INTERNET EXCHANGE IN BANGLADESH AND FUTURE IMPROVEMENT OF ITS INFRASTRUCTURE

Zahurul Mohammad Tamim
Student ID: 01101044

Department of Computer Science and Engineering
Section 1.01 December 2005

INTERNET EXCHANGE IN BANGLADESH AND FUTURE IMPROVEMENT OF ITS INFRASTRUCTURE

A Thesis

Submitted to the Department of Computer Science and Engineering

of

BRAC University

by

ZAHURUL MOHAMMAD TAMIM

Student ID: 01101044

In Partial Fulfillment of the

Requirements for the Degree

of

Bachelor of Science in Computer Science

December 2005

DECLARATION

I hereby declare that this thesis is based on the results found by myself. Materials of work found by other researcher are mentioned by reference. This thesis, neither in whole nor in part, has been previously submitted for any degree.

Signature of Supervisor	Signature of Author

ACKNOWLEDGEMENT

At the end of our thesis, we acknowledge all the persons who have cooperated and worked together with us. Without their help, this report writing would have been absolutely impossible. So our earnest tanks to them for their collaboration and support.

We would like to thank Yousuf Mahbubul Islam Ph.D. our honorable Supervisor for guiding and helping to make this thesis report. This thesis gives us the chance of visiting different government and nongovernmental organizations and working with real life, which helped us to acquire clear insight about the real life systems. It was a thoroughly enjoyable experience.

Our special gratitude and admiration goes to M. Hakikur Rahman Ph.D. project coordinator, SDNP and Riton Kumar Roy, System Administrator, BDIX, SDNP. We are really grateful to them for their assistance and guidance that they have given us during our thesis. We are also thankful to Sumon Ahmed Sabir, Director, BDCOM Online Ltd, and Joint Secretary, ISP association, Bangladesh and to Dr. Javed I. Khan, BRAC University for giving us important suggestions and guidance that they have given in collecting data. We are thankful to all of the stuffs of BDIX and my classmates- for continuous support, assistance, and inspiration.

However, we are the only responsible for the errors and omissions in the report, if any.

ABSTRACT

Internet Exchange or IX is an Inter connection between ISPs in Bangladesh intended to keep local traffic local. There are over 70 ISPs in Bangladesh, which are providing Internet service to the customers. However there are only about 15 ISPs that are connected with IX. If all the ISPs are connected with IX then we can prevent the loss of huge bandwidth, minimize the cost, and promote better and faster Internet service to the customers. Our project would include study of the amount of local traffic vs. foreign traffic. It would then work out the benefits of keeping local traffic local. We would study the work of BTRC, ISPs, SDNP and BIDS. We would invite members of the ISP Association and show the results of our study. Output of the project would be individual ISP benefits, national benefits and possible recommendations for BTRC.

CONTENTS AT A GLANCE

TABLE OF CONTENTS

Chapter 4: Background of Technical Overview for IX…………………17

Chapter 6: Savings through BDIX ………………………………………….53

LIST OF FIGURE:

1. Chapter 1

Introduction

1.1. Introduction:

Internet exchange is an inter connection between different ISPs in a country intended to keep local traffic local. The goal is to interconnect the Internet Service Providers in a country so that they can exchange local traffic locally without routing through the VSAT with other country. A direct interconnection between two networks is naturally the most efficient way for those networks to exchange data. If two networks in the same area need to exchange data but are not connected to a local exchange point, they will instead pass their data through their upstream Internet providers or by VSAT.

This incurs added costs and delays, as upstream providers charge by capacity or utilization, and the traffic will often be passed through other cities before reaching its destination. The largest costs facing any ISP are the upstream capacity costs of connections. Peering arrangements at an exchange reduce the need to send IP traffic through a bandwidth of upstream provider. One single point connection to an exchange point may reduce the need for multiple connections. A physical network infrastructure is operated by a single entity with the purpose to facilitate the exchange of Internet traffic among Internet Service Providers. This exchange of national or international IP traffic on an Internet Exchange is generally known as Peering. Peering is the exchange of traffic between ISPs. In order to settle the terms to which this exchange takes place ISPs use peering agreements which often do not include an exchange of money. This helps to reduce the costs of IP Traffic in a significant way.

The operational costs for peering ISPs will be reduced through better utilization of expensive international bandwidth as well as optimal levels of domestic interconnectivity. To the extent the ISPs pass on the savings, the internet subscribers will be benefited. Additionally, subscribers will experience better Quality of Services in terms of response time and bandwidth for communicating within the country.

Most small and large ISPs (Internet Service Providers) will route traffic through this exchange to ensure security of the domestic traffic. This will allow IT applications in areas such as education, entertainment and healthcare. It helps the government to ensure that computer costs come down rapidly and Internet access devices are within the reach of all.

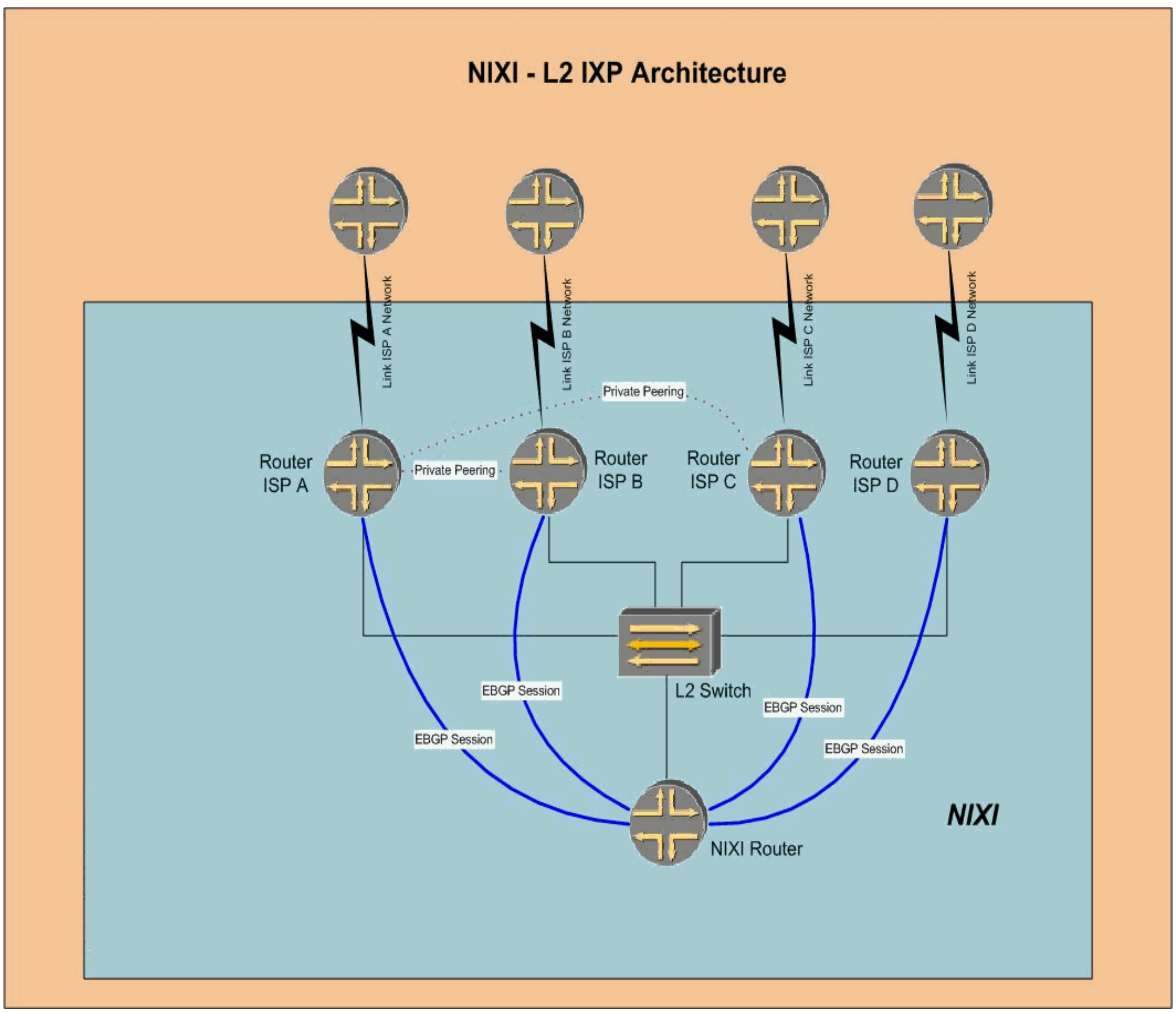


Figure1.1: National Internet Exchange, India.

This is a simple picture of National Internet Exchange, India (NIXI). In this example there are four ISPs, their down stream customers and their upstream Bandwidth provider. Here

All ISPs brings their own Router

All ISP will do EBGP with NIXI Router

All ISP will use Leased Lines to connect their network to their Router at NIXI

Fr private Peering and Transit ISPs connect directly to other ISPs without using L2 switch of NIXI.

1.2. Current situation of Internet structure in Bangladesh

The history of Internet is not that old even in developed countries. The people of Bangladesh had to remain in dark about it for a long time because of the non-availability of the service in this part of the globe. The main obstacle to start the service was to have data circuits to a suitable overseas location. However in this condition a few young talents started dial-up e-mail service and made it commercially available for public use. In late 1995 the government of Bangladesh invited applications to subscribe the Very Small Aperture Terminal (VSAT) data circuits.

On June 4, 1996 the VSAT base data circuit was commissioned for the first time in the country. Upon VSAT commissioning, internet connectivity was established and its services were made available to the public. After internet was launched, the June 1996 National polls results were made available of the world using World Wide Web. This was the first ever usage event of its kind in Bangladesh. The effort was appreciated in many corners especially among the Bangladeshis living abroad. Later many popular events were put on the web servers to make these available on the Internet. Couple of these included ‘cyclone warning’ and ‘disaster information’, cricket match, a wide variety of business information, local educational programmes, entertainment and many more.

In the early nineties, Bangladesh had access to email via dialup to Bulletin Board Systems (BBS) of a few local providers. The combined Internet users of all the email only service providers was not more then 500. Users were charged by the kilobyte, and mail was transferred from the BBS service providers to the rest of the world by International dialup using UUCP.

In June 1996, the Government allowed VSAT's to be operated in the Private Sector, provided solely by the Government owned Telephone Operator, BTTB. Only a handful of ISPs were connected within the first year. However, more liberal Government policies followed in the subsequent years which led to a rapid expansion of this industry, eventually resulting in over 180 registered ISP's by 2005. ISPs are currently regulated by the Bangladesh Telecommunication Regulatory Commission through the Bangladesh Telecommunications Act.

1.3. Growth of Local Internet

The mid 1996 introduction of Internet did not instantly create a market. At the end of year, there were only two ISPs in the country and the number of users was close to one thousand only. The year 1997 recorded a tremendous growth. The total number of ISPs was more than a dozen and the clientele growth was ten times higher than that of the previous year. Afterwards, a few new ISPs started their venture recording a proportionate growth in number of users. As of now the total number of ISPs in the country is approximately more than eighty and there are approximately 1.5 million users connected to them. In last couple of years, some wide variety of Internet services popularly known as value added services, were introduced. These include Internet fax and internet telephony services.

1.4. Sample example how IX works

This is a simple example how Internet Exchange works. Let say there are four end user or four customers. Person 1 and 2 or P1 and P2 are connected with ISP A. Person 3 and 4 or P3 and P4 are connected with ISP B. The ISPs are connected to the Global Internet through VAST. ISP A and B are not far apart one from another. Let say they are situated differently in two building which are neighbor. Both of them their individual networks. Individually they can transfer their autonomous data or internal data one customer to another internally. For that purpose they don't need the VSAT.

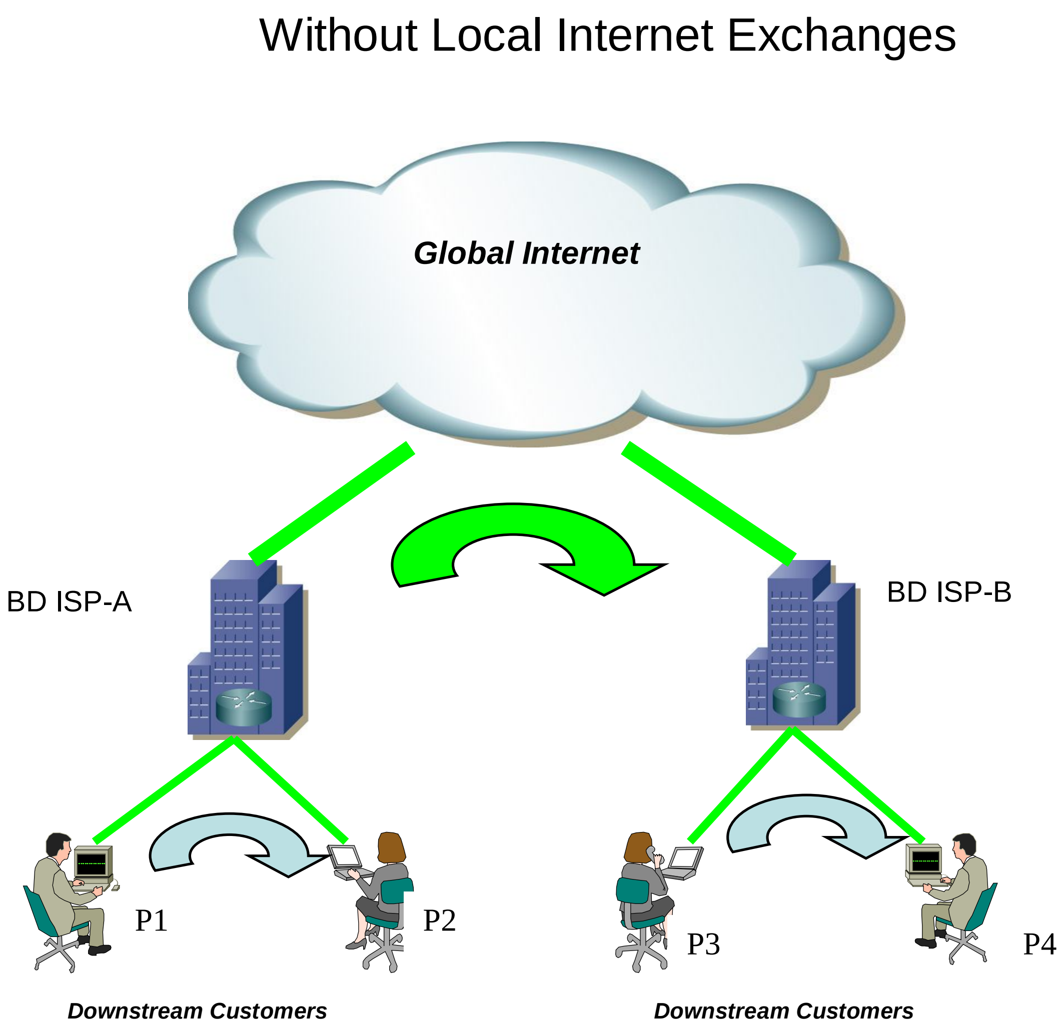


Figure 1.2: Without Internet Exchange

P1 wants to send a data to P2. When he sends the data it goes to the ISP A and ISP A sees that the data destination is its own ISP. So the ISP sends the data to person P2 without using the VSAT or other external links.

Now P1 wants to send a data to P3 who are so nearer to each other. When P1 sends the data, it goes to the ISP A. ISP A sees that it is not their internal data. The destination of the data is ISP B. ISP A and ISP B is connected with Global Internet. So ISP A sends the data to ISP B via VAST. The data passes a long route and traveled a long distance and finally reaches to ISP B. There creates several problems. Like:

- The data traveled a long distance.
- The ISP has limited amount of Bandwidth. The data occupied the Bandwidth unusually.
- The ISP has to pay a huge amount of money to the Bandwidth provider.
- The internal data goes out side the country. So it hampers the security purpose.
- The data is received by ISP B after a long period of time.
- Higher latency
- No control over traffic flows
- Complicated Traffic engineering
- Higher setup and management cost
- No open peering policy among the ISPs

These problems can be solved if the ISP A and ISP B are connected with IX or Internal Exchange. The following picture shows how the transaction can be done by IX.

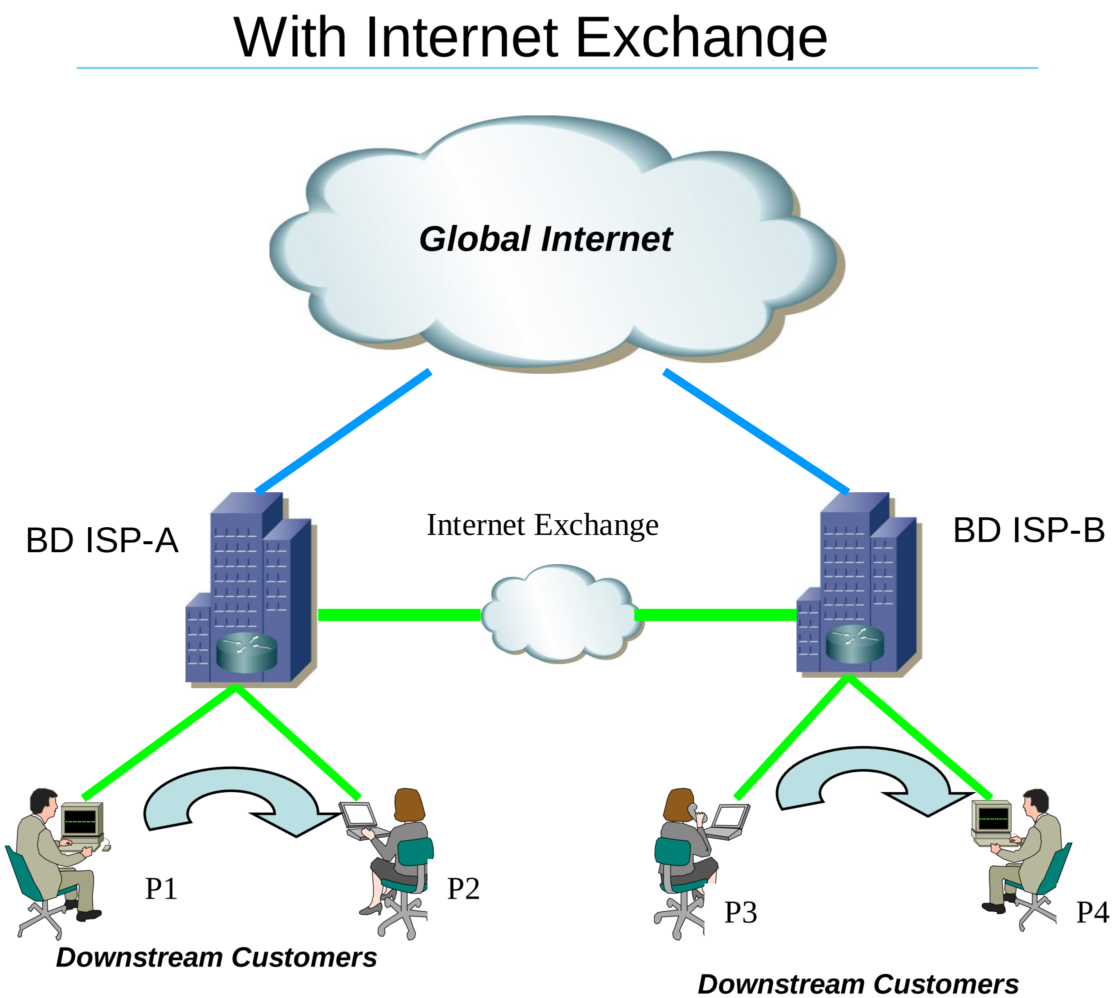


Figure 1.3: With Internet Exchange

Now if customer P1 sends a data to P3, then the data goes to ISP A. ISP A sees that the destination of the data is ISP B. So the ISP A sends the data to IX, without sending the data to Global Internet via VSAT. IX sends the data to ISP B and ISP B sends the data to the P3. This is the very simple way how data is transferred among the in country ISPs through IX.

