On Designing Application-Specific Overlays

Parijat Dube, Zhen Liu, Sambit Sahu, and Jeremy Silber
IBM T. J. Watson Research Center
Yorktown Heights, NY, 10598
pdube,zhenl,sambits,jeremy.silber@us.ibm.com

Abstract. Experimental studies have shown that the performance of overlay networks is largely dependent on the limited access link bandwidth at the end-system nodes. We provide a theoretical framework for understanding the impact of different performance metrics on the overlay network construction. We define different metrics to express the sensitivity of various types of applications to packet loss. Using simple queueing models of access links at the end-system nodes, we obtain explicit characterization of the metrics in terms of the forwarding degree of the tree, the depth of the tree, network load and node parameters (bandwidth and buffer). These expressions can be exploited to obtain qualitative and quantitative insights into the sensitivity of overlay performance to tree topology and network parameters.

Keywords: overlay networks, last mile problem, application specific loss metrics, queueing model.

1 Introduction

1.1 Problem Motivation

End-system (application level) multicast has been advocated [3, 4] as a practical alternative to IP-multicast for supporting large scale group communication due to its ease of deployment that demands no changes to existing network layer protocol. Typically, a forwarding overlay tree is constructed using participating end-system devices and (optionally) forwarding servers to distribute content. While such overlay based communication is attractive due to the ease of deployment, one has to address performance concerns that arises due to end-system forwarding. One of the performance concern is due to the limited access link bandwidth to the end-systems that may cause packet losses and increased delay in packet delivery.

While several studies [3, 6, 5, 1, 2] have proposed intelligent overlay construction accounting for such last mile constraints, these have primarily focussed on bounding long term averages for end-to-end loss and delay. The performance of several enterprise applications supporting group communications, such as on-demand video conferencing, VoIP, netmeetings have varying degree of degradation with data loss. For example, while minimization of average packet loss is desirable for large file transfer, videoconferencing applications may require bounds on the number of packets that are lost by any of the group members. Then there is concern about long term averages and short term performance; and bursty losses and spread out losses for same average loss probability. Thus it is imperative to define and support application specific performance metrics for the satisfactory performance of different applications running on overlay networks. The definition of performance metrics for supporting group communication on overlay networks and in particular the metrics associated with data loss may involve both subjective as well as objective approaches as are traditionally followed for unicast VoIP sessions.

1.2 Our Contributions

We begin by defining loss metrics for overlay-based group communication to capture a wide range of application specific requirements. The first metric, $c_1$, is a measure of the maximum packet loss probability among all the nodes. For the second metric, $c_2$, we are interested in the expected correct delivery of all
packets to all the nodes in the overlay. While first metric may be a useful performance measure for non-temporal applications like long term file transfers etc., the second metric may be a useful performance measure for temporal applications (with stringent packet loss requirements) like instant messaging, secure network services, content delivery applications including streaming media for e-learning and corporate communications and mission critical data distribution.

We model the limited bandwidth of access link path of each end-system using a simple queuing model that captures the interaction between the packet processing at the incoming and outgoing interfaces. We first propose a simple bufferless queueing model and later generalize the approach for finite buffer queueing models. The bufferless model gives explicit closed form expressions for loss metrics in terms of depth of the overlay tree, the fanout of nodes, and network parameters like node and link bandwidth. While the analysis using finite buffer model is not so explicit and expressions for loss metrics can be obtained in a recursive manner, the model is more realistic as it accounts for buffers at the input and output interfaces at end-systems and can be used to obtain queueing delays at nodes. Using the analytical results, we study the sensitivity of different loss metrics to the forwarding degree and depth of the tree.

The rest of the paper is organized as follows: Section 2 introduces two different performance metrics for packet loss. In Section 3 we motivate the queueing model and derive explicit expressions for loss at an individual node. We derive expressions for the $m$-ary overlay tree for different metrics for packet loss in Section 4. Section 5 illustrates the numerical results based on our analysis and studies the sensitivity of overlay performance. Section 6 summarizes the results and discusses our ongoing future works.

