Intrusion-Tolerant Cloud Monitoring and Control *

Daniel Obenshain
Johns Hopkins University
dano@cs.jhu.edu

Tom Tantillo
Johns Hopkins University
tantillo@cs.jhu.edu

Andrew Newell
Purdue University
newella@purdue.edu

Cristina Nita-Rotaru
Purdue University
crisn@cs.purdue.edu

Yair Amir
Johns Hopkins University
yairamir@cs.jhu.edu

1. INTRODUCTION

Clouds usually span several geographically separated data centers in order to ensure low-latency access to a set of clients (either external or internal to the cloud system) and to gain resilience against correlated faults. Each data center often has connectivity to multiple Internet backbones, providing similar accessibility and resiliency benefits.

An effective approach to ensure communication between multiple data centers and leverage the multiple backbones paradigm is to use an overlay messaging architecture, with one logical node at each data center. Each logical node may consist of tens of thousands of machines, allowing the cloud system to scale, while the relatively small number of logical nodes meets the scalability requirements of overlay messaging. Our experience shows that in a benign environment, overlay routing [2] is able to provide sub-second convergence as long as a path exists between source and destination [3]. In contrast to end-to-end packet recovery for reliable Internet communication, overlay routing enables hop-by-hop recovery of lost packets, resulting in more timely recovery and lower end-to-end loss and jitter [1].

Monitoring and control messaging services are critical for a cloud infrastructure as a cloud is remote to its administrators. Since administrators do not have constant physical access to the cloud infrastructure, management must be done through monitoring and control messages gathered from and passed to the elements of the cloud. Without these messages, the administrators lose the ability to manage and react to developing situations in the cloud. One of the most dire situations is when part of the cloud is being compromised. In such situations, monitoring and control messaging serves as the tool for administrators to diagnose and recover the system. In order for the messaging service to operate in such situations, it has to be intrusion-tolerant.

Constructing an effective overlay messaging system that overcomes intrusion is a challenge. For example, a compromised node can broadcast faulty overlay routing state, forcing all overlay messages to flow to it, where they are then dropped. Thus, a single compromise can stop all service in the cloud and entirely cripple the administrators' ability to monitor and control the cloud. In order to provide maximum intrusion tolerance against cloud compromises, there is a need for optimal resiliency for the monitoring and control messages. This would in turn provide the utmost reachability in the presence of compromised portions of the cloud and would indirectly benefit all cloud applications.

Our contribution. We propose an intrusion-tolerant overlay messaging service for cloud monitoring and control. Our service, Controlled Authenticated Overlay Flooding, combines authentication mechanisms and flooding with an overlay architecture to provide optimal resiliency in the presence of compromised nodes. While Internet flooding is not feasible because of the number of nodes and links involved, our approach is practical as it imposes a maximal overlay topology and confines the communication to that topology's overlay links, limiting the cost of flooding.

Below we provide more details about our service and discuss requirements for monitoring and control services.

*Partially funded by DARPA grant N66001-1-2-4014

2. CONTROLLED AUTHENTICATED OVERLAY FLOODING

Overlay routing algorithms have different trade-offs in terms of resiliency and cost. For example, an algorithm that uses k node-disjoint paths to route from source to destination is resilient against up to $k-1$ compromised nodes with a cost of k times the single-path routing costs. Overlay flooding has a higher cost, but is optimal in terms of resiliency (reachability) in the presence of compromised nodes. Specifically, overlay flooding ensures that messages will flow from source to destination, as long as there exists a path of honest nodes between source and destination.

The idea of using flooding as an intrusion-tolerant state dissemination mechanism was first proposed by Perlman[4, 5]. The proposed algorithm is intended for link state dissemination and does not perform well as a general messaging service.

Assumptions. We assume that compromised nodes are not able to prevent neighboring correct nodes from exchanging messages. That is, a compromised node cannot overwhelm the bandwidth resource to receive packets or the computational resource to verify received packets. This is a reasonable assumption as such volume of traffic can be easily detected and responded to by network administrators. We assume that compromised nodes may act arbitrarily, including dropping messages, delaying messages, and replaying messages.

Our Approach. We propose Controlled Authenticated Overlay Flooding which confines the overlay nodes to a maximal topology that defines which nodes are direct neighbors on the overlay. Correct nodes accept messages only from direct neighbors. The protocol relies on public key cryptography to ensure authentication, i.e. each overlay node has its own private key, the cloud administrator has an offline private key, and all overlay nodes know the corresponding public keys. Authentication ensures that all messages are

cryptographically authenticated and that no new nodes can join the network without an action of the cloud administrator. Signatures guarantee that messages cannot be altered or nodes impersonated.

