

UDC 004.75, 004.724.2

G. V. Poryev

National Technical University of Ukraine «Kyiv Polytechnic Institute»
Peremohy ave., 37, 02056 Kiev, Ukraine

Efficiency of the Overlay Network Clustering based on the Locality Metric

The method of estimating the efficiency of the overlay network clustering in the swarms of peer-to-peer networks is proposed. The brief overview of the locality metric as a topological distance substitute is given and using regional Internet registries as information sources for constructing the structural model of the network is substantiated.

Key words: *Internet, distributed networks, peer-to-peer networks, locality, clustering.*

Introduction

Since the emergence of the Internet as a truly global network publicly available for consumers and businesses, the number of Internet nodes has increased drastically. The initial phase of the commercial Internet development was marked mostly by spontaneous growth of its regional sub-networks. Service providers tended to build network linkage out of technical availability, investment amounts, marketing strategy, peering cost and many other factors none of which was related to the desired optimal structure of the regional Internet segments, let alone any scientific ground for such optimality.

Nowadays, the practice of designing the national segments in the technologically developed countries consists, in particular, in facilitating so called Internet Exchange Points (IX) [1]. To build an IX, several major service providers will dedicate a single, territorially localized network operations center with routers and auxiliary equipment. The network of each service provider will then connect to such a point. Other local providers may then also participate in the system, thereby reducing the maintenance cost of the links due to the aggregation of member networks in such a way that the traffic from one member network to another is only transferred through the IX [1].

The resulting difference in guaranteed bandwidth depending on whether or not the destination node belongs to the same IX can easily reach two orders of magnitude. Relying on the cost difference, the end-users could also use personal intelligent filters to minimize the «external» traffic thereby reducing their expenses should they depend on the amount of traffic consumed.

© G. V. Poryev

The beginning of 2010s seems to appeal to the ever-growing consumer demand for large amounts of traffic, for file-sharing applications and multimedia streaming in particular, which are commonly held to constitute the majority of Internet traffic today. As a result, Internet service providers offer more «unlimited» subscription plans, which are to guarantee (and to bill for) specific bandwidth rather than amount of traffic consumed. Despite the bandwidth being generally lower in such plans, the user may schedule his operations as to consume more traffic volume than previously.

Today's publications related to the peer-to-peer networks show little concern for the optimal overlay structures in terms of throughput and latency within the overlay and lowering inter-domain traffic amounts. Those who show, however, are rather inclined to dedicated observational infrastructures or otherwise actively intervening methods which are subject to numerous drawbacks of filtering and traffic policing.

For this reason, the problem of increasing the efficiency of the distributed networks on the existing transport infrastructure is of great importance.

Currently used approaches to this problem, especially those using traditional locality metrics, exhibit some drawbacks. Traditional locality metrics, such as ping response time and trace-route count are influenced by the link load and conditions; the dedicated overlay control and maintenance infrastructures are subject to administrative shortcomings, etc.

In this paper we consider the various ways to determine the extent of the overlay network clustering, which may serve as a good estimation of the efficiency of the underlying peer-to-peer network.

Using the CARMA metric for locality estimation

In our previous paper on this subject [2] we have pointed out several crucial considerations that are to be taken into account prior to deriving any efficiency estimation method.

First, the majority of the distributed network users in general and peer-to-peer networks in particular are concentrated behind the «last mile», so it is reasonable to assume that the differentiation of their links by the bandwidth speeds is not as much crucial as it is for backbone networks.

Second, network based applications often use global knowledge of all network nodes and distances between these nodes. This information is usually managed by a central instance or may be derived from external infrastructures.

Consequently, if there is no central instance or external infrastructures in the network, clients usually apply either ping or hop count methods to estimate node distances. Hereby the problem is, that even if the ping or hop count methods would provide reasonable and reliable results, there is no way to apply these methods to a pair of foreign (i.e. both of them are not belonging to the local node) IP addresses. If clients are interested in this metric, they explicitly have to request this information from the corresponding nodes, which causes a significant communication overhead.

Generally speaking, the construction of complex structures requires either additional communication between nodes (in decentralized P2P systems), or is not scalable due to the existence of a Single Point of Failure (in P2P systems with a central instance) managing relevant information, which is a big drawback in global scale networks. In

order to sidestep these drawbacks, in our previous works [3,4] we have employed decentralized P2P systems and have proposed the CARMA (Combined Affinity Reconnaissance Metric Architecture) model and metric which is calculated locally on each node. This metric is calculated given the remote IP address of the peer and that all information can be implicitly derived from it.

