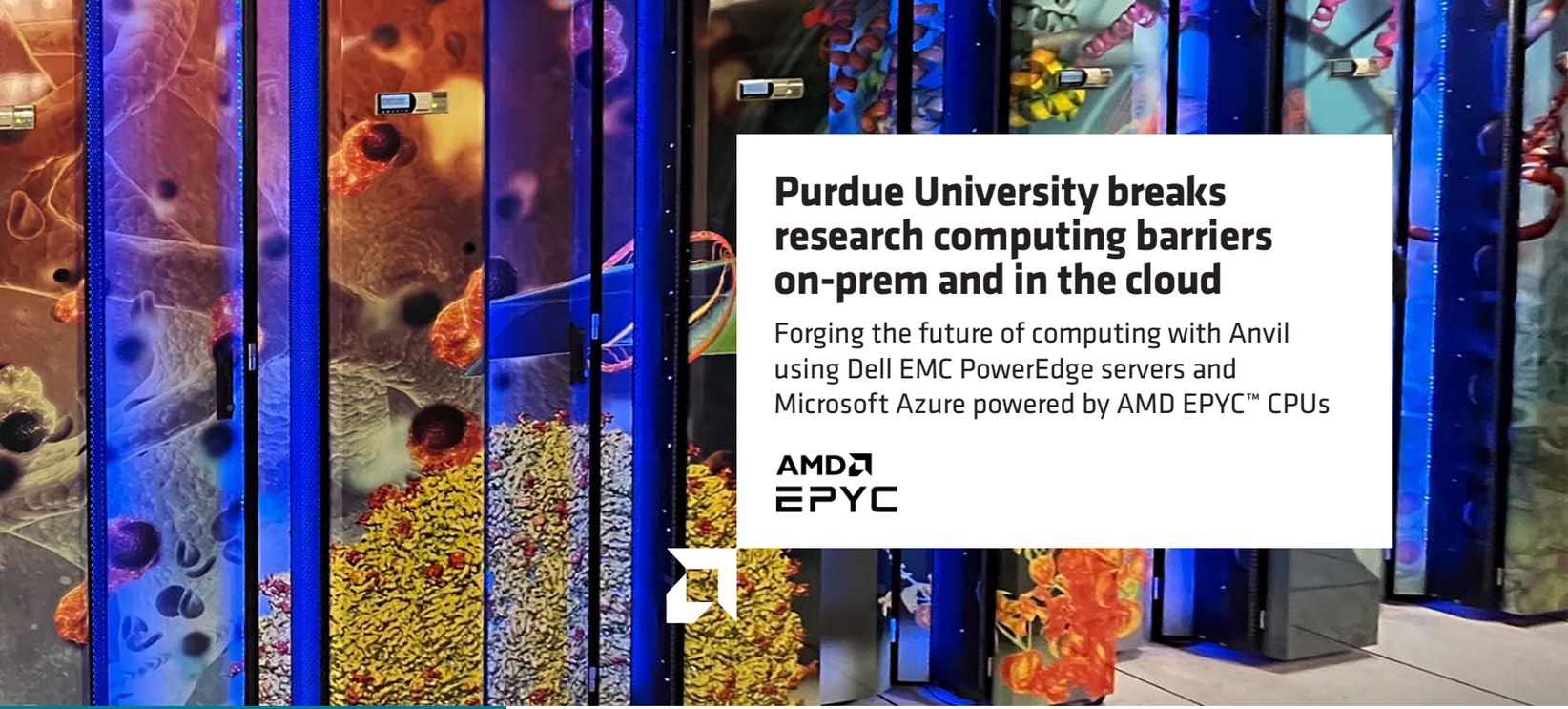

Purdue University breaks research computing barriers on-prem and in the cloud

Forging the future of computing with Anvil using Dell EMC PowerEdge servers and Microsoft Azure powered by AMD EPYC™ CPUs

AMD EPYC

CUSTOMER

PURDUE UNIVERSITY

INDUSTRY

Education

CHALLENGES

Deliver the most possible compute resources to the national US research community

