



GPU Server Portfolio Overview

Neil Truong – Product Manager, Supermicro
Jason Pai – Product Manager, Supermicro



GPU Solutions – HPC Optimized



“The most comprehensive product line in the Industry”

- ✓ High efficiency power supplies at full capacity
- ✓ Excellent thermal design
- ✓ Non-blocking air-flow
- ✓ Greatest performance layout
- ✓ No re-driver required; no latency

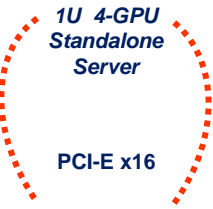


2016 – Next Gen GPU Innovation

Tesla S1070



1U 4-GPU Standalone Server



PCI-E x16



1U Twin™

2008 – GPGPU Where it started...

The fastest 1U server in the world



7U 10-blade 20-GPUs



4U 8-GPUs



2015 – 1U 4-GPU Optimized



GPU Server & Workstation



2013 – GPU FatTwin™

Integrated GPU Server



2009 – Hybrid Computing Pioneer



2011 – GPU Blades





X10 GPU Server Portfolio



Ratio:
GPU:CPU

GPU OPTIMIZED

TOWER

GPU




7048GR
4:2 (4U)

RACK

GPU



4028GR
8:2 (4U)



1028GQ
4:2 (1U)



2028GR
6:2 (2U)



1028GR
3:2 (1U)



1018GR/5018GR
2:1 (1U)

MULTI-NODE

FatTwin



F628G3/F628G2
3:2 (4U / 4Node)



F648G2
6:2 (4U / 2Node)

Blade



7128RG
2:2 (7U / 10Node)



Server Portfolio with Highest Rack-level GPU Density

Best-in-class technology designed for highly parallel applications to deliver ultimate performance, flexibility, and scalability

2

1018GR

Cost Effective



- Single Haswell/Broadwell CPU
- 8 DDR4 DIMMs
- 6x 2.5" HS HDD bays
- 2 Double-Width GPUs
- 1 x8 PCIe 3.0 slot
- 1x 1400W Platinum PWS

3

1028GR

Mainstream

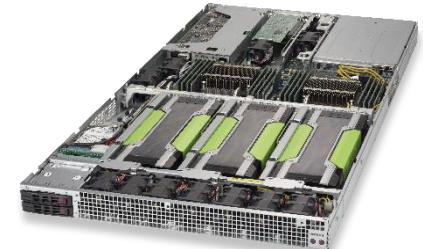


- Dual Haswell/Broadwell CPUs
- 16 DDR4 DIMMs
- 4x 2.5" HS HDD bays
- 3 Double-Width GPUs
- 1 x8 PCIe 3.0 slot
- 2x 1600W Platinum PWS

4

1028GQ

Parallel Optimized



- Dual Haswell/Broadwell CPUs
- 16 DDR4 DIMMs
- 4x 2.5" HS HDD bays
- 4 Double-Width GPUs
- Active/Passive GPUs
- 2 x8 PCIe 3.0 Slots
- 2x 2000W Platinum PWS



