

## NVIDIA Powers the World's Top 13 Most Energy Efficient Supercomputers

ISC -- Advancing the path to exascale computing, NVIDIA today announced that the [NVIDIA® Tesla® AI supercomputing platform](#) powers the top 13 measured systems on the new [Green500](#) list of the world's most energy-efficient high performance computing (HPC) systems. All 13 use NVIDIA [Tesla P100](#) data center GPU accelerators, including four systems based on the [NVIDIA DGX-1™ AI supercomputer](#).

- As Moore's Law slows, NVIDIA Tesla GPUs continue to extend computing, improving performance 3X in two years
- Tesla V100 GPUs projected to provide U.S. Energy Department's Summit supercomputer with 200 petaflops of HPC, 3 exaflops of AI performance
- Major cloud providers commit to bring NVIDIA Volta GPU platform to market

Advancing the path to exascale computing, NVIDIA today announced that the [NVIDIA® Tesla® AI supercomputing platform](#) powers the top 13 measured systems on the new [Green500](#) list of the world's most energy-efficient high performance computing (HPC) systems. All 13 use NVIDIA [Tesla P100](#) data center GPU accelerators, including four systems based on the [NVIDIA DGX-1™ AI supercomputer](#).

NVIDIA today also released performance data illustrating that NVIDIA Tesla GPUs have improved performance for HPC applications by 3X over the Kepler architecture released two years ago. This significantly boosts performance beyond what would have been predicted by Moore's Law, even before it began slowing in recent years.

Additionally, NVIDIA announced that its Tesla V100 GPU accelerators -- which combine AI and traditional HPC applications on a single platform -- are projected to provide the U.S. Department of Energy's (DOE's) [Summit supercomputer](#) with 200 petaflops of 64-bit floating point performance and over 3 exaflops of AI performance when it comes online later this year.

### NVIDIA GPUs Fueling World's Greenest Supercomputers

The Green500 list, released today at the International Supercomputing Show in Frankfurt, is topped by the new [TSUBAME 3.0](#) system, at the Tokyo Institute of Technology, powered by NVIDIA Tesla P100 GPUs. It hit a record 14.1 gigaflops per watt -- 50 percent higher efficiency than the previous top system -- NVIDIA's own [SATURNV](#), which ranks No. 10 on the latest list.

Spots two through six on the new list are clusters housed at Yahoo Japan, Japan's National Institute of Advanced Industrial Science and Technology, Japan's Center for Advanced Intelligence Project (RIKEN), the University of Cambridge and the Swiss National Computing Center (CSCS), home to the newly crowned fastest supercomputer in Europe, Piz Daint. Other key systems in the top 13 measured systems powered by NVIDIA include E4 Computer Engineering, University of Oxford, and the University of Tokyo.

Systems built on NVIDIA's DGX-1 AI supercomputer -- which combines NVIDIA Tesla GPU accelerators with a fully optimized AI software package -- include RAIDEN at RIKEN, JADE at the University of Oxford, a hybrid cluster at a major social media and technology services company and NVIDIA's own SATURNV.

"Researchers taking on the world's greatest challenges are seeking a powerful, unified computing architecture to take advantage of HPC and the latest advances in AI," said Ian Buck, general manager of Accelerated Computing at NVIDIA. "Our AI supercomputing platform provides one architecture for computational and data science, providing the most brilliant minds a combination of capabilities to accelerate the rate of innovation and solve the unsolvable."

"With TSUBAME 3.0 supercomputer our goal was to deliver a single powerful platform for both HPC and AI with optimal energy efficiency as one of the flagship Japanese national supercomputers," said Professor Satoshi Matsuoka of the Tokyo Institute of Technology. "The most important point is that we achieved this result with a top-tier production machine of multi-petascale. NVIDIA Tesla P100 GPUs allowed us to excel at both these objectives so we can provide this revolutionary AI supercomputing platform to accelerate our scientific research and education of the country."

### Volta: Leading the Path to Exascale

NVIDIA revealed progress toward achieving exascale levels of performance, with anticipated leaps in speed, efficiency and AI computing capability for the Summit supercomputer, scheduled for delivery later this year to the Oak Ridge Leadership Computing Facility, a DOE Office of Science User Facility at Oak Ridge National Laboratory.

