This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the ICC 2008 proceedings.

iREX MPO : A multi-path option for the iREX inter-domain QoS policy architecture

Ariffin Datuk Yahaya
School of Information & Computer Science
University of California, Irvine
Email: ariffin@ics.uci.edu

Tatsuya Suda
School of Information & Computer Science
University of California, Irvine
Email: suda@ics.uci.edu

***Abstract*—The inter-domain Resource Exchange (iREX) architecture uses economic market mechanisms to automate the deployment of end-to-end (E2E) inter-domain (ID) quality of service (QoS) policy among resource consumer and resource provider Internet Service Providers (ISPs). In iREX, each policy reservation is deployed on a single E2E ID path made up of the most "desirable" (i.e. cheapest and least congested) ISP resources.**

To accommodate ISPs that prefer redundancy when deploying ID QoS policy, in this paper we introduce an extension to the iREX architecture that gives an originating ISP a multi-path option (MPO) when deploying a reservation. MPO takes an initiating ISP's preference for redundancy and provides information about the available path options to achieve this preference in a distributed manner. Our simulation results show that while providing redundancy to the originating ISP using MPO does increase its resource costs in accordance to an ISP's preference, it only marginally increases overhead, and does not affect overall network performance – in fact the use of MPO lowers congestion.

***Index Terms*—multi-path QoS routing; path coupled charging; inter-domain QoS policy; resource allocation and management; network control by pricing; economics.**

I. INTRODUCTION

MANAGING an Internet domain's policy to offer Quality of Service (QoS) to select traffic flows is an important networking research area. Presently, any number of domain Internet Service Providers (ISPs) can create and install policy to selectively support multiple traffic flows with different QoS specifications within the domain(s) that they control. However, because no single ISP controls all the domains in the Internet, deploying end-to-end (E2E) inter-domain (ID) QoS traffic flows (*requirements*) must involve negotiating for, and propagating the policies to support the traffic flows' QoS specifications (QoS policy) with transit ISP domains that own ID QoS resources along the ID path to the destination.

A. iREX

A distributed method to automate the deployment of E2E ID QoS policy has been suggested by the inter-domain Resource Exchange architecture (iREX) [1] [2] [3]. iREX facilitates a distributed economic system where ISPs trade in network resources to deploy policy supporting aggregated individual ID QoS flows (i.e. jumbo flows [4]).

We use the term *resource* as an abstract ID network transport service defining ownership and transport responsibility starting from a domain's ingress border router, going through the domain and ending at a neighboring domain's ingress border router. iREX research explores the application of economics within network management automation to solve non-technical *human* problems like "ownership" and "trust".

B. The Need for Redundancy

Increasing the use of redundancy when deploying ID traffic and splitting aggregated user traffic to use multiple paths is preferable in terms of *reliability* because a non-redundant path makes for a single point of failure. Traditionally, ISPs have had a choice to use redundant policy driven paths rooted at the source based on the information presented by its first-hop neighbors.

Assuming the probability of any link in the Internet failing to be equal, splitting traffic by using more redundancy (i.e. beyond the first hop) will mean that less traffic will be affected by a link failure. However, increasing redundancy by splitting the deployment after the first hop is either difficult or impossible due to 1) the lack of information about the network away from the ISP's immediate area, and also 2) the lack of a mechanism to actually initiate such a deployment. For the same reasons, the current design of iREX is also limited to a single peer selected "most desirable" path when deploying reservations.

While we have shown in [2] that, in the worse case, iREX's fault tolerance mechanism can recover from resource failure in less than 1.2 seconds, this deployment re-routing mechanism is limited to the available alternate resources at the time of failure. Deploying policy with redundancy is a prevention measure that works to minimize the traffic that needs to be re-routed during a failure – further increasing deployment reliability.

C. Current Focus

In this paper, we introduce an extension to the iREX architecture that allows iREX ISPs to split a requirement into

---

This work is supported by the NSF through grants ANI-0083074, ANI-9903427 and ANI-0508506, by DARPA through grant MDA972-99-1-0007, by AFOSR through grant MURI F49620-00-1-0330, and by grants from the California MICRO and CoRe programs, Hitachi, Hitachi America, Hitachi CRL, Hitachi SDL, DENSO IT Laboratory, DENSO International America LA Laboratories, NICT (National Institute of Communication Technology, Japan), NTT Docomo and Novell.

multiple deployments at the time the reservation is made. The split deployments utilize different downstream sub-paths in a shape of two bushy trees joined at the leaves and rooted at both the originating and destination ISPs. iREX ISPs can increase the reliability of their deployments by using the new multi-path option (MPO) to "buy" redundancy when initiating a reservation and still rely on iREX's fault tolerance mechanism after the reservation is deployed.

To use iREX with MPO, the originating ISP specifies its "price tolerance" in terms of a percentage it is willing to pay more than the current cheapest path. We will show that there is a direct correlation of the price tolerance to the increase in redundancy.

