

# Toward a Scalable and Collaborative Network Monitoring Overlay

Vasco Castro, Paulo Carvalho, and Solange Rito Lima

University of Minho, Department of Informatics, 4710-057 Braga, Portugal  
e-mail: {pmc,solange}@di.uminho.pt

**Abstract.** This paper presents ongoing work toward the definition of a new network monitoring model which resorts to a cooperative interaction among measurement entities to monitor the quality of network services. Exploring (i) the definition of representative measurement points to form a network monitoring overlay; (ii) the removal of measurement redundancy through composition of metrics; and (iii) a simple active measurement methodology, the proposed model aims to contribute to a scalable, robust and reliable end-to-end monitoring. Besides the model proposal, a JAVA prototype was implemented to test the conceptual model and its design goals.

## 1 Introduction

Monitoring of large networks raises multiple challenges regarding scalability, robustness and reliability of measurements. It is known that monitoring systems where a single point is responsible for gathering and processing measurements obtained throughout the network suffer from severe scalability and robustness limitations. To address this problem, distributed solutions where monitoring data is collected and processed at each measurement point (MP) have been proposed. For instance, solutions based on active edge-to-edge measurements provide a straightforward way of measuring service quality, however, the potential interference of cross probing among boundary nodes on network behaviour needs to be carefully considered.

To reduce network overhead and improve spatial coverage, it is important to identify the most representative and critical network points in order to obtain an overall view of the network status involving only a subset of MPs. Resorting to composition of metrics between these MPs, i.e. through concatenation of partial metrics, the interference on network operation can be reduced, avoiding redundant measurements in overlapping links. The composition of metrics also allows observing trends, being more informative as a result of the underlying metric partitioning scheme.

In this context, this paper proposes a collaborative network monitoring overlay which resorts to the cooperation between representative MPs strategically located in the network to compute performance and quality metrics both intra-area and end-to-end. The aim is to pursue a flexible, scalable and accurate monitoring overlay solution that simplifies and systematises the cumulative computation of metrics by involving a subset of network nodes.

This paper is organised as follows: related work is discussed in Section 2, the proposed monitoring model and its components are described in Section 3, the model prototype is presented in Section 4 and the conclusions are summarised in Section 5.

## 2 Related Work

Active monitoring carried out on an edge-to-edge basis, i.e., between network boundaries, is particularly suitable for monitoring network performance and quality of service (QoS) [1]. This approach improves scalability by involving only edge nodes in the monitoring process, removing the complexity of monitoring tasks from the network core. Considering that in edge-to-edge probing, probes from distinct pair of edges may cross the

same links, hop-to-hop monitoring strategies try to avoid repeating probes in those links. However, capturing network behaviour combining hop-by-hop measures is not an efficient and easy solution as it involves : (i) a high-degree of metrics' concatenation; (ii) monitoring agents in all network nodes; and (iii) additional traffic in the network for reporting metrics to management stations. To reduce the amount of data exchanged between management stations and MPs, several solutions have been pointed out, namely the use of flow aggregation [2], statistical summarisation [3] and network thresholds crossing alerts [4].

Tomography concepts [5] continue to deserve significant attention for estimating distinct aspects of network behaviour, including QoS [6, 7]. In [8], network tomography is applied to the definition of a monitoring overlay to infer packet loss in all network nodes.

Taking in consideration the mentioned strategies, this study proposes a network monitoring overlay solution which resorts to representative MPs to compute performance and quality metrics both intra- and inter-area, with reduced overhead.

### 3 A Collaborative Monitoring Overlay

The proposed model relies on a collaborative participation of representative MPs acting as peers, each one contributing with a disjoint measure component to the evaluation of the global measure. Thus, end-to-end measurements are obtained through the aggregation of metrics calculated in each of the network areas involved.

