# Pathfinding architectures for interdomain Bandwidth Broker operation

Christos Bouras, Dimitris Primpas
Research Academic Computer Technology Institute, N. Kazantzaki Str., Patras University Campus, 26500 Rion, Patras, Greece &
Department of Computer Engineering and Informatics, University of Patras, 26500 Rion, Patras, Greece
TEL: +30 2610 {960375, 960316}
FAX: +30 2610 960358
E-MAIL: bouras@cti.gr, primpas@cti.gr

**Abstract - This paper deals with bandwidth brokers and their inter-domain operation. The basic issues for inter-domain operation are discussed and we try to approach the most demanding issues as the selection of the best inter domain routing path (pathfinding). Generally, we discuss two models for inter domain routing through bandwidth broker, analyzing their advantages and comparing them. Also, we simulated the second one (generalized distributed pathfinding) trying to evaluate its performance. Finally, we discuss the simulation's results and present how this model should be incorporated in normal operation routing of commercial ISPs or academic networks that works on a federated way.**

## 1. INTRODUCTION

The last years many critical applications have been appeared widely, demanding specific network characteristics in order to operate effectively. For this purpose, many service providers and researchers started studying Quality of Service issues. Currently, 2 basic architectures have been proposed by IETF, the Integrated Services (IntServ) and the Differentiated Services (DiffServ). The DiffServ [1] has been widely adopted and many mechanisms and QoS services have been implemented.

The most critical issue in the deployment of a QoS service is the ability to provision the network and provide the service in end to end basis. Otherwise, the deployment is quite complicated, as many parties should work and also there are many issues that can finally degrade the overall performance that the applications experiences. In this direction, the automatic provisioning and management of a network and ideally all the connected networks through pre-agreed interfaces is the next big challenge. This goal can be achieved through the bandwidth brokers [2]. The last years, several researchers work on the issue of bandwidth brokers, dynamic QoS provisioning and inter domain operation. Many research papers have been presented, focusing on various models and architectures [5][6][10][12][15]. In this paper, we focus on 2 proposed models for interdomain operation and we try to compare and evaluate them.

The rest of the paper is organized as follows. Section 2 describes the bandwidth brokers and the inter-domain operation and section 3 focuses on all the related issues regarding the successful end to end provisioning and efficient utilization of the resources. Section 4 describes the simulation approach and the simulation results. Next, section 5 describes the conclusions and section 6 is devoted to the future work that we intend to do in this area.

## 2. INTER DOMAIN BANDWIDTH BROKER ARCHITECTURES

A bandwidth broker is a "service" that provisions a backbone network and manages the supported QoS services in a domain. Bandwidth broker is also responsible for the inter-domain communication with the bandwidth broker of adjacent domains. Generally, a bandwidth broker contains several modules that are necessary for its transparent and efficient operation.

- *An inter-domain interface.* It is used for communication with adjacent BBs
- *An intra-domain interface.* It is used for communication with the service components that are located inside the domain that the BB controls
- *A routing table interface.* It is used so that the BB knows the network topology and reacts with routing protocols and routing paths
- *A user/application interface.* The scope of this interface is to allow the user and applications to send requests to the BB.
- *A policy manager interface.* This interface allows dynamic implementation of policy management or admission control strategies.
- *A network management interface.* It is used for network management and router's configuration, according to admission control and provisioning status.

A bandwidth broker keeps internal information of the managed domain, peering information and resources through its local provisioning system. Many alternative approaches on provisioning methods have been presented lately [5][6][12][16]. Also, the network dimensioning and call admission control algorithms are another important point in bandwidth brokers [10][14][15].

In general, when a request to the bandwidth broker is submitted, it contains source and destination points and also the requested bandwidth. Next, the basic module of the bandwidth broker should parse the source and destination addresses and try to identify the sub- networks where these addresses belong to. The source naturally resides in the ISP's network, the destination might well be in another ISP's domain. That domain might not be the next connected domain to the ISP and there might be one or more domains in between. If both the source and destination addresses are in the stub networks of the same ISP domain, the Broker that maintains the domain can find the ingress and egress routers by some simple lookup in the related. If the destination is in domain other than the source domain, then the Broker must identify the final domain by the peering information that it keeps. In addition and depending on the architecture of the inter-domain protocol, the bandwidth broker may also need to find the intermediate domains (if any) that will be traversed to the destination (the bandwidth broker should calculate the "best" path from source to destination through the intermediate domains).

