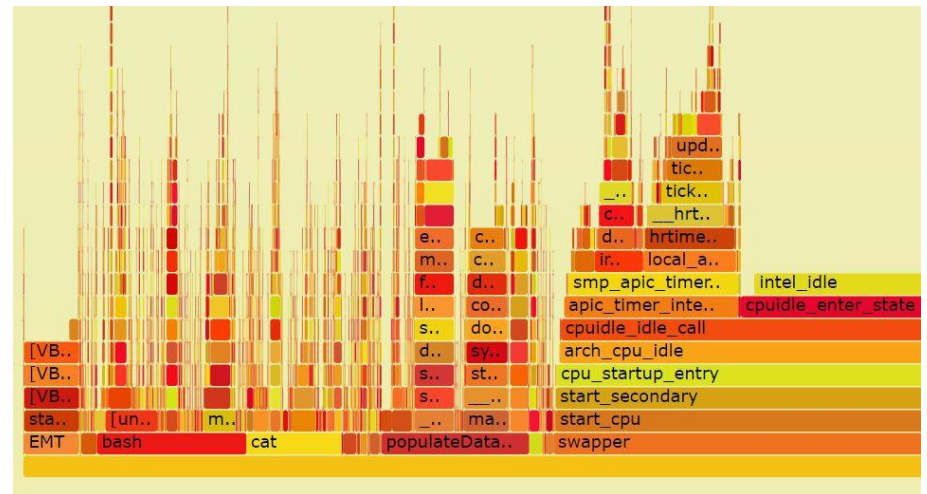




A Virtual Cluster Monitoring Toolkit for Bottleneck Analysis

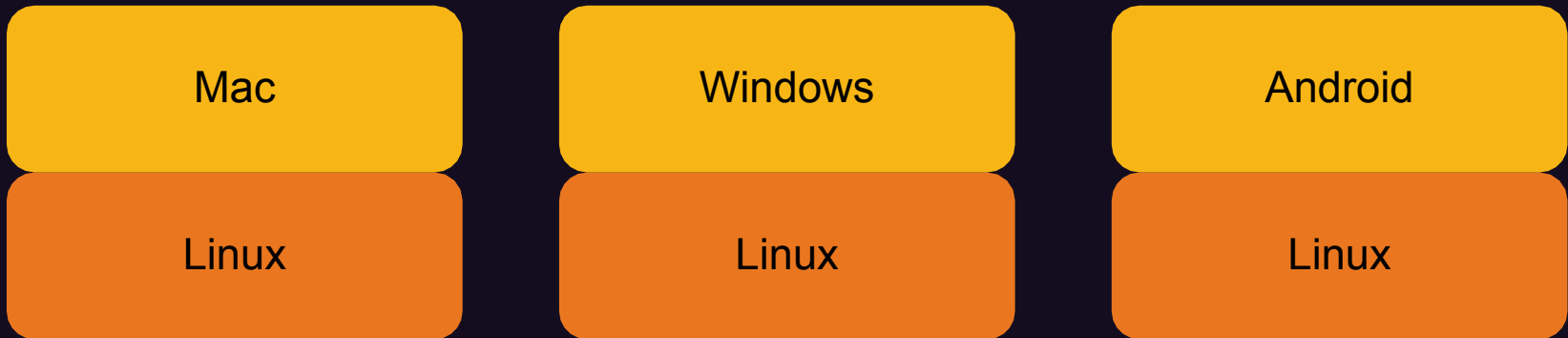
Natasha Frumkin and Christian Marquardt



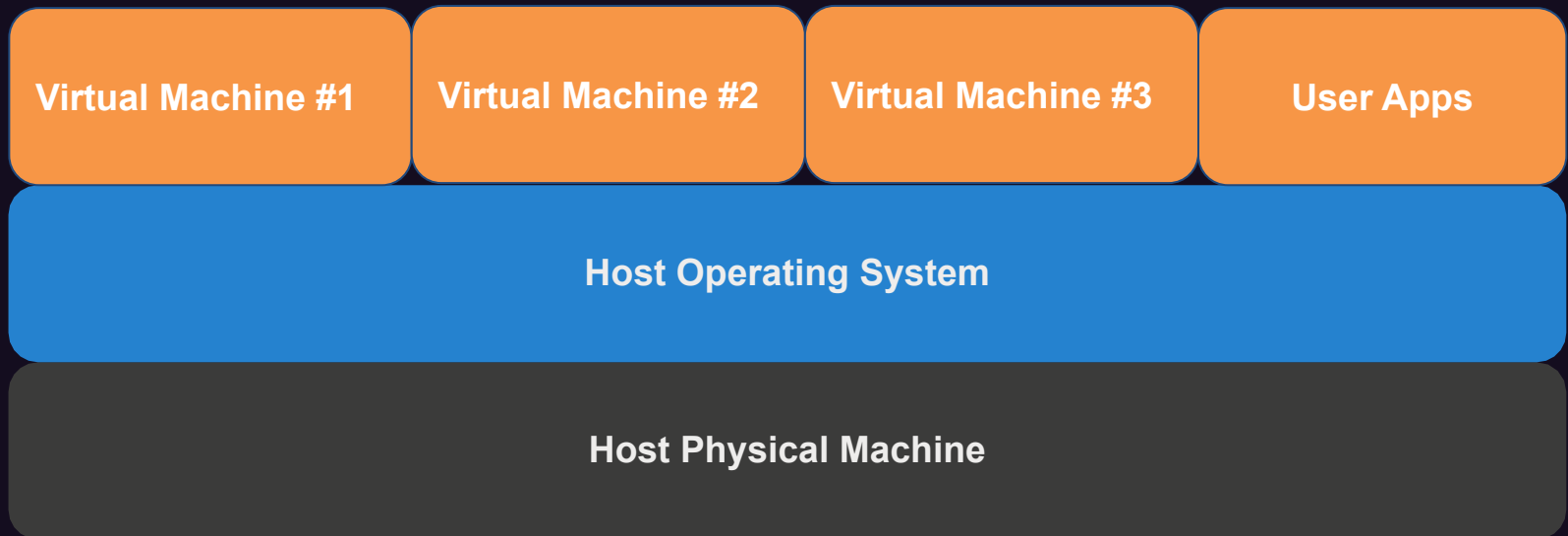
Host Call Stack for Virtual Node Provisioning

Virtualization:

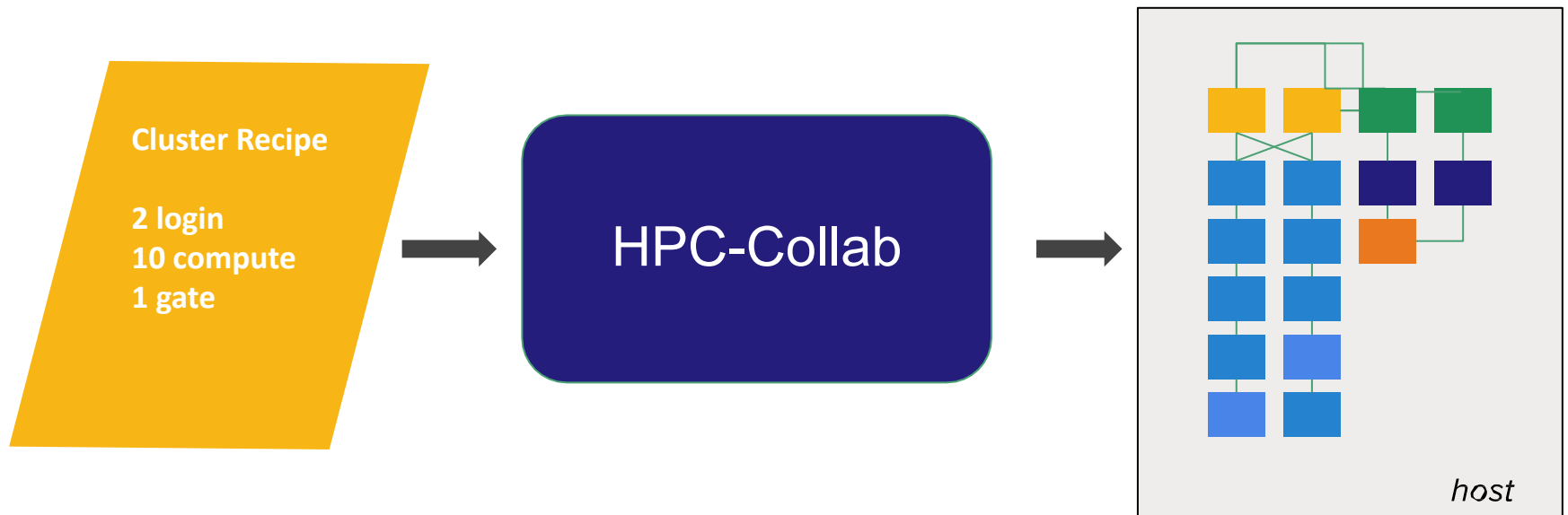
emulating one computer on another



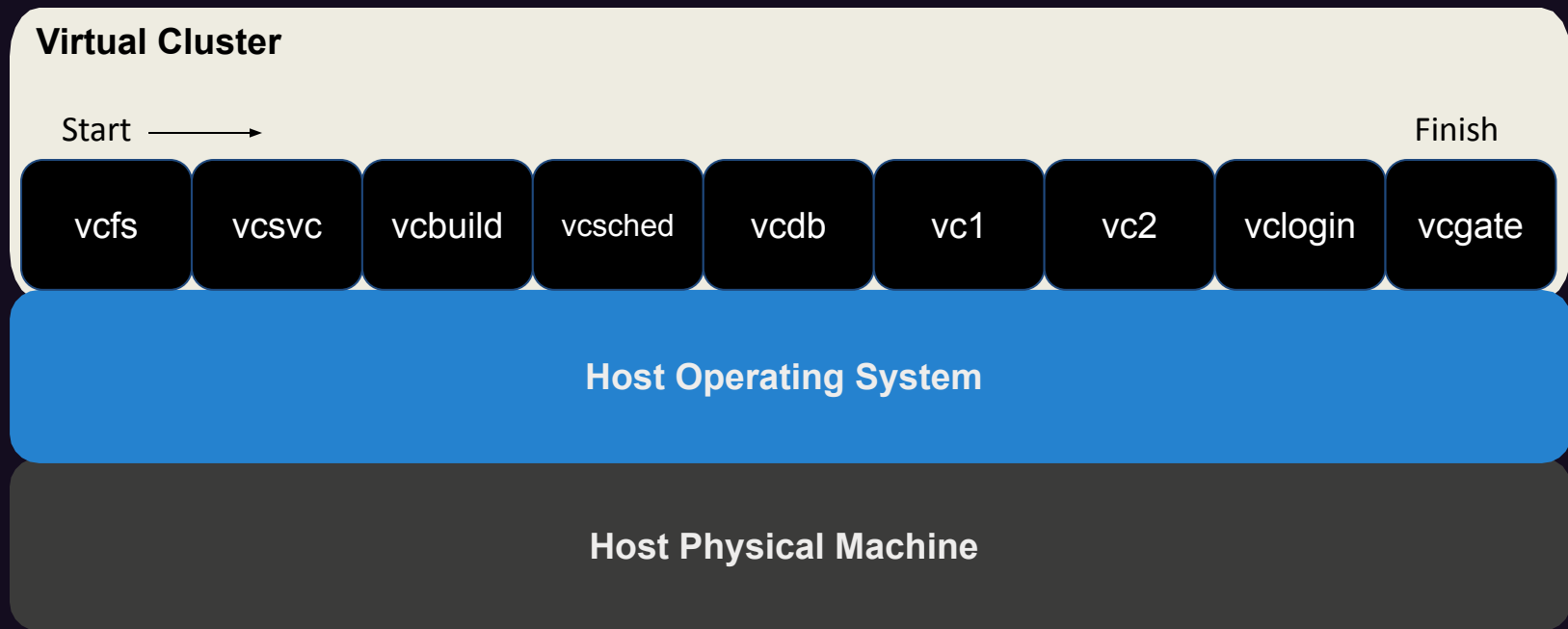
Virtualized HPC Clusters



HPC-Collab: a framework for designing virtual HPC clusters



Cluster Provisioning Timeline



Monitoring

- R/W to virtual disk
 - iostat
- Memory Usage
 - details in /proc/meminfo
 - vmstat
- CPU computation/power
 - perf
- Network performance
 - netstat
- Processor measuring
 - mpstat

```
te05:~$ iostat -x
Linux 3.10.0-1062.9.1.el7.x86_64 (te05)          07/27/2020      _x86_64_
avg-cpu:  %user   %nice %system %iowait  %steal   %idle
           0.36    0.04   4.06   0.05    0.00   95.48

Device:            rrqm/s   wrqm/s     r/s     w/s    rkB/s   wkB/s avgrq-sz
sda                  0.00     4.85     0.60    6.44   46.29  1135.90  336.16
```

*E.g. iostat
visualization*

```
te05:~/monitor/MonitorScripts$ cat /proc/meminfo
MemTotal:        65744940 kB
MemFree:         525388 kB
MemAvailable:   42037196 kB
```

*E.g. /proc/meminfo
visualization*

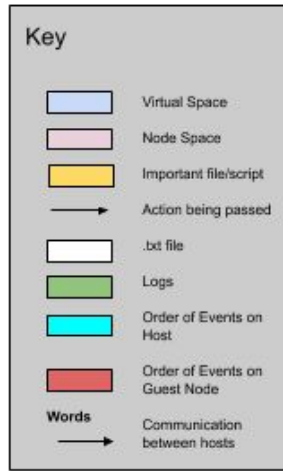
```
te05:~$ netstat -s | grep "total packets"
33624922 total packets received
te05:~$ netstat -s | grep "segments received" | grep -ve "bad"
33333616 segments received
te05:~$ netstat -s | grep "packets received" | grep -ve "total" |
256213 packets received
```

*E.g. Total, UDP, and TCP
packets and segments
received visualization*



How is our monitoring being done?

Control Flow Diagram for Monitoring Data on our Virtualized Cluster



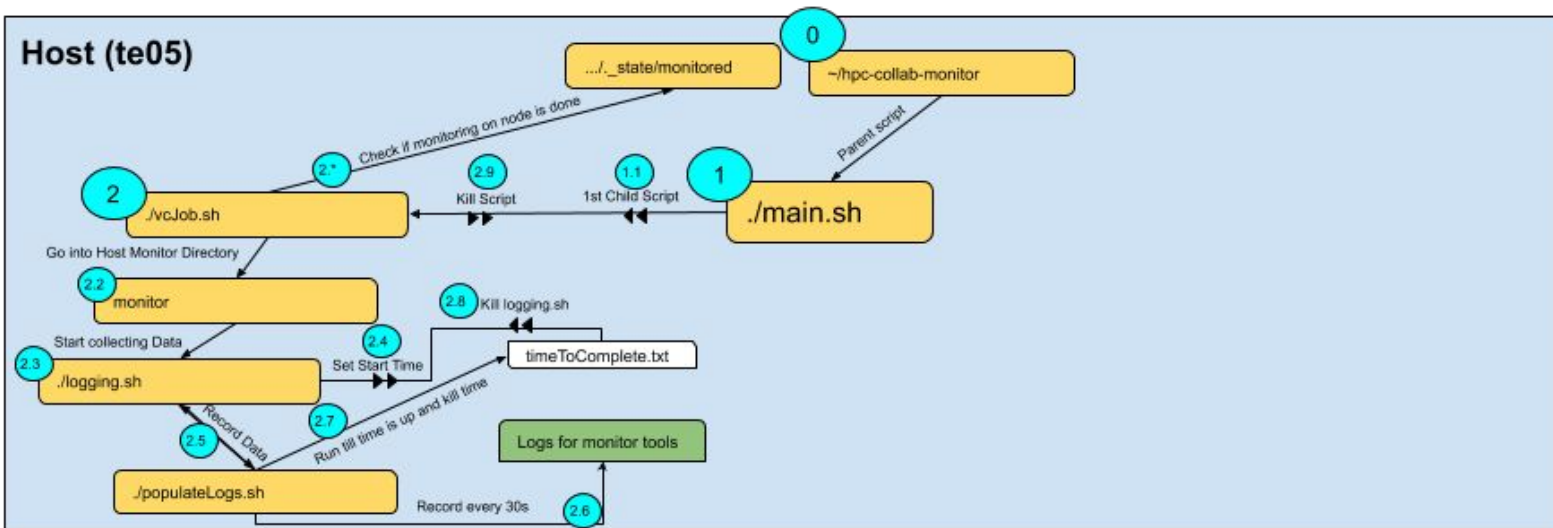
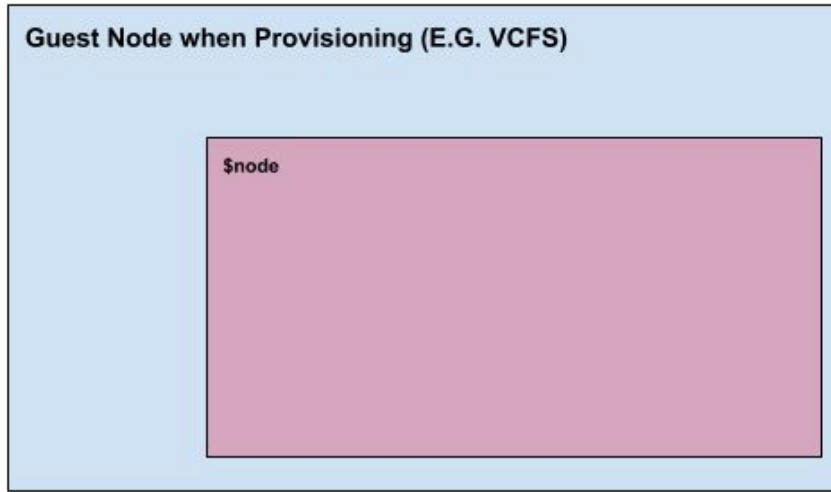
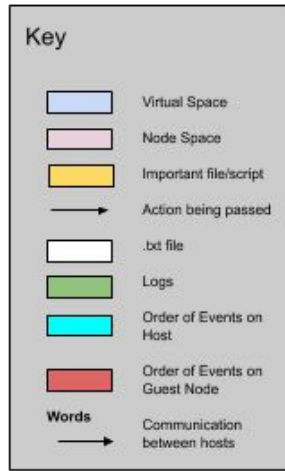
Guest Node when Provisioning (E.G. VCFS)

