Performance Analysis of Scientific Applications on GPUs Using MVAPICH2

Amit Majumdar, Mahidhar Tatineni, and Dongju Choi

San Diego Supercomputer Center

University of California San Diego

MVAPICH User Group (MUG) - 2013

Outline

- **Triton Shared Computing Cluster at SDSC**
- **Amber 12**
- **Performance on GTX680 and GTX780**
- **Future Work**

➢ *TSCC at SDSC*

TSCC at SDSC

- **TSCC is a computational cluster for research computing at UCSD**
- **Provides flexible usage and business model**
- **Collaborative system**
  - Majority of nodes are purchased and shared by users
  - These are *condo* owners
  - *hotel* nodes are available to condo owners and other users on a rental basis

TSCC at SDSC…continued

- **Condo model**
  - Participants purchase one/more nodes for shared cluster
  - Receives equivalent annual time allocation to run
    - Own nodes
    - Nodes contributed by others
    - Hotel nodes
    - Various kinds of job start time guaranteed e.g
    - Jobs of condo size/less start within 8 hours, run for unlimited time
    - Jobs extending to other users' condo nodes – 8-hour limit
    - Jobs using hotel nodes have 72-hour time limit
- **Researchers have the hotel option to pay /hour**

TSCC at SDSC…continued

General Computing Nodes	
Processors	Dual-socket, 8-core, 2.6GHz Intel Xeon E5-2670 (Sandy Bridge)
Memory	64GB (4GB/core) (128GB memory optional)
Network	10GbE (QDR InfiniBand optional)
Hard Drive	500GB onboard (second hard drive or SSD optional)
Warranty	3-years

GPU Nodes	
Host Processors	Dual-socket, 6-core, 2.3GHz Intel Xeon E5-2630 (Sandy Bridge)
GPUs	4 NVIDIA GeForce GTX 680 (GTX Titan upgrade available)
Memory	32GB (64GB/128GB memory optional)
Network	10GbE (QDR InfiniBand optional)
Hard Drive	500GB + 240GB SSD
Warranty	3-years

PDAF (shared resource; pay-as-you-go only)	
Processors	8-socket, 4-core AMD Shanghai Opteron
Memory	512 GB
Network	10 GbE

