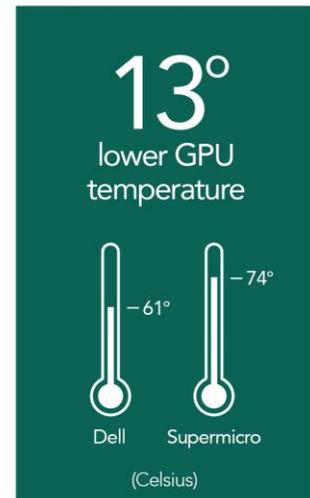## DELL™ POWEREDGE™ C4130 SERVER

POWERED BY THE INTEL® XEON® PROCESSOR E5-2600 v3 FAMILY

with four NVIDIA® Tesla™ K80 GPUs delivered cooler temperatures and improved performance

E5-2650 v3 **25%** better performance with Intel Xeon processors E5-2650 v3

E5-2670 v3 **28%** better performance with Intel Xeon processors E5-2670 v3

E5-2690 v3 **27%** better performance with Intel Xeon processors E5-2690 v3

13° lower GPU temperature

~74°
~61°

Dell   Supermicro

(Celsius)

versus Supermicro® 1028GR-TR configured with a maximum of three GPUs

plus, the added GPUs boosted performance of the Dell PowerEdge C4130 server by up to **9.37X**

Organizations running the more than 200 available high-performance computing (HPC) applications know that computational performance is their top priority for getting results faster. In recent years, one approach to achieving increased compute power has been to add graphics processing units (GPUs) that can handle the complex and demanding compute requirements, such as the NVIDIA Tesla K80 GPU accelerators, to server hardware designed to support them. The new Dell PowerEdge C4130 server provides a unique platform to support the accelerated performance of GPUs with a physical design intended to facilitate airflow and reduce overheating.

At Principled Technologies, we first compared the floating-point performance of an Intel processor-powered Dell PowerEdge C4130 without any GPUs (baseline) and then with the maximum amount of NVIDIA Tesla K80 GPU accelerators (maximum) in three different Intel Xeon processor configurations: pairs of E5-2650 v3, E5-2670 v3, and E5-2690 v3 processors. We found that adding the four GPUs to the Dell PowerEdge C4130 delivered up to 9.37 times more performance than the Dell PowerEdge C4130 without any GPUs. Improved performance from GPUs means getting more from HPC workloads.

We then compared the performance of the Dell PowerEdge C4130 with four GPUs against a Supermicro 1028GR-TR with the maximum amount of NVIDIA Tesla K80 GPUs (three) in all three processor configurations. We found that the Dell PowerEdge

C4130 consistently outperformed the Supermicro server solution, delivering up to 28 percent better performance. In addition, when we looked at GPU temperatures as an indicator of airflow through both the Dell and Supermicro servers, we found the GPU temperatures of the Dell PowerEdge C4130 were up to 13 degrees (Celsius) cooler than the Supermicro 1028GR-TR. These cooler running temperatures may help prevent deterioration of the silicon from overheating and may extend the lifetime of the chips.

For more information about the testing components, see Appendix A. For more information on the system configuration of each server, see Appendix B. For detailed steps on how we tested, see Appendix C.

GET MORE COMPUTING POWER WITH NVIDIA TESLA K80 GPUS

Organizations running HPC applications specialize in areas ranging from molecular dynamics to animation to weather forecasting. In recent years, the addition of GPUs to servers has seen increased performance for these organizations. The GPUs add computational power, which can offload or supplement the workload on the CPU.

We used Intel's version of the high-performance LINPACK (Intel HPL) benchmark to test the floating-point performance of the baseline configuration without any NVIDIA GPUs. We used NVIDIA's version of the high-performance LINPACK (NVIDIA HPL) benchmark to test the floating-point performance of the maximum configuration, which aggregates all available compute power of both GPUs and CPUs. Our tests measured CPU performance for the baseline configuration and GPU and CPU performance combined for the maximum configuration. The HPL benchmark solves a set of linear equations and returns the time required and the average floating point performance achieved in giga-floating-point operations per second (Gflops). In our tests, we chose a problem size of 121,344 and a block size of 1,024 to make full use of the memory and to maximize performance. Our testing occurred in a datacenter at 25 degrees Celsius.

We first tested the baseline configuration of the PowerEdge C4130 in three different processors. Having three processor configurations allowed us to see how changes in processor affected the floating-point performance. Next, we tested the maximum configuration of the PowerEdge C4130 in the same three Intel Xeon processor configurations. Figure 1 presents information for each Intel Xeon processor we used.

	Intel Xeon processor E5-2650 v3	Intel Xeon processor E5-2670 v3	Intel Xeon processor E5-2690 v3
TDP (watts)	105	120	135
Frequency (GHz)	2.3	2.3	2.6
Cores	10	12	12

Figure 1: CPU information for the processors we used in testing.

Our tests showed that adding GPUs to each processor configuration could dramatically increase floating-point performance. Figure 2 provides the normalized Gflops performance numbers for baseline and maximum GPU configurations for each processor. The Dell PowerEdge C4130 maximum GPU configuration delivered 9.3 times, 7.6 times, and 7.0 times more floating-point performance than the baseline configuration with Intel Xeon processors E5-2650 v3, E5-2670 v3, and E5-2690 v3, respectively.