SOLUTION

Dell EMC PowerEdge C6525 servers and Microsoft Azure powered by 3rd Gen AMD EPYC™ processors

RESULTS

1 billion CPU core hours per year for national research projects, giving a significant boost to XSEDE's computing capacity

AMD TECHNOLOGY AT A GLANCE

2nd Gen AMD EPYC™ 7662 processors with 64 cores

3rd Gen AMD EPYC™ 7763 processors with 64 cores

TECHNOLOGY PARTNERS

DELL Technologies

Microsoft Azure

Science is increasingly driven by computation. Research domains are proliferating. With an increasing number of scientists wanting to harness the potential of computer-aided discovery, greater levels of processing power need to be shared more widely. Purdue University wanted to create a system with forward-looking mixed (cloud) computing and composable capabilities that could deliver the maximum amount of research computing capacity to the US academic community. With the help of Dell EMC PowerEdge servers powered by AMD EPYC™ processors, Purdue won a $10 million National Science Foundation (NSF) grant to provide the Extreme Science and Engineering Discovery Environment (XSEDE) with the Anvil supercomputer.

"The amount of computing that we were able to get for Bell on the [2nd Generation AMD EPYC™ processor] Rome-based system was just staggering compared to our previous systems. We got 45 percent more FLOPS per dollar."

Preston Smith, Executive Director, Research Computing at Purdue University

The ever-increasing demand for research computing

"There was a call for grant proposals from the NSF in 2019 for its Advanced Computing Systems and Services (ACSS) program," says Carol Song, Senior Research Scientist and Scientific Solutions Director at Purdue University, and the principal investigator of Anvil. "This is different from the traditional NSF research programs. It's about putting together and operating large systems to provide services to research users."

Over the last few years, research computing has expanded into domains beyond those traditionally associated with computation. "Historically, at Purdue, it would have been engineering and the college of science using these resources," says James Johnson, Operations Manager, Research Computing at Purdue University. "Fast forward to today, and every single school or college on campus is utilizing high performance computational power in some way."

"We had a good feel for what the general trend in research is, so we put in a proposal for Anvil to address three main challenges," continues Song. "One is the demand for computing cycles increasing every year. The NSF-funded XSEDE program systems are all oversubscribed all the time. The second, is the diversification of applications with more domains joining the HPC world, further increasing demand for computation and data-driven analytics. The third challenge is training the emerging workforce, both from the perspective of using HPC resources and running them."

However, the Purdue University Anvil system wasn't just going to be for internal use. The NSF grant requires that systems be shared externally as well. "Any proposals funded by the ACSS program are for the national community," says Song. "The system is allocated through XSEDE, the coordination body of a set of resources. These resources are in various institutions, with Purdue being one of the providers."

"One of the things that really made the Dell EMC PowerEdge C6525 platform with AMD EPYC processors so compelling is that it enables four 128-core nodes in 2U of rack space, which is extremely dense."

Preston Smith, Executive Director, Research Computing at Purdue University

A system aimed at providing compute power to the national academic community must deliver the maximum possible resources for the funding available. Purdue's previous experience with AMD EPYC processors on its "Bell" cluster made this a clear platform contender.

Researchers had already found Bell extremely easy to run their workloads on, with excellent performance. "The high memory bandwidth let the life scientists' bioinformatics applications solve bigger problems than they were able to do on previous systems," says Preston Smith, Executive Director, Research Computing at Purdue University. "The amount of computing that we were able to get for Bell on the [2nd Generation AMD EPYC™ processor] 'Rome'-based system was just staggering compared to our previous systems. We got 45 percent more FLOPS per dollar."