The next sections will cover an overview of the current iREX architecture (II), details of the new MPO scheme (III), our simulation design (IV), and our numerical analysis (V). We close with a related work (VI) and a conclusion (VII) section.

II. iREX ARCHITECTURE

The iREX architecture is based on the "Posted Price Competition" economic model in which providers independently choose prices that are publicly communicated to consumers on a take-it-or-leave-it basis as characterized by Abbink and Brandts [5]. In the ID QoS context, domains are both providers and consumers at the same time because while domains have ID QoS resources that they can "sell", they also need to "buy" resources to deploy their own ID QoS requirements.

The iREX architecture contains a set of routing, resource reservation, deployed resource fault tolerance and resource reputation score maintenance protocols. Domains that support the iREX protocol form a loose community (*iREX market*) that exists for the sole purpose of trading in ID network resources; members of the iREX market cooperate in a competitive manner to facilitate ID resource selection by maintaining information about the desirability of resources within the iREX market, and by supporting the deployment of E2E ID QoS policy. With iREX, an ISP's profitability and long-term existence is determined by this economic market.

A. Selling: iREX Path Vectors

To evaluate resource *desirability*, iREX ISPs use "real" *resource price* in the form of "*monetary unit* per *time unit* per *bundle of resource*" and *resource reputation score* in the form of "*number of complaints* against a resource". Domains choose desirable resources to form chains of ID resources leading to destination domains in the form of path vectors (*iREX path vectors*). iREX path vectors use the current *total price* per unit time as a routing metric. The total price per unit time to a destination is determined by adding the price of each component ID resource used in forming the iREX path vector to that destination.

iREX ISPs decide on a selling price for each of their available ID QoS resources based on how its customers are projected to use (*demand*) the available resource. The domains then incorporate these prices into the current cheapest known iREX path vectors and advertise the path vectors to their neighbors. Domains receiving these advertisements evaluate and filter the received iREX path vectors by first excluding those that use resource with bad reputation scores, and then selecting the path vectors with the cheapest total price to each destination domain.

By the periodic advertisement of known iREX path vectors and the filtering of received advertisements, iREX path vectors formed using the cheapest reputable ID QoS resources propagate to all domains within the iREX market. iREX ISPs have total autonomy to determine their own ID QoS deployment path(s) using source routing, making total convergence for iREX path vectors unnecessary (in contrast to BGP's [6] use of path vectors).

iREX path vectors are predicated on economics where network resources are treated as a scarce good. iREX ISPs are assumed to always act selfishly to balance two priorities, 1) the ISP's responsibility towards its currently active customers, and 2) the ISP's desire to compete and make money selling ID QoS resources. When an ISP advertises an updated resource price, the distributed iREX path vector filtering mechanism will evaluate the new price, and the resource may either be more or less desirable when compared to other competing ID resources advertised by other domains.

Domains support the deployment of E2E ID QoS to earn revenue, and maintain good quality service to maintain untarnished reputations so that they are not blocked from supplying the market.

iREX assumes that each different QoS traffic specification is standardized into its own resource commodity. Multiple iREX markets may exist; each specializing in a QoS traffic specification and maintaining a set of iREX path vectors formed using the most desirable resources with this specification.

B. Buying: Resource Reservation

To deploy ID QoS policy using iREX, a consumer domain first identifies the entire ID path by referring to the current iREX path vector leading to the destination domain. If the total price is acceptable to the consumer ISP, it will initiate ID QoS policy deployment with the domains along the identified path by reservation request signaling.

Resource prices quoted and advertised by provider ISPs can fluctuate continuously leading to dynamics in the iREX path vector, but a successful reservation request freezes the reserved ID path and its associated prices for the duration of the reservation. Multiple ID QoS policies using different ID paths may be concurrently deployed for the same origin-destination pair depending on the iREX path vector at the time of each policy's reservation.

Each ISP that agrees to participate in the deployment of an ID QoS policy maintains the policy by monitoring resources to the next hop being used. Upon detecting a resource failure, an ISP will recover the policies affected by the fault either by suggesting a redeployment path directly to the consumer ISP, or by signaling a failure to the previous hop domain used by the policy so that it may help find an alternate path to recover the fault. In this way, should a fault occur, a source domain will either receive a suggestion to reroute or a fault signal, and can then make a decision on recovering the affected policies.

C. iREX economics: Congestion Avoidance

Lower demand for an ID QoS resource will result in an abundance of that resource, which is accompanied by a lowering of the risk of impacting the ISP's current customers and a lowering of revenue as domains in the iREX market use cheaper alternate resources. To increase revenue, the ISP will seek to increase the demand for its ID resources by lowering the price of the resource.

Higher demand for an ID QoS resource will result in resource scarcity, which is accompanied by the risk that a ISP's current customers will receive decreased QoS and complain. To compensate for the increased risk, the ISP will increase the price of its ID resources. Increasing the price also effectively decreases demand.

Lower prices accompany lower resource use, and higher prices accompany higher resource use, therefore choosing reputable resources that are cheaper translates directly to choosing conforming resources that are less congested. In this manner, economics and reputation are used to dynamically change iREX path vectors to include the cheapest reputable resources, which also translates into using the least congested conforming resources.

