Utility Maximization for Resolving Throughput/Reliability Trade-offs in an Unreliable Network with Multipath Routing

Vladimir Marbukh
National Institute of Standards and Technology
100 Bureau Drive, Stop 8920, Gaithersburg,
MD 20899-8920, USA
Tel: 1 301 975 2235
E-mail: marbukh@nist.gov

ABSTRACT

This paper proposes a framework for balancing competing user (i.e., application) level requirements by resolving the corresponding trade-offs in a distributed system with limited resources. Assuming that each user's preferences can be characterized by some utility function, the goal of balancing competing requirements for each user as well as across different users is to maximize the aggregate utility. The framework assumes a presence of Intelligent Plane, which isolates users from details of the network properties and mechanisms of implementation of the user level requirements. The Intelligent Plane performs the following tasks: (a) maps the user level requirements into the network resource requirements, (b) maps the resource congestion prices into prices of the user level requirements, and (c) maps the user willingness to pay for the user level requirements into payments for the specific sets of resources. Once payments for the specific sets of resources are identified, the resources are allocated to the users by a "TCP-friendly" algorithm. The paper discusses this framework for a particular case of balancing user requirements for throughput and survivability in an unreliable network, where survivability is achieved through redundancy, e.g., using multipath routing.

Keywords

Distributed system, resource allocation, elastic user, multipath routing, throughput, reliability trade-offs, pricing, intelligent plane.

1. INTRODUCTION

Since network resources are shared by multiple users (i.e., applications) and performance of each user is typically characterized by multiple competing criteria, network management includes the following two major tasks: (a) making the best use of the allocated resources for each user by resolving the trade-offs among competing user criteria, and (b) sharing resources among different users. Framing the goal of network management as the aggregate utility maximization subject to the capacity constraints, where the aggregate utility is the sum of the individual user utilities, has been proposed in [1]. This framework is based on the concept of elastic users, capable of adjusting their behavior in response to congestion pricing signals. Papers [2]-[3] have developed a distributed scheme for aggregate utility maximization in a case when user utilities are expressed in terms of the link bandwidths. This scheme interprets Lagrange multipliers associated with capacity constraints as congestion costs of the corresponding resources. These costs are communicated to the elastic users, who adjust their resource requirements or willingness to pay for the resources by maximizing the individual net utilities. Figure 1 illustrates this scheme.

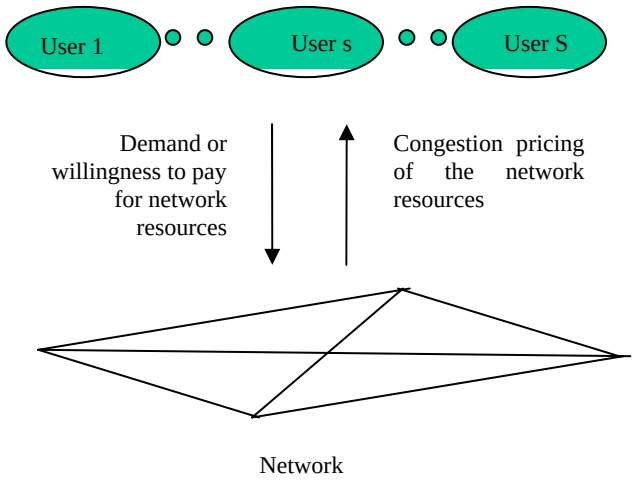

Fig. 1. Users directly responding to resource pricing

However, assumption [2]-[3] that user utilities are expressed in terms of the network resources may be too restrictive. Typically, users more naturally can express their preferences in terms of the user level requirements, such as rates and Quality of Service (QoS) parameters, rather than network level parameters, such as required bandwidth. Mapping user level requirements into network level resource requirements as well as mapping congestion resource pricing signals into pricing of the user level requirements depend on the specific network properties as well as specific implementation of the user level requirements. In the Internet with a dumb core and intelligent applications concentrated at the network edges this mapping can be performed by intelligent applications through probing.

Several recent proposals, starting with [4], argued in favor of relieving users from the burden of such probing by moving some intelligence to a separate "Intelligent Plane" (IntPlane). The IntPlane sits between the users and the network and hides the details of the network properties and user level requirements implementation mechanisms from the users. The advantages of such enhanced architecture include user convenience, possibility of optimization of the resource allocation and security considerations [4]. This paper proposes the functionality for the IntPlane as a mapping mechanism, which is shown on Figure 2.

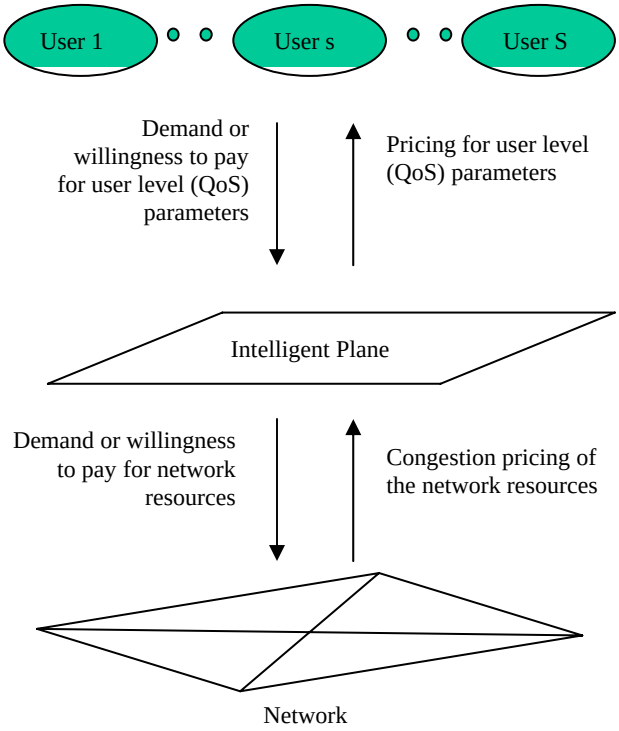

Fig. 2. Intelligent Plane as a mapping mechanism

Each elastic user attempting to maximize its individual net utility informs the IntPlane on its relative marginal utilities and "willingness to pay" for the network resource. The IntPlane performs the following tasks: (a) given the amount of the network resources allocated to each user, the IntPlane optimizes the balance among competing user level requirements for each user, (b) maps user willingness to pay into payments for specific sets of the network resources, and (c) communicates to the user the aggregate congestion cost of the resources allocated to the user. Once the willingness to pay for the specific sets of resources is identified, the resources are allocated to users by a TCP-type algorithm. The "payments" may either represent real funds, or be simply a parameter in the TCP-type protocol [5]. To ensure capability of this scheme to operate in a competitive (non-cooperative) environment, the resource allocation should be proportionally fair, meaning that resources are allocated to the users proportionally to the payments [2]-[3]. Proportional fairness ensures that both schemes, based on the direct user payments for the resources and user payments for the QoS, result in the same resource allocation and user payments [6].

This paper discusses possible implementation and benefits of the proposed enhanced architecture in a case of providing reliable services in an unreliable network. The reliability is achieved through redundancy by reserving extra bandwidth to protect against link capacity variability due to fading and mobility, and using multipath routing to protect against link failures. The packet level implementation of the redundancy scheme can be based on the route diversity coding [7]. Benefits of multipath routing for load balancing and protection against network element failures have been known for a long time [8]. However, research on load balancing, protection and restoration for wire-line and wireless networks has been mostly concentrated on evaluation of various performance and survivability metrics of certain multipath routing schemes [6]. While providing quantification of improving survivability with increase in redundancy through consuming more network resources, this research leaves aside the problem of balancing survivability and economic efficiency for each user as well as across different users. Conventional practical solutions, which offer users a limited set of choices with respect to survivability, attempt to resolve these trade-offs within a centralized framework by assigning the corresponding service classes. A price based market framework shifts choices regarding requested services, including survivability levels, to the users, assuming that users are aware of the available services and their prices [10]-[11].

The paper is organized as follows. Section II describes a model of the unreliable network and implementation of the reliable throughput. Section III introduces user utility of obtaining certain QoS and formulates the corresponding aggregate utility maximization framework. Section IV briefly extends decentralized aggregate utility maximization framework [2]-[3] to a case when each user is aware of mapping its QoS requirements into the requirements for the network resources. The decentralization is based on congestion pricing of the resources and elastic users responding to these pricing signals by maximizing their individual net utilities expressed in terms of the requested network resources. Section V develops a decentralized aggregate utility maximization framework assuming that users are unaware of the network properties and implementation of the user level requirements. The decentralization is based on proportionally fair pricing of the user level requirements and elastic users responding to these pricing signals by maximizing their individual net utilities expressed in terms of the user level parameters. The mapping between user

and network level parameters is done by the IntPlane. Section VI considers some examples and discusses the implication. Finally, conclusion briefly summarizes the proposed framework and proposes directions for future research.
.

