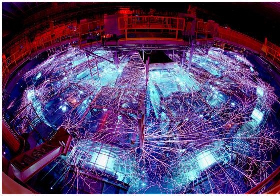
Exceptional service in the national interest

Sandia National Laboratories



Lessons Learned from 10k Experiments to Compare Virtual and Physical Testbeds

Jonathan Crussell, Thomas M Kroeger, David Kavaler, Aaron Brown, Cynthia Phillips

August 12th, 2019

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Goals

Discover where and how virtual and physical testbeds differ

- “Virtualization artifacts”

Goals

Discover where and how virtual and physical testbeds differ

- "Virtualization artifacts"

Methodology:

- Run representative workloads on physical and virtual testbeds
- Collect, compare, and contrast metrics
  - Application-, OS-, and network-level

Lessons Learned

Methodology and experimental results presented in previous work:

Jonathan Crussell, Thomas M. Kroeger, Aaron Brown, and Cynthia Phillips. Virtually the same: Comparing physical and virtual testbeds. In 2019 International Conference on Computing, Networking and Communications (ICNC), pages 847–853, Feb 2019.

This work focuses on lessons learned in four areas:

- Tool Validation and Development
- Instrumentation
- Data Collection and Aggregation
- Data Analysis

Methodology & Results

ApacheBench fetching fixed pages from an HTTP server

- Over 10,000 experiments across three clusters
- Over 500TB of data (without full packet captures)
- Varied payload size, network drivers, network bandwidth
- Large variations in network driver *offloading* behavior
- Near-identical sequences of system calls

Leverage *minimega* toolset (see `http://minimega.org` for details)

Tool Validation and Development

Lesson: Using a testbed toolset for experimentation requires substantial effort and consideration to put tools together in a workable and validated fashion.

Tool Validation and Development

```
bash$ bash run.bash
USAGE: run.bash DIR ITER DURATION CONCURRENT VMTYPE
    DRIVER NCPUS OFFLOAD RATE NWORKERS URL NREQUESTS
    INSTRUMENT
bash$ bash sweep.bash /scratch/output params.bash >
    sweep-params.bash
bash$ head -n 1 sweep-params.bash
bash /root/run.bash /scratch/output/ 1 360 1 kvm e1000 1
     on 1000 1 http://10.0.0.1/ 100000 true
bash$ sort -R sweep-params.bash | parallel -j1 --eta -S
    $(python igor.py --heads jc[0-9])
```

Tool Validation and Development

Running thousands of repetitions:

- Handle all edge cases (rare bug in *minimega*'s capture API)
- Clean up all state (failed to unmount container filesystem)

```
bash$ mount | grep megamount | head -n 5
megamount_5562 on /tmp/minimega/5562/fs type overlay
megamount_5566 on /tmp/minimega/5566/fs type overlay
megamount_5611 on /tmp/minimega/5611/fs type overlay
megamount_5752 on /tmp/minimega/5752/fs type overlay
megamount_5774 on /tmp/minimega/5774/fs type overlay
bash$ mount | grep megamount | wc -l
96
```

Tool Validation and Development

Toolset improvements:

- Add snaplen option to capture API
- Add standalone C2 server

```
# On VMs
minimega -e cc exec mkdir /que
minimega -e cc background protonuke -serve -http
minimega -e cc recv /miniccc.log

# On physical nodes
rond -e exec mkdir /que
rond -e bg protonuke -serve -http
rond -e recv /miniccc.log
```

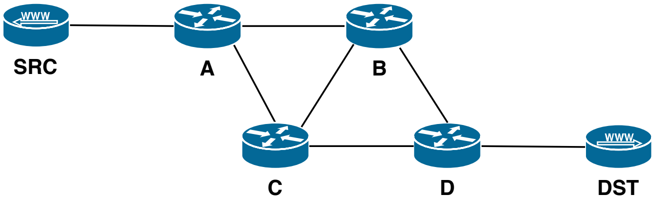
Instrumentation

Lesson: Instrumentation is invaluable but it is often manually added, expensive, and experiment-specific. Integrating more forms of instrumentation into the toolset could help researchers to more rapidly develop experiments.