## 2 Application Specific Performance Metrics for Packet Loss

Figure 1 illustrates a motivational application scenario deployed in the Internet using overlay network. We shall be concerned with packet loss in this study to show how different metric for packet loss may influence the construction of an overlay network and assume delay to be dominated by propagation delay over links and (assuming the nodes to be in reasonable proximity) is given in terms of the hop count. The finiteness of the buffers allows us to account for queueing delays at overlay nodes. The study involving queueing delay in addition to propagation delay is the subject of our ongoing work. We define the following performance metrics:

- $c_1$: minimizing $c_1$ amounts to maximizing the fraction of packets that are correctly received by the worst node (the worst node being the one with the largest end to end packet loss probability).
Fig. 2. Various components in an end-to-end path

- $c_2$: minimizing $c_2$ amounts to maximizing the expected delivery of all packets to all the nodes in the overlay.

Next we mathematically characterize these metrics. Let us consider the case of transmission of application packets by the source node to a group of $N-1$ nodes. For the $i$th node we define the random variable $Y_i(n)$ as: $Y_i(n) = 1 (0)$, if the $n$th packet is received by node $i$ (otherwise). Define $\tilde{Y}_i(n) = 1 - Y_i(n)$. We consider $\{\tilde{Y}_i(n), n \geq 1\}$ to be i.i.d. (the distribution we shall determine in our analysis).

- $c_1 = \max_i E[\tilde{Y}_i] = \max_i p_i$, where $p_i$ is the probability that the packet is not received by node $i$.
  
  Thus minimizing $c_1$ amounts to minimizing the loss probability of the worst node.

- $c_2 = E[\max_i \tilde{Y}_i]$: minimizing $c_2$ loss amounts to maximizing the probability of correct delivery of each packet to each node. This is an important metric for temporal, mission-critical group communication applications where it is highly desirable that all the nodes are receiving all the information.

Observe that $c_1 < c_2$ (by Jensen’s inequality), thus $c_2$ is a much stringent loss requirement.

3 Queueing Model to Capture Access Link Sharing

Figure 2 illustrates the various components involved in the forwarding of data from one host to another using overlay network. First host 1 sends data to its output interface, which is transmitted to the (DSL/Cable/Dialup) access hub using the uplink. The access hub forwards to the other end over the high speed Internet that is received at host 2’s (DSL/Cable/Dialup) access hub. This data is then forwarded to host 2 on its “last mile” downstream link.

We consider the queueing model depicted in Figure 3 to abstract different components in the data forwarding path in an overlay. The first queue represents forwarding over the upstream link that is shared across all the data forwarding paths from host 1 to its children nodes. We approximate the service rate available to one receiving host (in this case host 2) as $r_{out}/m$, where $m$ is the forwarding degree of host 1. We also assume that a $B_{out}$ amount of buffer is allocated for buffering data in the kernel/device for forwarding data on this stream. The next service station with infinite server model accounts for the propagation delay from one access hub to the other end of the access hub over the Internet. There is no queueing and loss at this station as the Internet is order of magnitude faster compared to the last mile. Next queue with service rate $r_{in}$ represents the forwarding from access hub to the host 2 over the downstream last mile link. The queue in host 2 represents the processing overhead with service rate $u_s$ and buffer capacity $B_{in}$ at the input interface of host 2.

Given that the processing overhead is much smaller compared to the forwarding rate on the outgoing interface, we represent each node with a single queue with a buffer size $K \equiv B_{in} + B_{out}$ with the processing

\[1\] $p_i$ is basically the loss probability of the path from the source to the $i$ node.
rate $r/m$ (with $r$ being the link bandwidth and $r \approx r_{\text{out}}$). We consider the problem in a homogeneous setting for analytical tractability, i.e., each node has same forwarding capacity and forwarding fanout of every node is assumed to be same for every participating nodes. Thus we limit our illustrations to regular $m$-ary tree configuration for overlay network topology. Let $d$ be the depth of a topology and $N$ represent the number of end-system nodes in the overlay network. We represent the end-to-end delay as the number of hops from the source to the destination node. Note that we are concerned with regular $m$-trees in our analysis. An $m$–tree is characterized by a constant branching ratio of $m$ and the depth of the tree is the smallest integer $d$ that satisfies:

$$N \leq 1 + m + m^2 + \ldots + m^d$$

Thus $d = \lceil \frac{\ln(N(m-1)+1)}{\ln m} - 1 \rceil$ when the delay is expressed as the hop distance.

