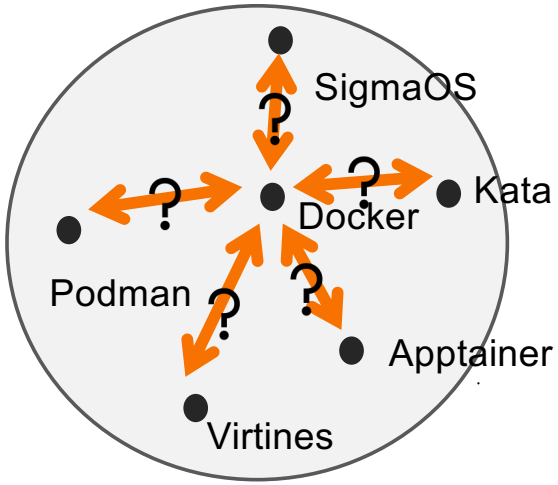
Comparing Isolation Mechanisms with OSmosis

PLOS 2025, Seoul, South Korea, Oct 13, 2025

Sid Agrawal, Shaurya Patel, Arya Stevinson, Linh Pham, Ilias Karimalis, Hugo Lefeuvre, Aastha Mehta, Reto Achermann, Margo Seltzer

Systopia Lab, University of British Columbia, Vancouver, Canada

No way to compare different isolation mechanisms



What should the comparison tell us?

- Can one task affect another task?
  - Sharing resources, or pools of resources
- How do tasks depend on each other?
  - Providing services to each other
  - Ability to control each other

Two View Points

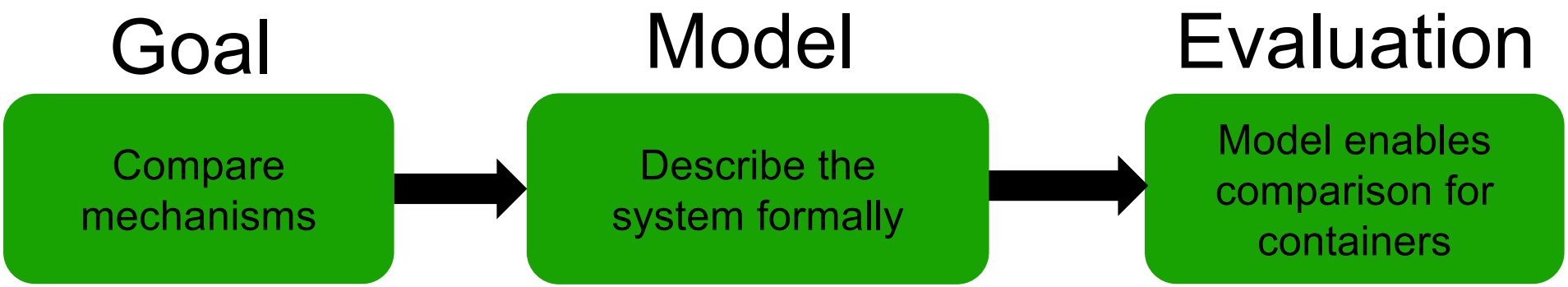
Trusted Computing Base (TCB)

- A set of domains upon which a protection domain relies
- Typically of interest to developers

Impact Boundary (IB)