Figure 2: Normalized LINPACK floating-point performance for the baseline and maximum configuration of the PowerEdge C4130 using multiple Intel Xeon processor configurations. Higher numbers are better.
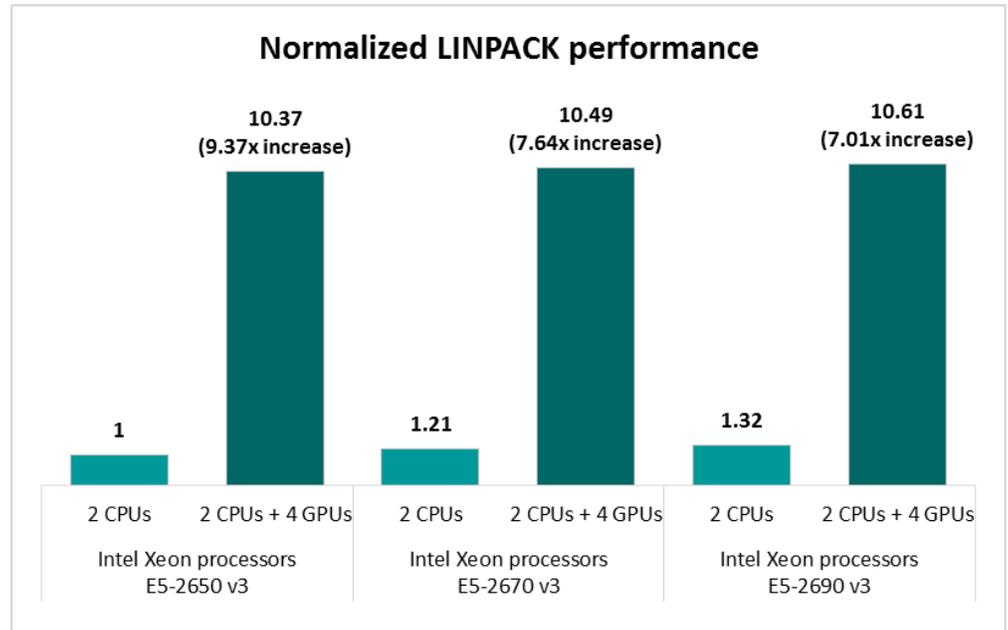

This increase in compute power can put servers in a dangerous position. Running too long at higher-than-normal temperatures can diminish workload performance for these HPC applications and can decrease hardware life. The design of the Dell PowerEdge C4130 positions the GPUs to receive significant airflow for cooling. In addition, the design includes more openings that allow air to move out of the server, helping air warmed by GPUs and processors to leave.

We found that peak GPU temperatures for the Dell PowerEdge C4130 maximum configuration were generally consistent with each of the three Intel Xeon processors (see Figure 3). All three maximum GPU temperatures were within a standard safe-operation temperature range.[1]

---

[1] NVIDIA GPUs typically have maximum operating temperatures of approximately 105°C. nvidia.custhelp.com/app/answers/detail/a_id/2752/~/nvidia-gpu-maximum-operating-temperature-and-overheating

Peak GPU temperatures
Maximum GPU configuration

Figure 3: Peak GPU temperature for the maximum configurations of the Dell PowerEdge C4130 with each of the three Intel Xeon processors.
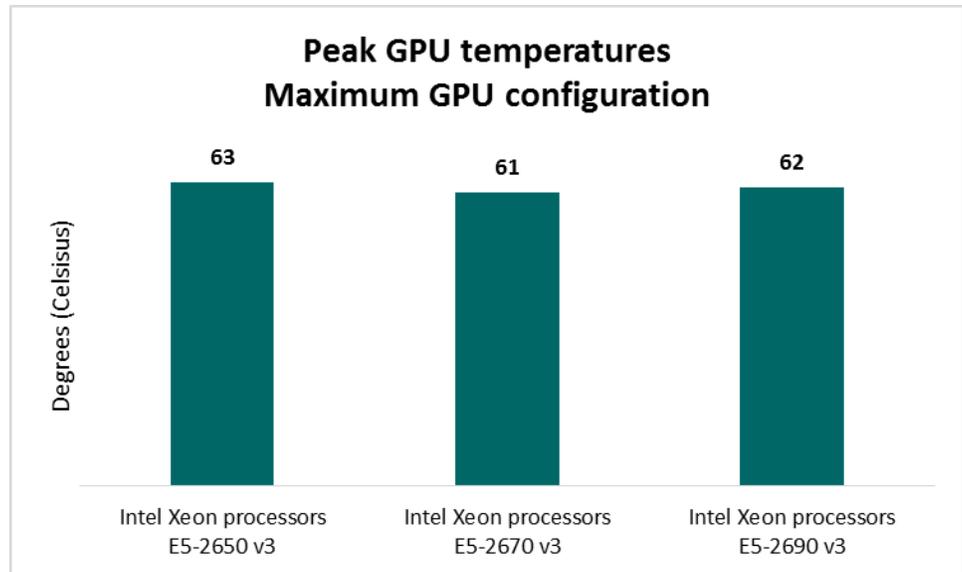
If your organization needs the power of GPUs to run graphic-heavy design applications or complex scientific simulations, the Dell PowerEdge C4130 can provide the hardware to support your demanding workloads. Keeping the silicon of GPUs at a reasonable temperature, as the PowerEdge C4130 did, could extend the lifetime of the chips, while reducing diminished performance and power inefficiency.