Host (te05)



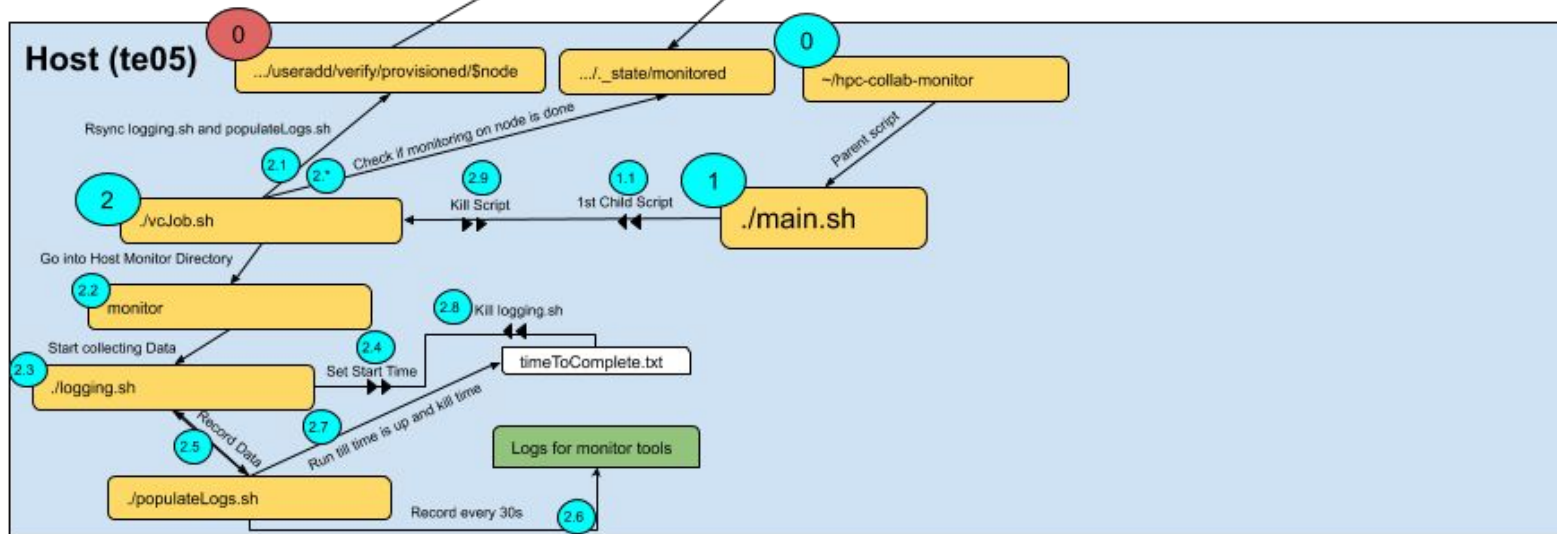
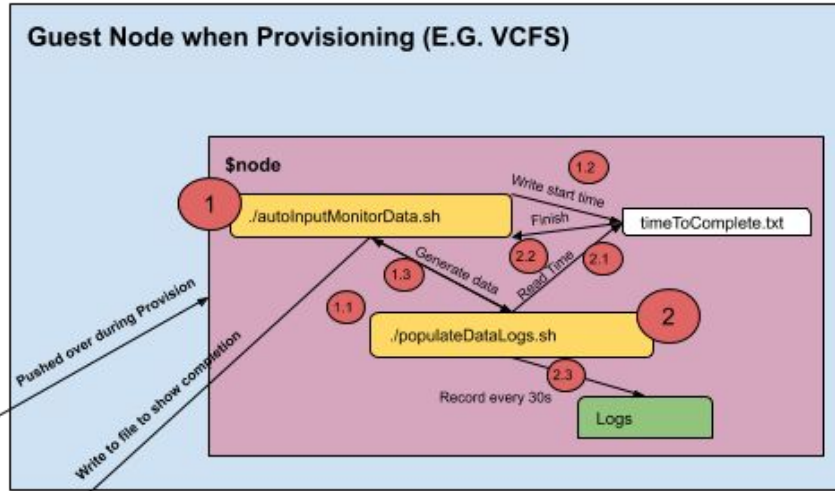
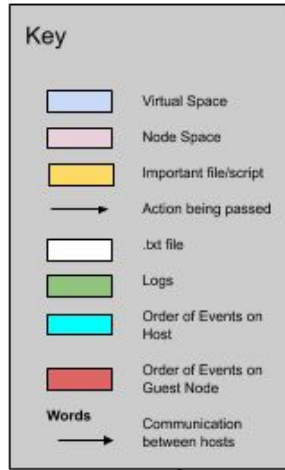
1. Start Provisioning and Monitoring

