The DAIDALOS approach to IP multicast Inter-domain QoS control

Emiliano Guainella, Claudio Sansone, Barbara Angelini and Noemi Angelini

Abstract—**During last years, mass interest services, like IP-TV, are increasing in number, thus paving the way to the wide deployment of the multicast routing technology. This paper describes a framework to control QoS for a multicast service involving a multi-domain environment. The QoS control is performed with the choice of the best available inter-domain multicast route in order to statistically respect end-to-end connection requirements. Extensive simulations confirmed the effectiveness of the proposed solution.**

Index Terms— **IP multicast, QoS, inter-domain, protocol independent multicast**

I. INTRODUCTION

IN the last few years the research community, network operators and content/service providers are investing a lot of resources on multicast and broadcast technologies. The reason of this interest is twofold: (a) there are more and more services with a mass interest (e.g. IPTV, video streaming, carousel application delivery) that could be consumed potentially by a larger number of users (whereas at present the number of receivers is limited by the bandwidth of the service provider) (b) broadcast and multicast technologies are becoming available for fixed and mobile users thanks to the early deployment of the IP Multicast backbones and the first trails on mobile broadcast networks like DVB-T/H (Digital Video Broadcasting Terrestrial and Handheld), UMTS MBMS (Multimedia Broadcast and Multicast Services), S-DMB (Satellite Digital Multimedia Broadcast), and the Asian T-DMB (Terrestrial DMB, derived from the DAB, Digital Audio Broadcasting) coupled with wireless (2/2.5/3G) interaction links ([1]–[2]). These conditions paved the way to studies on the control of *QoS* for *IP multicast* transmissions.

IP Multicast is a bandwidth-conserving technology that reduces IP traffic by simultaneously delivering a single stream of information to multiple recipients in the same or different networks, whereas broadcast technology delivers the same information to all users of a network. In details, when a data stream has to be delivered to many users, a multicast-aware network (and routers) is in charge of the stream routing in such a way that (i) it reaches all recipients, (ii) uses efficiently network resources avoiding multiple packets passing on the same path.

The scope of the authors' work was to specify an architecture and an inter-domain multicast protocol that, cooperating with the DAIDALOS QoS network entities, could assure the best performance as for end-to-end content and delivery quality.

This paper is organized as follows: section II describes the QoS framework of the DAIDALOS project; section III describes the technological challenges of the multicast and broadcast technologies; section IV describes the proposed QoS-aware multicast protocol; section V illustrates the simulation results; section VI summarizes our work.

II. END-TO-END DAIDALOS QOS ARCHITECTURE

This section gives a brief description of the network entities of the DAIDALOS QoS architecture permitting the end-to-end QoS assurance. The environment considered in the project is made up of many administrative domains, each one composed by a MPLS fiber optic core network that connects a large number of heterogeneous access networks [2]. These elements are:

--*QoS Broker* (QoSB): It has the task to admit new traffic and to configure consequently network elements (e.g. routers) within its domain. The QoSBs of adjacent domains interact in order to set up dynamic agreements (SLA/SLS) to handle QoS requirements of transiting traffic. It retrieves information on network status from the monitoring platform.

--*Policy Based Management System* (PBMS): It contains a database with rules and policies for network and fault management.

--*Central Monitoring System* (CMS): It has the task of monitoring the network performance, by active and passive measurements. Depending on policies, those measurements

---

Manuscript received February 6, 2005. The work described in this paper is based on results of IST FP6 Integrated Project DAIDALOS. DAIDALOS receives research funding from the European Community's Sixth Framework Programme. Apart from this, the European Commission has no responsibility for the content of this paper. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

Emiliano Guainella is with the Computer and System Sciences Department, University of Rome "La Sapienza", Via Eudossiana 18, Rome, 00184 Italy (phone: +390644585444; fax: +3906233200487; e-mail: guainella@dis.uniroma1.it).

Claudio Sansone is with the Information and Communication Science Department, University of Rome "La Sapienza", Via Eudossiana 18, Rome, 00184 Italy (e-mail: sansone@infocom.uniroma1.it).

Barbara Angelini and Noemi Angelini are with the Computer and System Sciences Department, University of Rome "La Sapienza", Via Eudossiana 18, Rome, 00184 Italy (e-mail: uor_team@dis.uniroma1.it).

can be performed between network elements or for aggregate paths. It sends alarm to the policy server in case of link failure or excessive congestion.

