HEC FSIO Workshop, Arlington, VA, August 9-12, 2009

Automatic Identification of I/O Bottleneck and Run-time Optimization for Cluster Virtualization

Xubin He

hexb@tntech.edu

Tennessee Tech UNIVERSITY

Stephen L. Scott

scottsl@ornl.gov

OAK RIDGE National Laboratory

Motivation

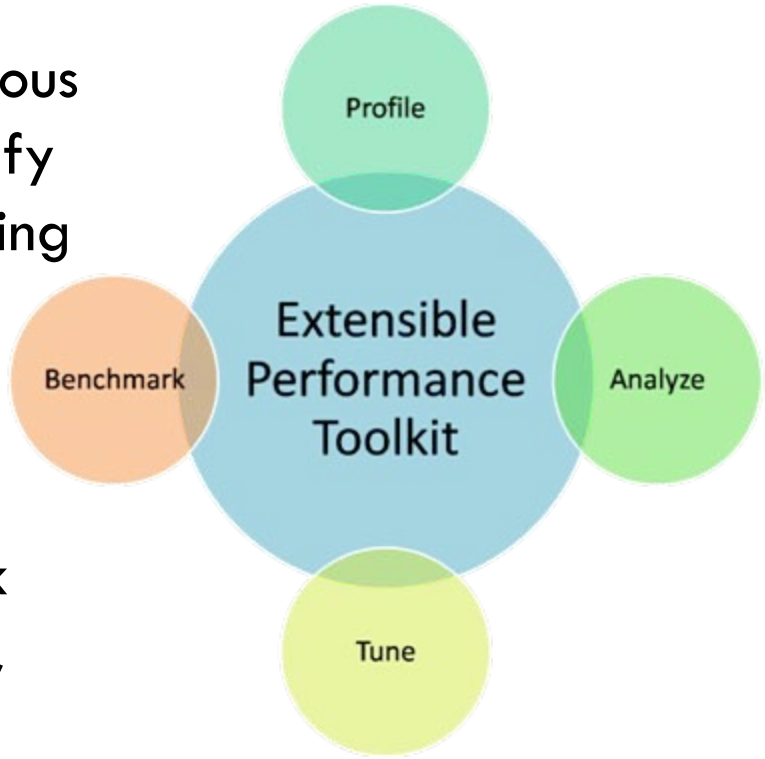
- Extending virtualization technology into high-performance, cluster platforms generates exciting new possibilities (runtime environments tailored for user/app, live migration, dynamic job consolidation, efficient hw utilization…).
- The question remains, however, whether the benefit of a virtualized cluster environment outweighs the potential hazards of additional overhead.
- This concern is especially critical for I/O-bound HPC applications, which have been notoriously difficult to make efficient within virtual machines.

- *To ensure applications continue to run efficiently on these virtualized platforms, it is critical to systematically and routinely identify and correct bottlenecks and optimize the resource utilization*.

Objective

- Our research objective is develop fundamental techniques for virtual clusters that not only facilitate rigorous performance studies, but also identify places where performance is suffering and then optimize system to lessen impact of such bottleneck.

- Specific goals: automatic bottleneck identification, run-time optimization, extensible framework.