1.5. Some Internet Exchanges or National Access Point (NAP):

• USA – Major NAPs

MAE-West California, MAE-East Wash. DC operated by WCOM

Chicago NAP operated by Ameritech

New York NAP operated by Sprint

Nap of the Americas – operated by Terremark

• China - Major NAPs

TerreNAP (Beijing),

ShangHai IX (SHIX)

• UK - Major NAPs

MaNAP,

LINX,

LoNAP,

ScotIX

• Japan - Major NAPs

JPIX,

Media Exchange (TTNet),

NSPIXP,

NSPIX2,

NSPIX3

• Korea - Major NAPs

KINX, KIX, KTIX

• Taiwan – Major NAPs

TWIX

• Singapore - Major NAPs

SingTel IX

• HKSAR –

HKIX,

ReachIX,

Pilhana

• Bangladesh – One NAP or Internet Exchange

BDIX (Bangladesh Internet Exchange)

1.6. Current situation of IX in Bangladesh

BDIX (Bangladesh Internet Exchange) has implemented Internet Exchange in Bangladesh, which is a not-for-profit partnership between Internet Service Provider Association Bangladesh and SDNP (Sustainable Development Networking Program) Bangladesh to establish a physical interconnection for its members to exchange domestic traffic through BDIX.

BDIX is a pilot project of SDNP under the Sustainable Environment Management Program (SEMP) of Ministry of Environment and Forest, funded by (UNDP) United National Development Program and executive by (BIDS) Bangladesh Institute of Development Studies.

Operation of BDIX has started in September 2004 and Bangladesh Government through BTRC has given the permission to set up the National IX in Bangladesh. Currently BDIX has 16 members.

1.7. The works ISP Association in Bangladesh

In 1998 the Internet Service Providers Association of Bangladesh was organized for the general purpose of improving business conditions of Internet Service Providers operating in Bangladesh by such activities as serving the common business interest of its Members by promoting higher business standards; disseminating technical, legal and other information to its members; operating as a trade association to benefit the community of Internet Service Providers in the Bangladesh and their customers; working for the enactment of laws and rules advancing the common business interests of its Members; and performing such other function as are customary among trade associations.

NOTE: For the list of the members of ISPAB please see Appendix A

1.8. Problems of current situation in Internet Sector

All the major ISPs of Bangladesh are providing the internet service using the satellite links VSATs (Very Small Aperture Terminal). Data circuits using such setup have a natural time delay of more than 500 milliseconds to cross at least two satellite hops and to reach the down end of the link. ISPs elsewhere use optical fiber links for internet connectivity. The optical fiber runs over the surface of the earth so the points there can be connected using much shorter length in contrast to propagation length of satellite transmissions. Moreover as optical fibers use the frequency of light so that higher band width is available from a single link than that of satellite. In this local ISPs are lagging.

All the ISPs of Bangladesh are not aware with the Internet Exchange. Few ISPs are now connected with BDIX, which provide Internet Exchange in Bangladesh. BDIX is a pilot project of Sustainable Development Network Program (SDNP) under the Sustainable Environment Management Program (SEMP) of Ministry of Environment and Forest, funded by (UNDP) United National Development Program and executive by (BIDS) Bangladesh Institute of Development Studies.

Through BDIX, the country saves a very few amount f money due to a very small amount of local traffic. The amount of local traffic is too small because all of the ISPs or major quantity of ISPs is not connected with IX or BDIX.

Because of the high cost of the satellite links, most of the ISPs have single VSAT circuits with their upstream provider either at Hong Kong or Singapore. The Bangladesh Telegraph and Telephone Board BTTB, a state monopoly controls and acts as only local VSAT operator. BTTB, as of now is the sole provider of landline phone as well. Processing of need requests for VSAT data circuit & phone lines through BTTB, often takes long time and extra effort.

2. Chapter 2

Methodology of Work

It was needed to gather some knowledge about Internet Exchange (IX) of different countries in the beginning stage of our pre-thesis which started from middle of June 2005. For this, we read various documentations, thesis papers related with this topic. At the starting stage we had to visit various ISPs in Dhaka. They talked with us about how ISPs can be benefited by IX, how our internet infrastructure can be benefited by IX etc. We continued this knowledge gathering process for two months which was finished on mid August 2005. Main part of our thesis started from early September in Bangladesh Internet Exchange (BDIX). BDIX is a pilot project of SDNP under the Sustainable Environment Management Programme (SEMP) of Ministry of Environment and Forest (MoEF), funded by United National Development Programme (UNDP) and executive by Bangladesh Institute of Development Studies (BIDS).

2.1 Plan of work

We divided our work into several phases. Phases are:

1. Gathering basic idea of IX
2. Understanding structure of IX in Bangladesh
3. Study of MRTG graph
4. Calculation the bandwidth of an IX member (daily, weakly, monthly)

Configurations:

1. BGP routing configuration
2. Server routing configuration
3. Installation procedure
4. Connection configuration of a new ISP

5. Requirements (BDIX, ISP)
6. BGP routing configuration
7. Configuration of MRTG graph

Finding out costing and saving:

1. Bandwidth calculation
2. Cost calculation
3. Comparison between IX & non IX
4. Cost comparison
5. Profit calculation

Management:

1. Cost management
2. Study and requirement of SDNP policies
3. Future planning of BDIX
4. Find out the limitations of BDIX
5. Study of a ISP which is not connected through IX

2.2 Places Visited

We visited Bangladesh Telecommunication Regulatory Commission (BTRC) and Internet Service Provider Association Bangladesh (ISPABA) to know their policy about ISPs in Bangladesh and about their concern for BDIX. After policy part we visited a number of ISPs and talked with their system administrators about how an ISP can be benefited from BDIX. Some of its were already connected and some got interested after talking with us. ISPs, we visited were- BDCOM, BD MAIL, BOL, SDNP, CONNECT BD, DRUTI, SPECTRA SOLUTIONS LTD. (Ctg), PROSHIKA ONLINE.

NOTE: See Appendix A to see the places where we visited for data collection

2.3 Work In BDIX:

After visiting all these ISPs our main thesis work started in BDIX which includes learning technical stuffs for understanding how IX works. For this, we had to learn regular routing along with advanced routing like Border Gateway Protocol (BGP) routing. This learning process was on hand training. In BDIX we got opportunity to work not with router simulator but with real router and switch. We had the opportunity to visit the BDIX control room if necessary at any time for technical purpose.

2.4 Seminars attended:

During our thesis we have attended three seminars related to Internet Exchange. The first seminar was held at BRAC University by Dr. Javed I. Khan, on BGP routing and Internet Exchange. The second and third seminar was held at BIDS. SDNP arranged two days long inaugural ceremony and the installation of the Root Server –F on 9th and 10th December 2005 at the conference room of BIDS. All of the seminars helped us to earn a huge knowledge about Internet Exchange, its current situation in the world and situation of IX in Bangladesh.

3. Chapter 3

Overview of Bangladesh Internet Exchange

For the first time ever in Bangladesh an effort to locally manage local internet traffic was made through the introduction of Bangladesh Internet Exchange (BDIX), a not-for-profit partnership between Sustainable Development Networking Programme (SDNP) Bangladesh. BDIX establishes a physical interconnection between ISPs to exchange internet traffic.

Bangladesh still depends on satellite bandwidth sharing, which causes any transfer of files or mails between local ISPs to go through the channels of foreign servers, taking up more bandwidth and resulting in a huge expenditure of foreign currency. The local server at BDIX will facilitate faster net access for any local traffic by providing faster connection between BDIX member ISPs.

The bandwidth project was taken up in August 2004 by the Sustainable Environment Management Programme (SEMP) of the Ministry of Environment and Forest (MoEF), which is funded by UNDP and executed by Bangladesh Institute of Development Studies (Bids) and Internet Service Providers of Bangladesh.

Using Cisco routers to channel the local traffic, BDIX expects to manage all local internet traffic locally and thus enhance user compatibility and lower bandwidth cost. At present, fifteen local ISPs are connected through the BDIX network. Most of these ISPs are using DSL or Radio-link connections to amend the quality of usage.

There are the lists of ISPs that are connected with BDIX:

SL/NO	ISP NAME	IX-IP	AS NUMBER	CONNECTION DATE
01	**SDNP**	198.32.167.17	9825	27-08-2004
02	**BDCOM ONLINE LTD**	198.32.167.18	24122	27-08-2004
03	**BANGLADESH ONLINE (BOL)**	198.32.167.20	9230	06-09-2004
04	**INFORMATION SERVICE NETWORKLIT.(ISN)**	198.32.167.19	9832	04-09-2004
05	**LINK3 TECHNOLOGIES LTD.**	198.32.167.21	23688	27-09-2004
06	**ACCESS TEL**	198.32.167.22	17469	04-11-2004
07	**DAFFODIL ONLINE**	198.32.167.23	24308	28-11-2004
08	**AFTAB IT**	198.32.167.24	24307	28-11-2004
09	**RANKS-ITT LTD**	198.32.167.25	23991	28-11-2004
10	**BIJOY ONLINE LTD**	198.32.167.26	65002	29-11-2004
11	**PROSHIKA**	198.32.167.27	3758	03-03-2005
12	**BRAC BDMail**	198.32.167.28	24342	03-04-2005
13	**BTTB**	198.32.167.30	17494	30-11-2005

Table 3.1: Members of BDIX

The exchange server membership at present is free of cost. Thus BDIX does not plan on acquiring any finance from its members as its project is well established and runs on full-comprehension. But they are now taking only Tk5000 per month from the member ISPs for the maintenance and other purposes.

The expected time frame on such matters may well range from two to five months or a little more, while BDIX acquire further troubleshooting skills on the exchange server.

The creators of the exchange server plan to establish a national data center within a short period. The net exchange server will then include web server, local newspaper server, chat server and net filter server. This data center will also include a server for academic and research institutes of the country, corporate networks, financial sectors, and possibly all the local ISPs of Bangladesh. The submarine cable is expected to enter the country through Cox's Bazaar, which is one reason why BDIX has chosen Cox's Bazaar to be one of the major routing nodes of the exchange server.

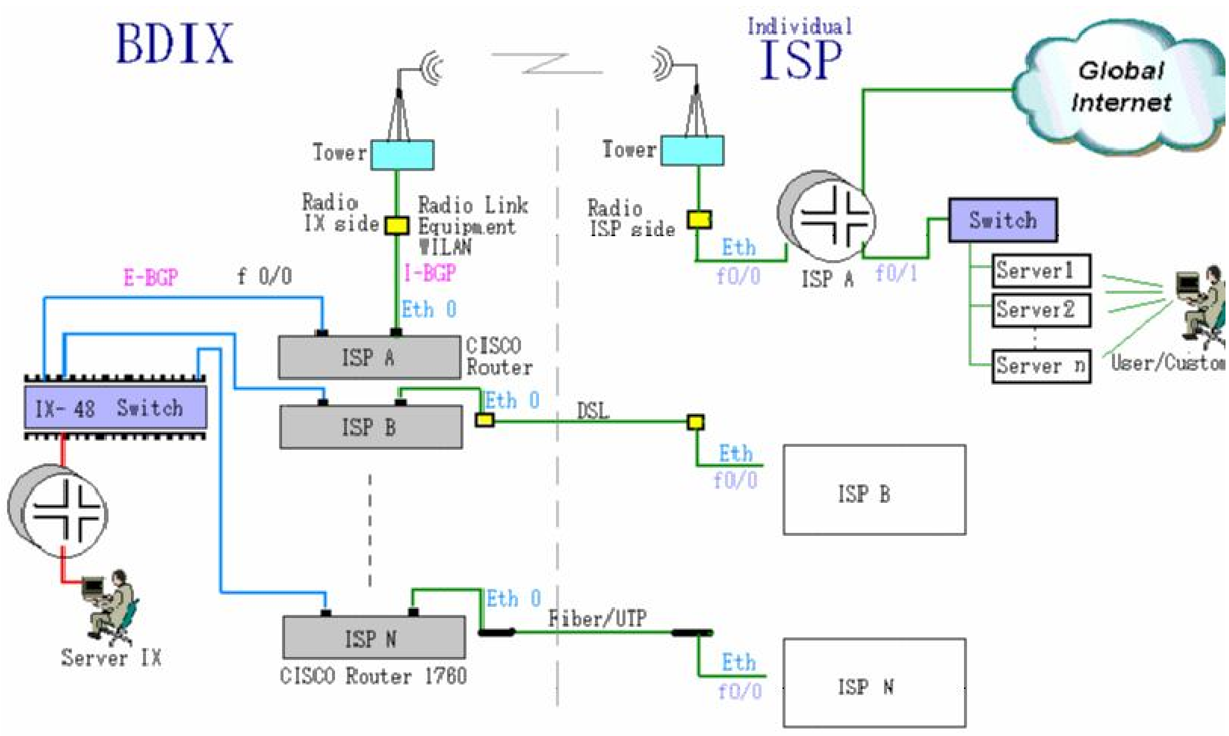


Figure 3.1: Picture of how BDIX works

4. Chapter: 4

Background of Technical Overview for IX

Some technical concepts are required for installation of Internet Exchange and maintenance it. There are some protocols and routing configuration as well for implementing Internet Exchange. All of them are describe bellow.

4.1 Bandwidth:

For digital signals and by extension from the above, the word bandwidth is often used to mean the amount of data that can be transferred through a digital connection in a given time period (in other words, the connection's bit rate). The more correct term for this, however, is throughput. It is usually measured in bits or bytes per second. A huge amount of bandwidth now a day measured in kilo bytes (Kb) or even in mega bytes.

The measure of bandwidth in internet exchange is an important thing. The cost and time calculation of an ISP depends on how much bandwidth that is using. In Bangladesh ISP usually use 2/2 kbps to 24/24 Mbps bandwidth.

4.2 TCP

TCP is a connection-oriented transport protocol that sends data as an unstructured stream of bytes. By using sequence numbers and acknowledgment messages, TCP can provide a sending node with delivery information about packets transmitted to a destination node. Where data has been lost in transit from source to destination, TCP can retransmit the data until either a timeout condition is reached or until successful delivery has been achieved. TCP can also recognize duplicate messages and will discard them appropriately. If the sending computer is transmitting too fast for the receiving computer, TCP can employ flow control mechanisms to slow data transfer. TCP can also communicate delivery information to the upper-layer protocols and applications it supports.

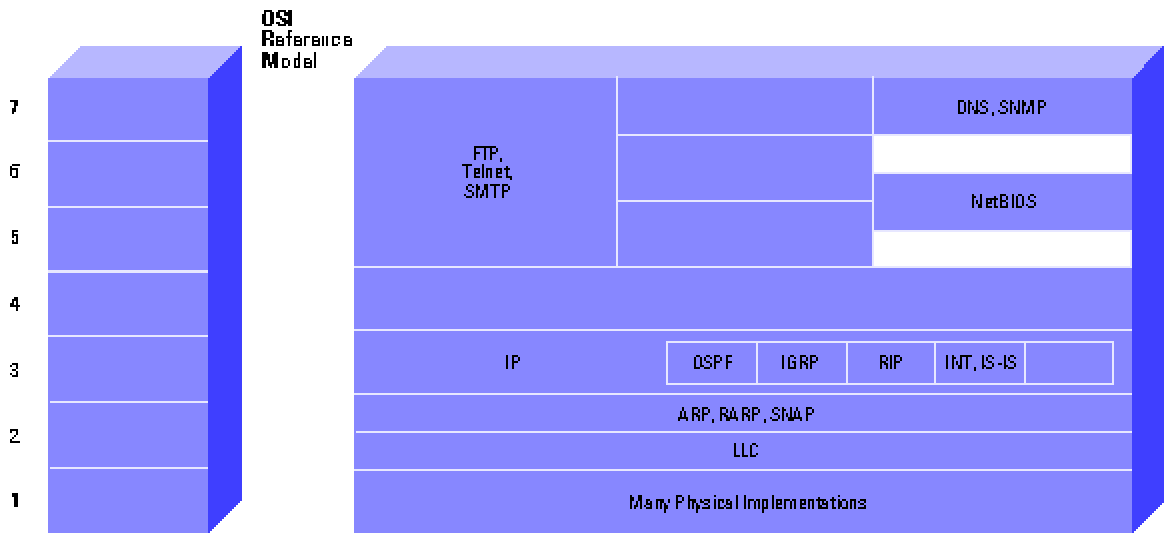


Figure 4.1: Relationship of the Internet Protocol Suite to the OSI Reference Model

4.3 IP

The Internet Protocol (IP) is a data-oriented protocol used by source and destination hosts for communicating data across a packet-switched internetwork. Data in an IP internetwork are sent in blocks referred to as packets or datagrams (the terms are basically synonymous in IP). In particular, in IP no setup of "path" is needed before a host tries to send packets to a host it has previously not communicated with.

The Internet Protocol provides an unreliable datagram service (also called best effort); i.e. it makes almost no guarantees about the packet. The packet may arrive damaged, it may be out of order (compared to other packets sent between the same hosts), it may be duplicated, or it may be dropped entirely. If an application needs reliability, it is provided by other means, typically by upper level protocols transported on top of IP.

Packet switches, or internetwork routers, forward IP datagrams across interconnected layer 2 networks. The lack of any delivery guarantees means that the design of packet switches is made much simpler. (Note that if the network does drop, reorder or otherwise damage a lot of packets, the performance seen by the user will be poor, so most network elements do try hard to not do these things - hence the best effort term. However, an occasional error will produce no noticeable effect.)

IP is the common element found in today's public Internet. The current and most popular network layer protocol in use today is IPv4; this version of the protocol is assigned version 4. IPv6 is the proposed successor to IPv4; the Internet is slowly running out of addresses, and IPv6 has 128-bit source and destination addresses, providing more addresses than IPv4's 32 bits. Versions 0 through 3 were either reserved or unused. Version 5 was used for an experimental stream protocol. Other version numbers have been assigned, usually for experimental protocols, but have not been widely used.

4.4 IP addressing and routing

Perhaps the most complex aspects of IP are addressing and routing. Addressing refers to how end hosts are assigned IP addresses and how subnetworks of IP host addresses are divided and grouped together. IP routing is performed by all hosts, but most importantly by internetwork routers, which typically use either interior gateway protocols (IGPs) or external gateway protocols (EGPs) to help make IP datagram forwarding decisions across IP connected networks.

4.5 Routing

Routing is the means of discovering paths in networks; typically communication networks, to provide paths through the network fabric along which information can be sent. In computer networks, the data is split up into packets, each handled individually. Circuit-based networks, such as the voice telephone network, also perform routing, to find paths for telephone calls.Routing is usually directed by routing tables, which maintain a record of the best routes to various network destination locations.

Routing directs forwarding, the passing of logically addressed packets from their local sub network toward their ultimate destination. In large networks, packets may pass through many intermediary destinations before reaching their destination. Routing and forwarding both occur at layer 3 of the OSI seven-layer model.

4.5.1 Dynamic routing basics

If a designated path is no longer available, the existing nodes must determine an alternate route to use to get their data to its destination. This is usually accomplished through the use of a routing protocol using one of two broad classes of routing algorithms; distance vector algorithms and link state algorithms, which together account for nearly every routing algorithm in use on the Internet.

4.5.2 Distance vector algorithms

Distance vector algorithms use the Bellman-Ford algorithm. When this used, the link between each node in the network is assigned a number, the cost. Nodes will send information from point A to point B via the path that results in the lowest total cost (i.e. the sum of the costs of the links between the nodes used).

The algorithm is very simple. When a node first starts, it only knows of its immediate neighbours, and the direct cost to them. (This information, the list of destinations, the total cost to each, and the next hop to send data to to get there, is the routing table, or distance table.) Each node, on a regular basis, sends to each neighbour its own current idea of the total cost to get to all the destinations it knows of. The neighbouring node(s) examine this information, and compare it to what they already 'know'; anything which represents an improvement on what they already have, they insert in their own routing table. Over time, all the nodes in the network will discover the best next hop for all destinations, and the best total cost.

When one of the nodes involved goes down, those nodes which used it as their next hop for certain destinations discard those entries, and create a new routing table. This is then passed to all adjacent nodes, which then repeat the process. Eventually all the nodes are updated, and will then discover new paths to all the destinations which are still reachable.

4.5.3 Link state algorithms

When link state algorithms are used, the fundamental data each node uses is a map of the network, in the form of a graph. To produce this, each node floods the entire network with information about what other nodes it is connected to, and each node then independently assembles this information into a map. Using this map, each router then independently determines the best route from it to every other node.