The entity in charge of managing inter-domain multicast is the Multicast Inter-Domain Entity (MIDE), which is strategically located inside the "Inter-Domain Rendezvous Point" (IDRP), the collector for every multicast transmission involving multiple domains. These entities are described in details in section IV.

III. IP MULTICAST: MAIN ISSUES

This section gives a brief overview on multicast technology and presents some technological challenges: QoS control and inter-domain environment handling.

The IP multicast packet is characterized by a multicast address in its header field. This special address refers to a large number of listeners, who belong to the same multicast group and want to receive simultaneously the same content.

In principle, depending on the position of the multicast listeners, it is possible to compute the optimum distribution tree that minimizes the cost of routing in terms of occupied links in the whole network. For the high degree of dynamicity of multicast groups, this may entail an unacceptable computational complexity. Moreover, calculating each time a new tree, it means to change the routing scheme thus causing transitory loops, data losses and jitter increases. Therefore, a sub-optimal resolution to construct a spanning tree is often considered and implemented in multicast protocols. The simplest sub-optimal approach is to build the spanning tree adding a single node at the time using, for example, the shortest path or the minimum cost path. This approach is used by the most widespread protocol for an intra-domain environment: the Protocol Independent Multicast – Sparse Mode (PIM-SM [4]).

PIM-SM uses a special router, called Rendezvous Point (RP), to share a single multicast tree among a large number of users and to discover the presence of new sources. Indeed, the receivers send IGMP or MLD JOIN requests ([5]–[6]) to their access routers, which forward PIM JOIN packets towards the RP. The sources, after a registration with the RP, send multicast data to the RP that redistributes packets to all receivers through the multicast tree.

As for QoS support, both PIM-SM and the other common routing protocols ([7]) support only best effort traffic. The term *QoS* has intrinsically a lot of slight different meanings, and especially in case of IP multicast it is important how QoS should be interpreted. For this reason, the assumptions on the QoS model considered in this work are:

--The content quality can be differentiated with the use of multi-layer multicast. This is possible thanks to the technological advances of scalable audio and video coding that permits to split a multimedia stream to a base layer with basic quality, that can be enriched with enhancement layers thus heightening the quality ([8]).

--The delivery quality, for as regards IP multicast, depends on how the multicast flows are handled within the network: for example, as in the scope of this work, with the use of the Diffserv model ([9]).

A QoS-aware multicast protocol, objective of this work, should guarantee for each joining user a path that satisfies its QoS requirements in terms of bandwidth, delay, jitter and loss. For this reason, efficient resource management functions are necessary (congestion control, connection admission control, path selection). At the same time, the multicast protocol should aim at maximizing the number of successful joins, minimizing the tree cost, minimizing the joining time, and being scalable to larger networks.

QoS-aware protocols existing in literature (e.g. QoSMIC and QMRP, [10]–[11]) are characterized by Multiple-Path Routing (MPR), namely the possibility, for a new member of a multicast group, to have multiple candidate paths to reach the nearest multicast router of the tree. In fact, having multiple paths is a crucial condition to be sure to find one with available resources that can satisfy members' requirements. Unfortunately, the effectiveness of the multiple-path search is limited by message overhead and existing MPR protocols are not able to scale to an inter-domain environment.

To scale the whole internet, multicast routing protocols need to be hierarchically divided in two main categories: (i) intra-domain protocols, which operate within an Autonomous System (AS), also called Administrative Domain (AD), and (ii) inter-domain protocols, which operate among ASs. The Border Gateway Multicast Protocol (BGMP, [12]) was the standard proposal for inter-domain IPv4 multicast. In addition to a lack of any QoS support, the BGMP assumes the existence of symmetric inter-domain links, which may conduct to the construction of sub-optimal trees. To solve this issues extensions of BGP (namely BGPv4+ [13]) make symmetric routes for multicast traffic.

To support inter-domain multicast in IPv6, the IETF proposes to use PIM-SM. To discover RP of other domains, its address must be embedded within the IP multicast address with a special algorithm [14]. Another option is the using of PIM Source Specific Multicast (PIM-SSM), a special instance of PIM-SM with a single source.

