DELIVERABLE 4.1:  BROADBAND BOTTLENECK ANALYSIS

Identifier: Deliverable D4.1
Class: Report
Version: V04
Version Date: 30/6/2008
Distribution: Public
Responsible Partner: TID, Pedro A. Aranda Gutiérrez
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADSL</td>
<td>Asymmetric Digital Subscriber Line</td>
</tr>
<tr>
<td>ATM</td>
<td>Asynchronous Transfer Mode</td>
</tr>
<tr>
<td>AS</td>
<td>Autonomous System</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>BH</td>
<td>Busy Hour</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol</td>
</tr>
<tr>
<td>DFI</td>
<td>Deep Flow Inspection</td>
</tr>
<tr>
<td>FPGA</td>
<td>Field Programmable Gate Array</td>
</tr>
<tr>
<td>FTP</td>
<td>File Transfer Protocol</td>
</tr>
<tr>
<td>HTTP</td>
<td>HyperText Transfer Protocol</td>
</tr>
<tr>
<td>ICMP</td>
<td>Internet Control Message Protocol</td>
</tr>
<tr>
<td>IDS</td>
<td>Intelligence Distributed System</td>
</tr>
<tr>
<td>IEPM</td>
<td>Internet End-to-end Performance Measurement</td>
</tr>
<tr>
<td>IP</td>
<td>Internet Protocol</td>
</tr>
<tr>
<td>IPv4, IPv6</td>
<td>Internet Protocol version 4, and 6 respectively</td>
</tr>
<tr>
<td>IPPM</td>
<td>IP Performance Metrics Routing</td>
</tr>
<tr>
<td>ISP</td>
<td>Internet Service Provider</td>
</tr>
<tr>
<td>LIM</td>
<td>Line Interface Module</td>
</tr>
<tr>
<td>NIMI</td>
<td>National Internet Measurement Infrastructure</td>
</tr>
<tr>
<td>PDH</td>
<td>Plesiochronous Digital Hierarchy</td>
</tr>
<tr>
<td>P2MP</td>
<td>Point-to-point services</td>
</tr>
<tr>
<td>P2P</td>
<td>Peer-to-Peer</td>
</tr>
<tr>
<td>PL</td>
<td>PacketLogic</td>
</tr>
<tr>
<td>POTS</td>
<td>Plain Old Telephony System</td>
</tr>
<tr>
<td>RIR</td>
<td>Regional Internet Registries</td>
</tr>
<tr>
<td>RNMON</td>
<td>Remote Network MONitoring</td>
</tr>
<tr>
<td>RPSL</td>
<td>Routing Police Specification Language</td>
</tr>
<tr>
<td>RTT</td>
<td>Round Trip Time</td>
</tr>
<tr>
<td>SDH</td>
<td>Synchronous Digital Hierarchy</td>
</tr>
<tr>
<td>SELT</td>
<td>Single-ended line testing</td>
</tr>
<tr>
<td>SIP</td>
<td>Session Initiation Protocol</td>
</tr>
<tr>
<td>SLA</td>
<td>Service Level Agreement</td>
</tr>
<tr>
<td>SMB</td>
<td>Server Message Block</td>
</tr>
<tr>
<td>SNMP</td>
<td>Simple Network Management Protocol</td>
</tr>
<tr>
<td>SONET</td>
<td>Synchronous Optical Network</td>
</tr>
<tr>
<td>(R)STP</td>
<td>(Rapid) Spanning Tree Protocol</td>
</tr>
<tr>
<td>TCP</td>
<td>Transmission Control Protocol</td>
</tr>
<tr>
<td>UDP</td>
<td>User Datagram Protocol</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice on IP</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

ABBREVIATIONS .................................................................................................................................. 2

1 INTRODUCTION: STATE OF THE ART ............................................................................................... 5
  1.1 EXISTING METHODS FOR NETWORK ANALYSIS AND TRAFFIC MEASUREMENTS .................. 5
  1.2 EXISTING TOOLS ........................................................................................................................... 5
    1.2.1 Measurement tools .................................................................................................................... 5
    1.2.2 Data analysing tools .................................................................................................................. 9
    1.2.3 Routing Repositories ................................................................................................................ 13
      1.2.3.1 Repository Structure ........................................................................................................... 13
      1.2.3.2 Repository applications ..................................................................................................... 14
  1.3 CHALLENGES FOR BOTTLENECK ANALYSIS ............................................................................. 16
    1.3.1 Last-mile problems .................................................................................................................... 16
    1.3.2 Congestions and bottlenecks caused by inter-domain routing issues ..................................... 17
    1.3.3 DSL measurement tools applications ......................................................................................... 18
      1.3.3.1 Ericsson Loop Qualification and Monitoring (LQ&M) tool .................................................. 18
      1.3.3.2 Alcatel 5530 Network Analyzer ........................................................................................ 18
      1.3.3.3 DSL Expresse™ .................................................................................................................. 18
  2 RELATED PROJECTS ............................................................................................................................ 19
    2.1 IST PROJECTS ............................................................................................................................... 19
      2.1.1 INTERMON .............................................................................................................................. 19
      2.1.2 MOME Database ..................................................................................................................... 19
      2.1.3 LOBSTER for traffic network passive monitoring ................................................................. 20
    2.2 ROUTING REPOSITORIES .......................................................................................................... 20
    2.3 CAIDA WORK ON INTERNET MEASUREMENT DATA CATALOGUE .................................... 21
    2.4 BART METHOD FOR MEASURING AVAILABLE BANDWIDTH ................................................. 21
  3 FURTHER WORK .................................................................................................................................. 23
    3.1 IDENTIFYING BOTTLENECKS IN BROADBAND MULTI-SERVICE NETWORKS ...................... 23
    3.2 PROPOSING SOLUTIONS FOR ELIMINATING BOTTLENECKS .............................................. 23
      3.2.1 Solutions for DSL .................................................................................................................... 23
EXECUTIVE SUMMARY

The new broadband network challenge is based on strategic planning upon understanding and considering the requirements from the user and application perspectives. Service providers have requested the need of an integrated measurement and modelling scenario whose long-term goal is the Quality of Service monitoring and analysis at application level. Different aspects would be considered impacting on the broadband infrastructure performance:

- Network provisioning
- Resource use and planning
  - Topology and routing
  - Traffic Engineering
  - Abnormal events

This document is part of a series of works aiming to providing a feedback loop to optimise resource management and planning decisions.

The bottleneck analysis allows operators to identify which links or network parts are responsible for congestions. This may lead incumbent network operators to rethinking new dimensioning methods. Besides, when congestion is due to routing events or because of false dimensioning in an external network, it will help finding the service provider responsible for the possible violation of SLAs.

This report specifically presents the available broadband bottleneck analysis tools. It covers the state of the art in traffic analysis tools and points out which ones could be applied to bottleneck identification as well as which potential add-ons would enhance them to accomplish this objective more efficiently.
1 INTRODUCTION: STATE OF THE ART

1.1 Existing methods for network analysis and traffic measurements

Internet measurement implementation is the key for the next generation Internet design. There are two types of tools that allow Internet traffic measurement for its analysis

Active measurement tools: methods that involve adding traffic to the network for the purposes of measurement. For example:

- **Ping**: Sends ICMP ECHO_REQUEST and captures ECHO_REPLY. It is useful for measuring RTTs. Only sender needs to be under experiment control.
- **OWAMP**: A daemon running on the target which listens for and records probe packets sent by the sender. It is useful for measuring one-way delay. It requires both sender and receiver to be under experiment control and also requires synchronized clocks or a method to remove clock offset.
- **Traceroute**: Is useful for determining the path from a source to a destination. It uses the TTL (Time To Live) field in the IP header in a clever but distorted way. A large scale measurement system called skitter uses traceroute to discover network topology. However, Traceroute presents some drawbacks:
  o Unstable paths and false edges
  o Aliases → IP addresses are for interfaces and not routers; routers typically have many interfaces, each with its own IP address. IP addresses of all router interfaces are aliases. Traceroute results require resolution of aliases if they are to be used for topology building.
  o Measurement Load → Traceroute inserts considerable load on network links if attempting a large-scale topology discovery.