Control Flow Diagram for Monitoring Data on our Virtualized Cluster



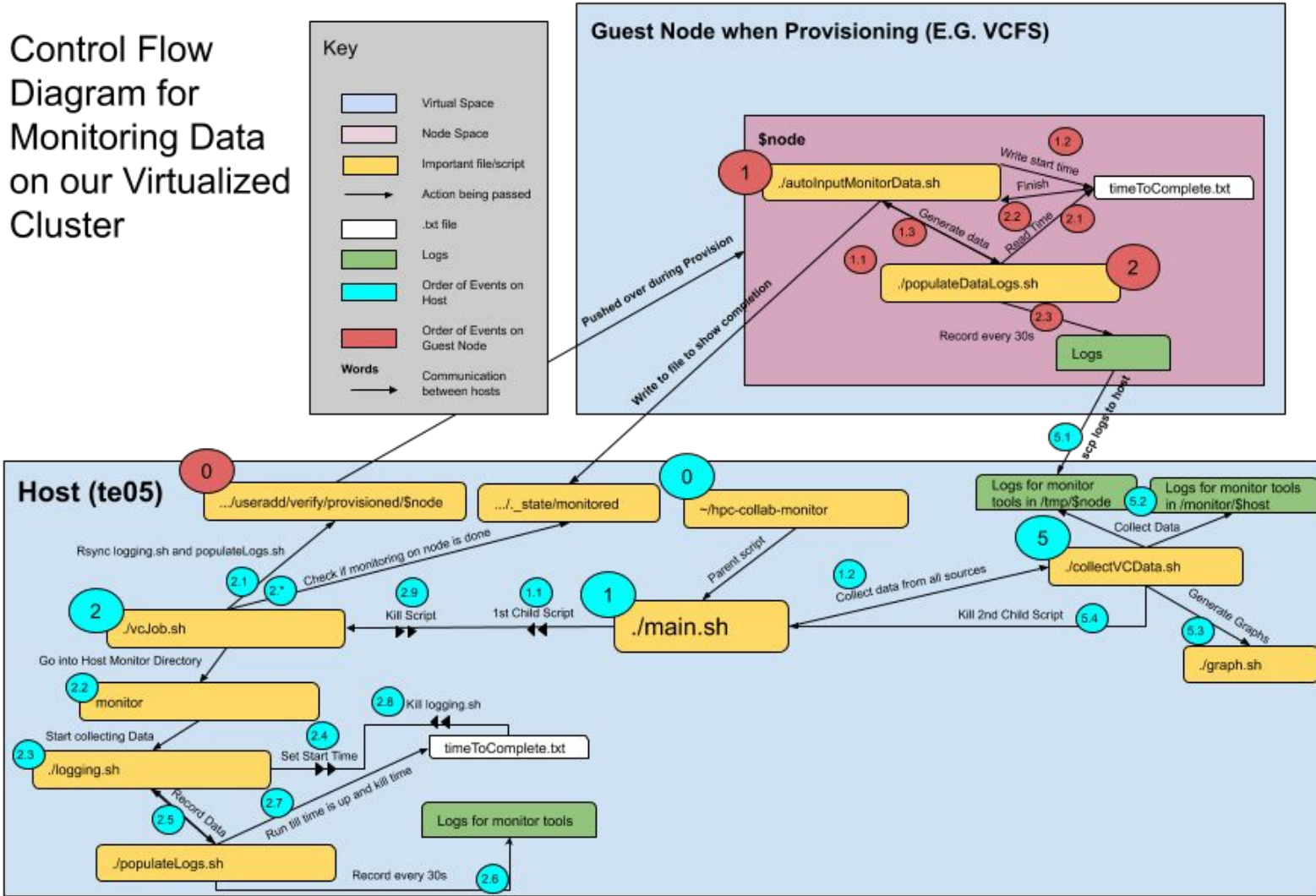
2. Rsync and Monitor on Virtual Nodes

Control Flow Diagram for Monitoring Data on our Virtualized Cluster



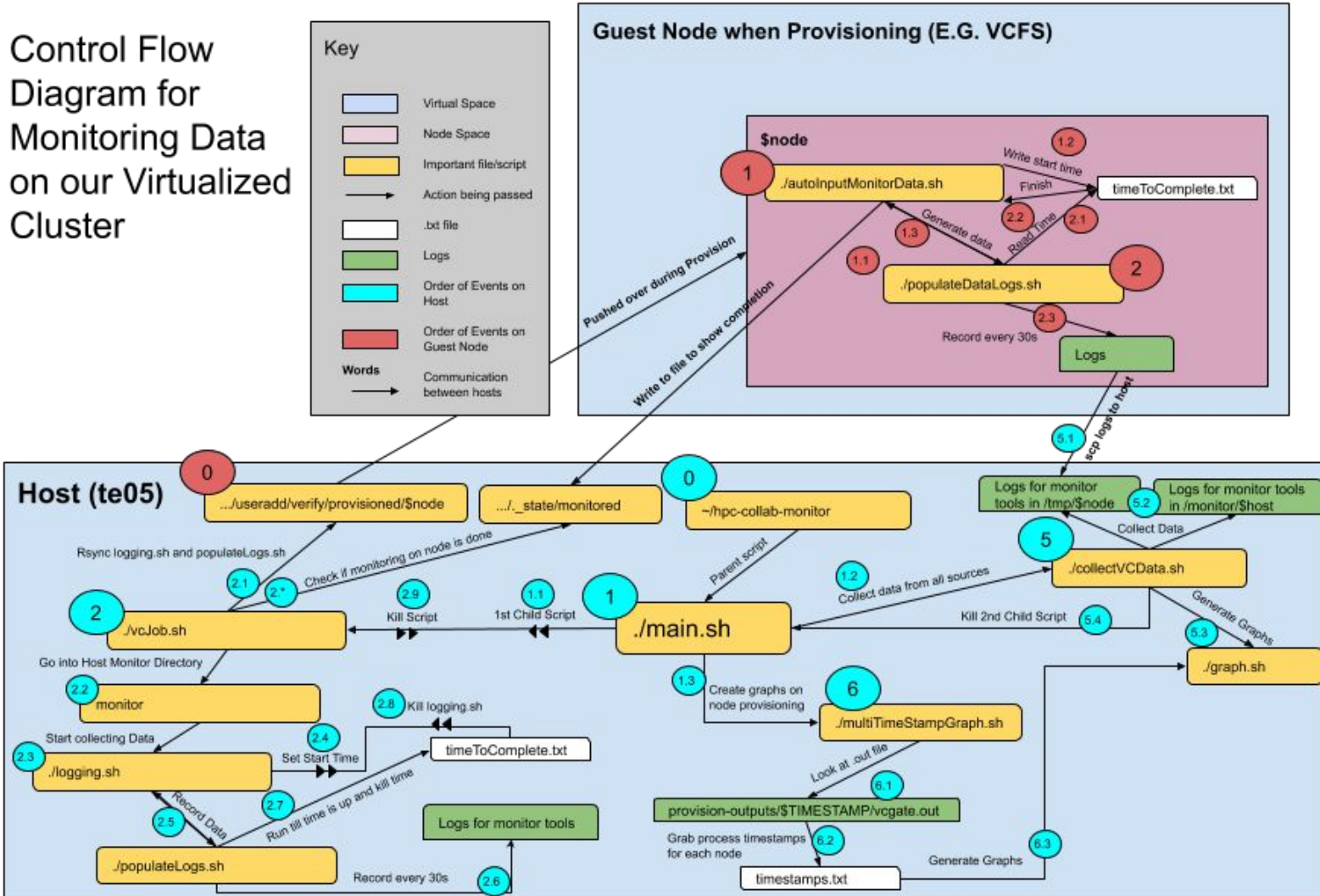
3. Collect Data via scp and Graph

Control Flow Diagram for Monitoring Data on our Virtualized Cluster



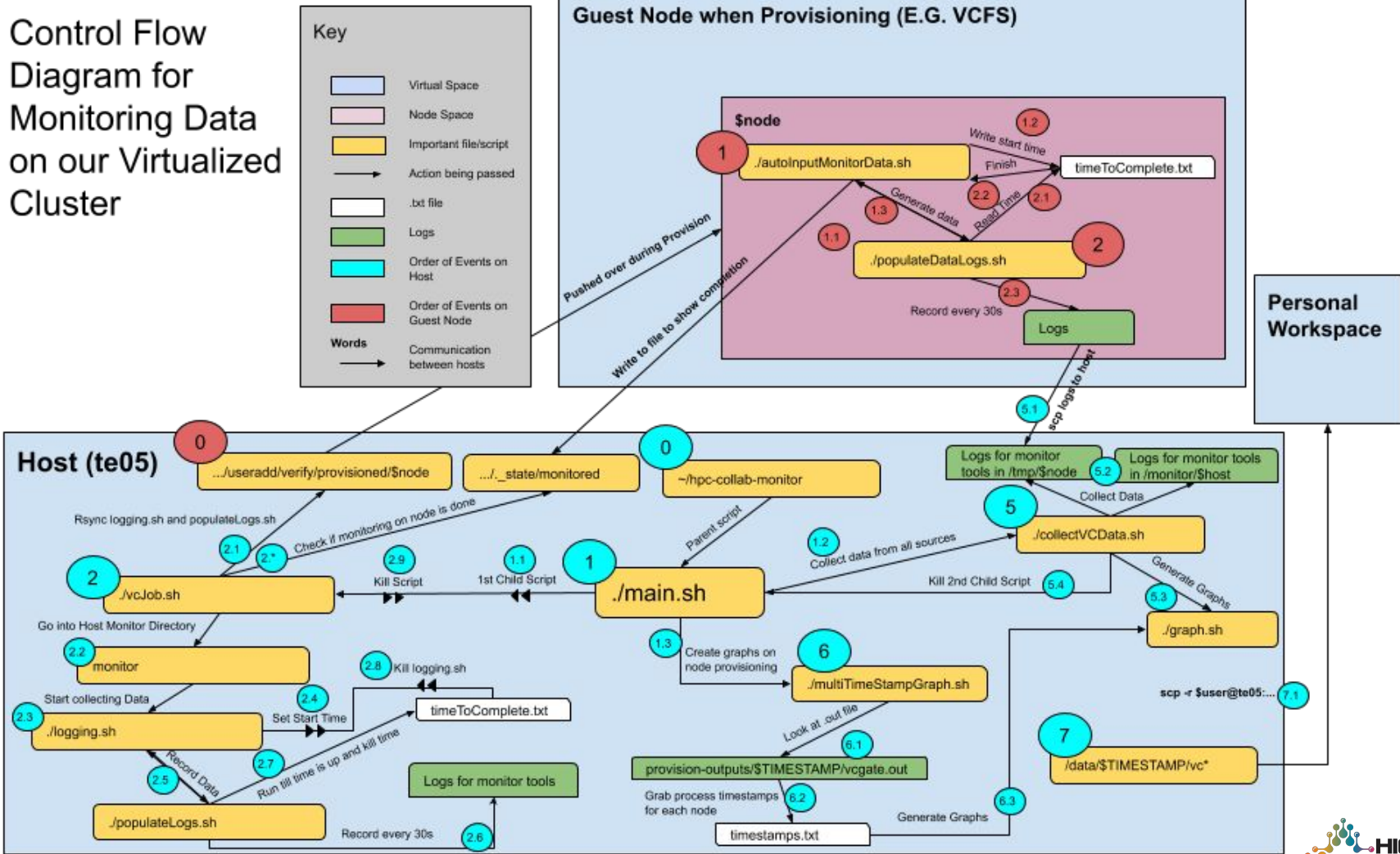
4. Process Graphing per Virtual Node

Control Flow Diagram for Monitoring Data on our Virtualized Cluster



