When an administrator, Alice, is running a data center and observes suspicious behavior, she must investigate whether the behavior is legitimate or malicious, as seen in the network infiltration attempt in Figure 1. To do so, a network forensic system is required in order to answer questions about the possible system breach. Past work has formalized this problem as a matter of data provenance, creating a network provenance system that gives Alice explanations as to the origins of network messages. We present a novel design for a high-fidelity network provenance system by merging network activity with host-layer context.

Host-Layer Provenance

Host-layer provenance produces records that detail every event that occurs with respect to system objects inside an individual machine. This provenance scheme tracks every object from the time it comes into existence up until its current state. Provenance data generation and verification occurs within the kernel, preventing rogue applications from tampering with the trusted computing base of the provenance monitor.

In our architecture, provenance monitors communicate via a Provenance Verification Points (PVPs) to securely collect provenance for network events. This even enables provenance collection from hosts under attacker control.

Without this host context, it is impossible to fully explain the cause of these network events.

In future work, we intend to reconcile this network-layer provenance with the host-based approach, creating provenance of exquisite detail.

Conclusion

By combining host- and network-layer provenance mechanisms, we are able to generate records with extreme levels of detail, providing administrators with an all-encompassing view of their system. Furthermore, with the aid of Software-Defined Networking, network-layer provenance removes the need for individual machines to comply with the host-layer provenance protocol, which allows for greater coverage of heterogeneous systems.