Passive measurement tools: Methods that capture traffic generated by other users and applications to build the topology. Routeview repository collects BGP views (routing tables) from a large set of ASs. Similarly, OSPF LSAs can be captured and processed to generate router graphs within an AS

- Large set of AS-AS, router-router connections can be learned by simply processing captured tables. However, especially using BGP views, there could be potential loss of cross-connections between ASs which are along the path. Secondly, route aggregation and filtering tends to hide some connections. Also, if multiple connections between ASs exist then they will be shown as a single connection in the graph.

1.2 Existing tools

1.2.1 Measurement tools

**Adtech AX4000**

The Adtech AX/4000 Broadband Test System is a hardware solution provided by Spirent. Adtech AX/4000 is capable of using many different interfaces and supporting nearly all kind of protocols and services in only one chassis. The test device can be used as traffic generator as well as analyzer, being able to compute delay measurements through switches or networks. It is possible to analyse heterogeneous networks since a large number of protocols are supported.

**Agilent Network Analyser and Distributed Network Analyser**

Agilent is a measurement and instrument company that formerly was a division of HP. For this reason Agilent manufactures many products that descend from HP's original product lines. The Agilent Network Analyser replaces the HP Agilent Advisor equipment which is no longer manufactured. It offers the following comprehensive features:

- Data can be accessed in multiple ways including standard web browser or instrument manager.
- Brings greater power for collecting and analysing Voice, Video and Data in real-time over multiple technologies such as Ethernet, ATM and IP.
- Triple Play Plus - Data, Voice, Video, Mobile all in one test solution for faster more efficient technology deployment.
- Supports any available LAN or WAN/ATM LIM1.
- LIMs are hot-swappable and interchangeable.
- Full-rate capture from any LIM up to 1 Gb/s.
- 256 MB capture memory per LIM.
- 32 real-time hardware capture filters.

**BART**

BART stands for Bandwidth Available in Real Time (see 2.4 for further information). It is an active tool for network performance analysis that uses inter-packet strain measurements between two terminals, one that sends probes and another one that receives them: When the network that links the two terminals becomes congested, the rate of arriving packets slows down, thus the measured inter-packet strain, $\varepsilon$, increases. The rate of packet injection, $u$, that makes $\varepsilon > 0$ corresponds to the value of available bandwidth and the bottleneck capacity is determined by the relationship $\Delta u / \Delta \varepsilon$. BART is still undergoing field test, as part of TRAMMS project and Ericsson has initiated the process to standardize the system in the ITU-T.

![BART Diagram](image)

**Bprobe**

Bprobe is a tool for measuring network bandwidth. It can be used for estimating the maximum bandwidth in the bottleneck of a link (i.e. the *available* bandwidth in the slowest link along a given path) when it is uncongested. The uncongested bandwidth is the maximum transmission rate that could be achieved by a connection in the absence of any competing traffic.

**Cprobe**

Cprobe is a software tool that estimates the current congestion along the bottleneck link within a path in the Internet. In particular, Cprobe measures the available bandwidth along a path, where the available bandwidth is the transfer rate available to an application at any time instant.

**Cisco routers**

Cisco Systems, Inc. is a multinational corporation worldwide leader in networking for the Internet. Cisco Internet Protocol-based networking solutions have become the foundation of the Internet and most corporate, education, and government networks around the world. Cisco routers provide solution techniques for monitoring links and for gathering data passing through the equipments. It provides an extremely broad line of solutions for transporting data, voice and video on IP networks. Cisco is the developer of the popular tool NetFlow for collecting IP traffic, and the memory usage saving solution Sampled NetFlow which uses sampling techniques not to waste a lot of memory in storing the collected data.

**DAG cards**

The DAG cards from Endace Measurement Systems are relatively low cost PCI-based passive monitoring cards. They are a combination of hardware design using FPGAs technology and software, based on a programmable chip. They are available in various models with interfaces for Ethernet, SONET, SDH or PDH. They ensure 100% packet capture regardless of interface type, packet size or network loading. The cards can be used as traffic generators, however the main use for the cards are passive monitoring where it captures the headers of each packet. A GPS receiver can be connected directly to the cards and a 64bit timestamp is appended to each captured packet. The precision of the timestamp is around 100ns.

More recent DAG cards feature the Endace Data Stream Manager (DSM) which provides hardware-based traffic filtering, data stream replication, and CPU load balancing functions. These effectively

---

1 Line Interface Modules
accelerate application performance, enabling Network Managers to be confident that every packet has been inspected and reported on, and that nothing has been missed.

**Fluke Networks OptiView Ethernet Pro Gigabit Integrated Network Analyser**

The OptiView is a handheld network tester for 10, 100 and 1000 Mbps Ethernet. It combines a number of network monitoring capabilities, such as testing network readiness for new applications, new technologies and new infrastructure deployment, including protocol analysis at all seven network layers, device and network discovery, SNMP device monitoring and RMON2 traffic analysis.

**InMon sFlow**

sFlow is a technology for monitoring traffic in data networks containing switches and routers. InMon produces several solutions based on sFlow for network monitoring, including:

- sFlow probe: An sFlow probe capable of continuously monitoring application level traffic flows at 1 Gbps.

Examples of the applications of sFlow data are:

- Detecting, diagnosing, and fixing network problems.
- Real-time congestion management.
- Understanding application mix and changes.
- Usage accounting for billing and charge-back.
- Audit trail analysis to identify unauthorized network activity and trace the sources of DoS attacks.
- Route profiling and peering optimisation.
- Trending and capacity planning.

**Netperf**

Netperf is a tool that can be used to measure networking performance. The main goal of this tool is to measure bulk data transfer and the performance of requests and responses using either TCP or UDP. The tests available include:

- TCP and UDP unidirectional transfer and request/response over IPv4 and IPv6 using the Sockets interface.
- TCP and UDP unidirectional transfer and request/response over IPv4 using the XTI interface.
- Fore ATM API.
- HP HiPPI Link Level Access.
- Link-level unidirectional transfer and request/response using the DLPI.
- Unix domain sockets.
- SCTP unidirectional transfer and request/response over IPv4 and IPv6 using the sockets interface.

Netperf works on several Operative System, including Windows, Linux and Unix.

**NIMI**

The NIMI is a software version system for building and deploying network measurement infrastructures. The infrastructures built with NIMI can be used to perform fault-diagnosis and performance measurements.

The objectives of the NIMI project are the study the particularities of Internet, and the evaluation of its current performance, in order to understand Internet behaviour and looking to improve its performance. It can be used for measuring network characteristics like, for example, packet loss rate, transmission delay and then get QoS parameters values for HTTP and FTP, for example. Furthermore, NIMI can be used to

---

2 Remote Network MONitoring
3 Denial of Service
4 X/Open Transport Interface
5 Data Link Provider Interface
6 Stream Control Transmission Protocol
7 National Internet Measurement Infrastructure
compare different implementations of IPPM metrics, such as bulk transfer capacity. Finally, it can track multicast connectivity, and monitor the performance of commercial and high speed research networks.