IV. QOS-AWARE INTER DOMAIN MULTICAST PROTOCOL

This section describes the proposed QoS-aware Multicast Inter-Domain routing protocol (MID), which extends PIM functionalities allowing the construction of inter-domain trees and guaranteeing the respect of QoS requirements of next generation services.

We introduced a Multicast Inter-Domain Entity (MIDE) in each domain to search and test multiple paths from domains already in the multicast group (tree-domains) towards a new joining-domain.

The path search is carried out in two steps. First, some tree-domains are selected as "searching-domains". Second, each searching-domain chooses the best paths between itself and the joining-domain and checks the available resources along these paths. The receiver-domain selects and allocates the best feasible one among the tested paths. Both searching-domain and path choices are based on the multi-objective optimization theory and take in consideration AS topology, tree topology, available resources, achievable QoS (in terms of delay, jitter and loss), and tree cost (calculated as sum of all the tree links).

To test available paths, the MIDE collects information from other domain's entities:

- AS connectivity information from BGP Border routers by configuring the MIDE as a BGP speaker and BGP UPDATE messages.
- Policies and SLAs from PBMS, used to filter routes and estimate the allowed traffic to/from neighbor domains.
- Multicast resource allocation and aggregated edge-to-edge measurements from CNQoSB and CMS.

As previously said, the MIDE coincide with the IDRP, which is the local Rendezvous Point, present in each domain, where all PIM messages interesting inter-domain transmissions are forwarded to by local access routers. This process allows receivers to discover the sources in other domains. In fact, each source MIDE (located in source domain) registers itself in the RP MIDE associated to the multicast group (that could be located in a foreign domain), creating an association between its IP address and the assigned multicast address. Moreover each RP MIDE has a set of unique inter-domain multicast addresses from which can be derived its IP address, thanks to the address embedding technique [15]. Therefore, from a multicast address, the USER MIDE can obtain the relative RP MIDE IP address, whereas the RP MIDE can obtain the relative source IP address.

Fig.1 illustrates the interdomain multicast signaling for source registration and user join. To begin an inter-domain multicast session, a source (through the access router) sends a PIM REGISTER message to its MIDE (indicated in figure as Source-MIDE). The Source-MIDE discovers the RP MIDE address from the multicast address and forwards a MID REGISTER message to the corresponding RP-MIDE. The RP-MIDE registers the source and sends a PIM JOIN request to the source MIDE, forwarded to the source server completing the session initialization.

A new user joins to the inter-domain multicast group sending an MLD_Report (Join) packet to its Access Router (AR). The MLD_Report reports an admission control request to ANQoSB, that can't be immediately answered for the lack of QoS information. The AR recognizes *inter-domain* multicast address by the SCOPE field introduced by IPv6 ([14]), and it sends immediately a PIM JOIN packet towards the MIDE of its domain (User-MIDE).

The User-MIDE extracts the RP MIDE IP address embedded in the multicast address and sends a MID JOIN packet to the RP-MIDE. The RP-MIDE replies with QoS requirements associated to the group. At this time, an interface with the ANQoSB permits to admit the multicast service in case there are enough resources. The ANQoSB can then configure queues in the AR.

The RP-MIDE selects some searching domains in the tree and sends a MID SEARCH_PATH packet with QoS requirements to each selected Tree-MIDE. Notice that even the RP-MIDE could be selected as searching-domain.