The algorithm used to do this, Dijkstra's algorithm, does this by building another data structure, a tree, with the node itself as the root, and containing every other node in the network. It starts with a tree containing only itself. Then, one at a time, from the set of nodes which have not yet been added to the tree, it adds the node which has the lowest cost to reach an adjacent node which is already in the tree. This continues until every node has been added to the tree.

This tree is then used to construct the routing table, giving the best next hop, etc, to get from the node itself to any other node in the network.

Depending on the relationship of the router relative to other autonomous systems, various classes of routing protocols exist:

- Ad hoc network routing protocols appear in networks with no or little infrastructure. For a list of a couple of the proposed protocols, see the Ad hoc protocol list

- Interior Gateway Protocols (IGPs) exchange routing information within a single autonomous system. Common examples include:
  - IGRP (Interior Gateway Routing Protocol)
  - EIGRP (Enhanced Interior Gateway Routing Protocol)
  - OSPF (Open Shortest Path First)
  - RIP (Routing Information Protocol)
  - IS-IS (Intermediate System to Intermediate System)
- Exterior Gateway Protocols (EGPs) route between separate autonomous systems. EGPs include:
  - EGP (the original Exterior Gateway Protocol used to connect to the former Internet backbone network -- now obsolete)
  - BGP (Border Gateway Protocol: current version, BGPv4, was adopted around 1995)

4.6 Autonomous System

An autonomous system (AS) is a group of networks under a single administrative control, which could be a computer, a division within a company, or a group of companies. An Interior Gateway Protocol refers to a routing protocol that handles routing within a single autonomous system. IGPs include RIP, IGRP, OFPS, AND IS-IS. Today there is only active EGP: the Broader Gateway Protocol (BGP). BGP is used to route traffic across internal backbone between different autonomous systems.

4.7 AS Number

Autonomous System Numbers (ASNs) are globally unique numbers that are used to identify autonomous systems and which enable an AS to exchange exterior routing information between neighboring ASes. An AS is a connected group of IP networks that adhere to a single and clearly defined routing policy.

4.7.1 How to get AS Number

When an ISP wants an AS number, then it will have to apply for AS number to APNIC. APNIC will then check some specifications of the ISP to justify their application. After a successful clarification, APNIC gives them an AS number.

4.8 Multi Router Traffic Grapher

4.8.1 What is MRTG

The Multi Router Traffic Grapher (MRTG) is a tool to monitor the traffic load on network links. MRTG generates HTML pages containing graphical images which provide a live visual representation of the traffic. It was originally developed by Tobias Oetiker and Dave Rand to monitor router traffic, but has developed into a tool that can create graphs and statistics for almost anything.

4.8.2 Feature of MRTG

Some of the important features of MRTG is listed bellow

- MRTG measures 2 values (I for Input, O for Output) per target.
- MRTG gets its data via an SNMP agent, or through the output of a command line.
- It typically collects data every five minutes (it can be configured to collect data less frequently).
- It creates an HTML page per target that features 4 graphs (GIF or PNG images).
- The results are plotted vs. time into day, week, month and year graphs, with the I plotted as a full green area, and the O as a blue line.
- The program automatically scales the Y axis of the graphs to show the most detail.

- MRTG adds calculated Max, Average and Current values for both I and O to the target's HTML page.
- MRTG can also send warning emails if targets have values above a certain threshold.

4.8.3Plotting MRTG Graph

The MRTG CFG maker program creates configuration files for network interfaces only, simultaneously tracking two object IDs: the NIC's input and output data statistics. The MRTG program then uses these configuration files to determine the type of data to record in its data directory. The index maker program also uses this information to create the overview, or Summary View Web page for the MIB OIDs you're monitoring.

This Summary View page shows daily statistics only. It has to click on the Summary View graphs to get the Detailed View page behind it with the daily, weekly, monthly, and annual graphs. Some of the parameters in the configuration file refer to the detailed view; others refer to the Summary View.

If it is needed to monitor any other pairs of OIDs, you have to manually create the configuration files, because CFG maker isn't aware of any OIDs other than those related to a NIC. The MRTG and index maker program can be fed individual OIDs from a customized configuration file and will function as expected if you edit the file correctly.

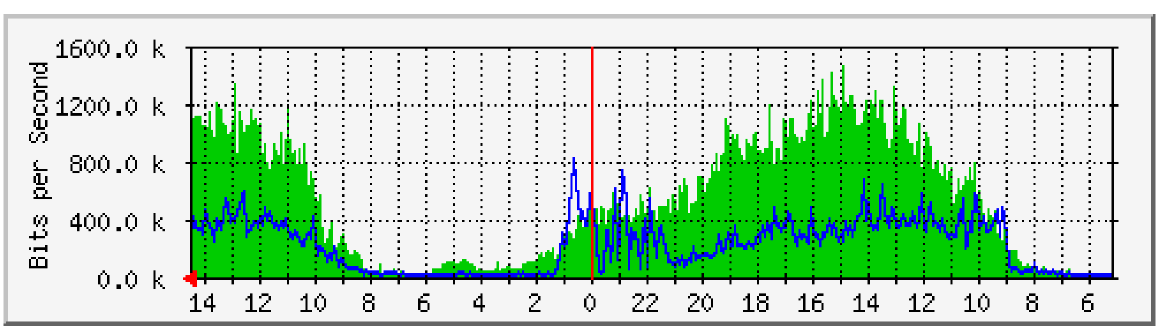


Figure 4.2: This is a simple MRTG graph of an ISP of BDIX:

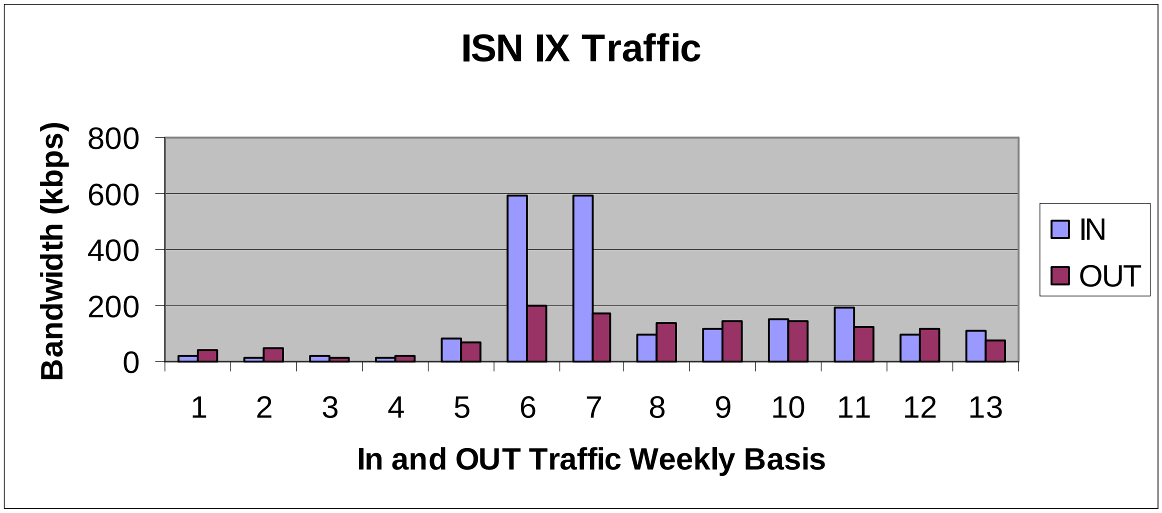


Figure 4.3: This is a simple graph of an ISP of BDIX:

4.9 RIP

The Routing Information Protocol (RIP) is one of the most commonly used Interior Gateway Protocol (IGP) routing protocols on internal networks (and to a lesser extent, networks connected to the Internet), which helps routers dynamically adapt to changes of network connections by communicating information about which networks each router can reach and how far away those networks are. Although RIP is still actively used, it is generally considered to have been obsoleted by routing protocols such as OSPF and IS-IS, although EIGRP, a somewhat more capable protocol in the same basic family as RIP (destination-vector routing protocols).

4.9.1Technical Details of RIP:

RIP is a distance-vector routing protocol, which employs the hop count as a routing metric. The maximum number of hops allowed with RIP is 15. Each RIP router transmits full updates every 30 seconds by default, generating large amounts of network traffic in lower bandwidth networks. It runs above the network layer of the Internet protocol suite, using UDP port 520 to carry its data. A mechanism called split horizon with limited poison reverse is used to avoid routing loops. Routers of some brands also use a holddown mechanism known as heuristics, whose usefulness is arguable and is not a part of the standard protocol.

In many current networking environments RIP would not be the first choice for routing as its convergence times and scalability are poor compared to OSPF or IS-IS, and the hop limit severely limits the size of network it can be used in. On the other hand, it is easier to configure.

4.10 Loop back Interface

Loopback address is a special IP number (127.0.0.1) that is designated for the software loopback interface of a machine. The loopback interface has no hardware associated with it, and it is not physically connected to a network. A loopback interface is a virtual interface that resides on a router. It is not connected to any other device. Loopback interface are very useful because they will never down, unless the entire router goes down. This helps in managing routers because there will always be at least one active interface on the routers, the loopback interface.

4.10.1Configuring Loopback Interfaces:

To create a loopback interface, it is need to do enter configuration mode for the interface

Router(config)#interface loopback (number)

The only option on this command is to specify a number between 0 and 2.147,483,647. Cisco IOS software gives plenty of loopback interfaces, if anyone wants to use all of them. When entering this command, Cisco IOS software automatically the loopback interface, place into interface configuration mode and remove the interface from shutdown mode. When that is complete, only need to assign an IP address to the interface. The criteria for the IP address of the loopback interface are as follows:

Create loopback interfaces on all routers using IP addess 192.168.1.1/24

Router(config)# interface loopback 0
Router(config)#
Line protocol on Interface Loopback0 , change state to up.
Router(config)#ip address 192.168.1.1 255.255.255.0

The router automatically removes the loopback from shutdown state; shows the console message indicating that the interface is up. So it is not required to type “no shutdown”.

4.10.2 Open Shortest Path First (OSPF)

4.10.3 What is OSPF

OSPF (Open Shortest Path First) is a router protocol used within larger autonomous system networks in preference to the Routing Information Protocol (RIP), an older routing protocol that is installed in many of today's corporate networks. Like RIP, OSPF is designated by the Internet Engineering Task Force (IETF) as one of several Interior Gateway Protocols (IGPs).

Using OSPF, a host that obtains a change to a routing table or detects a change in the network immediately multicasts the information to all other hosts in the network so that all will have the same routing table information. Unlike the RIP in which the entire routing table is sent, the host using OSPF sends only the part that has changed. With RIP, the routing table is sent to a neighbor host every 30 seconds. OSPF multicasts the updated information only when a change has taken place.

Rather than simply counting the number of hops, OSPF bases its path descriptions on "link states" that take into account additional network information. OSPF also lets the user assign cost metrics to a given host router so that some paths are given preference. OSPF supports a variable network subnet mask so that a network can be subdivided. RIP is supported within OSPF for router-to-end station communication. Since many networks using RIP are already in use, router manufacturers tend to include RIP support within a router designed primarily for OSPF.

4.10.4 OSPF Protocol Overview

Open Shortest Path First (OSPF) is a recent entry into the Internet interior routing scene. Sanctioned by the IETF, it is intended to become Internet's preferred interior routing protocol. OSPF is a link-state routing protocol with a complex set of options and features. Not all of these features are available on all implementations, but some of its advantages are:

- **Scalability**:
    - OSPF is specifically designed to operate with larger networks. It does not impose a hop-count restriction and permits its domain to be subdivided for easier management.
- **Full subnetting support:**
    - OSPF can fully support subnetting, including VLSM and non-contiguous subnets.
- **Hello packets:**
    - OSPF uses small "hello" packets to verify link operation without transferring large tables. In stable networks, large updates occur only once every 30 minutes.

- **TOS routing**:
    - OSPF can route packets by different criterion based on their Type of Service (TOS) field. For example, file transfers could be routed over a satellite link while terminal I/O could avoid such high delays. This requires cooperative applications on the end systems.
- **Tagged routes**:
    - Routes can be tagged with arbitrary values, easing interoperation with EGPs, which can tag OSPF routes with AS numbers.

OSPF has some disadvantages as well. Chief among them are its complexity and its demands on memory and computation. Although link-state protocols are not difficult to understand, OSPF muddles the picture with plenty of options and features.

OSPF divides its routing domain into *areas*. Area 0, the backbone, is required. This divides interior routing into two levels. If traffic must travel between two areas, the packets are first routed to the backbone. This may cause non-optimal routes, since inter area routing is not done until the packet reaches the backbone. Once there, it is routed to the destination area, which is then responsible for final delivery. This layering permits addresses to be consolidated by area, reducing the size of the link state databases. Small networks can operate with a single OSPF area, which must be area 0.

OSPF divides networks into several classes, including point-to-point, multiaccess, and non-broadcast multiaccess. A serial link connecting two routers together would be a point-to-point link, while an Ethernet or Token Ring segment would be a multiaccess link. A Frame Relay or X.25 cloud would be classified as non-broadcast multiaccess.

Multiaccess networks (like Ethernet) use a designated router (DR) to avoid the problem of each router forming a link with every other router on a Ethernet, resulting in a N^2 explosion in the number of links. Instead, the DR manages all the link state advertisements for the Ethernet. Selecting the DR requires an election process, during which a Backup Designated Router (BDR) is also selected. OSPF provides a *priority* feature to help the network engineer influence the choice of DR and BDR, but in practice this is difficult. Link layer multicasting is also used, if available, to avoid broadcasts and better target routing updates.

Non-broadcast multiaccess networks (like X.25) also use the designated router concept, but since broadcasts (and presumably multicasts) are not supported, the identity of neighboring routers must be specified manually. A DR on such a network without a complete list of neighbors will cause a loss of connectivity, even though the network is otherwise functional.

OSPF's primary means of verifying continuing operation of the network is via its Hello Protocol. Every OSPF speaker sends small hello packets out each of its interfaces every ten seconds. It is through receipt of these packets that OSPF neighbors initially learn of each other's existence. Hello packets are not forwarded or recorded in the OSPF database, but if none are received from a particular neighbor for forty seconds, that neighbor is marked down. LSAs are then generated marking links through a down router as down. The hello timer values can be configured, though they must be consistent across all routers on a network segment.

Link state advertisements also age. The originating router read vertices an LSA after it has remained unchanged for thirty minutes. If an LSA ages to more than an hour, it is flushed from the databases. These timer values are called *architectural constants* by the RFC.

OSPFs various timers interact as follows:

- If a link goes down for twenty seconds, then comes back up, OSPF doesn't notice.
- If a link flaps constantly, but at least one of every four Hello packets make it across, OSPF doesn't notice.
- If a link goes down for anywhere from a minute to half an hour, OSPF floods an LSA when it goes down, and another LSA when it comes back up.
- If a link stays down for more than half an hour, LSAs originated by remote routers (that have become unreachable) begin to age out. When the link comes back up, all these LSAs will be re-flooded.
- If a link is down for more than an hour, any LSAs originated by remote routers will have aged out and been flushed. When the link comes back up, it will be as if it were brand new.

4.11 Border Gateway Protocol

4.11.1 BGP-4 Protocol Overview:

Border Gateway Protocol Version 4 (BGP-4) is current exterior routing protocol used for the global Internet. BGP is essentially a distance-vector algorithm, but with several added twists. BGP uses TCP as its transport protocol, on port 179. On connection start, BGP peers exchange complete copies of their routing tables, which can be quite large. However, only changes (deltas) are then exchanged, which makes long running BGP sessions more efficient than shorter ones.

BGP's basic unit of routing information is the BGP path, a route to a certain set of CIDR prefixes. Paths are tagged with various path attributes, of which the most important are AS_PATH and NEXT_HOP.

One of BGP-4's most important functions is loop detection at the Autonomous System level, using the AS_PATH attribute, a list of Autonomous Systems being used for data transport. The syntax of this attribute is made more complex by its need to support path aggregation, when multiple paths are collapsed into one to simplify further route advertisements. A simplified view of AS_PATH is that it is the list of Autonomous Systems that a route goes through to reach its destination. Loops are detected and avoided by checking for your own AS number in AS_PATH's received from neighboring Autonomous Systems.

Every time a BGP path advertisement crosses an Autonomous System boundary, the NEXT_HOP attribute is changed to the IP address of the boundary router. Conversely, as a BGP path advertisement is passed among BGP speakers in the same AS, the NEXT_HOP attribute is left untouched. Consequently, BGP's NEXT_HOP is always the IP address of the first router in the next autonomous system, even though this may actually be several hops away. The AS's interior routing protocol is responsible for computing an interior route to reach the BGP NEXT_HOP. This leads to the distinction between Internal BGP (IBGP) sessions (between routers in the same AS) and External BGP (EBGP) sessions (between routers in different AS's). NEXT_HOPs are only changed across EBGP sessions, but left intact across IBGP sessions.

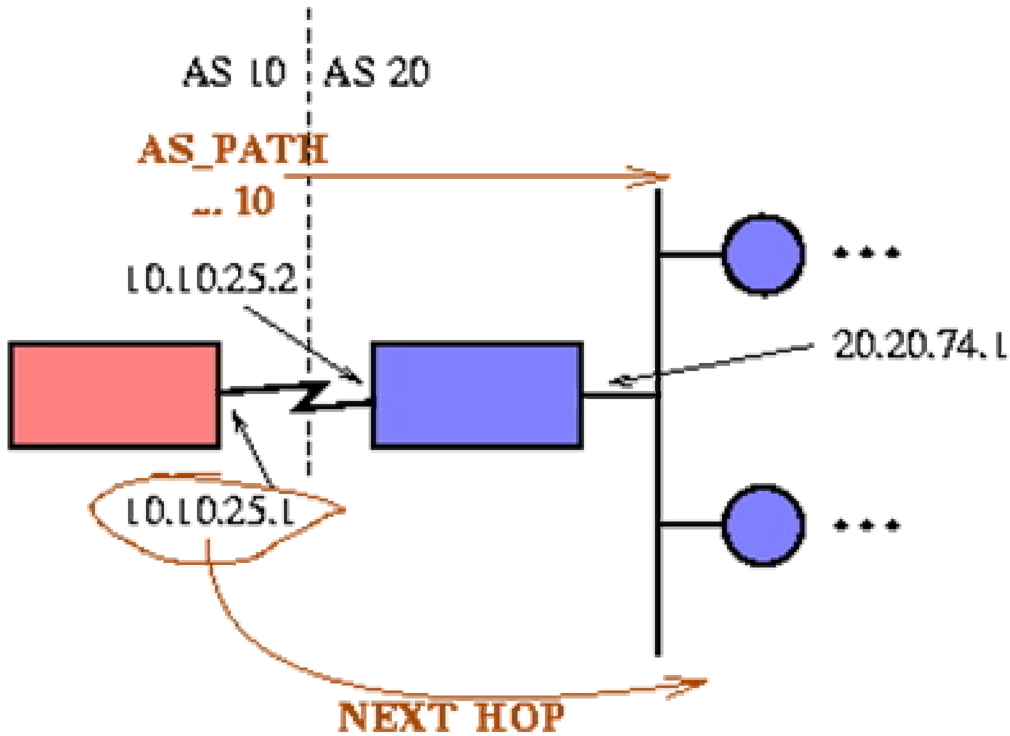


Figure 4.4: Different AS routing

The two most important consequences of this design are the need for interior routing protocols to reach one hop beyond the AS boundary, and for BGP sessions to be fully meshed within an AS. Since the NEXT_HOP contains the IP address of a router interface in the *next* autonomous system, and this IP address is used to perform routing, the interior routing protocol must be able to route to this address. This means that interior routing tables must include entries one hop beyond the AS boundary. Furthermore, since BGP does not relay routing traffic from one Interior BGP session to another, (only from an Exterior BGP session to an IBGP session or another EBGP session) BGP speakers must be fully meshed. When a BGP routing update is received from a neighboring AS, it must be relayed directly to all other BGP speakers in the AS. Do not expect to relay BGP paths from one router, through another, to a third, all within the same AS.

4.12 Peering and Transit:

4.12.1 Peering

Peering is the exchange of traffic between ISPs. In order to settle the terms to which this exchange takes place ISPs use peering agreements which often do not include an exchange of money. This helps to reduce the costs of IP Traffic in a significant way. One of the largest costs facing any ISP nowadays is the upstream capacity costs of connections. Peering arrangements at an exchange reduce the need to send IP traffic through a bandwidth upstream provider. One single connection to an exchange point (such as BD-IX) may reduce the need for multiple connections. Each member might have a different peering policy, and this policy may differ depending on the ISP that is negotiating with them.

