A Trial Experience on Management of MPLS-Based Multiservice Networks

Eduardo Grampín[1], Javier Baliosian[2], Joan Serrat[3],
Gonzalo Tejera[1], Federico Rodríguez[1], and Carlos Martínez[1]

[1] Universidad de la República, Montevideo, Uruguay
{grampin,gtejera,rodrigue,carlosm}@fing.edu.uy
[2] Network Management Research Centre, Ericsson Ireland
javier.baliosian@ericsson.com
[3] Universitat Politècnica de Catalunya, Barcelona, Spain
serrat@tsc.upc.edu

Abstract. This article presents a component-based, distributed management system for Multiprotocol Label Switched (MPLS) multiservice networks[1]. Delivery of "triple play" multimedia services to the broadband residential user is a demanding challenge. The complexity is increased by the requirement of preserving Quality of Service (QoS) assurance for legacy connectivity services to the enterprise segment over the same infrastructure. New technologies are being introduced in the access, aggregation and core networks. Management applications must be aware of these advances and shall evolve accordingly. The proposed management architecture benefits from the capabilities of the MPLS Control Plane, in conjunction with a traditional management approach to provision QoS-aware services. This hybrid solution pursues short connectivity setup times by means of Control Plane signalling, with Traffic Engineering capabilities provided by the management framework. The system is being prototyped on a trial metropolitan testbed. Simulation results show that an advantageous trade-off between speed and resource optimisation is feasible.

1 Introduction

The convergence of video, voice and data services over the Internet occurs in parallel with the convergence of the historical Internet and Telecommunications management frameworks into object-oriented, distributed management paradigms. Convergent multiservice networks are build over several technology layers that must be tied up together in order to provide "triple-play" services to the broadband end user. ADSL "first mile" user access, ATM, Gigabit Ethernet and/or SDH aggregation and core networks with proper dimensioning and Traffic Engineering techniques enable the delivery of voice, Interactive TV and High Speed Internet services. Furthermore, these networks must also support, among others, traditional best effort Internet and VPN services for the Corporative segment, augmented with Datacenter (IT) services. The overwhelming operational complexity of such multiservice networks requires the

[1] This work is part of a project undertaken by Universidad de la República (UdelaR) and ANTEL, the public Telco of Uruguay, for development of a Metropolitan Multiservice Network with MPLS over optical transport, funded by Programa de Desarrollo Tecnológico (PDT)

T. Magedanz, E.R.M. Madeira, and P. Dini (Eds.): IPOM 2005, LNCS 3751, pp. 191–201, 2005.

appropriate integration at the Data, Control and Management Planes. Traditional layered non-integrated systems are simply unable to cope with the multiservice demand. IP/MPLS can provide some degree of integration at the Data and Control Plane, while modern management paradigms can complement such integration at the Management Plane.

This paper presents the preliminary results obtained by an ongoing project aimed at the study of the technological, operational and management problems arising in the deployment of a multiservice MPLS-based network. A component-based, distributed management system named GERMINA, which stands for “GEstor de Red del grupo MINA” [1] (MINA research group Network Management System), is being developed to face the O&M challenges. GERMINA comprises the Element and Network Management Layers (EML, NML), with a light Service Management interface that enables future development of the top layers of the traditional ITU-T TMN hierarchy.

The paper is structured as follows. Section 2 describes the GERMINA management system, while Section 3 reviews relevant related work. Section 4 presents early evaluation results, and the last section is devoted to make some concluding remarks and to point out future work.

2 GERMINA Management System

The evolution path towards the next generation IP over optical networks is supported by the MPLS Control Plane, which provides functionality to solve path computation and path establishment (i.e., the provisioning process); however, there are alternative proposals to support connectivity provisioning based on the Management Plane. For instance, the IST WINMAN project specified an Integrated Network Management System (INMS) for providing IP over WDM connectivity services, mostly using management functions supported by the Control Plane wherever applicable, as presented in [2].