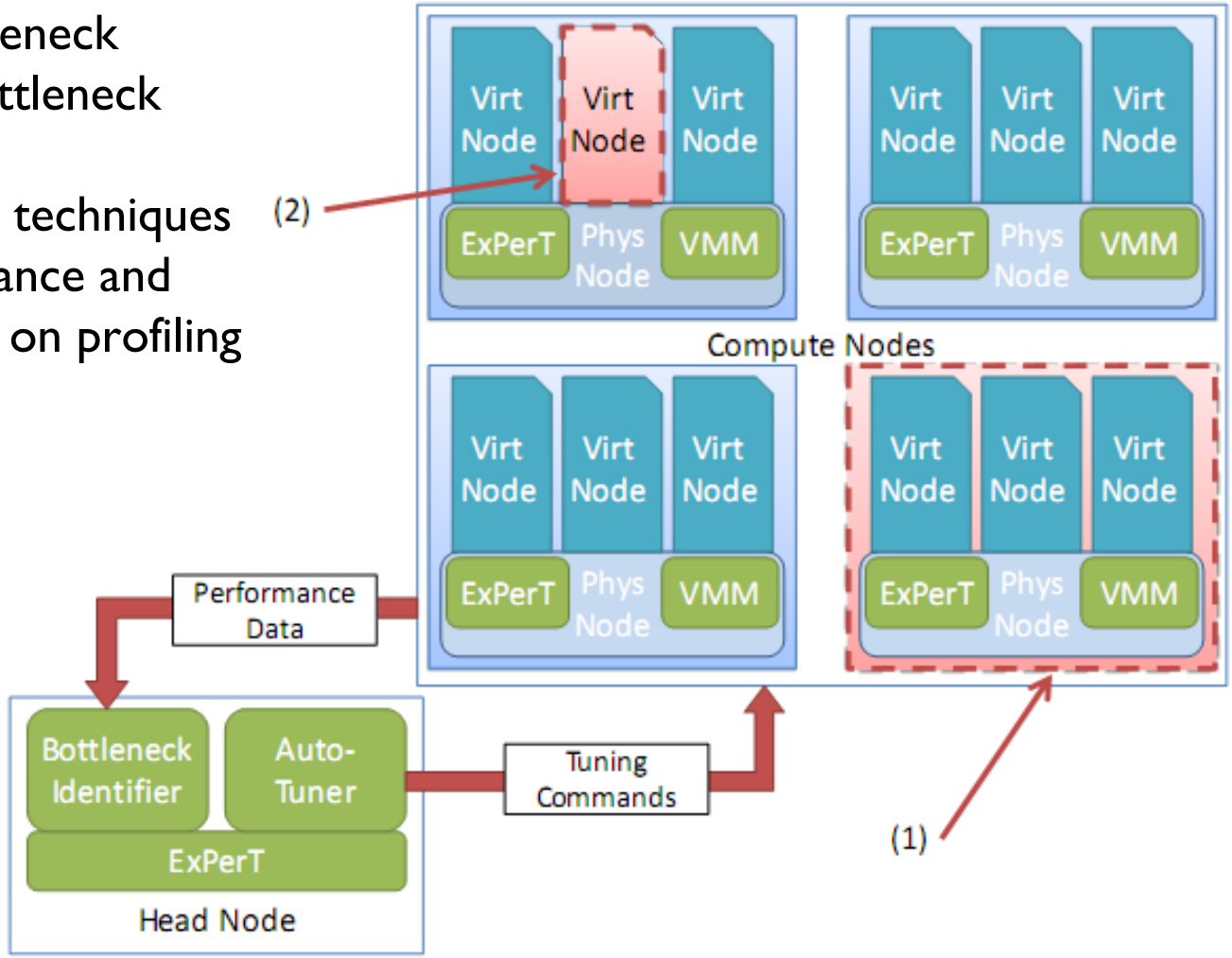
Automatic Bottleneck Identification

- Bottleneck:
  - Over-utilization of a specific system resource resulting in the underutilization of all other system resources.
- Hands-free approach utilizing clustering techniques from past measurements to automatically yield performance profile for different, common types of workloads.
- Given the standard performance profile, a statistically strong deviation from the nearest neighbor would indicate a performance anomaly.

Tuning and Load Balancing on a Virtual Cluster

(1) Physical node bottleneck
(2) Virtual machine bottleneck

We will investigate techniques to achieve load balance and optimization based on profiling information



Runtime Optimization

- We can approach self-optimization in two ways: post-mortem and at run-time.

- Run-time optimization techniques
  - Live migration/load balancing
    - Process-level
    - Virtual machine level
  - Dynamic parameter tuning
    - Changing VMM properties (RAM, # cores, etc.), configurable driver parameters…

Extensible Framework

- Coordinated benchmarking and profiling
- Transparently deploy distributed profiling jobs
- Accommodate distributed virtual systems
- Allow easy extensibility to incorporate existing data gathering tools (eg. Sysstat, vmstat) and profiling tools (eg. Oprofile).
- Plug in in to existing cluster frameworks (OSCAR)
- Provide common data collection interface
- Provide data analysis and presentation tools

