

Efficient Data Structures for Tamper-Evident Logging

Scott A. Crosby Dan S. Wallach

Rice University

Presented at USENIX Security 2009

Reliance on logs



Assume the adversary doesn't tamper with the logs.

Towards Scalable Cluster Auditing through Grammatical Inference over Provenance Graphs

Wajih Ul Hassan, Mark Lemay,[‡] Nuraini Aguse, Adam Bates, Thomas Moyer*

DeepLog: Anomaly Detection and Diagnosis from System Logs through Deep Learning {mind, lifeifei, guineng, svivek}@cs.utah.edu

Yang Ji, Sangho Lee, Evan Downing, Weiren Wang, Mattia Fazzini, Taesoo Kim, Alessandro Orso, and Wenke Lee Georgia Institute of Technology 1 INTRODUCTION

RAIN: Refinable Attack Investigation with On-demand

Inter-Process Information Flow Tracking

Intrusion Recovery Using Selective Re-execution

Taesoo Kim, Xi Wang, Nickolai Zeldovich, and M. Frans Kaashoek MIT CSAIL

The districte the advicency design FUTEN, consider the following stands, when we sell use is a uning exemple district and obtains a reas, shell. To make he read, so the moves the last how controls from the system to Six them executes success a brokels we first the system in the same or account for sever and to FPP sections a first for account or when the properties allowed her to execute orbitary commends with HTTP. For the movement of new factors in downship and that the control for the sections of these backers in downship and that the control for the section of the s

ProTracer: Towards Practical Provenance Tracing by Alternating Between Logging and Tainting

gunes; andit logging and provinence propagation for arbitral 1 Audit logging [16], [27], [25], [27], [29], [34]–[56], [39] records exents during costem execution and then consolly

Transparent Web Service Auditing via **Network Provenance Functions**

MPI: Multiple Perspective Attack Investigation with Semantics Aware

Dongyan Xu

Xiangyu Zhang Purdue University

Abstract

Traditional auditing techniques generate large and inac curate causal graphs. To overcome such institutions, researchers processed to leverage execution particulation, as transcription and the processed of the second processed of the second particulation and the processed of the second particulation of the processed o tan, Pertraca and it reserves spaces. The economics results show that our technique generates cleaner attack graphs with rich high-lazed sensories and has much lower speci and time as arbenda, when compared with the escut loop based partitioning techniques BEEP and ProTeaser.

understand the strack including its root cames and strack-cer costs. Such inspection is certical for fundly recycles in the cost of the cost of the cost of the cost of the extensions [15], so, 69, 50, 50 of multi-hosting and record-ing impostant system level exists [26], with copartisons, and then consultating these costs during an colline investigation process. The consultations were analyst bytes, between two processes such as a process version as a charge of the process brough exp. Journal instrument process was an approximate process. The consultations will be the control of the cost Hosticas, these text-injunes suffer from the depositions of the cost of the productive with many dipotent and their processes during its lifetime all length only of such and be in this cost of the productive with many dipotent and their processes during its lifetime all length only in our propriet in the cost of the cost of the cost of the productive with a cost of the cost of the cost of the cost of the length of the cost of the cost of the cost of the cost of the length of the cost of the cost of the cost of the cost of the length of the cost of the cost of the cost of the cost of the length of the cost of the cost of the cost of the cost of the length of the cost of the cost of the cost of the cost of the length of the cost of the length of the cost of the length of the cost of the length of the cost of the cos

band systems such as FEFF [4] and ParTraca [46] true demonstrated great potential, they restrictored execution based on event bandling loops. That is, each intention of an event bandling loop is considered a unit. Despite its generality, such a partitioning scheme has inhumal lim-lations. (c) Event loop intentions are too low level and 1 Introduction

Trousmost tracking is critical for stack invocation, consists of the control beyindrated uses to the local and proposally for Advanced Ferences (Photon (ATT)) that it to back by optimization seek to allow provinces and steep comments and secretics. Aft tracks do they again a long control of time with 1,000 profile, and bence are difficult to detect and investigate. A provenance tracking system records the causality of system objects (e.g., files) and subjects (e.g., files). Note that the adults of the causality of system objects (e.g., files) and subjects (e.g., fil