Messages flow in the overlay as follows: a node that receives a message from a source will forward it to all of its neighbors. Upon receiving a message from a neighbor, an overlay node forwards the message to all of its other neighbors. The overlay node that is serving the destination forwards the message to that destination the first time it receives that message. The overall cost of the protocol is that each message is sent on each link at least once, and often twice (in both directions). To reduce the likelihood of the same message being sent on the same link in both directions, nodes delay the forward operation by a small, randomized period of time, and do not forward a message if it was already received from the other side.

Our protocol enforces flow control and limits each node's ability to overwhelm the system. Each node enforces limits on all the traffic that flows through it and maintains fairness. As a result, any traffic that flows from a compromised node to an honest node will be constrained.

Practical considerations. Our experience shows that practical topologies that provide good coverage for a global cloud consist of tens of nodes and hundreds of edges. For example, in an overlay of 50 well-placed overlay nodes with each node connected to five to ten direct neighbors, each message will be sent on between 125 to 250 links. Since the average path in such a global network is about eight overlay links, this protocol is about 15 to 30 times more expensive than a regular link state protocol with shortest path routing.

Based on the experience of the LTN cloud [3], the combined traffic of monitoring and control messages is very small compared to the overall traffic in the cloud. Control messages are sent infrequently, resulting in a negligible amount of traffic. Monitoring messages are usually sent periodically, consuming in the order of 0.1% of total traffic. As a result, using the above example, we expect the Controlled Authenticated Overlay Flooding algorithm to utilize about 3% of overall cloud traffic when used for monitoring and control. This represents a small overhead in exchange for optimal monitoring and control resiliency in the face of intrusions.

3. MONITORING VERSUS CONTROL

Monitoring and control messages have different properties and require different guarantees. This calls for different variations of Controlled Authenticated Overlay Flooding. The challenge we aim to overcome is to ensure the availability of resources (bandwidth, computation and memory) to each source despite the existence of malicious nodes which aim to consume network resources.

Monitoring. Monitoring messages are usually periodic status messages, where some are more important than others. When network resources are limited, it is necessary to ensure that the most important monitoring messages are delivered in a timely manner. In order to ensure the availability of resources, each overlay node enforces strict source-based fairness on each of its links to mitigate the ability of compromised nodes to overburden the network.

Source-based fairness allows the source to determine which of its messages are more important than others, knowing that whenever there is contention among its messages, the least prioritized messages will be dropped. Still, fairness ensures that each source is guaranteed at least $\frac{1}{N}$ of the available out-bound network resources on each link of each correct overlay node (N being the number of overlay nodes in the network). Therefore, on each link, messages generated by any specific malicious source node cannot consume more resources than messages generated by any specific correct source node, unless that correct source node has received all the resources it requires on that link.

Control. Control messages change the state of the cloud and therefore require reliable delivery. For example, the maximal topology can be dynamically changed by the system manager by issuing a newer topology signed with its private key. In order to make this communication reliable, and since overlay nodes have only finite memory, back-pressure must be used to stop correct source nodes from generating messages beyond what the network can absorb.

An overlay node stores a control message until it knows that all neighboring nodes have that message or it receives an end-to-end acknowledgment, indicating that the destination has received that message. These end-to-end acknowledgments serve the added purpose of clearing out messages stuck in intermediate overlay nodes, either due to network problems or malicious neighbors that refuse to acknowledge them.

For reliable delivery, flow-based fairness rather than source-based fairness needs to be maintained. By flow we mean all messages sourced at a specific overlay node that are destined to another specific overlay node. With source-based fairness, a malicious destination can block flows from a correct source by refusing to acknowledge the messages from that source destined to itself, causing that source's allocated memory at the intermediate nodes to fill up. With flow-based fairness, a malicious destination can only block flows destined to itself (as long as there is at least one path of correct nodes from source to destination). Furthermore, in case of link contention, an overlay node guarantees that it will not forward more than a flow's fair share, ensuring that malicious nodes cannot adversely consume network resources beyond their direct neighbors.

4. CONCLUSION

We have described a flooding-based messaging service for monitoring and control of cloud infrastructure that overcomes intrusions. By constraining the maximal topology and ensuring fairness on each link, the overhead associated with this service is practical for global deployments.

5. REFERENCES

[1] Y. Amir and C. Danilov. Reliable communication in overlay networks. In *Proceedings of the IEEE DSN*, pages 511–520. Citeseer, 2003.

[2] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris. Resilient overlay networks. *ACM SIGCOMM Computer Communication Review*, 32(1):66–66, 2002.

[3] LTN global communications. `http://www.ltnglobal.com/`. Accessed: 5/2/2012.

[4] R. Perlman. *Network layer protocols with byzantine robustness*. PhD thesis, Massachusetts Institute of Technology, 1989.

[5] R. Perlman. Routing with byzantine robustness. *report no: TR-2005-146 [Online]. Available: http://research. sun. com/techrep/2005/smli_tr-2005-146. pdf*, 2005.