2. MODEL

Subsection A defines two user $s \in S$ QoS parameters: the reliable throughput μ_s and the corresponding reliability exponent γ_s. Subsection B introduces a "fair" bandwidth sharing with controlled portions of link bandwidths allocated to different users. This bandwidth sharing allows for implementation of the reliable throughput by creating a "safety margin" for the fluctuating instantaneous user throughput. Subsection C describes an approximation for the reliability exponent used in the remainder of the paper.

2.1 User level parameters

Consider a network with link capacities c_l being subject to variability due to fading, mobility, node and link reliability, etc. Each network user $s \in S$ is uniquely identified by its origin-destination and user level Quality of Service (QoS) requirements. Presence of several users with the same origin-destination models different types of applications with the same origin-destination, e.g., voice and video. We assume that link capacity fluctuations occur on such fast timescale that they cannot be completely absorbed by the network management actions. Due to these fluctuations, link capacities c_l are in effect random variables and thus it may be difficult or even impossible to guarantee a fixed bandwidth (throughput) to a user. Instead it may be more natural to view the instantaneous aggregate throughput x_s for a user $s \in S$ as a random variable. Due to possible large fluctuations in the instantaneous aggregate throughput x_s users may prefer to characterize their requirements in terms of the pair (μ_s, γ_s) of the "reliable" aggregate throughput μ_s and the corresponding reliability exponent γ_s quantifying the confidence level that the instantaneous throughput x_s does not deteriorate below μ_s, where

$$\gamma_s = -\log\left(\frac{P\{x_s \le \mu_s\}}{P\{x_s \le \tilde{x}_s\}}\right) \quad (1)$$

and the average aggregate bandwidth reserved for user s is $\tilde{x}_s$. Figure 3 illustrates that creating a "safety margin" $\Delta_s = \tilde{x}_s - \mu_s$ increases confidence that the instantaneous throughput x_s would not deteriorate below μ_s.

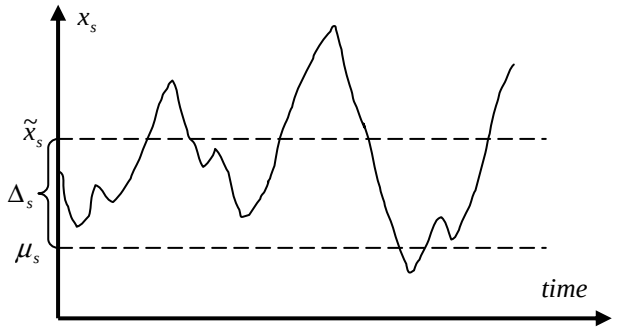

Fig. 3. Reliable aggregate throughput

Note that besides reserved average aggregate throughput $\tilde{x}_s$ and reliable throughput μ_s, reliability exponent γ_s also depends on: (a) probability distribution of the random link capacities c_l, (b) mechanism for sharing of the instantaneous link bandwidth among different users, (c) implementation of the reliable throughput μ_s, given resources allocated to user s, and (d) bandwidths reserved on specific routes. This paper assumes that random link capacities c_l are jointly statistically independent for all links $l \in L$. Assumptions (b)-(d) are described in the next two subsections.

This paper assumes that each user s instantaneous aggregate throughput x_s can be approximated by a normally distributed random variable with average $\tilde{x}_s$ and standard deviation σ_s, and thus reliability exponent (1) is

$$\gamma_s = -\log \Phi\left(\frac{\tilde{x}_s - \mu_s}{\sigma_s}\right) \quad (2)$$

where

$$\Phi(\xi) = \sqrt{\frac{2}{\pi}} \int_{-\infty}^{\xi} \exp\left(-\eta^2/2\right) d\eta \quad (3)$$

Note that approximation (2)-(3) neglects small probability event that the bandwidth is negative. In a case of high reliability requirements: $\gamma_s \to \infty$, reliability exponent (2) can be asymptotically approximated as follows:

$$\gamma_s \sim \frac{1}{2\sigma_s^2}(\tilde{x}_s - \mu_s)^2 \quad (4)$$

2.2 Reliable Throughput

We assume that each user s is allocated a certain controlled portion ϕ_{ls} of the link l bandwidth $\tilde{c}_l$, or equivalently, the average bandwidth $\tilde{x}_{ls} = \tilde{c}_l \phi_{ls}$, where average capacity of a link l is $\tilde{c}_l = E[c_l]$, and $\phi_{l\Sigma} \overset{def}{=} \sum_{s \in S} \phi_{ls} \le 1$. The instantaneous bandwidth allocated to a user s on a link l is a random variable $x_{ls} = c_l \phi_{ls} = (c_l / \tilde{c}_l)\tilde{x}_{ls}$. In a case of small variability in the link capacities it is convenient to introduce "small" random variables $\xi_l = 1 - c_l / \tilde{c}_l$ with zero averages $E[\xi_l] = 0$, so that the instantaneous bandwidth allocated to a user s on a link l is

$$x_{ls} = (1 - \xi_l)\tilde{x}_{ls} \tag{5}$$

In a particular case of a link failure model, when operational link l has capacity $c_l = \hat{c}_l$ and failed link has capacity $c_l = 0$ it is convenient to introduce binary random variables $\delta_l = 0$ if link l is operational and $\delta_l = 1$ otherwise, so that the instantaneous link l bandwidth is $c_l = (1 - \delta_l)\hat{c}_l$, and $\xi_l = \delta_l - \bar{\delta}_l$, where $\bar{\delta}_l = E[\delta_l]$. In this particular case the instantaneous bandwidth (5) is $x_{ls} = (1 - \delta_l)\hat{x}_{ls}$, where $\hat{x}_{ls} = \hat{c}_l \phi_{ls}$.

The rest of this subsection discusses implementation of the reliable throughput μ_s, given the instantaneous link bandwidths x_{sl} allocated to user s. Given vector $X_s = (x_{sl}, l \in L)$, the maximum achievable user s instantaneous aggregate throughput is $x_s = \sum_{l \in M_s^*(X_s)} x_{sl}$, where $M_s^*(X_s)$ is the corresponding min-cut. This paper assumes a suboptimal implementation of the reliable throughput, based on the route diversity coding [4] and shown on Figure 4.

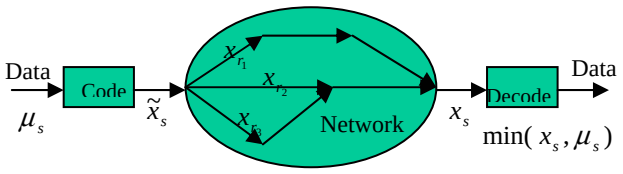

Fig. 4. Route diversity coding

In this implementation, after adding redundant bits and coding, user $s \in S$ data stream of rate μ_s is transformed into stream of higher rate $\tilde{x}_s \ge \mu_s$. This resulting stream is split into flows $\tilde{x}_{sr}$ over feasible routes $r \in R_s$ with the same origin-destination:

$$\tilde{x}_s = \sum_{r \in R_s} \tilde{x}_{sr} . \tag{6}$$

User s instantaneous throughput, i.e., rate of the user stream received at the destination, is

$$x_s = \sum_{r \in R_s} x_{sr} , \tag{7}$$

where the instantaneous throughput over route $r \in R_s$ is

$$x_{sr} = (1 - \xi_r)\tilde{x}_{sr} \tag{8}$$

and the normalized variability of a route r capacity is characterized by random variable

$$\xi_r = 1 - \prod_{l \in r} (1 - \xi_l) . \tag{9}$$

The reliability exponent (1) quantifies the possibility of reconstructing user s data stream at the destination [4]. Note that formula (9) is based on the assumption that link capacities fluctuate at much faster time scale than time needed for a packet to reach its destination. In the opposite extreme case the normalized variability of a route r capacity is characterized by random variable $\xi_r = 1 - \max(1 - \xi_l),\ l \in r$. Our analysis can be easily carried out for this case also.

Calculation of the reliability exponent (1) is comparatively simple in a case when routes $r \in R_s$ do not have overlapping links. In this case the aggregate instantaneous throughput (7) is a sum of jointly statistically independent random variables since ξ_r are jointly statistically independent random variables for $r \in R_s$. When routes $r \in R_s$ do have overlapping links, calculation of the reliability exponent (1) is generally a difficult problem [9].