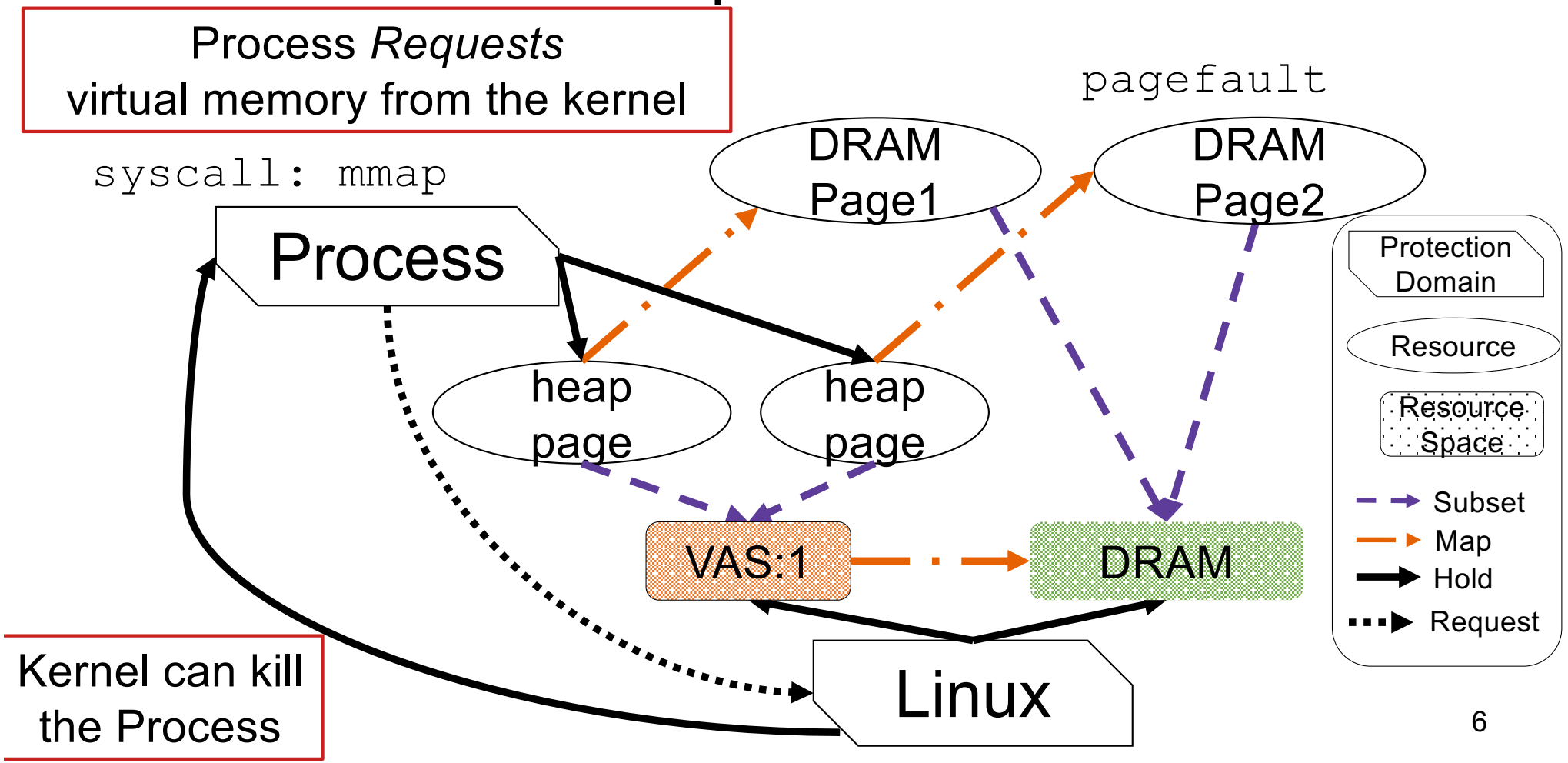
- A set of domains affected by a faulty or malicious domain
- Typically of interest to infrastructure providers

Contributions



OSmosis Model Graph of a Linux Process

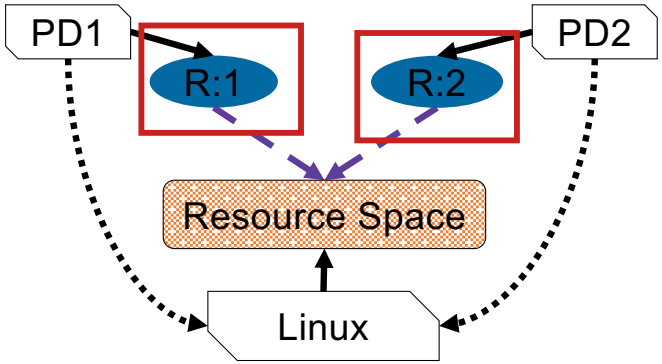
Process *Requests*
virtual memory from the kernel