26th USENIX Security Symposium 1111

High Accuracy Attack Provenance via Binary-based Execution Partition

Kyu Hyung Lee Xiangyu Zhang Dongyan Xu nputer Science and CERIAS, Purdue University, West Lulayette, IN 47907, USA

Advancional important apport of cyber altock foresite is interested to substitute of the matter. It is still understant the prostument of apportant content at Reference is to content to the matter. It is still understant the prostument of apportant content in Reference is the content to the matter. It is still the store that is done in any adjusting and like, in Reverce the parts inclined to depend the store of the matter and the store of the postument capabilities to depositions of capabilities and making attack forestingline to depositions of the postument capabilities and making attack forestingline to the content of the postument capabilities and making attack forestingline to the content of the postument capabilities and making attack forestingline the content of the postument capabilities and making attack forestingline and the store of the postument capabilities and the store of the postument capabilities and the postument capabilities and the store of the postument capabilities and the store of the store of the postument capabilities and the store of the postument of the postument capabilities and t

SLEUTH: Real-time Attack Scenario Reconstruction from COTS Audit Data*

Md Nahid Hossaid¹, Sadogh M. Milajordi², Junao Wang¹, Birhann Esbete², Rigel Gjomeme², R. Sekar¹, Scott D. Stoller¹, and V.N. Verkatakrishnan²

We present an approach and systems for mal-time recon-mentation of attack scenarios on an entrepyine hor. To meet the scatability and mal-time needs of the problem, we develop a plutform-neural, multi-memory broad, ca-petaciesy graph themseline of multi-log data. We their present officient, tag-water themselings for uttack deter-tion and reconstruction, including source identification and impact analysis. We the dreshep methods to reveal the last actions of fluids by contractions of community of

USENIX Association

some now, are another initially weaked excessly, and the stack.

I Introduction

We are witnessing a rigid excluding in angeled optionation (Principle dedicated and Persistent Theories Content of Section 19 and 1

Form storage and analysis. How can we store the millions of suscede from event streams efficiently and have also sittens sitt through this data in a matter of science?

28th USENIX Security Symposium 487

LogGC: Garbage Collecting Audit Log

Backtracking Intrusions

Figure 1: Audit log growth.

Can we trust the logs?



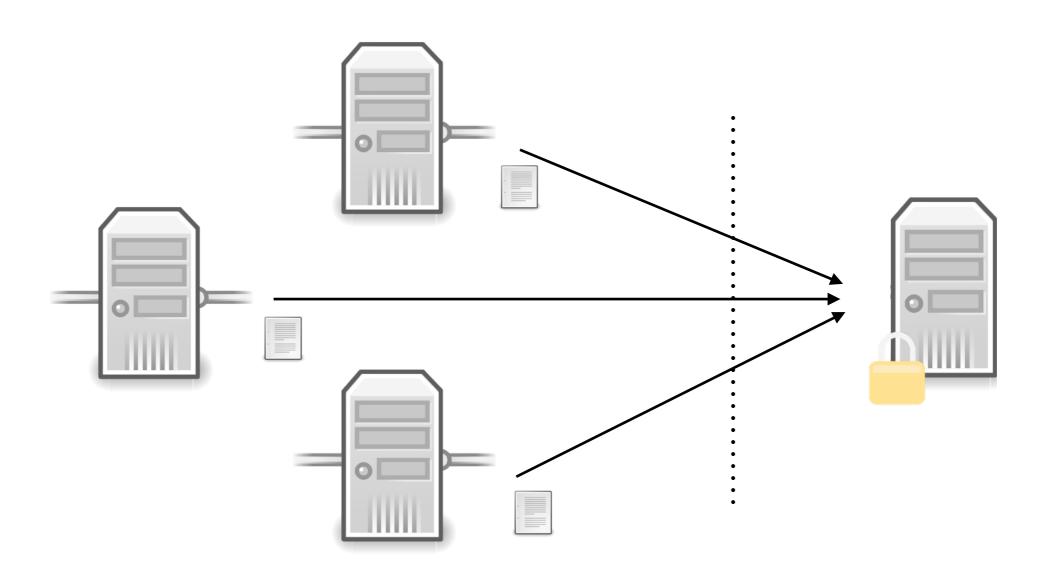
The attacker may modify the log file to cover their traces!

Goal: An event, once correctly logged, cannot be *undetectably* hidden or modified.

Industry practice



Send logs to a **trusted** central server



This paper



Allow the central server to be untrusted.