2.3 Approximation for Reliability Exponent

We approximate the reliability exponent (1) by the leading term in the asymptotic expansion (4):

$$\gamma_s = \frac{1}{2\sigma_s^2} (\tilde{x}_s - \mu_s)^2 , \tag{10}$$

where user s average aggregate throughput is

$$\tilde{x}_s = \sum_{r \in R_s} \tilde{x}_{sr} \tag{11}$$

the variance of the aggregate throughput is

$$\sigma_s^2 = \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \tilde{x}_{sr_1} \tilde{x}_{sr_2} \quad (12)$$

the "normalized correlation" between route $r_1, r_2 \in R_s$ capacities is characterized by

$$\theta^2_{r_1 r_2} = \sum_{l \in r_1 \bigcap r_2} \theta_l^2 \quad (13)$$

and the normalized variance of the link l capacity is $\theta_l^2 = E\left[\xi_l^2\right]$. Note that matrix $\Theta_s = \left(\theta^2_{r_1 r_2}\right)_{r_1, r_2 \in R_s}$ is symmetric and positive: $\theta^2_{r_1 r_2} = \theta^2_{r_2 r_1}$ and $0 \le \theta^2_{r_1 r_2} \le \min\left(\theta^2_{r_1}, \theta^2_{r_2}\right), \ \forall r_1, r_2$. Also note that expression (10) can be obtained from a Gaussian link model in a large deviation regime of high reliability [12]. For a particular model of link failures the normalized variance of the link l capacity is $\theta_l^2 = \bar{\delta}_l\left(1 - \bar{\delta}_l\right)$, where the probability of link l failure is $\bar{\delta}_l \in [0,1]$.

Reliability exponent (10) can be expressed as follows:

$$\gamma_s = \frac{1}{2}\left(1 - \frac{1}{\omega_s}\right)^2 \left(\sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} \right)^{-1} \quad (14)$$

in terms of the user s redundancy factor, i.e., the number of bits transmitted per a bit of the "payload" [7],

$$\omega_s \overset{def}{=} \frac{1}{\mu_s} \sum_{r \in R_s} \tilde{x}_{sr} \quad (15)$$

and portions of load routed on feasible paths $r \in R_s$ are

$$\alpha_{sr} \overset{def}{=} (\omega_s \mu_s)^{-1} \tilde{x}_{sr} \quad (16)$$

where

$$\omega_s \ge 1 \quad (17)$$

$$\sum_{r \in R_s} \alpha_{sr} = 1. \quad (18)$$

Given load allocation vector $\alpha_s = (\alpha_{sr}, r \in R_s)$, the upper limit on reliability exponent (14), achieved as $\omega_s \to \infty$, is

$$\hat{\gamma}_s = \frac{1}{2}\left(\sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} \right)^{-1} \quad (19)$$

Given upper limit (19), the minimum redundancy (15) required to achieve reliability exponent γ for user s is

$$\omega = \left(1 - \sqrt{\gamma / \hat{\gamma}_s}\right)^{-1} \quad (20)$$

If routes $r \in R_s$ do not have overlapping links, formula (12) takes the following form

$$\sigma_s^2 = \sum_{r \in R_s} \left(\theta_r \tilde{x}_{sr}\right)^2 \quad (21)$$

and thus, formula (14) simplifies as follows:

$$\gamma_s = \frac{1}{2}\left(1 - \frac{1}{\omega_s}\right)^2 \left(\sum_{r \in R_s} \theta_r^2 \alpha_{sr}^2 \right)^{-1} \quad (22)$$

where we simplified notations as follows: $\theta_r^2 = \theta_{rr}^2$.

Given redundancy factor ω_s, one may attempt to maximize the reliability exponent (14):

$$\gamma_s^* = \max_{\alpha_{sr} \ge 0} \gamma_s \quad (23)$$

subject to constraints (18).

Theorem 1. Given redundancy factor ω_s and network properties represented by matrix Θ_s, solution to optimization problem (22)-(23), (18) is

$$\gamma_s^* = \frac{1}{2}\left(1 - \frac{1}{\omega_s}\right)^2 \left(\sum_{r_1, r_2 \in R_s} t^2_{r_1 r_2} \right)^{-1} \quad (24)$$

and is achieved for load allocation

$$\alpha_{sr}^* = t^2_{r_1 r_2} \Big/ \sum_{r_1', r_2' \in R_s} t^2_{r_1' r_2'} \quad (25)$$

where symmetric and positive matrix $T_s = \left(t^2_{r_1 r_2}\right)_{r_1, r_2 \in R_s}$ is the inverse to Θ_s: $T_s = \Theta_s^{-1}$.

Proof. The optimal load allocation is determined by solution to the following optimization problem:

$$\min_{\alpha_s} \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} \quad (26)$$

subject to constraints (18). The Lagrangian for (25), (18) is

$$L = \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} + \lambda\left(1 - \sum_{r \in R_s} \alpha_{sr}\right)$$

and, due to convexity, the corresponding necessary and sufficient Kuhn-Tucker conditions form the following linear system [13]:

$$\partial L / \alpha_{sr} = \sum_{r' \in R_s} \theta^2_{rr'} \alpha_{sr'} - \lambda = 0 \quad (27)$$

where Lagrange multiplier λ is determined by (18). This ends the proof.

The following statements directly follow from Theorem 1.

Corollary 1. Given the network properties represented by matrix Θ_s, the upper limit on the reliability exponent (10), achieved as redundancy factor $\omega_s \to \infty$, is

$$\hat{\gamma}_s^* = \frac{1}{2}\left(\sum_{r_1, r_2 \in R_s} t_{r_1 r_2}^2 \right)^{-1} \tag{28}$$

Corollary 2. If routes $r \in R_s$ do not have overlapping links, the maximal reliability exponents (24) is

$$\gamma_s^* = \frac{1}{2}\left(1 - \frac{1}{\omega_s}\right)^2 \left(\sum_{r \in R_s} \theta_r^{-2} \right)^{-1} \tag{29}$$

the optimal load allocation (25) is

$$\alpha_{sr}^* = \theta_r^{-2} \Big/ \sum_{r' \in R_s} \theta_{r'}^{-2} \tag{30}$$

and the upper limit (28) is

$$\hat{\gamma}_s^* = \frac{1}{2}\left(\sum_{r \in R_s} \theta_r^{-2} \right)^{-1} \tag{31}$$

3. GOAL OF NETWORK MANAGEMENT

Subsection A introduces individual user utility of obtaining service parameters (μ_s, γ_s). Subsection B formulates the aggregate utility maximization framework [1] for a particular case of balancing competing requirements for reliable throughput and the corresponding reliability for each user as well as across different users.

3.1 User Utilities

Let $h_s(x, \mu)$ be a function, monotonously increasing in both arguments $0 \le \mu \le x < \infty$. Consider elastic user s whose satisfaction of obtaining service with parameters (μ, γ) is characterized by a utility function

$$U_s(\mu, \gamma) = u_s(\mu) v_s(\gamma), \tag{32}$$

where function $u_s(\mu)$ is a conditional average over the aggregate rate x_s:

$$u_s(\mu) = E_{x_s}\left[h(x_s, \mu) \middle| x_s > \mu \right], \tag{33}$$

and function $v_s(\gamma)$ is monotonously increasing for $0 \le \gamma < \infty$. Note that under large deviation regime of high reliability, conditional average in (32) can be approximated by the corresponding unconditional average. Figures 5 and 6 sketch typical utility functions $u_s(\mu)$ and $v_s(\gamma)$ respectively.

Definition (32)-(33) is quite flexible, covering a wide range of possibilities. Consider some particular cases. User s having "hard" requirements on the reliability parameter $\gamma_s \ge \gamma_s^{\min}$ is characterized by utility function (32)-(33), where

$$v_s(\gamma) = \chi(\gamma - \gamma_s^{\min}), \tag{34}$$

and step-wise function is $\chi(\gamma) = 1$ if $\gamma > 0$, and $\chi(\gamma) = 0$ if $\gamma \le 0$.