III. iREX WITH THE MULTI-PATH OPTION

MPO for iREX is a distributed way to increase redundancy by deploying reservations that utilize downstream sub-paths. To add this capability, we modified the iREX reservation packet to include a new *MPO field* carrying a floating point number signifying the source ISP's "price tolerance". An ISP's *price tolerance* is a percentage that the ISP is willing to pay more than the current cheapest price for alternate paths to the destination. Other modifications include new protocol functionality to communicate available alternate path information (MPO advisory) to, and "multi-path" decisions from the source ISP.

iREX source ISPs with a preference for policy deployments with redundancy and a *willingness to pay* for it, can use MPO to convey this preference to resource owner ISPs. Actually getting and *having to pay* for a preferred level of redundancy depends on the existence of available paths at intermediate ISPs that satisfy the price tolerance, and on the willingness of intermediate ISPs to participate. In this section we will go through an overview of MPO (III-A) and work through an example (III-D) of its use.

A. iREX Reservation with MPO

Keeping within iREX's protocol ideology of respecting each ISP's autonomy and preference to decide its risk tolerance, market aggressiveness and level of user satisfaction, multi-path using MPO occurs as a series of real-time "negotiations".

During an iREX reservation using MPO, a source ISP (consumer) first chooses a path for QoS deployment by referring to its current iREX path vectors, which will give it information on the neighboring domain that is currently advertising the cheapest path to the destination. At this time the consumer may also decide to split the QoS deployment according to the available information from multiple single hop neighboring domains.

Once the consumer has chosen a deployment path to the destination, the consumer then decides on a preference for downstream redundancy for the path. The consumer conveys this preference by filling the MPO field with its price tolerance before initiating the reservation.

B. Selling with MPO

Upon receiving an iREX reservation packet and each resource owner ISP (provider) along the source chosen deployment path to the destination will first perform the usual iREX process of checking price and resource availability, then it will check the MPO field. If the MPO field is a non-zero, the provider has the option to participate and search for available alternate paths to suggest to the source ISP.

To participate, the provider calculates a maximum price that the consumer is willing to pay by increasing the reservation price from itself to the destination by the percentage specified by the MPO field. The provider then searches its path database (from all its neighbors) for a list of paths to the destination that are within the calculated maximum price.

If the provider chooses not to participate in the consumer's MPO request or does not find any alternate paths within the maximum price criteria, the provider then continues processing the reservation normally. However, if the provider chose to participate and found some alternate paths, the current reservation is put on hold and information about the alternate paths and their associated prices are sent directly to the consumer in an *MPO Advisory* message.

There is a twofold economic incentive for a provider to participate in MPO. Firstly a provider remains competitive by participating because *not* participating may cause multi-homed consumer domains with a high preference for redundancy to select an alternate path bypassing the non-participating provider. Secondly, by participating in MPO a provider is able to gain surplus by selling expensive links that were previously deemed uncompetitive and excluded from the iREX market.

C. Buying with MPO

A consumer receiving an *MPO Advisory* message has to make a decision on whether to split its reservation or to ignore the advisory. To ignore the advisory, the consumer sends an *MPO Advisory Reject* message back to the sender, at which point the sender will continue processing the original reservation. To accept the advisory, the consumer has to decide on a reservation split among the available paths.

The original reservation is resized by sending a *Reserve Update* message reflecting the new reservation size and path along the original path; when the *Reserve Update* message reaches the sender of the *MPO Advisory* message, it will be converted into a *Reserve* message and reservation continues along the original path with the new size. To deploy the remaining split requirements, the consumer simply issues new reservations using those paths.

Any *Reserve Update* or *Reserve* message generated from an *MPO Advisory* will also include information on which domain issued the advisory to prevent a reservation to be split by the same domain more than once. We expect a consumer to split reservations based primarily on physical and logical constraints (e.g. by domain user or flow segment) and secondarily on the prices of the available paths (i.e. bigger flows on cheaper paths).

D. An MPO Example

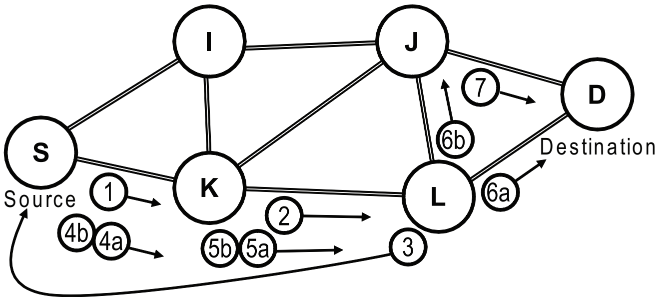

Fig. 1: iREX:MPO Scheme

We will now use the network of ISP domains in Fig. 1 to illustrate an MPO example. In our example, S is initiating a reservation along the path SKLD with the MPO price tolerance set to $x\%$ using *Reserve* message **1**. Note that for this example 1) we will only be showing the "reserve" (forward) part of the reservation process and not the "confirm" (reverse) process, and 2) the messages in Fig. 1 are numbered with respect to the reservation's progress.