The later (the investigation of the best path for the end to end communication, taken into account the possible SLAs between domains as well as traffic engineering characteristics) is an important and challenging issue during the processing of such requests. Many researchers have studied traffic engineering issues on intra-domain and inter domain bandwidth brokers [7][8]. In this paper, we discuss 2 models that approach the end to end traffic engineering (pathfiding) on interdomain operation.

## 3. PATHFINDING APPROACHES

### *3.1 Centralized Pathfinding model*

The first one is called centralized and according to this model, the decision about the routing of the request, in particular the domains and the internal paths that the traffic will traverse in order to reach the destination domain is made by the source domain. In order to take this decision, a central provisioning system is necessary that will keep topology, peering (SLAs between ISPs) and technology information. In particular, the ISPs that take part in this model, announces their topology and resource status to a common database that is used for bandwidth broker's operation. All ISPs should keep the relevant information in this database always synchronized. Otherwise, the bandwidth broker will work on inconsistent data and can lead to incorrect paths and reservations. The implementation of such a database is an open issue and several approaches have been presented [12].

Next, after a request will be submitted, the source domain makes "queries" asking for the requested resources across the paths. It follows this procedure sequentially until it finds all paths from source to destination that have the requested resources. Next it decides regarding the best one according to criteria that have been specified in the bandwidth broker. The criteria should be defined by the ISPs and should indicate the cost that they have in order to provide the resources.

This centralized pathfinding model can be parallelized with the operation of RSVP-Traffic engineering protocol on a single managed domain [3]. The RSVP-TE on a single MPLS enabled domain provisions the network resources and provides Label Switched Paths (that can use the resources) using various selection criteria in case of many candidates. Generally it is close to the centralized pathfinding model but not applicable, as the interdomain traverses independently managed networks and therefore the end to end operation of a RSVP-TE is not possible. Also, the RSVP-TE operates only on the runtime of network devices and its operation can not be exported for usage on offline or complemented tools. Additionally, the same operation can be succeeded through other proposed approached [7][8], but they still have the limitation that eliminates the freedom of each domain (announces its local policy) or requires common network mechanisms that are not applicable on independently managed federated networks.

### *3.2 Distributed Pathfinding model*

The second model is "peer to peer" and actually the source domain sends the requests that have destination to another domain in every adjacent domain, under the condition that there are available resources from source to the respective egress points. Next, every domain receiving such a message checks if it is the destination domain or it is immediately attached to it. If the answer is positive then the bandwidth broker checks if it can guarantee the requested resources from the ingress point to the destination of the egress point to destination domain. If there are the necessary resources, then the message is forwarded there. In case that the domain is an intermediate and not attached to the destination, then it broadcasts the request to all adjacent domains, which it has the requested resources to reach them (from its ingress point to the egress point to next domain). From the broadcasting have been isolated the domains from which it has received the message in order to avoid loops. This procedure is repeated and finally all the possible routes between the source and destination are declared to the source domain. Then, the source domain decides the best routing according to specific criteria that have been declared in the pathfiding model. In case there are not any paths with the requested resources due to SLAs violation, it is applicable to the model to request SLA negotiation between the domains and finally decide. Also, in case that the destination is unreachable due to network failure, the module comes in a deadlock, as there is not returning paths even rejected due to insufficient resources. Therefore, the module should have an expiration time and in case that this period passes, the module understands that destination is unreachable.

The complexity of this distributed pathfinding module is quite large and depends on the current topology between ISPs that the module is applied. Actually, this problem resides on Depth-first search (DFS) on a graph, where it is formatted in every request and uses as root the source domain [9]. Time complexity of DFS algorithm is proportional to the number of vertices plus the number of edges in the graphs it traverse.

Additionally, a very important issue in the whole operation is the definition of criteria that should be applied on the selected paths (if the model provides more than one) in order to decide the best one. The proposed model uses a combination of criteria that are the minimum hops, the minimum SLA fulfillment and the total cost that have different weight as each one criterion influence differently the overall factor. As total cost is considered the summary of the costs from all the traversed domains. The cost of each domain is declared by the ISP itself and indicates the actual cost of the ISP to provide the resources according to its internal topology and provisioning system. The aim of the formula to decide the best path from the available ones is to perform load balancing and low cost according to the ISP's premises. The model, as described above, assumes symmetric SLA on transmit and receive direction between domains, but it also works when the SLAs are asymmetric.