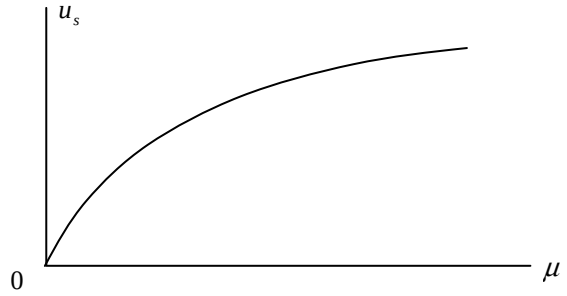

Fig. 5. Typical user utility of the reliable throughput

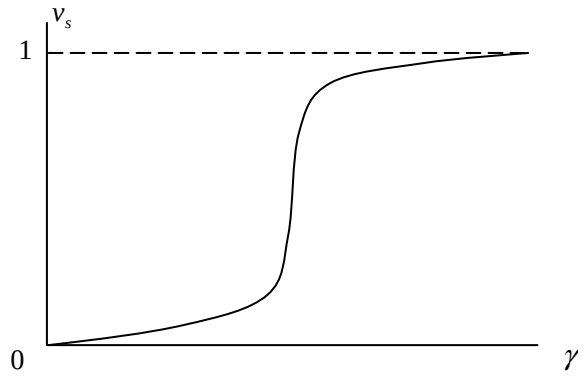

Fig. 6. Typical user utility of the reliability exponent

User s, elastic with respect to the reliable throughput μ, is characterized by utility (32)-(33), where function $h_s(x, \mu) \equiv u_s(\mu)$ does not depend on the actual random aggregate throughput $x \in [\mu, \infty)$ and depends only on the reliable aggregate throughput $\mu \in [0, \infty)$. A particular case (32)-(34) with $\gamma_s^{\min} = 0$ describes an elastic user concerned with the average throughput: $U_s = u_s\left(E[x_s]\right)$. A particular case of (32)-(34) with $\gamma_s^{\min} = 0$ and function $h_s(x, \mu) \equiv u_s(x)$ independent of the reliable throughput $\mu \in [0, \infty)$ describes an elastic user whose satisfaction is characterized by the average utility of the instantaneous aggregate throughput: $U_s = E[u_s(x_s)]$.

3.2 Aggregate Utility Maximization Problem

S. Shenker has proposed [1] aggregate utility maximization to be the objective of network management. In our particular case the aggregate utility maximization framework takes the following form:

$$\max \sum_s U_s(\mu_s, \gamma_s) \tag{35}$$

over user level requirements $(\mu, \gamma) = (\mu_s, \gamma_s, s \in S)$ and vector $\tilde{X} = (\tilde{x}_{sr} : s \in S, r \in R_s)$ subject to constraints (10), link capacity constraints

$$\tilde{y}_l \le \tilde{c}_l, \tag{36}$$

flow non-negativity constraints: $\tilde{x}_{sr} \ge 0$ and constraints on the reliable throughput $0 \le \mu_s \le \tilde{x}_s, \ s \in S$, where the link l load is

$$\tilde{y}_l = \sum_s \sum_{r: l \in r \subseteq R_s} \tilde{x}_{sr} \tag{37}$$

Optimization problem (35) is equivalent to the following optimization problem

$$\max_{\mu, \gamma, \tilde{X}} W \tag{38}$$

subject to the same constraints except (37), where the "social welfare" is

$$W = \sum_s U_s(\mu_s, \gamma_s) - \sum_l f_l(\tilde{y}_l) \tag{39}$$

and appropriately selected penalty functions $f_l(y)$ may quantify the congestion penalty in terms of delays or packet loss as link utilization approaches link capacities [3]. For packet networks it is often assumed [14]

$$f_l(y) = y/(\tilde{c}_l - y). \tag{40}$$

A particular case of optimization problem (38)-(39), when each user s specifies its service requirements (μ_s, γ_s) correspond to the following traffic engineering problem:

$$\min_{\tilde{x}_{sr} \ge 0} \sum_l f_l\left(\sum_s \sum_{r: l \in r \subseteq R_s} \tilde{x}_{sr} \right) \tag{41}$$

subject to constraints

$$\sum_{r \subseteq R_s} \tilde{x}_{sr} \ge \mu_s + \left(2\gamma \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \tilde{x}_{sr_1} \tilde{x}_{sr_2} \right)^{1/2}, \forall s \in S \tag{42}$$

4. USERS ALLOCATE BANDWIDTH

This section assumes that each $s \in S$ (a) is aware of the network properties quantified by matrix Θ_s, and (b) capable of finding the optimal balance (μ_s^*, γ_s^*) between competing requirements for the reliable throughput μ_s and the corresponding reliability exponent γ_s by maximizing the individual utility, given allocated bandwidths $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$:

$$\gamma_s^* = \arg\max_{\gamma \ge 0} \left\{ u_s(\tilde{x}_s - \sigma_s \sqrt{2\gamma}) v_s(\gamma) \right\} \tag{43}$$

$$\mu_s^* = \tilde{x}_s - \sigma_s \sqrt{2\gamma_s^*} \tag{44}$$

Once optimization (43)-(44) is performed and thus individual utilities with respect to the bandwidth

$$\tilde{U}_s(\tilde{X}_s) = u_s(\tilde{x}_s - \sigma_s \sqrt{2\gamma_s^*}) v_s(\gamma_s^*) \tag{45}$$

are identified, the aggregate utility maximization problem (38)-(39)becomes

$$\max_{\tilde{X}} \left\{ \sum_s \tilde{U}_s(\tilde{X}_s) - \sum_l f_l\left(\sum_s \sum_{r: l \in r \subseteq R_s} \tilde{x}_{sr} \right) \right\} \tag{46}$$

Note that under hard constraints on the reliability (34) the optimal operating point (43)-(44) is

$$(\mu_s^*, \gamma_s^*) = \left(\tilde{x}_{s\Sigma} - \sigma_s \sqrt{2\gamma_s^{\min}}, \ \gamma_s^{\min} \right) \tag{47}$$

and thus individual utility (45) is

$$\tilde{U}_s(\tilde{X}_s) = u_s(\tilde{x}_s - \sigma_s \sqrt{2\gamma_s^{\min}}) v_s(\gamma_s^{\min}) \tag{48}$$

This section describes distributed algorithms to aggregate utility maximization (46), assuming that user $s \in S$ utility function $\tilde{U}_s$ is known only to this user. The algorithms are the straightforward extension of algorithms proposed in [2]-[3] for a case $\tilde{U}_s = \tilde{U}_s\left(\sum_{r \in R_s} \tilde{x}_{sr} \right)$, and assume that elastic users respond to congestion price of the bandwidth. Subsection A describes algorithm based on users adjusting bandwidth requirements in response to bandwidth prices. Subsection B describes algorithms based on users adjusting their willingness to pay for bandwidth in response to rates charged for the bandwidth.

4.1 Users Adjusting Bandwidth Requests

Consider the following individual optimization problem for a user s attempting to maximize its individual net utility:

$$\max_{\tilde{x}_{sr} \ge 0} \max_{\gamma \ge 0} \left\{ u_s(\tilde{x}_s - \sigma_s \sqrt{2\gamma}) v_s(\gamma) - \sum_{r \in R_s} d_r \tilde{x}_{sr} \right\} \tag{49}$$

where the route r price is:

$$d_r = \sum_{l \in r} f_l'(\tilde{y}_l) \qquad (50)$$

the link l price $f_l'(\tilde{y}_l)$ is a derivative of the congestion penalty function for this link $f_l(\tilde{y}_l)$, and the link load $\tilde{y}_l$ is given by (37). Solving individual optimization problem (49)-(50) by each user $s \in S$ also maximizes the aggregate utility (46) if the link prices are "right", meaning that derivatives $f_l'(\tilde{y}_l)$ are calculated at the optimal link l load $\tilde{y}_l = \tilde{y}_l^{opt}$, $\forall l$.

Kuhn-Tucker necessary conditions for a vector $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$ to solve (49) are as follows [13]:

$$\sqrt{2\gamma}\sigma_s^{-1} \sum_{r' \in R_s} \theta_{rr'}^2 \tilde{x}_{sr'} = 1 - \frac{d_r}{u_s'} \ \ if \ d_r \le u_s' \qquad (51)$$

$$\tilde{x}_{sr} = 0 \ \ if \ d_r > u' \qquad (52)$$

where $u_s'(\mu) = du_s(\mu)/d\mu$ is the derivative of the user s utility at the point of this user reliable throughput $\mu = \mu_s$ and σ_s is given by (12). If user utilities $\tilde{U}_s(\tilde{X}_s)$ are concave, (51)-(52) are also the corresponding sufficient conditions [13]. In this case, user s optimal response to the pricing signals d_r is requesting bandwidth vector $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$, which solves system (51)-(52) and thus maximizes its individual net utility (49)-(50).