4.12.2 Transit

Transit is the business relationship whereby one ISP provides (usually sells) access to all destinations in its routing table. Another way it is a carrying traffic across network to network, usually for a fee. Traffic and prefixes originating from one AS is carrying across an intermediate AS to reach their destination AS.

4.12.2.1 ISP Transit Issues:

There are some Transit issues among the ISP for Transits. None can violate those issues. The issues are:

I) Only announced default to one ISP BGP customers unless they need more prefixes.

II) Only except the prefixes which one ISPs customer is entitled to originate.

III) If an ISP customer has not told to the ISP that he is providing transit, doesn’t except anything else.

4.12.3 Exchange point:

Common interact location where several ASes exchange routing information and traffic.

4.12.4 Simple Configuration: ISP Transit Providers:

This is a simple example how routers are configured in different autonomous system. There are four routers; Router A, Router B, Router C and Router D. Router A is in different autonomous system which AS number is AS 107. Router D is in different autonomous system which AS number is AS 109. The transit between A and D is given by another autonomous system which AS number is 108 and Router C and B are in this autonomous system.

AS 107 and AS 109 are stub/customer ASes of AS 108. They may have their own peering with other ASes. They must have minimal routing table desired, minimum complexity required.

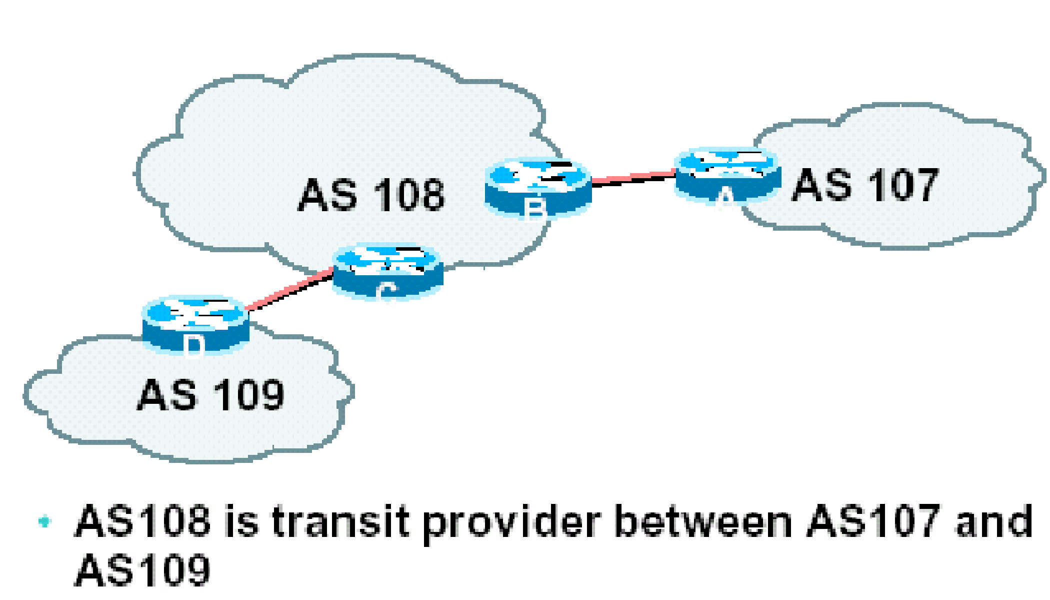


Figure 4.5: Transit providing between A and D

<u>Now the configuration of Router A, B, C, and D are following:</u>

- **Router A Configuration**

```
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor 222.222.10.2 remote-as 108
 neighbor 222.222.10.2 prefix-list upstream out
 neighbor 222.222.10.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list upstream permit 221.10.0.0/19
!
ip route 221.10.0.0 255.255.224.0 null0
```

- **Router B Configuration**

```
router bgp 108
 neighbor 222.222.10.1 remote-as 107
 neighbor 222.222.10.1 default-originate
 neighbor 222.222.10.1 prefix-list Customer107 in
 neighbor 222.222.10.1 prefix-list default out
!
ip prefix-list Customer107 permit 221.10.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

- **Router B announces default to Router A, only accepts customer /19**

- **Router C Configuration**

```
router bgp 108
 neighbor 222.222.20.1 remote-as 109
 neighbor 222.222.20.1 default-originate
 neighbor 222.222.20.1 prefix-list Customer109 in
 neighbor 222.222.20.1 prefix-list default out
!
ip prefix-list Customer109 permit 219.0.0.0/19
ip prefix-list default permit 0.0.0.0/0
```

- **Router C announces default to Router D, only accepts customer /19**

- **Router D Configuration**

```
router bgp 109
 network 219.0.0.0 mask 255.255.224.0
 neighbor 222.222.20.2 remote-as 108
 neighbor 222.222.20.2 prefix-list upstream out
 neighbor 222.222.20.2 prefix-list default in
!
ip prefix-list default permit 0.0.0.0/0
ip prefix-list upstream permit 219.0.0.0/19
!
ip route 219.0.0.0 255.255.224.0 null0
```

4.12.5 Simple Configuration: ISP Exchange Point:

There is a simple example of create an ISP Exchange point (IX) having 6 routers; Router A, Router B, Router C, Router D, Router E, and Router F; where Router A's AS number is 107, where Router B's AS number is 108, Router C's AS number is 109, Router D's AS number is 110, Router E's AS number is 111 and Router F's AS number is 112. All the routers are connected with a single exchange point or single Internet Exchange point. They all have layer-2 Ethernet switch. Each ISP peers with the others.

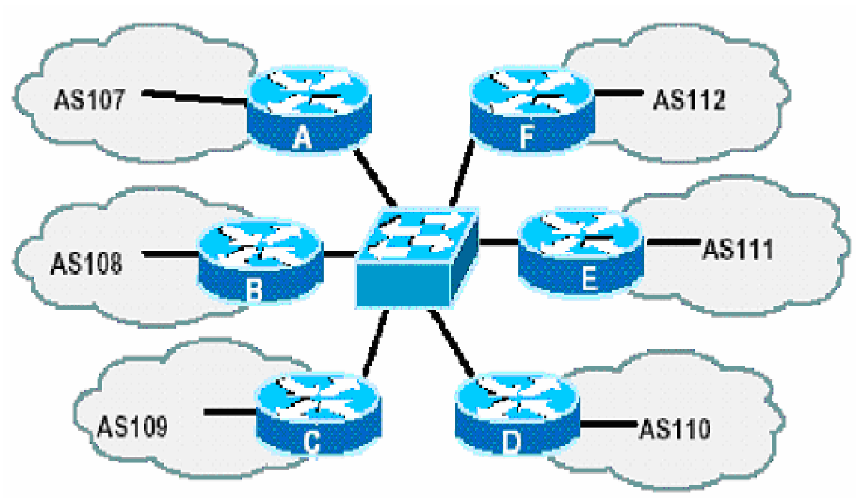


Figure 4.6: Exchange point of 6 ISPs.

Router A Configuration:

```
interface fastethernet 0/0
 description Exchange Point LAN
 ip address 220.5.10.2 mask 255.255.255.224
 ip verify unicast reverse-path
 no ip directed-broadcast
 no ip proxy-arp
 no ip redirects
!
router bgp 107
 network 221.10.0.0 mask 255.255.224.0
 neighbor ixp-peers peer-group
 neighbor ixp-peers soft-reconfiguration in
 neighbor ixp-peers prefix-list myprefixes out
neighbor 220.5.10.2 remote-as 108
neighbor 222.5.10.2 peer-group ixp-peers
neighbor 222.5.10.2 prefix-list peer108 in
neighbor 220.5.10.3 remote-as 109
neighbor 222.5.10.3 peer-group ixp-peers
neighbor 222.5.10.3 prefix-list peer109 in
neighbor 220.5.10.4 remote-as 110
neighbor 222.5.10.4 peer-group ixp-peers
neighbor 222.5.10.4 prefix-list peer110 in
neighbor 220.5.10.5 remote-as 111
neighbor 222.5.10.5 peer-group ixp-peers
neighbor 222.5.10.5 prefix-list peer111 in
neighbor 220.5.10.6 remote-as 112
neighbor 222.5.10.6 peer-group ixp-peers
neighbor 222.5.10.6 prefix-list peer112 in
```