In order to implement this model, each domain has the freedom to use any open source or proprietary provisioning tool for its network. The only requirements are the implementation of the overall module that will provide the synchronization of the operation and also a common format of the data that each domain provides. Additionally, this format is proposed to be based on XML as the overall pathfinding module can work using web services. A proposed XML schema is presented in Figure 1. In addition, Figure 2 describes the architecture of the model in the interdomain environment and Figure 3 presents the distributed pathfinding model in pseudo code.

```
<fathfinding-report>
  <domain>DATA</domain>
  <ingress_point>DATA</ingress_point>
  <outgress_point>DATA</outgress_point>
  <resources>DATA</resources>
  <cost>DATA<cost>
  <destination_reached>YES/NO</destination_reached>
</fathfinding-report>
```

Figure 1: XLM schema for pathfinding module

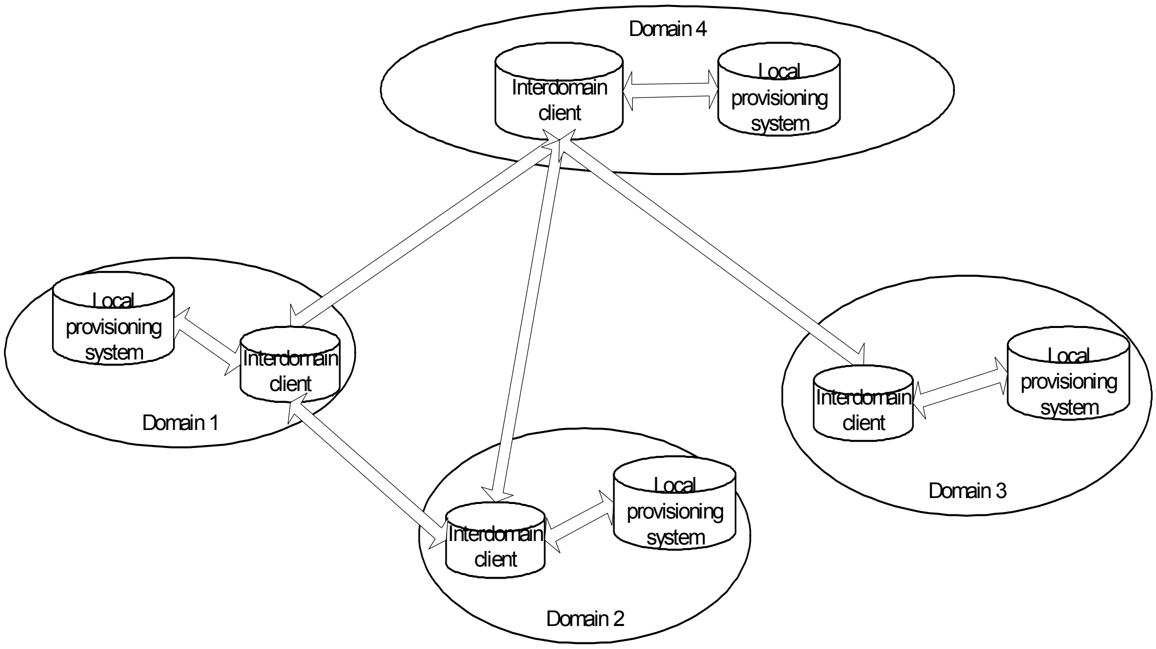


Figure 2: An interdomain approach using Distributed Pathfinding

The source bandwidth broker receives a request for a destination outside its domain.

It checks for resources across the routing paths to every adjacent domain (source, destination).

For each path {

  If the answer is positive {

```
    Forwards a request to the adjacent domain
  } else {
    failed
  }
The next bandwidth broker performs the same check with source the interconnection point with the first domain and destination the original destination or the interconnection points with other domains except the first one. If the bandwidth broker receives the same request from the same domain second time it rejects it (as it means that we reached a cycle in the network)
The source bandwidth broker finally receives the answers.
For all answers {
  It checks the answers according to pathfinding priorities - criteria and sorts the answers
  The bandwidth broker selects the first answer in the sorting list as the final path
}
If all answers are negative, search if it is due to interdomain  SLA violation and stores for every path all violated SLAs, order by minimum violation
For each violated SLA {
  It asks for dynamic negotiation and waits a response.
  If answer is positive then break, and therefore the request is admitted
}
```