Other management systems presented in Section 3, confirm that the intelligent Control Plane provides an important part of the operational solution in MPLS networks, which shall be complemented with well-known, trusted management techniques and tools. In this regard, a set of definitions and tools that provide support for Operation and Management (OAM) of MPLS networks are becoming available [3].

GERMINA is founded over JacORB [4], which provides the framework for the integration of components, developed by different groups of people spread in time and space. The Network and Element Management (NML-EML) interface is modelled following the Multi Technology Network Management (MTNM) Information Model [5], augmented with the MPLS extensions developed by the aforementioned IST WINMAN project.

2.1 Scenarios

This section presents some representative scenarios in order to state the general principles of GERMINA system operation to justify the system architecture depicted in Fig. 1. , which will be described later in Section 2.2:

- Service Provisioning
- SLA surveillance

- Offline Reoptimisation
- LSP Provisioning triggered by the Control Plane

Service Provisioning

Service Provisioning is triggered by the Service Manager (SM), as a result of the interaction with customer care applications, which defines service attributes, including specific Service Level Agreement (SLA). This SLA is translated by the SM into a Service Level Specification (SLS) in order to be implemented by the provisioning processes. Once the service specification is complete, a provisioning request is issued to the Configuration and Performance Manager (CPM), which proceeds with the phase of resource reservation and connectivity setup.

In order to fulfil the requested Quality of Service (QoS) it is necessary to solve a Constraint-Based Routing (CBR) problem for each request. This process is performed by the Routing and Management component (RMA) upon request of the CPM. The RMA has detailed knowledge of network topology, which is gathered running the Internal Gateway Protocol (IGP) as a peer on the network. The resources selected to satisfy the ongoing connectivity request are reserved in the internal CPM Inventory Component.

The provisioning process progresses in the Management Plane side towards the Element Manager (EM), which configures the required SubNetwork Connections (SNCs) over the technology neutral managed objects. This configuration is translated into technology/device specific commands to implement the changes in the network by means of a suitable protocol such as SNMP and/or Command Line Interface (CLI).

Finally, once the changes have been implemented, the upper layers of the management framework become aware of the topology changes by CORBA notifications issued by the EM (i.e. the Inventory is updated by these notifications).

SLA Surveillance

The Performance and Fault Management processes of the EM trigger an SLA surveillance process when given thresholds are reached in relevant parameters and/or faults are detected in network elements. Several mechanisms can be used to detect these anomalies, depending on equipment peculiarities. Most vendors implement SNMP Agents embedded in network devices with trap generation capabilities under certain faulty conditions and/or when thresholds are reached in relevant objects (i.e. percentage of link bandwidth usage). Other possibilities include external monitoring and measuring by dedicated appliances (e.g. RMON probes).

Performance degradation is detected at the Network Management Layer by means of notifications issued by the EM. On arrival, the Performance Manager first consolidates the information, and then checks the conformance of configured SLSs. If a service violation is detected the CPM may react in several ways; if (1) a set of policies are defined, the CPM will initiate a reconfiguration process in order to meet the SLS parameters, else (2) the upper management layer (SM) will be notified in order to proceed with a service re-establishment. A complete implementation shall also trigger commercial processes such as discount and report to the affected client(s).

Offline Re-optimisation

The objective of this process, which can be periodically scheduled by network administrators (e.g. once a week), is to re-compute the optimum resource usage for a known traffic demand. Given an optimum distribution of resources at a given time (e.g. after offline re-optimisation), network states evolve out of the optimum due to dynamic resource assignment. The CBR process computes feasible resource distribution for each request, but in general a global optimum is not reachable under this dynamic demand regime.