**Ntop**

Ntop is a passive network monitoring application based on libpcap. It is available both for Unix and Windows. It shows the network usage, similar to what the popular top Unix command does. Ntop captures simultaneously packets from one or more network cards and stores traffic information in memory.

Ntop users can use a web browser to navigate through Ntop traffic information and get a dump of the network status.

Traffic information is maintained on a hash table and kept sorted per host. Ntop is able to provide several information about captured traffic including the ability to:

- Sort network traffic according to many protocols
- Show network traffic sorted according to various criteria
- Display traffic statistics
- Show IP traffic distribution among the various protocols
- Analyse IP traffic and sort it according to the source/destination
- Display IP Traffic Subnet matrix and Report IP protocol usage sorted by protocol type.

**OC3MON**

OC3MON is a passive network monitoring machine. It collects network traffic from an OC3 link, providing it to the Coralreef suite for latter processing. It consists of a common personal computer equipped with interfaces for monitoring several protocols. It has two ATM network cards for monitoring both directions of the link, and can monitor Ethernet links too.

**PacketLogic (PRO)**

PacketLogic (PL) is commercial real-time hardware/software solution used mainly for traffic surveillance, traffic shaping or as a firewall. The PL has a statistics database that collects and stores data. It records the average amount of traffic in the inbound and outbound directions as well as the total traffic for all nodes in the network. The data is averaged over 5 minute periods.

PacketLogic covers three product segments – DFI (deep flow inspection), Network Traffic Management, and Service Management. DFI, which is an evolution of DPI (deep packet inspection), is a way to identify and classify the traffic in a network. PacketLogic's DFI engine DRDL™ has been rewarded unparalleled accuracy in identifying the application protocols. The identification is done port-independent based on a multitude of criteria, including the five-tuple, bi-directional information, payload, number of packets/bytes, etc. This makes it possible to relate multi-session protocols with e.g. control and data sessions like FTP, SIP and Direct Connect. Granular traffic details, e.g. filename, URL, chat channel, SIP Caller ID and direction, are aggregated during the identification process.

PacketLogic is a transparent Layer 2 appliance. This offers a lot of benefits like higher capacity, lower latency and better security (transparent as a cable). But it also simplifies installation since no network definitions have to be changed when deploying the system. This design also makes high-availability installations simple since standard L2 fail-over like Rapid Spanning Tree (rSTP) can be used. The system will forward all traffic by default.

The active user and application traffic properties are highly detailed and granular, thus ensuring accurate identification of users and applications. The Statistics module offers the ability to search by connection during a defined time interval – by application protocol, destination, origin and many more criteria for each user and each application.

Traffic identification is based on packet content (deep packet inspection and deep flow inspection) instead of port definitions. PL uses the self-developed Datastream Recognition Definition Language (DRDL) to identify different application protocols. The PL can identify more than 700 application protocols.

The identification process is connection-oriented, which means that each established connection between two hosts is matched to a certain application protocol. When a new connection is established, usually detected by some hand shaking procedure, the identification of this connection begins. The identification algorithm searches for specific patterns, called application signatures, in the connection. The patterns are found in the IP header and application payload. The identification algorithm uses a tree-
like structure of patterns. When the identification starts, the algorithm is in the root of the tree. When certain patterns occur in the traffic, the algorithm moves down in the tree. When it reaches one of the leaves, the full identification of the connection is completed.

Most connections are bidirectional. The PL uses the traffic in both directions in the identification process. Also, different types of connections need to be tracked during different periods of time. For example, a Bit Torrent connection only needs to be tracked when it is established. Once it is identified as a Bit Torrent connection, the tracking of the connection can stop. However, an HTTP connection needs to be tracked until it is finished, since HTTP may be used to tunnel other application protocols, i.e. some file-sharing applications and VPNs.

**PingER**

PingER (Ping End-to-end Reporting) is a project for active measurement of IP-based networks. It is an IEPM (Internet End-to-end Performance Measurement) project focused on monitoring end-to-end performance of Internet links. The PingER project is mainly related to network performance between laboratories, universities, etc. collaborating on experiments.

The PingER analysis defines five metrics: Packet loss, RTT, unreachability, unpredictability, and quiescence. The measurement tools are simple and broadly available (ping, traceroute); they use ICMP packets allowing a large number of network nodes to take part in measurements but this may also cause troubles since they are exposed to network hackers. Therefore a potential disadvantage for the project might rise since an increasing number of network nodes could either process them at a lower priority level or not process them at all. Thus the metrics are in some cases usable for the ICMP flows only and real application performance is thereof hard to derive. The measurement results have been accessible via a web interface in hourly, daily, and monthly reports since the project launch.

**RIPE TTM**

The RIPE NCC is an independent, non-profit membership organisation that supports the infrastructure of the Internet through technical co-ordination in its service region. Its Test Traffic Measurements (TTM) Service measures key parameters of the connectivity between a site and other points on the Internet, in particular routing vectors, one-way delay and packet-loss, IPDV, bandwidth. TTM makes use of equipments dedicated to generate a small amount of traffic. TTM metrics and methodologies comply with current standards in RFC2330 and RFC2678 through RFC2681.

**Sting**

Sting is a TCP-based network measurement tool. It measures the packet loss rate between a source host and some target host on both the forward and the reverse paths using TCP messages. It differs from other related tools, like ping or traceroute, in that the latter ones do not distinguish between the forward and the reverse path. Another big difference is that Sting only needs that the host runs a TCP-based service, such as a Web server. Besides, Sting has been tested with a large number of different operating systems like Windows 95/NT, Solaris Linux, FreeBSD and AIX.

**Surveyor**

Surveyor is a measurement infrastructure that is based on standardized work carried out by the IETF’s IPPM WG. Surveyor measures the performance of the Internet paths among participating organizations. The project is also developing methodologies and tools to analyse the performance data. Surveyor currently measures end-to-end unidirectional delay, packet loss, and route information along Internet paths. It is deployed at about fifty higher education and research sites around the world.

**Top Layer IDS Balancer**

The Top Layer IDS Balancer is an equipment very useful when dealing with high speed links. It allows splitting a 1 Gbps line into several 100 Mbit lines so that several lower speed monitoring equipments can be used to monitor it.

### 1.2.2 Data analysing tools

**CoralReef**

CoralReef is a suite of Internet traffic data collection and analysis tools. It needs not any additional network infrastructure and does not increase the traffic inside the network because it is a passive...
monitoring system. It neither interferes with any existing equipment within the network due to the same reasons. When dealing with optical networks, an optical splitter is used to divert a small fraction of the fiber content to the monitor device. It is compatible with a large number of protocols and can manage and process the relevant information of the data. It supports system networks interfaces from libpcap, and analysis can be done from real time data or from trace files. It includes programming APIs for C and Perl, and applications for analysis, capture and web report generation.

**Skitter**

The skitter tool facilitates the active collection of IP-level topology data on the Internet, by using an ICMP based tool to probe the Internet from 19 hosts around the world, allowing collection of a "snapshot" of a network path traversed between a source and a destination. It measures forward paths and RTT and tracks persistent routing changes. Skitter can also return such information as Intermediate RTT and ICMP errors encountered during a measurement.

**NeTraMet - a Network Traffic Flow Measurement Tool**

NeTraMet is a traffic monitor that provides interfaces to traffic collection and to collect packet and flow data. It is the first implementation of the RTFM Working Group's Measurement Architecture. This is sketched in RFC 1272, "Internet Accounting Background," and has three components:

- **meters**, small hosts that measure traffic flowing on a segment
- **meter readers**, which retrieve information from meters
- **managers**, which instruct meters as to which flows they should measure and meter readers as to which meters they should collect from.

