A Call to Arms for Management Plane Analytics

Aditya Akella
Microsoft Research and UW-Madison

Ratul Mahajan
Microsoft Research

ABSTRACT

Over the last few decades, the networking community has developed numerous techniques for understanding *how* real networks behave through analyzing their data and control planes. In this paper, we call upon the community to similarly develop techniques to analyze the network management plane, that is, activities that underlie network design and operation. Such analytics can shed light on *why* a network behaves as observed and the relative merits of different management practices. While the management plane is often not directly observable, we argue that many relevant aspects can be inferred through data that most networks already gather (e.g., snapshots of configurations). Using preliminary analysis of such data from many large networks, we demonstrate the feasibility and the value of management plane analytics.

Categories and Subject Descriptors

C.2.3 [**Network Operations**]: Network management

Keywords

Management plane, network configuration, analytics

1. INTRODUCTION

The networking community does not need to be reminded about the value of measuring real networks. It has developed a suite of tools and creative techniques that discover how the network functions even when the network does not directly reveal that information. For instance, *traceroute*, a bread and butter tool, is a clever "hack" in which a basic router functionality—dropping packets with expired time-to-live (TTL) and sending an error message to the source—is exploited to discover the paths that packets take to their destination. Other examples include techniques to infer link characteristics [7], available bandwidth along a path [8], loss rate and re-ordering [11], network topology [18], and so on.

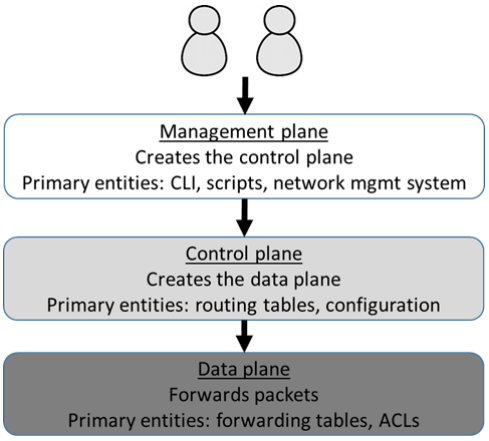

Figure 1: The three network planes

However, the vast majority of the measurement work thus far focuses on inferring aspects of the network's data plane (e.g., routing paths) or control plane (e.g., routing weights). Little work has gone into measuring and understanding the network's management plane. Figure 1 illustrates the three planes and their relationship. The data plane forwards packets. The control plane generates the data plane using configuration files, or control programs in the case of SDN, and control plane protocols such as OSPF and BGP. The management plane is composed of practices and protocols that generate the control plane based on the network's policies.

Despite the importance of the management plane to a well-functioning network, we have limited insight into how real networks are designed and operated today. Even basic facets are unknown. For instance, how heterogeneous are the networks in the hardware and software they use; how often configuration changes are made; what fraction of changes are done manually versus automatically; what types of changes are common (e.g., ACLs, routemaps, VLANs, etc.); or how many devices are impacted by a typical change. While it is well-known that network management is onerous and error-prone, beyond broad reasons like configuration languages

Hotnets '14, October 27–28, 2014, Los Angeles, CA, USA.

being low-level, it is not known which management practices in particular have a higher risk of causing failures.

In this paper, we call upon the research community to make a concerted, systematic effort to analyze the management plane of today's networks. The high-level goals of such *management plane analytics* are: i) infer a given network's management practices; ii) infer which practices lead to better operational health (e.g., fewer failures); and iii) develop a predictive model of a network's operational health, based on its management practices, to help with what-if analysis and improve management practices.

Management plane analytics is key to designing a better management plane, one that reduces the burden on operators and reduces failures. We are inspired by how research into software engineering practices, also called "empirical software engineering," has helped improve the quality of software and reduced the number of bugs [4]. We expect a similar positive impact from research into network management practices. Now is a particularly relevant time for this undertaking because, in the form of SDN, the community is engaged in re-architecting networks. A detailed understanding of the strengths and weaknesses of the current management plane will help inform the design of an SDN-based management plane as well.

A primary hurdle in uncovering management practices is that they are not directly logged by most networks. However, we posit that it is possible to infer them from other data that is already logged by most networks. In particular, this data is: 1) snapshots of device configurations that are taken and archived using popular tools such as RANCID [16] and HPNA [20]; and 2) logs of alerts and trouble tickets. This data is indirect and noisy, but useful information can be extracted from it. This particular challenge is similar to what is addressed by existing data- and control-plane measurement tools, and it offers a pragmatic alternative to waiting for perfect data sources to become available.