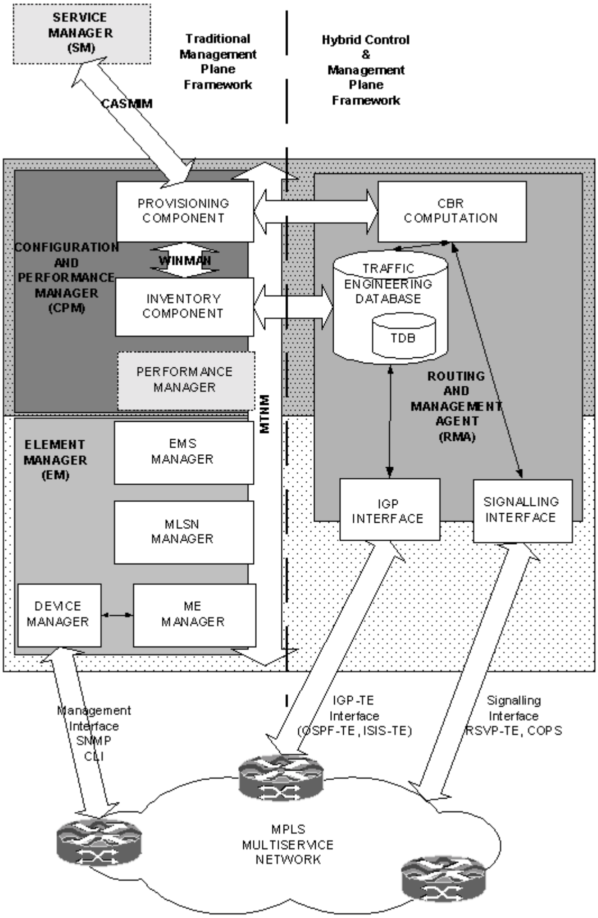

Fig. 1. System Architecture

Offline re-optimisation is performed using a snapshot of the Topology Database and taking into account global optimisation objectives, which are defined by network administrators in the form of objective functions and administrative constraints (i.e. policies), considered by the CBR algorithm. Once the global optimum is computed by the RMA, the system shall reconfigure network connectivity services. Care must be taken to avoid service disruption, using a "make before break" procedure.

LSP Provisioning Triggered by the Control Plane

The *Service Provisioning* scenario described a traditional Management Plane-initiated provisioning process. GERMINA is capable of LSP setup upon request of an ingress

LSR. The request is received by the RMA using its *Signalling Interface*. There are two optional approaches for signalling, namely the request/reply transaction between the RMA and the ingress LSRs, as described in Section 2.2.3. Once a path has been computed for a particular request, the ingress LSR receives a reply from the RMA and LSP setup proceeds using the standard Control Plane signalling.

An interesting feature of this approach is that the RMA can compute both the downstream and the upstream paths. Thus, bidirectional LSP setup is enabled, triggering standard signalling in both the ingress and the egress LSR.

2.2 Description of System Components

The system implements basic Element (EML) and Network Management (NML) layers. The existing functionality comprises Configuration, Performance and limited Fault Management (basically provided by the EML).

Note that the left side of the Fig. 1. presents a "traditional" distributed, component-based management framework founded over the mentioned CORBA bus, while the right side shows the RMA modules, which implement the aforementioned hybrid connectivity setup approach. The RMA management interfaces (through the CBR and Topology DataBase modules) enable the interaction with the rest of the framework for the fulfilment of the functionality needed in the above described scenarios. The hybrid approach can work in an autonomous fashion and will be described in Section 2.2.3. The following sections describe the GERMINA components in detail.

2.2.1 Service Manager

This module has been partially implemented as a proof of concept of the NML and EML functionality, using a minimal version of the CASMIM SML-NML interface [6]. The SM invokes methods provided by the NML related to Subnetwork Connection lifecycle:

- createSNC();
- activate SNC();
- deactivateSNC();
- deleteSNC();

2.2.2 Configuration and Performance Manager

This module is driven by a *Provisioning Component* (PC) which makes use of the *Inventory Component* (IC) and the *Routing and Management Agent* (RMA) components to fulfil the Configuration Management processes. PC and IC have been developed by coordinated BSc Projects [7][8], while the *Performance Manager* (PM) component is partially developed, aiming to handle network-level performance policies using the functionality provided by the EM and under the surveillance of the SM objectives. Note that PC and IC have an additional interface that uses the Information Model developed by the IST WINMAN project for object persistency management.