It exports collected data via SNMP. NeTraMet is configurable using a standardized programming language defined by the IETF. It can perform analysis such calculating RTT and packet size distributions. NeMaC, a combined manager and collector program, is supplied with NeTraMet. It downloads rules to meters, and collects data from them.

**AutoFocus**

AutoFocus is an offline passive tool for analysing and visualizing traffic over a link. The description of the traffic mix in a link is carried out by textual reports and time series plots that are computed automatically. The traffic is clustered by a pentuple, composed by the source and destination IP addresses, the source and destination port addresses and the protocol field. To get the description of a traffic mix, the user has to selected the corresponding traffic clusters. There are separate reports that measure the traffic in bytes, packets and flows. RRDtool is used to plot the time series of the traffic mix graphing each category in a different colour. The time range that AutoFocus produces reports and plots is from half an hour intervals to several weeks. It accepts two types of input: packet header traces and NetFlow data. Both types of input can be sampled, but AutoFocus only compensates for the sampling in the reports that measure the traffic in bytes and packets, and not in those measuring the traffic in flows.

**Beluga**

Beluga is a tool for real time graphing RTTs and packet lost of an end host, dividing RTTs in total RTT and per-hop RTT.

**Flow-tools**

Flow-tools is a library and a collection of programs used to record, filter, print and analyse flow records from NetFlow enabled devices. The flow-tools library provides an API for development of custom applications for NetFlow. Reports are limited to simple text but many of the reports can be handled by tools like gnuplot to generate graphs. A Perl and Python interface have been contributed and are included in the distribution. All the applications run under most UNIX like operating systems. They use pipelines to communicate between them; therefore the use of intermediate files is not needed. Besides, zlib compression is supported to save space.

Flows are exported from a router in a number of different configurable versions. A flow is a collection of key fields and additional data. The flow key is \{srcaddr, dstaddr, input, output, srcport, dstport, prot, ToS\}.

---

8 Realtime Traffic Flow Measurement
The most common architecture when using Flow-tools is to have multiple PCs located near the Netflow enabled devices. These PCs store the flow records until they are periodically reported to a centralized computer that will perform the analysis and archive the flow records.

**cflowd: Traffic Flow Analysis Tool**

Is a flow analysis tool for analyzing Cisco Netflow information. It is composed of modules for collection, storage and analysis in the C++ language and with an available library to make custom applications. It allows data collection, analysis by ISPs in support of capacity planning, trends analysis and characterization of workloads in a network service provider environment.

There are three major components in cflowd:

- **cflowdmux.** - Is the program that acts as the receiver of flow-export data from one or more Cisco routers. It writes raw packets into shared memory and allows clients to have access to this data.
- **Cflowd.** - Create summaries of these raw data and stores them in tabular forms containing:
  - AS matrix
  - Net matrix
  - Port matrix
  - Interface matrix
  - Nexthop table
  - Protocol table
- **Cfdcollect.** - All the information from cflowd is gathered by cfdcollect, storing it in regular intervals and producing time-series data.

Besides, cflowd has utilities that help to run and maintain systems when it is running.

Other areas where cflowd may prove useful include usage tracking for Web hosting, accounting and billing, network planning and analysis, network monitoring, developing user profiles, data warehousing and mining, as well as security-related investigations.

**FlowScan - Network Traffic Flow Visualization and Reporting Tool**

FlowScan analyzes and reports on Internet Protocol flow data exported by routers. It was written in Perl language and produces graph images that provide a continuous, near real-time view of the network border traffic. The suite contains three modules:

- a flow collection engine (a patched version of cflowd)
- a high performance database (Round Robin Database – RRD)
- a visualization tool (RRDtool)

The notion of flow profiling was introduced within the network research community in order to better understand the nature of Internet traffic. Flow profiling offers a pragmatic compromise between having too many or too few collected data.

FlowScan collects and processes raw flows exported from routers, analyzes and reports on NetFlow format data collected using CAIDA’s cflowd flow tool. FlowScan examines each flow and maintains counters based upon that flow’s classification. Counter values are stored using RRDtool, a database system for time-series data. Finally, FlowScan uses visualization capabilities of both RRDtool and other front-ends to report on the processed flow data, and it may be configured to either archive or discard the raw flows after processing.

**Cricket**

Cricket is a time series visualization tool intended for network managers. It is available under the GNU GPL license. It was expressly developed to help network managers to visualize and understand the traffic on their networks. It regularly polls data from various sources and stores them in a round robin data store. Graphs are produced on demand and are displayed in a browser.

Cricket is configured from a set of config files called config tree. The config tree sets what types of data have to be collected, how to get it and from which targets it should collect data.

---

9 General Public License
It is written entirely in Perl language and can be configured to send email messages when certain conditions set by the user are detected.

**Wireshark**

Wireshark is a wide spread freeware tool that provides network protocol analyzing functionalities for Unix and Windows. It was formerly known as Ethereal, but due to trademark issues its name was changed. Wireshark offers the ability to examine data from live network both, in-line and off-line. Captured data can be stored on a hard disk and reloaded to the application for further analysis and examination. It uses libpcap for capturing packets on any available interface and storing the captured information.

It also offers the ability to browse the captured data, perform some statistics for the data stream and to investigate in detail each single packet. It has a wide range of filters that can be applied over the captured data to improve the browse experience, and also it is possible to apply own defined filters.

**OpenView**

OpenView was a family of products developed by Hewlett Packard to provide complete management solutions to business networks in many different levels. In 2007 the entire HP OpenView portfolio was rebranded under the strengthened HP Software name.

OpenView products offer capabilities to monitor resource usage among the different network elements. There are also products that monitor network traffic and report measurements with regard to specific services. The data is managed in an efficient way and a large number of results are produced in various formats. Some of the reports may simply show a network failure or may provide data to support functions such as monitoring and accounting.

**Ipanema QoSma**

QoSma is a commercial system distributed by Ipanema Technologies to manage application service level, monitoring and optimising IP networks. It is based on points of measurement across the network with central management. The QoSma consists of three elements:

- **ipengine**, a proprietary hardware with an interface for GPS receiver and another optional Ethernet port for out-of-band management.
- **ipagent** is software running on ipengine. It entrusts traffic measurement, traffic organization and time synchronization. It can be measured one-way delay, packet lost, jitter, bandwidth and throughput.
- **ipbossthe** is the central manager of the architecture and runs on central server. It handles the whole system, provides data collection and storage and generates reports to the user. It includes also interface with other network management tools via SNMP that has made it very popular.

**Nettimer**

Nettimer is a software tool to measure the bandwidth of a bottleneck link along an path. Bandwidth is measured by passive techniques both in the forward and backward direction. It also can monitor the Internet usage and log this information, allowing monthly reports and filtering the logs by date and/or connection's names.

To perform bandwidth measurements in both directions, forward and backward, two capture hosts are needed. Measurements are based on the time interval between reception of successive packets sent back to back from the sender. Such measurements are carried out in real time. The bottleneck bandwidth is the smallest bandwidth of all link's belonging to the path, and it is a key measure for any application performance.

**snort**

snort is a lightweight free software open source NIPS\(^{10}\) and NIDS\(^{11}\). It uses the libpcap packet capture library and can perform packet logging and real time traffic analysis on IP networks. It can be used as a packet sniffer or packet logger too, like Wireshark.

It uses a rule-driven language to perform protocol analysis or content searching/matching and is commonly used for blocking or detecting a variety of attacks, such as buffer overflows, stealth port

---

\(^{10}\) Network Intrusion Prevention System

\(^{11}\) Network Intrusion Detection System
scans, web application attacks, SMB probes and OS fingerprinting attempts. This software is mostly used for intrusion prevention purposes, dropping attacks as they are taking place. There are also several user interfaces available to help using snort software, like Sguil or Sourcefire.