When K receives *Reserve* message **1**, it decides to participate and searches for alternate paths fitting the price tolerance criteria, but not finding any, simply records the reservation and propagates the reservation according to the initial path SKLD by sending *Reserve* message **2** to L.

Upon receiving *Reserve* message **2**, L also decides to participate and it finds alternate path LJD fitting the price tolerance criteria – L then sends S *MPO Advisory* message **3**. S decides on a split for the requirement being reserved and first sends *Reserve Update* message **4a** to resize the initial reservation, and then initiates a new reservation for the residual requirement with *Reserve* message **4b** on the path SKLJD learnt from the advisory.

When K receives *Reserve Update* message **4a**, it looks up and resizes the previous reservation and continues the reservation by propagating the reservation update according to the initial path SKLD by sending *Reserve Update* message **5a** to L. K also receives *Reserve* message **4b**, and still having no alternate paths to satisfy the price tolerance criteria, K records and propagates the reservation on the new path SKLJD by sending *Reserve* message **5b** to L.

When L receives *Reserve Update* message **5a**, it propagates the update on path SKLD by sending *Reserve Update* message **6a** to D. Having no previous record of the reservation referred to by this message, D will convert the received *Reserve Update* message **6a** into a *Reserve* message and process it as such and the forward reservation process completes on path SKLD.

Backtracking a little in our example, L also received *Reserve* message **5b**, which L recorded and propagated by sending *Reserve* message **6b** to J. Even though the MPO field was set, L did not process it because it noticed that *Reserve* message **5b** was a result of its advisory. The forward process of the reservation on path SKLJD completes without further splitting (i.e. J did not have any alternate paths that fit the criteria) with *Reserve* message **7** to the destination D.

Using MPO, the source S has now completed the reservation process on both paths SKLD and SKJD. In practice, with the help of other ISP domains, there may be other possible splits including ones that originate at S itself. Note that while S may deploy split reservations that originate at S by only using information from direct neighbors K and I, without information that MPO facilitates from L, S had no way of knowing that the downstream sub-path LJD was available.

IV. SIMULATION DESIGN

The iREX simulator (available at [7]) models a user specified network topology of ISP domains with each domain having two main goals: 1) to fulfill a set of generated ID QoS demands for its local domain users, and 2) to sell available resources to other domains. The simulator does packet level simulation for control packets used for iREX and BGP signaling, and flow level simulation for the deployment of QoS flows. In order to evaluate the iREX architecture with MPO we modified the iREX simulator by adding MPO capabilities to the current functionality.

For our simulation, each domain has a maximum 1 minute advance knowledge of its own current ID QoS demands, and reservations need to be initiated ad-hoc with other domains using iREX signaling. Domains within the iREX simulation have no knowledge outside of their own domain other than those received through BGP and iREX signaling. Each ISP prices its resources independently with a goal to minimize its risk and maximize its revenue. Domains also respond to other domains' reservation requests. To set the price of a domain's resources, the simulator uses a Squared (i.e. $price = (current\ used\ bandwidth)^2$) price function that conveys an ISP's aversion to risk as explained in section II-A. Signaling resource advertisements for iREX path vector maintenance is done at 3 minute intervals.

To simulate MPO using real world topologies we chose the Very High Performance Backbone Network Service (vBNS) and L3 ISP topologies with each point of presence representing an ISP domain. ISP domains are assumed to be connected with OC48 optical fiber links to its neighbors and the length of each link is calculated to be the actual beeline distance between the cities. Fig. 2 and Fig. 3 illustrate the chosen topologies. The L3 topology was chosen in addition to the vBNS topology to show the effect of adding ID links to the same cities.

ID QoS demands within the simulator are viewed as "bundles" of traffic sized at 0.2% of line speed (about 4.8mb/sec)

Fig. 2: vBNS Topology

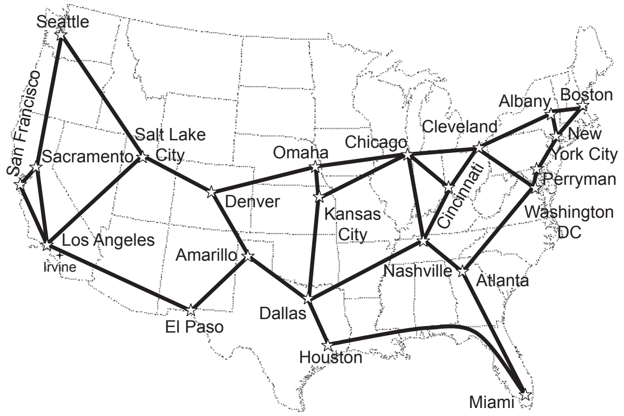

Fig. 3: L3 Topology

with a 5 minute average reservation duration. The total projected bandwidth usage (*traffic load*) is determined according to a percentage of each domain's actual total egress capacity in the vBNS topology from 0% to 100% in 4% steps – we used the same traffic on the L3 topology. To generate domain level reservation requests, we used a simple Poisson arrival model with parameters derived from $M/M/\infty$ analysis.