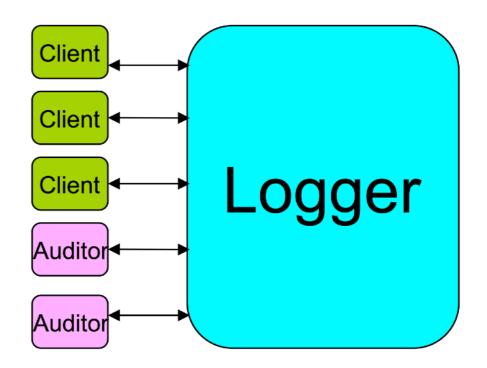
Ingredients:

- 1. Auditing
- 2. History Tree

High-level design



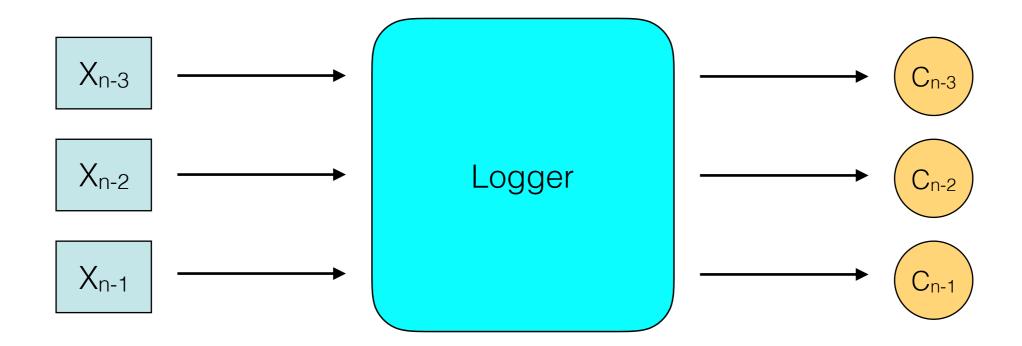
- Logger (central server)
 - Stores logs
- Clients
 - Generate logs
- Auditors
 - Verify the correct operation of the logger



Logger



- Logs come in
- Commitments go out



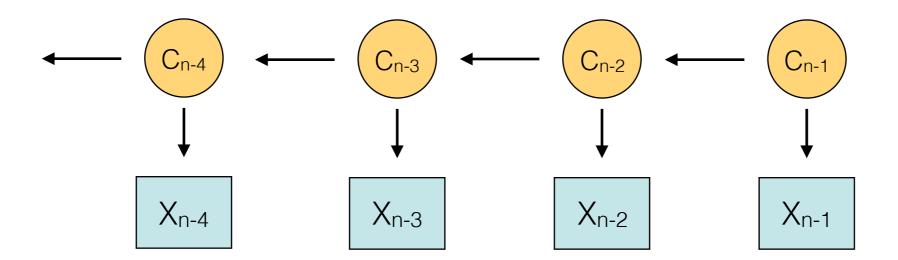
Commitments



 Each commits to the entire past. Example construction [Kelsey, Schneier]:

$$- C_n = H(C_{n-1} || X_n)$$

They are signed by the logger



We don't trust the logger!

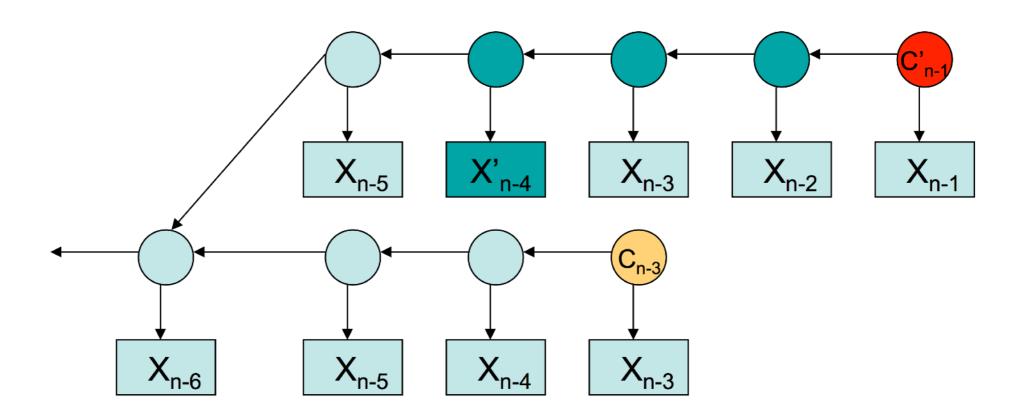


- Does Cn-3 really contain Xn-3 ?
- Do Cn-2 and Cn-1 commit the same historical events?
- Is the event at index i in the log defined by $\binom{c_n}{c_n}$ really $\binom{x_i}{x_i}$?