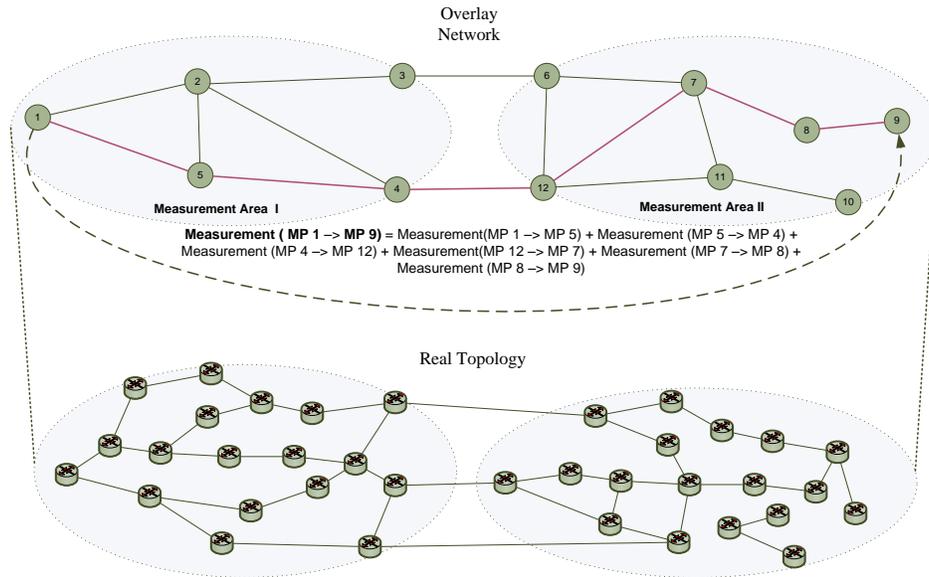


Fig. 1. Example of measurement between different administrative areas

Figure 1 illustrates the monitoring overlay network and the underlying physical topology. The overlay network consists of representative MPs and these are the only players taking part in the measurement process. Each MP in the overlay is expected to store the measurements to its neighbouring MPs. Thus, measurement data is distributed and stored throughout the overlay network. Based on a monitoring request, each MP in the measurement path provides the required measures for aggregation in order to calculate a set of metrics between any specified MPs. Distributing measurement data over several MPs also enables a rapid recovery of the measurement process by bringing alternative MPs in the pro-

cess of rebuilding the measurement path in case of routing or network topology changes. Note that these changes do not necessarily imply a change in the overlay topology.

The proposed model allows measurements at two levels: Intra-area and Inter-area. *Intra-area* measurements are carried out on a regular time basis to ensure that MPs in the same area have a clear view of network status and quality of service. A MP may, at any-time, send or exchange measurement data between itself and any other MP within its area. Thus, by retrieving data from multiple MPs in the area and using composition of metrics, it is possible to calculate the value of a metric for a given measurement path. *Inter-area* measurement is performed through the composition of the metrics resulting from intra-area measurements. Conversely to intra-area operation, this type of measurement does not need to be performed continuously, but on request. This process can be triggered, for example, by an application signalling process to assess the communication path before establishing an end-to-end session crossing different network areas.

**Model operation** - Initially, a monitoring entity sends a message to the initial MP indicating that it needs to obtain a set of metrics between a pair of MPs (see example in Figure 1). This MP, after receiving the request, sends it in the overlay network as a packet measurement request. Each MP in the overlay path will intercept this packet and attach measurement data between itself and the upstream MP, before sending it to the downstream MP. This process is repeated until the destination is reached, i.e., each MP will successively attach its measurement data along the overlay. The final MP or the destination, upon receiving the packet measurement request, will proceed similarly, sending subsequently the resulting message back to the initial MP with all collected measurements. At this point, the initial MP is able to compose the required metrics in order to obtain end-to-end measures. This operation can assume distinct cumulative functions (additive, multiplicative, max-min, etc.) depending on the nature of the metric being evaluated.

In practice, this process can be considerably simplified as area border MPs (e.g. MP 12 in Figure 1) may already have up-to-date measurements from the remaining measurement path. This allows an immediate reply from that MP to the measurement requester, reducing measuring latency significantly. This process can be further improved through proper proactive metrics dissemination among inter-area MPs.

One challenge of the present model is to identify the representative MPs. Although several works target this topic [8, 9], this aspect requires further study. Other relevant aspect currently under study is focused on solutions to avoid measurement data fragmentation.

