# Privacy Preserving Distributed Network Outage Monitoring

Mentari Djatmiko*†, Dominik Schatzmann‡, Arik Friedman*, Xenofontas Dimitropoulos‡ and Roksana Boreli*†
* NICTA, Sydney, Australia, {mentari.djatmiko, arik.friedman, roksana.boreli}@nicta.com.au
† UNSW, Sydney, Australia
‡ ETH Zurich, Zurich, Switzerland, {dominik.schatzmann, fontas}@tik.ee.ethz.ch

*Abstract*—**Troubleshooting network outages is a complex and time-consuming process. Network administrators are typically overwhelmed with large volumes of monitoring data, like NetFlow data, and are often "left alone", fighting problems with very basic debugging tools, like *ping* and *traceroute*. Distributed network traffic monitoring and intelligent correlation of data from different Internet locations are highly valuable for analysing the root cause of network outages. However, correlating measurements across domains is presently largely avoided due to privacy concerns. A possible solution to this problem is secure multi-party computation (MPC). In this work, we propose a distributed mechanism based on MPC for privacy-preserving correlation of traffic measurements from multiple networks, towards network outage diagnosis. We first outline an MPC protocol that can be used to analyse the scope (local, global, or semi-global) and impact of network outages across multiple domains. Then, we use NetFlow data from a medium-sized ISP to evaluate the performance of our protocol. Our preliminary findings indicate that correlating data from several dozens of parties is feasible in real-time, with a delay of just a few seconds. This underlines the scalability, and potential for real-world deployment of our scheme. Finally, we apply our scheme to a known connectivity issue involving a large European Internet Exchange Point (IXP) and demonstrate that our approach enables to easily distinguish between local, global, and semi-global outages. In our study, 81.54% of the 3,408 reported outages were local, and 18.46% affected between 2 and 5 organizations.**

Internet outages caused by fiber cuts, power failures, routing problems, or prefix hijacking attacks, may disrupt critical services, and hence it is essential to troubleshoot them in a timely manner. However, troubleshooting outages is challenging due to the difficulty in determining their root cause. An important part of root cause analysis is diagnosing the scope and impact of an outage. Identifying, for example, if an outage is a local or global problem is typically the very first step of troubleshooting. However, answering even this simple question is non-trivial with existing network monitoring tools. In this context, correlating data from multiple domains is very useful.

Recent work has identified passive one-way traffic measurements as a new and promising approach for monitoring network outages [1], [2]. The key idea is that observing a large number of unsuccessful (one-way) connections and the concurrent absence or decline of successful (two-way) connections to a specific remote destination (IP address, prefix, or a collection of prefixes, i.e., autonomous system (AS)) indicates a network outage. Intuitively, one-way traffic monitoring enables the operators to passively observe and use the regular outgoing traffic of a network as probe traffic for outage detection. One-way traffic monitoring enables prioritization of outages based on their impact, i.e., the number of IP addresses that were affected by an event, and detects only the outages that actually affect the clients of a monitored network [1], [2].

Correlating one-way traffic measurements from multiple domains allows determining whether an outage is a local or global problem, i.e., the scope of the outage. A local outage is only detected by a single domain, while a global outage is detected by multiple domains. The scope of an outage is essential for troubleshooting since it helps to narrow down the potential cause of the outage. Moreover, to facilitate troubleshooting, it is also be beneficial to obtain information about the number of distinct domains that cannot reach a specific destination and the total number of hosts that were affected by the outage.

However, such correlation is not presently possible. Internet service providers are competitive organizations that protect the commercially sensitive operational information about the performance of their network and generally do not disclose it. In addition, distributed monitoring of network outages based on passive network traffic measurements of one-way traffic requires exchanging information about contacted services, IP addresses, prefixes, and ASs. This information is generally sensitive and its exchange is often prohibited by privacy laws.

Secure multi-party computation (MPC) [3] provides a possible solution to this problem. MPC is a cryptographic approach that enables multiple entities to jointly compute a mathematical function where the input data are provided by a number of input peers. MPC is suitable for the problem as it provides a formal guarantee on the privacy of the input and intermediate data, and the correctness of the computation result. Until recently, the research domain of MPC was almost exclusively of theoretical interest, with very limited real-world use, due to the high computational overhead. This has changed in the last few years, as efficient MPC protocols and frameworks, like the SEPIA MPC library [4], provide reasonable running times that are feasible for real-world applications.

In this work, **we show that we can improve troubleshooting processes** with additional information about the scope and impact of outages **using efficient MPC protocols**. We first **introduce an MPC mechanism to correlate one-way traffic measurements** about outages from multiple domains. The goal of our scheme is to provide additional information

978-1-4799-0056-5/13/$31.00 ©2013 IEEE

| Input peers | Bloom filter size | |
|---|---|---|
| | 3890 | 17290 |
| 10 | 0.236 sec | 0.704 sec |
| 30 | 0.210 sec | 0.832 sec |
| 100 | 0.409 sec | 1.523 sec |

TABLE I
ESTIMATED COMPUTATIONAL RUNTIME.

about the number of domains and the total number of hosts that cannot reach a specific destination.