BETTER PERFORMANCE FROM THE DELL POWEREDGE C4130 WITH COOLER NVIDIA TESLA K80 GPUS

After realizing the performance benefits of adding NVIDIA Tesla K80 GPU accelerators to the Dell PowerEdge C4130, we compared the maximum configuration of the PowerEdge C4130 to a Supermicro 1028GR-TR server with the maximum number of supported NVIDIA Tesla K80 GPU accelerators in the three Intel Xeon processor configurations. The PowerEdge C4130 can house four NVIDIA Tesla K80 GPU accelerators simultaneously, while the Supermicro 1028GR-TR can hold only three. Not only did the Dell PowerEdge C4130 maximum configuration outperform the Supermicro 1028GR-TR maximum configuration, but it also kept the GPU temperatures cooler for all three processors. Figure 4 shows GPU temperatures for both solutions in each processor configuration.
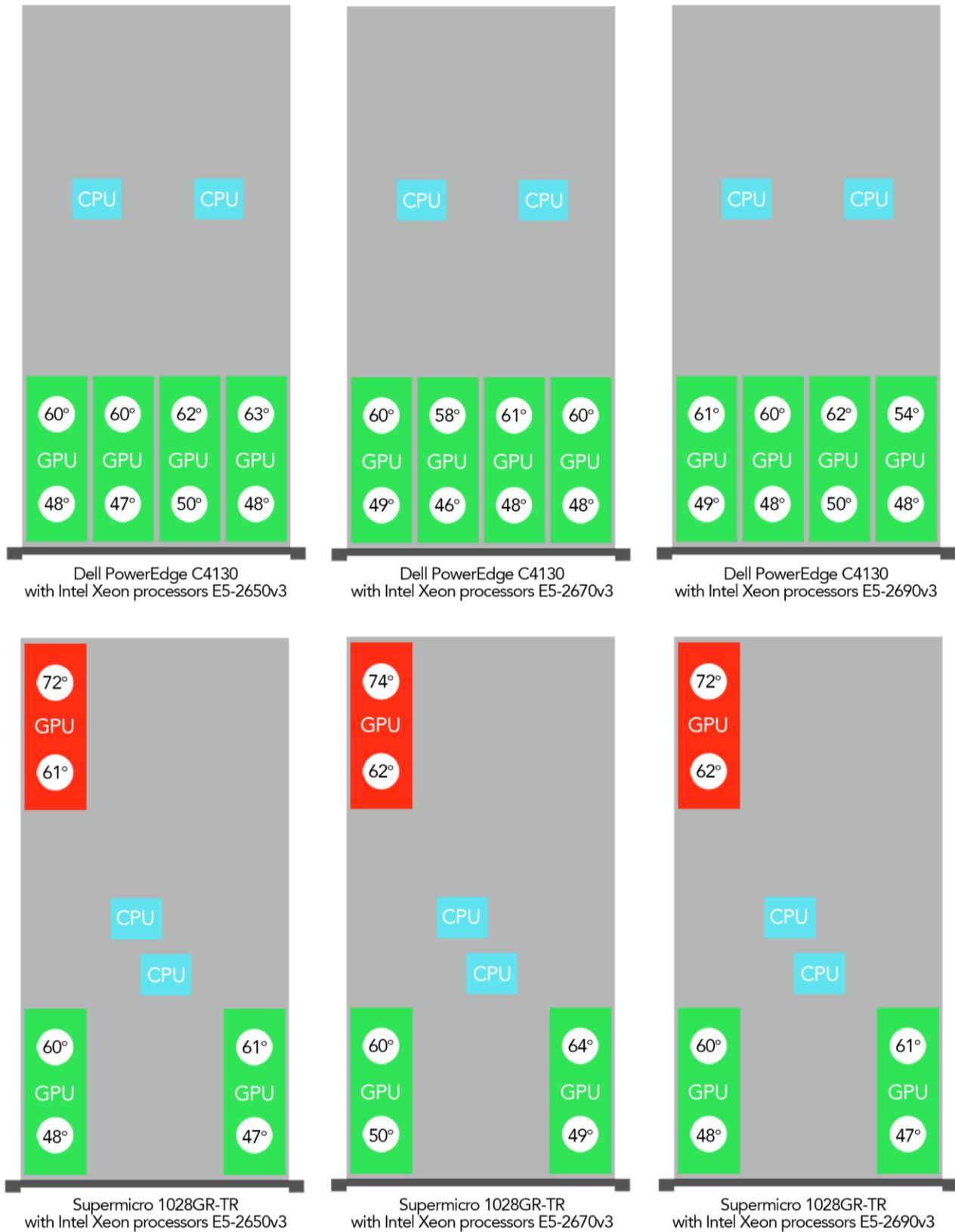
Figure 4: GPU and CPU placement in the two servers we tested showing sensor temperatures with multiple Intel Xeon processor configurations.

In the following sections, we show how the maximum configurations for each solution performed in each of the processor configurations we tested.

Intel Xeon processor E5-2650 v3

Floating-point performance

We found that the Dell PowerEdge C4130 maximum configuration achieved 25 percent greater floating-point performance than the Supermicro 1028GR-TR maximum configuration. Figure 5 shows the normalized Gflops achieved by the two solutions with Intel Xeon processors E5-2650 v3.
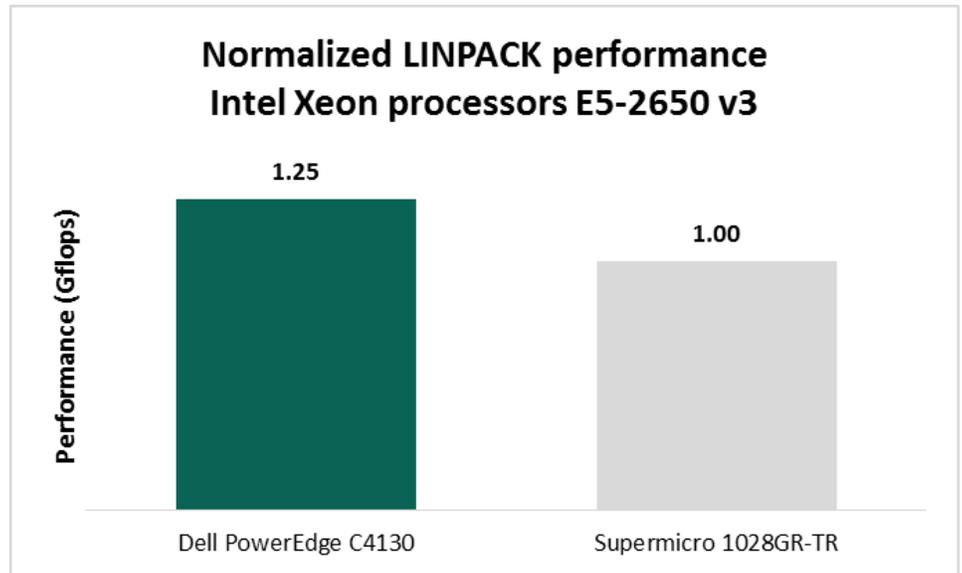
Figure 5: Normalized LINPACK floating-point performance for the maximum configuration of each system. Higher numbers are better.