**Windmill**

Windmill is a tool for network performance measurement. Its particularity is that this tool tries to correlate lower-layer protocol events with the behaviour of the application-level protocols. The way Windmill achieves it is by mimicking itself as the end host of each packet that belongs to the stream of interest, which can be specified by the users. Windmill executes the lower-layer protocols every time a packet arrives. Moreover, each user can define experiments for specific traffic characteristics and Windmill will dynamically delete and insert experiments and monitor traffic under real time conditions.

### 1.2.3 Routing Repositories

The challenge of inter-domain scenarios lies in the fact that each domain is under independent management. Even in the case of collaboration between ISPs, key network policy aspects are kept under secret, as they are the source of differentiation between suppliers and therefore, part of ownership.

Any Quality of Service analysis tool relies on the existence of network state information for a network segment in a time period around the precise event. Two of the tools that rely on different data repository types are: InterRoute, that works with public BGP-4 routing data repositories, and the INTERMON database, whose topographical components build a private measurement repository that is, in fact, a measurement multidisciplinary repository available to all INTERMON tools.

One of the main INTERMON project lessons was the huge amount of network measurement data needed to investigate sophisticated network phenomena. Collecting these data for each new research is highly inefficient. But, since there are so many projects which produce and rely on network measurements, the logical consequence is trying to connect them all and allow their measurements to be shared. This is the goal of MOME group. MOME connect logically a significant number of projects that are making Internet measures and using those measurement data. The MOME synergistic effects reduce the redundant measurement number made in Internet and therefore the stress produced by the measured data is reduced.

Finally, the huge routing data amount would also help enhancing the development of the new IPv6 Internet. Simulations that implement these improvements can show their ability to cope with actual network situations and assess the benefits of its development during the design phase.

#### 1.2.3.1 Repository Structure

Topological data and BGP-4 protocol specific data are normally in this category and they are considered to be confidential by Service Providers. Rarely partial views are shared and publicly available, obtained from Internet Exchange Points by projects like Oregon Route Views and RIPE-NCC Routing Information Service. Other topological data can be obtained by means of network topography, but these data rely in traceroute traffic that is normally blocked by Service Providers for security reasons.

Repositories store data in the form of global Internet routing tables, obtained through collaboration and routing tables exchange for a large number of AS, with which it is possible to obtain the network topology.

There are different Data Organizations that store routing information:

**Oregon RouteViews:**

- 5 Collection points, multiple peers at each point
- RIB files logged roughly every 2 hours
  - Ex: rib.20051024.1602.bz2
- Update files logged roughly every 15 minutes
  - Ex: updates.20051024.1906.bz2

**RIPE RIS**

- 15 collection points, multiple peers at each point
- RIB files logged every 8 hours
  - Ex: Bview.20051024.1600.gz
• Update intervals every 5 minutes
  o Ex: updates.2005.1024.2130.gz

Besides, a few minor changes have been produced over the time like:
• Change filenames to UTC in 2003 (RV)
• Change in update interval times (RIPE)

Then, even if they are minor changes, they produce enough differences to make external management and monitoring scripts rather complicated. Thus, a uniform data organization has been proposed to get the following requirements:
• Uniform data interface where user specifies monitoring points, collection points, peers and time
• Tool retrieves corresponding data
  o Time period need not match RIB/Update dump times
  o Produces a starting RIB file from time period start
  o Retrieves all update files for the time period
  o Converts all times into UTC
  o Breaks data out by peer
• Example
gedata –m rv –c oreg –s 2005.09.03.162301 –e 2005.10.23.193806
• The resulting File Structure is
  mp/cp/peer/year.month
  o Top level directories specify monitoring point (mp)
  o Second level directories specify collection point (cp)
  o Third level directory specify peers
  o Fourth level directory is year and month
• Data Directory contains daily RIBs and Updates
  o updates.YYYY.MM.DD
  o rib.YYYY.MM.DD
• Example:
  o rv/oreg/12.0.1.63/2005.10/rib.2005.10.01.gz
  o rv/oreg/12.0.1.63/2005.10/updatos.2005.10.01.gz

This is a more consistent and easy to use structure for scripts.

1.2.3.2 Repository applications

BGP-4 Routing Data

BGP routing repositories have been established for projects such as RIPE-NCC and Oregon Routeviews. Data is difficult to handle and process. An interface with a relational database is needed to focus on routing research rather than on algorithms to analyze routing data.

The data stored in routing repositories can be classified using the following criteria:
• Address family (for example IPv4 vs IPv6)
• Routing Registry type:
  o Registries dump in the routing table
  o Announcement
  o Remove
  o State change for a router

INTERMON application focuses on IPv4 routing related registers, is:
• IPv4 routing table registries dump
• IPv4 prefix announcement
• IPv4 prefix remove

Those three routing registries share a common structure that is completed in the routing table announcement and dump cases with additional data sets and both share this structure.

**BGP Policies. Internet Routing Registries**

The robustness of the Internet relies heavily on the robustness of BGP routing. BGP is the glue that holds Internet Autonomous Systems together: It is the common language of the routers that interconnect AS. The robustness of BGP and our ability to manage it effectively is hampered by the limited global knowledge and lack of coordination between AS. One of the few efforts to develop a globally analyzable and secure Internet is the creation of the Internet Routing Registries (IRRs). IRRs provide a voluntary detailed repository of BGP policy information. The IRR effort has not reached its full potential because of two reasons:

• extracting useful information is far from trivial, and
• its accuracy of the data is uncertain

There is a tool called Nemecis (Network ManageMent and Configuraion System) to take the IRR to the next level in order to extract and infer useful information from IRR, with the ultimate goal to use this information to model, manage and protect Internet routing.

There exist a number of tools that helps measuring actual BGP routing (see 1.2.1). But a tool to bridge the gap between intended policy (configuration) and actual routing does not exist. Internet Routing Registries (IRR), contain the policy of a large number of networks, expressed in a high level language, RPSL. These registries are considered to be useless and outdated, based primarily on empirical evidence. Thus, as far as we know, a tool that can analyze these policies and check their validity or freshness does not really exist.

Furthermore, the registries are maintained manually and in a voluntary basis to a large extent, the policies remaining as simple text. Then, analysing IRR is not a trivial task. Main difficulties are:

• RPSL is very flexible, so policies can be very complex,
• there can be many different ways to express the same policy,
• the registries can contain inaccurate, and incomplete data.

However, the information of IRR is important to understand and interpret Internet routing, since routing tables are not sufficient to understand the intended policies.

Despite recent efforts, we do not have a complete model for BGP, and the problem of robustness has not been solved. Several measurement studies show that the Internet behaviour is not fine-tuned and it can be destabilized by cascading effects because

1. Internet routing relies, to a large extent, on trust. Network nodes usually assume that the information received from their neighbours is correct. This is not always the case, since misconfigurations, bugs in the software of the routers and malice may occur.
2. There have been very few (public) efforts to automate monitoring and management of BGP.
3. The issue of secure and robust Internet routing is an open problem despite significant research efforts and studies.

How can we automate the management of the Internet and set and its safety? This is the general problem we attempt to address. Within the framework of MOMENT project, we focus on three major problems in BGP research and its management:

• The lack of detailed knowledge and accurate models for Internet at BGP level
• The lack of tools to analyze the configuration of an Autonomous System, and to check whether the registered policy matches the intended policy
• The requirements for an automatic mechanism to detect abnormal routing behaviour.