Fig. 4 shows a $m = 2$ tree with $N = 3$. As illustrated in Figure 4, the arrival process at the downstream node is thinned due to losses at upstream nodes. Let the input arrival rate to node $A$ is $\lambda$ and $p$ is the loss probability of a data unit at $A$. By flow conservation the average departure rate from $A$ is $\lambda(1-p)$. Thus its children $B_1$ and $B_2$ will each have the average input arrival rates as $\lambda(1-p)$.

### 3.1 Bufferless Model: Loss and Delay Expressions

We first approximate the network of nodes between the source and the leaf node as a tandem of $M/M/1$ queues with zero buffer. We take the mean size of a packet as $1/\mu$. If a node has $m$ children and the node bandwidth is shared in a processor sharing manner (at packet level) we approximate the service rate at each of the $m$ connections by $r/m$. Thus the mean service time of a packet is $m/\mu$. In this approximation the input to queue 0 is Poisson with rate $\lambda_0 \equiv \lambda$, to queue 1 is Poisson with rate $\lambda_1 \equiv \lambda(1 - P_0)$ and similarly the input to queue $j$ is Poisson with rate $\lambda_j \equiv \lambda \prod_{k=0}^{j-1}(1 - P_k)$, where $P_k$ is the loss probability at a bufferless $M/M/1$ queue (representing the $k$th node in the path between a node and the root node) with arrival rate $\lambda \prod_{k=0}^{j-1}(1 - P_k)$ and service rate $m/\mu$. Let $\rho = \frac{m \lambda}{\mu}$.

By simple $M/M/1$ formulation with zero buffer the loss probability of a packet at node 0 is given by:

$$P_0 = \frac{m \rho}{1 + m \rho}$$

Fig. 4. A $m = 2$ tree with $N = 3$. 

\begin{equation}
\rho = \frac{m \lambda}{\mu}
\end{equation}
In general for $0 \leq j \leq d$, we get

$$P_j = \frac{m \rho}{1 + (j + 1)m \rho}. \quad (1)$$

Observe that because of symmetry, the loss probability at all nodes at level $j$ is given by $P_j$, $0 \leq j \leq d$.

### 3.2 Finite Buffer Model: Loss and Delay Expressions

To capture the finiteness of the buffers at overlay node and hence the packet losses and queueing delays at the overlay nodes we model a node as $M/M/1/K$ queue instead of bufferless as in the previous model. Again we approximate the input to queue 0 as Poisson with rate $\lambda_0 = \lambda$, to queue 1 Poisson with rate $\lambda_1 = \lambda(1 - \pi_0)$ and similarly the input to queue $j$ is Poisson with rate $\lambda_j = \lambda \Pi_{k=0}^{j-1} (1 - \pi_k)$, where $\pi_k$ is the loss probability at a $M/M/1/K$ queue (representing the $k$th node) with arrival rate $\lambda \Pi_{i=0}^{k-1} (1 - \pi_i)$ and service rate $\frac{\mu c}{m}$. From queueing theory, we have $\pi_0 = \frac{(m \rho)^k}{\sum_{j=0}^{\infty} (m \rho)^j}$. The expressions for $\pi_j$, $1 \leq j \leq d$ can be obtained similarly by accounting for losses at upstream nodes.

### 4 Expressions for Different Performance Metrics

We shall obtain expressions for $c_1$ and $c_2$ using the loss probability relations obtained in previous section. We shall restrict our analysis to bufferless model and obtain explicit closed form expressions.

#### 4.1 Expression for $c_1$

With any packet we associate at each node $j$, $0 \leq j \leq d$, in tandem, an indicator function $\bar{X}_j$. $\bar{X}_j = 0$ if the packet is lost at the node and is 1 otherwise. Also given $\bar{X}_{j-1} = 0$ we take $\bar{X}_i = 0$, $i \geq j$.

