A Heuristic for IP Summarization in perfSONAR Performance Monitoring Architecture

Marcos Portnoi, *Member, IEEE*
Department of Computer and Information Sciences
University of Delaware
mportnoi@udel.edu

Martin Swany
Department of Computer and Information Sciences
University of Delaware
swany@cis.udel.edu

ABSTRACT

A distributed, hierarchical distributed service for computer networks might rely in several instances, located in different hierarchical layers. A distributed directory service, for example, might be comprised of upper level listings, and local directories. The upper level listings contain a compact version of the local directories. Clients desiring to access the information contained in local directories might first access the high level listings, in order to locate the appropriate local instance. One of the keys for the competent operation of such service is the ability to properly summarize the information, which will be maintained in the upper level directories. We analyze the case of the Lookup Service the perfSONAR performance monitoring architecture, which implements IPv4 summarization in its functions. We propose an experimental method, or heuristic, to achieve the summarizations, based on the PATRICIA tree. We further apply the heuristic on test sets and contemplate the results.

Categories and Subject Descriptors
E.1 [**Data Structures**]: Graphs and Networks.

General Terms
Algorithms, Experimentation, Management, Performance.

Keywords
IP, summarization, tree, patricia tree, trie, subnet.

1. INTRODUCTION

Certain distributed computer network services work by dispersing resources and data among several instances of the service, in a hierarchical manner. One such service can be a hierarchical directory service, where lower level instances maintain local data and publish the data to higher-level instances. The higher level instances are responsible to keep a compact listing of all the data administrated by lower level instances, and must be able to indicate which particular instance maintains specific data. The higher level instances, therefore, hold a summary of all the lower level counterparts.

For this mechanism to operate efficiently, the lower level instances must summarize their data and publish this summarization to the upper level layer. The kind of data being summarized might accommodate different techniques for this procedure, some of them based on heuristics. One such example is the Lookup Service in perfSONAR architecture [4], which resembles a distributed directory. Its lower level instances summarize the controlled data and forward the compacted version to the upper layer. One kind of published data is IPv4 addresses.

This paper describes the work done in producing a heuristic to generate IP addresses summarizations in the realm of perfSONAR architecture.

2. OUR WORK

We demonstrate, in the research, a heuristic applied to construct a summarized set from an original set of IPv4 addresses. This original set can be host addresses.

As motivators, we outline the needs of perfSONAR, a performance monitoring architecture. This architecture includes a distributed directory service, where other services register to and clients perform queries to find other services and performance data. This directory service resembles a DNS system. To operate adequately, the service relies on lower level instances, which publish a summary of their controlled information to the upper level instance layer. By employing summarization, the service administers the volume of information that is published into the network. Moreover, resources, such as memory and storage, can be constrained.

We further describe the mechanics of IP summarization and the techniques utilized by our proposed heuristic to obtain a final summarization. Essentially, our heuristic assembles a special data structure, called PATRICIA tree; then, it selects nodes from this tree based on specific metrics and configured thresholds. Namely, we use *distance, density*, and *minimum subnet mask*. This selection follows particular interpretations of appropriate summarizations.

The heuristic was applied to test sets comprised of IP addresses collected from real networks that adopt perfSONAR architecture. The results illustrate compressions rates below 30% (the smaller, the more compressed) and the capability of adjusting the compression rate by means of a *granularity* setting. It was also verified that the heuristic performance is evidently dependent on the arrangement of IP addresses in the original data set.

3. REFERENCES

FREDKIN, EDWARD. "Trie Memory", Communications of the ACM, vol. 3, issue 9, p. 490-499, Sep 1960.

HANEMANN, A. *et al*. "PerfSONAR: A Service Oriented Architecture for Multi–Domain Network Monitoring", In

"Proceedings of the Third International Conference on Service Oriented Computing", Springer Verlag, LNCS 3826, pp. 241–254, ACM Sigsoft and Sigweb, Amsterdam, The Netherlands, December, 2005.

INFORMATION SERVICES USE CASES: Finding Closest MP. https://spaces.internet2.edu/display/ISWG/Use+Cases

IP SUMMARIZATION. Available at: http://code.google.com/p/perfsonar-ps/wiki/IPSummarization

KNUTH, D. E. "The Art of Computer Programming", vol. 3: Sorting and Searching. Addison-Wesley, 1973.

MORRISON, DONALD R. "PATRICIA - Practical Algorithm to Retrieve Information Coded in Alphanumeric", Journal of the ACM, 15(4), p. 514-534, Oct 1968.

PERFSONAR. Available at: http://www.perfsonar.net

SOAP version 1.2 specification: http://www.w3.org/TR/soap12.

SWANY, MARTIN. "An Extensible Schema for Network Measurement and Performance Data". http://forge.gridforum.org/sf/go/doc15649?nav=1.

TIERNEY, BRIAN *et al*. "Instantiating a Global Network Measurement Framework", LBNL Technical Report LBNL-1452E, January 2009. http://acs.lbl.gov/~tierney/papers/perfsonar-LBNL-report.pdf.

ZURAWSKI, Jason *et al*. "Hierarchically Federated Registration and Lookup within ehe Perfsonar Framework". In 10th IFIP/IEEE International Symposium on Integrated Network Management, 2007 (IM '07), p. 705-708, May 2007.