To show the performance of MPO, we configured all domains in the simulation to always accept multi-path suggestions resulting from a preset MPO price preference and compared it to a reference simulation using the same configuration but without using MPO. When any consumer domain receives a multi-path suggestion, it splits the current reservation into the available number of paths according to a reverse ratio with respect to the prices of the paths (i.e. cheaper paths get more reservation bandwidth).

Only the 20% preset MPO price tolerance configuration data set is included due to space limitation. The selection of a price tolerance level is quite subjective, but we chose the 20% configuration data set because it is representative of our general findings at other MPO price tolerance levels. The results presented in the next section (V) were averaged from 44 independent simulation runs each using a different set of generated traffic that collectively total more than 100 million simulated reservation requests. The high number of simulated reservations were necessary to achieve a level of convergence because at a price tolerance of 20%, the MPO data set was only a small subset of the data generated by the iREX simulator.

V. NUMERICAL ANALYSIS

We have three basic goals in this section: to show how well MPO increases redundancy (i.e. using multiple sub-paths), to show the effect MPO has on the network in general, and to show the overheads of using MPO.

A. Increasing Redundancy

To show how well MPO increases redundancy, we present two metrics, the MPO Ratio and the Number of Unique Paths.

The MPO Ratio metric is defined as the number of QoS traffic flows deployed as a result of an MPO advisory divided by the total number of deployments. Fig. 4 shows the MPO ratio of iREX while varying traffic load. We use the MPO Ratio metric to show how well ISP preferences for MPO are being fulfilled.

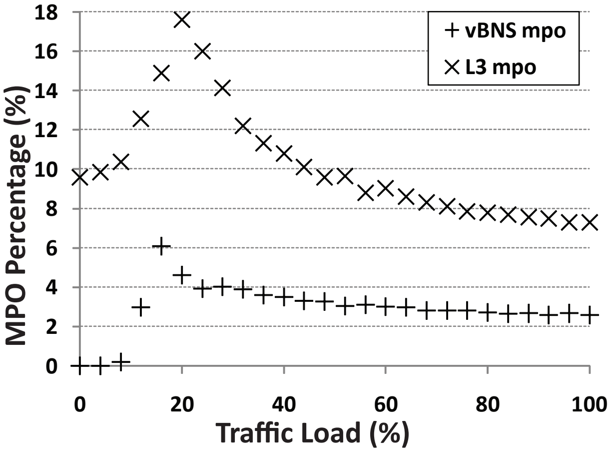

Fig. 4: MPO Ratio

Even though the MPO price tolerance was set to 20%, for the vBNS topology, traffic flows resulting from MPO peaked at only about 6% of all traffic when traffic load was 16% and then leveled off to an average of 3%; this is because of the limited connectivity within the vBNS topology. In comparison, the peak for the L3 topology which has increased connectivity is about 18% at about 20% traffic load. In both topologies, MPO Ratio is higher when traffic is not congested (i.e. at lower traffic loads) because as traffic loads increase, MPO traffic is also competing for scarce resources with other deployments.

"State" information held by ISPs supporting MPO will increase proportionately to MPO use, and the MPO Ratio metric also indirectly represents this overhead increase. A deterrent that counters the increase of this overhead is the price that the source domain has to pay for each redundant deployment.

Number of Unique Paths is defined as the number of unique deployed paths going from point A to point B. By unique we mean that the path has at least one segment or sub-path link different from other paths. In this case, we looked at deployments from Los Angeles (i.e. A) to Boston (i.e. B) on both topologies and tabulated the number of unique paths across the total number of deployments.