Instrumentation

Two forms of instrumentation:

- Verify functionality of environment
- Understand and evaluate experiments



Integrating the former into the toolset could simplify experiments

Instrumentation

- Mismatch between capture locations
- Dropped events for containers

```
bash$ tcpdump -i eth0 -w foo.pcap
tcpdump: listening on eth0, link-type EN10MB (Ethernet),
     capture size 262144 bytes
9001 packets captured
9000 packet received by filter
1 packets dropped by kernel
bash$ sysdig -w foo.scap
<no indication of dropped events>
```

Instrumentation

Many ways to instrument experiment at many levels, mostly by hand

- Application-level: RPS, Jitter, ...
- OS-level: System calls, utilization, perf stats, ...
- Network-level: Flow statistics, bandwidth, ...

Use to understand anomalies

- e1000 stalls
- Performance increase

Data Collection and Aggregation

Sandia National Laboratories

Lesson: Testbeds can provide a wealth of data to researchers but should do more to streamline the process of collecting and aggregating it into a usable form.

Data Collection and Aggregation

How to extract instrumentation data from VMs?

- C2 has limits on file size
- VMs write to qcow2, host extracts
- Future: mount guest filesystem?

How to aggregate data?

- Dump per-iteration data into SQLite database
- Combine SQLite databases after all iterations complete

Data Analysis

Sandia National Laboratories

How to reduce storage?

```
total packets:              5        total packets:              5
ack pkts sent:              4        ack pkts sent:              5
pure acks sent:             2        pure acks sent:             2
sack pkts sent:             0        sack pkts sent:             0
dsack pkts sent:            0        dsack pkts sent:            0
max sack blks/ack:          0        max sack blks/ack:          0
unique bytes sent:         72        unique bytes sent:        486
actual data pkts:           1        actual data pkts:           1
actual data bytes:         72        actual data bytes:        486
rexmt data pkts:            0        rexmt data pkts:            0
rexmt data bytes:           0        rexmt data bytes:           0
zwnd probe pkts:            0        zwnd probe pkts:            0
zwnd probe bytes:           0        zwnd probe bytes:           0
outoforder pkts:            0        outoforder pkts:            0
pushed data pkts:           1        pushed data pkts:           1
SYN/FIN pkts sent:        1/1        SYN/FIN pkts sent:        1/1
req 1323 ws/ts:           Y/Y        req 1323 ws/ts:           Y/Y
adv wind scale:             7        adv wind scale:             7
========================== <15 lines omitted> ==========================
missed data:                0 bytes  missed data:                0 bytes
truncated data:             0 bytes  truncated data:           352 bytes
truncated packets:          0 pkts   truncated packets:          1 pkts
data xmit time:         0.000 secs   data xmit time:         0.000 secs
idletime max:             1.0 ms     idletime max:             0.7 ms
throughput:             38482 Bps    throughput:            259754 Bps
```

Data Analysis

Sandia National Laboratories

How to reduce storage?

- *tcptrace* produces 78 statistics per flow
- Compute summary statistics over all flows for iteration
- Compare mean of means across iterations and parameters

Data Analysis

Sandia National Laboratories

Lesson: Testbeds allow for many repetitions but care should be used when analyzing the data, especially in conflating statistical significance with practical importance.

Data Analysis

Hypothesis testing

- Everything seems significant with many iterations
- But practically important? (*e.g.* 0.1% decrease in latency)

Multiple comparisons

- Comparing many metrics can result in significance by chance

What's next?

Experiments with contention

- Run N client/server pairs on the same hardware
- Generates Nx the data
- Surprising performance **improvement** when N is small (<12)

Questions/Comments?

Lessons:

- Using a testbed toolset for experimentation requires substantial effort and consideration to put tools together in a workable and validated fashion.
- Instrumentation is invaluable but it is often manually added, expensive, and experiment-specific. Integrating more forms of instrumentation into the toolset could help researchers to more rapidly develop experiments.
- Testbeds can provide a wealth of data to researchers but should do more to streamline the process of collecting and aggregating it into a usable form.
- Testbeds allow for many repetitions but care should be used when analyzing the data, especially in conflating statistical significance with practical importance.

Presenter: Jonathan Crussell
jcrusse@sandia.gov