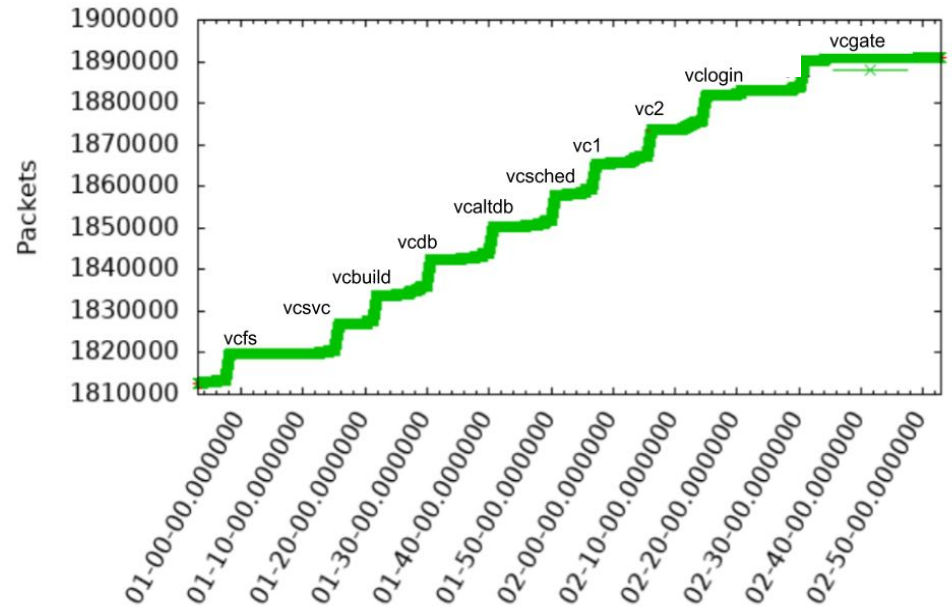
5. scp onto Personal Workspace

Control Flow Diagram for Monitoring Data on our Virtualized Cluster

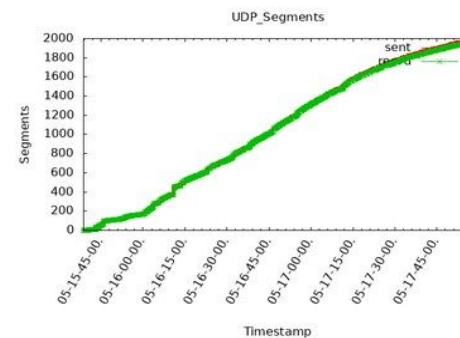
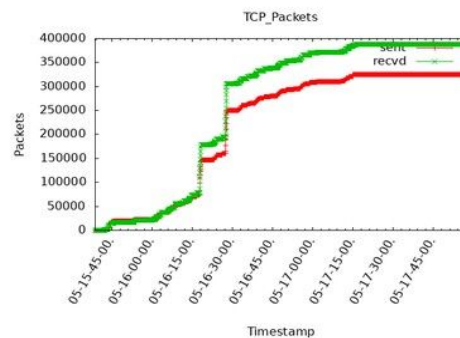
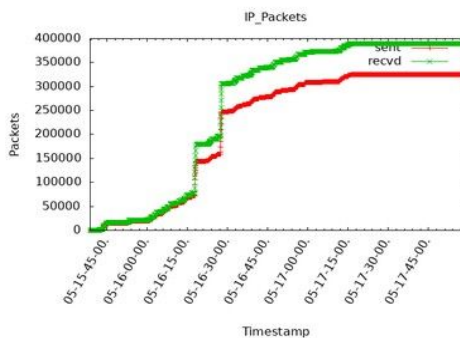
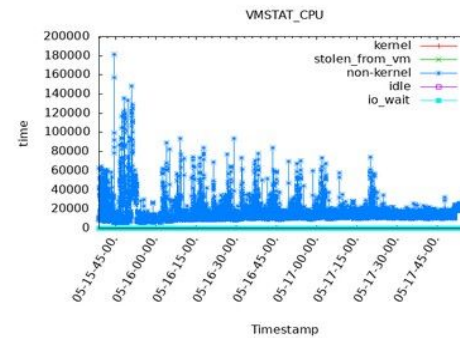
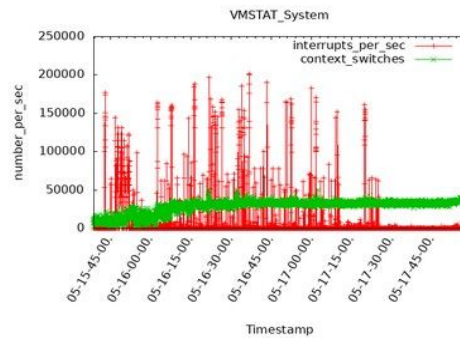
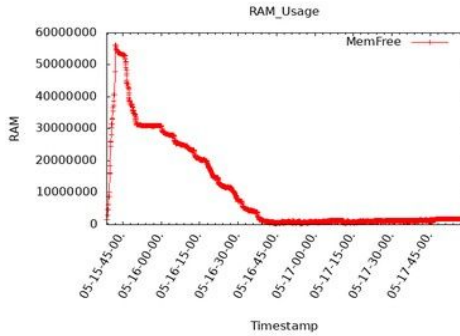
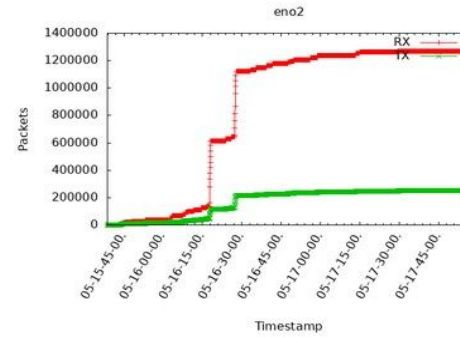
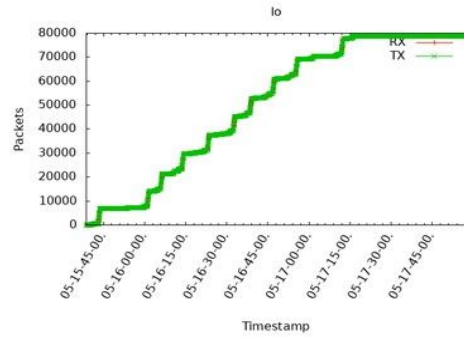
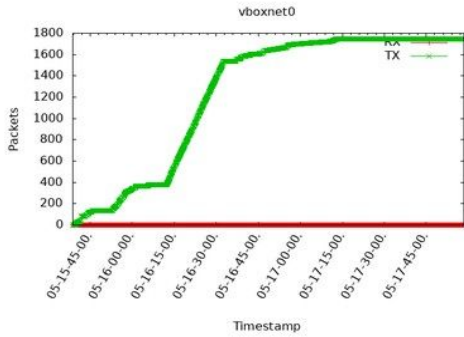