Generally, optima in (46) and (49) are achieved when some flows are zero: $\tilde{x}_{sr} = 0$ for some $r \in R_s$, $s \in S$. In fact, this situation is typical in presence of "high cost", e.g., highly congested or very "long" routes, when optimal solution is not to use these "expensive" routes. For example, conventional shortest path routing uses only one, "optimal" route. Given $\mu \ge 0$, define a subset of feasible routes participating in user $s \in S$ transmission:

$$R_s(\mu) = \{r : u_s'(\mu) > d_r, r \in R_s\} \qquad (53)$$

Consider two routes $r_1, r_2 \in R_s(\mu)$, which do not have overlapping link with each other or with any other route $\forall r \in R_s(\mu)$: $r_i \bigcap r = \varnothing$, $i = 1,2$. In this case we have from (51):

$$\frac{\tilde{x}_{sr_1}}{\tilde{x}_{sr_2}} = \left(\frac{\theta_{r_2}}{\theta_{r_1}}\right)^2 \frac{1 - d_{sr_1}/u_s'}{1 - d_{sr_2}/u_s'} \qquad (54)$$

It follows from (54) that if two routes $r_1, r_2 \in R_s(\mu)$ have the same cost: $d_{r_1} = d_{r_2}$, then the user transmission rate on these routes should be inversely proportional to the variances of the fluctuating bandwidths of the corresponding routes:

$$\tilde{x}_{sr_1} / \tilde{x}_{sr_2} = \left(\theta_{r_2} / \theta_{r_1}\right)^2 \qquad (55)$$

This conclusion that load allocation among several routes of the same cost should send more traffic on the better quality routes while preserving routing diversity is intuitively plausible.

In a case of hard reliability constraints (34) when feasible routes $r \in R_s$ do not have overlapping links, the optimal flow vector $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$ can be identified explicitly. Indeed, in this case Kuhn-Tucker equations (51) take the following form:

$$\tilde{x}_{sr} = \frac{\sigma_s}{\sqrt{2\gamma}\theta_r^2} \max\left(0,\ 1 - \frac{d_r}{u_s'}\right) \qquad (56)$$

Summarizing (56) over $r \in R_s$ we obtain:

$$\tilde{x}_s = \frac{\sigma_s}{\sqrt{2\gamma}} \sum_{r \in R_s} \frac{1}{\theta_r^2} \max\left(0,\ 1 - \frac{d_r}{u_s'}\right) \qquad (57)$$

Substituting (57) into (44) we obtain the following expression for σ_s:

$$\sigma_s = \frac{\mu}{\sqrt{2\gamma}} \Bigg/ \left[\sum_{r \in R_s^*(\mu)} \left(1 - \frac{d_r}{u_s'}\right) \frac{1}{\theta_r^2} - 2\gamma \right] \qquad (58)$$

Substituting σ_s into (56) we obtain the following expression for the flows $\tilde{x}_{sr}, r \in R_s(\mu)$:

$$\tilde{x}_{sr} = \frac{\mu}{\sum_{r' \in R_s(\mu)} \frac{1}{\theta_{r'}^2}\left(1 - \frac{d_{r'}}{u_s'}\right) - 2\gamma} \left(1 - \frac{d_r}{u_s'}\right) \frac{1}{\theta_r^2} \qquad (59)$$

Substituting (59) into right-hand side of the following necessary condition for optimality in (49)

$$u_s' = \frac{1}{\mu} \sum_{r \in R_s} \tilde{x}_{sr} \qquad (60)$$

we obtain a quadratic algebraic equation for the derivative u_s', yielding the reliable throughput $\mu = \mu_s$. After that, flows are determined by (59).

4.2 Uses Adjusting Willingness to Pay

Solving individual optimization problem (49) by each user results in a decentralized maximization of the aggregate utility assuming convexity and "right" link prices. Formula (50) can be used as a basis for finding the right prices by a distributed algorithm [2], when users declare their requirements for bandwidth $\tilde{X}_s = (\tilde{x}_{sr})$, then "the network" informs users on the route costs (50),

then users adjust their bandwidth requirements, etc. This subsection describes a distributed algorithm for finding the "right" prices, based on the user willingness-to-pay. This algorithm, being a straightforward extension of the corresponding algorithm [3], probably better fits into existing Internet architecture.

Consider a situation when, given bandwidth vector $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$, each user s determines its willingness to pay w_{sr} for bandwidth on each route $r \in R_s$ by maximizing its individual net utility:

$$\max_{w_{sr} \geq 0} \left\{ \tilde{U}_s \left(\sum_{r \in R_s} (w_{sr} / p_{sr}) \right) - \sum_{r \in R_s} w_{sr} \right\} \tag{61}$$

where p_{sr} is the rate charged by the network for a unit of bandwidth on route $r \in R_s$. After user s informs the network on the vector $(w_{sr}, r \in R_s)$, the network, running Transmission Control Protocol – Active Queue Management (*TCP-AQM*) protocol [5], adjusts bandwidth vector $\tilde{X}_s = (\tilde{x}_{sr}, r \in R_s)$ according to the following system of differential equations:

$$\dot{\tilde{x}}_{sr} = \tilde{k}\left(w_{sr} - \tilde{x}_{sr} d_r\right) \tag{62}$$

Assuming that each user s monitors its rates $\tilde{x}_r$ on routes $r \in R_s$ and instantaneously adjusts parameters w_{sr} by solving optimization problem (61) the user willingness-to-pay is

$$w_{sr} = \tilde{x}_{sr} \frac{\partial \tilde{U}_s(\tilde{X}_s)}{\partial \tilde{x}_{sr}} \tag{63}$$

Consider rate of change of the social welfare (39) with time:

$$\dot{W} = \sum_s \sum_{r \in R_s} \frac{\partial W}{\partial \tilde{x}_{sr}} \dot{\tilde{x}}_{sr} \tag{64}$$

Substituting (62)-(63) into right-hand side of (64) we obtain

$$\begin{aligned} \dot{W} &= \sum_s \sum_{r \in R_s} \left(\frac{\partial \tilde{U}_s}{\partial \tilde{x}_{sr}} - d_r \right) \left(\frac{w_{sr}}{\tilde{x}_{sr}} - d_r \right) \tilde{x}_{sr} = \\ &= \sum_s \sum_{r \in R_s} \left(\frac{\partial \tilde{U}_s}{\partial \tilde{x}_{sr}} - d_r \right)^2 \tilde{x}_{sr} \geq 0 \end{aligned} \tag{65}$$

Thus social welfare (39) is a Lyapunov function for the dynamic system (62)-(63). Note that since social welfare (39) may have multiple local maxima for streaming applications [1], inequality (65) only implies that the bandwidth adjustment process (62)-(63) converges to the local maximum of the social welfare (39).

5. QOS PRICING

This section proposes algorithms for aggregate utility maximization (35) assuming that users are unaware of the network layer parameters. These algorithms assume presence of the IntPlane, which isolates users from the network properties and QoS implementation mechanisms. Subsection A considers implementation and pricing of the service parameters (μ_s, γ_s) by the IntPlane, given price of the bandwidth $(d_r, r \in R_s)$. This setting may describe a case of "fat" links carrying traffic from a large number of users, so that the link costs can be considered fairly stable. Subsection B describes a cross-layer, distributed algorithm for aggregate utility maximization,. The algorithm is based on user willingness to pay for service parameters (μ_s, γ_s) and results in proportionally fair pricing. Subsection B also demonstrates that under certain, rather restrictive, conditions this algorithm maximizes the aggregate utility.

5.1 Users Requesting QoS

Consider a user $s \in S$ individual optimization problem

$$\max_{\mu, \gamma \geq 0} \left\{ U_s(\mu, \gamma) - \mu D_s(\gamma) \right\} \tag{66}$$

where the price of a unit of reliable throughput for user s is

$$D_s = \frac{\tilde{d}_s}{1 - \sqrt{\gamma / \hat{\gamma}_s}} \tag{67}$$

the price of a unit of the average throughput for user s is

$$\tilde{d}_s = \sum_{r \in R_s} d_r \alpha_{sr} \tag{68}$$

the upper limit on the reliability exponent $\hat{\gamma}_s$ is given by (19), cost of a route r is d_r and vector $\alpha_s = (\alpha_{sr}, r \in R_s)$ characterizes implementation of user s requirements.

Given implementation of all user level requirements $\alpha = (\alpha_s, s \in S)$, maximization individual net utility (66) by each user $s \in S$ also maximizes the aggregate utility:

$$\sum_s U_s(\mu_s, \gamma_s) - \sum_l f_l \left(\sum_s \frac{\mu_s}{1 - \sqrt{\gamma_s / \hat{\gamma}_s}} \sum_{r: l \in r \subset R_s} \alpha_{sr} \right) \tag{69}$$

over user level requirements $(\mu_s, \gamma_s : s \in S)$ if the route costs are

$$d_r = \sum_{l \in r} f_l \left(\sum_s \frac{\mu_s}{1 - \sqrt{\gamma_s / \hat{\gamma}_s}} \sum_{r: l \in r \subset R_s} \alpha_{sr} \right) \quad (70)$$

The problem of joint maximization of the aggregate utility (69) over user level parameters $(\mu_s, \gamma_s : s \in S)$ and implementation $(\alpha_{sr}, r \in R_s, s \in S)$ can be decomposed into (a) maximization of individual net utility (66) by each user $s \in S$, and (b) minimization of the cost of implementation of user $s \in S$ requirements by the IntPlane:

$$\tilde{D}_s^* = \min_{\alpha_r \geq 0} \frac{1}{1 - \sqrt{2\gamma \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2}}} \sum_{r \in R_s} d_r \alpha_{sr} \quad (71)$$

subject to constraints (18).

