLEC:	Competitive Local Exchange Carrier (e.g., Level3, GNAPS)
DLEC:	Data Local Exchange Carrier (e.g., Covad)
POP:	Point of Presence
RAS:	Remote Access Server
ATM:	Asynchronous Transfer Mode
SONET:	Synchronous Optical Network
DOCSIS:	Data Over Cable Service Interface Specification

Small Provider Backbone

Figure 1 illustrates how a small provider might design its network. The box marked "Central POP" is the central site where the provider has access to the Internet. The boxes marked "Remote POP" represent off-site locations housing RAS gear or customer-dedicated connections.

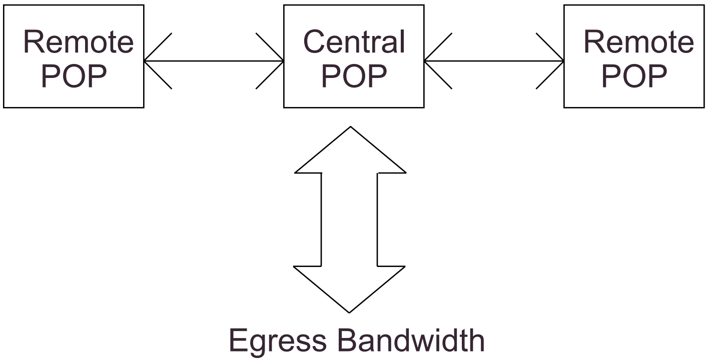

Figure 1: Small Provider Network

The characteristics of a "small" provider are centered around the following:

- One egress point for traffic
- Limited peering
- Small routers, not hierarchical
- No or limited redundancy

Of course, limiting cost is what usually drives these attributes. While multiple egress points are desirable from a reliability standpoint, redundant access to the Internet is simply beyond most small providers. The size of the paid egress is likely to be measured in T1s, not T3s or DS3s.

Peering is another area that most providers won't be able to afford, or qualify for, except for very specific situations. A small ISP may utilize private peering in facilities, but likely won't use public peering.

A small provider probably uses smaller routers, with limited port counts and functionality. For example, the Cisco 2500/2600 series routers would be used in most places except for the provider's hub (where servers might be located, for example), where a larger router like the Cisco 3600 router could be used.

There is normally no redundancy engineered in a small provider's network. The cost and complexity is beyond the small provider (and even can be too much for larger providers as well).

Larger Provider Backbone

Figure 2 illustrates how a larger provider might design its network. The boxes marked "Core" indicate the core routers/nodes of the network. Each core node usually has two or more connections to other core nodes in the providers network, forming the backbone of the provider's network. The boxes labeled "Border" indicate remote POPs that terminate customer connections. In the case of a "traditional" ISP, customer connections might be leased lines running at T1 or T3 speeds. In the case of a dialup ISP, the border routers are facilities with RAS gear serving dialup customers. In the case of a cable modem ISP, the border routers are cable head ends where traffic exits the cable provider's network and enters the Internet. (See the DOCSIS Web page for more information on this topic.) Egress can take the form of peering points or paid bandwidth. Egress points are normally on the provider's core network, where fast routers and interconnects are located.

The characteristics of a larger provider might be the following:

- Multiple egress points
- Multiple peers
- Large routers set up hierarchically
- Some redundancy