Visualization

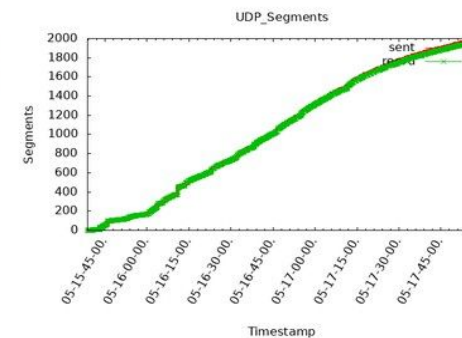
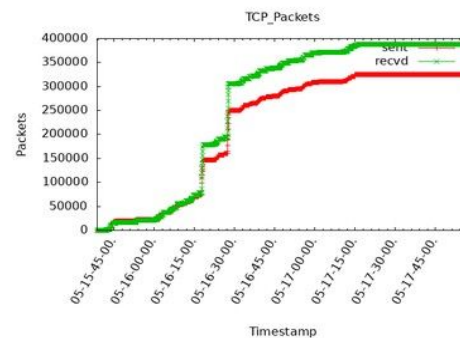
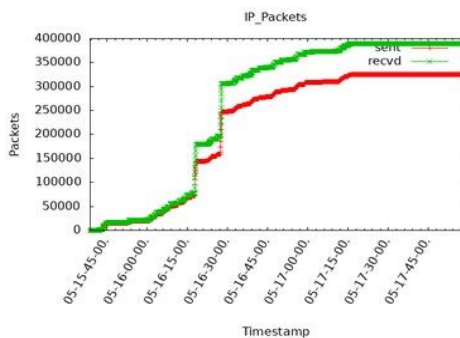
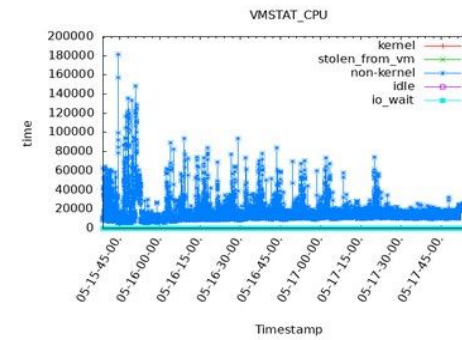
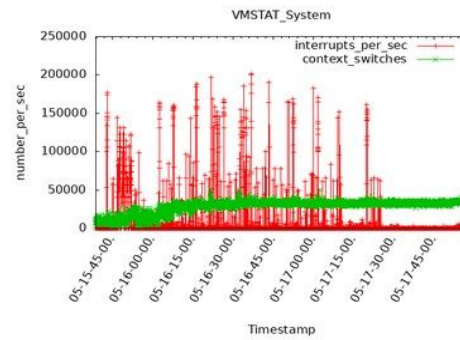
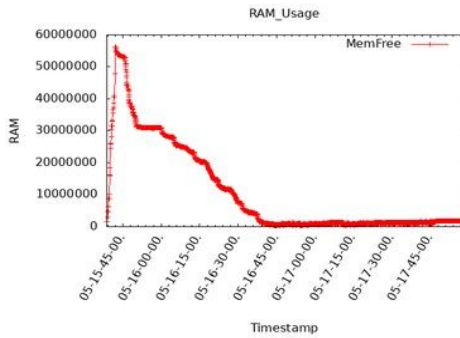
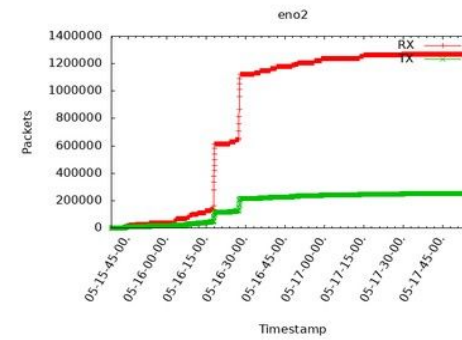
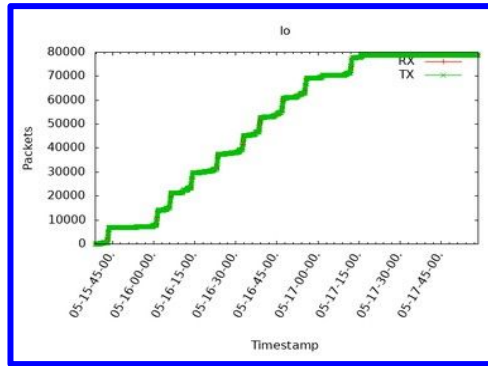
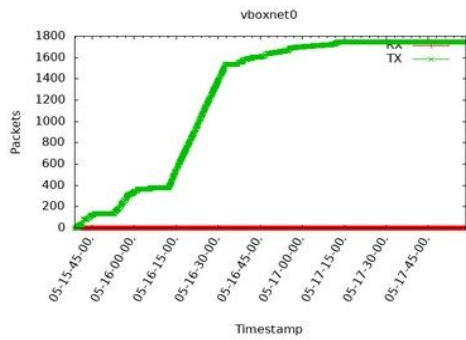
23-23-50-50.442871,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-50.625916,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-50.812695,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-50.988788,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-51.164090,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-51.336390,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-51.508727,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-51.683689,eno2,87405394,21932989,eth0,511099,465478,lo,4772040
23-23-50-51.876905,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.063875,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.237081,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.409386,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.585503,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.770240,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-52.947159,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-53.137041,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-53.326510,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-53.499186,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-53.675928,eno2,87405395,21932989,eth0,511099,465478,lo,4772040
23-23-50-53.862442,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.035580,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.217395,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.397952,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.574922,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.749636,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-54.918804,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.095100,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.287360,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.461422,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.636280,eno2,87405397,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.812771,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-55.981673,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-56.150279,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-56.329525,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-56.512956,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-56.685355,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-56.862438,eno2,87405398,21932990,eth0,511099,465478,lo,4772040
23-23-50-57.038874,eno2,87405398,21932990,eth0,511099,465478,lo,4772053
23-23-50-57.220796,eno2,87405398,21932990,eth0,511099,465478,lo,4772066
23-23-50-57.387691,eno2,87405398,21932990,eth0,511099,465478,lo,4772066
23-23-50-57.558571,eno2,87405398,21932990,eth0,511099,465478,lo,4772066
23-23-50-57.733944,eno2,87405399,21932990,eth0,511099,465478,lo,4772066
23-23-50-57.917462,eno2,87405399,21932990,eth0,511099,465478,lo,4772066
23-23-50-58.104429,eno2,87405399,21932990,eth0,511099,465478,lo,4772066



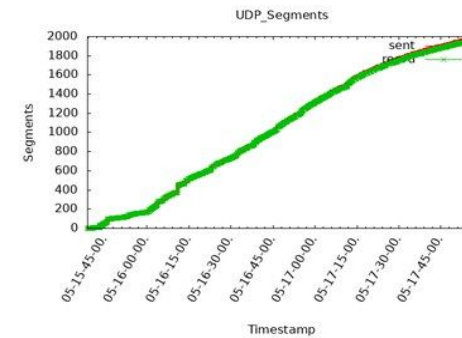
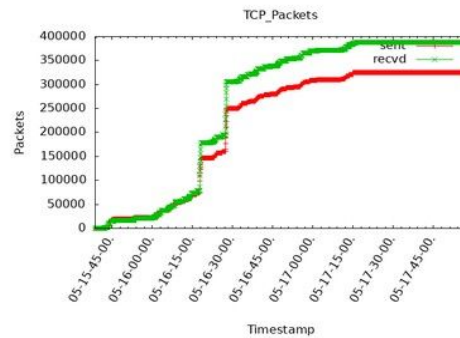
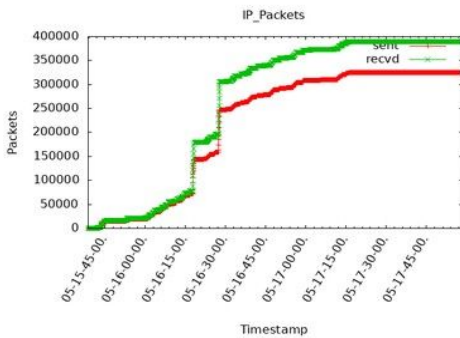
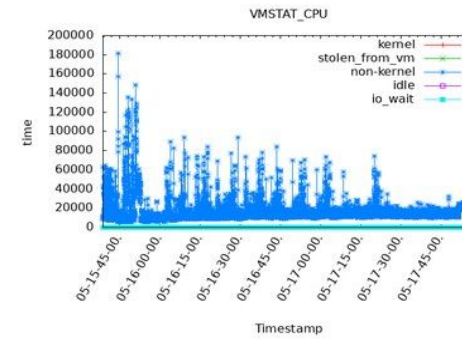
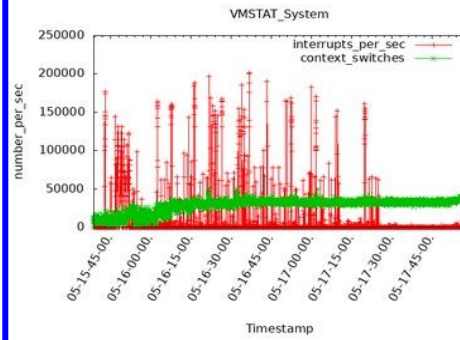
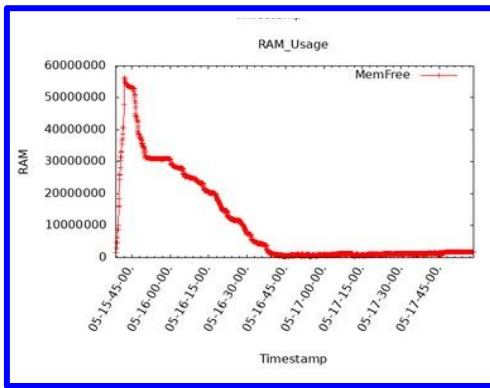
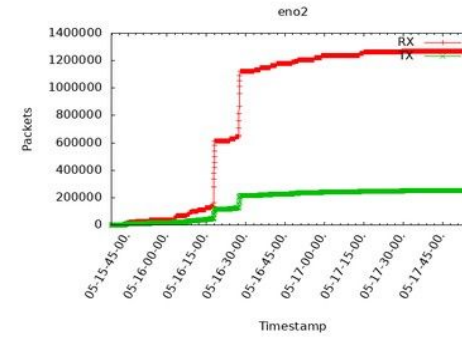
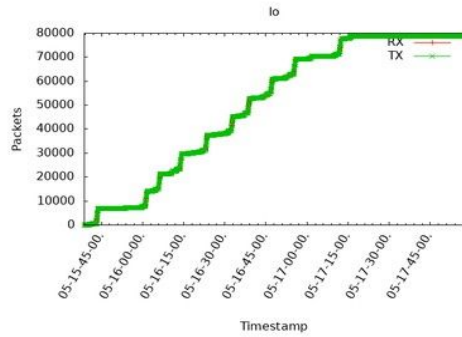
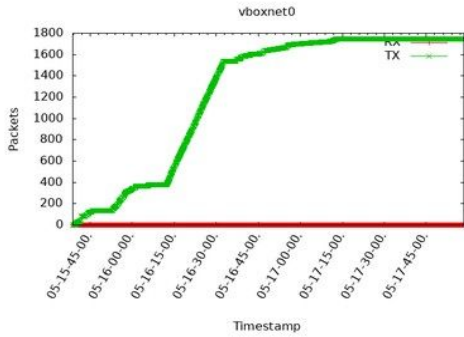
Anatomy of a Dashboard



