Implementing Secure Isolated Containers in an Operating System Kernel

Aleksandar Andrejevic

About *Monolithium*

- A new monolithic hobby operating system project, built with simplicity in mind
- Written entirely from scratch
- Designed to be modular, keeping the balance between the complexity of each module, and the complexity of communication between them
- Still very incomplete, though it has a working kernel

Why Simplicity?

- Less code means less bugs
- Simple code is easier to analyze, leading to improved bug discovery
- With proper modularization, a simple program can also be full-featured and equipped to solve a wide variety of problems

Resource Management Through Objects

- “Everything is an object”
- Resource management modules depend on a common module for object management, which handles:
  - Reference counting
  - Retrival by name or other criteria
  - Creation and disposal
- Objects are grouped in namespaces

Kernel-based Virtualization

- Object namespaces are almost completely isolated
  - Only objects in the current namespace are seen
  - Namespaces can have an associated filesystem path prefix
- One namespace = one container
- It is possible to nest namespaces, and therefore run containers within containers

Kernel-based Virtualization

- The root user of the root namespace can access all objects in the system
- The root user of every other namespace can only access all the objects in that namespace as well as all its nested namespaces
- Proper virtualization requires a very fine-grained object access security system

Implementation

- The proof of concept of virtualization using object namespaces will be implemented in Monolithium
- Certain other operating system kernels also manage their resources through objects
  - It may be possible to apply the same mechanism to implement containers for some of them

Conclusions and Future Work

- Kernels that use object-based resource management could implement kernel-level virtualization in a simple and clean manner
- The greatest obstacle to implementing this idea is the incompleteness of *Monolithium*.