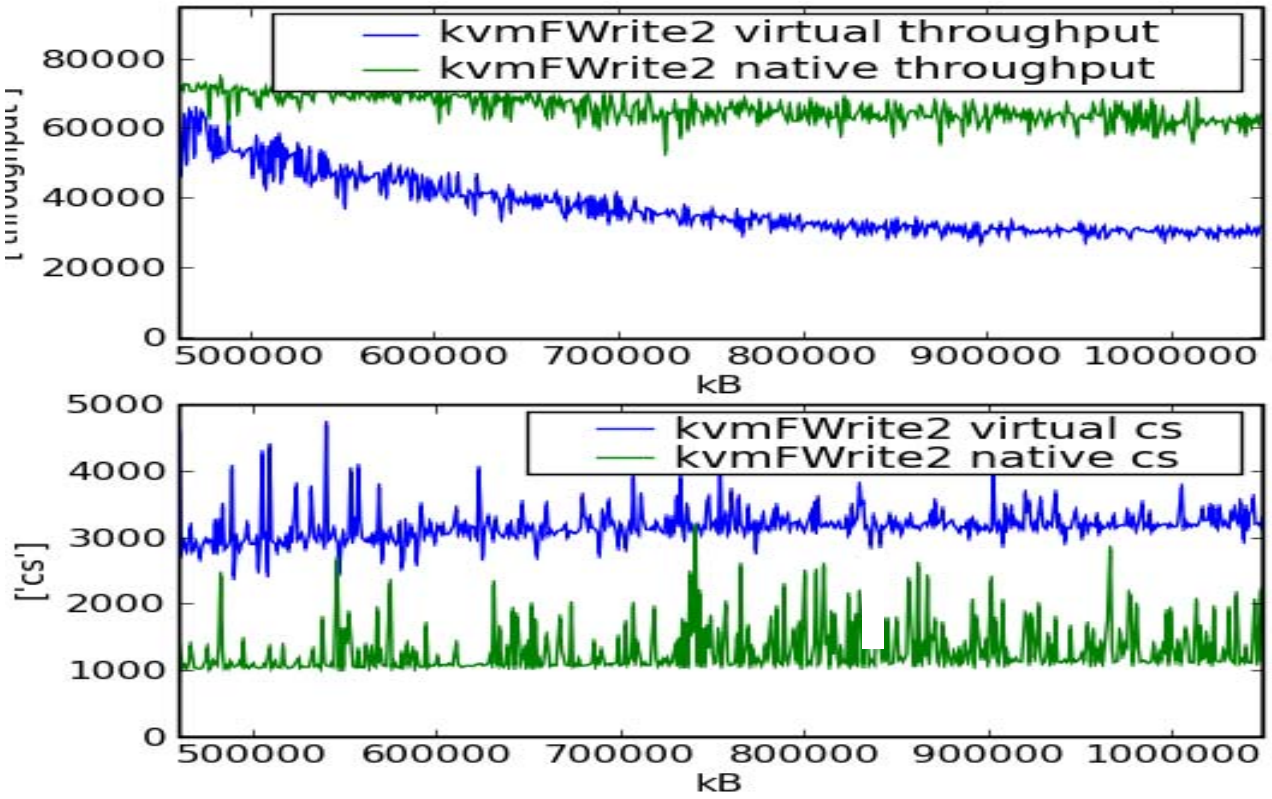
Case Study: Virtual Disk I/O for KVM [EuroPar'2009]

- I/O efficiency in virtualized environments, especially with respect to disk I/O, remains little understood and hardly tested.

- To show the promise of the framework, we conducted some preliminary studies on understanding the virtual I/O behaviors of KVM.