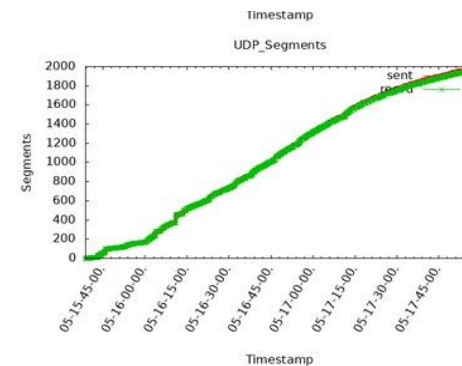
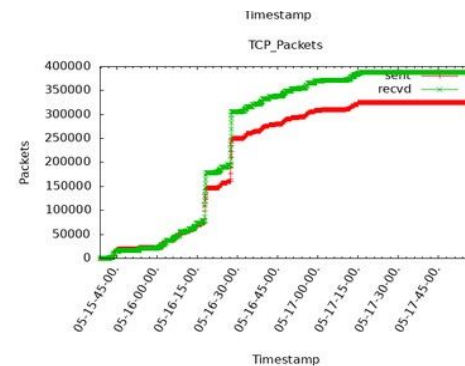
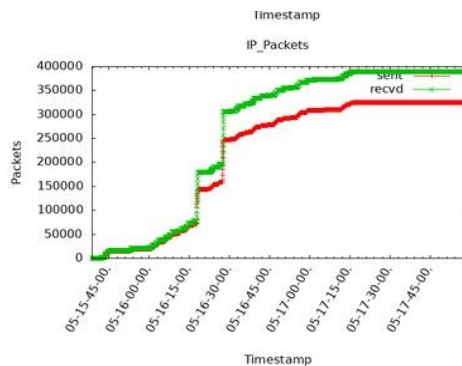
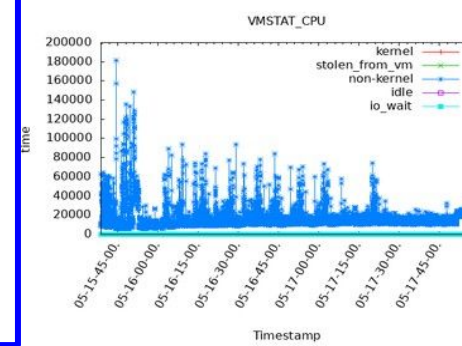
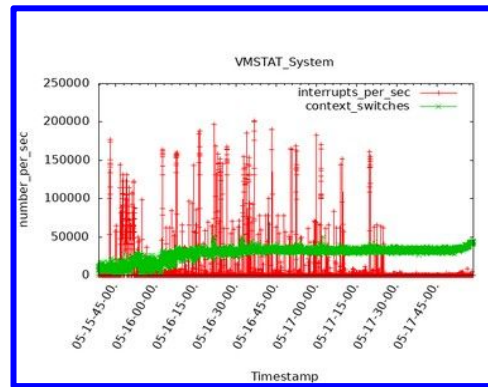
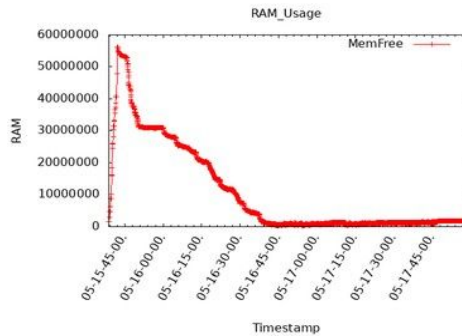
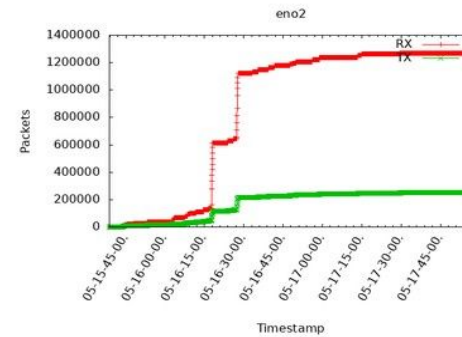
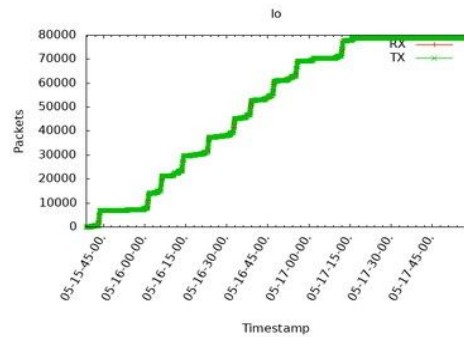
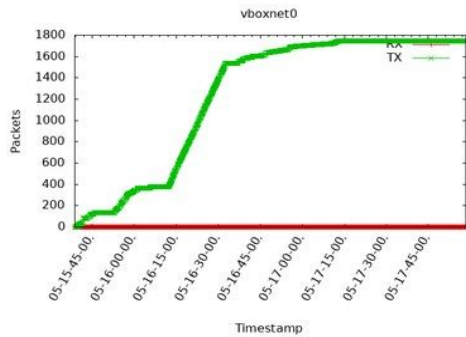
Anatomy of a Dashboard