Figure 3: The distributed pathfinding model in pseudo code

### 3.3 *Comparison*

Both models need updated information for the network's condition and its policy. The first algorithm requires every domain to announce that information centrally, where all the other domains "query" it to process inter-domain requests. The second algorithm eliminate this situation (announcement of network's condition and operation), by engages the bandwidth broker server of each domain to answer about the possible paths with available resources. Therefore the operation of each network remains internally and every request that may pass through a domain is processed by domain's bandwidth broker. On the other hand, the second algorithm has a bigger complexity and also needs more time to answer a request as all the possible routes should be checked though DTS search by asking domains' bandwidth brokers. Therefore, the response time includes the transmission delay of the requests for path finding between the bandwidth brokers of the domains and their execution time locally in each domain. The first algorithm has a big advantage in this issue, as all the information about all domains and their operation is stored locally in a database and therefore all the alternative routing paths can be found by searching the new graph that is produced by adding all domain's topology, the available resources as well as the SLAs between adjacent.

As a conclusion, the first algorithm has better response time but needs to announce internal information periodically to keep the global "provisioning database" updated. The second algorithm keeps that internal information internally but the process of every request engages all the bandwidth brokers of the involved domains. As this model targets on interdomain operation where each domain is an independent ISP, the second one (distributed) is more suitable due to the fact it retain the independence of the management and local policy.

## 4. SIMULATION AND ASSUMPTIONS

In order to study more the distributed pathfinding module, we simulated its operation and applied it on random requests (almost 170) between domains in the 3 topologies presented in Figure 4. The measurement that we tried to approach is the exchanged packets per request, as they indicate the overall response time, using the assumption that the response time of each domain's bandwidth broker is quite similar. According to the results (see Figure 5), the number of packets that the pathfinding module needs is quite small (the average numbers of packets for pathfinding queries when random requests are generated in the 3 above topologies are 3, 7 and 8 packets respectively) and depends on the overall topology and the location in the topology of the source and destination domain on every request. The figure presents the exchanged packets between the domains and not the internal packets in each domain (if any), as it depends on the provisioning service that each domain has select to use. After the pathfinding module, the algorithm returns the available paths that reaches the destination domain (and their characteristic), waiting for the final decision about the preferred path that will be produced after setting up the pathfinding's criteria. The overall request's process will finish when the appropriate path has been selected and the relevant domains have been informed to reserve the resources and perform the necessary configuration on network devices.

Comparing the results for the needed packets (communication) for pathfinding in the 3 topologies, we notice that the average number of exchanged packets increases proportionally to the complexity of topology (and in particular when it contains many links that lead to cycles). This result was expected due to the fact that the distributed pathfinding module is based on DFS algorithm [9].