Featuring [Tesla V100](#) GPU accelerators, Summit is projected to deliver 200 petaflops of performance -- compared with 93 petaflops for the world's current fastest system, China's TaihuLight. Additionally, Summit is expected to have strong AI computing capabilities, achieving more than 3 exaflops of half-precision Tensor Operations.

"AI is extending HPC and together they are accelerating the pace of innovation to help solve some of the world's most important challenges," said Jeff Nichols, associate laboratory director of the Computing and Computational Science Directorate at Oak Ridge National Laboratory. "Oak Ridge's pre-exascale supercomputer, Summit, is powered by NVIDIA Volta GPUs that provide a single unified architecture that excels at both AI and HPC. We believe AI supercomputing will unleash breakthrough results for researchers and scientists."

### Volta on Every Major Cloud

The extreme computing capabilities of the V100 GPU accelerators will be available later this year as a service through several of the world's leading cloud service providers. Companies that have [stated](#) their enthusiasm and planned support for Volta-based services include Amazon Web Services, Baidu, Google Cloud Platform, Microsoft Azure and Tencent.

### Volta: Ultimate Architecture for AI Supercomputing

To extend the reach of Volta, NVIDIA also announced it is making new [Tesla V100 GPU accelerators](#) available in a PCIe form factor for standard servers. With PCIe systems, as well as previously announced systems using [NVIDIA NVLink™ interconnect technology](#), coming to market, Volta promises to revolutionize HPC and bring groundbreaking AI technology to supercomputers, enterprises and clouds.

Specifications of the PCIe form factor include:

- 7 teraflops double-precision performance, 14 teraflops single-precision performance and 112 teraflops half-precision performance with NVIDIA GPU BOOST™ technology
- 16GB of CoWoS HBM2 stacked memory, delivering 900GB/sec of memory bandwidth
- Support for PCIe Gen 3 interconnect (up to 32GB/sec bi-directional bandwidth)
- 250 watts of power

NVIDIA Tesla V100 GPU accelerators for PCIe-based systems are expected to be available later this year from NVIDIA reseller partner and manufacturers, including Hewlett Packard Enterprise (HPE).

"HPE is excited to complement our purpose-built HPE Apollo systems innovation for deep learning and AI with the unique, industry-leading strengths of the NVIDIA Tesla V100 technology architecture to accelerate insights and intelligence for our customers," said Bill Mannel, vice president and general manager of HPC and AI at Hewlett Packard Enterprise. "HPE will support NVIDIA Volta with PCIe interconnects in three different systems in our portfolio and provide early access to NVLink 2.0 systems to address emerging customer demand."

More information on the NVIDIA Tesla supercomputing platform is available at [www.nvidia.com/tesla](http://www.nvidia.com/tesla).

Certain statements in this press release including, but not limited to, statements as to: the benefits, impact, performance and availability of the NVIDIA Tesla AI supercomputing platform, NVIDIA Tesla P100 GPUs, NVIDIA Tesla V100 GPU accelerators, the Summit supercomputer, AI, HPC, Volta, a PCIe form factor of Tesla V100 GPU accelerators and planned partner support are forward-looking statements that are subject to risks and uncertainties that could cause results to be materially different than expectations. Important factors that could cause actual results to differ materially include: global economic conditions; our reliance on third parties to manufacture, assemble, package and test our products; the impact of technological development and competition; development of new products and technologies or enhancements to our existing product and technologies; market acceptance of our products or our partners' products; design, manufacturing or software defects; changes in consumer preferences or demands; changes in industry standards and interfaces; unexpected loss of performance of our products or technologies when integrated into systems; as well as other factors detailed from time to time in the reports NVIDIA files with the Securities and Exchange Commission, or SEC, including its Form 10-Q for the fiscal period ended April 30, 2017. Copies of reports filed with the SEC are posted on the company's website and are available from NVIDIA without charge. These forward-looking statements are not guarantees of future performance and speak only as of the date hereof, and, except as required by law, NVIDIA disclaims any obligation to update these forward-looking statements to reflect future events or circumstances.

© 2017 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, Tesla, NVIDIA DGX-1 and NVIDIA NVLink are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. Features, pricing, availability and specifications are subject to change without notice.

#### **Media Contacts**

Kristin Bryson  
+1 203 241 9190  
[kbryson@nvidia.com](mailto:kbryson@nvidia.com)