It should be noted that although the information stored in Regional Internet Registries (RIR) databases may to some degree resemble topological junctions available from BGP, its purpose never was to maintain the real-time track of the actual precise Internet topology. The CARMA model is designed to estimate (not measure or calculate) the relative locality class, therefore it permits some tolerance towards latency or inconsistencies of its sources, namely RIRs.

If an application has access to the BGP information which reflects the present link status it is possible to estimate distance metric with high precision, however, the BGP is only used for the multi-homed routers in the backbone networks and in the geographically distributed enterprise networks. Hence for the task of estimating the topological locality we can use the database files published by the RIRs, which contain all the necessary data concerning the IP address ranges, autonomous systems definitions and their relation.

It should be noted that the definitions in the RIR databases are administrative rather than purely technical, and their public availability is the consequence of the openness and transparency policy of the Internet governing structures.

In our previous work [2] we have proposed to use the RIR database as the source of information to build a topological structural model.

When all database files are processed, the resulting incomplete graph reflects the Internet segment topology as close as it could be done without having access to BGP information. It is not necessary to devise any graph-walking algorithm to calculate the locality class, because the purpose of CARMA is to estimate the affinity of two given nodes, not calculating an exact hop count.

Deriving the efficiency rate

One of the motivating forces for the emergence of the peer-to-peer networks in general, and file-sharing networks in particular, was the need to relieve load on the publishing node from the overwhelming numbers of requests for the same content particle. This is the principal drawback of the «client-server» architecture which shows itself in the scenarios when the equipment of the client nodes and server nodes do not differ significantly, especially concerning the ability to establish direct links among themselves.

Up until recently, the prevailing practice of relieving the aforementioned stress was the deployment of the so-called CDN (content delivery networks), naturally distributed regionally and topologically over the target audience area.

In the modern times, however, the peer-to-peer technologies are gaining more widespread usage as content delivery means which do not demand corporate investments because of the transferring the communication costs to the customers themselves.

Let us consider the overlay network of a single swarm concerning the parameters important for the estimation of the integral efficiency of the network.

Let $i \in [0, C]$ be the index and total number of content particles in the current swarm, and $n \in [0, N]$ is the index and total number of the nodes in the swarm. Then the binary function used to indicate the presence of the i -th particle of the content \bar{C} on the n -th node may be expressed as:

$$p(n, i) = \begin{cases} 1: C_i \in \bar{C}, \\ 0: C_i = \emptyset. \end{cases} \quad (1)$$

When the initial seeding has started from the 0th node, the situation is:

$$p(0, i) \equiv 1, \quad \forall n > 0: p(n, i) \equiv 0. \quad (2)$$

In the process of seeding the particles are transferred from the seeding node to the other nodes.

Now we can formulate the principle of designing the efficiency rate for the overlay structure, given the particle distribution across the nodes. In the worst case (equal to the «client-server» case), the channel of the seeder node is used to transfer repeatedly the same particle over and over on all other nodes simultaneously. In the best case, every content particle is uploaded from the seeder node to the other node, different for each particle.

As an intermediate indicator Q_i we shall take the number of nodes which has the i -th particle:

$$Q_i = \sum_{n=0}^N p(n, i) k_n. \quad (3)$$

Where k is the «behavioral balancing» setting specific for each node and determining how upload traffic is given priority over download traffic.

Obviously, in the aforementioned worst case, unlike the best case, the Q_i values will have greater dispersion, which, in turn, may be estimated as follows:

$$E = \sqrt{\frac{\sum_{i=0}^C \left(\frac{\sum_{l=0}^C Q_l}{C} - Q_i \right)^2}{C(C-1)}}. \quad (4)$$

It should be noted that the efficiency of the overlay structure of the peer-to-peer network is determined not only and not mostly by the efficiency of the seeder node's channel utilization, because the initial conditions important for it do matter only for the period required to transfer the first content particles.

Subsequently, the individual data exchange speeds between swarm nodes begin to play an important role. Now the swarm clustering affects the exchange speeds more than anything else, besides the bandwidth limitation of the hardware or the subscription plan.

Generally, if the swarm nodes do not use any sorting algorithm to determine the peer priority in the queue by any characteristic, it is safe to assume that the locality class distribution among their queue will be roughly equal or at least the same as for the whole swarm.