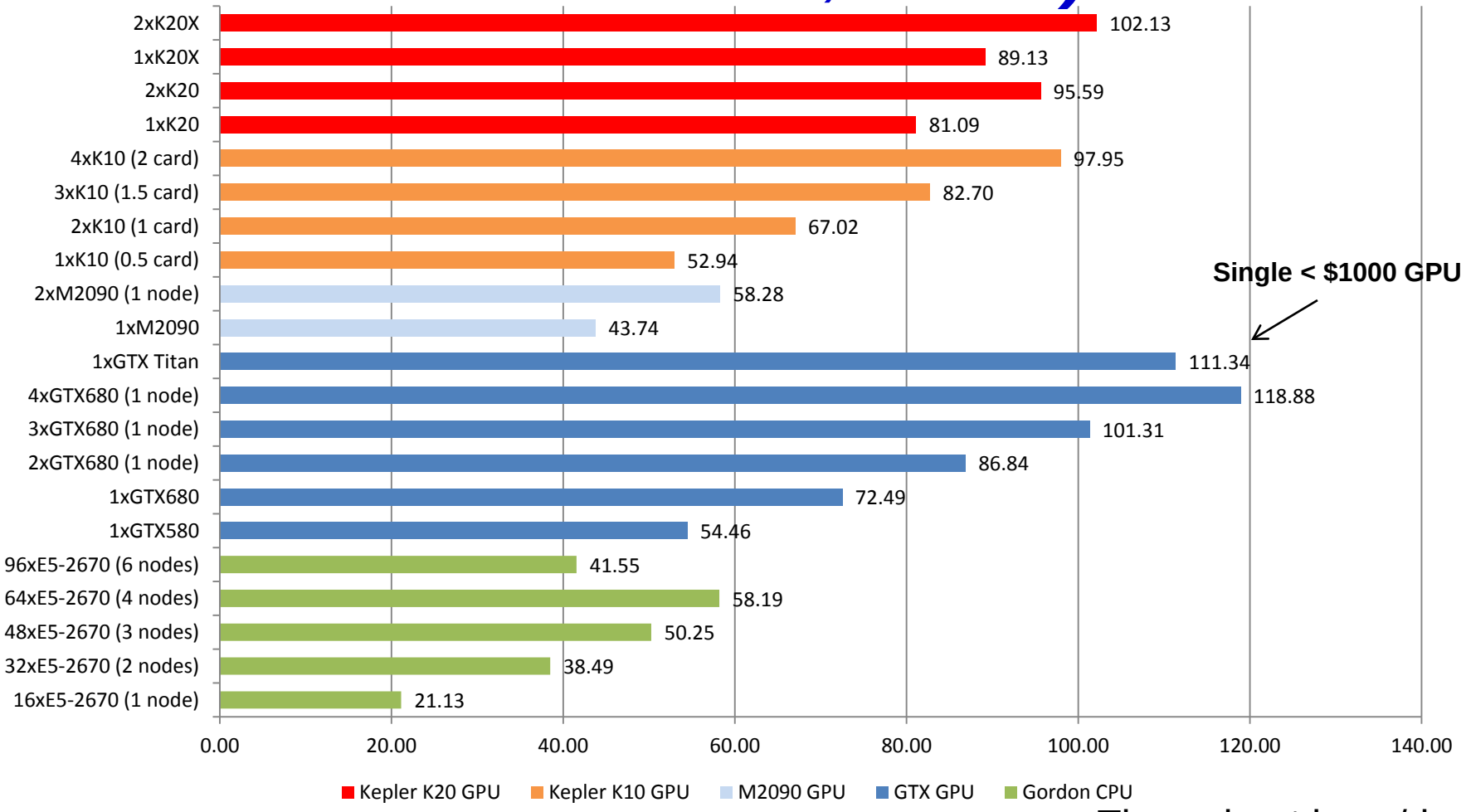
- ***Amber***

Amber 12

- **Significant fraction of GPU users on TSCC use Amber.**
- **Amber:**
  - A set of molecular mechanical force fields for the simulations of biomolecules
  - A package of molecular simulation programs
- **Used by many 1000s of users world wide**
  - Researchers, graduate students, postdocs, industry
  - Any performance improvement has immense impact on many users
- **Amber on GPUs provide 'cheap' way to do significant science on desktops, or small clusters**

Explicit Solvent Performance (JAC-DHFR NVE Production Benchmark

Amber GPU v12.3, Jul 2013)



Slides provided by Ross Walker, SDSC

Performance Example (TRP Cage)