2.2.3 Routing and Management Agent

The RMA is a logically centralized entity intended to perform computation of network paths under a given topology with QoS and administrative constraints, and without assuming previous knowledge of future demand. In other words, the RMA is a Constraint Based Routing (CBR) capable entity, and behaves like a Path Computation Server (PCS) with the following additional and innovative features:

- The RMA gathers network topological information as an IGP peer in a given domain.
- Path computation requests can be signalled by ingress LSRs using the standard RSVP-TE signalling protocol [9], or using a request/reply protocol, as being defined by the IETF Path Computation Element (PCE) Working Group [10].
- Path signalling is performed by ingress LSRs using the standard RSVP-TE signalling protocol.
- The RMA can be integrated in a management distributed component-based environment using well-known frameworks such as CORBA, J2EE, .NET, etc.
- The RMA hardware and software architecture is a distributed computational system based on High Performance Computing (HPC) principles, with virtually unbounded resources.

Traditional *offline CBR* performs path computation outside network elements, in a PCS. Such process takes as input a known, static traffic matrix (result of existing connections and connectivity requests) and, based on a detailed and accurate topology map (built with information gathered from the network), computes the *optimal* network paths for that given traffic matrix.

On the other hand, *online CBR* is a routing mechanism embedded on the network elements intelligence. Such a routing process receives as input dynamic traffic requests and has no knowledge of future requests. Given this traffic request and based on dynamic (and possibly incomplete) network state it computes a *feasible* path for that request.

CBR was traditionally used in transport networks to accommodate relatively static traffic demand; on the other hand, traditional IP networks used to offer best-effort service, relying on hop by hop routing without QoS guarantees. The shift from traditional best-effort connectionless IP networks to multiservice MPLS-based connection-oriented networks with QoS guarantees and dynamic traffic demand imposes the challenge to solve the CBR problem with the accuracy of the offline process but the timing restriction of the online solution. CBR with more than one constraint is a well-known NP-complete problem, and therefore, the necessary computation power is unbounded. This prevents the introduction of full CBR capabilities into network devices, since they have scarce computational resources mainly devoted to packet forwarding. Moreover, different implementations of CBR algorithms lead to the impossibility of fulfilling network-wide TE objectives.

The RMA is a routing and signalling peer node in the network, but avoiding traffic forwarding. Also, it is assumed that in order to fulfil its main objective, the RMA has enough computation power to solve the CBR problem in near real time. This entity, while enabling a Control Plane-based provisioning, can be used as a complementary TE tool by management applications, using its interface towards the Management Plane.

In the tested RMA prototype, path computation can be requested using the standard RSVP-TE signalling, or by means of a request/reply protocol, in line with the aforementioned PCE WG definitions. The well-known COPS protocol [10] has been tested for this role, which additionally supports the exchange of policy information between the LSRs and the RMA. Simulation results shown in Section 4 reveal that both signalling alternatives are feasible. Further explanation of the RMA architecture can be found in [12].

2.2.4 Element Manager

This module is composed by the EMS (Element Manager System), MLSN (Multi Layer Subnetwork) and ME (Managed Element) managers, developed by the aforementioned BSc projects. Note that these modules are technology-independent, while the ME module implements an extension API that is used by the Device Manager (DM), which actually implements the technology-specific communication with managed elements.

The DM is an extension of the system presented in [13]. Its functionality is tailored for IP/MPLS networks. Besides the existing Cisco-specific parser, support for standard MPLS-MIB is being built in the DM. Also, the Linux nodes are being augmented with an MPLS-MIB compliant SNMP Agent using Net-SNMP extension modules [14].