Many of the largest networks today have been in operation for years, if not decades, and have well established management processes. Given this fact, and that many of these networks are mission critical, conducting management plane analytics is both necessary and eminently feasible.

To demonstrate the value and feasibility of management plane analytics using existing data sources, we conduct a preliminary analysis across hundreds of data center networks. We show that network heterogeneity and how often the network is changed negatively contributes to its operational health, while the size of the network and the extent of automation appear to have minimal impact. We also show that more research is needed to develop predictive models of network operational health.

2. MANAGEMENT PLANE ANALYTICS

In this section, we first describe commonly available data sources that can be used for management plane analytics, and we then outline a simple framework for conducting this analysis systematically.

2.1 Data sources

We focus on the following two data sources that in our experience are commonly available today.

1. Configuration snapshots.

Network operators track changes in device configuration for a variety of reasons, for instance, to help them debug configuration errors by comparing current configuration to older ones. They are aided in this task by network management systems (NMS) such as RANCID and HPNA. NMSes take configuration snapshots either periodically (e.g., every night) or when device configuration changes. Many types of devices send a syslog alert when their configuration changes; NMSes subscribe to this feed and pull the latest configuration when they see this alert. Snapshots are archived in a database or a version control system. Each snapshot includes the configuration text and additional useful information (also extracted by NMSes), such as timestamps of configuration changes, the login information of the entity that made the change, device model and firmware version.

Configuration files are technically part of the control plane, not management plane (Figure 1). But our observation is that analyzing the information in successive snapshots allows us to infer management plane activities.

2. Trouble ticket logs.

The second data source is the history of network alerts and failures, also known as trouble ticket logs. Each log entry has a mix of unstructured and structured information. The former includes syslog details and communication (e.g., emails and IMs) between operators that occurred to diagnose the issue. Structured information can include aspects of the issue such as the device that failed or the type of root cause. While unstructured information is commonly present, the presence and quality of the structured information varies across networks.

Of course, other data sources exist that can provide insight into management practices. For instance, some networks have documents that specify the desired network design, policies, and how changes should be made; similarly, operators could be surveyed to gain insight into management practices. We focus on the data sources above because they can provide up-to-date, quantitative information (e.g., specification documents may be out of date, and surveys may not provide quantitative information). However, when other reliable sources of data are available, they may be used to extract richer inferences or to cross-check inferences made using data sources above.

2.2 An analysis framework

We outline a simple framework for management plane analytics using the data sources above. This framework is only

Management practices
Design practices
Number of devices and links
Topology characteristics (e.g., whether a fat tree, network diameter)
Types of devices (e.g., switches, routers, load balancers)
Hardware and software heterogeneity
Operational practices
Rate of change
Size of changes (#devices, #config lines)
Nature of changes (e.g., adding nodes, editing ACLs)
Mode of change (automated, manual)

Operational health
Rate of alerts and tickets
Common root causes (e.g., configuration error, h/w failure, s/w bug)
Impact of failures (e.g., availability loss, duration)

Table 1: Example metrics for management plane analytics.

one possible way to analyze the data—it may not even be the best way—but it has helped us narrow down the space of possible analyses. The overall space is huge and there are many other promising ways to slice it.

Our framework is based on inferring two categories of metrics and then correlating them across time for a given network or across different networks. The two categories of metrics and example metrics for each are shown in Table 1. Our goal here is not to be exhaustive, but rather to illustrate the rich set of metrics that can be gathered and the rich analysis that this permits.

The first category is management practices, which we subdivide into two types based on the timescales of activities. The first subtype, which we call design practices, represents slower timescale activities of designing and provisioning the network, selecting suitable devices, etc. It can be captured using metrics such as the number of devices and links, levels of hierarchy in the topology, network diameters, and hardware and software heterogeneity. More advanced metrics for capturing the complexity of a network's design and configuration, such as, the heterogeneity in reachability policies imposed across end-point pairs [2] can also be considered.

While many networks do not explicitly log and track all design practices, we can infer the metrics from configuration snapshots. For instance, topology can be inferred from interface configurations in device configuration snapshots; layer-3 links can be inferred using subnet assignments of interfaces (two ends of a link share a /30 prefix), and layer-2 interface configuration often has comments about which device they connect to. Richer information can be obtained by constructing more descriptive models of a network's configuration; e.g., the number of configuration templates in use [2] can shed light on hierarchy in the topology, and "routing process graphs" [13, 2] can illuminate the existance of administrative divisions in the network [13]. While potentially noisy, such inferences are good enough for the types of aggregate analyses that we are proposing.