From our analysis we get $E[X_0] = (1 - P_0)$, $E[X_1|X_0 = 1] = (1 - P_1)$ and in general $E[X_j|X_i = 1, 1 \leq i \leq j - 1] = (1 - P_j)$, where $P_j$’s are given by (1). Let $\bar{p}(j)$ be the probability that a packet is received at a node at depth $j$, $0 \leq j \leq d$. $\bar{p}(0)$ is probability that a packet is not lost at the root node in an $m$-tree.

We have $\bar{p}(0) = 1 - P_0 = \frac{1}{1 + m \rho}$, $\bar{p}(1)$ is given by:

$$\bar{p}(1) = P(\bar{X}_0 = 1, \bar{X}_1 = 1) = E[\bar{X}_0]E[\bar{X}_1|\bar{X}_0 = 1] = (1 - P_0)(1 - P_1) = \frac{1}{1 + 2m \rho}.$$ 

Continuing this way, the probability that a packet correctly reaches a node at depth $d$ is given by

$$\bar{p}(d) = \frac{1}{1 + (d + 1)m \rho}.$$ 

Thus the maximum end-to-end loss probability (as a function of $m$ and $d$), $p(d) \overset{\text{def}}{=} 1 - \bar{p}(d)$, is $\frac{(d+1)m \rho}{1 + (d + 1)m \rho}$.

For a complete tree $d + 1 = \frac{\ln(N(m-1)+1)}{\ln m}$.

Since it is an $m$–tree, all leaf nodes at depth $d$ (this is the same as all leaf nodes if its a regular $m$–tree) will have same probability that a packet from the root will reach them correctly. Thus to calculate $c_1 := \max_{1 \leq i \leq N} p_i$, we look at the end-to-end loss probability seen by the leaf nodes and hence $c_1$ is simply the end-to-end loss probability from the source to any of the leaf nodes, given by:

$$c_1 = \frac{1}{1 + \left(\frac{m \rho \ln(N(m-1)+1)}{\ln m}\right)^{-1}}. \quad (3)$$

$^2$ This is because if the packet is lost at an upstream node it will never reach the downstream node as so is considered to be lost at downstream nodes.

$^3$ We assume that a packet is received at any node $j$ correctly if it is not lost at node $j$ and at any of the succesive upstream nodes in the path from the root till node $j$. 


4.2 Expression for $c_2$

To calculate $c_2$ we first obtain the distribution of $\max_{i=\{1,\ldots,N\}} \tilde{Y}_i$. Observe that $\tilde{Y}_i = \tilde{Y}_{i-1} + (1 - \tilde{Y}_{i-1})X_i$. Thus we have $\max(\tilde{Y}_i, \tilde{Y}_{i-1}) = \tilde{Y}_i$. Continuing in this way we shall observe that $\max_{j=\{N-m^d, N-m^d+1, \ldots, N\}} \tilde{Y}_i$, i.e., we should only look at the max over set of leaf nodes. To make the analysis simple, we shall first obtain the expression for $c_2$ for a binary tree with depth 2 and later generalize it to any $m$-ary tree with depth $d$. Thus we will first obtain the distribution of $\max_{i=\{1,\ldots,7\}} \tilde{Y}_i$.

Let random variable $X_i = 1 - X_i$ for $i = 1, \ldots, 7$. From our analysis in previous section, we have $p_1 \equiv P_0$, $p_2 = p_3 \equiv P_2$, $p_4 = p_5 = p_6 = p_7 \equiv P_3$. After some algebra we get:

$$\max_{i=\{1,\ldots,7\}} \tilde{Y}_i = \max(\tilde{Y}_4, \tilde{Y}_5, \tilde{Y}_6, \tilde{Y}_7)$$

$$= \max(X_1 + (1 - X_1)X_2 + (1 - X_1)(1 - X_2), X_1 + (1 - X_1)X_3 + (1 - X_1)(1 - X_3))$$

$$= \max(X_1 + (1 - X_1)X_2 + (1 - X_1)(1 - X_2)Z_2^{(1)}, X_1 + (1 - X_1)X_3 + (1 - X_1)(1 - X_3)Z_2^{(2)})$$

