Fat-Tree Overlays for Scalable Multisource Multicast

Leiserson’s fat-trees
- Minimal and scalable diameter
- Maximal and scalable bisection bandwidth

Resemble Leiserson’s fat-trees on the overlay

Data Forwarding in FatNemo
- Increased degree of channels

FatNemo
- Increase cluster size as one ascends the tree
- Clustered by bandwidth capacity

Narada
- Overcast
- Overlay nodes are highly transient

Overcast handles failures well

Conventional Trees
- Depend on reliability of non-leave nodes
- Overlay nodes are highly transient
- Median session time 1 hr - 1 min.

Forwarding under Failures

Our Approach
Fat Trees for Root Bottleneck Problem
We bypass the root bottleneck problem found in conventional tree-based systems by emulating Leiserson’s fat-trees. The resulting overlay fat-trees have peers with higher bandwidth capacity located higher up in the hierarchy.

Our Goal
Enabling large scale multisource multicast applications.

The Challenge
Handling highly transient populations and leverage/respect heterogeneity.

Two-metric Approach: Sufficient bandwidth & minimized latency
The overlay fat tree promotes the optimal peer based on the latency metric which also has sufficient bandwidth to the next higher layer. The algorithm is completely decentralized.

scalability of tree-based protocols

<table>
<thead>
<tr>
<th>Protocol</th>
<th>E[Layers]</th>
<th>E[Outdegree]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nice</td>
<td>log₂(N)</td>
<td>4.5 E[Layers] = 4.5 log₂(N)</td>
</tr>
<tr>
<td>Nemo</td>
<td>log₂(N)</td>
<td>1.5 E[Layers] = 1.5 log₂(N)</td>
</tr>
<tr>
<td>FatNemo</td>
<td>2.5 + (2.2 + 8 ln(N))/2.2</td>
<td>≤ E[Layers] = 2.5 + (2.2 + 8 ln(N))/2.2</td>
</tr>
</tbody>
</table>

PlanetLab – Failures
1 Publisher, 50 end hosts, high bandwidth scenarios

Simulation – Failures
1 Publisher, 256 and hosts, high bandwidth scenarios

PlanetLab – Failures
1 Publisher, 50 and hosts, high bandwidth scenarios

FatTree Resilience

FatTree Performance

FatTree Scalability

Simulation – Delivery Latency
6 publishers, 256 end hosts, high bandwidth scenarios

Simulation – Delivery Latency CDF
6 publishers, 256 end hosts, low bandwidth scenarios

Lowest delivery latency with multiple publishers

PlanetLab – Delivery Latency
1 Publisher, 50 and hosts

Low delivery latency

Simulation – Outdegree CDF
8 publishers, 256 and hosts, high bandwidth scenarios

FatNemo is highly scalable

Scalability is limited by forwarding responsibility

Ideally most peers should work in this region

Outdegree defines the forwarding load of the peers

High outdegrees leads to overloaded peers