Cost minimization (71) subject to constraint (18) can be carried out as follows. Consider optimization problem:

$$\tilde{\theta} = \min_{\alpha_{sr} \geq 0} \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} \quad (72)$$

subject to constraints

$$\sum_{r \in R_s} d_r \alpha_{sr} \leq \tilde{d} \quad (73)$$

and constraints (18). Note that this optimization problem intends to maximize the bound on the reliability exponent (19) subject to upper constraint on the average route cost, or, equivalently, to minimize the average route cost subject to lower bound on the reliability exponent (19). The Largangian for this optimization problem is [13]:

$$L = \sum_{r_1, r_2 \in R_s} \theta^2_{r_1 r_2} \alpha_{sr_1} \alpha_{sr_2} - \\ - \lambda_1 \left(\tilde{d} - \sum_{r \in R_s} d_r \alpha_{sr} \right) + \lambda_2 \left(1 - \sum_{r \in R_s} \alpha_{sr} \right) \quad (74)$$

where the corresponding Largange multipliers are λ_1 and λ_2. Optimization problem (72)-(73), (18) is convex and thus, the necessary and sufficient conditions for a vector $\alpha_s = (\alpha_{sr}, r \in R_s)$ to be a solution to this optimization problem are as follows [13]:

$$\sum_{r' \in R_s} \theta^2_{rr'} \alpha_{sr'} = \lambda_2 - \lambda_1 d_r \ \ if \ \ \lambda_2 - \lambda_1 d_r > 0 \\ \alpha_{sr} = 0 \ \ if \ \ \lambda_2 - \lambda_1 d_r \leq 0 \quad (75)$$

where Largange multipliers are λ_1 and λ_2 are determined from (18) and (73).

It can be shown that $\lambda_1, \lambda_2 \geq 0$, and thus the structure of the solution to (72)-(73), (18) is as follows. Without loss of generality, assume that all $K = \dim R_s$ routes $r \in R_s$ are arranged in M mutually exclusive groups $G_m, m = 1,..,M$ so that all routes in the same group have the same cost $\tilde{d}_m$ and cost increases as the group number increases:

$$\underbrace{\overbrace{d_1 = .. = d_{K_1}}^{G_1}}_{=\tilde{d}_1} < \underbrace{\overbrace{d_{K_1+1} = .. = d_{K_2}}^{G_2}}_{=\tilde{d}_2} < ... < \underbrace{\overbrace{d_{K_{M-1}} = .. = d_K}^{G_M}}_{=\tilde{d}_M}$$

Routes within each group are numbered arbitrarily. To avoid trivialities we further in the paper assume that $\theta_r > 0, \ \forall r \in R_s$. Since solution to system (75) is

$$\alpha_i = \sum_{j=1}^{K_m} t^2_{mij} \max(0, \lambda_2 - \lambda_1 d_j), \quad (76)$$

where matrix $T_m = \left(t^2_{mij} \right)_{i,j=1}^{K_m}$ is inverse to the matrix $\Theta_m = \left(\theta^2_{ij} \right)_{i,j=1}^{K_m}$, at the optimum the load is spread over feasible routes from groups $G_i, i = 1,..,m$ and m is determined by conditions: $\tilde{d}_m \leq \lambda_2 / \lambda_1, \ \tilde{d}_{m+1} > \lambda_2 / \lambda_1$. Substituting (76) into conditions (73) and (18) results in explicit, though elaborate, expressions for Lagrange multipliers λ_1 and λ_2. Thus, the complexity of solving optimization problem (72)-(73), (18) lies in inverting matrices $\Theta_m, m = 1,..,M$.

Once solution $\alpha = \alpha(\tilde{d})$, $\tilde{\theta} = \tilde{\theta}(\tilde{d})$ to optimization problem (72)-(73), (18) is found, solution to optimization problem (71), (18) is $\alpha^{opt} = \alpha(\tilde{d}^{opt})$, where $\tilde{d}^{opt}$ and D_s^{opt}, solve the following optimization problem:

$$\tilde{D}_s^* = \min_{\tilde{d}} \frac{\tilde{d}}{1 - \sqrt{2\gamma \tilde{\theta}(\tilde{d})}} \quad (77)$$

subject to constraint

$$\tilde{d}_1 \leq \tilde{d} \leq \tilde{d}_M \quad (78)$$

It can be shown that (76)-(77) is a convex optimization problem, which can be solved by fixed points as follows:

$$\tilde{d} = \frac{\sqrt{2} - 2\gamma\theta}{\lambda_1 \gamma \theta}$$

where λ_1 is the Lagrange multiplier in (74). It is also possible to show existence of M constants

$$0 \overset{def}{=} \eta_0 < \eta_1 \leq \eta_2 \leq .. \leq \eta_M \leq \eta_{M+1} \overset{def}{=} 1 \quad (79)$$

such that if $\gamma/\hat{\gamma}_s^* \in [\eta_{m-1}, \eta_m)$ then $\alpha_i^* > 0,\ i = 1,..,K_m$ and $\alpha_i^* = 0,\ i = K_m + 1,..,K$. Once optimal load split is identified, the redundancy factor is given by (20) and the optimal reliable throughput is determined by solution to the corresponding individual optimization problem.

In a case of a user s concerned only with the average throughput: $\gamma_s \to 0$, solution to (71)-(18) sends entire traffic on minimum cost routes $r \in G_1$. If there are several minimum cost routes: $\dim G_1 \geq 2$, a situation of minimum equal cost multipath arises. The optimal load split among minimum cost routes $r \in G_1$ is

$$\alpha_k = \left(\sum_{i,j=1}^{K_1} t_{1ij}^2 \right)^{-1} \sum_{j=1}^{K_1} t_{1kj}^2 \quad if \quad k = 1,..,K_1 \qquad (80)$$

$$\alpha_k = 0 \quad otherwise$$

and the redundancy factor is $\omega = 1$. In another extreme case of very reliability sensitive user s: $\gamma \to \hat{\gamma}_s^* - 0$, the optimal load split among feasible routes $r \in R_s$ is given by (25), and redundancy factor is given by (20).

5.2 Users Willingness to Pay for QoS

We assume that user s is charged for service (μ, γ) a price proportional to the reliable throughput μ

$$P_s = p_s \mu \qquad (81)$$

where rate p_s is some increasing functions of the reliability exponent γ. Given service (μ, γ) and price structure (81), user s (a) determines and communicates to the IntPlane its willingness to pay for the service $w = w_s$, where

$$w_s = \arg\max_{w \geq 0} \{ U_s(w/p_s, \gamma) - w \} \qquad (82)$$

and (b) estimates and communicates to the IntPlane the relative importance of its competing requirements for the reliable throughput μ and reliability exponent γ quantified by its relative marginal utility

$$g_s(\mu, \gamma) = \frac{\partial U_s}{\partial \gamma} \Big/ \frac{\partial U_s}{\partial \mu} \qquad (83)$$

Based on this information and being aware of the network properties quantified by matrix Θ, the IntPlane performs the following tasks: (a) maximizes user s utility $U_s(\mu, \gamma)$, given bandwidth vector $\tilde{X}_s$, as follows:

$$\dot{\mu}_s = k \left\{ 1 + g_s(\mu_s, \gamma_s) \left[\partial \gamma_s / \partial \mu_s \right]_{\tilde{X}_s} \right\} \qquad (84)$$

where $k > 0$ is some constant, and (b) allocates portions

$$\pi_{sr} = \frac{\tilde{x}_{sr}}{\mu_s} \frac{\partial \mu_s}{\partial \tilde{x}_{sr}} \qquad (85)$$

of user s payment (82) to "pay" for the route $r \in R_s$ bandwidths, where the reliable throughput is

$$\mu_s = \tilde{x}_s - \sigma_s \sqrt{2\gamma}\,. \qquad (86)$$

Combining (85) with (86) we obtain

$$\pi_{sr} = \left(1 - \sqrt{2\gamma} \tilde{x}_{sr} \theta_r^2 \sigma_s^{-1} \right) \frac{\tilde{x}_{sr}}{\mu_s} \qquad (87)$$

It is easy to verify that

$$\frac{1}{\mu_s} \sum_{r \in R_s} \tilde{x}_{sr} \frac{\partial \mu_s}{\partial \tilde{x}_{sr}} \equiv 1 \qquad (88)$$

and thus the proposed payment scheme is proportionally fair [2]-[3]. Once payments $w_{sr} = w_s \pi_{sr}$ are identified, the flow vector $\tilde{X} = (\tilde{x}_{sr})$ is adjusted by a "TCP-type" load allocation algorithm (62).

