Evaluation of the HPC Challenge Benchmarks in Virtualized Environments

Vince Weaver

ICL Lunch Talk

8 July 2011

VHPC'11 Paper

6th Workshop on Virtualization in High-Performance Cloud Computing

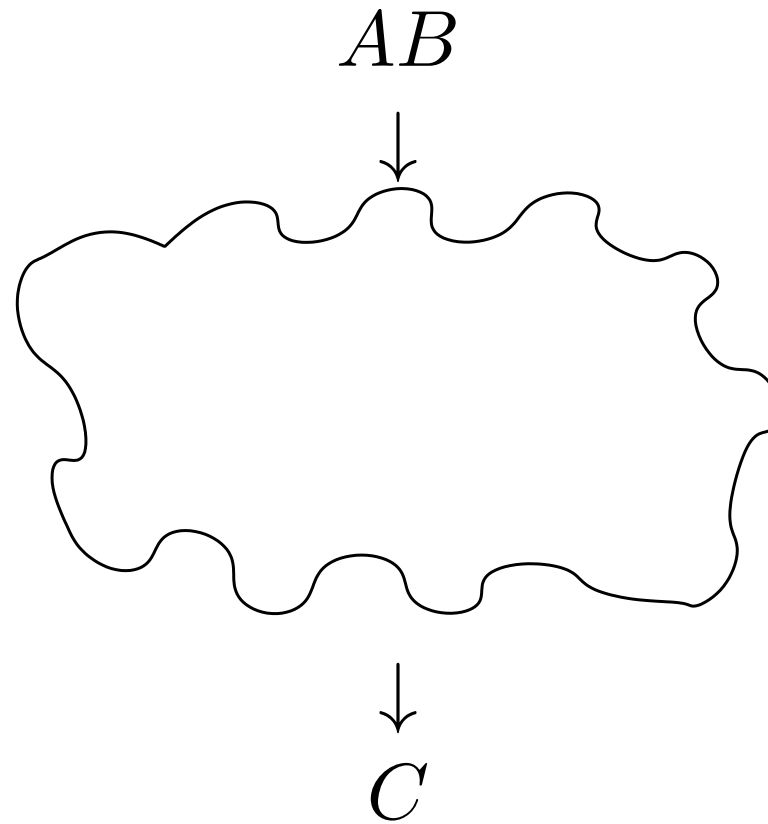Piotr Luszczek, Eric Meek, Shirley Moore, Dan Terpstra, Vince Weaver, Jack Dongarra