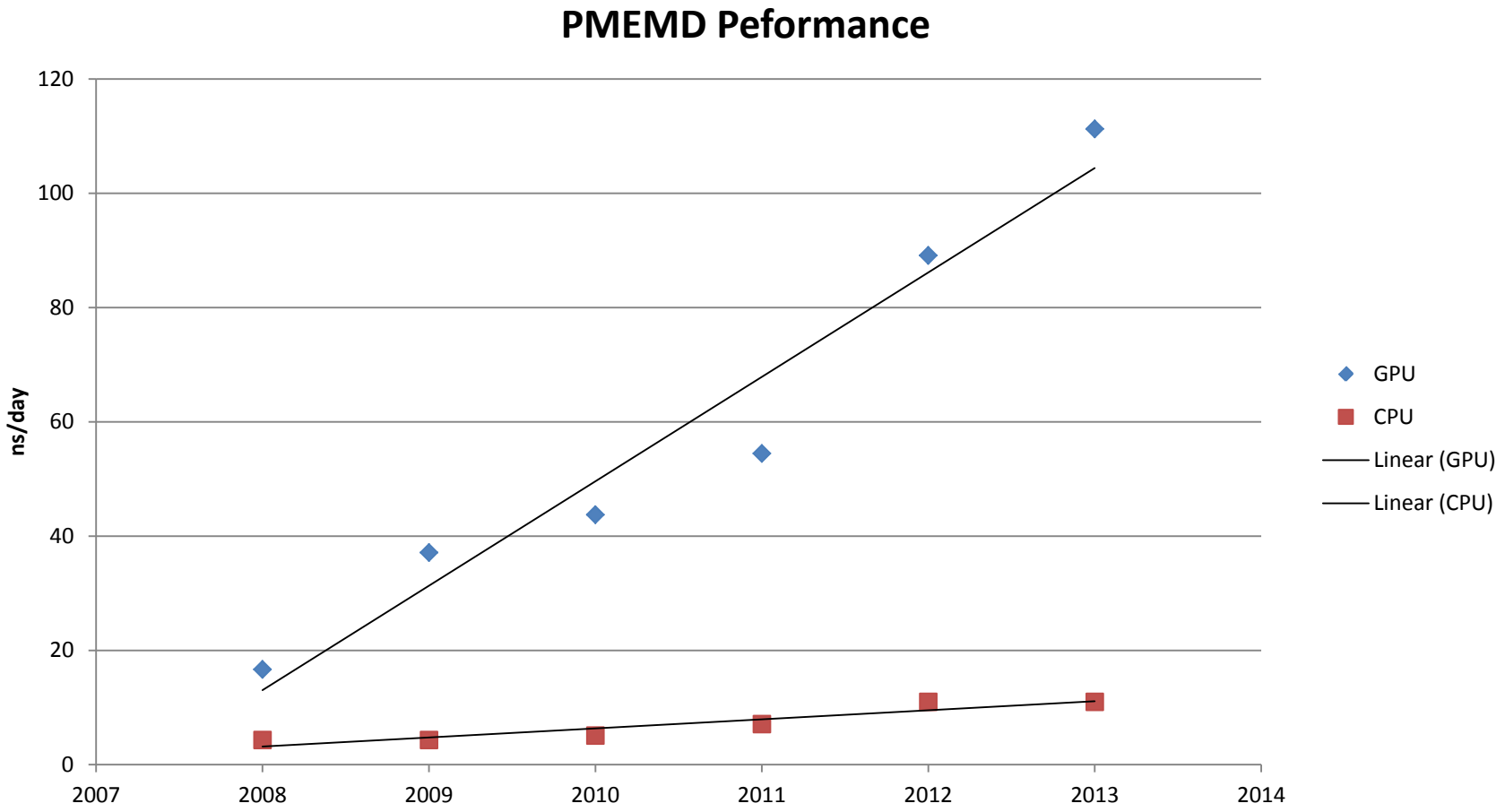
CPU 2x8xE5-2670

```
[10:50][caffeine:0.00][rcw:2009_11_13_Benchmark_for_SC09_Talk]$
```