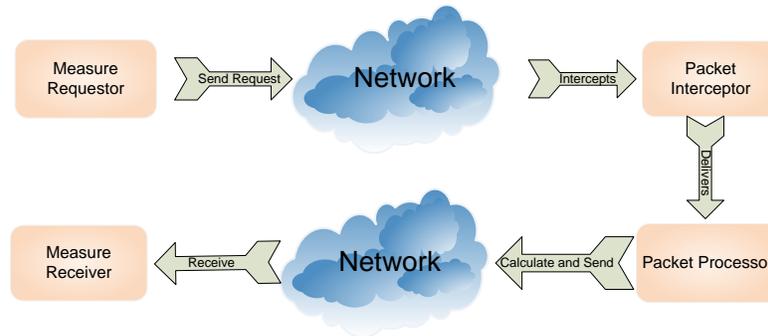
## 4 The Implemented Prototype

The model prototype was implemented in Java and MySQL, being the measurement primitives structured in XML. The implementation includes four main components: (i) the “Measure Requester”; (ii) the “Packet Interceptor”; (iii) the “Measure Processor”; and (iv) the “Measure Receiver”, interacting as illustrated in Figure 2.

**Measure Requester** - This component is responsible for initiating the measurement process between two MPs. In the developed prototype, this is a command line application that receives as parameters, the source and destination MP, and the set of metrics to measure.

**Packet Interceptor** - This component is responsible for capturing measurement packets. These packets are differentiated in the network through the use of `router alert` option within IPv4 header, avoiding packet processing at upper protocol layers. In a Linux router, this can be accomplished resorting to proper `iptables` packet filtering (it requires the extension `xtables-addons`). Captured packets are taken from kernel to user space (through `libnetfilter_queue`) for processing at MPs. The use of `router alert` option avoids the use of explicit MP addressing, allowing for a more flexible overlay topology definition.

**Packet Processor** - This component is responsible for processing and concatenating measurement data. Once a packet request is intercepted at an MP, this component detects the new request, validates it and appends the required metrics to the measurement packet. This



**Fig. 2.** Interaction of Model Components

process involves identifying the latter upstream MP before adding its measurement contribution. Then, the component builds an IP packet setting the `router alert` option, updates the data payload accordingly and sends the packet to the downstream MP. Once the last MP is reached, the "Packet Processor" opens a TCP connection to the initial MP for sending the aggregate measurement outcome.

**Measure Receiver** - When the measurement process starts, a measurement packet request is issued and, simultaneously, the request is stored in a database, remaining in listening mode on an UDP port. Upon receiving the corresponding measurement result, this component updates the database for the corresponding request.

## 5 Conclusions

This paper has presented ongoing work toward the definition of a network monitoring overlay which resorts to a cooperative interaction among representative MPs to monitor the quality of network services. In the proposed model, measurement overhead and redundancy are reduced through the composition of metrics from non-overlapping measurement paths, both intra- and inter-area. This aspect along with the ability to accommodate network topology changes aim to contribute to a scalable and flexible end-to-end monitoring solution. A JAVA prototype has also been implemented to test the conceptual design goals of the model, being currently under evaluation in a virtualised network environment.

## References

1. Habib, A., Khan, M., Bhargava, B.: Edge-to-edge measurement-based distributed network monitoring. *Computer Networks* **44** (2004) 211–233
2. Lin, Y.J., Chan, M.C.: A scalable monitoring approach based on aggregation and refinement. *IEEE JSAC* **20** (2002)
3. Asgari, A.H., Egan, R., Trimintzios, P., Pavlou, G.: Scalable monitoring support for resource management and service assurance. *IEEE Network* **18**(6) (2004) 6–18
4. Wuhib, F., Stadler, R., Clemm, A.: Decentralized service-level monitoring using network threshold crossing alerts. *Communications Magazine, IEEE* **44**(10) (2006) 70–76
5. Vardi, Y.: Network Tomography: Estimating Source-Destination Traffic Intensities from Link Data. *Journal of the American Statistical Association* **91**(433) (1996) 365–377
6. Gu, Y., Jiang, G., Singh, V., Zhang, Y.: Optimal probing for unicast network delay tomography. In: *INFOCOM'10*, Piscataway, NJ, USA, IEEE Press (2010) 1244–1252
7. Arya, V., Duffield, N., Veitch, D.: Temporal Delay Tomography. In: *INFOCOM*. (2008) 276–280
8. Chen, Y., Bindel, D., Katz, Y.H.: Tomography-based Overlay Network Monitoring. In: *ACM SIGCOMM Internet Measurement Conference (IMC)*, ACM Press (2003) 216–231
9. Ratnasamy, S., Handley, M., Karp, R.M., Shenker, S.: Topologically-Aware Overlay Construction and Server Selection. In: *INFOCOM*. (2002)