```
!
ip route 221.10.0.0 255.255.224.0 null0
!
ip prefix-list myprefixes permit 221.10.0.0/19
ip prefix-list peer108 permit 222.0.0.0/19
ip prefix-list peer109 permit 222.30.0.0/19
ip prefix-list peer110 permit 222.12.0.0/19
ip prefix-list peer111 permit 222.18.128.0/19
ip prefix-list peer112 permit 222.1.32.0/19
!
```

Router B,C,D, E and F Configuration:

Configurations of other routers are as same as configuration of Router A.

Ethernet port Configuration:

Use ip verify unicast reverse-path.

AS 107 needs to know all the prefixes its peers are announcing, new prefixes require the prefix-list to be update. There is an alternative solution. That is: use the Internet Routing Registry to build prefix list and use AS path filters.

4.12.6 Technical Structure of BDIX:

For connecting with IX, an ISP should fulfill some criteria to ISPAB and BDIX. Then they are permitted to connect with BDIX. Several equipments are also required for BDIX side and for ISP side. Some of the equipments are provided by BDIX and some of them should be provide by the ISP on their own.

Technical Specifications of Equipment (BDIX)

Server (DELL)	02
UPS 2200 VA (APC)	02
Workstation (DELL)	05

Managed Switch (CISCO)	01
Rack (RAXXES)	01
Cisco router (1760)	16
Radio Link equipment (WILAN)	16
Managed Switch	01
Cisco router	02
Sun Server	01

Technical Specifications of Equipment (BDIX)

Cisco Router	01
Ethernet port	01
Linux PC Server	01
Radio/DSL modem	01

ISP connection mode	**Mode**
SDNP	UTP
BDCOM	UTP
ISN	DSL
BOL	DSL
LINK3	Radio
AccessTel	DSL
Daffodil	DSL
Aftab IT	Radio
Rangs IT	Radio
Bijoy Online	Radio
ProshikaNet	Radio Link
Agni System Ltd.	Radio Link
Brac Bdmail Ltd	Radio Link
Metro Link	Fiber Link
Ektoo Ltd.	Radio Link
Akij Online Ltd	Radio Link

ISPs are basically connected with Internet Exchange (IX) server through radio link, DSL modem, Fiber Optic/UTP etc. The responsibility of an ISP themselves is to carry a link to the IX switch in order to get connected. The rest part which is installation and router configurations are responsibilities of IX technicians.

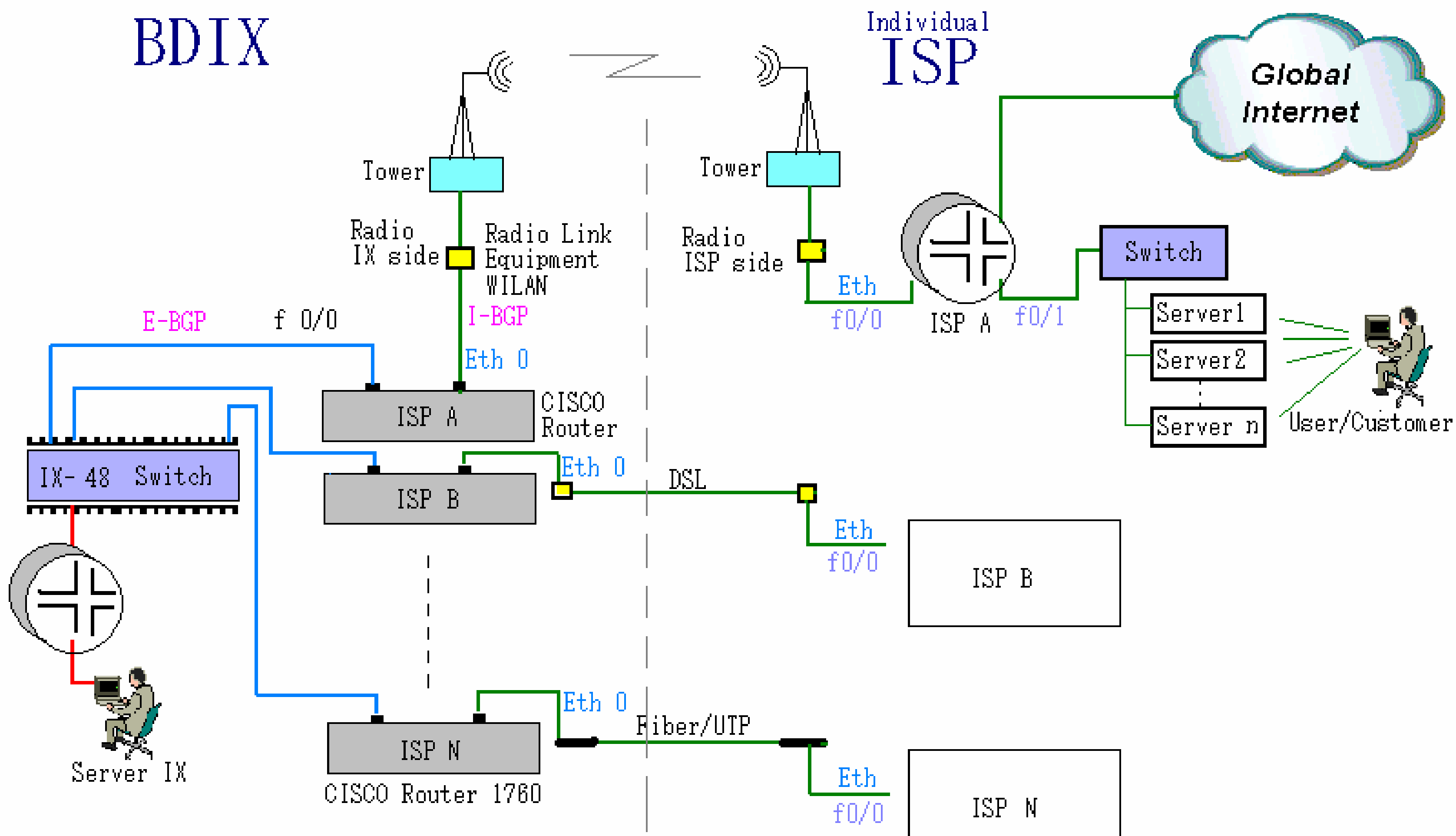


Figure 4.7 : How BDIX connected with ISPs

Lets see the connection break down from an individual ISP's point of view where Radio Link is been used to get connected with IX. In an ISP, they use VSAT, Router, Switch for managing Internet connection. So, a Switch is in connection with Router through Fast Ethernet0/1 port. Another Ethernet0/0 port is connected with Radio tower for caring link up to IX part. Similarly, there is a Radio tower also in IX part for establishing the link. The link gets connected with CISCO Router in IX part through Ethernet0/0 port. One important thing is, the link from ISP Router to CISCO router, F0/0 port is in same Autonomous System (AS). So, during Border Gateway Protocol (BGP) session, Internal BGP (IBGP) protocol will be applied. The CISCO Router is connected with IX Switch through F0/1 port. Here, External BGP session (EBGP) will be used cause the IX
Switch has different AS no. Finally, the ISP is connected to IX and able to transfer local data locally without using unnecessary Bandwidth.

Using DSL Modem:

In this pattern, there are two DSL Modems in ISP and IX part. A P2P connection is established between two DSL Modems. ISP Modem is connected with ISP Router throughF0/0 port. IX Modem is connected with CISCO Router through its F0/0 port. This is in same AS.F0/1 CISCO port is connected with IX Switch as this system is in different AS.

Chapter 5:

Benefits of Internet Exchange

An Internet Exchange can bring revolutionary changes in Internet infrastructure of a country. It can serve various purposes and many can be beneficiaries from this. Benefits of IX include keeping local traffic local, giving local content and serving as national data center etc. Basically, these are the benefits which are directly linked with IX. Besides these, there are other indirect benefits of IX. For example, e-commerce can start with help of IX. Many unemployed people will be able to start making money of it. Unemployment rate will decrease and many social problems will be solved which usually take place due to depression created from unemployment.

5 Benefits

Benefits of Internet Exchange are listed below.

5.1 Open Peering Policy Peers

A peering agreement is a bilateral business and technical arrangement in which two connectivity providers agree to accept traffic from one another (and from one another's customers, and their customers' customers).

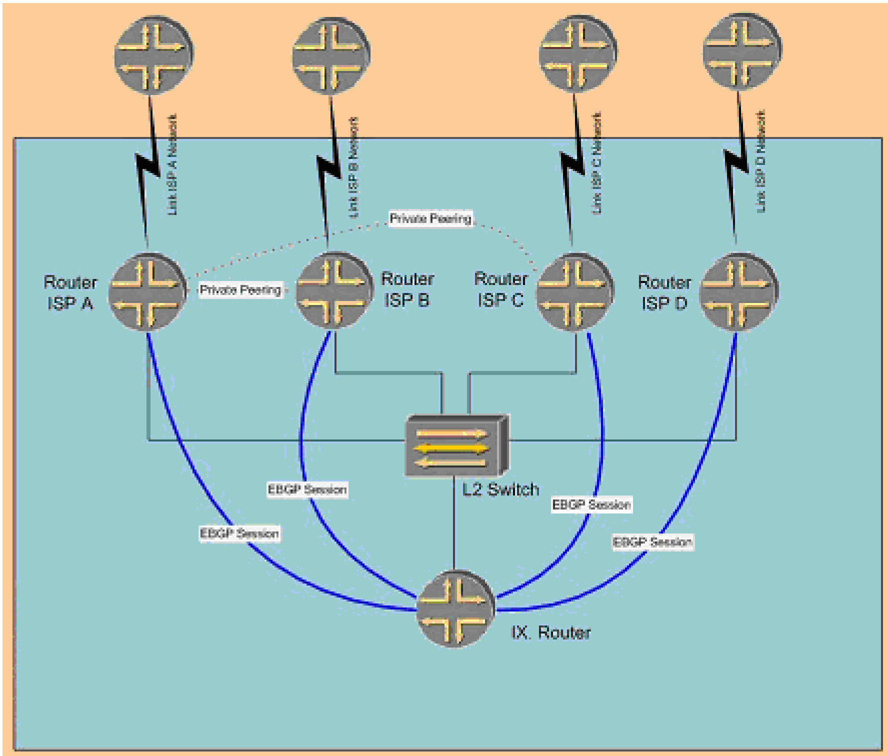


Figure 5.1: Peering among ISPs

In a peering agreement, there is no obligation for the peer to carry traffic to third parties. There are no cash payments involved – rather, it is more like barter, with each ISP is trading direct connectivity to its customers in exchange for connectivity to the ISP's customers. So, by peering through IX it is possible for ISPs to reduce the traffic cost.

5.2 Reduce Transit Cost

A transit agreement is also a bilateral business and technical arrangement, but one in which the transit provider agrees to carry traffic from the customer to third parties, and from third parties to the customer. The customer ISP is thus regarded as an end point for the traffic; the transit provider serves as a conduit to the global Internet. Generally, the transit provider will undertake to carry traffic not only to/from its other customers but to/from every destination on the Internet.

Transit agreements typically involve a defined price for access to the entire Internet.

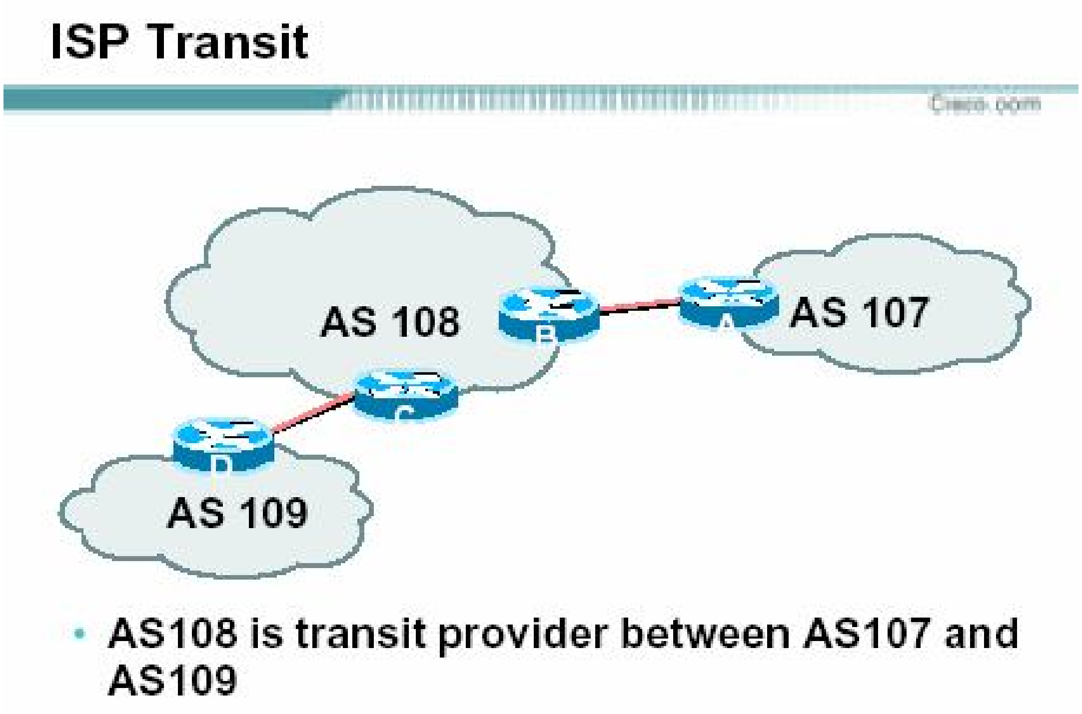


Figure 5.2: Transit

If anyone who is a customer of very small size ISP sends an e-mail from HRC to BOL then his ISP is bound to sign a transit agreement with either a local ISP or upstream provider where this small ISP will have to carry 100% cost of inbound and outbound bandwidth. Solution is joining in IX.

5.3 Better Control over Traffic Flows

Without IX in any country, every ISP uses its own gateway to get connected with Internet. So there is no provision to have control over passing packets. Cyber criminals take this opportunity for maintaining criminal activities. If all ISPs are connected with IX then its possible for IX technicians to sort out every packets passing through them and stopping the crime.

5.4 Lower Latency

In performance measurement of any network, latency is a very important metric. People get very impatient if the latency is very high of packet transferring. Suppose, a BOL user sends a mail to BDCOM customer, the packet firstly goes to the immediate upstream provider of BOL through VSAT, and then it goes to server which is in USA or Canada. Lastly it returns back to Bangladesh through another path to BDCOM customer. This whole time consuming and costly procedure wouldn't take place if BOL and BDCOM are connected with IX. They would exchange their local traffic very easily through IX in very short time than the other procedure. So, through IX it is possible to achieve faster packet transferring with local ISPs.

5.5 Traffic Engineering

Traffic Engineering is basically used for Load Balancing between different paths among hops. It is a bit tough task to load balance for ISPs because measuring load in different paths requires huge statistical data and proper utilization of appropriate algorithms. But joining IX can give them a solution. If an ISP join IX then it is possible for them to manage the Load Balancing stuffs because IX can provide them the packet monitoring facility. So with this opportunity ISPs can make their idle routes busy and busy routes a bit less busy as well.

5.6 Saving International Bandwidth

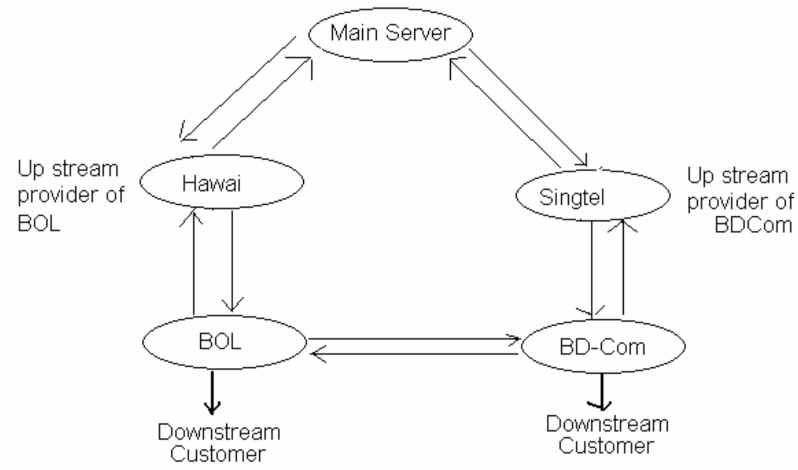


Figure 5.3: Route of Bandwidth

This is one of the most important facilities which IX gives to ISPs. International Bandwidth costs huge money for our ISPs. This huge cost indirectly comes to our customers also. If a BOL customer sends an email to a BDCOM customer then this packet firstly goes to upstream provider of BOL which is in Singapore and then goes to Canada or USA main server and then comes back to Bangladesh to BDCOM customer. This whole process can be stopped if BOL and BDCOM are connected to each other through IX. If ISPs are connected to IX then it is very possible for them to save more or less 320 Kbps average bandwidth monthly per ISP which costs 100000 taka according their whole Bandwidth consumption. It means an ISP can sell bandwidth of another extra 100000 taka. In this point they can give some incentives to customers. They can lower the Internet usage price for customers which will encourage those who are still out of this Information Highway.

5.7 Lowering Pressure on Core

Presently World Wide Web (WWW) is maintained by USA. Highest number of root servers is located in USA which is six in number and the upstream providers are also from that region. If a BOL customer sends an email to a BDCOM customer then this packet firstly goes to upstream provider of BOL which is in Singapore and then goes to. Canada or USA main server and then comes back to Bangladesh to BDCOM customer. This whole process can be stopped if BOL and BDCOM are connected to each other through IX. If so, then it will be possible to pass packets between them in fastest possible time and to reduce pressure from the Internet core as well. So, if every ISP and every country can keep their local traffic local through then the Internet core will be less pressurized.

5.8 Media Content provision

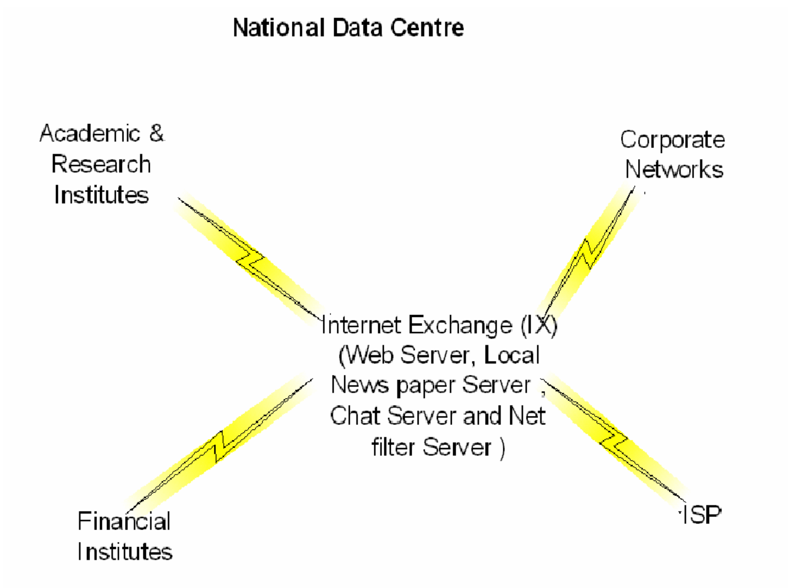


Figure 5.4: Media hosting in IX

Present situation is our local web sites are hosted in different countries. Our local job sites, daily news paper sites which customers frequently hit are hosted out side country. If these famous sites can be hosted in IX (IX is capable of doing so) then customers will be able to access these sites almost in no time. By doing so, our internal hit will increase in dramatic way. In consequence of this, our local traffic will increase which means the more local traffic the more bandwidth savings. Now a days people love to chat. Famous local chat servers are hosted outside. We can host these chat servers in IX so that people can get connected with server very fast. One problem is getting disconnected from server for low bandwidth. In this case, this sort of problem will not take place. Educational institutions can get connected with each others through IX and can use it as central data center.

5.9 Enable e-business Transaction

At this moment, e-commerce hasn't started vastly in our country. Some companies are operating but not in large scope. If someone starts business through IX he will be benefited by getting his site hosted in IX very cheaply, customers will get faster access to his site. So there is no problem in infrastructure level. Young entrepreneurs will be encouraged to introduce new business plans based upon IX. A huge business opportunity will open from which Government can also be benefited. The more general people will earn, the more Govt. will earn. Foreign investors will come to invest in a proved system like IX. We can have a look upon India. Their economy has experienced a boom due to spread of Internet through three IXs for which huge foreign investors have invested over there. Right now, almost no world famous IT companies are left who haven't open their office in India. IT field is one of the biggest fields for Indian Govt. from which they earn a huge amount of money. They can be a model for our company.

5.10 Benefits of BDIX

Through BDIX all the members of IX in Bangladesh are getting the following benefits:

a) Reduce transit cost
b) Lower latency
c) Better control over traffic flows

d) Setting up local Internet exchanges for intra-country or intra-city traffic is very

e) important for faster Internet development within that country or city

f) Enable e-business transaction (acquisition, large customer requirement)

g) Traffic engineering (asymetric routing)

h) Minimize cost of setup, management and operation

i) Reduces the loading to the Internet cores

j) Easy Content provision (media, services)

k) Open peering policy peers

l) Saving of International Bandwidth

Chapter 6

Savings through BDIX

6.1 Individual ISP Bandwidth savings:

Example (a): *BRAC BDMail Network (BBN)*

BRAC BDMail is one of the most reputed Internet service providers in Bangladesh. Here is some important notes for BBN:

- Internet Exchange IP (IX IP):
    - 198.32.167.28
- Autonomous System number:
    - 24342
- Internal Bandwidth:
    - Approximately 12 Mbps duplex
- Bandwidth provider:
    - Hazison (Hongkong)
    - Hawai (USA)
- Connected via:
    - VSAT
- Bandwidth charge:
    - 32000 taka for 64 Kbps duplex
- Date of connection with BDIX:
    - 03-04-2005
- Connection cost for IX:
    - Free
- Maximum speed:
    - 100.00 Mbps

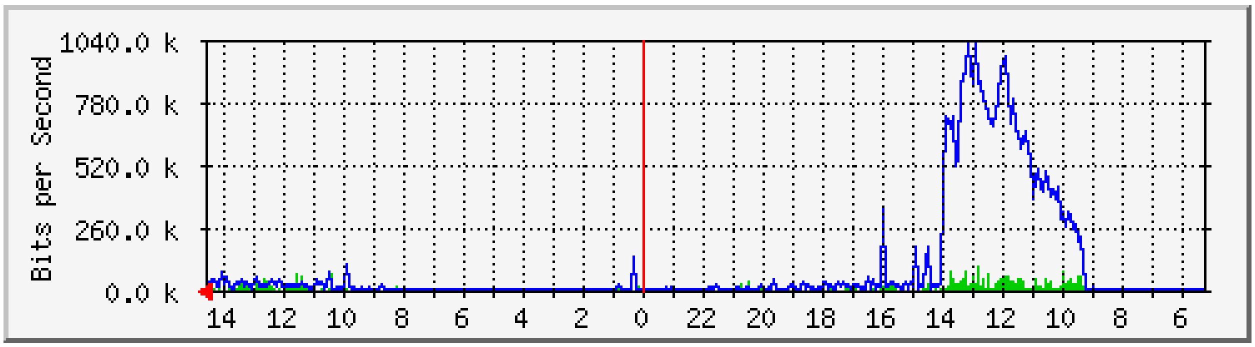
MRTG Graph of BBN:

The statistics were last updated **Sunday, 11 December 2005 at 14:35**, at which time **'switch.dac.sdnbd.org'** had been up for **3 days, 5:11:10**

Green lines are Downstream (In) Bandwidth

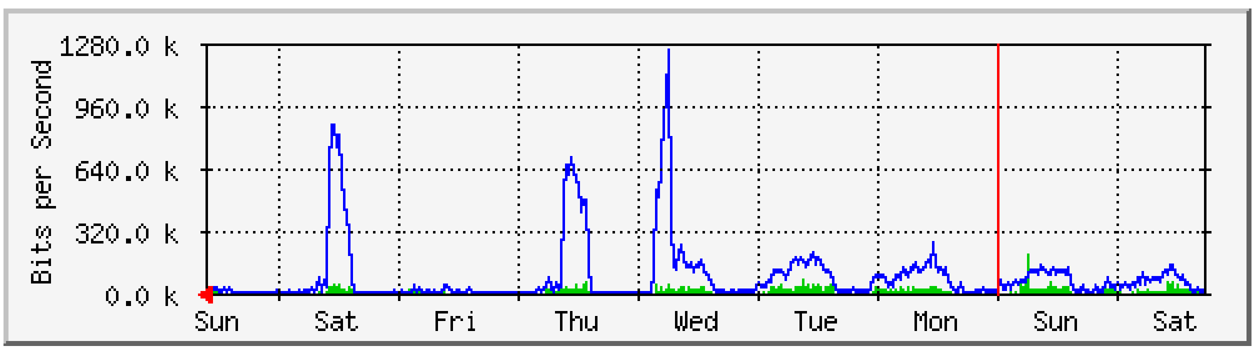
Blue lines are Upstream (Out) Bandwidth

Figure 6.1: MRTG Graph of BBN: 'Daily' Graph (5 Minute Average)



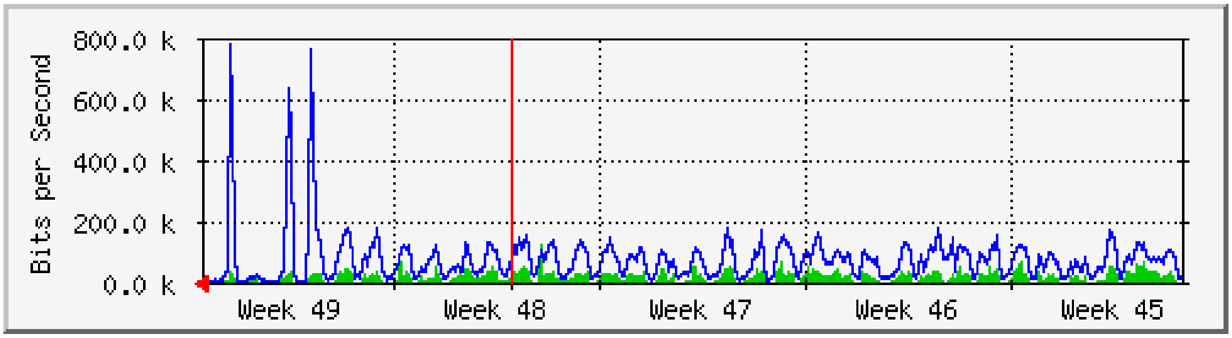
Max In:	108.0 kb/s (0.1%)	Average In:	13.2 kb/s (0.0%)	Current In:	14.6 kb/s (0.0%)
Max Out:	1026.3 kb/s (1.0%)	Average Out:	99.6 kb/s (0.1%)	Current Out:	32.5 kb/s (0.0%)

Figure 6.2: MRTG Graph of BBN: `Weekly' Graph (30 Minute Average)



Max In:	206.8 kb/s (0.2%)	Average In:	18.6 kb/s (0.0%)	Current In:	13.1 kb/s (0.0%)
Max Out:	1245.8 kb/s (1.2%)	Average Out:	88.9 kb/s (0.1%)	Current Out:	42.5 kb/s (0.0%)

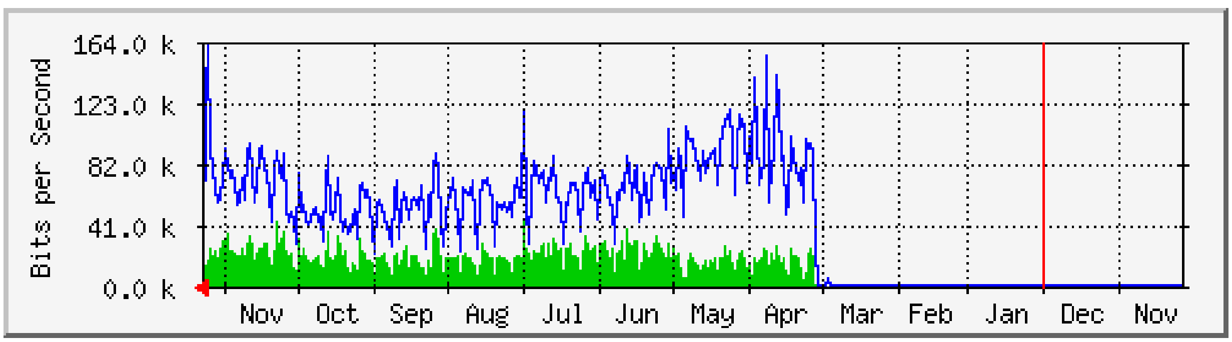
Figure 6.3: MRTG Graph of BBN: Monthly' Graph (2 Hour Average)



Max In: 129.1 kb/s (0.1%) Average In: 25.2 kb/s (0.0%) Current In: 23.1 kb/s (0.0%)

Max Out: 780.1 kb/s (0.8%) Average Out: 78.2 kb/s (0.1%) Current Out: 25.6 kb/s (0.0%)

Figure 6.4: MRTG Graph of BBN: `Yearly' Graph (1 Day Average)



Max In: 44.6 kb/s (0.0%) Average In: 22.0 kb/s (0.0%) Current In: 7448.0 b/s (0.0%)

Max Out: 163.7 kb/s (0.2%) Average Out: 68.1 kb/s (0.1%) Current Out: 11.8 kb/s (0.0%)

Monitoring the graphs:

- Highest Bandwidth:
    - 1280 kbps approximately through IX (On average)
- Cost of 1280 Kbps is 640000.00 Taka
- Extra sell is 640000.00 Taka per month!

Example (b): *Bangladesh OnLine (BOL)*

BOL is another most reputed Internet service provider in Bangladesh. Here are some important notes for BOL:

- Internet Exchange IP (IX IP):
    - 198.32.167.20
- Autonomous System number:
    - 9230
- Internal Bandwidth:
    - Approximately 12 Mbps duplex
- Bandwidth provider:
    - Hazison (Hongkong)
- Connected via:
    - VSAT
- Bandwidth charge:
    - 28500 taka for 64 Kbps duplex
- Date of connection with BDIX:
    - 06-09-2004
- Connection cost for IX:
    - Free
- Maximum speed:
    - 100.00 Mbps

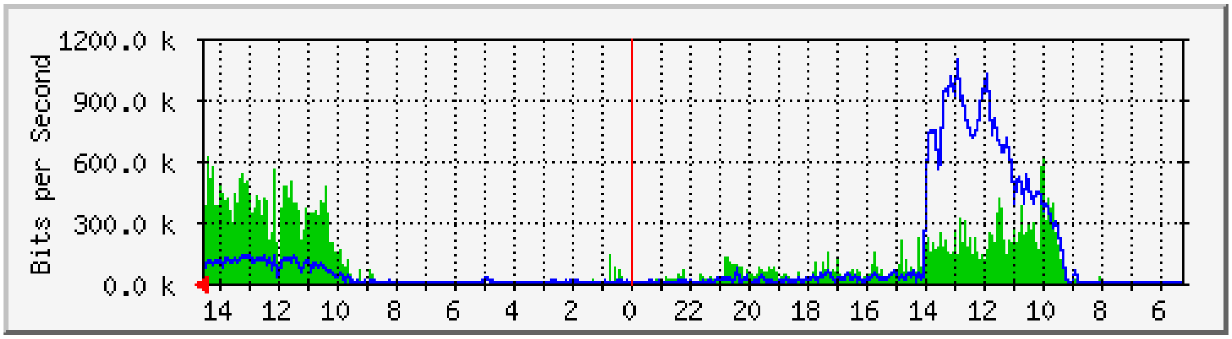
MRTG Graph of BOL:

The statistics were last updated **Sunday, 11 December 2005 at 14:35**, at which time **'switch.dac.sdnbd.org'** had been up for **3 days, 5:11:10**.

Green lines are Downstream (In) Bandwidth

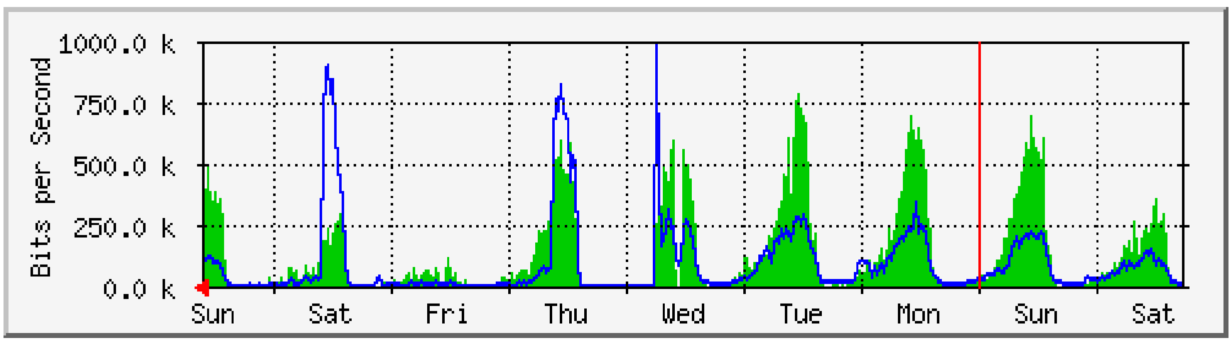
Blue lines are Upstream (Out) Bandwidth

Figure 6.5: MRTG Graph of BOL: 'Daily' Graph (5 Minute Average)



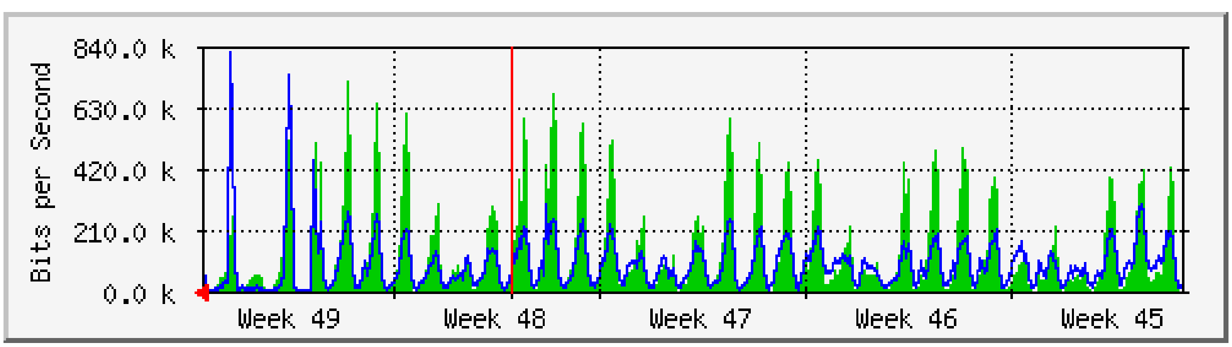
Max In:	628.6 kb/s (0.6%)	Average In:	112.8 kb/s (0.1%)	Current In:	559.4 kb/s (0.6%)
Max Out:	1093.1 kb/s (1.1%)	Average Out:	117.7 kb/s (0.1%)	Current Out:	119.5 kb/s (0.1%)

Figure 6.6: MRTG Graph of BOL: `Weekly' Graph (30 Minute Average



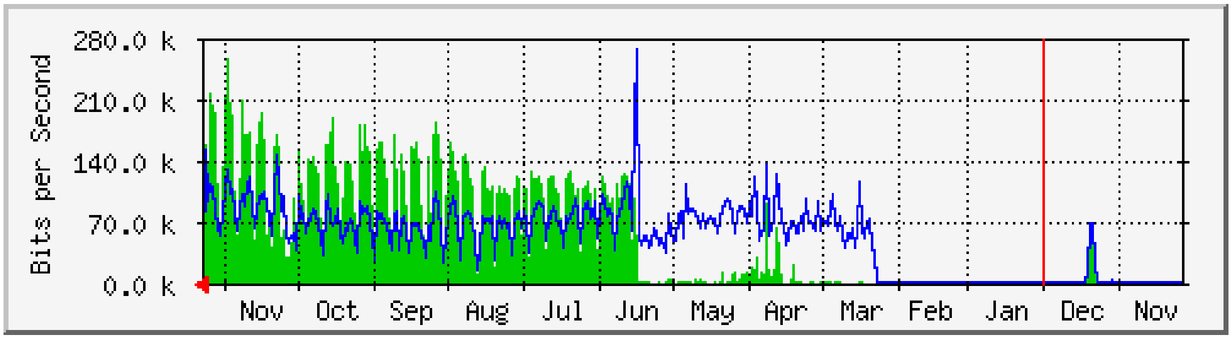
Max In:	793.2 kb/s (0.8%)	Average In:	147.4 kb/s (0.1%)	Current In:	481.9 kb/s (0.5%)
Max Out:	991.1 kb/s (1.0%)	Average Out:	95.9 kb/s (0.1%)	Current Out:	109.1 kb/s (0.1%)

Figure 6.7: MRTG Graph of BOL: Monthly' Graph (2 Hour Average)



Max In:	725.8 kb/s (0.7%)	Average In:	144.2 kb/s (0.1%)	Current In:	351.2 kb/s (0.4%)
Max Out:	817.7 kb/s (0.8%)	Average Out:	95.4 kb/s (0.1%)	Current Out:	93.5 kb/s (0.1%)

Figure 6.8: MRTG Graph of BOL: `Yearly' Graph (1 Day Average)



Max In:	259.3 kb/s (0.3%)	Average In:	57.2 kb/s (0.1%)	Current In:	36.7 kb/s (0.0%)
Max Out:	267.8 kb/s (0.3%)	Average Out:	55.5 kb/s (0.1%)	Current Out:	11.9 kb/s (0.0%)

Monitoring the graphs:

- Highest Bandwidth:
  - 1200 kbps approximately through IX (monthly average)
- Cost of 600 Kbps is 540000.00 Taka
- Extra sell is 540000.00 Taka

6.2 **National Bandwidth savings:**

There are 16 ISPs connected through right now. But there are around 80 active ISPs in Bangladesh.

NOTE: For the list of current Bandwidth Monitor Graph for Traffic Class of BDIX see Appendix B

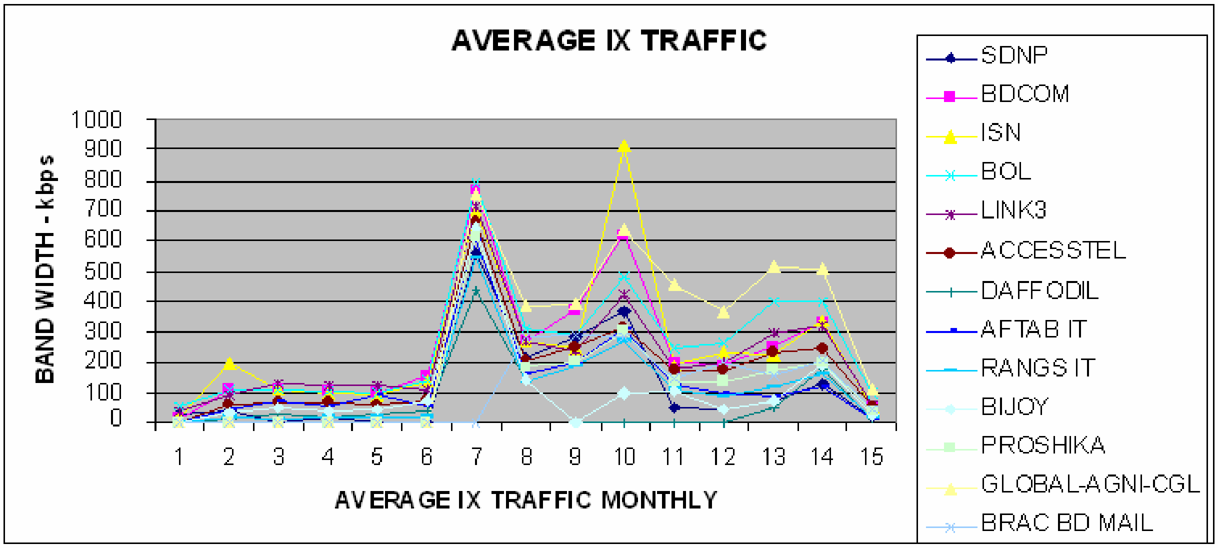


Figure 6.9: Average Internet Exchange monthly traffic

From the graph:

Monthly saving of an ISP highest 1.6 Mbps

Monthly saving of an ISP average 1.0 Mbps

Any ISP can sell extra 800 Kbps of Bandwidth to customer per month.

So approximately more sale will be 350000.00 Taka

National savings is approximately 10500000 Taka per month.

Note: National savings is approximately 10500000 Taka per month recently. But the savings will increase to about 50000000.00 taka per month with in few months. Because some ISPs are going to connect within two month, thus internal traffic will increase to approximately 10 Mbps so far.

Chapter 7

Future Improvements of IX infrastructure in Bangladesh

One improvement can be whether to have single or multiple switch equipment. Financial constraints may mean that the IXP can only afford a single switch, and indeed many successful IXPs have started this way. This may be the only solution initially available to the IXP; however, some observations with regard to the use of multiple switches are discussed here to provide some assistance in planning for the future growth of an IXP, even where the IXP has had to be established with modest resources.

The major benefit of building a multiple switch IXP is redundancy. This may be purely physical redundancy by using more than one switch from a single vendor, or physical and 'genetic' redundancy by using switches from two or more vendors. Single vendor redundancy can help to ensure the operation of the IXP in the case of a hardware failure but not certain software (or firmware) failures that may be common to all devices from that vendor. Multiple vendor redundancy can potentially help in both failure modes; it is unlikely that a software bug causing the failure of one vendor's equipment will also affect another vendors' equipment. Against this some other factors should be considered; equipment from a single vendor may provide better economies of scale in purchasing, spares holding, maintenance and management than that of two or more vendors; also interoperability issues are more likely with equipment from multiple vendors.

Multiple switch IXPs can offer better continuity of service than a single switch IXP. IXPs with multiple switches often offer, allow, or require members/customers to have two or more connections to two or more physically separate switches. In this situation, should one switch fail the connected member/customer network has an alternative route to the IXP. Continuity of service can also be provided during routine maintenance or upgrade of any one switch device. This, does however, result in increased hardware costs, greater management overhead, and more cost for members/customers, who will require a second router interface and more cabling from their infrastructure to the IXP.

Internet Exchange can offers a unique environment for organizations looking to leverage their marketing opportunities, enabling companies to undertake a totally distinctive selling alternative that is proven to increase campaign awareness, sales, and customer retention. Partnering with Internet Exchange helps educate customers, drive more sales through websites or stores and increases both customer retention and future buy rate.

BDIX is now in its starting point. At this moment, it is operating upon single switch. But using redundancy in switch, IX technical persons can provide more Availability.

7 Improvement required for BDIX

BDIX needs some improvement of its resource and management for better contribution to the ISP as well as the Internet User in Bangladesh. They have some short comes of resource, man power, work place, policies and fund. Some of the improvements are describe bellow:

7.1 Collector Router

To assist the IXP and members in troubleshooting, BDIX can provide a router with which all members peer and announce their routes. The router listens to, or 'collects' these announcements, but does not announce any routes itself, hence some IXPs use the term 'collector' router for this equipment. IXP staff and member ISPs have user accounts on this router, enabling them to have a central 'view' of the IXP, independent of the 'view' through their own connection.

7.2 Transit Router

Where an IXP has server equipment hosting, for example, their web site and email, and possibly some staff requiring Internet access, a router with full Internet connectivity is obviously required. (With care in configuration, this function can be combined with that of a 'collector' router.) BDIX can think about it for some extra facilities.

7.3 Security

IXPs often become critical to their members' businesses, so it is important that the site of the BDIX is as secure as possible. The housing space must be equipped with sophisticated security systems, some elements to consider when looking for space are: 24x7 security manning, CCTV coverage (inside and out), and multiple level access control (site, building, room, rack/cage).

7.4 Power Redundancy

Much modern switching and routing equipment has the capability of redundant Power Supply Units. To take advantage of this feature there should be provision of at least two power supplies. There should also be on site generators, with battery back-up/switch over to protect against total supply failure. The IXP may also wish to consider having its own non-interruptible power supply.

7.5 Manpower

BDIX perform its operation at Pantapath Feroz Tower. Currently they have 5 stuffs. They are worked every working day through 9 to 5 pm. So, if there is some problems occur in the holiday then the BDIX will be down until the next working day coming. So, to prevent this type of extreme situation some action should be taken. For that it needs more manpower to prevent those problems.

Chapter 8

F-Root Server:

A machine that has the software and data needed to locate name servers that contain authoritative data for the top-level domains (e.g., root servers know which name servers contain authoritative data for com, net, ch, uk, etc.). The root servers are, in fact, name servers and contain authoritative data for the very top of the Domain Name System (DNS) hierarchy. Currently, technical specifications limit the number of root servers to 13, from root server ‘A’ to ‘M’. There is a mirror F-Root Server in Bangladesh at Bangladesh Internet Exchange (BDIX). These machines are currently located around the globe, in the U.S., the U.K., Sweden, and Japan. Root nameserver is a DNS server that answers requests for the root namespace domain, and redirects requests for a particular top-level domain to that TLD's nameservers. The root servers hold the list of addresses for the authoritative servers for the top-level domains. Every name lookup must either start with an access to a root server, or use information that was once obtained from a root server.

All domain names on the Internet actually end in a . (period) character -- that is, technically, bracuniversity.net is actually hosted on the domain "www.bracuniversiy.net." This final dot is implied, as modern DNS software does not actually require that the final dot be included when attempting to translate a domain name to an IP address. The empty string after the final dot is called the root domain, and all other domains (i.e. .com, .org, .net, .uk, etc.) are contained within the root domain.

When a computer on the Internet wants to resolve a domain name, it works from right to left, asking each nameserver in turn about the element to its left. The root nameservers (which have responsibility for the . domain) know about which servers are responsible for the top-level domains. Each top-level domain (such as .bd) has its own set of servers, which in turn delegate to the nameservers responsible for individual domain names (such as wikipedia.org), which in turn answer queries for IP addresses of subdomains or hosts.