Normalized LINPACK performance
Intel Xeon processors E5-2650 v3

GPU temperature

We found the peak GPU temperature for the Dell PowerEdge C4130 maximum configuration was 9 degrees cooler than the peak GPU temperature of the Supermicro 1028GR-TR (see Figure 6).
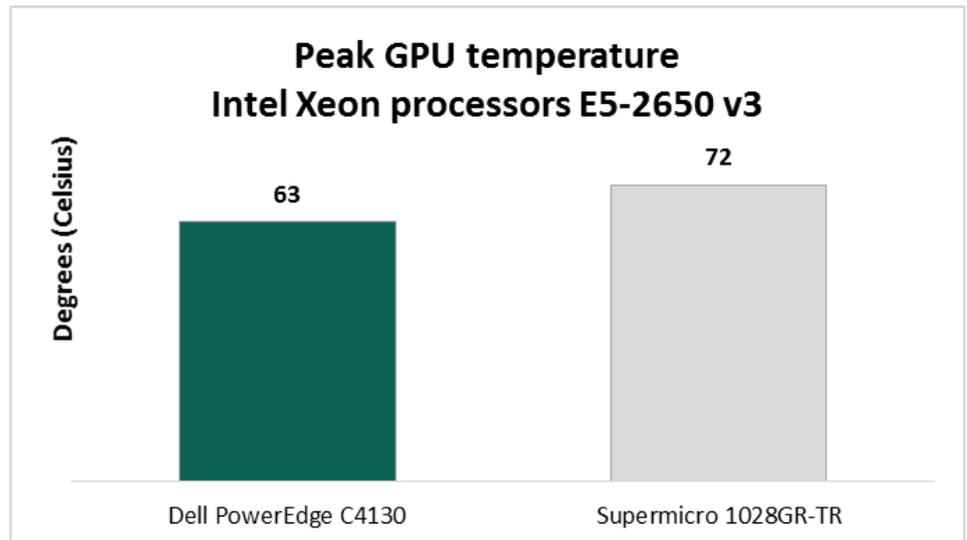
Figure 6: Peak GPU temperature for the maximum configuration of both systems. Lower numbers are better.

Peak GPU temperature
Intel Xeon processors E5-2650 v3

Intel Xeon processor E5-2670 v3

Floating-point performance

We found that the Dell PowerEdge C4130 maximum configuration achieved 28 percent greater floating-point performance than the Supermicro 1028GR-TR maximum configuration. Figure 7 shows the normalized Gflops achieved by the two solutions with Intel Xeon processors E5-2670 v3.
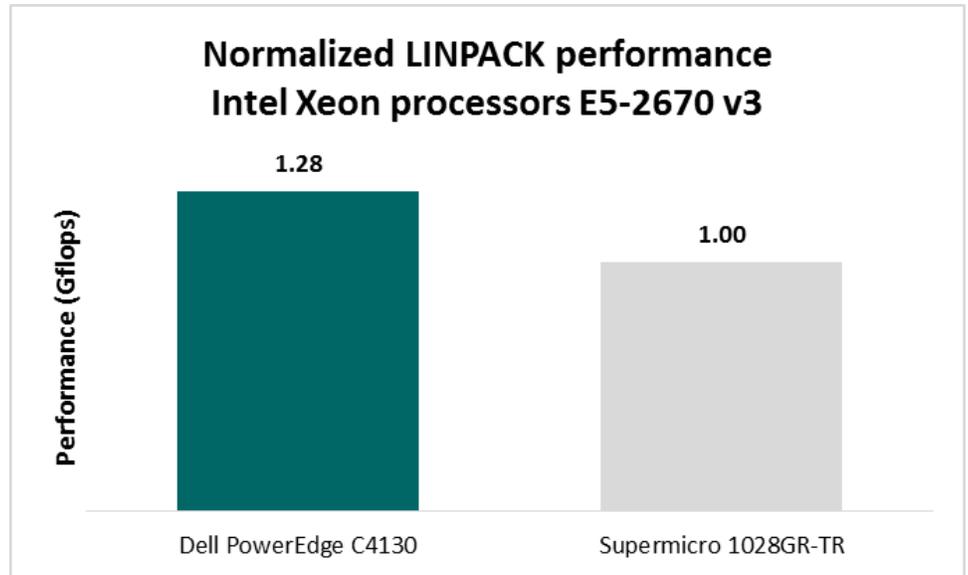
Figure 7: Normalized LINPACK floating-point performance for the maximum configurations of each system. Higher numbers are better.

GPU temperature

For the maximum configuration, the peak GPU temperature of the Dell PowerEdge C4130 was 13 degrees cooler than the Supermicro 1028GR-TR solution (see Figure 8).
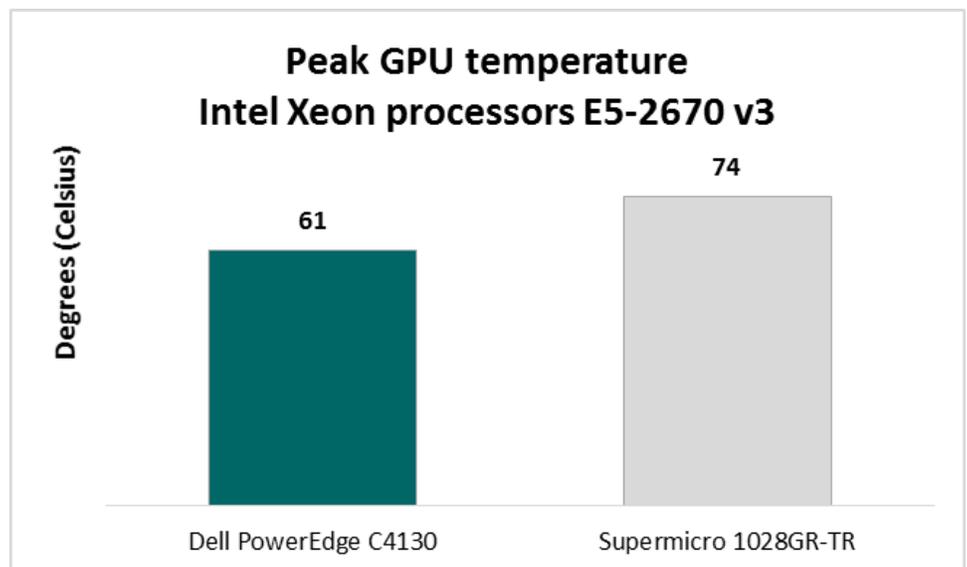
Figure 8: Peak GPU temperature for the maximum configuration of both systems. Lower numbers are better.