Example: log forks



 What if the logger rolls back the log and adds on different events?



Solution: Auditing



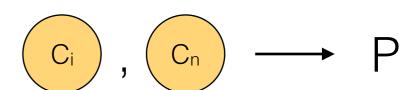
- Check the returned commitments
 - For correct event lookup
 - For consistency

Two kinds of audits



Membership auditing

- i, C_n \longrightarrow X_i , P
- Verify proper insertion
- Lookup historical events
- Incremental auditing



 Prove consistency between two commitments

Who does what?

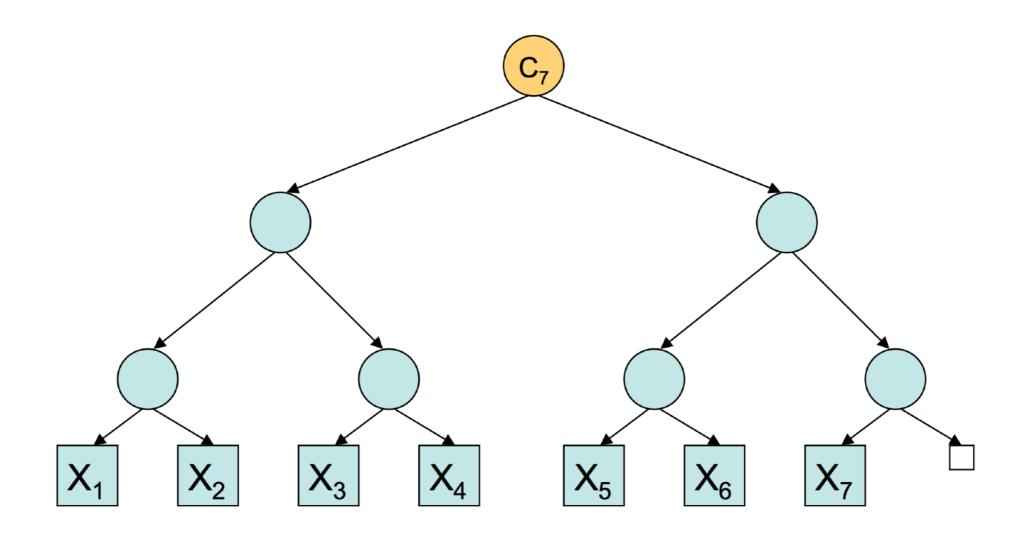


- Clients must redistribute their received commitments from the logger to auditors.
- A host can be both client and auditor at the same time.
- Auditing strategies are not discussed in detail.

