Thoth: Efficiently enforcing data confidentiality and integrity in data retrieval systems

Eslam Elnikety, Anjo Vahldiek, Aastha Mehta, Deepak Garg, Peter Druschel
MPI-SWS

1. Policy-Compliant Search Engine

Personal E-mails → Personal E-mails → Indexer → Index
Public → Company → Indexer
Internal → User preferences → Legal mandates → Privacy policies

Problem: compliance despite bugs, misconfiguration, or operator errors

2. Thoth

Thoth: a safety net against bugs and other errors

Policies: specifications in a high-level declarative policy language to control confidentiality and integrity of data.

Enforcement: per-task policy enforcement logic (PEL). PEL monitors tasks’ inputs and outputs and enforces policies regardless of bugs or errors outside the PEL.

3. Policies

Read/Write/Destroy: Who can read/write/destroy, and when?

Declassify: where can data flow?
Provenance: where can data come from?

4. Enforcement

Implementation alternatives

PEL
VM
VMM
PEL
Process
OS

Enforcement rules

• Each task (VM or a process) is tagged with the policies of all inputs it consumes.
• PEL enforces that all inputs’ declassification is satisfied by the output’s policy.
• PEL enforces that the output’s provenance is satisfied by all the inputs’ policies.

5. Example: Search Engine

Log: Admin read David’s email; subject: “..”

Guarantees with limited TCB:

• Enforcing system policies; results are delivered only if the policies associated with inputs and outputs are satisfied.