Anatomy of a Dashboard

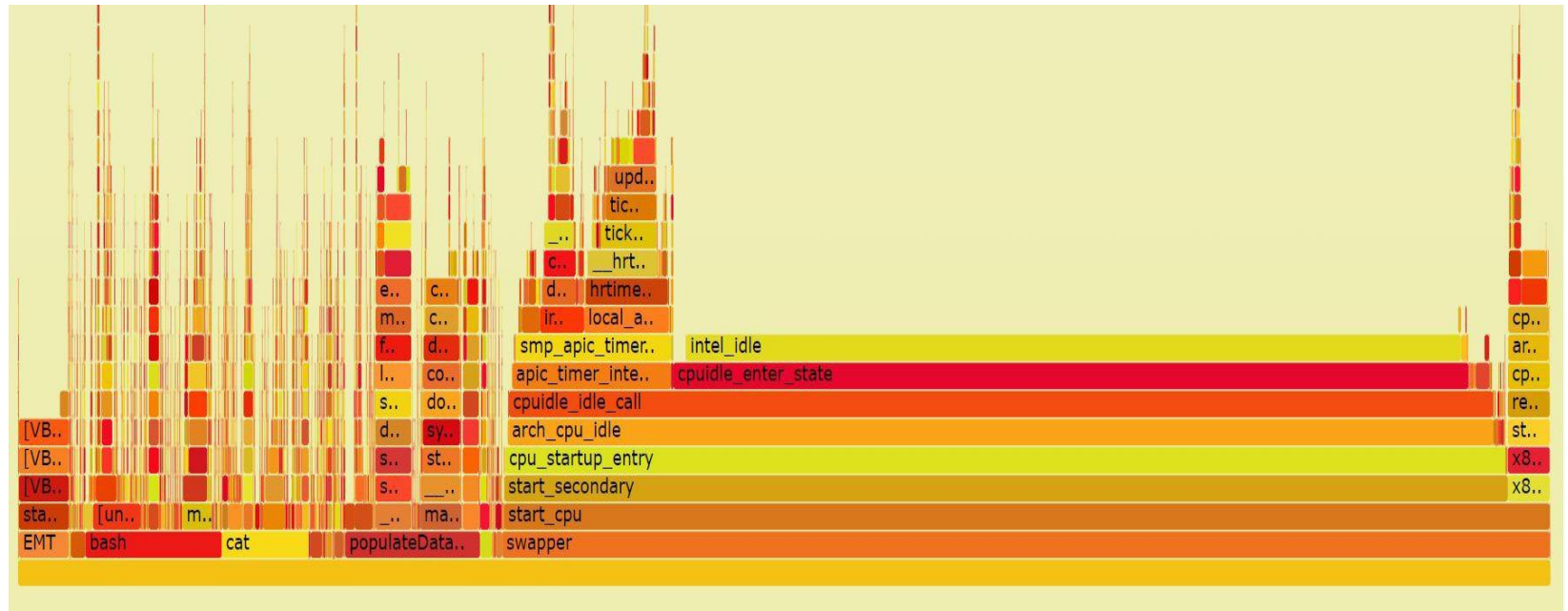


Anatomy of a Dashboard



Perf-events: a fine-grained performance analysis tool

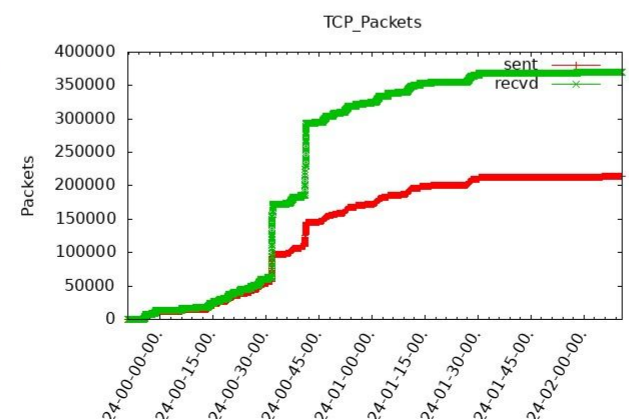
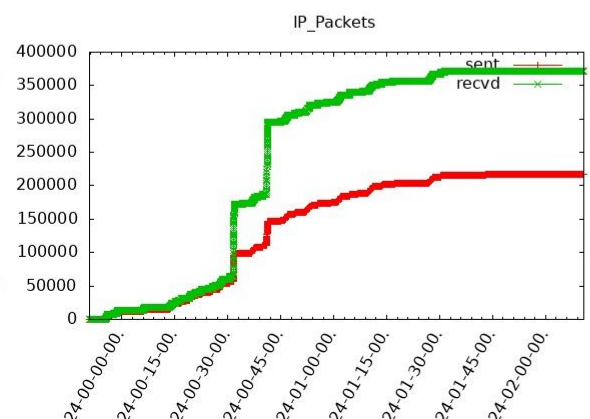
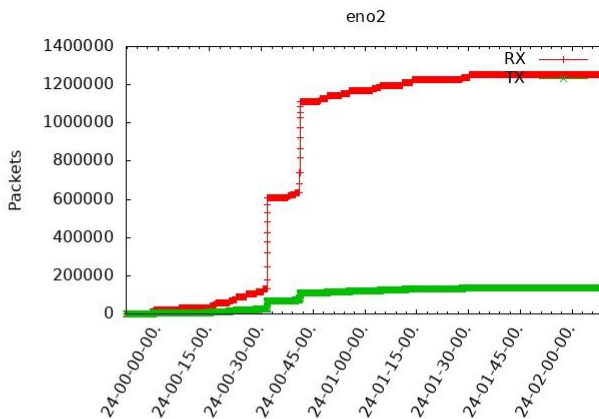
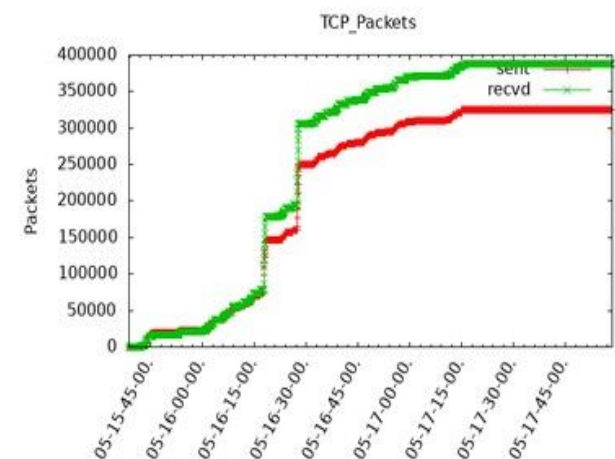
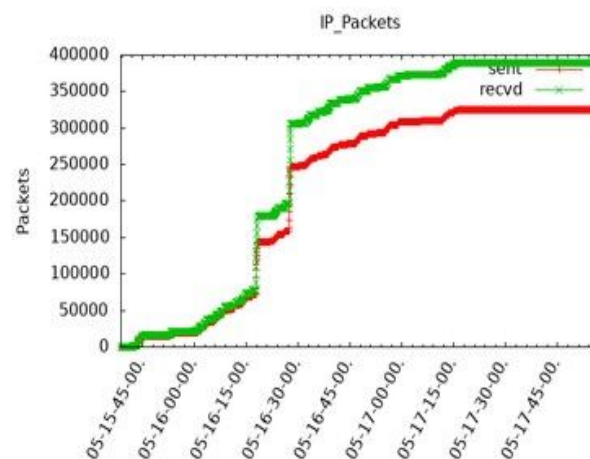
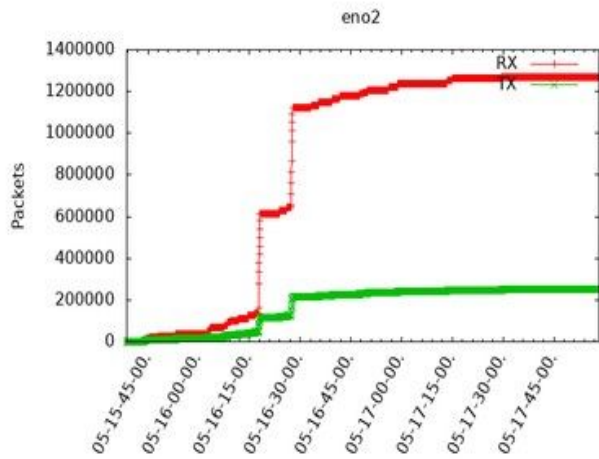
Call Stack



Stack Profile Population (alphabetically)

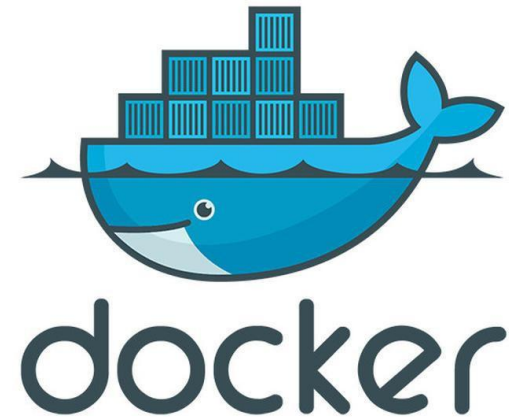


VCBuild: RPMs **cached** vs. **uncached**



Further Directions

- Additional Functionality
 - Perf and Flame Graphs
 - Collectd
- Quantifying Virtualization
 - Overhead
 - Signatures
- Containerization
 - Faster provisioning
- Scalability concerns
 - Larger cluster
 - Clearer Results
- Code Release & Contribution
 - Enable reproducibility



Version 1.97	Sequential Output						Sequential Input				Random Seeks		
	Size	Per Char		Block		Rewrite		Per Char		Block			
		K/sec	% CPU	K/sec	% CPU	K/sec	% CPU	K/sec	% CPU	K/sec	% CPU	/sec	% CPU
te05	128408M	766	99	137384	17	52638	8	1858	96	182797	13	340.4	7
	Latency	22112us		488ms		1621ms		36716us		157ms		818ms	

e.g Bonnie++ Benchmark run on host



Thanks for watching!



Name: Natasha Frumkin
Email: nfrumkin@utexas.edu



Name: Christian Marquardt
Email: cmarquardt45@gmail.com

References

- <http://www.brendangregg.com/perf.html>
- <https://www.ibm.com/developerworks/library/l-linuxuniversal>
- [Github: https://github.com/hpc/hpc-collab](https://github.com/hpc/hpc-collab)

Additional Materials

Virtual Cluster ("VC") Recipe