No more names can be used because of protocol limitations, but the C, F, I, J and K servers now exist in multiple locations on different continents, using anycast announcements to provide a decentralized service. As a result most of the physical, rather than nominal, root servers are now outside the United States. As a result there is a mirror of F-root server in Dhaka at BDIX.

NOTE: To see Appendix A for Root Servers list

8.1 Domain Name System:

The Internet Domain Name System (DNS) is a distributed Internet hierarchical lookup service primarily used to translate between domain names and Internet Protocol (IP) addresses. For example, when a user enters http://www.itu.int in a web browser, the browser first dispatches a DNS query for www.itu.int. The DNS responds with the Internet Protocol (IP) address, in this case 156.106.192.163. Packets and routing on the Internet are all based on these numeric addresses, and the DNS thus forms the bridge between the lower-layer IP and TCP protocols and the upper-layer applications such as web browsing and electronic mail.

The DNS service consists overall of DNS data, name servers, and a protocol used to retrieve data from the servers. Clients of the DNS can be applications such as web browsers or mail transfer agents and even other name servers. Simple text database records called resource records are placed into millions of files called *zones*. Zones are kept on *authoritative* name servers distributed around the Internet, which answer queries according to the DNS network protocols. In contrast, *caching* servers simply query the authoritative servers and cache any replies. Most servers are authoritative for some zones and perform a caching function for other DNS information. The DNS software implementation known as Berkeley Internet Name Domain (BIND), produced by the Internet Software Consortium (ISC), is the most commonly used domain name server on the Internet.

To understand the DNS hierarchy, it is useful to examine the structure of Internet host names. The last portion of a host name, such as .int, in the case of the www.itu.int (ITU's web site), is the top-level domain to which a host belongs. There are a set of generic top-level domains (gTLDs) such as .com, .net, and .org, as well as country code top-level domains (ccTLDs) such as .be for Belgium, .cn for the People's Republic of China, .mx for Mexico, and .us for the United States, .bd for Bangladesh, etc. Other top-level domains such as .int, .gov, .mil and .edu do not neatly fit into either of these classifications — they could be considered "chartered" gTLDs since they have entrance requirements before registration is accepted. The .int gTLD, for example, is currently restricted to intergovernmental organizations and .edu is restricted to educational institutions.

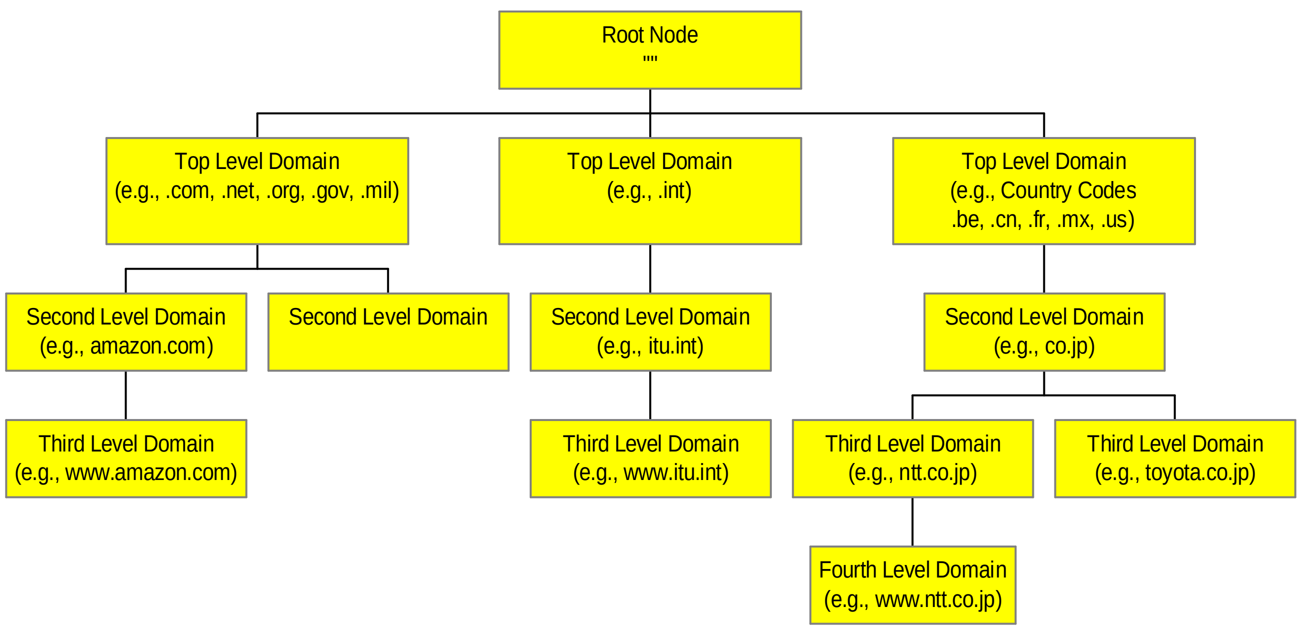


Figure 8.1: DNS Hierarchy

8.2 Root Server System:

The root node of the Internet name space consists of a single file, the root zone file. The root zone file contains pointers to the master (primary) and slave (secondary) servers for all top-level domains (e.g. gTLDs, ccTLDs).

The master or primary server is the definitive source of data for a DNS zone. This is where all changes to the zone's contents are made. The DNS protocol provides an automatic mechanism for propagating the contents of a zone to slave or secondary servers. The provision of secondary servers provides improved reliability and robustness and prevents having a single point of failure. If one name server for a zone fails or is unreachable, there will be other name servers for the zone that can be queried instead. Usually a name server will only give up on an attempt to resolve a query when all the known servers for the zone have been tried and none have responded.

At the top of the DNS database tree are 13 root name servers. Until very recently, the primary root server was “a.root-servers.net” with 12 secondary name servers, named “b.root-servers.net through “m.root-servers.net”. However, following recent changes in the root server network architecture, a separate and hidden master server is the definitive source of information carried by all thirteen root name servers. The final authority for change control of the root zone file (e.g. addition of new top-level domains or modification of existing top level domains) is held by the United States Department of Commerce.

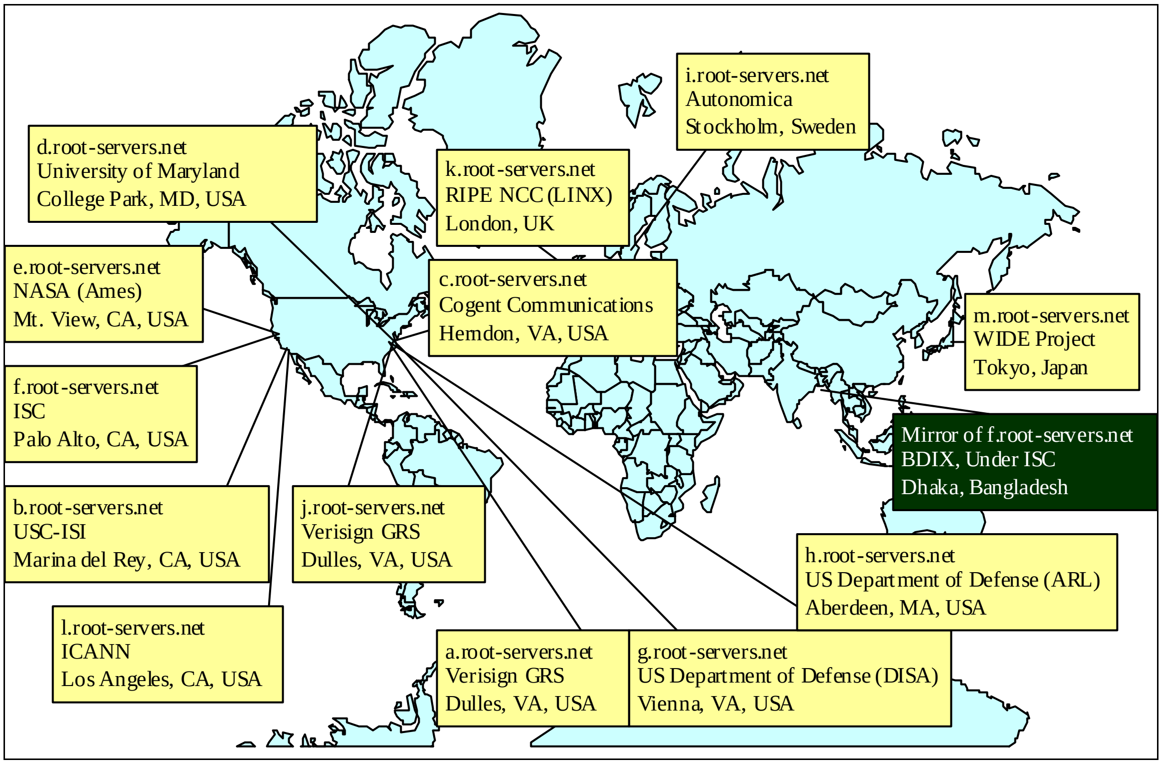


Figure 8.2: Location of 13 DNS Root Name Servers

An example can be given of a DNS lookup to find the Internet Protocol (IP) address of ITU's website: www.itu.int as follows: When a server looks up www.itu.int, it will go to the root name servers. They would then return a referral to the .int name servers. The local server then queries one of them for www.itu.int. A server for .int then returns a referral to the itu.int name servers. The server repeats the query for www.itu.int a third time, this time to one of the itu.int name servers, which gives the answer. This iterative process is known as resolving.

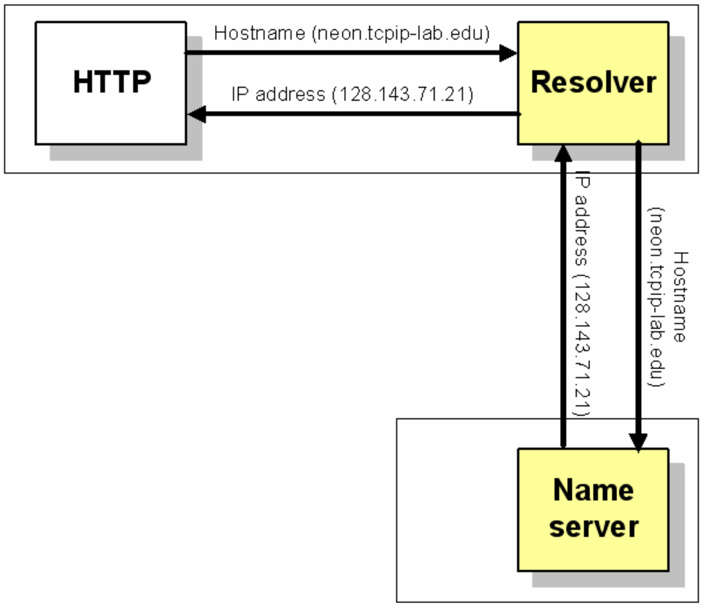


Figure 8.3: Resolving Domain name

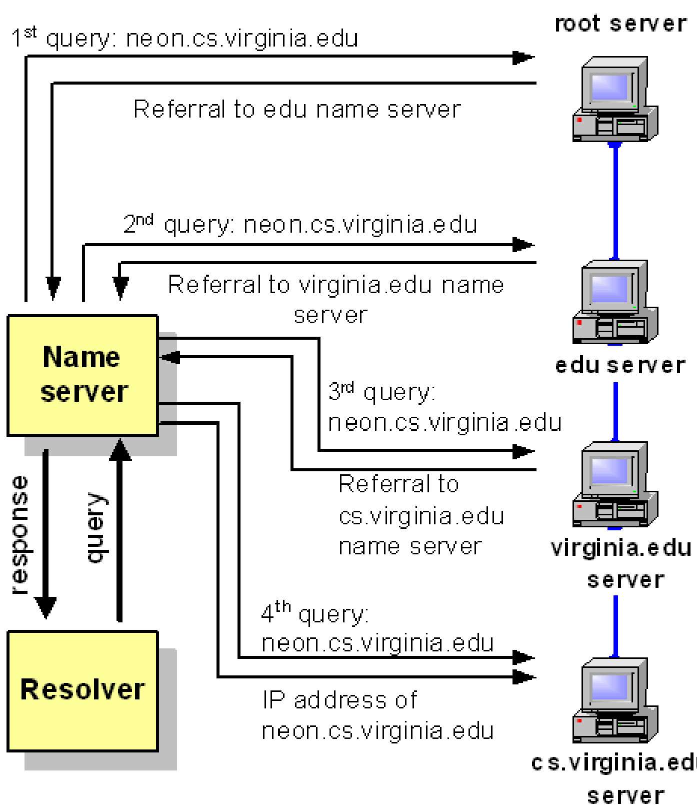


Figure 8.4: Domain name Query

The answers a name server gets when it is resolving queries are cached and used to speed up subsequent lookups. For example, if the name server that looked up www.itu.int were then asked to lookup up mail.itu.int, it would query the itu.int name servers directly and not start resolving the query again from the root name servers.

8.3 Root Server Anycasting:

“Anycasting” is a recently developed technique for “cloning” one server in multiple locations, all of which respond to the same IP address and all of which contain identical data. This technique has recently been applied to the F root server, allowing any suitably provisioned location to have a root server.

In the first half of 2003, several announcements have been made by the Internet Software Consortium for plans to anycast the F root server in new locations, including in New York City, Los Angeles, Madrid, Rome, Auckland, and Asia-Pacific sites (the latter in cooperation with APNIC). RIPE NCC has also recently published information on the potential anycasting of the K root server, which is hosted in the United Kingdom of Great Britain and Northern Ireland.

8.4 F root server in Bangladesh:

Sustainable Development Networking Program (SDNP) Bangladesh pleased to announce that a mirror of the Internet F-root nameserver has been deployed in Bangladesh on 9th December 2005. The Root Server- F is one of the 13 Root Servers around the world to facilitate operation of the Internet. This F-Root Server will enhance the Internet service quality of all Bangladeshi Internet users, and especially for the subscribers of the BDIX connected ISPs. The installation of this root server in Bangladesh has been made possible by SDNP Bangladesh and Internet Service Provider Association of Bangladesh (ISPAB) with financial and logistical support from Asia Pacific Network Information Centre (APNIC) and

Internet Systems Consortium (ISC).

8.5 Benefits of Root Server

The latter is of particular benefit for developing countries and/or isolated nations. A brief summary of the potential benefits of localized national root service include:

National infrastructure protection and self-sufficiency:

Domain name system resolvers will continue to function in a predictable way even during a loss of international connectivity, since they have access to local root name services;

Anycast:

One benefit of using anycast for root name services is that it solves a number of security issues and permits a better global distribution of root name service. It also permits a reduction in router and link resources, as standard IP routing protocols will deliver packets over the shortest path to the closest available host. This benefit keeps traffic in a local or regional context and thereby reducing the use of expensive international links.

Performance:

Resolvers will have a much lower latency (and, often, a less congested) path to a local root name server than to a root name server that must be accessed across international transcontinental links. This improves the average speed of DNS resolution, since recursive queries aimed at the root will get answered more quickly.

Without F-root server in Bangladesh to resolved the name (any domain name) need 800ms but after setup F-root server now times only 4 ms. So it is too much faster for caching domain name and we told one step advanced of ICT in Bangladesh. The query result are given

#dig @f.root-servers.net

;; Query time: 4 msec

;; SERVER: 192.5.5.241#53(f.root-servers.net)

;; WHEN: Sun Dec 11 14:04:10 2005

;; MSG SIZE rcvd: 436

Costs:

At a national level, it permits a reduction in router and link resources, as standard IP routing protocols will deliver packets over the shortest path to the closest available host. This benefit keeps traffic in a local or regional context and thereby reducing the user of expensive international links.

Resilience:

If a denial-of-service attack is launched at the F root name server in some other part of the world, the traffic will land at a different instance of F and hence the local community will not see any effects of the attack. Conversely, a locally originated attack will hit just the local F root node, and leave the rest of the world's F root name service unaffected. This makes the root name server infrastructure as a whole more resilient.

Emergency Response:

An indirect benefit is that in having a distributed set of sinks for attacks traffic makes it easier for the root server operators to identify and isolate the source of an attack. The quicker the source of an attack can be identified, the sooner the attack can be stopped.

Chapter: 9

Policies for Internet Exchange

9.1 Peering Policies of BDIX:

1. The BDIX at SDNP is a layer two Internet Exchange over Ethernet. All participants connecting to BDIX are exchanging data via an Ethernet, which is a layer two technology. All routers on the Ethernet are allowed to peer by exchanging routing tables through the route server set up by BDIX using BGP4.

2. The main purpose of the BDIX is for routing of intra-Bangladesh traffic (the participants are allowed to exchange traffic among their peers or downstream network(s) within the country) at free of charge.

3. BDIX is a settlement-free interconnection point. In other words, no settlement is to be paid by the peering participants for the incoming and outgoing traffic.

4. All participants should not filter traffic or routing table entries to or from any other participants.

5. The data links to BDIX and the routers located at BDIX must be supported (or paid) by the corresponding participants. At present BDIX has 16 routers, of which they will be given to 16 ISP's at "first come first connect" basis without any charge for the router at BDIX end. ISP must provide routers and there end equipment until these resources exhausts. BDIX has also 16 Radios, and can be given with the same policy.

6. The routers distributed under the pilot programme located at BDIX will be managed by SDNBD technical team. Individual Routers provided by ISPs may be managed separately under technical supervision of SDNP Bangladesh.
7. Each participant must have its own global Internet connectivity through other

Internet access provider(s) which is independent of BDIX facilities.

8. The BDIX must not be used as the primary connection to the global Internet.

9. IX is going to setup a route server on BDIX so that the participants can only peer with this server. They can also peer with other participants directly as long as the arrangement does not violate any of the guidelines/policies stated here.

10. All ISPs at BDIX, no matter how large or how small, are considered to be as equal and all ISPs must peer with each other via the route server.

11. Each participant must have an Autonomous System (AS) by itself with a globally unique AS number assigned by InterNIC, RIPE-NCC, APNIC or their sub-registries. The source addresses of the data traversing the BDIX must be officially assigned by InterNIC, RIPE-NCC, APNIC or their sub-registries.