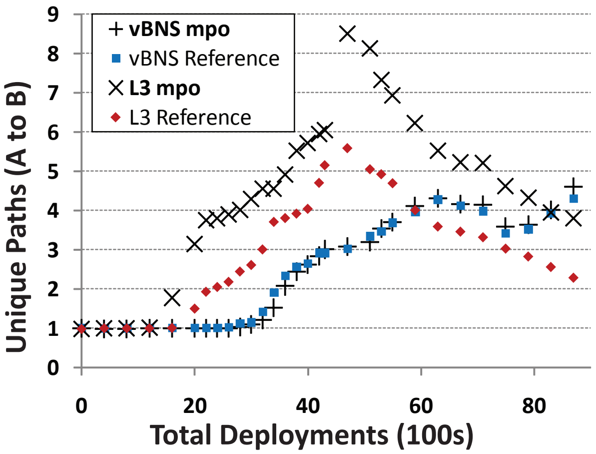

Fig. 5: Unique Paths

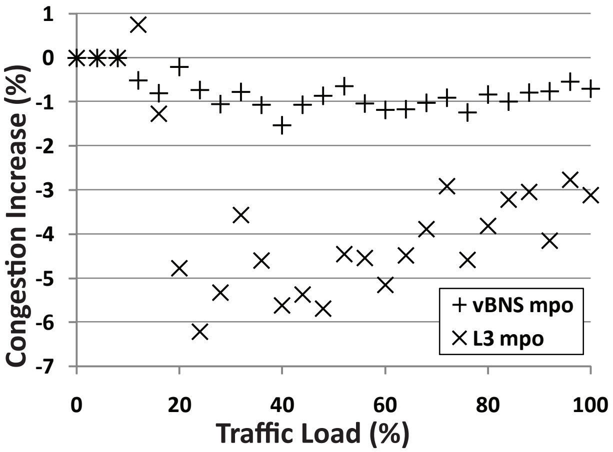

Fig. 6: Congestion

Fig. 5 shows the number of unique paths used with respect to the number of deployments for each topology – both with and without using MPO (denoted by "Reference" in the legend). Note that while not exact, there is a correlation between the total number of deployments (i.e. the horizontal axis) to the traffic load. We chose the total number of deployments over that of traffic load because we placed priority on showing a more detailed picture of the situation.

For the L3 topology, we can clearly see that when MPO is used, there is an increase of unique paths of about 20% on average. However, the vBNS topology does not show this very clearly with the MPO deployments seeming to track the Reference (i.e. not using more unique paths). This is because while successful MPO deployments guarantee reservations using redundant paths, the paths used for redundancy may already be in use by other deployments since iREX itself is also seeking out and using available paths when deploying single path deployments – this causes a problem for this metric since in the vBNS topology, paths are not as plentiful as the L3 topology. We show this metric to give another perspective to the MPO and non-MPO policy deployment mix illustrated by the previous metric (MPO Ratio in Fig. 4).

B. *Effect on Network*

To show the effect of MPO on the network, we present the Congestion metric. We define a link as congested when more than 50% of its bandwith is in use. The Congestion metric compares simulation results for deployments using MPO to a reference that does not by showing the increase (i.e. $MPO\ congestion - non\ MPO\ congestion$) in the congestion experienced by the network when MPO is in use.

Fig. 6 shows the Congestion metric for both the vBNS and L3 topologies showing congestion increase on the vertical axis while varying traffic load. Both the vBNS and the L3 topologies show a negative increase (decrease) of congestion when using MPO. Cross-referencing this metric to the MPO Ratio metric in Fig. 4, note that there is a correlation between high MPO usage and low congestion.

There appears to be some erratic data at the lower traffic loads of the L3 topology, this is caused by the price function operating in the less responsive lower parts of the x^2 price curve coupled with the light load since we used the vBNS traffic on the L3 topology.

C. *Overheads*

To show MPO overheads we present three metrics: Resource Price, Control Overhead and Setup Time.

Resource Price is defined as the increase of the price of deployments (i.e. $MPO\ price - non\ MPO\ price$) to the source ISP while varying traffic load. Each traffic requirement was uniquely numbered during the simulation and only requirements that were deployed in both the MPO and non-MPO configurations were selected for this comparison. Note that this metric is highly dependent on the chosen price function.

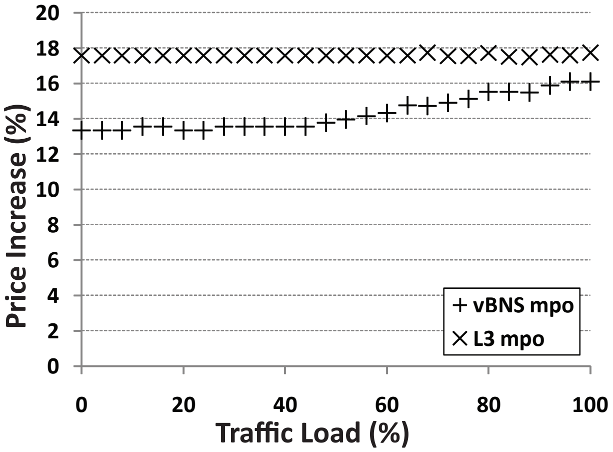

Fig. 7: Resource Price

Fig. 7 shows the Resource Price metric for both the vBNS and L3 topologies with prices increased an average of about 15% and 18% for the vBNS and L3 topologies respectively. The price increase reflects MPO's behavior in choosing more expensive paths to create redundancy. The vBNS price increases with traffic load. This reflects the prices of resources tracking scarcity.