**Nemecis** is a methodology and a tool for addressing these issues. Its goal is providing a framework for the analysis of RPSL policies. It can be used either during the configuration phase, or in the operation phase: During the configuration phase it can check the registered policy for correctness. During the operation phase, it can check whether the intended policy matches the actual routing. This way, Nemecis
can reduce the time to discover and fix routing problems. Most importantly, it can start monitoring how Internet routing works. In fact, this tool is among the first public tools to analyze the IRR policies.

RIPE, has as a long term goal to validate the policies that AS register, and thus increase the robustness of BGP.

1.3 Challenges for bottleneck analysis

1.3.1 Last-mile problems

The strong development of broadband communications during last decades has been provided by DSL access technologies, specially in Europe. These technologies allow one to create high-speed digital links to end-customers over the copper wires that once were used exclusively for plain phone service (POTS). The most common DSL transmission technologies deployed today are ADSL2 and ADSL2plus (ADSL2+).

To deliver triple play services (voice, data and video), operators can still use xDSL technologies instead of investing huge amounts in deploying fiber to the home (FTTH). This set of services, based on the convergent IP layer, does not only demand a high bit rate, but also puts strong constraints on the bit error rate, delays and jitter, as QoS parameters, specially considering the imminent introduction of the VoIP service: The quality of a streaming video service is quite sensitive to jitter and bit errors and real time services, like voice, gaming, etc. require robust network with rapid mechanism to solve failures and avoid long delays.

Then triple play success can not only be based on over dimensioning metro and access networks but special care should be taken to guarantee QoS by watching bottlenecks and avoiding misconfiguration problems. Besides a good routing control allows the operator to set traffic engineering mechanisms as tools for saving network resources while keeping QoS.

In addition to the change of the type of the services, also the transmission technologies themselves have changed. The evolution path from ADSL, via ADSL2+ to VDSL2 meant a substantial increase of the used physical bandwidth on the telephony twisted pairs. This increase on itself invited a whole spectrum of new effects from disturbances.

Over the last couple of years, it has become apparent that a substantial percentage of the xDSL links (say 5%, which implies very large absolute numbers) are sensitive to all sorts of external noise, leading to bit error problems. In contrast to internal (crosstalk) noise, external noise is the noise that originates from outside the cable. An example of this is the repetitive electrical impulse noise (REIN) stemming from faulty or poorly designed electric devices that can be found at the Customer Premises side operating with AC line voltage. Another example is (time-varying) radio frequency interference (RFI) from AM broadcast stations. A short description of all these (physical) disturbances is included below:

- **Crosstalk.** Crosstalk refers to the noise interference between channels. In the DSL world, the interference between nearby cables can have a negative impact on the performance of the affected cable(s). An example of crosstalk is when you hear another voice in the background when having a telephone conversation.

- **Repetitive Impulse Noise (REIN).** The repetitive electrical impulse noise (REIN) can affect communication systems and can comprise a plurality of impulses, each impulse having a duration, wherein impulses are separated by regular time periods or intervals. Electric power lines carrying AC power are one mechanism that can cause REIN, wherein the REIN impulses typically appear with twice the frequency of the AC power.

- **Radio Frequency Interference (RFI).** These interferences are cause by powerful external sources which induce noise in the DSL line. This interference in Digital Subscriber Line (DSL) networks is of particular concern in urban areas with many strong radio sources nearby. Examples of these radio frequency sources are: amateur (Ham) radios, AM broadcast stations and RF heating systems.

These external disturbances are difficult to predict and affect the DSL lines very differently. In many cases, those disturbances are not even permanent which make them difficult to detect. However, they could cause important problems when it comes to service degradation, being the first bottleneck in the network.
1.3.2 Congestions and bottlenecks caused by inter-domain routing issues

**BGP Routing – State of the Art**

The Border Gateway Protocol (BGP) is the routing protocol used to exchange routing information across the Internet. It makes it possible for ISPs to connect to each other and for end-users to connect to more than one ISP. BGP is the only protocol that is designed to deal with a network of the Internet's size, and the only protocol that can deal well with having multiple connections to unrelated routing domains.

BGP has proven to be scalable, stable and provides the mechanisms needed to support complex routing policies.

The Border Gateway Protocol is an inter-Autonomous System routing protocol. The primary function of a BGP speaking system is to exchange network reachability information with other BGP systems. This network reachability information includes information on the list of Autonomous Systems (ASs) that reachability information traverses. This information is sufficient to construct a graph of AS connectivity from which routing loops may be pruned and some policy decisions at the AS level may be enforced.

BGP4 provides a set of mechanisms for supporting Classless Inter-Domain Routing (CIDR). These mechanisms include support for advertising a set of destinations as an IP prefix and eliminating the concept of network "class" within BGP. BGP version 4 also introduces mechanisms which allow aggregation of routes, including aggregation of AS paths.

Routing information exchanged via BGP supports only the destination-based forwarding paradigm, which assumes that a router forwards a packet based solely on the destination address carried in the IP header of the packet. This, in turn, reflects the set of policy decisions that can (and can not) be enforced using BGP. BGP can support only the policies conforming to the destination-based forwarding paradigm.

A unique AS number (ASN) is allocated to each AS for use in BGP routing. The numbers are assigned by IANA and the Regional Internet Registries (RIR), the same authorities that allocate IP addresses. There are public numbers, which may be used on the Internet and range from 1 to 64511, and private numbers from 64512 to 65535, which can be used within an organization.

AS numbers are currently 16-bit integers, but deployment of 32-bit AS Numbers (4-Byte ASN) is starting because of expected ASN shortage in the 16-bit range around the year 2010.

**BGP Routing Bottleneck**

The primary responsibility of an Internet Service Provider (ISP) is to provide transit service from its set of customers to the remainder of the Internet and to bring traffic from its own upstream providers and peers destined for its customers. The interface from the ISP to the customers, upstream providers, and peers is through a set of border routers of the ISP.

This responsibility is balanced with an objective of the ISP to minimize the resources used on its network in carrying transit traffic. The ISP wishes to get traffic “on its way” toward its ultimate destination as quickly as possible.

A poorly designed selection of border routers for the flows of traffic through the ISP can result in numerous problems. On one hand, ingress and/or egress traffic from/to neighbours may exceed the capacity of the selected border routers and its links, causing the ISP to fail to meet its responsibility. On the other hand, underutilization of the potential capacity at border routers, or carrying traffic across the ISP network longer than necessary, results in inefficient use of costly resources of the ISP. However, there is anecdotal evidence to suggest that the peering links at the border routers are often bottlenecks in the Internet, so it is important that these links be utilized efficiently.

ISPs today have few tools or algorithms to deal with this problem. Policies governing inter-domain routing and border router/edge link selection are carried manually by means of pure intuition, ad-hoc methods and constant tuning.

The objective is to determine an optimal selection of outgoing links and their border routers to be used for the egress of transit traffic where the selection minimizes the provider’s network utilization and balances the load of traffic flows exiting the service provider across the selected egress links by respecting capacity constraints. By respecting capacity constraints, the goal is to optimize the utilization of these outgoing links too.
1.3.3 DSL measurement tools applications

As shown in previous section, the DSL lines could become an important bottleneck in the network. Due to the varying characteristics of the DSL lines and of the disturbances on the lines, measurement and diagnostics tools become really important for network operators.

1.3.3.1 Ericsson Loop Qualification and Monitoring (LQ&M) tool

Loop Qualification and Monitoring (LQ&M), designed by Ericsson, is a concept that today includes the assessment of a loop characteristics and capability and the verification of the correct provisioning and operation of the overlaying broadband service(s) mainly by means of measurements, in dynamic way. It follows and respects the standards of the different ADSL/2/2+ (and soon VDSL2) technologies and it is designed to be naturally integrated into the operators' workflow.