3 Related Work

MPLS augments traditional best-effort IP with traffic engineering capabilities, adding functionality to network devices (i.e. constraint-based path computation), which need ever growing resources to cope with this complexity in terms of processing time. Finding optimal, or even feasible network paths that meet certain constraints is a challenging task for head-end routers. Some traffic engineering systems have been proposed to assist network devices in path computation and resource assignment in MPLS networks. Their relevant characteristics are briefly described hereafter.

RATES [15] presents a component-based, expandable management architecture, comprised of the following major components:

- Explicit Route computation.
- COPS Server for communication with head-end routers.
- Data Repository.
- Network Topology and State Discovery component, running OSPF.

Some of the ideas presented by this proposal have been considered in the design of GERMINA; in particular the usage of OSPF to gather network information is an important contribution.

Wise<TE> Traffic Engineering server for MPLS networks [16], presents a complete analysis of the caveats of IP-based traffic engineering, the limitations of CBR running on network nodes and other issues, but reaches the limited conclusion that the solution is an offline traffic engineering tool. Wise <TE> major components are the following:

- Traffic Measurement and Analysis Server, which performs collection, characterization and analysis of traffic information.
- Routing Advisor for Traffic Engineering, a planning tool with CBR capabilities.
- Resource Monitoring Server, a topology and configuration information repository
- Policy Server, which configure policies on network devices using vendor specific device agents.

This proposal is similar to the RATES system; its major contribution regarding the GERMINA architecture is, again, the dynamic gathering of network states to feed the traffic engineering database.

Both proposal are founded over a CORBA distributed environment, and present similar ideas. Based on previous work, the design of a new system is advised to be component-based, with monitoring capabilities, and with participation in the IGP process. The communication with the network devices is largely dependent on specific routers capabilities; anyhow, COPS is widely deployed and is a feasible choice, as shown in Section 4. Path computation is non-standard by nature, and each proposal implements algorithms that claim to fulfil certain performance objectives. A new proposal shall be flexible in this respect to support different capabilities and do not be restricted by computational resources.

4 Proof of Concept Results

GERMINA is a growing system composed of several disjoint modules tight together by a common CORBA framework. Functional testing of the Provisioning and Inventory Components have been conducted; results can be found in [7] and [8].

Stimulating results have been obtained by simulation of the RMA component, using ns-2 simulator [16] extensions for RSVP-TE [18] and COPS [20], with Waxman topologies generated by the BRITE tool [18]. Fig. 2. shows the summarizing results for the establishment of LSPs using the Management Plane, the Control Plane and the hybrid (i.e. the RMA) approaches.

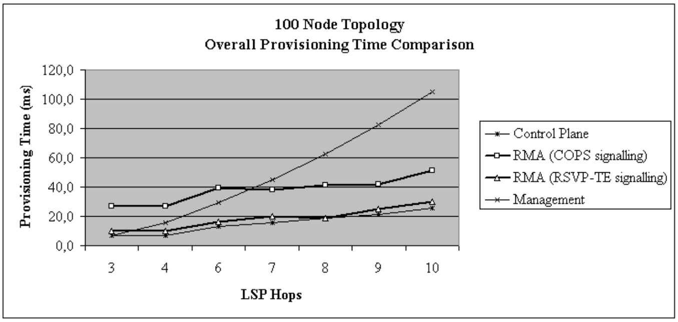

Fig. 2. LSP Setup Time as a function of LSP hops

Provisioning time for the Management Plane and the Control Plane approaches impose upper and lower bounds respectively. The RMA using RSVP-TE signalling shows a performance very close to the Control Plane lower bound, while the RMA with COPS signalling timing shows a slight degradation due to the complexities added by the session establishment between the ingress LSRs and the COPS server (i.e. the RMA). Note that in this case, bidirectional LSP setup is achieved, so there is a trade-off between a slower response time and an enhanced functionality.