Slides provided by Ross Walker, SDSC

Historical Single Node / Single GPU Performance

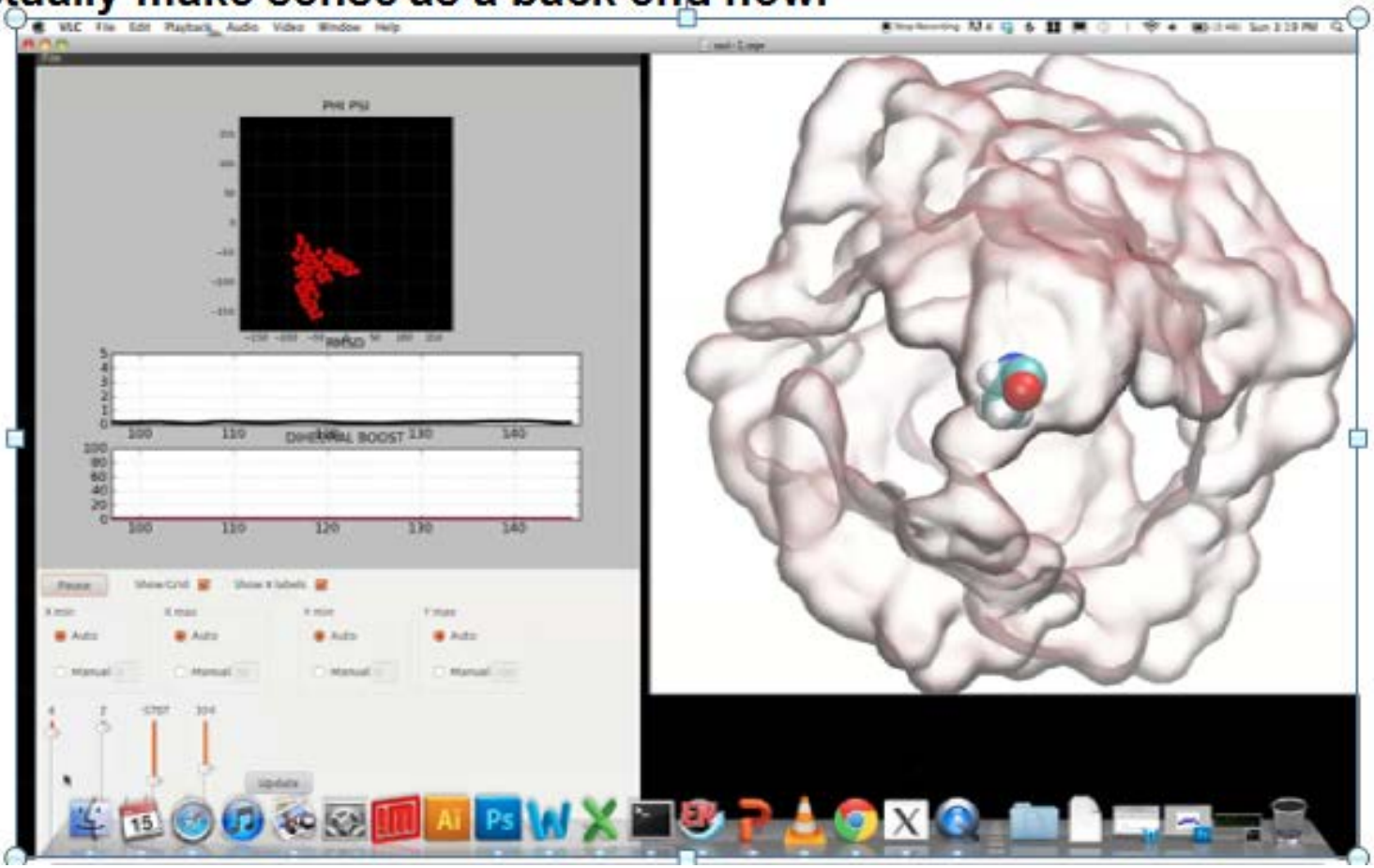
(DHFR Production NVE)



Slides provided by Ross Walker,SDSC SDSC

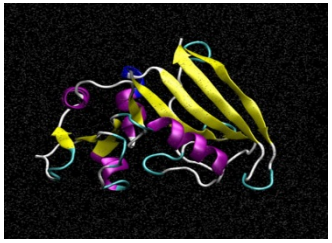
Interactive MD?

Single nodes are now fast enough, GPU enabled cloud nodes actually make sense as a back end now.



Slides provided by Ross Walker, SDSC

Cost Comparison



4 simultaneous simulations, 23,000 atoms, 250ns each, 5 days maximum time to solution.