Intel Xeon processor E5-2690 v3

Floating-point performance

We found that the Dell PowerEdge C4130 maximum configuration achieved 27 percent greater floating-point performance than the Supermicro 1028GR-TR maximum configuration. Figure 9 shows the normalized Gflops achieved by the two configurations with Intel Xeon processors E5-2690 v3.
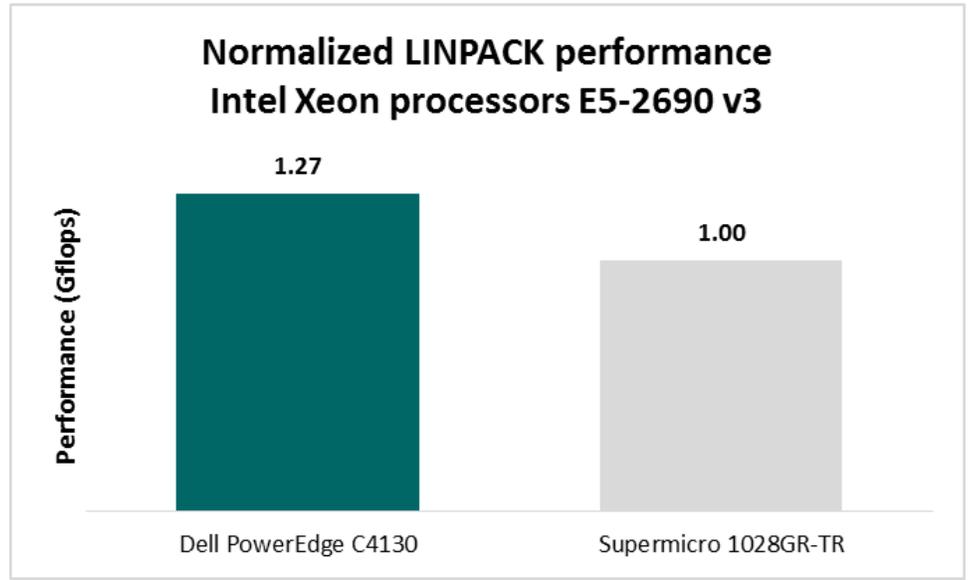
Figure 9: Normalized LINPACK floating-point performance for the maximum configurations of each system. Higher numbers are better.

Normalized LINPACK performance
Intel Xeon processors E5-2690 v3

GPU temperature

For the maximum configuration, the peak GPU temperature of the Dell PowerEdge C4130 was 10 degrees cooler than the Supermicro 1028GR-TR solution (see Figure 10).
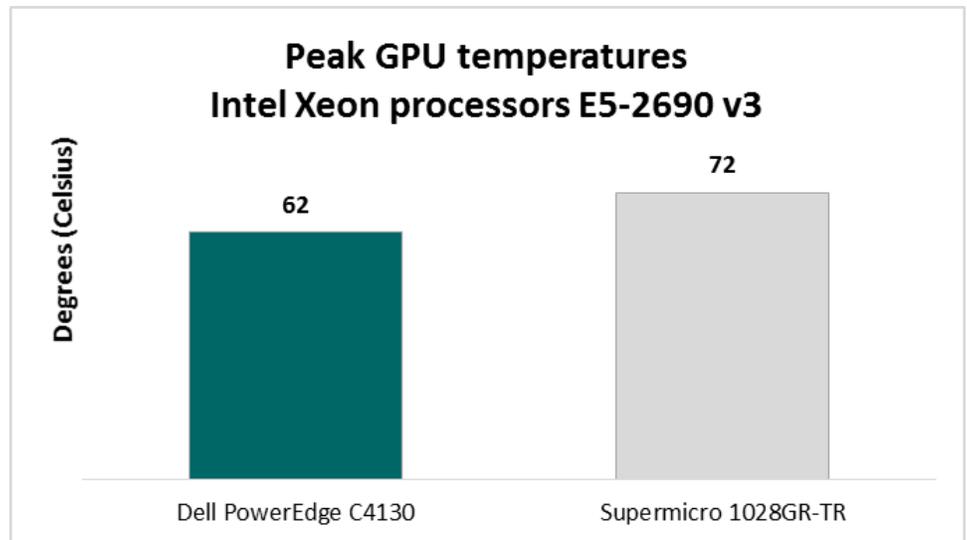
Figure 10: Peak GPU temperature for the maximum configuration of both systems. Lower numbers are better.

Peak GPU temperatures
Intel Xeon processors E5-2690 v3

CONCLUSION

Choosing a server solution that supports additional GPUs is a great option for offloading your HPC workloads and maximizing server performance. We found that the maximum configuration of the Dell PowerEdge C4130 with four NVIDIA Tesla K80 GPU accelerators delivered up to 9.3 times more performance than the Dell PowerEdge C4130 without GPUs. In addition, choosing the right server solution to support GPUs for your HPC workloads can deliver additional performance while maintaining reasonable GPU temperatures. We found that the maximum configuration of the Dell PowerEdge C4130 with four NVIDIA Tesla K80 GPU accelerators delivered up to 28 percent better performance than the maximum configuration of the Supermicro 1028GR-TR with three NVIDIA Tesla K80 GPU accelerators. In our testing of internal temperatures, we found the peak GPU temperature of the Dell PowerEdge C4130 in the maximum configuration to be up to 13 degrees cooler than the Supermicro 1028GR-TR maximum configuration.