Consider a particular situation, when (a) relaxation of the user level parameters (84) is much faster than relaxation of the allocated bandwidths (62), i.e., $k >> \tilde{k}$ and thus:

$$\left[\frac{\partial U_s}{\partial \mu_s} \right]_{\gamma_s} \left[\frac{\partial \mu_s}{\partial \gamma_s} \right]_{\tilde{X}_s} + \left[\frac{\partial U_s}{\partial \gamma_s} \right]_{\mu_s} \equiv 0\,, \qquad (89)$$

(b) each user s instantaneously adjusts and informs the IntPlane on its willingness to pay w_r (82):

$$w_s = \mu_s \frac{\partial U_s}{\partial \mu} \qquad (90)$$

(c) the IntPlane instantaneously allocates each user s payment (90) into payments $w_{sr} = w_s \pi_{sr}$ for the bandwidths on specific routes $r \in R_s$.

Consider rate of change of the social welfare (39) with time:

$$\dot{W} = \sum_s \left(\frac{\partial W}{\partial \mu_s} \dot{\mu}_s + \frac{\partial W}{\partial \gamma_s} \dot{\gamma}_s \right) \qquad (91)$$

Due to our assumptions

$$\dot{W} = \sum_s \sum_{r \in R_s} \left(\frac{\partial U_s}{\partial \mu_s} \frac{\partial \mu_s}{\partial \tilde{x}_{sr}} - d_r \right) \left(\frac{w_{sr}}{\tilde{x}_{sr}} - d_r \right) \tilde{x}_{sr} \qquad (92)$$

where

$$w_{sr} = \tilde{x}_{sr} \frac{\partial U_s}{\partial \mu_s} \frac{\partial \mu_s}{\partial \tilde{x}_{sr}} \qquad (93)$$

Substituting (86) and (93) into (92) we obtain that the proposed adaptation algorithm increases the social welfare:

$$\dot{W} = \sum_s \sum_{r \in R_s} \left(\frac{\partial U_s}{\partial \mu_s} \frac{\partial \mu_s}{\partial \tilde{x}_{sr}} - d_r \right)^2 \tilde{x}_{sr} \geq 0 \qquad (94)$$

and in a convex case maximizes the social welfare.

Note that proportional fairness of this scheme is a result of property (88) of approximation (10). For more general trade-offs than (10) property (88) may not hold, and thus ensuring of the proportional fairness may require more complicated pricing structure than (81).

6. EXAMPLES

This Subsection A looks at benefits of multi-path routing. Subsection B considers a case of feasible routes without overlapping links.

6.1 Benefits of Multi-path Routing

In a case of a single-path routing, when user traffic must be routed on a single path, the optimal route and the corresponding price of a unit of the reliable throughput under approximation (10) are determined by solution to the following optimization problem

$$D_{r_s^*(\gamma)} = \min_{r \in R_s} D_r(\gamma) \qquad (95)$$

where the price of a unit of the reliable throughput on a route r is

$$D_r(\gamma) = d_r / (1 - \theta_r \sqrt{2\gamma}) \qquad (96)$$

Figure 7 sketches the price of a unit of the reliable throughput on a fixed route (96), the price of optimal single-route implementation (95) (fat curve), and the price of optimal implementation using multipath routing (71) as functions of the reliability parameter γ.

Figure 7 assumes a typical situation, when higher quality routes are more congested due to higher demand: $d_{r_1} < d_{r_2} < d_{r_3}$, while $\theta_{r_1} > \theta_{r_2} > \theta_{r_3}$. In a case of a single-path routing, when user reliability requirements for γ are low, the least congested, low quality route r_1 should be used. As user reliability requirements increase, the user traffic should be carried on more congested, higher quality route r_2. As user reliability requirements keep increasing, the user traffic should be shifted to the most congested route r_3 having the highest quality. Sufficiently high user reliability requirements cannot be met with a single-path routing.

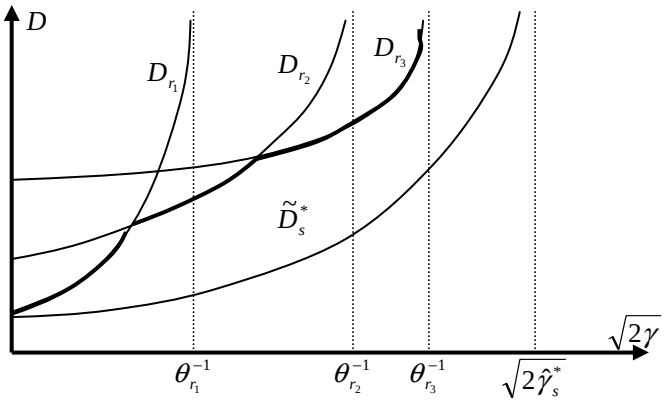

Fig. 7. Price of a unit of the reliable throughput

Since, according to (95)-(96), maximal reliability exponent user s can achieve with a single path routing is

$$\breve{\gamma}_s^* = (1/2) \max_{r \in R_s} \theta_r^{-2}, \qquad (97)$$

it follows from (31) that this user can increase its reliability exponent with multi-path routing without overlapping links up to

$$\Gamma_s = \left(\sum_{r \in R_s} \theta_r^{-2} \right) \min_{r \in R_s} \theta_r^2 > 1 \qquad (98)$$

times. Gain (98) increases with increase in the routing diversity. Beneficial effect of multi-path routing on load balancing manifests itself in reduction of the average price of the unit of reliable throughput. Generally, this beneficial effect increases with increase in the user reliability requirements. Note that multi-path routing does not have beneficial effect for a user not concerned with reliability $(\gamma = 0)$, since in this case optimal implementation is based on the minimum congestion cost routing.

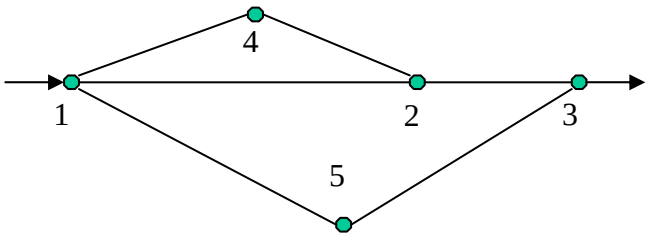

Fig. 8. Network topology

To get feeling of equal cost multi-path routing consider a network shown on Figure 8. The network has three feasible routes $r_1 = (1,2,3)$, $r_2 = (1,4,2,3)$, and $r_3 = (1,5,3)$ with the same congestion costs: $d_1 = d_2 = d_3 = d$, and matrix

$$\Theta = \begin{pmatrix} \theta^2 & \chi\theta^2 & 0 \\ \chi\theta^2 & \theta^2 & 0 \\ 0 & 0 & \theta^2 \end{pmatrix} \quad (99)$$

where parameter $\chi \in [0,1]$ characterizes overlapping between routes r_1 and r_2. In this case the optimal load split (80) is as follows: $\alpha_1 = \alpha_2 = \frac{1}{3+\chi}$, $\alpha_3 = \frac{1+\chi}{3+\chi}$.

If $\chi = 0$, i.e., equal cost routes r_1, r_2 and r_3 do not overlap, the optimal allocation splits load equally among these three routes: $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$. If $\chi = 1$, i.e., matrix (99) describes a network with just two equal cost routes $r \equiv r_1 \equiv r_2$ and r_3, the optimal loads allocation splits load equally among these two routes: $\alpha_r = \alpha_3 = 1/2$.

6.2 Routes without Overlapping Links

To illustrate our results, consider a case of K feasible routes without overlapping links: $\Theta = diag\left(\theta_1^2, \theta_2^2, .., \theta_K^2\right)$, where without loss of generality we assume that $\theta_1 \geq \theta_2 \geq \theta_K$, i.e., route r_1 has lower quality than route r_j if $1 \leq i < j \leq K$.