The selected Tree-MIDEs check the available resources,

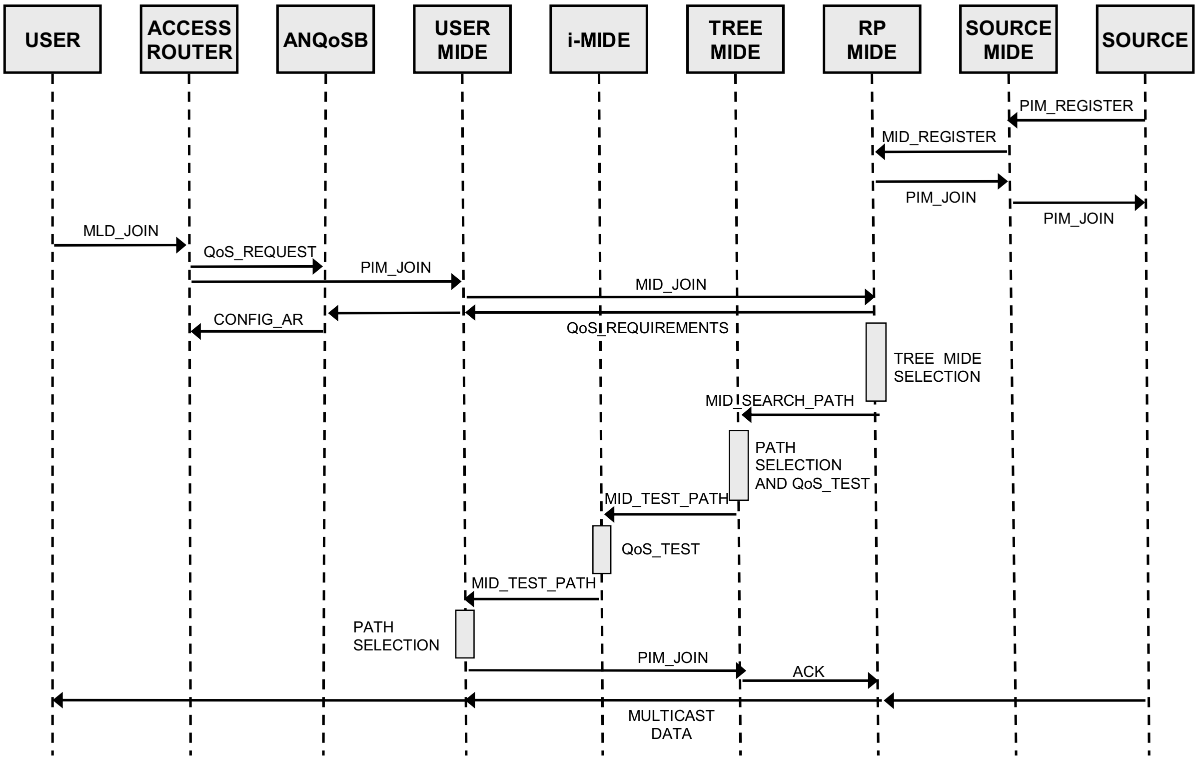

Fig. 1. Multicast Signaling for source registration and user join

choose the best paths towards the User-MIDE and send a MID TEST_PATH message on each selected route. Each MIDE along a tested path (i-MIDE) executes a QoS evaluation test and, if the test results positive (that is the QoS provided in the path towards the next domain is enough for the service requirements), it updates and forwards the MID TEST_PATH message to the next domain.

All the MID TEST_PATH packets are collected by the User-MIDE. If no candidate path with adequate QoS assurances is found, further paths could be tested from the same or new searching-domains. Otherwise, the User-MIDE selects the best path among the feasible candidates and adds it to the multicast tree by a back-reverse PIM JOIN message sent to the Tree-MIDE. The PIM JOIN is routed on the selected path by adopting IPv6 route select feature or by forwarding it MIDE-by-MIDE in the path. If the path-allocation eventually fails (for packet loss or network conditions changing), another candidate path could be allocated or other paths could be tested.

When the user join procedure is completed, the Tree-MIDE informs the RP-MIDE (ACK) and the user is able to receive the multicast content (MULTICAST_DATA).

The described join procedure is completely executed only for the first join request in a domain. The following requests in a tree-domain require only a PIM JOIN request towards the User-MIDE and so they are much faster.

V. SIMULATIONS

This section describes the simulations done to evaluate the performance of the proposed protocol. We implemented the MID protocol using the OPNET Modeler 9.0 and compared its performances with the PIM-SM protocol. The tests were performed on over 100 node networks generated with Waxman model [16], with graph average degree (average number of links per node) ranging from 3 to 6. The network load is modeled as the average percentage of used resources in every domain over the maximum allowed, calculated with uniform or normal probability distributions. The more resource is used and the worse performance is provided by a domain. Table I summarizes the main adopted assumptions and parameters' values for the tests.

Fig.2 shows some results of our tests for two possible graph degrees (4 and 6). As performance metrics we used join success probability, end-to-end delay, average tree cost, join latency and control message overhead.