The second subtype of management practices, which we call operational practices, represents day-to-day activities for operating the network. These activities can be captured using metrics such as rate of configuration changes, size of changes in terms of number of configuration lines or devices, types of changes, and modality of changes. As we show below, these metrics can be inferred from configuration snapshots.

By studying the metrics above, we can meet the first goal of management plane analytics, which is to infer and understand management practices. But it does not help meet the second goal of understanding which practices are more effective. For this goal, we leverage a second category of metrics, which represents operational health of the network. It can be captured using metrics such as the rate of tickets, types of problems found, and the impact of experienced failures. These metrics can be extracted either directly from trouble ticket logs if structured information is reliably populated or inferred by analyzing unstructured information [15].

Given both categories of metrics, we can now infer which management practices can lead to better operational health by correlating the metrics across space (different networks) or across time for the same network. For instance, by correlating the rates of change and rate of tickets across different networks, we can determine if changing a network more frequently correlates with more faults. A similar correlation analysis can be done by slicing a network's data temporally. For example, we can compute the rate of change and rate of tickets for different months in the past, and see if months with more changes experienced higher failure rate in general. Many such correlations are possible, each providing valuable insight into the nature and impact of management practices in today's networks.

While valuable, the above does little to show causation: we could infer that the configuration change rate is positively correlated with the rate of faults, but we cannot conclude from it that reducing change rate will reduce the fault rate. Understanding causal relationships is key to developing accurate predictive models for a network's susceptibility to management issues or failures, which is the final goal of management plane analytics. However, deriving such relationships and building suitable models is non-trivial. In particular, we must systematically control for various confounding factors to establish causality with reasonable confidence. Of course, we must also systematically identify all plausible confounding factors.

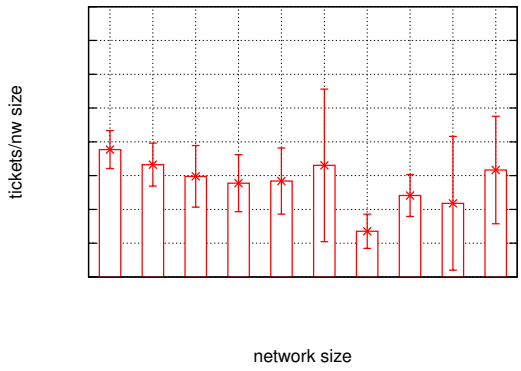

Figure 2: Is network size correlated to operational health?

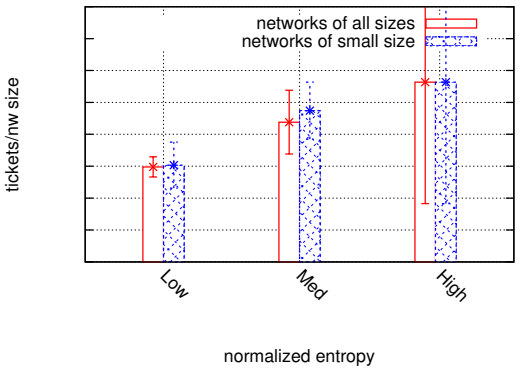

Figure 3: Is network heterogeneity correlated with operational health?

In the next section, we present an illustration of management plane analytics. Via a preliminary study, we exemplify the insights that it can offer and show empirically some of the underlying challenges in developing predictive models. We discuss richer analysis methods in §4.

3. PRELIMINARY ANALYSIS

Our study is based on data center networks owned and operated by a large online service provider (OSP). The OSP manages multiple data centers across the world. The devices within a data center are organized into multiple networks (or domains). Different networks host different applications and have different sets of operators responsible for them. Each network is built according to one of a handful of *architectures*, where an architecture indicates the number of different device roles in a network and the physical and logical interconnections to employ between devices within and across a role. The networks range in size from tens to hundreds of devices from several major vendors. In all, we study hundreds of networks that are spread around the world.

For each network, we obtain trouble ticket logs over a fourteen month period. We also obtain configuration snapshots of each device over the same period. The snapshots were taken by a network management system (NMS), which also tracks the role the device plays in the network and its vendor and model.