Traditional HPC

$$AB$$
$$\downarrow$$

$$\downarrow$$
$$C$$

Cloud-based HPC

AB

$\downarrow$

$\downarrow$

C

Cloud Tradeoffs

Pros

- No AC bill
- No electricity bill
- No need to spend $$$ on infrastructure

Cons

- Unexpected outages
- Data held hostage
- Infrastructure not designed for HPC

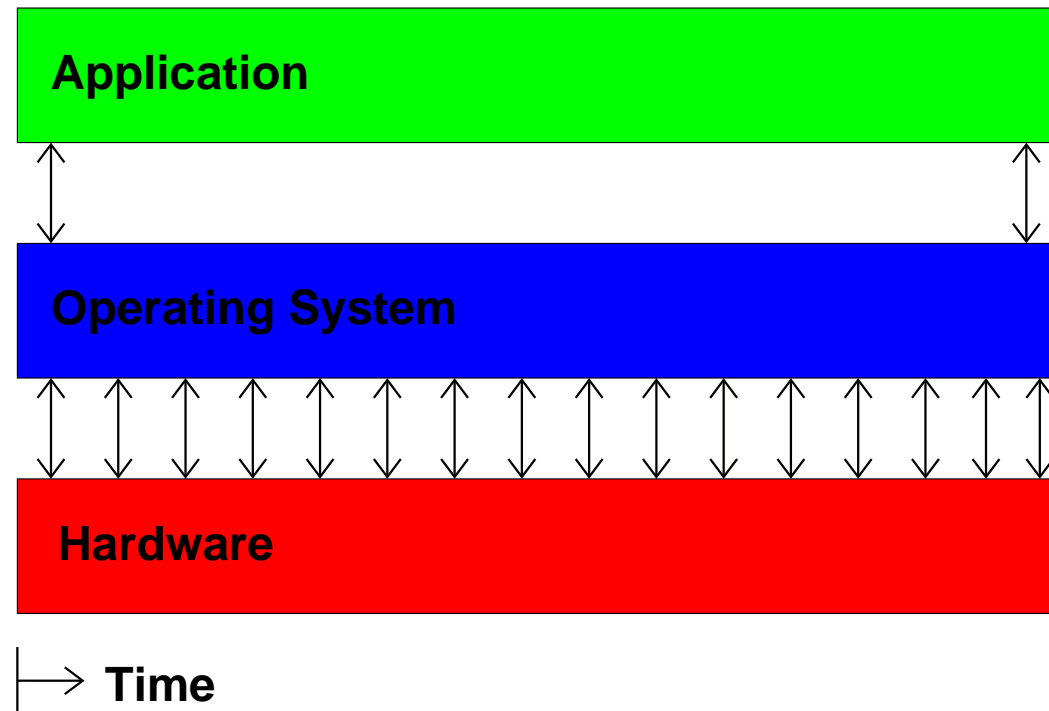# Measuring Performance in the Cloud
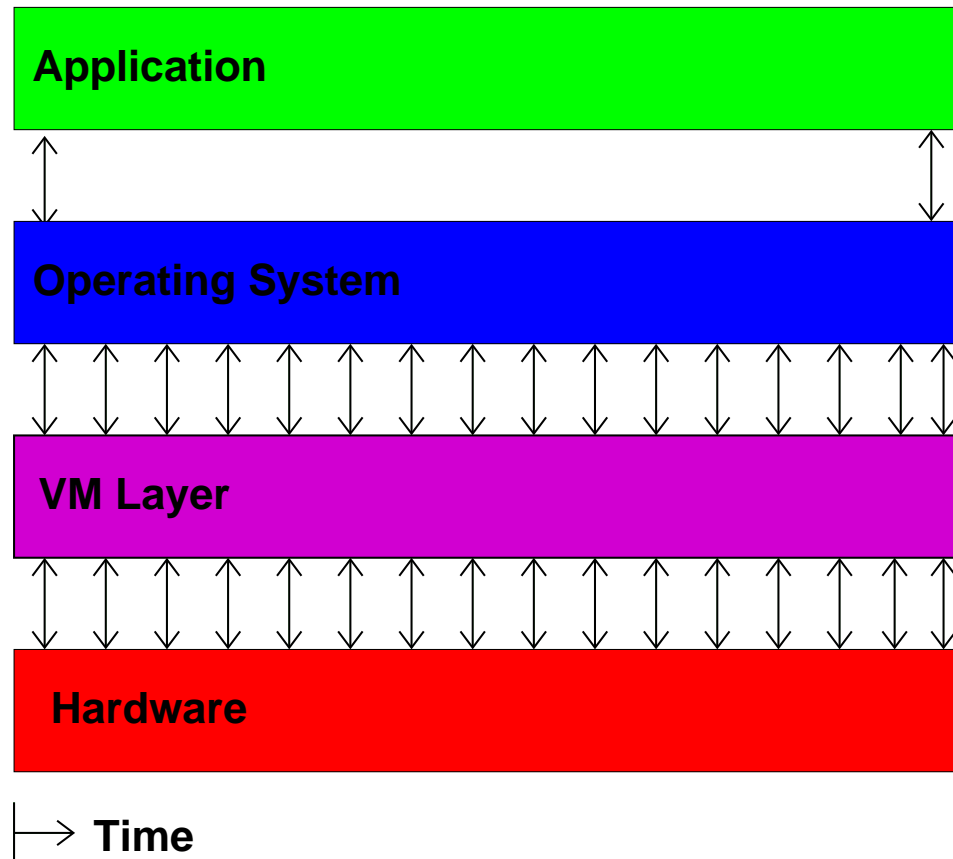
First let's just measure runtime