The input to our computation is the output of the FACT outage monitoring system [1]. FACT identifies remote destinations, i.e., IP addresses, IP prefixes or ASs, that are unreachable from local clients. To compute the number of domains that cannot reach a specific destination, each input peer prepares an input multiset where the elements of the set represent remote entities, e.g., IP addresses, IP prefixes or ASs, that are unreachable from an input peer. Then we perform a multiset union operation to find aggregate multisets. The number of occurrences of an element in an aggregate multiset gives the number of input peers that cannot reach a specific destination. We also use multiset union to compute the total number of hosts that cannot reach a specific destination. In this case, an input peer supplies a multiset that represents the number of local hosts that cannot reach a destination. The number of occurrences of an element in the aggregate multiset gives the total number of hosts that cannot reach a specific destination.

We use the SEPIA library to realize our scheme. SEPIA provides an efficient multiset union operation using counting Bloom filters [5]. A counting Bloom filter is a space-efficient probabilistic data structure which uses multiple hash functions to map a multiset into a fixed size array of counters. Essentially, a counting Bloom filter transforms multiset union into a fixed size array summation. The multiset union operation is very efficient [5] since it only requires an *addition* operation which in MPC has low computational overhead.

Second, **we evaluate the feasibility of this approach** by estimating the computational runtime of the proposed scheme using input data from a real network. In more detail, we focus on the worst case scenario of aggregating data about unreachable hosts. Aggregating data about IP prefixes or ASs results in a much smaller input to the MPC computation. Based on the collected data (as described in the following paragraph), the maximum size of the input is 1,729 unreachable hosts, with an average of 597.52 and a standard deviation of 274.76. The MPC computation of the multiset union protocol of SEPIA is known to scale very well even for larger input sizes [5]. As illustrated in Table I, it requires less than 2 seconds to aggregate counting Bloom filters with more than 17,000 counters from 100 input peers, using three privacy peers. This finding suggests that the computational overhead for correlating information about outages will easily scale to dozens or even hundreds of input domains.

Third, we **demonstrate the utility of correlating data from multiple domains** to analyze outages by applying our scheme to a known incident. For this purpose, we use unsampled NetFlow data collected from the border routers of the SWITCH network [6]. SWITCH is a medium-size ISP which connects approximately 30 Swiss universities, government institutions, and research laboratories to the Internet. To simulate multiple domains we partitioned the data set into the six largest customers and analyzed those data sets independently with FACT. In more detail we study the data from 25th of March 2010, where after a scheduled maintenance by AMS-IX, traffic from the IXP toward SWITCH was partially blackholed as discussed in more detail in [1]. Figure 1 shows the impact of the outage on the individual domains measured by the number of BGP prefixes that were unreachable from 1 to 6 domains over time. Note that we only include BGP prefixes which are unreachable from at least 2 hosts in each domain. The graph shows that many remote destinations are unreachable from only one domain, which indicates that these events are likely to be local issues or false positives. Furthermore, we see that 18.46% affected between 2 and 5 organizations.

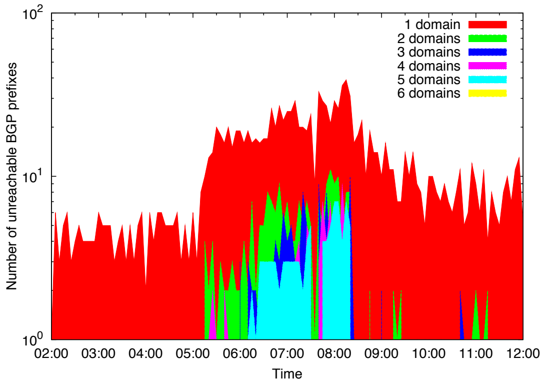

Fig. 1. Unreachable BGP prefixes detected by 1 to 6 domains.

In conclusion, we propose a mechanism to correlate network outages detected using flow-based measurements from multiple domains to facilitate troubleshooting. Our scheme is based on MPC to fulfil the privacy requirement. As future works, we will refine and analyse the error introduced by our mechanism.

## REFERENCES

[1] D. Schatzmann, S. Leinen, J. Kögel, and W. Mühlbauer, "FACT: flow-based approach for connectivity tracking," in *Proceedings of the 12th international conference on Passive and active measurement*, ser. PAM'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 214–223.

[2] E. Glatz and X. Dimitropoulos, "Classifying internet one-way traffic," in *Proceedings of the 2012 ACM conference on Internet measurement conference*, ser. IMC '12. New York, NY, USA: ACM, 2012, pp. 37–50.

[3] R. Cramer, I. Damgaard, and J. B. Nielsen, "Multiparty Computation, an Introduction," May 2008.

[4] M. Burkhart, M. Strasser, D. Many, and X. Dimitropoulos, "SEPIA: privacy-preserving aggregation of multi-domain network events and statistics," in *Proceedings of the 19th USENIX conference on Security*. Berkeley, CA, USA: USENIX Association, 2010, p. 15.

[5] D. Many, M. Burkhart, and X. Dimitropoulos, "Fast private set operations with sepia," ETHZ, Tech. Rep. 345, March 2012.

[6] The Swiss Education and Research Network (SWITCH). http://www.switch.ch.