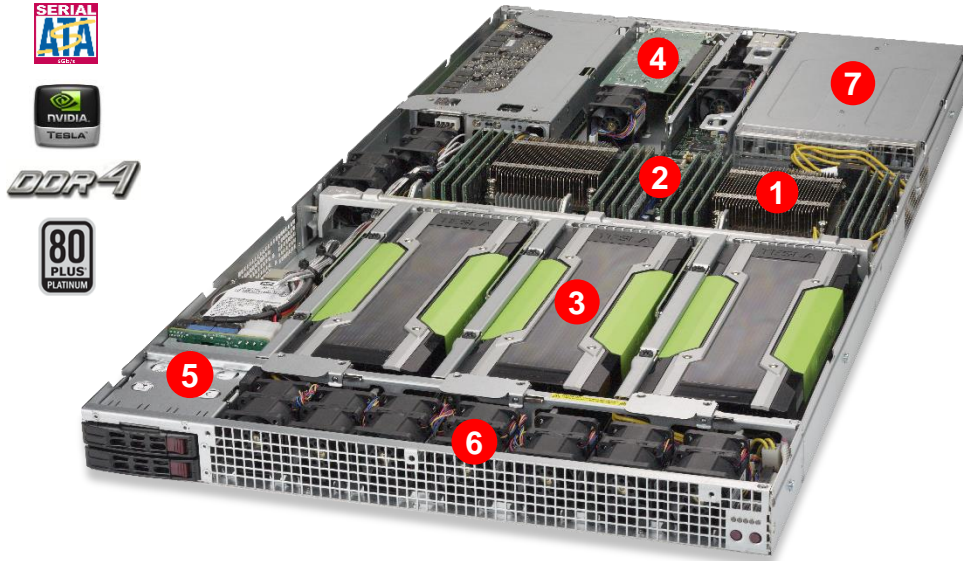
1993

2009

2015

5X

GPU: 1U DP SYS-1028GQ-TR(T)



- 1 Processor Support**
Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)

- 2 Memory Capacity**
16 DIMMs, up to 2TB ECC LRDIMM, 512GB ECC RDIMM, DDR4, up to 2400MHz

- 3 Expansion Slots**
4 PCI-e x16 Gen 3 for double-width GPU cards
2 x8 (in x16 slot) LP card

- 4 I/O ports**
1x VGA, 2x GbE or 2x 10GbaseT LAN, 2x USB 3.0, and 1x IPMI dedicated LAN port

- 5 Drive Bays**
2 hot-swap 2.5" drives bays; 4 total 2.5" HDD bays

- 6 System Cooling**
9 counter rotating fans with optimal fan speed control

- 7 Power Supply**
2000W Platinum Level efficiency redundant power supply

Motherboard: X10DGQ

Chassis: CSE-118GQETS-R2K03P

Key Features:

- Supports up to 4 double width GPU cards (including GTX)
- Redundant Platinum Level 2000W power supplies
- No GPU-Preheat
- Cost Optimized System

Key Applications:

- Oil & Gas
- Research & Scientifics
- VDI technology
- Computational Finance



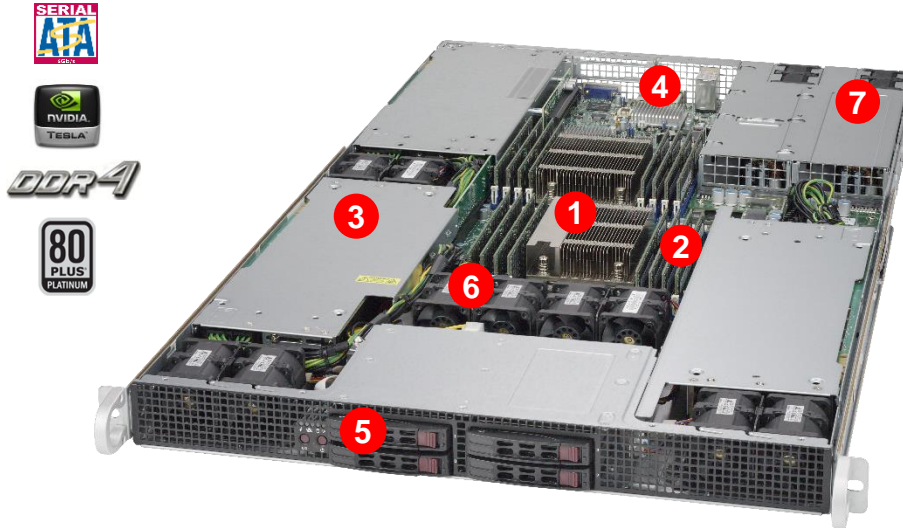
1993

2009

2015

5X

GPU: 1U DP SYS-1028GR-TR(T)



Motherboard: X10DRG-H(T)
Chassis: CSE-118GH-R1K66B

Key Features:

- Supports up to 3 double width GPU cards
- 4 Hot Swap 2.5" HDD bays
- 16 DIMM, up to 2TB memories
- Redundant Platinum Level 1600W power supplies

- 1 Processor Support**
Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)
- 2 Memory Capacity**
16 DIMMs, up to 2TB ECC LRDIMM, 512GB ECC RDIMM, DDR4, up to 2400MHz
- 3 Expansion Slots**
4 PCI-e x16 Gen 3 for double width GPU cards
1 x8 (in x16 slot) LP card
- 4 I/O ports**
1x VGA, 2x GbE or 2x 10GbaseT LAN, 2x USB 2.0, and 1x IPMI dedicated LAN port.
- 5 Drive Bays**
4 hot-swap 2.5" drive bays
- 6 System Cooling**
10 counter rotating fans with optimal fan speed control
1 air shroud
- 7 Power Supply**
1600W Platinum Level efficiency redundant power supply