Accordingly, the LQ&M tool includes the following modules:

- **Network Overview Tool (OVT):** It gathers DSL network statistics of status and performance. It derives statistical information on different hierarchical network levels (line, bundle of lines, nodes, regions, whole copper plant). It provides the current status and performance information to the operator and triggers the identification of problematic parts of a DSL network.

- **Fault Localization Tool (FLT):** It manages line testing for detailed troubleshooting of specific lines. It contains SELT, DELT, noise monitoring and DSL primitive monitoring. This helps operators troubleshooting lines, in a scheduled or real-time manner.

- **Service Behavior Tool (SBT):** It monitors the quality of services that runs over DSL. Physical QoS-related line primitives are monitored and alarms are fired if lines misbehave. Problematic lines are sorted and pinpointed by the tool for further investigation (by e.g. FLT).

- **Service Activation and Delivery Tool (SAT):** It supports DSL line installation and upgrading procedures. For installation purposes, SELT on several lines support the deployment process (installation). The line capability estimations support upgrading decisions taken to deploy new services on specific lines. DSL Automatic Optimization (DAO): It finds the optimum configuration for a line in terms of rate and delay, line stability and noise protection. This tool optimizes the usage of the copper plant resources on a per line basis.

1.3.3.2 Alcatel 5530 Network Analyzer

The Alcatel 5530 Network Analyzer (NA) is an advanced diagnostic software tool that automates the process of line quality assessment by using expert diagnostics to continuously monitor service quality at the physical layer, during pre-qualification and in-service.

For the pre-qualification the NA offers Single-ended line testing (SELT) and dual-ended line testing (DELT) capabilities. These features allow tests to be performed on the copper line to ensure that new services (like video) can be delivered effectively and reliably — before they are offered to the consumer.

Dynamic line management (DLM) is an in-service profile management technique aimed at minimizing customer complaints owing to line instabilities through dynamic profile adaptation. DLM increases line stability through supervision of the installed base for instabilities and automatic profile change. This change is done along three dimensions:

- Transmit PSD: on short loops to minimize crosstalk
- Interleave/Fast: to improve robustness to noise vs. latency
- Rate

1.3.3.3 DSL Expresse™

ASSIA inc.’s DSL Expresse™ Spectrum Management Center software implements Dynamic Spectrum Management (DSM) Level 1 functionality to automatically monitor and optimize entire DSL networks. Using the management interface described in the ATIS Dynamic Spectrum Management TR (DSM TR) and ASSIA's unique and patented expert systems algorithms; DSL Expresse™ dynamically adjusts line parameters to improve DSL performance with no interruption of customer service. It also automatically identifies and classifies line troubles, enabling proactive response by the service provider before the customer reports trouble and improved efficiency of field personnel. Integrated into the DSL Provider's Operations Support Systems (OSS) and Customer Care Centers with a flexible North Bound Interface, DSL Expresse™ supports the service provider's specific workflows, loop inventory and service models.
Optimization of DSL services provided by any DSLAM or Remote Terminal is supported by DSL Express™'s extensible DSL network element interface.

2 RELATED PROJECTS

2.1 IST projects

2.1.1 INTERMON

In order to enhance the inter-domain QoS analysis, in large-scale, multi-domain Internet infrastructures, the project INTERMON developed and demonstrated a scalable inter-domain QoS architecture. It included integrated monitoring, topological and geographical structure mapping, modeling, simulation, optimization and visual data-mining components using a common distributed QoS database with intelligent agents for management of components interworking and automated processing of different kind of inter-domain QoS information (inter-domain QoS, traffic, resource, events). The focus is to offer to Internet service providers (ISP), QoS enabled end system developers and application users, an integrated inter-domain QoS analysis architecture for the purpose of operative control, planning and optimization.

The aim is a novel scalable inter-domain QoS analysis architecture for large interconnected Internet domains (DiffServ, MPLS). The inter-domain QoS architecture is validated in field trials involving multiple QoS based domains.

The work description is:

- Definition of inter-domain QoS analysis framework for test and verification of inter-domain QoS of application classes (e.g. Internet telephony, videoconferencing, etc) in large interconnected QoS based domain infrastructures.

- Design of scalable inter-domain QoS analysis architecture with integrated monitoring, topological and geographical QoS structure mapping, modeling, simulation, optimization and visual data-mining components using common distributed QoS database.

- Implementation of a prototype of integrated inter-domain QoS analysis architecture:
  - Distributed inter-domain database with intelligent agents for management of tool interworking and automated information processing
  - Integrated inter-domain QoS monitoring toolkit including passive monitoring, active probing tools and event detection tools
  - Modeling toolkit including algorithms and tools (long term trend and short term experiments) for prediction of inter-domain QoS for application classes, network-performance models, event patterns, traffic profiles
  - Automated mapping of Internet topological and geographical structures including options for mapping of inter-domain QoS, traffic, resource, events and network-performance issues to autonomous systems (BGP routing), countries, equipment (multiple IP addresses map to the same router), and geographic location information (latitude/longitude co-ordinates)
  - Simulation environment for inter-connected QoS based domain (MPLS, DiffServ) infrastructures and QoS enabled application classes combined with optimization toolkit
  - Visual datamining environment (models, tools and advanced multidimensional techniques) for specific kind of inter-domain QoS and traffic analysis.

2.1.2 MOME Database

The MOME cluster offers a platform for knowledge and tool exchange and for co-ordination of activities in the field of IP monitoring and measurement between FP5 and FP6 IST projects and other European partners. Many projects include monitoring and measurement components for traffic engineering, billing and accounting, security, etc. To avoid double work/funding and to improve the strength by synergetic knowledge transfer, common activities in-between these projects are required.

This is advanced by interoperability testing of monitoring and measurement components of different projects and afterwards by collecting monitoring and measurement data in a common format in a measurement database. Collected data is made accessible for the community over the Internet via simple interfaces for further work, e.g. mathematical and statistical analyses. Monitoring and measurement related contributions to standardization bodies like IETF from the participating projects are coordinated by
the MOME cluster. The activities are supported by the organization of public workshops and conferences within the field of monitoring and measurement.

- **Measurement data**: Metadata to measurement data including links to the data sources are available from the MOME database. Once metadata is available, data analyses can be requested by the submitter at the MOME platform. There are different data types and several tools for capturing them:
  - Data types: packet traces, flow traces, QoS measurement results, HTTP traces and web based repositories
  - Tools: libpcap, tcpdump, DAG Gard, netflow, sflow, IPFIX, CMToolset or by collecting web server.

- **Measurement tools**: Measurement Tools stored in the MOME database are catalogued and tested for interoperability with other tools regarding their data and communication formats.
  - There are measurement tools to packet capturing, traffic flow, packet monitoring, application level monitoring, service monitoring, accounting, intrusion detection, sniffing, performance monitoring, connectivity checking, routing detection, topology detection, traffic visualization, traffic generation and bandwidth measurement.

### 2.1.3 LOBSTER for traffic network passive monitoring

Network traffic monitoring and measurement is increasingly regarded as an essential function for understanding and improving the performance and security of the cyber-infrastructure. With networking technologies and services evolving rapidly, as witnessed by the explosive growth of the World-Wide Web, peer-to-peer networks, and GRID, accurate network traffic monitoring is required to ensure the security and optimize the efficiency of the cyberspace.