To evaluate the *join success probability* we considered the respect of QoS Requirements (delay, jitter and loss) by the allocated paths. The MID adopts an admission control mechanism, therefore a branch is not allocated if it does not respect the QoS requirements or there are not enough resources. On the contrary, the PIM is not QoS-aware and so allocates a new tree branch even without respecting QoS requirements, but we consider this join as failed. Observing the join statistic, it is clear that MPR and QoS test adopted by MID increase the probability of a successful join.

TABLE I
MODELING ASSUMPTIONS AND PARAMETERS ADOPTED IN THE SIMULATIONS

Parameter	Values - Assumptions
Tested Protocols	PIM-SM, MID
Performance Metrics	Join success probability, end-to-end delay, tree cost, control overhead, join domain latency
Network and Load Scenario	
Graph model	Waxman model
Graph average degree	From 3 to 6
Node number	100
Source nodes (%)	10
Receiver nodes (%)	20 (for each group)
Load Distribution	Uniform and Normal (mean from 30 to 70 at step of 10, variance 30)
Source- Receiver Selection	Uniformly Distributed
QoS Requirements	
Maximum delay	500 ms
Maximum jitter	50 ms
Maximum loss	0.001 %

The *end-to-end delay* is the mean delay of the multicast data from the source to the receivers. It is a good metric to evaluate the "join quality" reached by the two tested protocols. The PIM allocates a new tree branch always on the shortest path, even if this path is congested and so presents high hop-by-hop delays. Our protocol is able to avoid congested domains using alternative paths with better performance obtaining smaller end-to-end delay.

The *tree cost* is evaluated as total number of the tree links. The tree cost graphic shows that our protocol obtains a lower tree cost than PIM. Indeed the MID often allocates new branches from on-tree domains, whereas the PIM always tries to join new domains to the source. For this reason the MID generally constructs trees with less branches and then a minor cost than the PIM's trees.

Note that in the graphics it is possible to observe very low values of delay and tree cost for high network load. This is due to the MID's admission control mechanism that limits the tree growth if there are not enough resources, and so avoid increasing of network congestion.

The proposed protocol has higher values of *joining delay* than PIM due to our multiple path searches. In PIM the branch creation is completed when the user's join request message reaches the first on-tree router. In MID the first join request in a domain has to reach the multicast source and wait for the end of path test procedure; the following join request of any on-tree domain will have the same delay as PIM. Therefore the MID requires more time to join a new domain than PIM, but this delay should be considered as a small cost to achieve a better QoS.

Finally, to test simultaneously multiple paths the MID sends more control messages than PIM. This cost should be observed in perspective and compared with the overall *overhead* incurred over the life of a multicast session because most of existing multicast protocols, and in particular PIM and MID, use a periodic refresh mechanism. The initial search overhead is a one time cost per domain (not per user in the domain), but in the complete life cycle of a multicast session it becomes negligible. The cost for any following join in a tree-domain and for route refresh is then the same for PIM and MID.

VI. CONCLUSIONS

In this paper the possible interaction between networks elements that will be present in Next Generation Networks and the multicast technology has been investigated. In that respect, an innovative multicast inter-domain entity (MIDE) and an inter-domain QoS-aware multicast protocol have been proposed. The performance of our Multicast Inter-Domain (MID) protocol have been evaluated with extensive OPNET Modeler simulations and compared with PIM. The simulation results show that MID achieves better join success probability, QoS and tree cost than PIM with a slight increment of the joining delay and initial message overhead.

REFERENCES

[1] 3GPP specification, Introduction of the MBMS in the Radio Access Network, TS25.346.

[2] DVB-CBMS working group, PDC in DVB-H: Technical Requirements v1.0.0 Rev.1, TM3095 Rev.2. Available: http://www.dvb.org

[3] DAIDALOS Deliverable 3.2, August 2004. Available: http://www.ist-daidalos.org

[4] PIM WG, "Protocol Independent Multicast - Sparse Mode (PIM-SM) Protocol Specification", IETF Draft <draft-ietf-pim-sm-v2-new-10.txt>, July 2004.

[5] B. Cain, S. Deering, I. Kouvelas, B. Fenner, A. Thyagarajan, "Internet Group Management Protocol, Version 3", IETF RFC 3376, October 2002.

[6] R. Vida, L. Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", IETF RFC 3810, June 2004.

[7] M. Ramalho, "Intra– an Inter– Domain multicast routing protocols: a survey and taxonomy", IEEE communications Surveys and Tutorials, First Quarter 2000.