Field evaluation of the GERMINA management system is being undertaken in a trial network. The core is composed by standard commercial MPLS routers, while the aggregation and access is comprised of Linux-based routers, as shown in Fig. 3. .

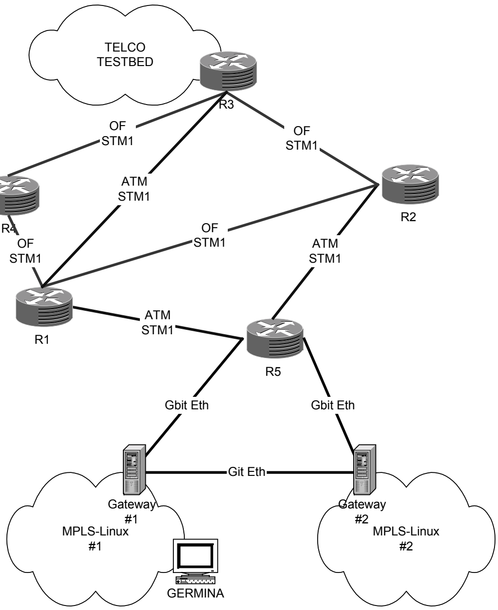

Fig. 3. Multiservice MPLS Network

The MPLS network runs the OSPF-TE routing protocol [21] and uses RSVP-TE signalling for the establishment of Traffic Engineered LSPs (TE-LSPs). The GERMINA management system is located in one of the access "clouds" and gathers network information using a Management Virtual Private Network built over the MPLS infrastructure. A set of tools for traffic generation and monitoring are used to emulate the users' behaviour, modelled using real network traces gathered from strategic nodes of the public IP network. Gathering and processing network traces for user characterization is one of the interesting outcomes of the project, motivating the usage of data warehousing and statistical techniques to cope with the enormous amount of data needed to be handled.

As seen on the picture, the infrastructure comprises point to point STM-1 optical links, ATM switched STM-1 and Gigabit Ethernet links. The core is composed of Cisco 7206 VXR routers and the Gateways are Intel servers equipped with PRO/1000 MF Dual Port adapters, running MPLS-Linux [22] and Quagga routing software [23]. As mentioned above, Net-SNMP tools are installed in the Linux nodes, which implement a subset of the MPLS-MIBs.

5 Conclusions and Future Work

This paper presents an ongoing project entailed to provide a solution for the O&M challenge imposed by new technologies and services being deployed in Service Providers' networks. The GERMINA management system provides alternatives for service provisioning using a standard distributed, component-based framework augmented with a hybrid concept, the RMA, which takes advantage of the MPLS Control Plane.

Simulation results for the RMA show promising performance in comparison with "pure" Control Plane LSP provisioning. This solution enables a global optimisation of network resources by means of an improved routing function outside LSRs, which also offload router processing, boosting the packet forwarding functionality. The RMA architecture is entirely based on standard protocols, and enables network administrators to control the routing strategies (i.e. network-wide TE objectives). Further validation is being conducted in the described trial testbed.

Future work involves the completion of the partially implement modules, conception of a mechanism to define network-wide TE policies for the CBR module, and trial of exact algorithms and heuristics to solve the CBR problem near real time.

An evolution towards a Service Oriented Architecture (SOA) is being evaluated as a foundation for the management building blocks, while aggregation and customer-specific issues shall be considered in the trial network. For example the usage of IP DSLAMs and multicast-enabled Ethernet aggregation over MPLS Virtual Private LAN Services (VPLS) and/or Virtual Private LAN (VLAN) are topics that will be evaluated.

Acknowledgements

The authors would like to express their gratitude to their co-workers in the Metropolitan Multiservice Network with MPLS over optical transport project at UdelaR, Uruguay, for their support and valuable comments.

