



Kodak: Snapshots for Distributed Systems

Nathan Burow, Purdue University

Project Mentors: Kasimir Gabert / 5638; Adam Vail / 5634

Problem Statement:

As distributed systems become larger, more prevalent, and more complicated, it has become impractical to study them by hand in an ad hoc manner. Kodak aims to solve this by providing an automated distributed system analysis platform. It provides a novel snapshot primitive that preserves network packets and other I/O, while scaling to thousands of nodes.

Existing snapshot techniques are slow (several seconds) and memory intensive (several gigabytes). This summer I worked on a novel KVM snapshot technique to overcome these problems.

Objective and Approach:

- Snapshot by saving the current page table, and marking the pages Copy on Write (COW)
 - Map process memory layout to specific pages
- Understand Virtio drivers to ensure complete and correct snapshots
 - Any ongoing I/O can be restored to exact prior state when reverting to a snapshot
- Ensure integrity of VM hard disk by using a RAM based file system that is similarly snapshotted

Impact and Benefits:

Kodak allows researchers to understand the complex interactions of a distributed system. Its snapshot primitive allows scientific testing of distributed systems. Analysts can fork the system and examine its behavior under different scenarios.

Results:

- Traced QEMU / KVM migrate command, determined structure of migrate file.
 - Instrumented writes to verify file structure
- Created a Kernel Module to log which pages are used during migration
 - Using this log across multiple VM states led to understanding the page table, and will enable our migration technique
- Determined Linux's TMPFS supports our snapshot technique
 - Ran a QEMU / KVM VM with the virtual hard disk in the TMPFS partition
 - Created a snapshot of the file system

