VAST: Visibility Across Space and Time

Matthias Vallentin, Vern Paxson
University of California at Berkeley

Motivation and Goals

Observation: Analysis of network activity is fragmented across *space* and *time* [3]

- ***Space***: Multiple heterogeneous data sources (e.g. log files from different applications highly vary in scope, quality, and expressiveness)
- ***Time***: Analysis of past activity greatly differs from analyzing future instances (e.g. processing UNIX daemon logs vs. configuring a firewall)

→ Unification of scattered analysis procedures holds promise to significantly improve the effectiveness of problem detection and forensic analysis

Core Idea: Database repository of site activity

- Application logs, Routers, Firewalls, IDS (Bro)
- *Policy-neutral* data

Vision: Operational deployment of VAST as single vantage point for arbitrary analyses of network activity

Application Scenarios

Network Troubleshooting:

- Chasing down errors in large-scale environments similar to finding *the needle in the haystack*
- Existing data volume exceeds resources to analyze it

→ Unified analysis helps in problem solving and codifies troubleshooting process for future instances

Forensics / Combating Insider Abuse:

- Reveal security incidents after they occur
- Difficult to understand full scope of insider attacks: Attacks often manifest as chain of authorized actions

→ Readily searchable archive of comprehensive activity helps to understand full breach implications

Detecting Intrusions:

- Most IDSs are either network-based or host-based

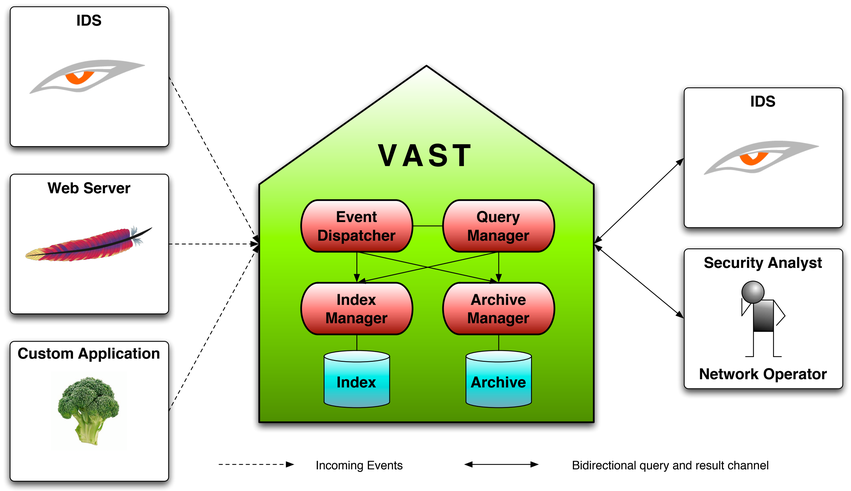
→ *Hybrid* synthesis of data from a wide range of sources substantially increases IDS efficacy

→ Synergetic effects enables potential for correlation

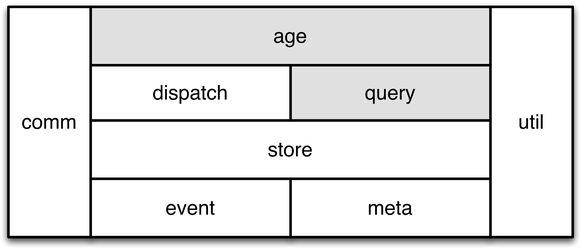
Design and Architecture

Unified Data Model: Bro's [1] rich-typed event model provides a generic abstraction for activity, enabling unification of heterogenous sources of information

Component-based Design: VAST is a distributed system based on components that can be scaled across multiple machines:



Clean Architecture: lower layers do not depend on higher layers



(grey layers not yet implemented)

Storage Engine: FastBit [2] provides efficient bitmap indexing technology to accelerate query processing

Concurrent Event Archival: Storage engine archives and indexes events in parallel to gain maximum benefit from modern multi-core CPUs

Challenges

Why not using an existing database?

- Traditional DBMSs exhibit poor query performance on high-dimensional data

→ Use bitmap indices to break *curse of dimensionality*

- Problem: OLAP-optimized DBMS cannot handle streaming data efficiently

→ Build custom streaming layer on top of FastBit

Handling Live Queries

- Unified querying: same language to access past data and install *triggers* to detect future activity
- High performance overhead to inspect each incoming event

→ Relay a copy of each incoming event to a dedicated component

→ Need an efficient event comparison mechanism to extract matches

Conclusions and Future Work

Summary:

- VAST: Intelligent database of network activity to analyze both past and future activity in a coherent fashion
- Devised architecture and first prototype implemented
- Next**:** Implement SQL-like query engine

Aggregation and Aging: Elevate events into higher semantic abstractions by condensing them into a more succinct form

→ Graceful degradation in terms of data reduction

Inter-site Analyses: Sharing of attack details today

- has highly informal character
- requires human-in-the-loop

→ Automate significant elements of cross-organizational security analyses via event-oriented analysis scripts

References:

[1] Vern Paxson. Bro: A System for Detecting Network Intruders in Real-Time. *Computer Networks*, 31(23-24):2435-2463, 1999.

[2] Kesheng Wu. Fastbit: An Efficient Indexing Technology for Accelerating Dataintensive Science. *Journal of Pysics: Conference Series*, 16:556-560, 2005.

[3] Mark Allman, Christian Kreibich, Vern Paxson, Robin Sommer, and Nicholas Weaver. Principles for developing comprehensive network visibility. *In Proceedings of the Workshop on Hot Topics in Security (HotSec)*, July 2008.