3.1 Design practices

The first question we attempt to answer via our framework is *which design practices correlate well with the operational health of a network?* We use the number of tickets observed for devices in a network, normalized by network size, as an indicator of its operational health.

We first consider the network's size. In Figure 2, we show the expectation of the normalized number of tickets observed for a network, as a function of size. Each bar represents a fixed-size bin, and we show 95% confidence intervals. For confidentiality, in this and other graphs in the paper, we do not show the absolute values of metrics that we study. But correlations that we seek for our analysis are still visible.

We can see in Figure 2 that smaller networks in general appear to have poorer operational health, but the weak negative correlation implies network size is not a good predictor.

Thus, we examine each network "one level deeper" using a normalized entropy metric that indicates *device-role heterogeneity*. The metric is computed for a given network as follows: $\frac{-\sum_{i,j} p_{ij} log_N p_{ij}}{log N}$, where p_{ij} is the fraction of devices of hardware model i that play role j in the network, and N is the size of the network. Higher values of this metric indicate greater heterogeneity, where the heterogeneity captures both the multitude of vendors used as well as the number of distinct device roles in a network's architecture.

In Figure 3, we show the expectation of the normalized number of tickets observed for a network, as a function of the normalized entropy binned in units of 0.33. That is, "low," "medium," and "high" heterogeneity refer, respectively, to normalized entropy in the range $[0, 0.33)$, $[0.33, 0.66)$, and $[0.66, 1]$. When applied to all networks, we find a clear positive correlation between heterogeneity and poor network health. To control for network size, we also show the results for small networks, and we observe similar evidence of a positive correlation.

Thus, *the degree of heterogeneity in a network is correlated with, and potentially impacts, its operational health.*

3.2 Operational practices

The next question we consider is *which operational practices are correlated with operational health of a network?* We consider how often a network's configuration is changed. In particular, we group all configuration changes that occur within a time window W of each other into a *change event*. We then compute the number of change events a network experiences per day, i.e., the *change event rate*.

To control for the impact of heterogeneity, we focus on networks whose normalized entropy lies within a certain small range. In Figure 4, we show operational health, as defined

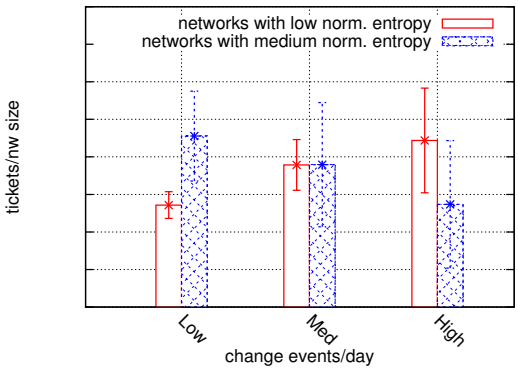

Figure 4: Is change event rate correlated with operational health?

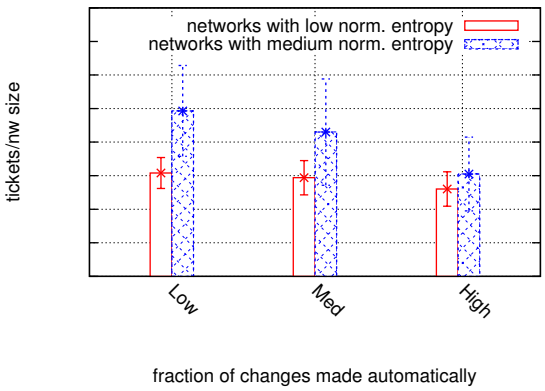

Figure 5: Does extent of automation correlate with operational health?

earlier, as a function of change event rate. The bars correspond to fixed-size bins. We see that within networks with low degree of heterogeneity, those with a higher change event rate see more trouble tickets. However, we do not see such positive correlation in networks with a medium degree of heterogeneity. One plausible explanation is that the extent of heterogeneity in the latter case may already be contributing significantly to the network's (poor) operational health. Thus, *change event rate is correlated with, and potentially impacts, operational health for homogeneous networks*.

Next we consider if a network's change events were dominated by "automatic" changes. These are changes made by automation scripts or cron jobs, which we can determine based on the user identifier included with each configuration snapshot. To elaborate, if all configuration changes in a change event were made by either a cron job or an automation account user ID, we consider the change event to be automatic. We compute the fraction of all change events that are likely to be automatic. This method underestimates automatic change events as the operators tell us that many changes that don't include automation account user IDs may also have taken place using other automation scripts.