[8] J Liu et al., "Adaptive video multicast over the internet", IEEE Computer Society, January-March 2003.

[9] S. Blake, D. Clarke et al., "An Architecture for Differentiated Services", RFC 2475, December, 1998.

[10] Shuqian Yan, M. Faloutsos, A. Banerjea, "QoS-aware multicast routing for the internet: the design and evaluation of QoSMIC", Transactions on Networking, IEEE/ACM, Vol. 10, Issue 1, pp. 54-66, February 2002.

[11] S. Chen, K. Nahrstedt, Y. Shavitt, "A QoS-aware multicast routing protocol", Selected Areas in Communications, IEEE Journal on , Vol. 18, Issue 12, December 2000.

[12] BGMP WP and D. Thaler, "Border Gateway Multicast Protocol (BGMP): Protocol Specification", IETF Draft <draft-ietf-bgmp-spec-06.txt>, January 2004.

[13] Marques, P., Dupont, F., "Use of BGP-4 Multiprotocol, Extensions for IPv6 Inter-Domain Routing", IETF RFC 2545, March 1999.

[14] B. Haberman and D. Thaler, "Unicast-Prefix-Based IPv6 Multicast Addresses", IETF RFC 3306, August 2002.

[15] P. Savola et al, Embedding the Rendezvous Point (RP) Address in an IPv6 Multicast Address, RFC 3956, November 2004

[16] B.M. Waxman, "Routing of Multipoint Connection", Selected Areas in Communications, IEEE Journal on, Vol. 6, Issue 9, 1988.

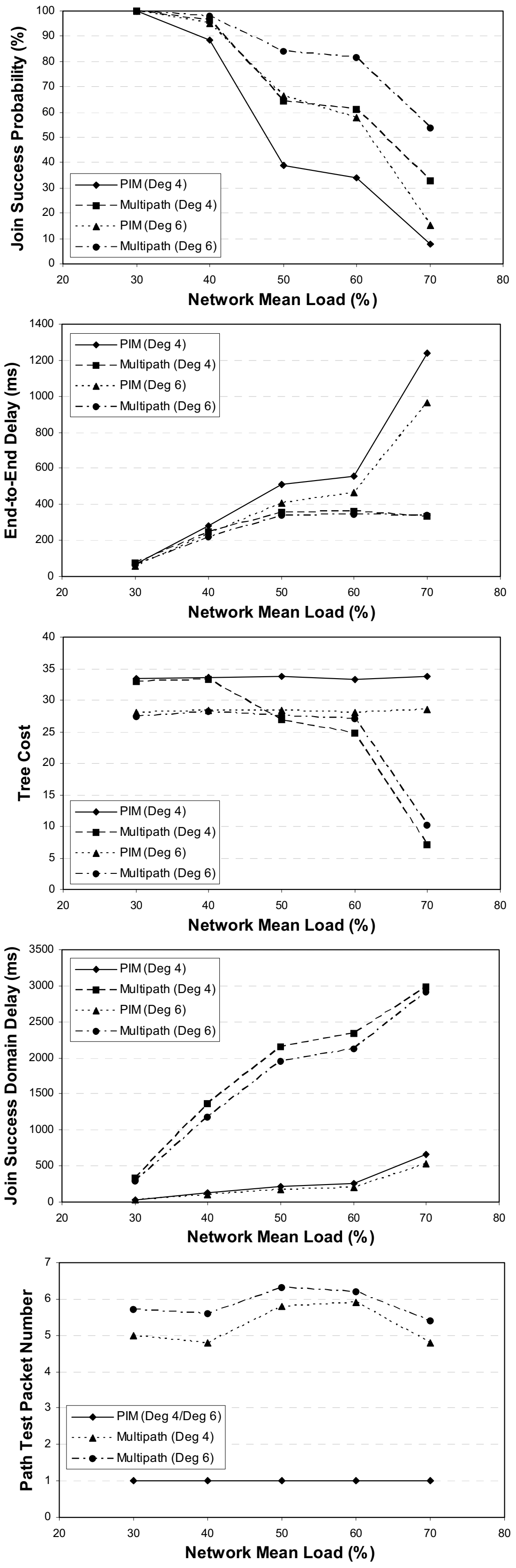

Fig. 2. Simulation Results