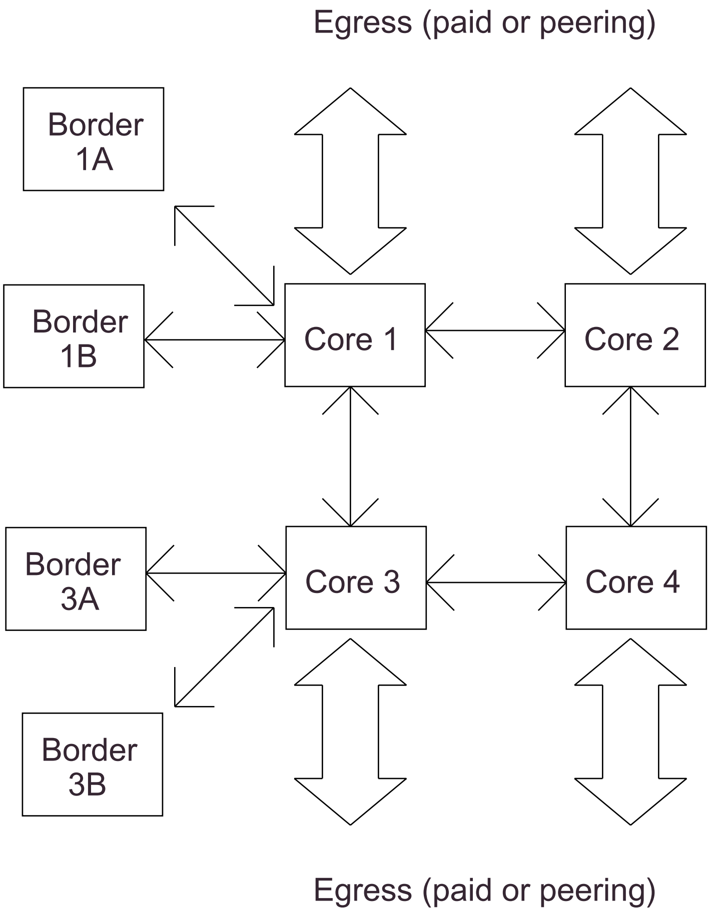

Figure 2: Large Provider Network

Cost is less of an issue for a larger provider. It will likely have multiple paid egress points for redundancy, at T3 speeds. A big service provider will have multiple peers at both private and public peering points. Large routers such as the Cisco 7000 series routers or Juniper Networks M-series will be used, set up in a border/core arrangement. The provider's backbone network will likely have some redundancy, so the loss of a single link or POP won't take down the entire network.

Most larger dialup ISPs have merged with telephone companies at this point.

The border and core router design is a common method used by larger providers to segregate their networks. Slower links terminate at the border routers, which send traffic to nearby core routers. Core routers have faster links along with peering and paid bandwidth connections. Border routers are where customer connections are normally terminated. An exception to this might be when the customer purchases high-speed bandwidth (OC3 and above) . Ordering a "fast" connection may cause the provider to terminate the connection at the faster core routers where higher speed line cards are available. This design results in faster access on and off the provider's network for the customer. As one might expect, "faster is better" in more ways than one.

The backbone protocol is likely to be ATM. While ATM is designed for voice and data networks, it is a mature technology and in wide use. Another option is IP over SONET, though this is normally utilized in OC12 and faster links. As previously mentioned, redundancy is engineered to the extent possible (within economic reason) in a larger provider. In some cases, additional markets are justified by providing additional network paths to certain POPs. Of course, the provider's service level agreements with its customers often dictate where and how certain links are provisioned for redundancy.

Small Provider Dial Network

In the case of a dialup ISP, a smaller provider will usually purchase T1 PRI line(s) in the markets they would like to reach. Often, connections from the local ILEC are purchased and terminated in the ISP's facilities in the region served by dialup. If the ISP chooses to use a CLEC, very often a large coverage area can be obtained from one POP location. For example, GNAPS serves a substantial part of Massachusetts, New Hampshire, Vermont, and Rhode Island from its location in Quincy, Massachusetts. The customer simply purchases appropriate PRI (and rackspace) and obtains coverage for the entire region and only has to site equipment in Quincy. If the same coverage was desired from Verizon, numerous PRIs would have to be ordered and facilities would be required in many Verizon POPs to obtain similar coverage.

Often, a small provider will have a small state or regional dial coverage. National/International coverage might be provided by a contract with a larger provider, if necessary. For example, both GRiC and IPASS provide national/international coverage.

Large Provider Dial Network

Large ISPs usually utilize DS3 PRI lines across the country. Using such high-bandwidth lines and associated equipment enables the provider to reduce costs. This is because DS3s and associated RAS equipment are cheaper by the port in larger capacities.