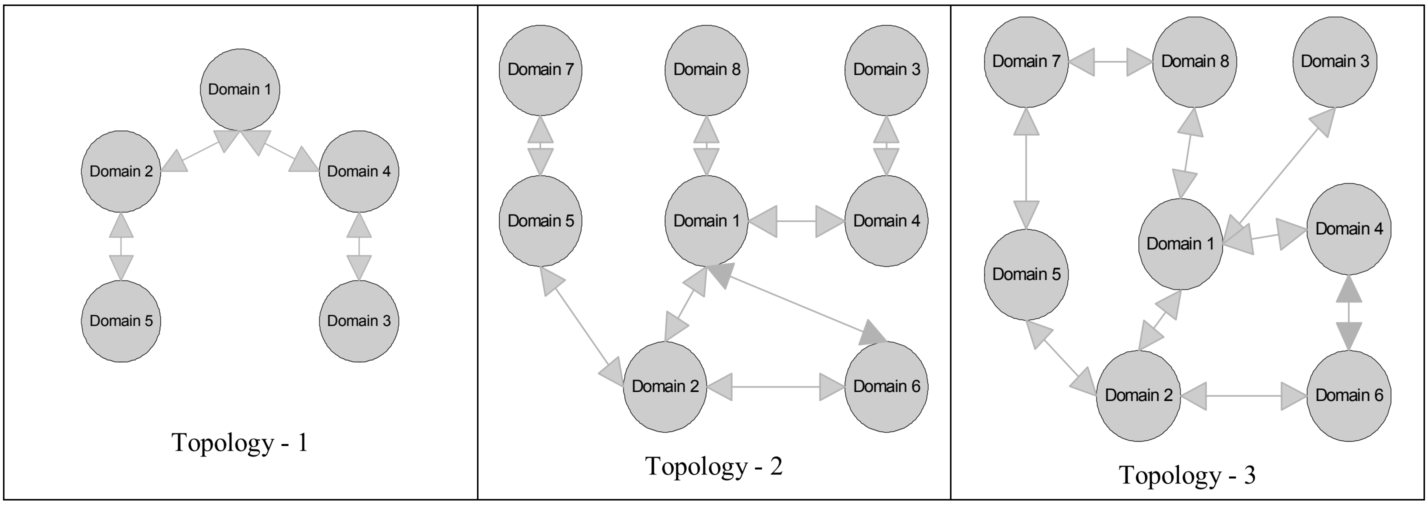


Figure 4: The simulated topologies

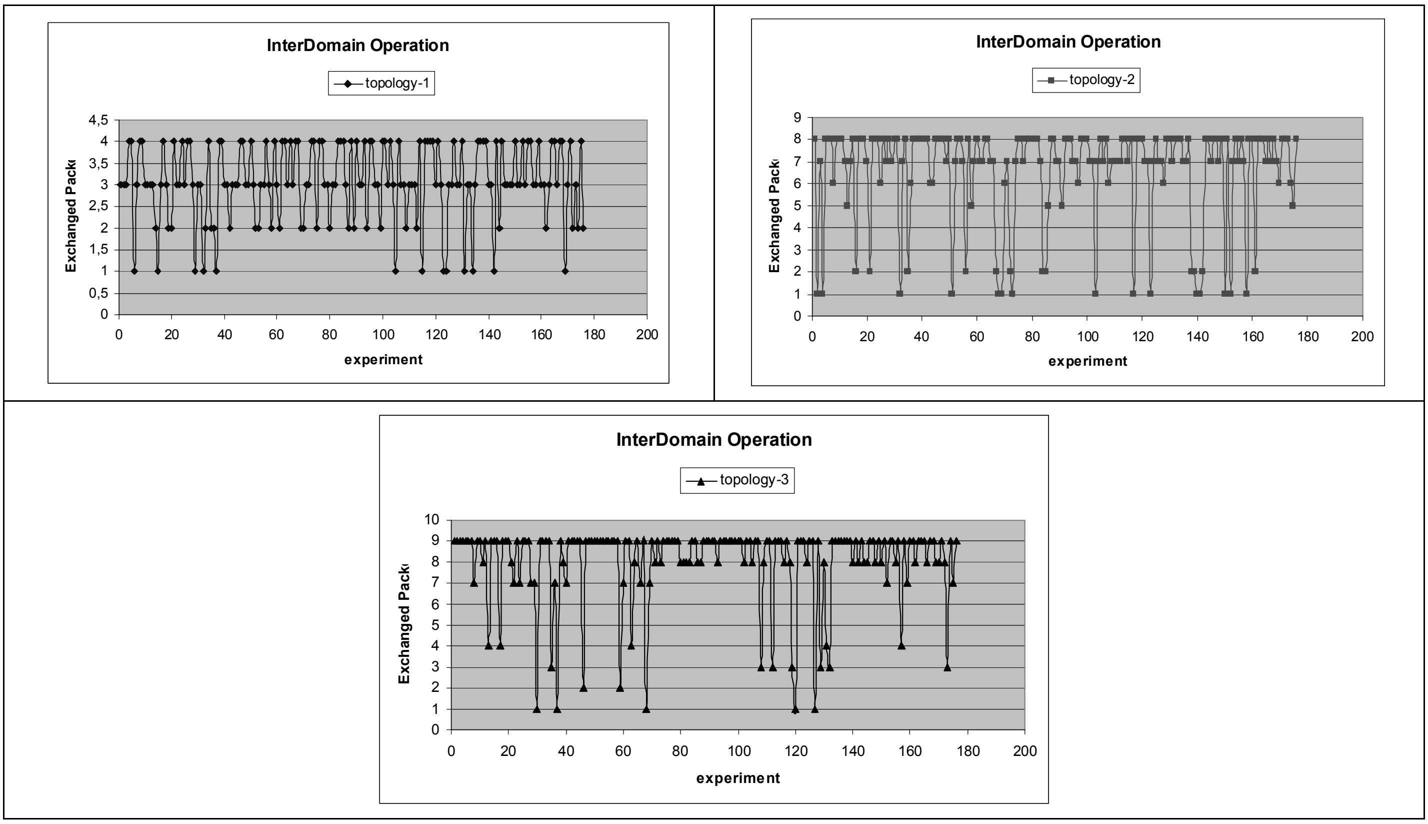


Figure 5: The exchanged packets for the pathfinding module

## 5. CONCLUSIONS - FUTURE WORK

This paper deals with the inter-domain operation of bandwidth brokers in order to perform end to end provisioning and therefore end to end guarantees through QoS services. Today's networks do not have yet automatic procedures to provide such functionality, but many approaches are under design or testing. The paper presents the relevant aspects for inter-domain operation of a bandwidth broker and focuses on pathfinding issue. In this paper we describe 2 models that can be used (the centralized and the peer to peer), we analyze and compare them. Also, we simulated the peer to peer model on 3 different topologies. The peer to peer model inserts a communication overhead as it exchanges many packets (the actual number depends on the topology and the location of source and destination domain). But this overhead can be characterized low as the exchanged data are small (the packets has very small size). On the other hand, the peer to peer model suites better on commercial ISPs and academic networks as it permits them to manage their network independently and interact through specific procedures based on standards (like web services and XML schemas).

This interdomain operation and approaches that described in this paper are applicable on both IP QoS bandwidth brokers as well in bandwidth brokers that provide resources on optical layer. It is obvious that the pathfinding module is similar on both cases instead of the other components of bandwidth broker, like provisioning, technology based reservation and establishment of guaranteed connections, that differ.

Finally, we plan to extend our work in this area mainly focusing on the implementation of the inter-domain operation on a existing bandwidth broker implementation on NS-2 simulator that we have done [11]. Then, we intend to perform a number of large scale tests in order to investigate the whole operation and performance.

## 6. REFERENCES


[1] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, *"An Architecture for Differentiated Services",* IETF RFC 2475, December 1998

[2] J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Bruijn, C. de Laat, M. Holdrege, D. Spence, *"AAA Authorization Application Examples"*, IETF RFC 2905, August 2000