Control Overhead is defined as the increase of control packets used in deployments (i.e. $MPO\ control - non\ MPO\ control$) while varying traffic load. Fig. 8 shows the Control Overhead metric for both vBNS and L3 topologies

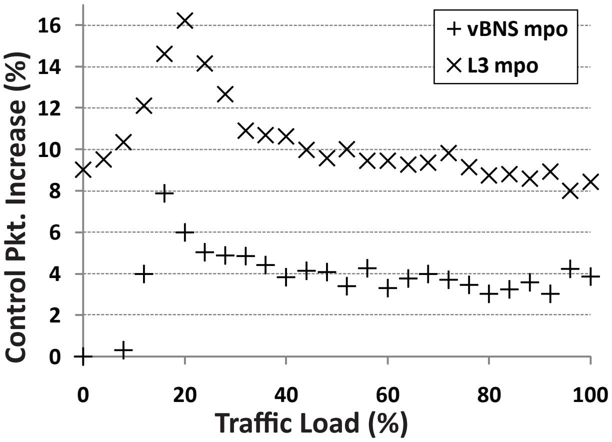

Fig. 8: Control Overhead

with increased control packets that closely track the MPO Ratio metric in Fig. 4. While a peak for the L3 topology of 16% may seem a little large, this may be offset by a 6% decrease in congestion shown by the Congestion metric in Fig. 6 at the same peak (i.e. at 20% traffic load). To add another numerical perpective we quote our work in [2] where we showed that the average control packet overhead for an iREX deployment was 5 and 8 for the vBNS and L3 topologies respectively; therefore, an increase of 16% would mean about 10 control packets per MPO deployment of a 4.8mb/sec bundle of traffic in the L3 topology.

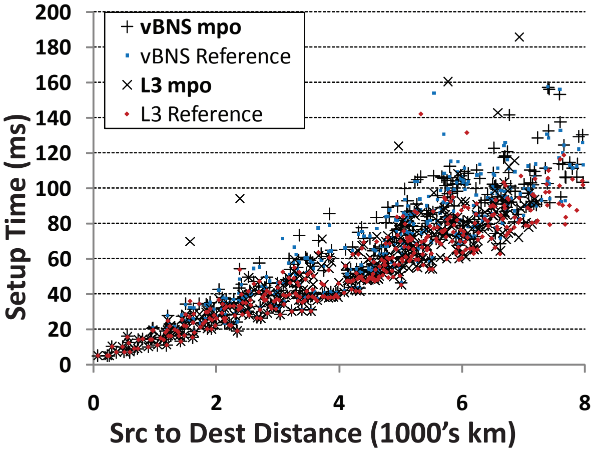

Fig. 9: Setup Time

Setup Time is defined as the time from the initiation of a reservation by the source ISP up to the point reservation is deployed, including any and all MPO negotiations. Fig. 9 shows the Setup Time metric for both the vBNS and L3 topology. In this figure we show the MPO values and also a non-MPO “Reference”. The rather small data point fonts on the figure is intentional to show overall grouping among the MPO and Reference values. We observe that MPO use has little effect on Setup Time.

There are some distinct points where the L3 setup times are about 20 milliseconds more that the grouping, this is because the increased number of links within the L3 topology allows for longer MPO paths with longer propagation times that take longer to deploy. We note that the worse case setup time is still under 200 milliseconds.

VI. RELATED WORK

Applying economics and the concept of pricing within networking has been studied in [8]–[11] but work in ID policy within an economic environment has been sparse. Fankhauser et al. [12] proposed an economics based SLA trading system, Koistenen et al. [13] proposed a protocol for peers to negotiate prices Wang et al. proposed RNAP [14] and Henkmann et al. proposed TDP [15], but in these systems, policy deployment is done bilaterally among neighboring peers whereas in iREX, the source domain deploys policy bilaterally with all domains involved in the deployment. Ambient Networks [16]–[18] allows a user to access multiple cooperative composite service providers, but iREX competitively filters resources and presents the “best” resources to a user.

Multi-path QoS routing has been studied in [19]–[24] and more, but work on management plane (i.e. policy) solutions assuming ISPs to be competitive has been sparse. Multipath choice in ID paths limited to the first hop has been commercially offered by [25] and [26], but iREX with MPO would offer more than just a choice of the first hop ISP. Bandwidth switching exchanges like Tradingcom Europe [27] are centralized services that operate similar to stock exchanges where ISPs trade excess capacity – iREX is a fully distributed architecture that can be used for similar purposes, but without the use of any centralized entity.

Preliminary work on iREX was previously published in [1], work defining the initial iREX architecture was previously published in [2], and work exploring the efficiency of iREX was previously published in [3].

VII. CONCLUSIONS & FUTURE WORK

We have presented a multi-path option within the iREX architecture as a distributed way to increase redundancy in ID QoS policy deployments and decrease ID QoS policy affected by resource failure. MPO takes an initiating ISP's preference for redundancy and provides information about the available options to achieve this preference – at a price the ISP is willing to pay.

Our numerical results conclude that 1) MPO use increases redundancy and decreases network congestion, 2) MPO deployments require 10 or less control packets and deploy in about the same time as non MPO deployments (i.e. less than 200ms), and 3) MPO use is more effective with higher connectivity and lower traffic loads.

An improvement to MPO that we would like to work on is for there to be a “joining” of each policy split when they meet on the way to the destination. This improvement would minimize ISP state overhead but requires a little more study on the lower layer implementation of IREX, which we have purposely left open at this point.