$$= X_1 + (1 - X_1)\left(\max\left(X_2 + (1 - X_2)Z_2^{(1)}, X_3 + (1 - X_3)Z_2^{(2)}\right)\right) = X_1 + (1 - X_1)Z_1^{(1)},$$

where $Z_2^{(1)} = \max(X_4, X_5)$, $Z_2^{(2)} = \max(X_6, X_7)$ and $Z_1^{(1)} = \max\left(X_2 + (1 - X_2)Z_2^{(1)}, X_3 + (1 - X_3)Z_2^{(2)}\right)$. Observe that all the random variables are binary valued taking 0 or 1 values. Thus $E[\max, Y_i] = E[X_1 + (1 - X_1)E[Z_1^{(1)}|X_1 = 0]$. We shall next find the conditional distribution of $Z_1^{(1)}$ given $X_1 = 0$. This requires finding the conditional distribution of $\max\left(X_2 + (1 - X_2)Z_2^{(1)}, X_3 + (1 - X_3)Z_2^{(2)}\right)$ given $X_1 = 0$. We shall first find the conditional expectation of random variables $X_2 + (1 - X_2)Z_2^{(1)}$ and $X_3 + (1 - X_3)Z_2^{(2)}$ given $X_1 = 0$. Observe that

$$E[X_2 + (1 - X_2)Z_2^{(1)}|X_1 = 0] = p_2 + (1 - p_2)E[Z_2^{(1)}|X_1 = 0, X_2 = 0]$$

Similarly we have:

$$E[X_3 + (1 - X_3)Z_2^{(2)}|X_1 = 0] = p_3 + (1 - p_3)E[Z_2^{(2)}|X_1 = 0, X_3 = 0]$$

Further we have $E[Z_2^{(1)}|X_1 = 0, X_2 = 0] = E[\max(X_4, X_5)|X_1 = 0, X_2 = 0]$. We shall next find the conditional distribution of $X_4$ and $X_5$ given $X_1 = 0$ and $X_2 = 0$. We have:

$$P(X_4 = 1|X_1 = 0, X_2 = 0) = p_4 = 1 - P(X_4 = 0|X_1 = 0, X_2 = 0)$$

Similarly we have:

$$P(X_5 = 1|X_1 = 0, X_2 = 0) = p_5 = 1 - P(X_5 = 0|X_1 = 0, X_2 = 0)$$

Thus we have from (6) and (7):

$$E[Z_2^{(1)}|X_1 = 0, X_2 = 0] = 1 - (1 - p_4)(1 - p_5).$$

Which implies from (4) and (8):

$$E[X_2 + (1 - X_2)Z_2^{(1)}|X_1 = 0] = p_2 + (1 - p_2)(1 - (1 - p_4)(1 - p_5)),$$

and similarly we can obtain from (5) and (8):

$$E[X_3 + (1 - X_3)Z_2^{(2)}|X_1 = 0] = p_3 + (1 - p_3)(1 - (1 - p_6)(1 - p_7)).$$
The symmetry of the results suggests that the arguments can be generalized to an \(m - ary\) tree with depth \(d\). Thus we have

\[
E[\max_i \hat{Y}_i] = 1 - \prod_{j=0}^{d-1} (1 - P_j)^{v(j)} = 1 - \prod_{j=0}^{d} \left( \frac{1 + jm \rho}{1 + (j+1)m \rho} \right)^{m^j}.
\]

From (12) we have: \(c_2 = 1 - \prod_{j=0}^{d} \left( \frac{1 + jm \rho}{1 + (j+1)m \rho} \right)^{m^j} \).

### 5 Numerical Studies

We first plot \(c_1\) and \(c_2\) as a function of the depth \(d\) and forwarding degree \(m\) of the tree for \(\rho = 0.001\) in Figs. 5 and 6 respectively and further in Figs. 7 and 8 for \(\rho = 0.01\). We quantitatively study the maximum number of users that can be supported in the overlay \(N_{max} \equiv \left( \frac{m^{d+1} - 1}{m^d - 1} \right)\), if the bounds on \(c_1\) and \(c_2\) are relaxed. Table 1 shows that relative increase in the maximum number of nodes that can be supported as \(c_1\) and \(c_2\) bounds are relaxed for \(d = 4\) and \(\rho = 0.001\). From the Table we observe that for a given depth as \(c_1\) is relaxed from 0.0079 to 0.039, \(N_{max}\) increases from 31 to 1111. The gain is more significant for large \(d\) as we observe from Fig. 5, as \(d\) increases the slope of the \(c_1\) vs. \(m\) curve increases.