Making audits cheap



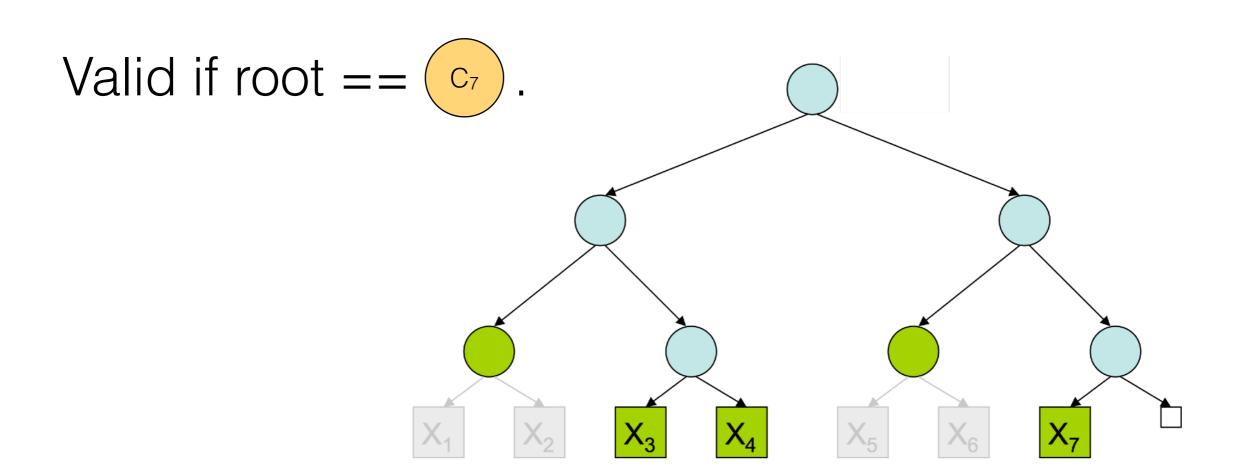
Logs are stored in a history tree



Membership Auditing



Given (3, C7) return (X3, P), where P is:



P takes O(log n) to build

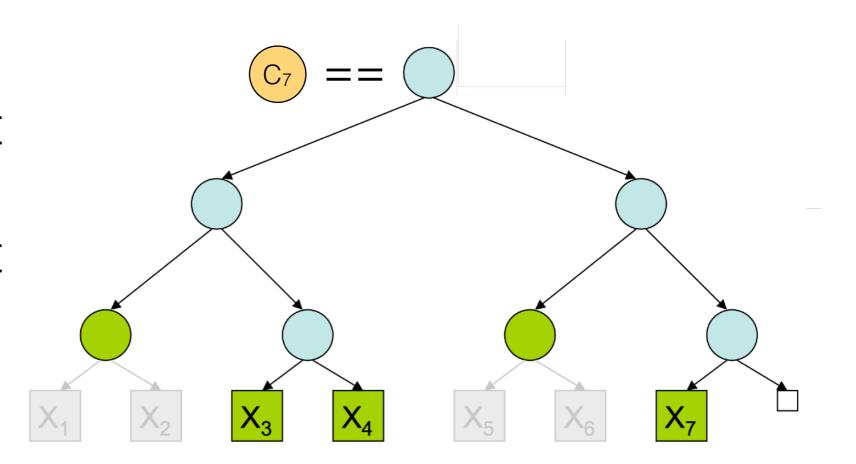
Incremental Auditing



Given (C3, C7) return (P), where P is:

Valid if:

- P is consistent
- with C7
- P is consistent
- with C₃



P takes O(log n) to build

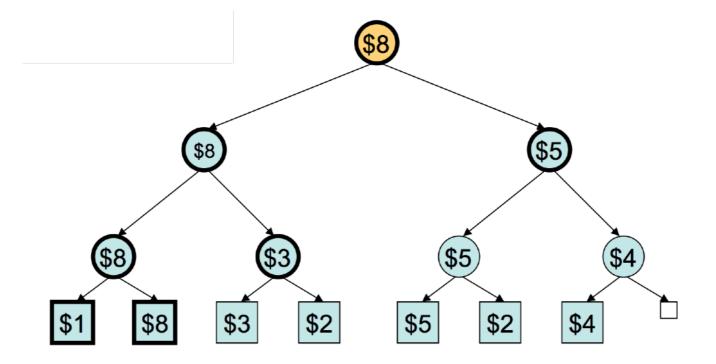
Merkle Aggregation



History trees can be extended to annotate events with attributes.

Application: support content searches.

Max()



Find all transactions over \$6

Performance



- Insert performance: 1,750 events/sec
 - 2.4%: Parse the log event
 - 2.6%: Insert the event to the tree
 - 11.8%: Get root commitment
 - 83.3%: Sign commitment

Proof generation:

- With locality (all events in RAM):
 - 10,000-18,000 incremental proofs/sec
 - 8,600 membership proofs/sec
- Without locality
 - 30 membership proofs/sec

Recap



- History trees allow the logger to store log events and generate integrity proofs efficiently.
- Other hosts (auditors) need to demand those proofs to ensure the logs are not tampered.
- Result: the logger can be untrusted (but at least one auditor needs to be honest).

Discussion



- No security analysis: what happens if a client colludes with the logger? What if the secret key of the logger is compromised?
- No full-system evaluation with multiple hosts.
 Network overhead? Overhead of redistributing commitments with gossip? Scalability?
- No auditing strategies are presented. What kind of audits, from whom and how often should be asked to the logger? What happens when tampering is detected? Lying auditors?