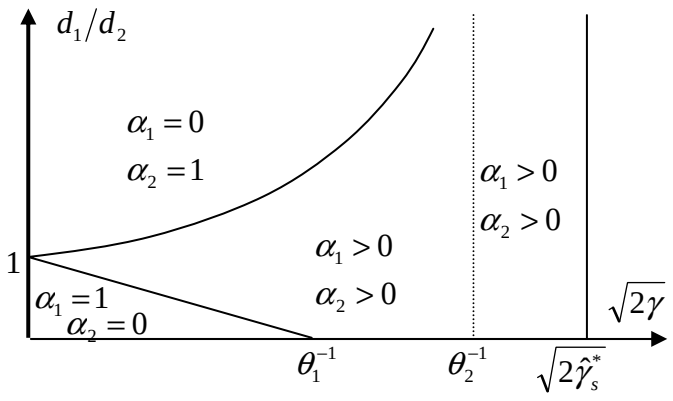

Fig. 9. Optimal route mixture, given route costs

Figure 9 sketches the phase diagram, given the route costs d_k, $k = 1,2$ and reliability exponent γ in a case of $K = 2$ feasible routes. This diagram shows three qualitatively different region with respect to the optimal route mixture (α_1, α_2), where α_k is the portion of the user traffic to be routed on path r_k, given route relative congestion costs d_1/d_2 and user reliability requirements γ. In the region $\alpha_1 = 1,\ \alpha_2 = 0$ entire user traffic should be sent over route r_1. In the region $\alpha_1 = 0,\ \alpha_2 = 1$ entire user traffic should be sent over route r_2. In the region $0 < \alpha_1, \alpha_2 < 1$ user traffic should be split between routes r_1 and r_2. Also note that the part of Figure 9, where $d_1/d_2 \leq 1$ represents a typical situation when lower quality route is less congested.

It is instructive to analyze the optimal route mixture as user reliability requirements γ or relative route congestion cost d_1/d_2 changes. Not reliability conscious user should use the minimum cost route. As user reliability requirements γ increase, multi-path routing becomes preferable until upper bound (28) is reached. Consider change in optimal connectivity as low quality route r_1 becomes more congested, i.e., as d_1/d_2 increases from zero to infinity. In this case optimal connectivity for not reliability sensitive user should change from single route r_1 to multi-path routing $r_1 \bigcup r_2$, and eventually to single high quality, less congested route r_2. Connectivity for moderately reliability sensitive user should change from multi-path routing $r_1 \bigcup r_2$ to single route r_2 since low quality route r_1 alone cannot provide required transmission reliability. Highly reliability sensitive user should be always connected over both routes: r_1 and r_2, since neither route alone can guarantee required transmission reliability. Generalization to case of an arbitrary number of feasible routes without overlapping links is straightforward.

Figures 10 sketches the phase diagram with respect to the optimal route mixture, given the average route capacities $\tilde{c}_k$, $k = 1,..,K$ and service parameters (μ, γ). Figure 10 assumes a typical situation when lower quality routes have higher capacity: $\tilde{c}_1 > .. > \tilde{c}_K$. In the region $A_k = (\mu, \gamma : \alpha_1, .., \alpha_k > 0; \alpha_{k+1} = .. = \alpha_K = 0\}$ routes $r_1, .., r_k$ are utilized while routes $r_{k+1}, .., r_K$ are not. As

service requirements (μ,γ) become more demanding, lower capacity, and thus more expensive, routes are utilized. Upper-right border of the region A_k also represents service requirements (μ,γ) having the same congestion cost. In a case of penalty function (40), the link l cost is $d_l = \tilde{c}_l/(\tilde{c}_l - \tilde{y}_l)^2$, where the average link load is $\tilde{y}_l$, and thus, the upper-right border of the region A_k represents service requirements (μ,γ) resulting in the same average delay on route r_{k+1}: $T = 1/\tilde{c}_{k+1}$ if $k = 1,..,K-1$, and $T = \infty$ if $k = K$.

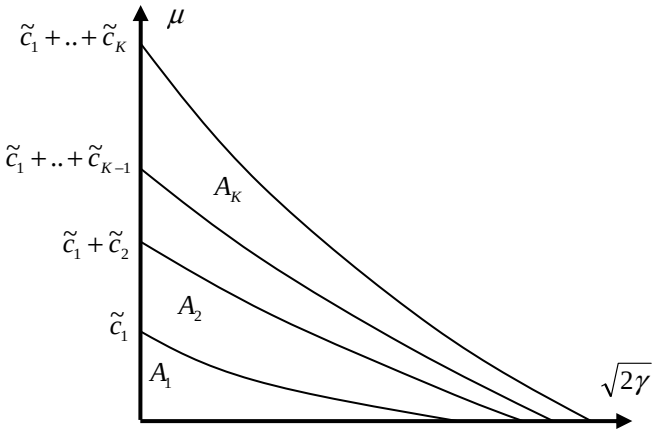

Fig. 10. Optimal route mixture, given route capacities.

7. CONCLUSION

This paper has proposed a framework for aggregate utility maximization in a distributed environment, where utilities are expressed in terms of application-level requirements. The framework assumes presence of the Intelligent Plane, which isolates users from the network layer. Numerous issues deserve further investigation, including the following: (a) Stability in presence of delays in feedback loops. (b) Property (88) ensures that pricing structure (81) results in proportionally fair resource allocation. Property (88) is a result of approximation (10) and may not hold in other situations, e.g., for a link failure model, when more sophisticated pricing schemes may be required to ensure proportional fairness [15]-[16]. (c) Possible generalization to a case when users are not only "buyers" but also "sellers" of the limited resources, such as in a case of a wireless multi-hop network, when intermediate nodes may expend their battery energy for relaying other users' traffic [17]-[18].

8. ACKNOWLEDGMENTS

Author appreciates illuminating comments by A. Nakassis.

9. REFERENCES

[1] S. Shenker, "Fundamental design issues for the future Internet," *IEEE JSAC*, 13 (1995) 1176-1188.

[2] F.P. Kelly, "Charging and rate control for elastic traffic," European Transactions on Telecommunications, Vol. 8, 1997, pp. 33-37.

[3] F.P. Kelly, A.K. Maulloo, and D.H.K. Tan, "The rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, pp. 237-252, vol. 409, 1998.

[4] D.D.. Clark, C. Partridge, J.C. Ramming, and J. Wroslawski, "A knowledge plane for the Internet," *ACM SIGCOMM*, 2003.

[5] S. H. Low, F. Paganini and J. C. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, 22(1), Feb. 2002, 28-43.

[6] V. Marbukh and R.E. Van Dyck, "On aggregate utility maximization by greedy ASs competing to provide Internet services," *Forty-Second Annual Allerton Conference on Communication, Control, and Computing*, Illinois, 2004.

[7] A. Tsirigos and Z.J. Haas, "Analysis of multipath routing – part I: the effect on the packet delivery ratio," IEEE Trans. *On Wireless Communications*, Vol. 3, No. 1, January 2004, pp. 138-146.

[8] N.F. Maxemchuk, "Dispersity routing, Proc. IEEE ICC, San Francisco, CA, 1975.

[9] A. Tsirigos and Z.J. Haas, "Analysis of multipath routing – part II: mitigation of the effects of frequently changing network topologies," *IEEE Trans. On Wireless Commun*, Vol. 3, No. 2, 2004, 500-511.

[10] J. MacKie-Mason, "What does economics have to do with survivability?" InfoSurv, June 1997. available at http://www-personal.umich.edu/~jmm/presentations/darpa-econ-tutorial.pdf.

[11] V. Marbukh, "Towards market approach to providing survivable services," *Conf. on Inf. Sciences and Syst.* Princeton Univ., 2004.

[12] F.P. Kelly, "Notes on effective bandwidth," In "Stochastic Networks: Theory and Applications" (*Editors F.P. Kelly, S. Zachary and I.B. Ziedins*) Royal Statistical Society Lecture Notes Series, 4. Oxford University Press, 1996. 141-168.

[13] S. Boyd and L. Vandenberge, Convex Optimization, Cambridge University Press, 2004.

[14] D. Bertsekas and R. Gallger, Data Networks, Prentice-Hall, 1992.

[15] C. Courcoubetis and R. Weber, Pricing Communication Networks, Wiley, 2003.

[16] V. Marbukh, "A knowledge plane as a pricing mechanism for aggregate, user-centric utility maximization," *ACM Performance Evaluation Review*, Volume 32, Issue 2, pp. 22-24, 2004

[17] J. Crowcroft, R. Gibbens, F. Kelly and S. Ostring, "Modelling insentives for collaboration in mobile ad hoc networks," *Performance Evaluation*, 57, pp. 427-439, 2004.

[18] P. Marbach and Ying Qiu, " Cooperation in Wireless Ad Hoc Networks: A Market-Based Approach," available at http://www.cs.toronto.edu/~marbach/publ.html.

[19] V. Marbukh, "On aggregate utility maximization based network management: challenges and possible approaches," *IEEE ICC 2004*