Since \(c_2\) is a much stringent requirement as the degree of the tree is increased to support a large population of users (see Table 1 and Fig. 6 and 8) the degradation in \(c_2\) is quite significant as compared...
Table 1. Table showing the relative increase in the maximum number of nodes that can be supported as $c_1$ and $c_2$ bounds are relaxed for $d = 4$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$N_{\text{max}}$</th>
<th>$c_1$</th>
<th>$c_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>31</td>
<td>0.0099</td>
<td>0.0579</td>
</tr>
<tr>
<td>3</td>
<td>121</td>
<td>0.0148</td>
<td>0.300</td>
</tr>
<tr>
<td>4</td>
<td>341</td>
<td>0.0196</td>
<td>0.7389</td>
</tr>
<tr>
<td>5</td>
<td>781</td>
<td>0.0244</td>
<td>0.9785</td>
</tr>
<tr>
<td>6</td>
<td>1555</td>
<td>0.0291</td>
<td>0.9999</td>
</tr>
<tr>
<td>10</td>
<td>1111</td>
<td>0.0476</td>
<td>1.000</td>
</tr>
</tbody>
</table>

![Fig. 7. The gain in $N_{\text{max}}$ ($= \frac{m^{d+1} - 1}{m^d - 1}$) with the relaxation in $c_1$ for bufferless model with $\rho = 0.01$.](image1)

![Fig. 8. The gain in $N_{\text{max}}$ ($= \frac{m^{d+1} - 1}{m^d - 1}$) with the relaxation in $c_2$ for bufferless model with $\rho = 0.01$.](image2)

...to the degradation in $c_1$. An interesting observation is that as the depth of the trees increases $c_2$ increases quite sharply as compared to $c_1$ suggesting that increasing the depth of the tree so as to increase the population of supported nodes might potentially violate the performance requirement of applications and the adverse effects are more severe on $c_2$ as compared to $c_1$.

We next plot similar curves for the finite buffer model by using our analytical results in Fig 9 and Fig. 10 for $c_1$ and $c_2$ respectively. The trends are similar to curves with bufferless model.

5.1 Network Simulations: Model Validation and Robustness

To validate our mathematical models and its assumptions we created a discrete event simulator. The simulator allows the creation of arbitrary networks of buffered nodes, exponential and deterministic service processes for packet forwarding, several packet arrival processes (Poisson, Markov Modulated Poisson (MMPP), and Constant Bit Rate (CBR)), and various network latency delays. Each node has its own buffer and input/output interfaces to its neighbors in the overlay, and all events occur according to a common timeline.

Each simulation data point reported represents an average over 10 or more simulation runs (with different random seeds), and with 10,000 packets sent per run. All simulations use a network size of 10,000 nodes. Figure 11 compares $c_1$ from the analysis and the simulation results for the Poisson arrival case with different load and node buffer size. We observe that our assumptions on the arrivals into the subsequent nodes in the overlay being Poisson match very closely with the simulation results.

Figure 12 illustrates the robustness of our model with MMPP arrival process for $\rho = 0.5, 0.1$. We compare the value of $c_1$ with MMPP arrival process obtained via simulation with the analysis using the Poisson arrival process. The arrival rate of the Poisson process is taken equal to the mean arrival rate of the MMPP process. We find a good match between simulation results for MMPP and our analysis with Poisson arrival.