Figure 5 shows operational health as a function of the fraction of change events that were automatic (binned in units of 0.33). We see that within networks of a given degree of heterogeneity, there is no clear relationship between the fraction of automated change events and the operational health. Thus, unlike the rate of change, *the extent of automation may not be a good predictor for a network's operational health*.

3.3 Toward a model of operational health

The first step to building a model of operational health is to identify the factors that matter. One way to judge if all relevant factors are being considered is to determine if the variability in the health metric (across time or across networks) can be collectively explained by the chosen factors.

For this analysis, we consider normalized ticket rate as the metric of operational health, along with network size, entropy, change rate, and extent of automation as potential factors. We then conduct analysis of variance (ANOVA) [1] which partitions the variance of the response variable—the operational health metric in our case—into components that can be attributed to different factors and residual variance that none of the factors capture. This analysis confirms the results above; it shows that network heterogeneity is a key contributory component (p-value < 0.05) and network size and extent of automation matter less.

More importantly, it also shows that the set of factors we consider is far from complete. They capture less than 95% of the variance in the normalized ticket rate. Thus, we clearly need to identify additional contributory factors. We then added two more factors into the ANOVA analysis—mean number of devices impacted by a change event and the architecture type of the network. We find that the first factor contributes but, surprisingly, the second one does not. However, much of the variance in normalized ticket rate still remains to be explained. We are in the process of identifying more contributing factors, starting with some of the metrics that we proposed in §2.

4. RESEARCH AGENDA

Our analysis above merely scratches the surface of what is possible through management plane analytics, and significant more work is needed before the promise of management plane analytics can be realized. We identify three broad areas of investigation.

New inference techniques.

So far, we have only made simple inferences from our data. But the data sources are much richer, and more valuable information can be extracted. A particularly promising direction is deeper analysis of successive configuration snapshots to infer the intent that underlies a configuration change. This requires reasoning about both changes within a device and across devices.

Further, while techniques exist to automatically classify

trouble ticket logs into likely root causes [15], little work has thus far gone into mapping faults to specific changes in the control plane made by the management plane. The challenge here is that faults due to a change can occur well after the change was applied (which is why we conducted a coarse-grained analysis to correlate changes to faults). A better understanding of the nature of changes and faults can help attribute faults to individual changes.

Making inferences from the outside.

Much of our focus above is on mining the two data sources described in §2. However, it may be possible to make useful inferences without direct access to that data by leveraging existing techniques that measure networks' data and control planes from the outside. Many techniques already exist to infer network characteristics such as topology (e.g., Rocketfuel [18]) and device types (e.g., nmap [14]). Operational health of the network too can be observed externally, e.g., by regularly probing for failures (e.g., Hubble [9]) and observing the control messages that a network sends (e.g., BGP announcements [12]).

That leaves us with the challenge of inferring management practices. These practices may be inferred by observing changes in the network's control plane. For instance, techniques exist to infer a network's control plane (e.g., routing weights and preferences) using data plane measurements and BGP announcements [17]. By repeatedly inferring the control plane, we can observe changes that the management plane makes to it.

The inferences above may not be perfect, but we believe that they can be done with usable reliability. If we succeed, it would mean that researchers would be able to conduct management plane analytics for commercial networks without active cooperation from the operators of those networks, greatly expanding the set of networks that can be studied.

From correlation to causality to practice.

While correlations that we suggest above are useful for gaining insights, they do not imply causality. It is not necessarily the case that a given network's operational health will improve if a management practice that is found to be correlated with poor operational health is changed. We need to establish causality before we can confidently recommend changes to certain management practices. For this goal of impacting practice, the networking community needs to develop a different set of scientific methods. There are two potential models.

The first one is observational studies in a large population. If data from such a large number of networks is available such that we can control for various confounding factors and conduct a counterfactual analysis [5], then we may be able to establish causality with (statistical) confidence. Implementing this model will likely require a shared repository of information on the three types of metrics from many networks; actual raw data is not necessarily needed. With community support, it is feasible to create such a repository, as was done for wireless traffic traces [6]. A key challenge here is understanding if all confounding factors have been accounted for. As shown above, techniques such as ANOVA [1] can prove useful here. Another key challenge is to normalize data contributed by different, independently-administered networks. For example, trouble tickets likely have different meanings across organizations; one way to normalize them is associate some uniform measure of impact with each ticket (e.g., some normalized duration of outage).