The probability of MPO successfully being used in a deployment is dependent on there being available paths at intermediate domains with prices within the ISP's price tolerance. This dependency is related to the price functions and MPO price tolerances being used by individual ISPs, and the interaction between them. In the future, we would like to study this interaction and its relation to iREX.

REFERENCES

[1] A. Yahaya and T. Suda, "iREX: Inter-domain QoS Automation using Economics," in *Proceedings of IEEE CCNC*, Las Vegas, Nevada, USA, January 2006.

[2] ——, "iREX: Inter-domain Resource Exchange Architecture," in *Proceedings of IEEE INFOCOM*, Barcelona, Spain, April 2006.

[3] A. Yahaya, T. Harks, and T. Suda, "iREX: Efficient inter-domain QoS policy architecture," in *Proceedings of IEEE Globecom*, San Francisco, California, USA, November 2006.

[4] W. Fang and L. Peterson, "Inter-as traffic patterns and their implications," *IEEE Global Internet Symposium*, vol. 5, pp. 68–75, December 1999.

[5] K. Abbink and J. Brandts, "Price competition under cost uncertainty: A laboratory analysis," *UFAE and IAE Working Papers 550.02*, 2002.

[6] Y. Rekhter and T. Li, "A border gateway protocol 4 (BGP-4)," IETF, RFC 1771, March 1995.

[7] iREX WEB, "http://netresearch.ics.uci.edu/irex/."

[8] J. MacKie-Mason and H. Varian, "Some economics of the internet," *10th Michigan Public Utility Conference*, March 1993.

[9] ——, "Pricing the internet," in Public Access to the Internet, Prentice-Hall, 1995.

[10] X. Yang, "Nira: A new internet routing architecture," in *ACM SIGCOMM FDNA 2003 Workshop*, Karlsruhe, Germany, August 2003.

[11] S. Shenker, D. Clark, D. Estrin, and S. Herzog, "Pricing in computer networks: reshaping the research agenda," *ACM Computer Communication Review*, vol. 26, no. 2, pp. 19–43, April 1996.

[12] G. Fankhauser, D. Schweikert, and B. Plattner., "Service level agreement trading for the differentiated services architecture," *Swiss Federal Institute of Technology, Computer Engineering and Networks Lab, Technical Report No. 59*, November 1999.

[13] J. Koistinen and A. Seetharaman, "Worth-based multi-category quality-of-service negotiation in distributed object infrastructures," *Hewlett Packard Software Technology Laboratory Technical Report HPL-98-51 (R. 1)*, July 1998.

[14] X. Wang and H. Schulzrinne, "RNAP: a resource negotiation and pricing protocol," in *Proceedings of NOSSDAV '99*, Basking Ridge, NJ, USA, June 1999.

[15] O. Heckmann, V. Darlagiannis, M. Karsten, R. Steinmetz, and B. Briscoe, "Tariff distribution protocol (TDP)," *Internet-Draft draft-heckmann-tdp-00.txt*, March 2002.

[16] C. P. Botham, A. L. Burness, P. L. Eardley, A. Eriksson, M. Jimenez-Abelleira, L. Loyola, and J. Rajahalme, "Inter-network routing in ambient networks," in *Proceedings of Mobile and Wireless Communications Summit, 2007. 16th IST*, Budapest, Hungary, July.

[17] O. Rietkerk and G. Huitema, "Business roles enabled by ambient networking to provide access for anyone to any network and service," in *Proceedings of Helsinki Mobility Roundtable, 2006*, Helsinki, Finland, June.

[18] L. Ho, J. Markendahl, and M. Berg, "Business aspects of advertising and discovery concepts in ambient networks," in *Proceedings of 2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications*, Helsinki, Finland, Sept.

[19] S. Plotkin, "Competitive routing of virtual circuits in atm networks," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1128–1136, August 1995.

[20] N. S. V. Rao and S. G. Batsell, "Qos routing via multiple paths using bandwidth reservation," in *Proceedings of IEEE INFOCOM*, San Francisco, USA, March 1998.

[21] I. Cidon, R. Rom, and Y. Shavitt, "Multi-path routing combined with resource reservation," in *Proceedings of IEEE INFOCOM*, Tel-Aviv, Israel, March 1997.

[22] Y. Jia, Y. Nikolaidis, and P. Gburzynski, "Multiple path qos routing," in *Proceedings of IEEE ICC*, Helsinki, Finland, June 2001.

[23] J. Shen, J. Shi, and J. Crowcroft, "Proactive multi-path routing," in *From QoS Provisioning to QoS Charging*, ser. Lecture Notes in Computer Science, G. Goos, J. Hartmanis, and J. van Leeuwen, Eds. Berlin Heidelberg: Springer, 2002, vol. 2511/2008, pp. 145–156.

[24] W. Xu and J. Rexford, "Miro: Multi-path interdomain routing," in *Proceedings of ACM SIGCOMM*, Pisa, Italy, September 2006.

[25] INTERNAP, "http://www.internap.com."

[26] AVAYA, "http://www.avaya.com."

[27] Tradingcom Europe, "http://www.tradingcomeurope.com."