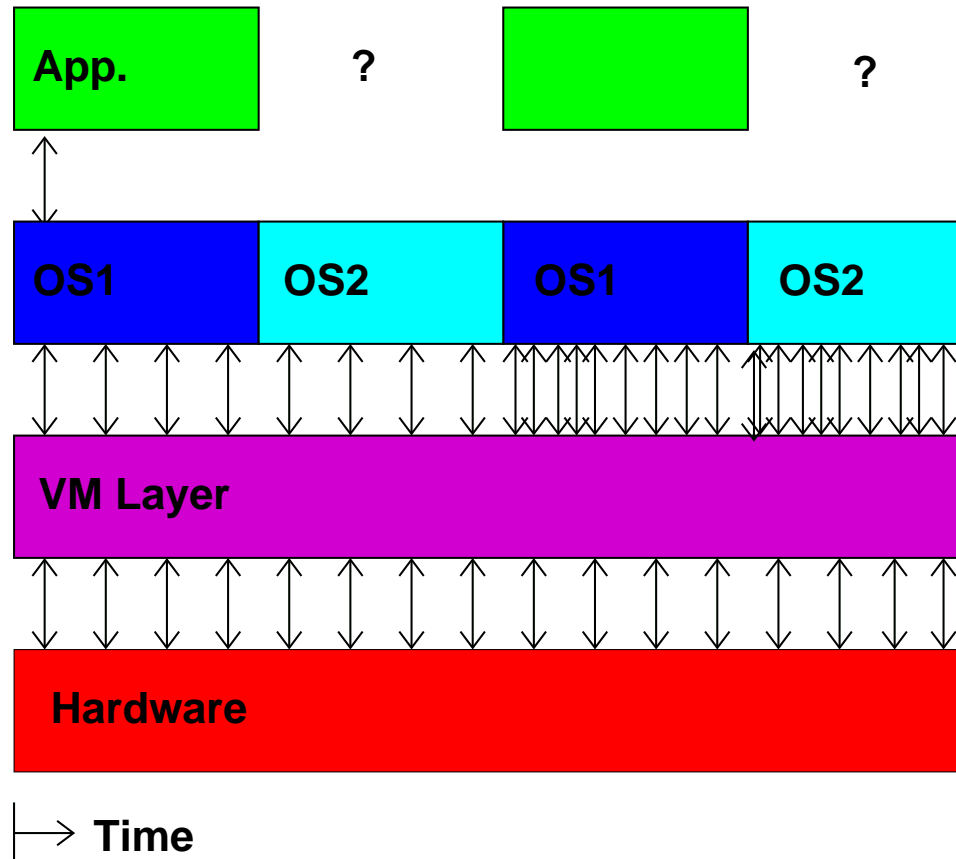This is difficult because in virtualized environments

Time Loses All Meaning

Simplified Model of Time Measurement

Then the VM gets involved

Then you have multiple VMs

So What Can We Do?

Hope we have exclusive access and measure wall-clock time.

Measuring Time Externally

- Ideally have local hardware access, root, and hooks into the VM system

- Otherwise, you can sit there with a watch

- Danciu et al. send UDP packet to remote server

- Most of these are not possible in a true "cloud" setup

Measuring Time From Within Guest

- Use `gettimeofday()` or `clock_gettime()`
- This might be the only interface we have
- How bad can it be?
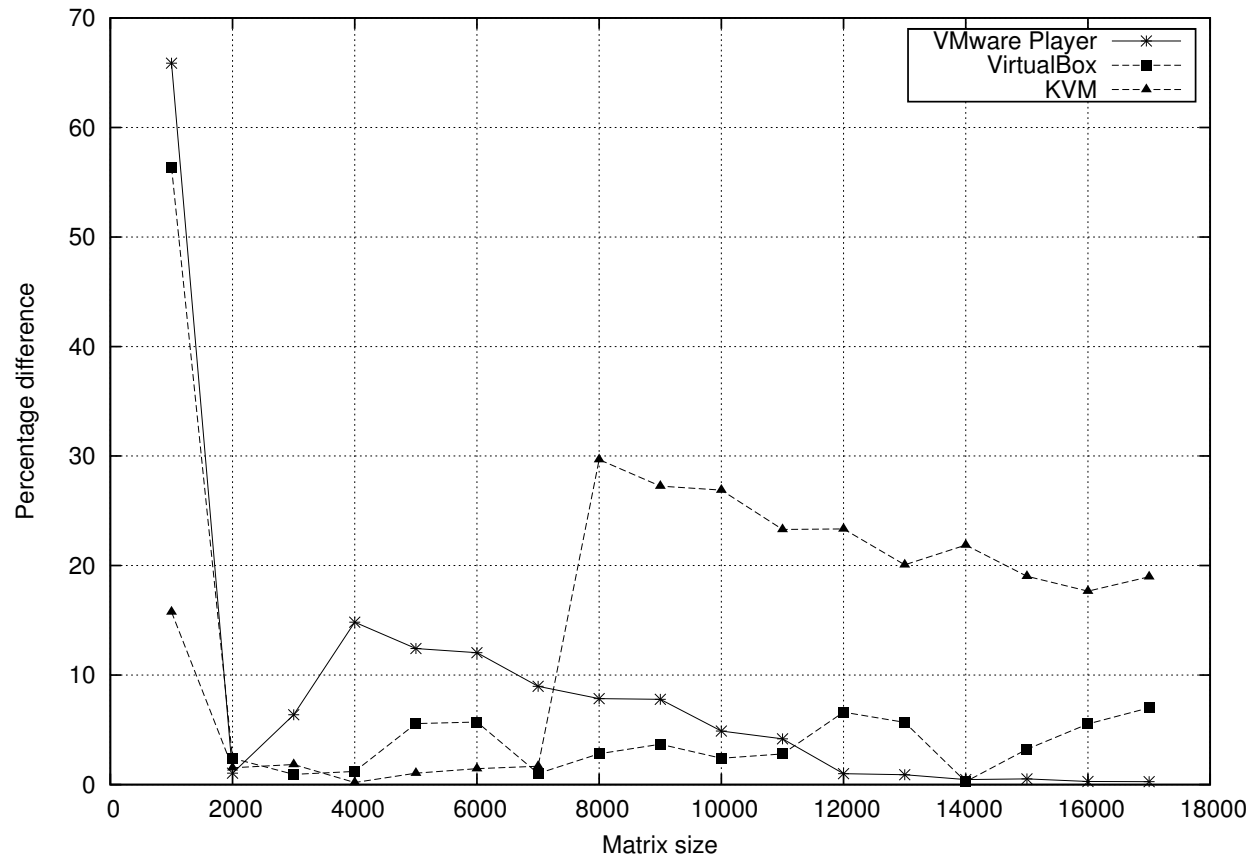
Our Experimental Setup