The added performance of NVIDIA Tesla K80 GPUs can be valuable for organizations running anything from advanced algorithms to rendering 3D graphics. The new Dell PowerEdge C4130 server supports up to four GPUs, providing the platform your organization needs in its datacenter to handle these compute-intensive workloads. The design of the PowerEdge C4130 helps lower internal GPU temperatures via internal airflow—bringing another benefit for your organization by potentially extending hardware and chip life.

APPENDIX A – ABOUT THE COMPONENTS

About Dell PowerEdge C4130

The Dell PowerEdge C4130 is, according to Dell, "designed to accelerate a range of demanding workloads including high-performance computing (HPC)," and powered by up to two Intel Xeon processors for the Intel Xeon processor E5-2600 v3 product family. The Dell PowerEdge C4130 houses up to four 300W double-width GPU accelerators or coprocessors in 1U of space, can offer up to 256GB of DDR4 memory, and features two rear PCIe® 3.0 slots and support for InfiniBand® FDR.

For more information about the Dell PowerEdge C4130, visit www.dell.com/us/business/p/poweredge-c4130/pd.

About NVIDIA Tesla K80 GPU accelerators

Based on the NVIDIA Kepler™ Architecture, Tesla accelerators are, according to NVIDIA, "designed to deliver faster, more efficient compute performance." They feature a dual GPU board with 24 GB of memory and 4,992 CUDA® parallel processing cores. For more information, visit www.nvidia.com/object/tesla-servers.html.

About LINPACK

The LINPACK benchmark runs a program that solves a system of linear equations to measure the floating-point rate of execution of a system. Often used to test the performance of supercomputers, LINPACK can help determine the peak performance of which a system is capable by using complex calculations to stress the processor. For more information about LINPACK, visit www.top500.org/project/linpack/.

APPENDIX B – SYSTEM CONFIGURATION INFORMATION

Figure 11 provides detailed configuration information for the test systems.

System	Dell PowerEdge C4130	Supermicro 1028GR-TR
Power supplies		
Total number	2	2
Vendor and model number	Dell D1600E-S0	Delta Electronics DPS-1600CB
Wattage of each (W)	1600	1600
Cooling fans		
Total number	8	10
Vendor and model number	Delta Electronics GFM0412SS-D	Delta Electronics GFC0412DS
Dimensions (h × w) of each	1" × 1"	1" × 1"
Volts	12	12
Amps	2.0	1.5
General		
Number of processor packages	2	2
Number of cores per processor	20, 24, 24	20, 24, 24
Number of hardware threads per core	2, 2, 2	2, 2, 2
System power management policy	Performance	Default (performance)
CPU		
Vendor	Intel	Intel
Name	Xeon	Xeon
Model number	E5-2650 v3, E5-2670 v3, E5-2690 v3	E5-2650 v3, E5-2670 v3, E5-2690 v3
Socket type	FCLGA2011-3	FCLGA2011-3
Core frequency (GHz)	2.3, 2.3, 2.6	2.3, 2.3, 2.6
Bus frequency (GT/s)	9.6, 9.6, 9.6	9.6, 9.6, 9.6
L1 cache	768 KB (all processors)	768 KB (all processors)
L2 cache	3,072 KB (all processors)	3,072 KB (all processors)
L3 cache	25, 30, 30 MB	25, 30, 30 MB
Platform		
Vendor and model number	Dell PowerEdge C4130	Supermicro 1028GR-TR
Motherboard model number	0797FV	X10DRG-H
BIOS name and version	Dell 0.3.5	American Megatrends 1.0
BIOS settings	Performance	Default (performance)

System	Dell PowerEdge C4130	Supermicro 1028GR-TR
Memory module(s)		
Total RAM in system (GB)	128	128
Vendor and model number	Hynix HMA41GR7MFR8N	Hynix HMA41GR7MFR8N
Type	PC4-2133P	PC4-2133P
Speed (MHz)	2,133	2,133
Speed running in the system (MHz)	2,133	2,133
Timing/Latency (tCL-tRCD-tRP-tRASmin)	11-11-11-33	11-11-11-33
Size (GB)	8	8
Number of RAM module(s)	16	16
Chip organization	Dual-sided	Dual-sided
Rank	Double	Double
Operating system		
Name	CentOS 7	CentOS 7
Build number	3.10.0-123.9.2.el7.x86_64	3.10.0-123.9.2.el7.x86_64
File system	XFS	XFS
Language	English	English
RAID controller		
Vendor and model number	Embedded Intel SATA Controller	Embedded Intel SATA Controller
Driver version	3.0	3.0
Cache size (MB)	N/A	N/A
Solid-state drives		
Vendor and model number	Dell 1H4WG	Samsung MZ-7PA2560/0D1
Number of drives	1	1
Size (GB)	60	256
Type	SATA SSD	SATA SSD
Ethernet adapters		
Vendor and model number	Intel I350 Gigabit NIC	Intel I350 Gigabit NIC
Type	Ethernet	Ethernet
Driver	5.0.5-k	5.0.5-k

Figure 11: System configuration information for the test systems.

APPENDIX C – HOW WE TESTED

On both the Dell PowerEdge C4130 and the Supermicro 1028GR-TR, we configured the BIOS settings for HPL performance and then installed minimal CentOS 7 onto the local storage. We installed Intel MKL 11.2 and used the included HPL 2.1 executable and libraries to run the baseline (processor only) configurations. For the maximum configurations, we installed NVIDIA's distribution of the HPL 2.1 executable, along with NVIDIA CUDA 6.5, OpenMPI 1.6.5 libraries, and the NVIDIA 340.32 GPU driver.

Installing and configuring the BIOS and operating system

On both servers, we configured the BIOS to run on physical cores only and in the best possible performance settings. We then created a single-drive RAID0 volume using the onboard storage controllers and installed the minimal installation of CentOS 7 onto the local storage. We then installed necessary packages and updated the kernel.