Typically applications have loss and/or delay requirements and an overlay supporting the application should be constructed such that it meets the loss and delay bounds while supporting the application. To
We consider three arrival processes, i.e., Poisson, MMPP and on-off CBR arrival at the origin source. We take $c_1$ bound ($\beta$) to be 0.35, and delay bound ($\Delta$) as 6, and $\rho = 0.1$. Table 2 illustrates the set of feasible degrees for loss bound ($P_\beta$), for delay bound ($D_\Delta$), and that meets both the loss and delay bounds ($P_\beta \cap D_\Delta$). We find that our model is able to predict the feasible degree of the overlay quite accurately for the MMPP and CBR case. We observe that for arrival processes other than Poisson (and

<table>
<thead>
<tr>
<th>Approach</th>
<th>$P_\beta$</th>
<th>$D_\Delta$</th>
<th>$P_\beta \cap D_\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Analysis</td>
<td>(2.4, ..., 10)</td>
<td>(2.4, ..., 12)</td>
<td>(2.4, ..., 10)</td>
</tr>
<tr>
<td>Poisson Simuls.</td>
<td>(2.4, ..., 10)</td>
<td>(2.4, ..., 12)</td>
<td>(2.4, ..., 10)</td>
</tr>
<tr>
<td>MMPP Simuls.</td>
<td>(2.4, ..., 10)</td>
<td>(2.4, ..., 12)</td>
<td>(2.4, ..., 10)</td>
</tr>
<tr>
<td>On-off CBR Simuls.</td>
<td>(2.4, ..., 20)</td>
<td>(2.4, ..., 12)</td>
<td>(2.4, ..., 10)</td>
</tr>
</tbody>
</table>

Table 2. The set of feasible degrees obtained from analysis and from simulations using different arrival processes and packet size distribution.

capture this, we next define the concept of a feasible set of forwarding degree of the tree. A feasible set of forwarding degree is a set, $\tilde{M} = \{\tilde{m}\}$ of degrees such that a regular tree constructed with degree $\tilde{m} \in \tilde{M}$ satisfies the loss and/or delay bounds. We next study the robustness of our model with different packet size distribution and arrival process. In particular we are interested in identifying the set $\tilde{M}$ of feasible degree for different scenarios such that the constructed tree satisfies bounds on $c_1$ and end to end delay (hop count).

Fig. 9. The gain in $m$ with the relaxation in $c_1$ for a given depth for the finite buffer, $K = 10$, $\rho = 0.4$.

Fig. 10. The gain in $m$ with the relaxation in $c_2$ for a given depth for the finite buffer, $K = 20$, $\rho = 0.1$.

Fig. 11. Model validation, $\rho = 0.1, 0.5$, $K = 6$

Fig. 12. Model robustness with MMPP arrival process with finite buffers, $K = 6$. 

We observe that for arrival processes other than Poisson (and
loss rate for MMPP), the exact values of maximum delay and loss do not match very closely with our simple model. However, we find that feasible value of the forwarding degree for the homogeneous case is relatively insensitive to the exact statistics of the arrival process and/or the packet size distribution. This is an important result as our simple model can be used as a guideline for constructing overlay even for more complicated arrival processes.

6 Conclusions

In this study, we provided a theoretical framework for understanding the impact of different performance metrics for packet loss on the overlay network construction. We identified two broad classes of performance metrics catering to different application-specific QoS requirements. Using simple queing models, we derived explicit expressions for these loss metrics in terms of the tree topology (forwarding degree, depth) and network parameters (load, link bandwidth, node buffers).

Using the analytical results, we studied the sensitivity of different loss metrics to the forwarding degree and depth of the tree. We observe different trends in the packet loss metrics as a function of the depth of the tree for a given forwarding degree of the node. Also we observe that as the degree of the nodes is increased to support a larger number of end-system nodes, the degradation in $c_2$ is quite significant as compared to that in $c_1$. These trends indicate that one has to construct an overlay tree differently based on application specific metrics. Our observations based on our analysis and numerical results provide useful design insights for constructing overlays meeting application specific performance requirements.

In this paper, we have focussed on the different performance metrics for the packet loss. However observe that the finiteness of the buffer at the nodes adds to the end-to-end queuing delay which was approximated by the hop count in the paper. Depending on the tree topology and the network load, the queuing delay can be a significant component of the end-to-end delay. We are presently studying the tradeoff between the end-to-end delay (including queueing delays) and loss metrics within our framework, which will potentially provide us guidelines for choosing the appropriate forwarding degree to support both delay and loss sensitive applications.

References