[3] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, G. Swallow, *"RSVP-TE: Extensions to RSVP for LSP Tunnels",* IETF RFC 3209, December 2001

[4] K. Nichols, V. Jacobson, and L. Zhang, *"A Two-Bit Differentiated Services Architecture for the Internet"*, IETF RFC2638, July 1999.

[5] Ibrahim Khalil and Torsten Braun "Implementation of a Bandwidth Broker for Dynamic End-to-End Capacity Reservation over Multiple Diffserv Domains", *4th IFIP/IEEEInternational Conference on Management of Multimedia Networks and Services (MMNS)*, Chicago, USA, Oct 29 - Nov 1, 2001

[6] Junseok Hwang, Rajesh Revuru "Inter-Domain Diffserv Dynamic Provisioning and Interconnection Peering Study Using Bandwidth Management Point - A Simulation Evaluation" *The 2003 International Conference on Information Systems and Engineering*

[7] I. Gojmerac, T. Ziegler, F. Ricciato, and P. Reichl, "Adaptive Multipath Routing for Dynamic Traffic Engineering," *in IEEE Globecom,* (San Francisco), Nov 2003.

[8] M. Menth, J. Milbrandt, and S. Oechsner, "Experience-Based Admission Control (EBAC)," in *IEEE Symposium on Computers and Communications (ISCC)*, (Alexandria, Egypt), June 2004.

[9] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms, Second Edition*. MIT Press and McGraw-Hill, 2001. ISBN 0262032937. Section 22.3: Depth-first search, pp.540–549

[10] C. Bouras, D. Primpas, "An admission control and deployment optimization algorithm for an implemented distributed Bandwidth Broker in a simulation environment", *4th International Conference on Networking – ICN 2005*, Reunion Island, France, April 17 - 21 2005, pp. 766 – 773

[11] Ceid DiffServ NS web site (2006), "DiffServ Quality of Service in NS", (http://ru6.cti.gr/diffserv-ns/default.htm), (Accessed 24 March 2006)

[12] QBone Bandwidth Broker Architecture, http://qbone.internet2.edu/bb/bboutline2.html

[13] Techniques for End-to-End Quality-of-Service control in multi-domain IP Networks, http://homes.dico.unimi.it/~pagae/NPTLab/camp.html

[14] C. Chhabra, T. Erlebach, B.Stiller, D. Vukadinovic *"Price-based Call Admission Control in a Single DiffServ Domain"*, TIK-Report Nr. 135, May 2002

[15] L. Burchard, H. Heiss, "Performance Evaluation of Data Structures for Admission Control in Bandwidth Brokers", April 2002

[16] UCLP software (User Controlled LightPath), http://www.canarie.ca/canet4/uclp/uclp_software.html