Key Applications:

- VDI technology
- HPC
- Machine Learning
- Computational Finance



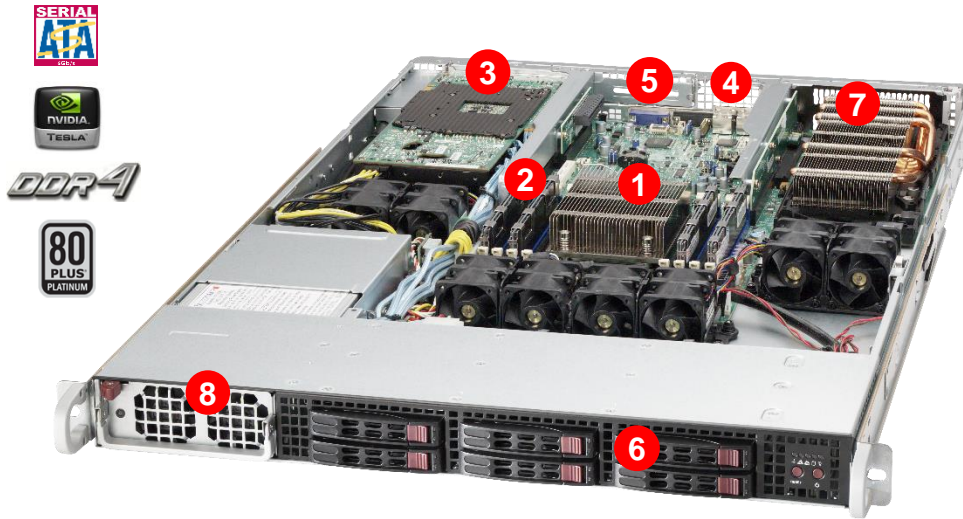
1993

2009

2015

5X

GPU: 1U DP SYS-1018GR-T/5018GR-T



- 1 Processor Support**
Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)

- 2 Memory Capacity**
8 DIMM, up to 1TB ECC LRDIMM, 256GB ECC RDIMM, DDR4, up to 2400MHz

- 3 Expansion Slots**
2 PCI-e x16 Gen 3 for double width GPU cards
1 x8(in x16) Gen 3 LP card

- 4 I/O ports**
1x VGA, 1x COM, 2x GbE LAN, 2x USB 3.0, and 1x IPMI dedicated LAN port

- 5 System Management**
On board BMC (Baseboard Management Controllers) supports IPMI2.0, media/KVM over LAN. (Dedicated LAN port for management)

- 6 Drive Bays**
1018GR: 6 Hot-Swap 2.5" Drive Bays
5018GR: 3 Hot-Swap 3.5" Drive Bays

- 7 System Cooling**
8 counter rotating fans w/ optimal fan speed control

- 8 Power Supply**
1400W Platinum Level efficiency power supply

Motherboard: X10SRG-F

Chassis: CSE-118GH-R1K66B

Key Features:

- Supports 2 double width GPU cards
- GPU optimized UP server
- 1400W Platinum Level super high efficiency power supply

Key Applications:

- Oil & Gas/Seismic
- Scientific/Data Mining



1993

2009

2015

5X

Server Portfolio with Highest Node-level GPU Density

Best-in-class technology designed for highly parallel applications to deliver ultimate performance, flexibility, and scalability

4

7048GR

Mission Critical



- 4U Chassis
- Dual Haswell/Broadwell CPUs w/ IPMI
- 16 DDR4 DIMMs
- 8x 3.5" HS HDD bays
- 4 Double-Width GPUs
- x16/x8/x4 – 4/2/1**
- 2x 2000W Titanium PWS

6

2028GR

Mainstream



- 2U Chassis
- Dual Haswell/Broadwell CPUs
- 16 DDR4 DIMMs
- 10x 2.5" HS HDD bays
- 6 Double-Width GPUs
- 1 x8 PCIe 3.0 slot
- 2x 2000W Platinum PWS