Kernel can kill the Process

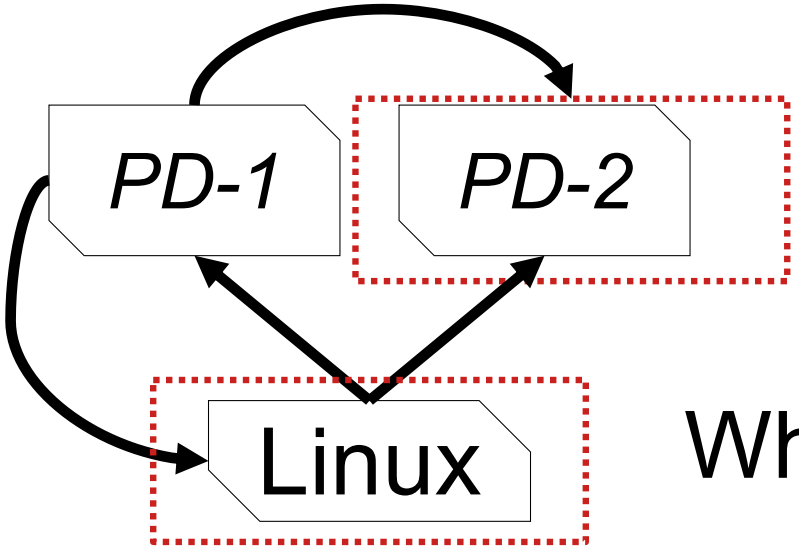
Finding Dependencies: Shared Resources

$$\textbf{SharedResources}(PD_x,\ \texttt{ResourceTypes},\ \texttt{AccessMode}) =$$
$$\{\ PD_i\ \mid\ PD_i \neq PD_x \wedge (\ BFS(PD_x,\ \{\texttt{hold},\ \texttt{map}\},\ \texttt{fwd},\ \texttt{ANY},\ \infty,\ \{\texttt{Resource}\},\ \texttt{ResourceTypes})$$
$$\cap$$
$$BFS(PD_i,\ \{\texttt{hold},\ \texttt{map}\},\ \texttt{fwd},\ \texttt{AccessMode},\ \infty,\ \{\texttt{Resource}\},\ \texttt{ResourceTypes})$$
$$\neq \emptyset\ )\ \}$$



Finding Dependencies: PD Control

Who can kill *PD-1*?



Who can be killed by *PD-1*?