Most larger dialup ISPs have merged with telephone companies at this point. The only possible exception to this (outside of WorldCom and other really big providers) is StarNet, which has managed to stay independent. Most other providers in this space (Concentric, Split Rock, Ziplink, etc.) have merged with CLECs or gone out of business. This consolidation of the industry is a testament to the cost of PRI lines, as consolidation reduces the cost of these lines on the balance sheet.

Larger ISPs often have POPs across the country. If the ISP is associated with a ILEC/CLEC, coverage outside of the home territory will be through another ILEC/CLEC. This will ensure that the provider has a wide coverage base.

Miscellaneous Topics

The <insert word here> Service Provider

These days, there is no shortage of differing types of service providers. Most <*insert word here*> service providers are a variant of the Web-hosting provider. These include application service providers and storage service providers, among others. These types of providers typically have a backbone network as outlined in the "Larger Provider Backbone," above. An important difference would be the fact that in a dial or dedicated environment, the direction of traffic is usually inbound, whereas in a Web-hosting environment, data flow will normally be outbound.

Staff

At Ziplink, four network engineers handled the backbone network which included approximately 70,000 ports. The RAS engineering staff consisted of approximately six full-time engineers. Of course, a Network Operations Center staff was available to both groups, in order to troubleshoot and perform simple fixes.

Central vs. Distributed Network Design

Some providers may not utilize a backbone network for some or all of their POPs. This means they simply purchase egress at every location where their RAS gear is located and forego the costs and headaches associated with running one's own backbone network. The downside of such a design is that the provider has little control over these individual connections and is at the mercy of the egress providers. Costs will be higher when the provider runs its own backbone network, as cross-country network links will usually be more expensive than buying egress at each POP.

Service Level Agreements

Service Level Agreements (SLAs) are formal definitions of the type of service the provider will give to the customer. SLAs tend to vary widely from provider to provider, and customer to customer, depending upon each party's particular business needs. Of course, a provider wants the most flexible SLAs as possible, while the customer wants 100% uptime no matter what extenuating circumstances may exist.

Wholesale Dial Providers

Many substantial end-user ISPs (such as MSN and Prodigy) have a small dialup network or none at all. Instead, they purchase access from a wholesale dial provider such as Level3 and let them manage the RAS gear, ports, and associated headaches. The end-user ISP purchases access in the form of ports, time (hours), and/or users.

Security/DoS attacks

No discussion of this topic would be complete without some mention of the security issues related to providers. Service providers are often the victims of attacks, as they lease fast connections to other providers. Many attacks take the form of denial-of-service (DoS) attacks, where an attacker stops an ISP's customers from being able to access the services they purchase by filling up the ISP's network connections. Distributed DoS attacks are a variant of the DoS attack, except the attacker mounts its attacks from multiple hosts. Detecting and mitigating these sorts of attacks are the topic of much current research. DoS attacks are stopped by implementing appropriate filters on egress routers. DoS attacks do not show signs of decreasing, at least in the near future.

A good source of information for learning about the service provider business in general is the ISP Planet home page listed in the references. Until next time, please send your questions and comments to me!

References

American Registry for Internet Numbers (ARIN): *http://www.arin.net/*
Avi Freedman's Multi-Homing page: *http://www.netaxs.com/~freedman/multi.html*
Cisco Systems: *http://www.cisco.com/*
DOCSIS starting point: *http://www.docsis.org/*
GlobalNAPS: *http://www.gnaps.com/*
GRiC: *http://www.gric.com/*
IPASS: *http://www.ipass.com/*
ISP Glossary: *http://isp.webopedia.com/*
ISP Planet: *http://www.isp-planet.com/*
Juniper Networks: *http://www.juniper.net/*
Level3: *http://www.level3.com/*
MAE Services and Facilities: *http://www.mae.net/*
MSN: *http://www.msn.com/*
Prodigy: *http://www.prodigy.com/*
Qwest: *http://www.qwest.com/*
StarNet/MegaPOP: *http://www.starnetusa.net/*
Verizon: *http://www.verizon.com/*
WorldCom Business Internet Dial:
http://www1.worldcom.com/us/products/access/dial/
WorldCom: *http://www.worldcom.com/*
WorldCom's Peering Policy: *http://www.uu.net/peering/*