RankSlicing: A decentralized protocol for supernode selection

Giovanni Simoni

Roberto Roverso

Alberto Montresor

ABSTRACT

In peer-to-peer applications deployed on the Internet, it is common to assign greater responsibility to supernodes, which are usually peers with high computational power, large amount of memory, or high network bandwidth capacity. In this paper, we describe a practical solution to the problem of supernode selection, that is the process of discovering the best peers in the network by some application-specific metric. We provide a distributed heuristic that allows to identify the best K nodes in the P2P overlay, by taking into consideration the realities of actual deployments, such as the presence of NATs. Our approach consists of an epidemic protocol which does not require new connections to be established, but rather relies on established connections, such as the ones provided by a NAT-resilient peer sampling framework. We support our claims with a thorough evaluation of our solution in simulation and in a real deployment on thousands of consumer machines.

CONTACT INFORMATION

Giovanni Simoni

giovanni.simoni@hivestreaming.com

Peerialism AB
Box 5207
10245 Stockholm
Sweden

Super-nodes are peers characterized by better performances, according to some application-specific metrics:
- CPU power
- Network bandwidth
- Available memory ...

Goal: selection of *k* nodes to promote as super-nodes.

Tackled requirements:

1. Stability of the result: avoid changes in the supernodes, unless better nodes join the network, supernodes change their characteristics or supernodes leave the network
2. Live computation: the computed supernodes set must be available all nodes as they are produced.
3. Live evaluation system: a quality measure on the computed result must be available in order to estimate the supernodes set stability.

Use cases

- Content replication on nodes having a good amount of memory
- Nodes acting as connectivity helpers
- Pre-fetching in live video streaming

The Algorithm

- Node capabilities: $cap_i = \langle x_i, y_i, z_i \rangle$
- Gossip descriptors: $\langle i, lc, age, cap_i \rangle$
- The View is a local approximationof supernodes set:
  - Room for *k* descriptors;
  - Send a sample of the view to random neighbour;
  - Emit and send a new local descriptor.
- Merging inbound sample with the local view:
  1. Remove duplicates
     - At most one descriptor per node;
     - Keep the freshest one (according to *lc*)
  2. Filter outdated nodes: *age* ≥ *PAL*;
  3. New view V': Truncate to best k elements (comparison over *cap* field of descriptors).

Live Evaluation

Ideal quality measure:

$$q = \frac{1}{|\Pi|} \sum_{p_i \in \Pi} \frac{|V_i \cap L|}{k}$$

Approximated quality measure:

$$q_{i,0} = \frac{|V_i \cap V_i'|}{k}$$

$$q_{i,1}^{(n)} = \alpha \cdot q_{i,1}^{(n-1)} + (1-\alpha) \cdot q_{i,0}^{(n)}$$

Parameter: $\alpha \in [0, 1]$

Evaluation and Deployment