References

1. MINA Academic Group: http://www.fing.edu.uy/inco/grupos/mina. Last visited: August 2005.
2. L. Raptis, G. Hatzilias, F. Karayannis, K. Vaxevanakis, E. Grampín, "An integrated network management approach for managing hybrid IP and WDM networks". *IEEE Network*, Volume: 17, Issue: 3, May-June 2003 Pages:37 – 43.
3. D. Allan, T.D. Nadeau, "A Framework for MPLS Operations and Management OAM", Internet Draft <draft-ietf-mpls-oam-frmwk-03.txt>. Work in Progress. Expiration Date: August 2005.
4. JacORB: http://www.jacorb.org. Last visited: August 2005.
5. TMF 608, "Multi-Technology Network Management Information Agreement NML-EML Interface." Version 2.0. TM Forum Approved, October 2001.
6. "Connection and Service Management Information Model (CaSMIM)" TMF 508,605 and 807 Public Evaluation, June 2001, Version 1.5
7. M. Giachino, C. González, J. Visca. "Component-based Network Management". BSc Graduation Project, April 2004. Universidad de la República, Uruguay.
8. M. Bartesaghi, B. Fagalde, F. Zubía. "Inventory Management". BSc Graduation Project, March 2004. Universidad de la República, Uruguay.
9. D. Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels". RFC 3209, December 2001.
10. IETF Path Computation Element (pce) Working Group. http://www.ietf.org/html.charters/pce-charter.html
11. D. Durham, J. Boyle, R. Cohen, S. Herzog, R. Rajan, A. Sastry, "The COPS (Common Open Policy Service) Protocol". RFC 2748, January 2000.

12. E. Grampín, J. Serrat. “Cooperation of Control and Management Plane for Provisioning in MPLS Networks”. IM 2005 Conference, Nice, France. May 15-19, 2005.
13. E. Grampín, J. Baliosian, J. Serrat. “Extensible, Transactional Architecture for IP Connectivity Management”. 3rd IEEE Latin American Network Operations and Management Symposium (LANOMS'2003). Iguaçu Falls, Brazil - Setiembre 4 - 6, 2003.
14. Net-SNMP: http://www.net-snmp.org/. Last visited: August 2005.
15. P. Aukia, M. Kodialam, P.V. Koppol, T.V. Lakshman, H. Sarin, B. Suter, “RATES: a server for MPLS traffic engineering”. *IEEE Network*, Volume: 14, Issue: 2, March-April 2000 Pages:34 - 41.
16. T.S. Choi, S.H. Yoon, H.S. Chung, C.H. Kim, J.S. Park, B.J. Lee, T.S. Jeong, “Wise<TE>: traffic engineering server for a large-scale MPLS-based IP network”. Network Operations and Management Symposium, IEEE/IFIP NOMS 2002, 15-19 April 2002.
17. The Network Simulator - ns-2: http://www.isi.edu/nsnam/ns/. Last visited: August 2005.
18. A. Medina, A. Lakhina, I. Matta, J. Byers, "BRITE: Boston university Representative Internet Topology gEnerator", http://www.cs.bu.edu/brite/ Last visited: August 2005.
19. C. Callegari, F. Vitucci, "RSVP-TE/ns network simulator", http://netgroup-serv.iet.unipi.it/rsvp-te_ns/. Last visited: July 2005.
20. Sayenko, T. Lahnalampi, "COPS for the NS-2 simulator", http://www.cc.jyu.fi/~sayenko/pages/en/projects.htm. Last visited: July 2005.
21. D. Katz, K. Kompella, D. Yeung, “Traffic Engineering Extensions to OSPF Version 2”, RFC 3630, September 2003.
22. MPLS for Linux project: http://sourceforge.net/projects/mpls-linux/. Last visited: July 2005.
23. Quagga Routing Software Suite: http://www.quagga.net/. Last visited: July 2005.