However, in the case where the locality metric is used for automated clustering of the overlay network, we can estimate the level of such clustering using the following expressions.

Let binary function

$$D(n_1, n_2) = \begin{cases} 1: n_1 \leftrightarrow n_2 \\ 0: n_1 \nleftrightarrow n_2 \vee n_1 = n_2 \end{cases} \quad (5)$$

denote the presence of the direct connection between nodes n_1 and n_2 . Then the binary function

$$L(n_1, n_2, f) = \begin{cases} 1: D(n_1, n_2) \neq 0 \wedge n_1 \cup n_2 \in \bar{L}_f \\ 0: D(n_1, n_2) = 0 \end{cases} \quad (6)$$

can be used to determine the presence of the nodes n_1 and n_2 to the locality class f . For each node the individual locality class distribution among directly connected peer lists will be expressed as

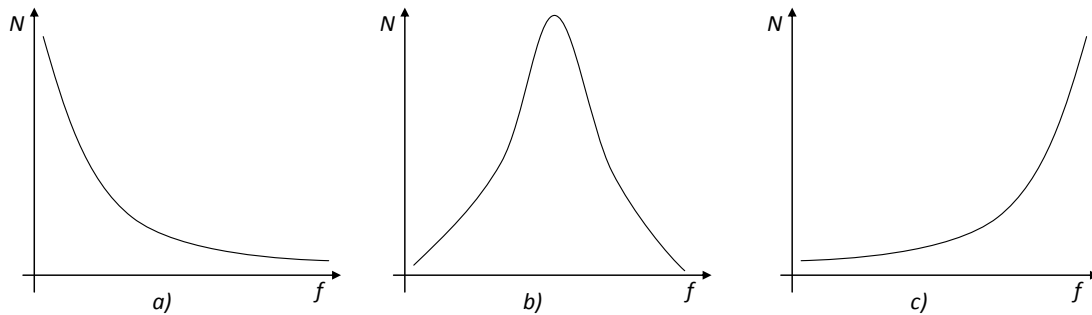
$$L'(n, f) = \sum_{i=0}^N L(n, i, f). \quad (7)$$

Whereas the overall locality class distribution across the whole swarm is

$$E'(f) = \frac{\sum_{n=0}^N L'(n, f)}{N}. \quad (8)$$

The behavior of the $E'(f)$ function determines the efficiency of the overlay network clustering in the following way (see figure).

Using the equations (4) and (8) the software operating on the nodes of the peer-to-peer network can independently estimate the efficiency of its overlay structure and decide on establishing new and breaking existing connections.



Typical locality class distribution across the swarm: a) high level of clustering (wanted);
b) average level of clustering (neutral); c) low level of clustering (unwanted)

Conclusions

The software and hardware developers working on the peer-to-peer solutions may take advantage of having control over the behavior of the network overlay structures in the real-time, thereby increasing their overall efficiency by managing response times and network throughputs. The leverage mechanism proposed as such control is the overlay network clustering of the swarms in peer-to-peer networks.

The clustering is based on the previously proposed topological locality estimation system with the pre-computed structural model of the national Internet segments, which will allow for automatic adaptive clustering of the overlay network using the peer list reordering based on the locality class. Unlike the traditional metric, proposed method at its first implemented layer does not use service traffic to directly measure topological distance and therefore can be used to estimate the locality even for non-local address pairs.

1. *Фурашев В.* Параметры украинского сегмента Internet как сложной сети / В. Фурашев, В. Зубок, Д. Ланде // Открытые информационные и компьютерные технологии: сб. науч. тр. — 2008. — Т. 40. — С. 235–242.

2. *Poryev G.* Using the Internet Registries to Construct Structural Model for Locality Estimation in the Overlay Networks / G. Poryev // Реєстрація, зберігання і оброб. даних. — 2011. — Т. 13, № 1. — С. 78–86.

3. *Poryev G.* Multi-Tier Locality Awareness in Distributed Networks // Інформаційні технології та комп'ютерна інженерія. — 2009. — № 3(16). — С. 13–17.

4. *Poryev G.* CARMA Based MST Approximation for Multicast Provision in P2P Networks / G. Poryev, H. Schloss, R. Oechsle // In Proceedings of the Sixth International Conference on Networking and Services (ICNS 2010) / IEEE. — Cancun, Mexico. — 2010. — P. 123–128.

Received 29.05.2011