12. SDNP is providing the space, electricity, air-conditioning and Ethernet connection to the BDIX for the routers located BDIX at pilot basis, until a membership arrangement be established within the next six months (starting Aug. 2004).

13. BDIX will not be responsible for any loss and damages to the participants caused by the operations of BDIX.

14. BDIX is not responsible for any illegal activities performed at any participants.

15. BDIX is a non profit entity. For the first six months, the services will be free. There may be a nominal membership fee of BD Tk. 5000/= from April 2005 to cover the cost of coordinating and managing BDIX activities.

16. Participant ISP must have legal license from Bangladesh Telecom Regularity Commission. Other entities line, universities / research institutes / corporate

bodies / IXPs may peer subjected to the fulfillment of the Peering policy 2.

17. Participant ISPs cannot share their IX connectivity to other ISPs, who are not the member of BDIX. BDIX updates route information for members of BDIX only.

<u>Recommendation for BDIX:</u>

18. All BDIX members must ensure that they suitably and proactively upgrade capacity from time-to-time so that they do not end up dropping traffic that other peers are exchanging with them.

19. By joining BDIX, each ISP saves a certain amount of money which they must spend for customers benefit or reduce installation or per month connection cost for customer.

20. BDIX should have the authority for screening of traffic passing through IX for national security purpose.

21. Right now, UNDP is funding the BDIX project which includes ISP connectivity at BDIX end through radio link facility also. But in future, when BDIX will have to run by itself then ISPs (those are interested to get connected) will have to bring their own equipments.

9.2 IX-IX Policies

1. In Bangladesh, in future if there is another IX then they must peer with each other.
2. The connecting cost at the end point should be beared by each individual IX.
3. They provide only domestic Internet traffic and no International traffic is allowed.

9.3 BTRC Policies for IX (Recommended)

1. Connectivity cost for the ISPs reduced immediately because it is the main obstacles right now for coming up new ISPs. Because a leased T1 link in the US costs about $ 2,000 per month and a T3 link about $ 40,000 per month whereas in Bangladesh a dedicated 64 Kbps link is $ 8,000.

2. BTRC should have the policy for monitoring the packets for national security purpose.

3. It is needed to review the spectrum allocations in a planned manner so that required frequency bands are available to the service providers.

4. Commercial VSAT service providers having ISP licence shall be permitted use of same hub station and remote station to provide Internet service directly to the subscribers. Further, this remote station shall be permitted to be used as a distribution point to provide Internet services to multiple independent subscribers. Necessary amendments in the licence agreement shall be carried out immediately. For all these process no extra charge should be charged.

5. BTRC should have intention to make available transponder capacity for VSAT services at competitive rates after taking into consideration the security requirements.

7. Any new ISP interested to get license should have to be connected to BDIX.

References

1. Riton Kumar Roy
System Administrator
BDIX, Sustainable Development Networking Programme, BIDS
Bangladesh
2. Sumon Ahmed Sabir
Director
BDCOM ONLINE Ltd.
4. BGP documentation – provided by SDNP

5. We got information about ISP association from www.ispab.org

6. Linux Home Networking
Grief Relief For The Home And Small Office
By Peter Harrison
www.linuxhomenetworking.com

7. Cisco certified Network Associate Study Guide
Rechard Deal
CCNA, CCNP, CCDP, CCDA
Tata McGraw-Hill Publishing Company Limited, NEW DELHI

8. BGP routing information, available in online at http://www.wikipedia.com, http://www.icann.org/general/crada-report-summary-14mar03.htm, : http://www.ietf.org/rfc/rfc1546.txt, http://www.cisco.com/public/cons/isp/essentials/ip-anycast-cmetz-03.pdf, http://www.ietf.org/rfc/rfc3258.txt, http://www.isc.org

9. Bandwidth saving information of different ISPs, available online at http://www.bdix.net

10. Some reference regarding thesis format are available at http://geocities.com/manzur_bd/fall2004

11. F-Root Server information, available in online at http://www.rootservers.org, www.isc.org.

12. A paper for Against All Odds, The Internet in Bangladesh by Larry Press The MOSAIC Group, Fairfax, VA, March 1999

13. Research paper VoIP Deregulation in Bangladesh
And BDIX Updates
APNIC 18, Nadi, Fiji, August 31, 2004
Presented by
Sumon Ahmed Sabir, Joint Secretary, ISP Association Bangladesh

14. Paper of Internet Exchange Points
Their Importance to Development of the Internet and
Strategies for their Deployment – The African Example

15. A tutorial for ISP Security
Vicky Shrestha
WorldLink Communications

Appendix

Appendix A

List of ISP Association Bangladesh Members

1) Access Telecom (BD) Ltd
2) Aftab IT Ltd
3) Agni Systems Ltd
4) Akij Online Ltd
5) Bangladesh Internet Exchange Ltd
6) Bangladesh Online Ltd
7) BDCOM Online Ltd
8) Bijoy Online Ltd
9) Brac BDMail Ltd
10) Broadband Telecom Services Ltd
11) BTS Communications (BD) Ltd
12) Connect BD Ltd
13) CPM Blue Online Ltd
14) Cyberx IT Ltd
15) Daffodil Online Ltd
16) Data Edge Ltd
17) Data Net Corporation Ltd
18) Dekko Airnet Ltd
19) Desh Online Services
20) Dhaka Com Ltd
21) Dominox Technologies Ltd
22) Dotcom Network Ltd
23) Drik Net Ltd
24) Global Access Ltd
25) Global Information Network Ltd
26) Global Link Telecom Ltd

27) Global Online Services Ltd
28) Grameen Cyber Net Ltd
29) Ichinet Ltd
30) Information Services Network Ltd
31) Insoft Systems Ltd
32) Intech Online ltd
33) International Computer Connections
34) Libra Computer Aid
35) Link3 Technologies Ltd
36) Maisha Technologies Ltd
37) etronet Bangladesh Ltd
38) MLB Net
39) Net Access Bangladesh
40) Poly IT Ltd
41) Pradeshta Ltd
42) Proshika Computer Systems
43) Ranks ITT Ltd
44) Reve Systems
45) Reza Telecom Bangladesh
46) Simni IT Induces Ltd
47) Sirius Broadband (BD) Ltd
48) Software Shop Limited
49) Spectra Net Ltd 0020
50) Square InformatiX Limited
51) Techno Online Ltd
52) Texas Electronics Ltd
53) Westec Ltd
54) Western Network Ltd
55) New Generation Graphics Limited

List of members of Internet Exchange of Bangladesh (BDIX):

SL/NO	ISP NAME	IX-IP	AS NUMBER	CONNECTION DATE
01	**SDNP**	198.32.167.17	9825	27-08-2004
02	**BDCOM ONLINE LTD**	198.32.167.18	24122	27-08-2004
03	**BANGLADESH ONLINE (BOL)**	198.32.167.20	9230	06-09-2004
04	**INFORMATION SERVICE NETWORKLIT.(ISN)**	198.32.167.19	9832	04-09-2004
05	**LINK3 TECHNOLOGIES LTD.**	198.32.167.21	23688	27-09-2004
06	**ACCESS TEL**	198.32.167.22	17469	04-11-2004
07	**DAFFODIL ONLINE**	198.32.167.23	24308	28-11-2004
08	**AFTAB IT**	198.32.167.24	24307	28-11-2004
09	**RANKS-ITT LTD**	198.32.167.25	23991	28-11-2004
10	**BIJOY ONLINE LTD**	198.32.167.26	65002	29-11-2004
11	**PROSHIKA**	198.32.167.27	3758	03-03-2005
12	**BRAC BDMail**	198.32.167.28	24342	03-04-2005
13	**BTTB**	198.32.167.30	17494	30-11-2005

Place visited during data collection

During our thesis we visited various government and non-governmental organizations. The lists of the organizations are listed below:

Government Organization:

1. SDNP
2. Bangladesh Institute of Development Studies (BIDS)
3. Bangladesh Telecommunication Regulatory Commission (BTRC)

Government Organization:

1. Internet Service Provider Association Bangladesh (ISPAB)

Private ISPs:

1. Bangladesh Online Ltd
2. BDCOM Online Ltd
3. Brac BDMail Ltd
4. Connect BD Ltd
5. Sirius Broadband (BD) Ltd
6. Ranks ITT Ltd
7. Agni Systems Ltd
8. Drik Net Ltd
9. Druti IT Ltd

Government ISP:

1. Sustainable Development Networking Programme Bangladesh (SDNP)

Root Server List:

There are currently 13 root name servers, with names in the form *letter*.root-servers.net where *letter* ranges from A to M:

Letter	Old name	Operator	Location
A	ns.internic.net	VeriSign	Dulles, Virginia, USA
B	ns1.isi.edu	ISI	Marina Del Rey, California, USA
C	c.psi.net	Cogent	distributed using anycast
D	terp.umd.edu	University of Maryland	College Park, Maryland, USA
E	ns.nasa.gov	NASA	Mountain View, California, USA
F	ns.isc.org	ISC	distributed using anycast
G	ns.nic.ddn.mil	U.S. DoD NIC	Vienna, Virginia, USA
H	aos.arl.army.mil	U.S. Army Research Lab	Aberdeen Proving Ground, Maryland, USA
I	nic.nordu.net	Autonomica	distributed using anycast

J		VeriSign	distributed using anycast
K		RIPE NCC	distributed using anycast
L		ICANN	Los Angeles, California, USA
M		WIDE Project	Tokyo, Japan

Older servers had their own name before the policy of using similar names was established.

No more names can be used because of protocol limitations, but the C, F, I, J and K servers now exist in multiple locations on different continents, using anycast announcements to provide a decentralized service. As a result most of the physical, rather than nominal, root servers are now outside the United States.

F Root-Server: List of types and location

Site Code	Location	IPv4/IPv6	Node Type
AKL1	Auckland, New Zealand	IPv4 and IPv6	Local Node
AMS1	Amsterdam, The Netherlands	IPv4 and IPv6	Local Node
BCN1	Barcelona, Spain	IPv4 and IPv6	Local Node
BNE1	Brisbane, Australia	IPv4	Local Node
CDG1	Paris, France	IPv4 and IPv6	Local Node

CGK1	Jakarta, Indonesia	IPv4	Local Node
DXB1	Dubai, UAE	IPv4	Local Node
GRU1	São Paulo, Brazil	IPv4	Local Node
HKG1	Hong Kong, China	IPv4	Local Node
JNB1	Johannesburg, South Africa	IPv4	Local Node
KIX1	Osaka, Japan	IPv4 and IPv6	Local Node
LAX1	Los Angeles, CA, USA	IPv4 and IPv6	Local Node
LCY1	London, UK	IPv4 and IPv6	Local Node
LIS1	Lisbon, Portugal	IPv4 and IPv6	Local Node
LGA1	New York, NY, USA	IPv4 and IPv6	Local Node
MAA1	Chennai, India	IPv4	Local Node
MAD1	Madrid, Spain	IPv4	Local Node
MTY1	Monterrey, Mexico	IPv4	Local Node
MUC1	Munich, Germany	IPv4 and IPv6	Local Node
NBO1	Nairobi, Kenya	IPv4	Local Node
PAO1	Palo Alto, CA, USA	IPv4 and IPv6	Global Node
PEK1	Beijing, China	IPv4	Local Node
PRG1	Prague, Czech Republic	IPv4 and IPv6	Local Node

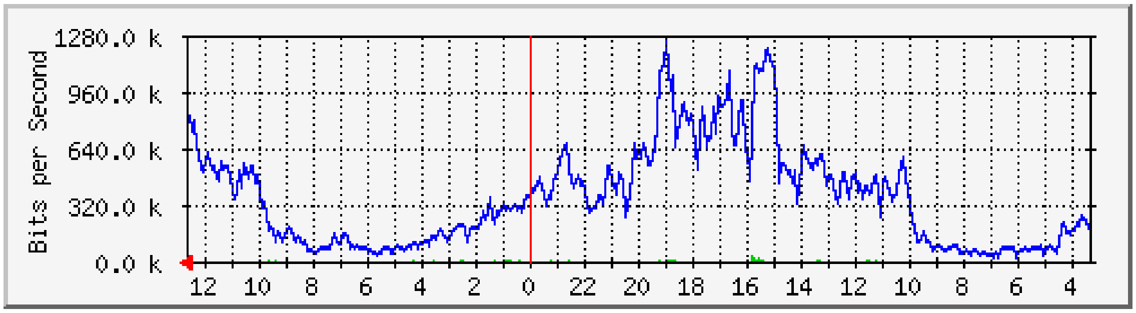
ROM1	Rome, Italy	IPv4	Local Node
SEL1	Seoul, Korea	IPv4 and IPv6	Local Node
SFO2	San Francisco, CA, USA	IPv4 and IPv6	Global Node
SIN1	Singapore	IPv4	Local Node
SJC1	San Jose, CA, USA	IPv4	Local Node
SVO1	Moscow, Russia	IPv4	Local Node
TLV1	Tel Aviv, Israel	IPv4	Local Node
TPE1	Taipei, Taiwan	IPv4	Local Node
YOW1	Ottawa, ON, Canada	IPv4 and IPv6	Local Node
YYZ1	Toronto, ON, Canada	IPv4	Local Node

Appendix B

List of all Bandwidth Monitor Graph for Traffic Class: BDIX

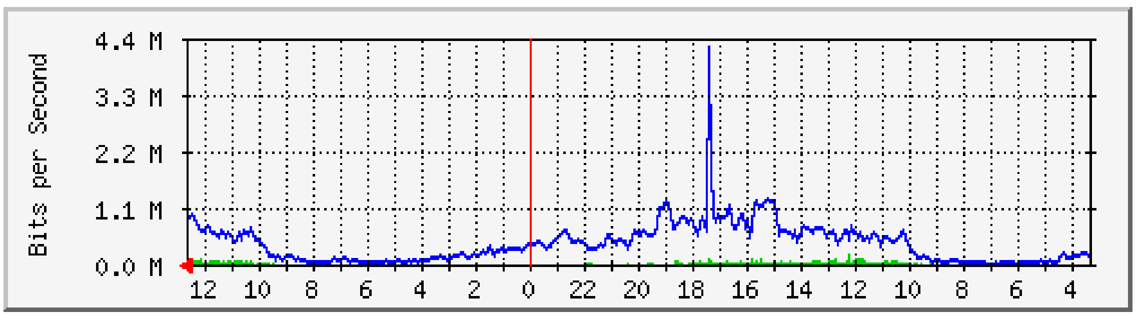
Bandwidth Monitor Graph for Traffic Class

47. SDNP Bangladesh (sdnbd.org)



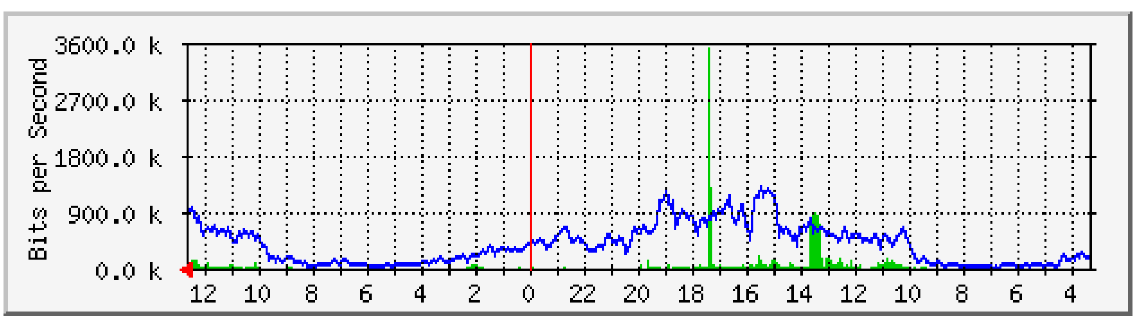
Bandwidth Monitor Graph for Traffic Class

18. BDCOM Online Ltd.(bdcom.com)



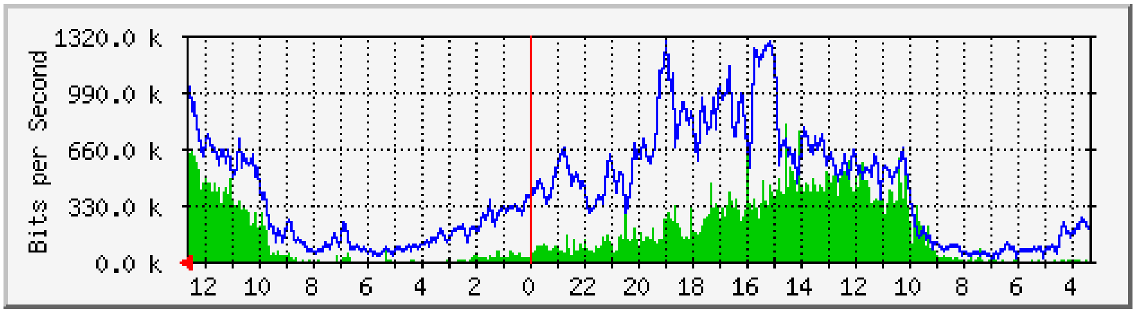
Bandwidth Monitor Graph for Traffic Class

19 ISN (bangla.net)



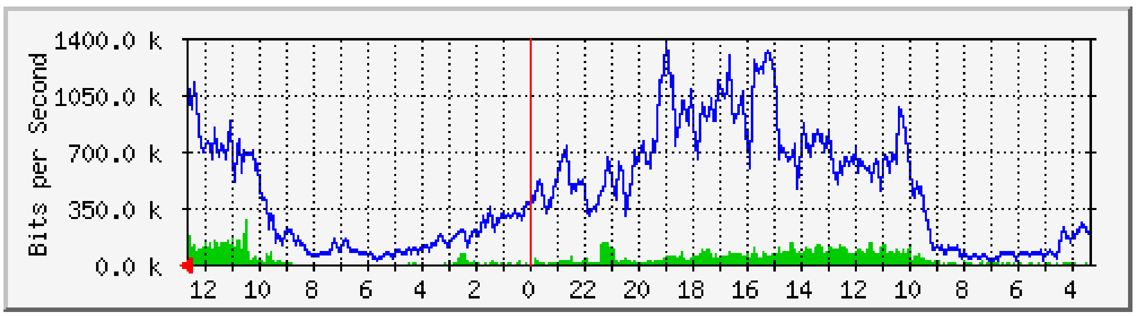
Bandwidth Monitor Graph for Traffic Class

20. BOL (bol-online.com)



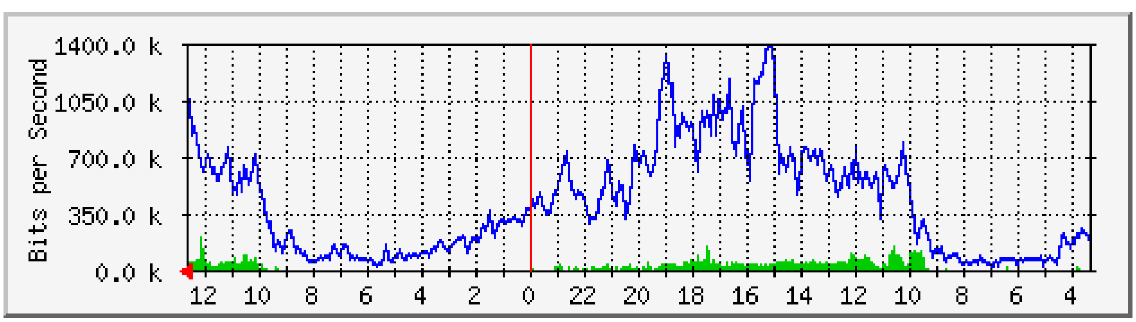
Bandwidth Monitor Graph for Traffic Class

21. Link3 (link3.net)



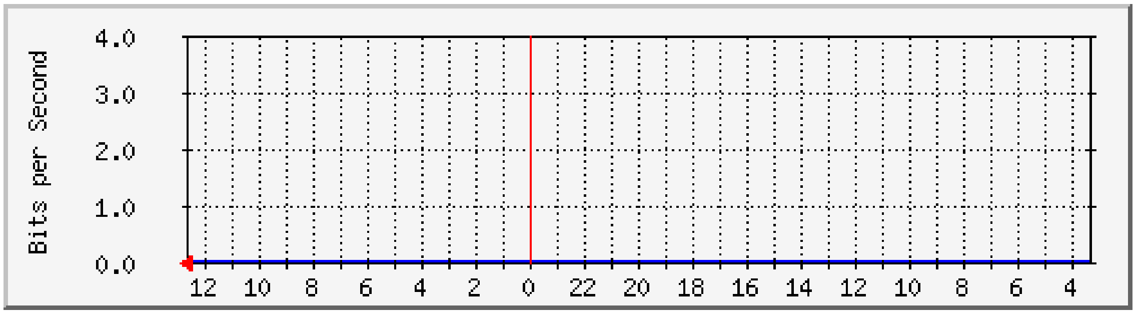
Bandwidth Monitor Graph for Traffic Class

22. Accesstel Ltd.(accesstel.com)



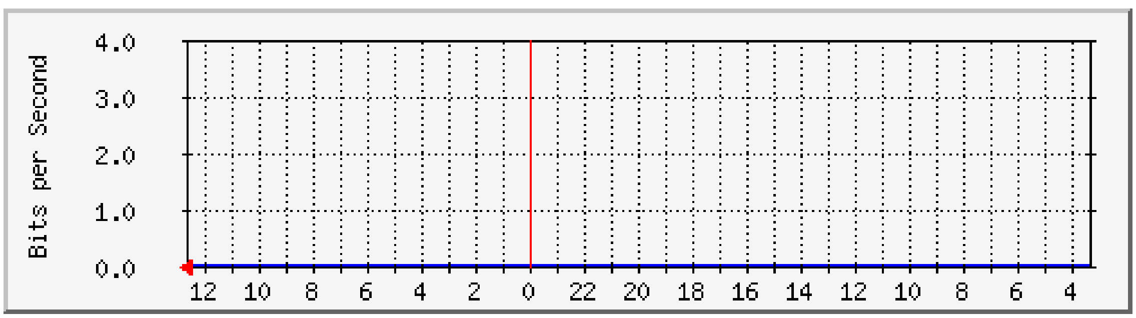
Bandwidth Monitor Graph for Traffic Class

23. Daffodil Online Ltd.(daffodilnet.com)



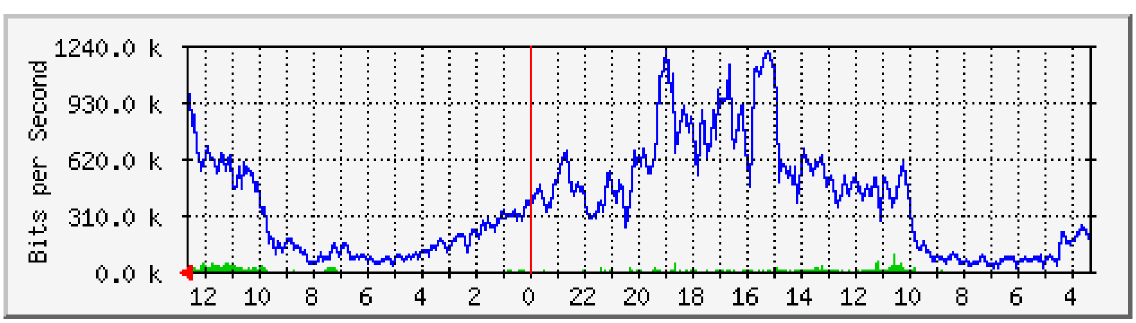
Bandwidth Monitor Graph for Traffic Class

24. Aftab IT Ltd.(aitlbd.net)



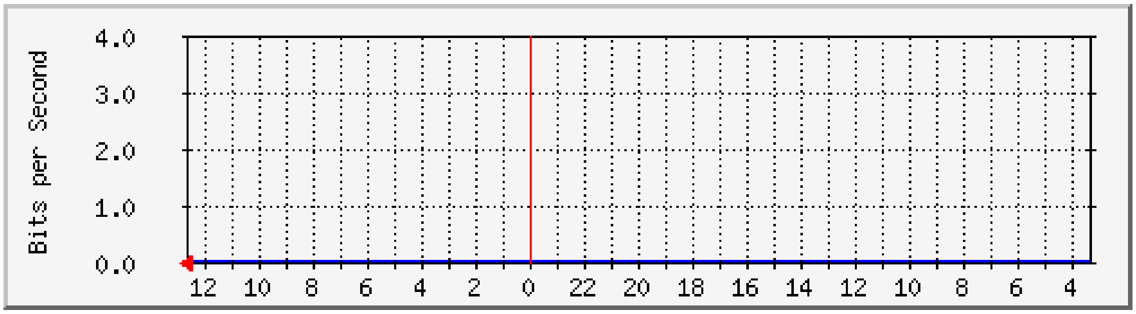
Bandwidth Monitor Graph for Traffic Class

25. Ranks IT Ltd.(ranksitt.net)



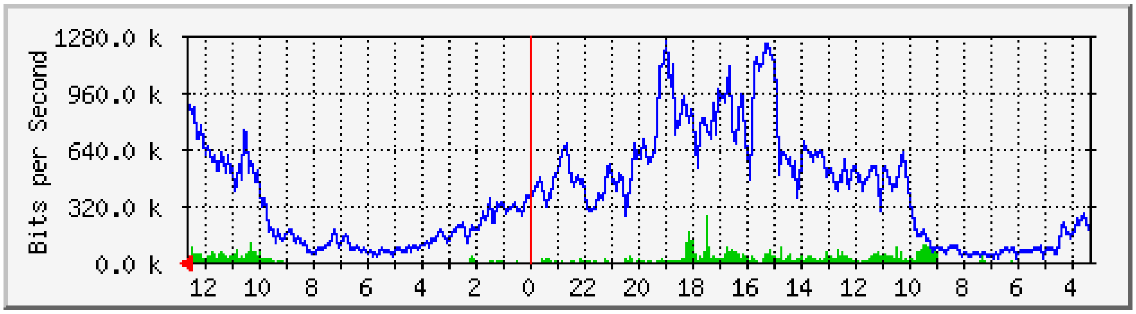
Bandwidth Monitor Graph for Traffic Class

26. Bijoy Online Ltd.(bijoy.net)



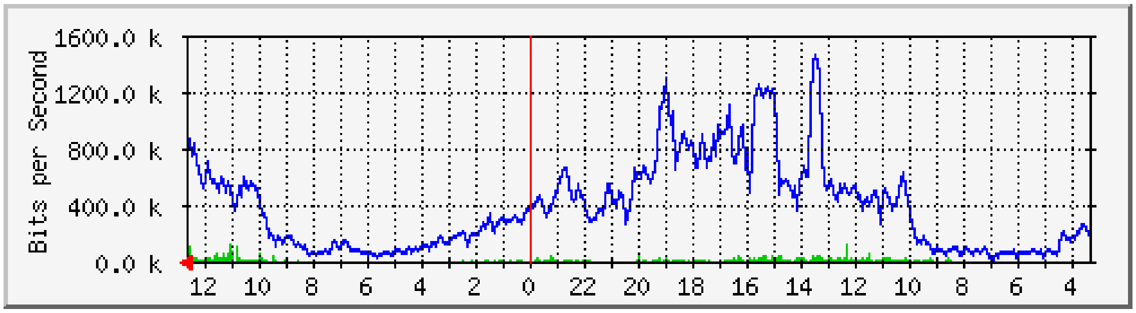
Bandwidth Monitor Graph for Traffic Class

27. ProshikaNet Ltd.(bdonline.com)



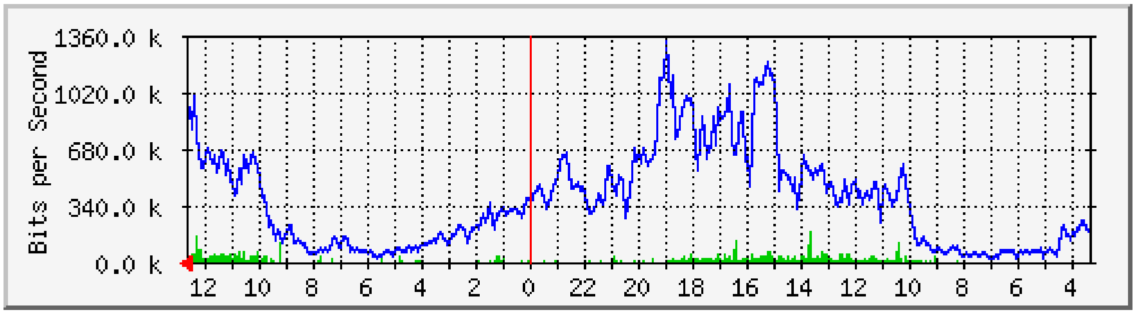
Bandwidth Monitor Graph for Traffic Class

28. BRAC BDMAIL (bdmail.net)



Bandwidth Monitor Graph for Traffic Class

30. BTTB (bttb.gov.bd)



Bandwidth Monitor Graph for Traffic Class

41. GLOBAL(252)-GCL(253)-Agni(254)