- 8-core Core i7, (dual 4-core 2.93GHz Xeon X5570)
- VMware Player 3.1.4, VirtualBox 4.0.8, KVM 2.6.35
- HPC Challenge Benchmarks, Open MPI
- Time measured by `gettimeofday()` invoked by `MPI_Wtime()`

Accuracy Drift

- Typical development model is to re-run app over and over again with slight changes while monitoring performance

- In virtualized environment, factors inherent in the virtualization might change runtime run to run more than any optimization tuning

Ascending vs Descending – HPL
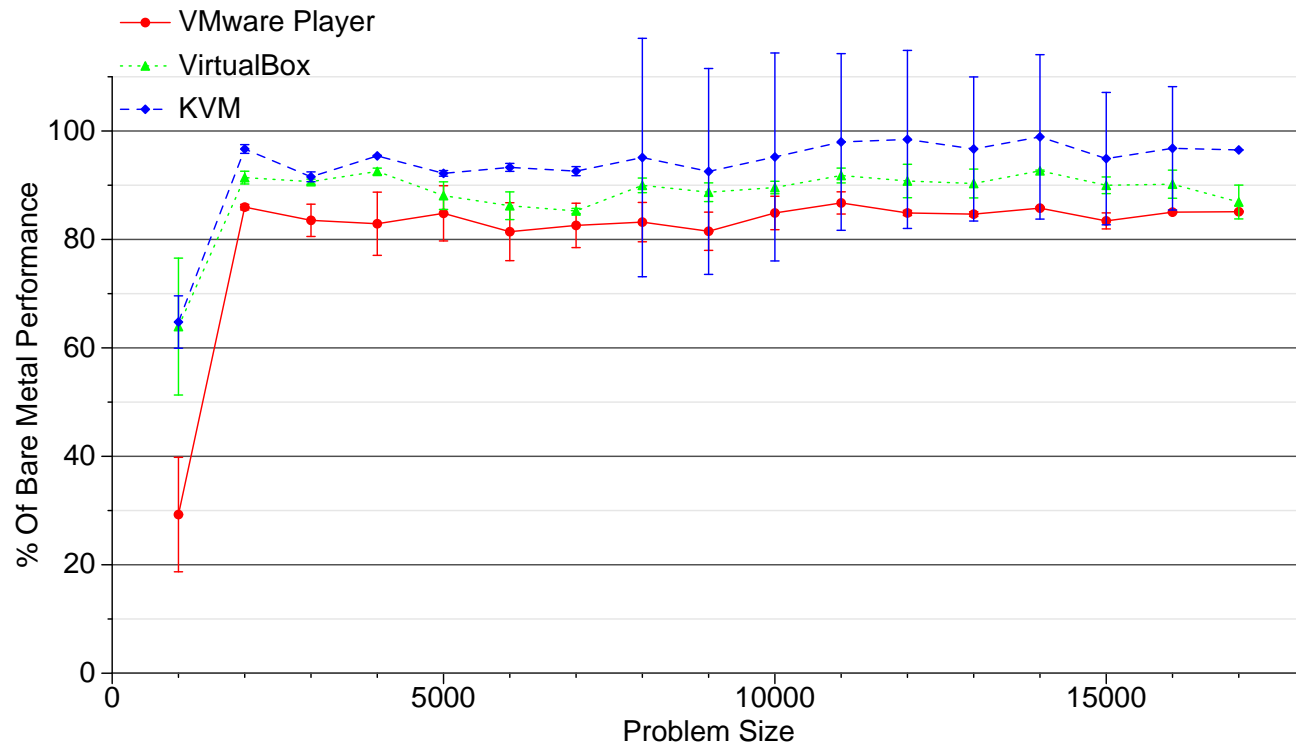
Bare metal showed no difference
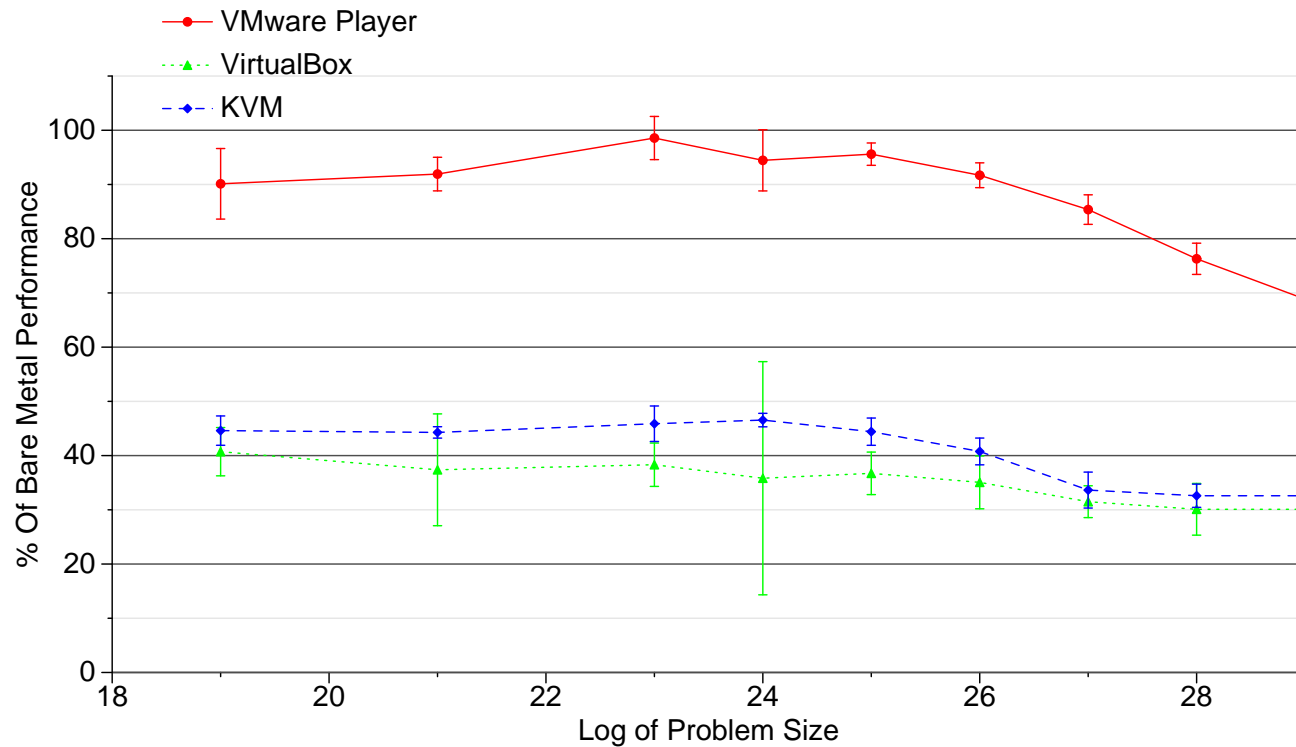
Performance Results

We use a relative metric, defined as:

$$\frac{\text{performance}_{\text{VM}}}{\text{performance}_{\text{bare metal}}} \times 100\%$$

HPL – Low OS/Communication Overhead

MPIRandomAccess – High OS/Communication Overhead

Conclusion

- Virtualization exacerbates the existing problem of accurate performance measurement

- Different workloads can stress the VM layer in drastically different ways

- Extra care needs to be taken to generate repeatable results

Future Work

- Validate internal time measurements with external ones

- More analysis of sources of VM overhead

- Performance of larger systems with off-node network activity

Future Work – PAPI-V

- "Improved" timer support. Direct wall-clock access?
- Virtualized performance counters
- Components for the virtualized hardware:
  Network Interfaces, etc.

Questions?

vweaver1@eecs.utk.edu