Configuring the BIOS

1. During POST, press the appropriate key to enter the BIOS menu (F2 on the Dell PowerEdge C4130, and Delete on the Supermicro 1028GR-TR).
2. In BIOS, navigate to the Processor menu and turn off Intel Virtualization Technology (Hyper-Threading).
3. On the Dell PowerEdge C4130, navigate to the System Profile menu and choose the Performance profile.
4. On both servers, save the BIOS settings and exit the BIOS menu.

Installing CentOS 7

1. Connect the installation media to the server. We used the virtual optical drive available on both servers' out-of-band management consoles.
2. Boot to the installation media.
3. At the splash screen, select Install CentOS 7 and press Enter.
4. For language, choose your desired language, and click Continue. We chose English (United States).
5. At the Installation Summary screen, configure the Date & Time to match your time zone.
6. Set the Software Selection to Minimal Install.
7. Set the Installation Destination to Automatic partitioning.
8. Configure the Network & Hostname for your testing network.
9. Click Begin Installation.
10. During the installation process, set the Root Password. We elected not to create another user for this setup.
11. Once the installation is completed, disconnect the installation media and click Reboot.

Updating the operating system

1. Begin an SSH session with the server, and log in as the root user.
2. To install the necessary packages for setup, enter the following command:

```
yum -y install kernel-devel libXi-devel libXmu-devel openmpi-devel pciutils
lm_sensors numactl
```

3. To update the kernel and included packages, enter the following command:

```
yum -y update
```

4. Reboot the server.

5. Download the NVIDIA driver package NVIDIA-Linux-x86_64-340.32.run, and run the package to install the NVIDIA 340.32 GPU driver.

6. Download the NVIDIA CUDA package cuda_6.5.14_linux_64.run, and run the package to install NVIDIA CUDA 6.5. When the installer prompts to install the NVIDIA GPU driver, select No.

7. Download the OpenMPI package openmpi-1.6.5-tar.gz, and extract the contents.

8. In the extracted location, run INSTALL to install OpenMPI 1.6.5.

9. Download the Intel Composer XE 2015 evaluation software package l_compxe_2015.0.090.tgz, and extract the contents.

10. In the extracted location, run install.sh to install the Intel MKL 11.2 software.

11. Reboot the server.

Installing and running the HPL benchmark

For the baseline (processor only) configurations, we used the HPL 2.1 executable included with Intel MKL 11.2. For the maximum configurations of GPU accelerators, we used a Dell-provided pre-compiled version of NVIDIA's HPL 2.1 benchmark. We created a series of bash scripts to run the benchmark.

Configuring the HPL benchmark for the baseline configuration

1. Navigate to /opt/intel/mkl/benchmarks/linpack.

2. Edit the OMP_NUM_THREADS environment variable in runme_xeon64 to reflect the number of physical cores in the system (20 for the Intel Xeon E5-2650 v3, and 24 for both the Intel Xeon E5-2670 v3 and the Intel Xeon E5-2690 v3).

3. Add the following lines to the beginning of runme_xeon64:

```
export LD_LIBRARY_PATH=
$LD_LIBRARY_PATH:/opt/intel/composer_xe_2015.0.090/mkl/lib/intel64/:/opt/int
el/composer_xe_2015.0.090/compiler/lib/intel64/

export PATH=$PATH:/opt/intel/composer_xe_2015.0.090/mpirt/bin/intel64/
```

4. Edit lininput_xeon64 to the following:

```
Sample Intel(R) Optimized LINPACK Benchmark data file (lininput_xeon64)
Intel(R) Optimized LINPACK Benchmark data
1      # number of tests
121344 # problem sizes
121344 # leading dimensions
1      # times to run a test
32     # alignment values (in KBytes)
```

Running the HPL benchmark for the baseline configuration

1. Navigate to /opt/intel/mkl/benchmarks/linpack.

2. Run runme_xeon64 to run the HPL 2.1 benchmark.

Configuring the HPL benchmark for the Dell PowerEdge C4130 maximum configuration (four GPUs)

1. Download the pre-compiled HPL 2.1 executable and library files to a directory on the server.

2. Create a file in the HPL folder called HPL.dat and edit it to the following:

```
HPLinpack benchmark input file
Innovative Computing Laboratory, University of Tennessee
HPL.out      output file name (if any)
6            device out (6=stdout,7=stderr,file)
1            # of problems sizes (N)
121344        Ns
1            # of NBs
1024         NBs
0            PMAP process mapping (0=Row-,1=Column-major)
1            # of process grids (P x Q)
2          Ps
4          Qs
16.0         threshold
1            # of panel fact
0 1 2        PFACTs (0=left, 1=Crout, 2=Right)
1            # of recursive stopping criterium
2            NBMINs (>= 1)
1            # of panels in recursion
2            NDIVs
1            # of recursive panel fact.
2 1 0        RFACTs (0=left, 1=Crout, 2=Right)
1            # of broadcast
3 0 1 2 4 5  BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4=Lng,5=LnM)
1            # of lookahead depth
0            DEPTHs (>=0)
1            SWAP (0=bin-exch,1=long,2=mix)
64           swapping threshold
1            L1 in (0=transposed,1=no-transposed) form
0            U  in (0=transposed,1=no-transposed) form
1            Equilibration (0=no,1=yes)
32            memory alignment in double (> 0)
```

3. For the configuration using Intel E5-2650 v3 processors (20 cores), create a file in the HPL directory called dell_hpl_2650, and edit it to the following:

```
#!/bin/bash

export LD_LIBRARY_PATH=$LD_LIBRARY_PATH: /root/hpl:/usr/local/cuda-
6.5/lib64:/root/openmpi-1.6.5/ompi/.libs/

nvidia-persistenced --persistence-mode
nvidia-smi -pm 1

export HPL_DIR=/root/hpl
export MKL_NUM_THREADS=2  #number of cpu cores per GPU
export OMP_NUM_THREADS=2  #number of cpu cores per process
export MKL_DYNAMIC=FALSE
```

```
export CUDA_DGEMM_SPLIT=.999
export CUDA_DTRSM_SPLIT=.999
export MKL_CBWR=AVX2

mpirun -np 8 /root/hpl/xhpl
```