The second model is controlled trials, where a network changes a particular management practice without changing anything else. By repeating with multiple networks and observing the impact on operational health, we can establish causality with (statistical) confidence. University and research networks are good candidates for such trials given their close relationship with the research community.

5. RELATED WORK

We are not the first to express interest in network management practices. Other researchers have investigated specific aspects of network management. For instance, references [3, 10, 19] study network configuration snapshots from a handful of campus and ISP networks to understand how configurations are changed over time when realizing various high-level tasks and/or as networks grow over time. These and other similar studies focus mainly on the first aspect of management plane analytics (inferring a network's management practices), and as such they do little to help improve management practices.

Researchers have also studied the network trouble ticket logs to infer root causes and common failure modes [15]. We advocate using such inference methods and correlating their inferences to management practices, to gain insight into which practices lead to fewer network faults.

6. CONCLUSIONS

We argued for systematic analysis of the management plane of today's networks to understand which practices are common and which ones lead to better outcomes (e.g., fewer faults). We observe that such analysis can be enabled by mining data sources that are already available, and we demonstrate its feasibility by conducting a preliminary analysis across hundreds of data center networks. Our analysis shows that the operational health is lower for networks that are more heterogeneous and for network where configuration changes are more frequent.

Acknowledgments We thank the operators who provided insight into the management practices for the networks that we study. We also thank Aaron Gember-Jacobson, Robert Grandl, Navendu Jain, Raajay Vishwanathan and the HotNets reviewers for feedback on this paper. This research was supported in part by NSF grants CNS-1302041, CNS-1314363 and CNS-1040757.

7. REFERENCES

[1] Analysis of variance. http://en.wikipedia.org/wiki/Analysis_of_variance. Retrieved 2014-07-15.

[2] T. Benson, A. Akella, and D. Maltz. Unraveling Complexity in Network Management. In *NSDI*, 2009.

[3] T. Benson, A. Akella, and A. Shaikh. Demystifying configuration challenges and trade-offs in network-based ISP services. In *SIGCOMM*, 2011.

[4] C. Bird, B. Murphy, N. Nagappan, and T. Zimmermann. Empirical software engineering at microsoft research. In *CSCW*, 2011.

[5] Impact evaluation. http://en.wikipedia.org/wiki/Impact_evaluation. Retrieved 2014-07-15.

[6] CRAWDAD. http://crawdad.cs.dartmouth.edu/.

[7] A. B. Downey. Using pathchar to estimate internet link characteristics. In *SIGCOMM*, 1999.

[8] M. Jain and C. Dovrolis. End-to-end available bandwidth: Measurement methodology, dynamics, and relation with TCP throughput. In *SIGCOMM*, 2002.

[9] E. Katz-Bassett, H. V. Madhyastha, J. P. John, A. Krishnamurthy, D. Wetherall, and T. Anderson. Studying black holes in the internet with hubble. In *NSDI*, 2008.

[10] H. Kim, T. Benson, A. Akella, and N. Feamster. The evolution of network configuration: A tale of two campuses. In *Internet Measurement Conference*, 2011.

[11] R. Mahajan, N. Spring, D. Wetherall, and T. Anderson. User-level Internet path diagnosis. In *SOSP*, 2003.

[12] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP misconfiguration. In *SIGCOMM*, 2002.

[13] D. Maltz, G. Xie, J. Zhan, H. Zhang, G. Hjálmtỳsson, and A. Greenberg. Routing design in operational networks: A look from the inside. In *SIGCOMM*, 2004.

[14] Nmap. http://nmap.org/.

[15] R. Potharaju, N. Jain, and C. Nita-Rotaru. Juggling the jigsaw: Towards automated problem inference from network trouble tickets. In *NSDI*, 2013.

[16] Really Awesome New Cisco ConfIg Differ (RANCID). http://www.shrubbery.net/rancid/, 2004.

[17] N. Spring, R. Mahajan, and T. Anderson. Quantifying the causes of path inflation. In *SIGCOMM*, 2003.

[18] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson. Measuring ISP topologies with Rocketfuel. *IEEE/ACM Transactions on Networking (ToN)*, 2004.

[19] Y. Sung, S. Rao, S. Sen, and S. Leggett. Extracting Network-Wide Correlated Changes from Longitudinal Configuration Data. In *PAM*, 2009.

[20] HP OpenView TrueControl Software. http://support.openview.hp.com/, 2004.