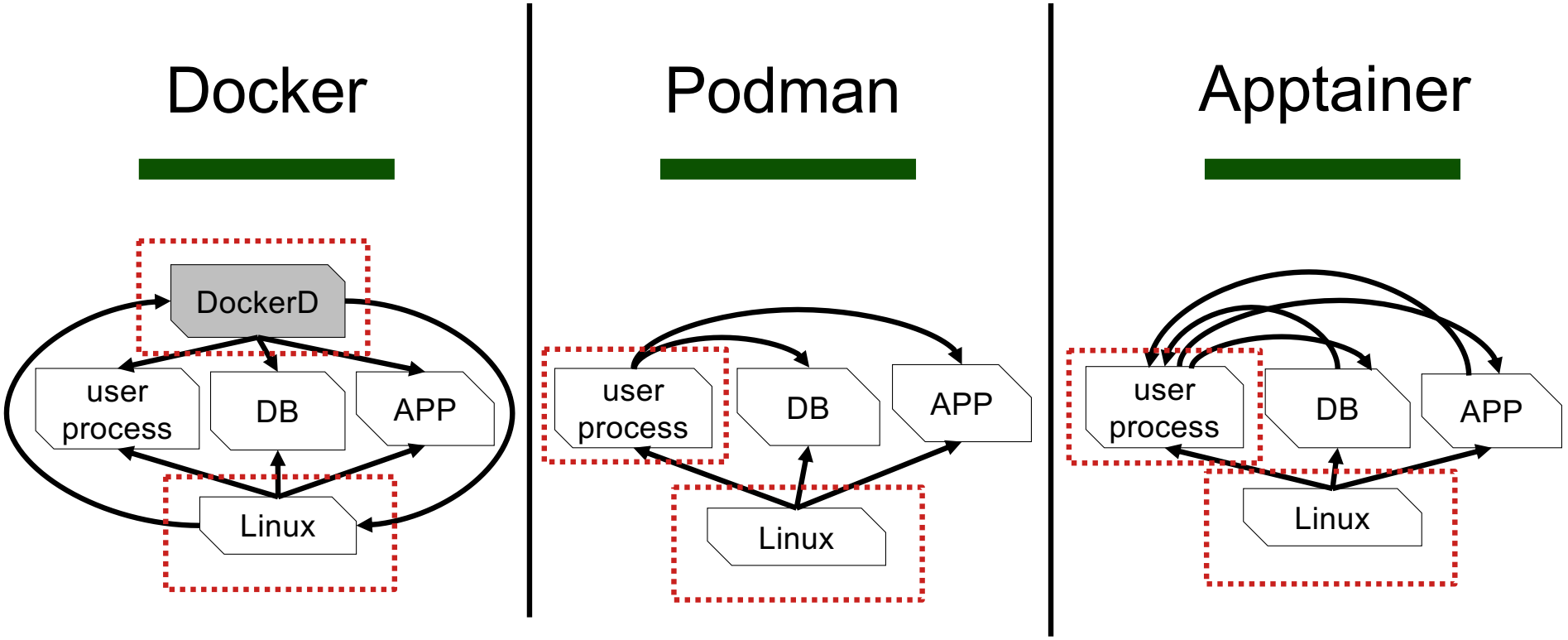
Precise TCB Definition as a Graph Query

$$\mathbf{TCB}(PD_x) = \texttt{SharedResources}(PD_x) \cup \texttt{CanControl}(PD_x)$$

Evaluation

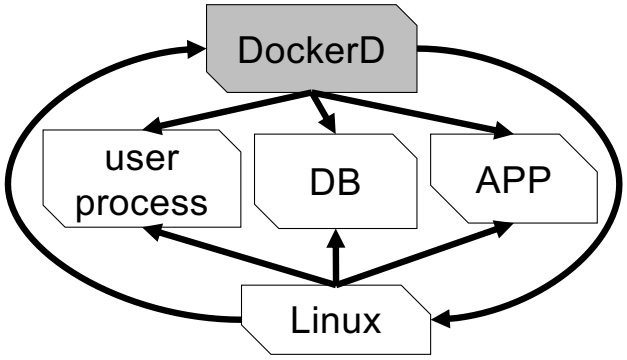
How does *TCB/IB* change from

Docker → Podman → Apptainer



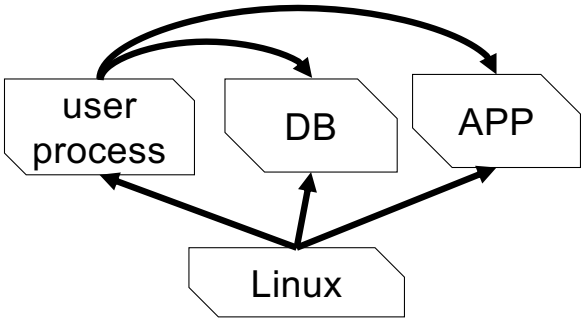
What is the *Trusted Computing Base* of the DB Process?