4. For the configuration using Intel E5-2670 v3 processors and the configuration using Intel E5-2690 v3 processors (24 cores each), create a file in the HPL directory called dell_hpl_2670_2690, and edit it to the following:

```
#!/bin/bash

export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/root/hpl:/usr/local/cuda-
6.5/lib64:/root/openmpi-1.6.5/ompi/.libs/

nvidia-persistenced --persistence-mode
nvidia-smi -pm 1

export HPL_DIR=/root/hpl
export MKL_NUM_THREADS=3  #number of cpu cores per GPU
export OMP_NUM_THREADS=3  #number of cpu cores per process
export MKL_DYNAMIC=FALSE
export CUDA_DGEMM_SPLIT=.999
export CUDA_DTRSM_SPLIT=.999
export MKL_CBWR=AVX2

mpirun -np 8 /root/hpl/xhpl
```

5. Run dell_hpl_2650 or dell_hpl_2670_2690 to run the HPL 2.1 benchmark for the appropriate configuration.

Configuring the HPL benchmark for the Supermicro 1028GR-TR maximum configuration (3 GPUs)

1. Download the pre-compiled HPL 2.1 executable and library files to a directory on the server.
2. Create a file in the HPL folder called HPL.dat and edit it to the following:

```
HPLinpack benchmark input file
Innovative Computing Laboratory, University of Tennessee
HPL.out      output file name (if any)
6            device out (6=stdout,7=stderr,file)
1            # of problems sizes (N)
121344        Ns
1            # of NBs
1024         NBs
0            PMAP process mapping (0=Row-,1=Column-major)
1            # of process grids (P x Q)
2          Ps
3          Qs
16.0        threshold
1            # of panel fact
0 1 2        PFACTs (0=left, 1=Crout, 2=Right)
1            # of recursive stopping criterium
```

```
2               NBMINs (>= 1)
1               # of panels in recursion
2               NDIVs
1               # of recursive panel fact.
2 1 0           RFACTs (0=left, 1=Crout, 2=Right)
1               # of broadcast
3 0 1 2 4 5     BCASTs (0=1rg,1=1rM,2=2rg,3=2rM,4=Lng,5=LnM)
1               # of lookahead depth
0               DEPTHs (>=0)
1               SWAP (0=bin-exch,1=long,2=mix)
64              swapping threshold
1               L1 in (0=transposed,1=no-transposed) form
0               U  in (0=transposed,1=no-transposed) form
1               Equilibration (0=no,1=yes)
32               memory alignment in double (> 0)
```

3. For the configuration using Intel E5-2650 v3 processors (20 cores), create a file in the HPL directory called `supermicro_hpl_2650` and edit it to the following:

```
#!/bin/bash

export LD_LIBRARY_PATH=$LD_LIBRARY_PATH: /root/hpl:/usr/local/cuda-
6.5/lib64:/root/openmpi-1.6.5/ompi/.libs/

nvidia-persistenced --persistence-mode
nvidia-smi -pm 1

export HPL_DIR=/root/hpl
export MKL_NUM_THREADS=3  #number of cpu cores per GPU
export OMP_NUM_THREADS=3  #number of cpu cores per process
export MKL_DYNAMIC=FALSE
export CUDA_DGEMM_SPLIT=.999
export CUDA_DTRSM_SPLIT=.999
export MKL_CBWR=AVX2

mpirun -np 6 /root/hpl/xhpl
```

4. For the configuration using Intel E5-2670 v3 processors and the configuration using Intel E5-2690 v3 processors (24 cores each), create a file in the HPL directory called supermicro_hpl_2670_2690 and edit it to the following:

```
#!/bin/bash

export LD_LIBRARY_PATH=$LD_LIBRARY_PATH: /root/hpl:/usr/local/cuda-
6.5/lib64:/root/openmpi-1.6.5/ompi/.libs/

nvidia-persistenced --persistence-mode
nvidia-smi -pm 1

export HPL_DIR=/root/hpl
```

```
export MKL_NUM_THREADS=4   #number of cpu cores per GPU
export OMP_NUM_THREADS=4   #number of cpu cores per process
export MKL_DYNAMIC=FALSE
export CUDA_DGEMM_SPLIT=.999
export CUDA_DTRSM_SPLIT=.999
export MKL_CBWR=AVX2

mpirun -np 6 /root/hpl/xhpl
```

5. Run supermicro_hpl_2650 or supermicro_hpl_2670_2690 to run the HPL 2.1 benchmark for the appropriate
   configuration.

ABOUT PRINCIPLED TECHNOLOGIES

Principled Technologies®

Principled Technologies, Inc.
1007 Slater Road, Suite 300
Durham, NC, 27703
www.principledtechnologies.com

We provide industry-leading technology assessment and fact-based marketing services. We bring to every assignment extensive experience with and expertise in all aspects of technology testing and analysis, from researching new technologies, to developing new methodologies, to testing with existing and new tools.

When the assessment is complete, we know how to present the results to a broad range of target audiences. We provide our clients with the materials they need, from market-focused data to use in their own collateral to custom sales aids, such as test reports, performance assessments, and white papers. Every document reflects the results of our trusted independent analysis.

We provide customized services that focus on our clients' individual requirements. Whether the technology involves hardware, software, Web sites, or services, we offer the experience, expertise, and tools to help our clients assess how it will fare against its competition, its performance, its market readiness, and its quality and reliability.

Our founders, Mark L. Van Name and Bill Catchings, have worked together in technology assessment for over 20 years. As journalists, they published over a thousand articles on a wide array of technology subjects. They created and led the Ziff-Davis Benchmark Operation, which developed such industry-standard benchmarks as Ziff Davis Media's Winstone and WebBench. They founded and led eTesting Labs, and after the acquisition of that company by Lionbridge Technologies were the head and CTO of VeriTest.