8

4028GR

Parallel Optimized



- 4U Chassis
- Dual Haswell/Broadwell CPUs
- 24 DDR4 DIMMs
- 24x 2.5" HS HDD bays
- 8 Double-Width GPUs
- 2 x8 PCIe 3.0 slot; 1 x4 PCIe 2.0 slot
- 4x 1600W Platinum PWS

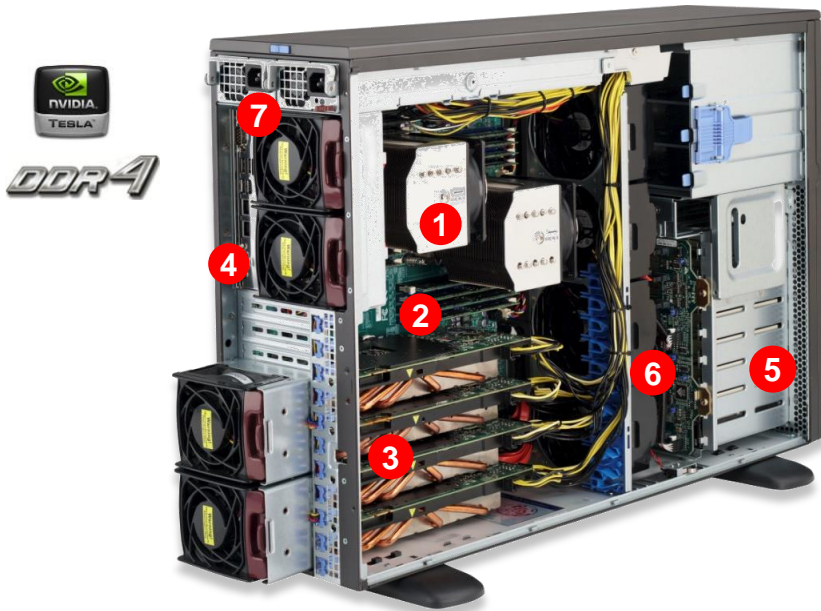


1993

2009

2015

GPU: 4U DP SYS-7048GR-TR



Motherboard: X10DRG-Q

Chassis: CSE-747TQ-R2K04B

Key Features:

- Supports 4 double width GPU cards
- Support up to 160W CPUs
- Optional GPU kits to support passive GPU
- 8x 3.5 inches HDD bays
- Supports Thunderbolt 2.0

1 Processor Support
Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)

2 Memory Capacity
16 DIMMs, up to 2TB ECC LRDIMM, 512GB ECC RDIMM, DDR4, up to 2400MHz

3 Expansion Slots
4 PCI-e 3.0 x16 for double width GPU cards, 2 PCI-e 3.0 x8 (1 in x16 slot), 1 PCI-e 2.0 x4 (in x8)

4 I/O ports
1x VGA, 2 COM, 2x GbE LAN, 6x USB 3.0, and 1x IPMI dedicated LAN port, Audio 7.1

5 Drive Bays
1x VGA, 2 COM, 2x GbE LAN, 6x USB 3.0, and 1x IPMI dedicated LAN port, Audio 7.1

6 System Cooling
4 heavy duty fans, 4 exhaust fans, and 2 active heat sink w/ Optimal Fan Speed Control

7 Power Supply
2000W Titanium Level efficiency redundant power supplies

Key Applications:

- Imaging and Computer Visio
- Simulation and Creation Design

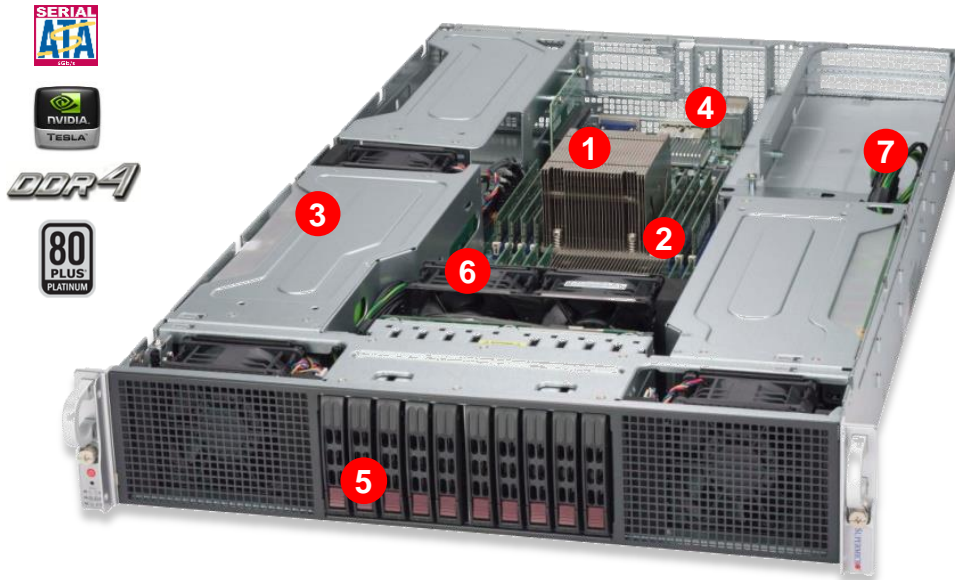


1993

2009

2015

GPU: 2U DP SYS-2028GR-TR(T)/TRH(T)



Motherboard: X10DRG-H(T)
Chassis: CSE-218GH-R2K03B

Key Features:

- Support up to 6 double width GPU cards
- 10 hot-swap 2.5" HDD bays
- 16 DIMMs, up to 2TB memory
- Platinum level 2000W power supply

1 Processor Support
 Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)

2 Memory Capacity
 16 DIMMs, up to 2TB ECC LRDIMM, 512GB ECC RDIMM, DDR4, up to 2400MHz

3 Expansion Slots
 -TR(T) : 4 PCI-e x16 Gen 3 for double width GPU cards
 1 x8 (in x16 slot) LP card
 -TRH(T) : 6 PCI-e x16 Gen 3 for double width GPU cards
 1 x8 (in x16 slot) LP card

4 I/O ports
 1x VGA, 2x GbE or 2x 10GbaseT LAN, 2x USB 2.0, and 1x IPMI dedicated LAN port.

5 Drive Bays
 10 hot-swap 2.5" drive bays

6 System Cooling
 5 counter rotating fans with optimal fan speed control
 1 air shroud

7 Power Supply
 2000W Platinum Level efficiency redundant power supply

Key Applications:

- Computational Finance
- HPC/Oil and gas
- Weather and Climate Analysis



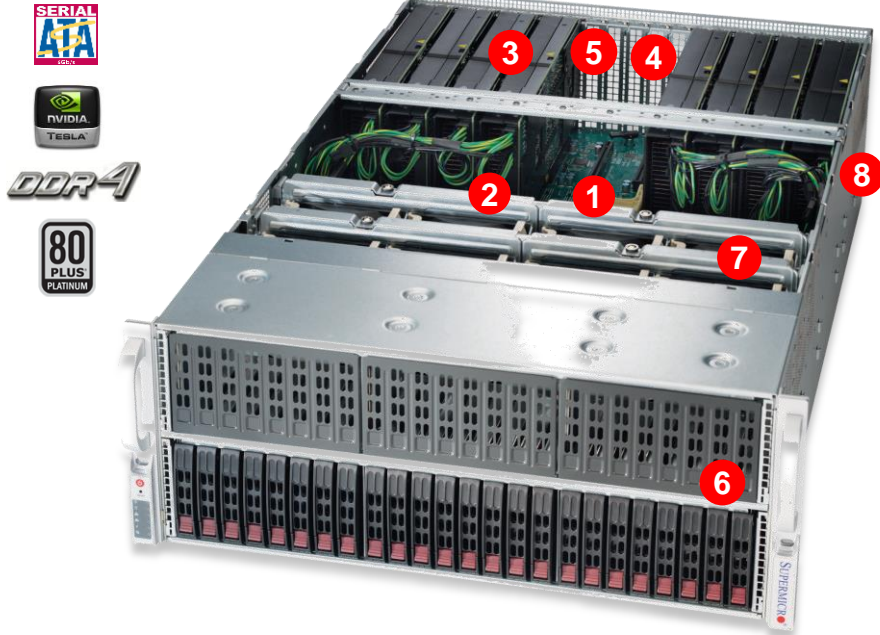
1993

2009

2015

5X

GPU: 4U DP SYS-4028GR-TR(T)