Docker



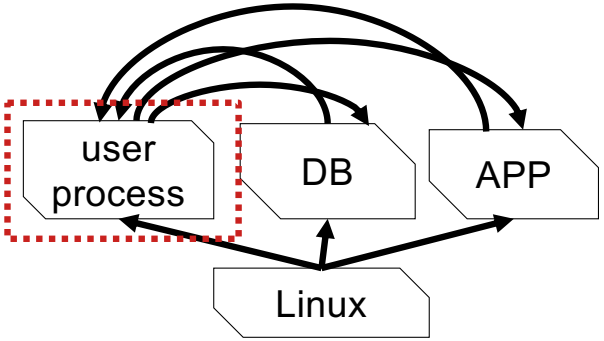
Ø

Podman



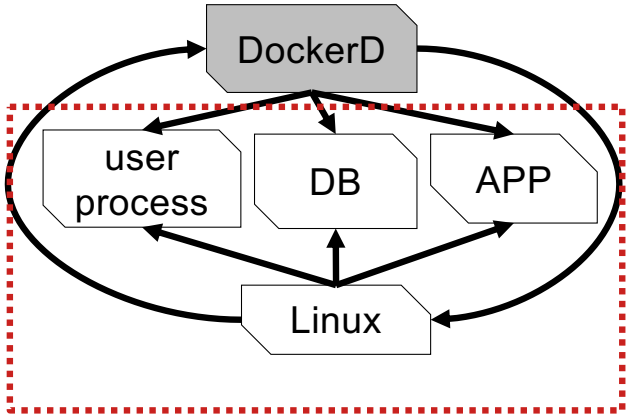
Ø

Apptainer

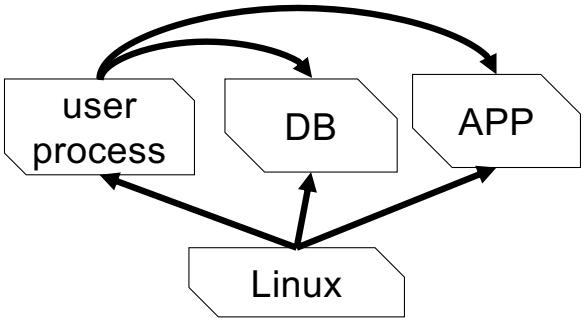


What is the Impact Boundary of the DB Process?

Docker

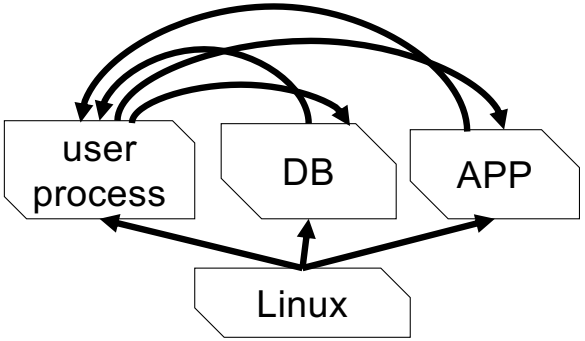


Podman



∅

Apptainer



∅

What is the *Impact Boundary* of the Docker Daemon?

Getting OSmosis State of a Linux System

- Linux has no central place to track resources.

- Our Approach
  - Gobble whatever info available in `/proc`
    - `/proc/pid/maps`: Address Space Layout
    - `/proc/pid/pagemap`: Virtual Memory to Physical Memory mapping
    - `/proc/pid/status`: Signal Masks, User ID, Namespace Information

- **Lintool** is our Python script that reads `/proc` and generates the graph

Key Takeaways

- **OSmosis model** enables precise isolation comparison
- **Lintool** enables the OSmosis model extraction from Linux
- We leverage both to compare container technologies
- Refer to the paper for additional examples and queries

Paper