UDC 004.75, 004.724.2

PEER-TO-PEER SWARM BEHAVIOR ON A UNIFORM MESH

Poryev G.V.
National technical university of Ukraine «Kyiv Polytechnic Institute», PhD, s.s.w.

Introduction

Today, the dominant traffic class on every major Internet backbone is peer-to-peer (p2p) traffic, most notably the file-sharing networks. According to CISCO VNI, p2p traffic is rivaling streaming multimedia, until recently the biggest traffic-wise contributor to Internet. It is therefore important to research into the area of large-scale p2p optimization, evaluation and architectural concepts. Our recent research had yielded the potential driving mechanism for such an optimization by implementing locality metric known as CARMA (Combined Affinity Reconnaissance Metric Architecture). In short, CARMA metric is derived from a model of national Internet segments built on open databases provided by the RIRs (Regional Internet Registries). The model allows for distinguishing of 8 classes of topological locality between arbitrary nodes represented as their IP addresses — from sub-range to horizon, expressed numerically from 0 to 7.

Generally speaking, CARMA metric may not only be used in p2p solutions. Whenever there is a distributed network of nodes that form overlay structure using the Internet and this network needs some kind of automatic clustering mechanism based on individual peer-selection, CARMA metric is the way to go, being not affected by any of the major drawbacks of previous techniques.

Problem

Providing a locality metric to be used in the peer reordering is just a half of the solution. It is necessary to have the reliable efficiency estimation technique to be able to ascertain the positive effect of using the proposed locality metric. The efficiency is supposed to be represented as a single value calculated from the characteristics of the nodes. The efficiency is therefore a subjective value, estimated from the "point of view" of each participating node relative to its directly connected peers, and, of course, is different for each node.

Conditions and solution

In order to build such efficiency estimation, we have designed an imitational model that represents major aspects of p2p swarm functioning.

Model contains a mesh of nodes participating in the same swarm and exchanging the same bulk of data, which is characterized by size and number of data chunks. Epoch of the model is iterated through so called "ticks", representing uniform time; for each node it takes exactly one tick to transfer one chunk of data to partner node if it can find the partner. Epoch begins when a single seeder node (having all chunks) is introduced and ends when every node has every chunk.

Though in real life the size of the swarm is unlikely to be less than node connection queue, for simplicity of the calculations and modeling we assume that in our model every node has a chunk bitmap of every other node. Therefore it can, on each tick, calculate the following parameters N_i — number of chunks downloaded by *i*-th peer node. S_{ei} and S_{ai} — number of chunks that *i*-th node has and given node has not and vice versa. The other parameters relevant to the efficiency estimation are: N — total number of nodes in the swarm; f_{max} — maximum locality class defined; f_i — locality class of *i*-th node related to local node; k — locally set download precedence quotient.

It should be noted, that equation must not be treated as having physical meaning nor as having any meaningful dimensions, for it is empiric in the sense that its values are dependent to the mesh size and number of chunks. However it is quite possible to introduce some kind of normalization, by calculating the efficiency at the last tick of the epoch. For example, mesh of 100 nodes with 100 chunks of data gives the maximum efficiency of around 2051.

The extensive research on this imitation model indicated that efficiency value abruptly rises to 0.7—0.8 within the 1/10 of the epoch length, which is about 300 ticks. Once critical point is reached, the value is slowly decreases roughly corresponding to the number of chunks transferred per tick.

Conclusions

The reliable estimation method of the peer-to-peer swarm efficiency is proposed based on the observed behavior of the imitational model. The method takes into account the individual distribution of the data chunks across mesh and can be used in any p2p-based networked solution.