Motherboard: X10DRG-O(T)+-CPU

Chassis: CSE-418GTS-R3200B

Key Features:

- Supports 8 double width GPU cards
- Support up to 160W CPU
- 24x 2.5" SSD/HDD bays
- 24 DIMMs, up to 3TB memory

- Processor Support**
Dual Xeon E5-2600 v4/v3 CPUs (Socket R3)
 - Memory Capacity**
24 DIMMs, 3TB ECC LRDIMM/RDIMM DDR4 2400MHz
 - Expansion Slots**
8 PCI-e 3.0 x16 for double width GPU cards
2 PCI-e 3.0 x8 (2 in x16 slots)
1 PCI-e 2.0 x4 (in x16)
 - I/O ports**
1x VGA, 2x Gbit LAN, 4x USB 2.0, and 1x IPMI dedicated LAN port
 - System management**
On board BMC (Baseboard Management Controllers) supports IPMI2.0, media/KVM over LAN with dedicated LAN for system management
 - Drive Bays**
24 hot-swap 2.5" drives bay
 - System Cooling**
8 heavy duty fans optimize to support 8 GPU cards
2 air shroud
 - Power Supply**
4 x 1600W (2+2) Platinum Level efficiency redundant power supply
- Dimensions**
H 17.2" (452mm) x W 7" (178mm) x D 29" (673mm)



1993

2009

2015

3rd Party Server Reviews



Rating:

97%

“As we have said before, the case used for the **7048GR-TR** Workstation is simply the best in quality, craftsmanship, and features. ... it is our go-to case every time. The **7048GR-TR** Workstation is designed for maximum uptime with hot-swappable drives and cooling fans, and includes dual redundant power supplies.”

— **TweakTown**

“Overall, for those looking to cram four GPU’s into a small 1U form factor for dense compute or even VDI applications, the Supermicro **1028GQ-TRT** is an excellent solution. With 10Gbase-T networking, the server is easy to integrate into existing datacenter infrastructure so long as the rack is able to handle higher-power rated gear.

... we find the **4028GR-TR** is a well designed system that has the ability to handle high performance work loads. Moving to a large 4U server allows larger capacity cooling systems to be installed that keep the system cool while running extreme work loads. This is a trade off vs smaller 1U systems which have higher density but operate at close to maximum heat load capacities.”

— **ServeTheHome**



Rating:

9.7

STAC-A2 Benchmarks

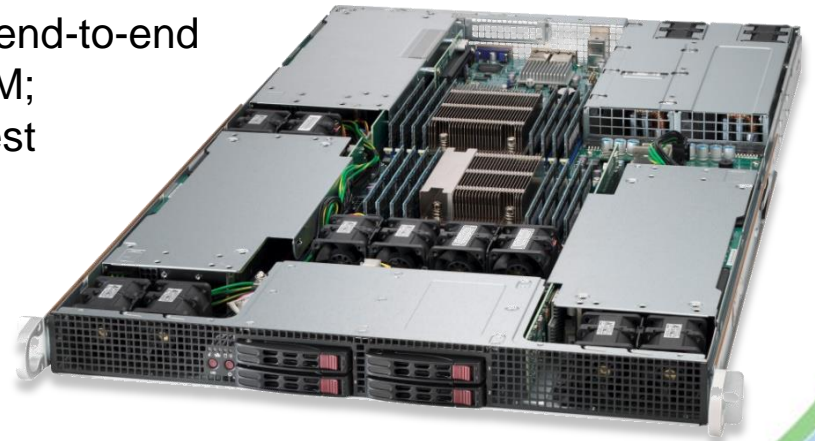
The STAC-A2 Benchmark suite is the industry standard for testing technology stacks used for compute-intensive analytic workloads involved in pricing and risk management. In all, the STAC-A2 specifications deliver nearly 200 test results related to performance, scaling, efficiency, and quality, which are detailed in this report.



Test System: Supermicro SYS-1028GR-TR server

World Record Results

Fastest warm time to date in the baseline end-to-end Greeks benchmark: GREEKS.TIME.WARM;
This was 1.27x the speed of the next fastest system, a 4-way Haswell-EX system (SUT ID: INTC150811).



1993

2009

2015

5X