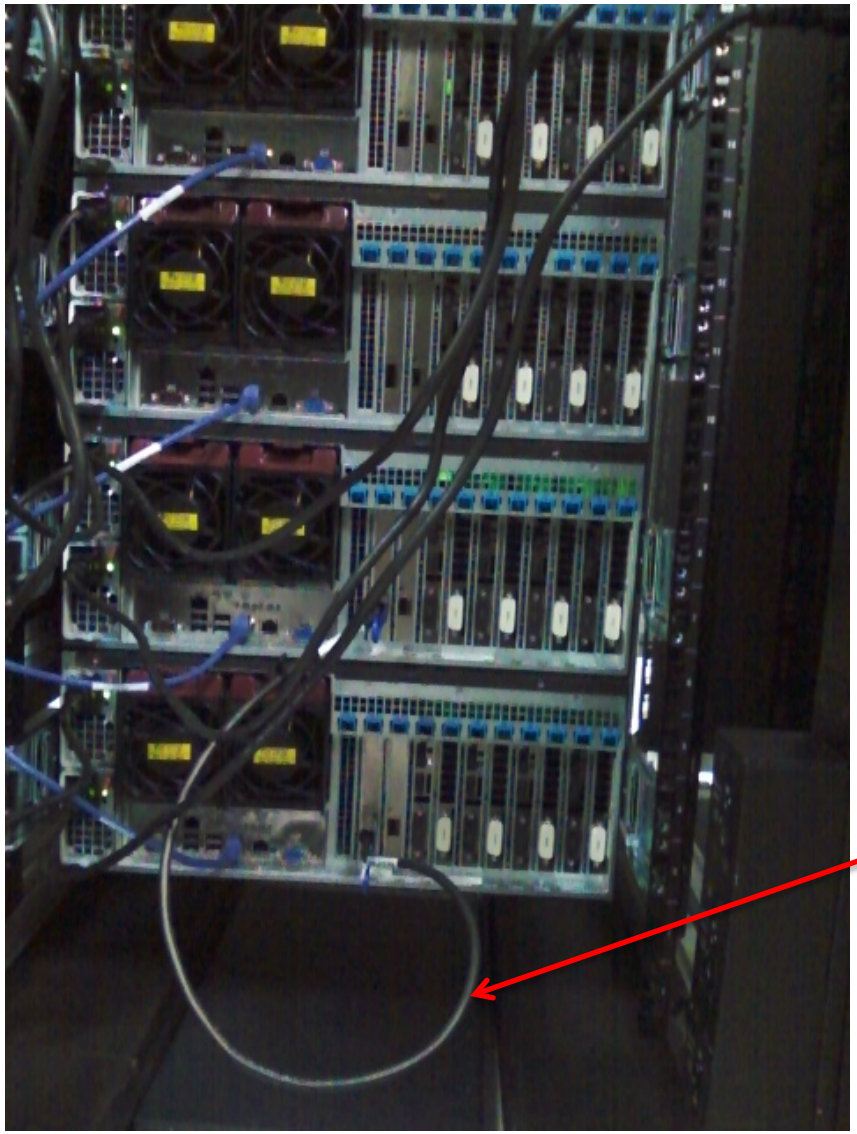
	Traditional Cluster	GPU Workstation
Nodes Required	12	1 (4 GPUs)
Interconnect	QDR IB	None
Time to complete simulations	4.98 days	2.25 days
Power Consumption	5.7 kW (681.3 kWh)	1.0 kW (54.0 kWh)
System Cost (per day)	$96,800 ($88.40)	$5200 ($4.75)
Simulation Cost	(681.3 * 0.18) + (88.40 * 4.98)	(54.0 * 0.18) + (4.75 * 2.25)
	$562.87	***$20.41***