Funded under 6th WFP, the main goal of the LOBSTER Project is to deploy a pilot advanced European Internet Traffic Monitoring Infrastructure based on passive monitoring sensors at speeds starting from 2.5 Gbps, and possibly up to 10 Gbps. Such an infrastructure serves as a catalyst that boosts the understanding of the Internet and leads to its better use in the long-run. Passive monitoring at such high speeds stresses significantly the computational, communication, and storage capabilities of the underlying monitoring sensor and poses several interesting research challenges.

Using the monitoring applications developed within the project, researchers and administrators are able to monitor the Internet traffic for gaining a better understanding of its performance, as well as to spot security incidents.

LOBSTER is based on passive network traffic monitoring that records all IP packets (both headers and payloads) that flow through the monitored link. Within the objectives of the project are:

- Deploy a pilot Internet Traffic Monitoring Infrastructure across Europe
- Organize stakeholders in the area of advanced Internet traffic monitoring
- Realize the appropriate data anonymizing tools that will prohibit unauthorized tampering with the original traffic data
- Develop novel applications enabled by the availability of the passive network traffic monitoring infrastructure
- Provide anonymized data traffic information on a regular basis

### 2.2 Routing repositories

Routing repositories provide historical memory to the Internet’s backbone where data is stored as in a router's routing table. Many useful tools have been developed around the information provided by the routing repositories:

- The **RIPE RIS** project has developed tools to check routing consistency
- A **whois** variant uses the most recent routing table snapshot to determine the AS identifier of the domain of a given IP address
- **BGPlay** shows the evolution in time of the connectivity to a certain network prefix
- **INTERROUTE**
  - It is a result of the IST-INTERMON project. It uses a tool to detect complex routing event patterns like human interventions, flapping …
It is used to establish the relationship between QoS outliers and routing events

2.3 CAIDA work on Internet Measurement Data Catalogue

Internet data remains one of the basic components of computer science network research. Despite its necessity, available data is limited by legal, social, and technical constraints on its collection and distribution. Thus, optimal distribution of knowledge about available data is a valuable service to the research community. To this end, CAIDA (see http://www.caida.org/home) has developed the Internet Measurement Data Catalog to:

- Provide a searchable index of available data
- Enhance documentation of datasets via a public annotation system
- Advance network science by promoting reproducible search

While the resource, legal, and privacy concerns limiting new Internet data collection efforts remain largely intractable, significant research could be promoted through more widespread use of existing data. To that end, CAIDA began developing an Internet Measurement Data Catalog (IMDC) – an index of existing datasets possibly available for research.

In addition to the obvious utility of locating datasets relevant to research projects, a large data catalog provides many other benefits to Internet research:

- The focus on specifically indexing data provides a forum for robust documentation of data collection procedures
- An independent repository of dataset information allows researchers sufficient space to describe the data collection process, resulting in better documentation of a dataset’s strengths and weaknesses that is accessible to paper reviewers and future users of that dataset.
- Researchers are able to find datasets that are relevant to the topic they wish to investigate – new scopes of research become possible

CAIDA Internet Measurement Data Catalog provides the research community with a persistent source of information about available Internet data. A flexible annotation system allows the catalog to incorporate pertinent details about current and future data types, as well as recording user comments and corrections for data objects. The benefits of the IMDC include encouraging users to make more datasets available, enhancing documentation of data collection methodologies and their impact on subsequent research, providing unique identifiers for data used in research studies, expanding the scope of research studies to include more comprehensive data across longer timescales, and perhaps most importantly, promoting reproducible research.

2.4 BART Method for measuring available bandwidth

BART is a new method for estimating the end-to-end available bandwidth over a network path. It estimates bandwidth quasi-continuously, in real-time. The method has also been implemented as a tool. BART relies on self-induced congestion, and repeatedly samples the available bandwidth of the network path with sequences of probe-packet trains, sent at randomized rates. BART requires little computation for each sample which makes it lightweight with respect to memory requirements, and adds only a small amount of probe traffic to the network path.

With restricted access to traffic statistics recorded by intermediate network nodes the estimation of available end-to-end bandwidth is only feasible by active probing of the network path. This kind of active measurement only requires access to the sender and receiver hosts. By injecting probe traffic into the network, and then analyzing the observed effects of cross traffic on the probes, BART can estimate the available bandwidth. As a side effect it also estimates the link capacity of the tight link.

Some of the features of BART are:

- It produces an estimate quickly
- Estimation stability can be traded for agility
- Tuning is largely automatic, that is there are few parameters that need manual adjustment. Nevertheless BART may be tuned according to the specific needs of the measurement application, such as agility vs. stability of the estimate; or to characteristics of the bottleneck link.
- The memory requirements are minimal, as only the previous estimate and the new measurement are needed to calculate the new estimate of the available bandwidth.

BART was developed within the Evalunet project which was funded by the Swedish Governmental Agency for Innovation Systems (2003–2006) and the Knowledge foundation (2004–2006). The partners in
the Evalunet project were Mälardalen University, the Swedish Institute of Computer Science, Ericsson AB and Gatorhole AB.
3 FURTHER WORK

3.1 Identifying bottlenecks in broadband multi-service networks

As part of the TRAMMS project, Ericsson has initiated a process aiming at standardizing the different components of BART by submitting a contribution to the ITU-T for consideration. The selected target study group within ITU-T is SG12 which is the lead study group on performance and quality-of-service in packet-switched networks. This kind of efforts are needed to help setting structures for identifying bottlenecks and events that may result in network congestions.

A nice example of this is the TRAMMS proposal, based on BART tool, that can be summarized as follows:

The figure shows how BART is being proposed to measure the end-to-end available bandwidth and bottleneck link capacity, using its probing scheme and kalman-based analysis, in an experiment using nodes placed in Kista, Sweden, and Bilbao, Spain. In fact, BART will also report on packet loss and estimate variance over time. No changes are needed to the algorithm or the implementation.

On the other hand, a repository, designed by two other partners of TRAMMS (TID and GCM) will be collecting statistics on BGP updates from any Internet Exchange Points located in-between the BART sender and receiver nodes.

From the estimates obtained from BART and the statistics gathered by the TID-GCM repository we will search for correlations between a BGP update and for example packet loss or increase in available bandwidth estimation variance: A routing misconfiguration-traffic troubles connection may be establish by this experimental procedure.

This kind of experimental approaches must be serve not only to standardization work but also to feed back theoretical improvements in the knowledge of Internet traffic and help understanding QoE issues. This side of the problem is also taken into account in TRAMMS (see http://wiki-tramms.celtic-initiative.org/index.php/Main_Page) and other related projects by European academic institutions (UAM, for example).

But the real goal is finding (at least proposing) mechanisms to detect such problems and solutions.

3.2 Proposing solutions for eliminating bottlenecks

3.2.1 Solutions for DSL

Even though the DSL technologies are mature and widely deployed, the DSL lines are not free of problems and not always perform as well as expected. As stated in previous sections of this document these lines are exposed to different types of disturbances which might cause the lines to fail and so become the first bottleneck in the network.
The usage of measurement, diagnosis and management tools for DSL networks is seen as a necessity in order to keep DSL performance under certain levels. These types of tools are also really important to reduce the reaction time needed to address different faults in which a technician is needed.

The tools presented above, were originally designed to qualify and monitor the DSL lines. Later on, dynamic management characteristics were added to this tools taking advantage of the qualifying and monitoring features. However, there is still room for improvements when it comes to dynamic reconfiguration and management on DSL lines.

Within the TRAMMS project, the Loop Qualification and Monitoring tool designed by Ericsson will be analyzed in order to propose improvements in the dynamic line management area. Besides that, TRAMMS will work on the open interfaces needed to interoperate with third party vendor equipment.