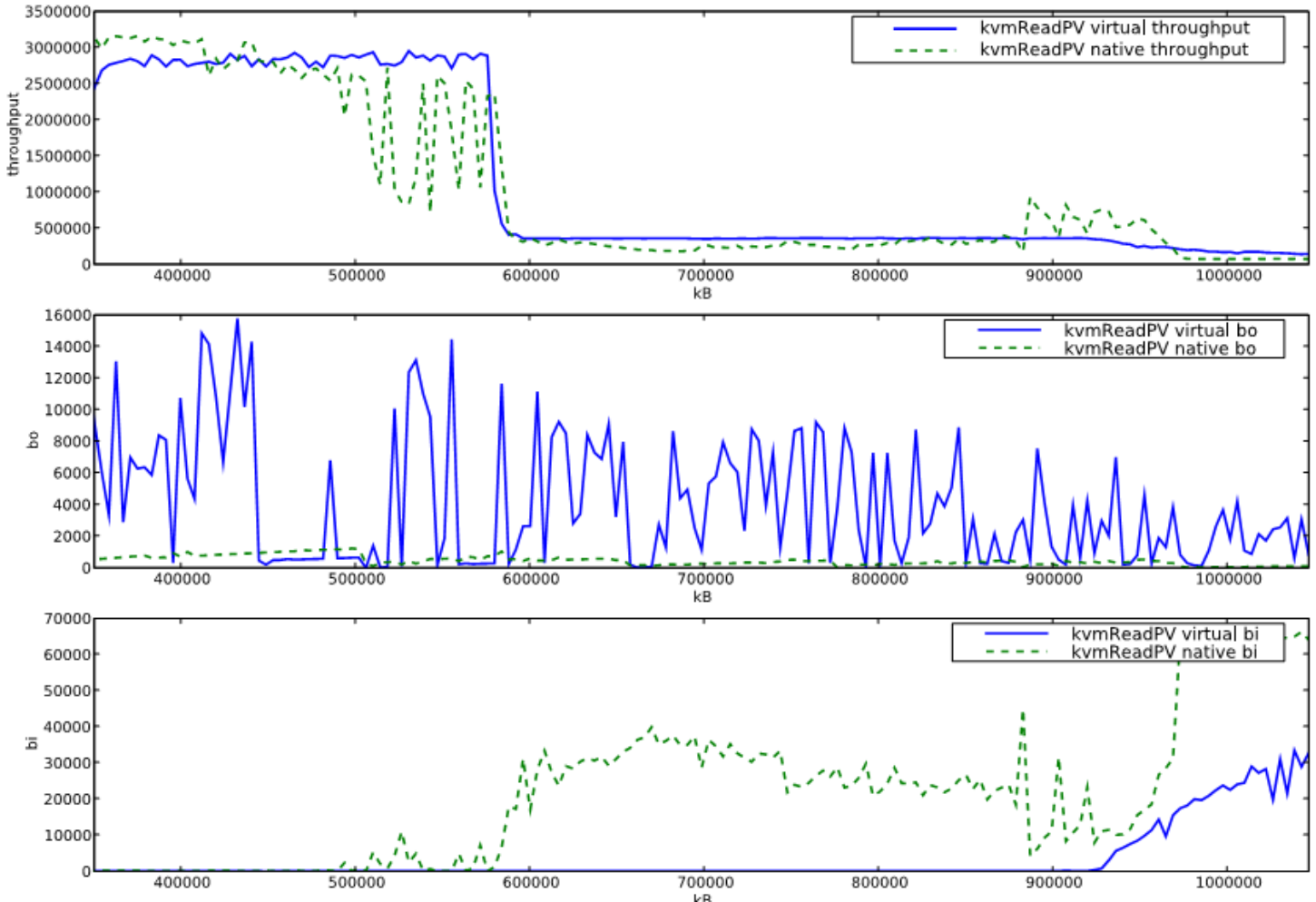
Context Switch

Using Iozone as our Bench and Vmstat as our *Profile*, we conducted tests while varying file sizes.



Disk I/O activities

- Virtualized caching behavior vs. “Bare Metal”



Research Tasks

- Comprehensive analysis of I/O in cluster virtualization (paravirtulaized I/O drivers, overprovision, impacts of different I/O scheduling and caching techniques…).

- Develop techniques to achieve:
  - Automatic bottleneck identification
  - Run-time self optimization

- Integrate above techniques into an extensible performance analysis framework, which will:
  - Collect and parse through large amounts of benchmarking data
  - Easily distribute profilers among remote compute nodes
  - Automate the bottleneck identification process
  - Support dynamic optimization and load-balancing

HEC FSIO Workshop, Arlington, VA, August 9-12, 2009

Automatic Identification of I/O Bottleneck and Run-time Optimization for Cluster Virtualization

Xubin He

hexb@tntech.edu

Tennessee Tech UNIVERSITY

Stephen L. Scott

scottsl@ornl.gov

OAK RIDGE National Laboratory