>25x cheaper AND solution obtained in less than half the time

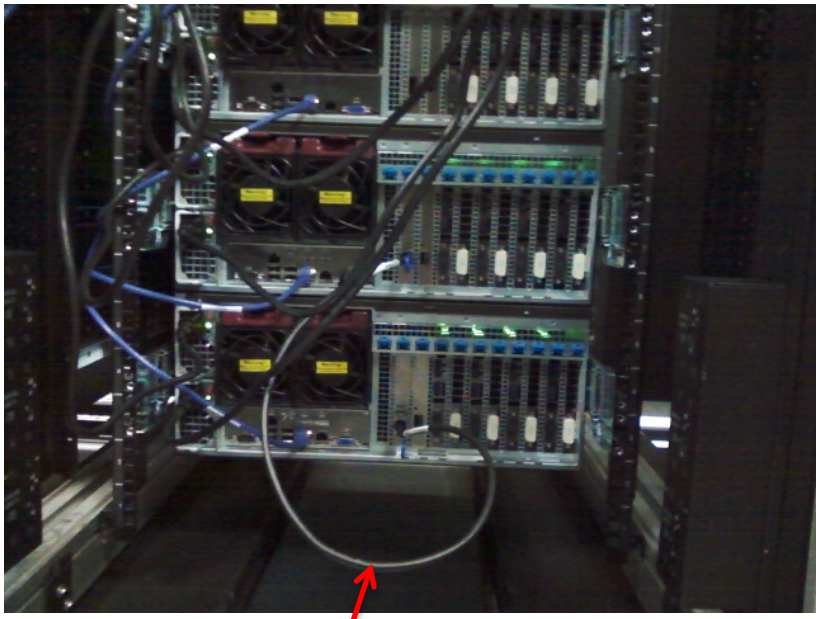
Slides provided by Ross Walker, SDSC

- ***Performance on GTX680 and GTX780***

TSCC GPU nodes connected with Mellanox cable



Back View



Mellanox cable
(no switch yet)

Front View

FACTOR_IX_PRODUCTION_NPT - 90,906 atoms PME

SET #1: TSCC Cluster, tscc-gpu-3-0, 3 GTX680 Cards per node, SINGLE NODE, MVAPICH2

1 x GTX680: | ns/day = 16.05 seconds/ns = 5383.93

2 x GTX680: | ns/day = 19.90 seconds/ns = 4340.91

3 x GTX680: | ns/day = 22.86 seconds/ns = 3779.03

SET #2: Stampede, 1 K20 per node, ECC on, Default MVAPICH2 (no MV2_USE_CUDA)

1 x K20: | ns/day = 17.82 seconds/ns = 4848.03

2 x K20: | ns/day = 21.85 seconds/ns = 3954.08*

3 x K20: | ns/day = 25.18 seconds/ns = 3431.06*

4 x K20: | ns/day = 27.86 seconds/ns = 3100.99*

*Multi-node

SET#3: TSCC Cluster, tscc-gpu-5-0, 4 GTX780 Cards per node, SINGLE NODE, MVAPICH2

1 x GTX780: | ns/day = 21.70 seconds/ns = 3981.62

2 x GTX780: | ns/day = 25.46 seconds/ns = 3393.15

3 x GTX780: | ns/day = 30.01 seconds/ns = 2879.27

4 x GTX780: | ns/day = 32.54 seconds/ns = 2655.46

CELLULOSE_PRODUCTION_NPT - 408,609 atoms PME

SET#1: TSCC Cluster, tscc-gpu-5-0, 4 GTX780 Cards per node

SINGLE NODE, MPVAPICH2

1 x GTX780: | ns/day = 5.17 seconds/ns = 16707.42

2 x GTX780: | ns/day = 5.54 seconds/ns = 15599.31

3 x GTX780: | ns/day = 6.67 seconds/ns = 12947.53

4 x GTX780: | ns/day = 6.88 seconds/ns = 12552.72

Link to AMBER Benchmark Page with comparisons:

http://ambermd.org/gpus/benchmarks.htm

Future Work

- Goal is to do GPUDirectRDMA for Amber 12
  - Has potential for performance improvement
  - GPUDirectRDMA supported on K20s
  - Stampede is potential machine

- GPUDirectRDMA not supported on GTX680 and GTX780