Faster performance, unparalleled compute density in the cloud and on-prem

Purdue performed benchmarks to see how much better the PowerEdge servers with 3rd Gen AMD EPYC processors would be than the servers with 2nd Gen AMD EPYC CPUs in the Bell cluster. Using 3rd Gen AMD EPYC virtual machines on Microsoft Azure, the test applications spanned typical research workloads, including GROMACS, LAMMPS, NAMD, OpenFOAM®, WRF®, DB12 and HPCG. "Going from 'Rome' to 'Milan' [2nd to 3rd Gen AMD EPYC CPUs], we were seeing from 9 to 39 percent faster performance [with Azure VMs]," says Smith. The Purdue tests were aided by the fact that Microsoft Azure was already offering 3rd Gen AMD EPYC HPC VMs (HBv3), which could be used for testing. "Migrating those software applications up into Azure was straightforward. We had great support from Microsoft."

"Whether we're building a system for campus or proposing something for the NSF, the value delivered by Dell Technologies, Microsoft, and AMD make it a no brainer and provide a compelling platform for the Anvil proposal."

Carol Song, Senior Research Scientist and Scientific Solutions Director at Purdue University, Principal Investigator of Anvil

These results implied that Dell EMC PowerEdge servers with 3rd Gen AMD EPYC processors would deliver Purdue the best performance for the money for the on-site portion of the deployment. "Whether it's federal money from the NSF or Purdue money, which is coming from the State of Indiana, it has to be proven that a proposal is the best possible one for the public good," says Smith. "One of the things that really made the PowerEdge C6525 platform with AMD EPYC processors so compelling is that it enables four 128-core nodes in 2U of rack space, which is extremely dense. We are still using the same facility that's been hosting high-performance computing at Purdue since the 1970s. Being able to fit modern density computing in that space is paramount. Having lots of these 4 by 128 core 2U units gives us a great deal of compute capacity and performance in a comparable amount of space and wattage."

"Going from Rome [2nd Gen AMD EPYC processors] to Milan [3rd Gen AMD EPYC processors], we were seeing from 9 to 39 percent faster performance."

Preston Smith, Executive Director, Research Computing at Purdue University.

Unparalleled science performance as a service

The Purdue proposal received $10 million in acquisition funding, alongside a $12.5 million grant for five years of operation and user support. "Anvil is a large CPU cluster with a thousand nodes of 3rd Gen AMD EPYC processors," says Song. "We have a sub portion with GPUs, and large 1TB memory nodes. Altogether, we are offering XSEDE one billion CPU core hours per year, giving XSEDE a boost with a lot of computing cycles. Our philosophy in the proposal was that we want to support lots of users and lots of moderate-sized jobs. Instead of just hero runs using huge resources, we are also focusing on the bulk of the computation that's happening on national systems today and in the future."

"We designed Anvil to fit a broad cross-section of research work," says Smith. "When we proposed the system, we found that a significant percent of both the jobs run and hours delivered on XSEDE, can fit on one Anvil node. With that goal in mind, having a thousand 128-core nodes allows us to fit lots of representative work from XSEDE onto Anvil." Song adds: "We see Anvil as a system forging the future of computing. We expect tens of thousands of researchers around the nation to be able to access it. We call it science as a service." With an eye on future workflows, Anvil's Dell PowerEdge compute nodes will be supplemented with burst capacity in Microsoft Azure, as well as leverage Azure's machine learning tools. Microsoft Azure will serve as a path for future scalability and reproducibility for applications developed within Anvil's Kubernetes®-based composable subsystem built within the cluster on Dell PowerEdge."

"High performance computing is going to be so critical to science, industry and lots of other sectors in the future," concludes Song. "We've got great plans for how we're going to use Anvil as a tool for training the next generation of the HPC workforce. Whether we're building a system for campus or proposing something for the NSF, the value delivered by Dell Technologies, Microsoft, and AMD make it a no brainer and provide a compelling platform for the Anvil proposal."

About Purdue

Purdue University was founded in 1869 under the Morrill Act of 1862. Nowadays, Purdue University has an enrollment of over 41,000 undergraduate and graduate students. Approximately 2,000 full-time faculty teach and conduct research in Purdue's 157 principal buildings with 377 total buildings on 2,307 acres. Purdue offers 6,700 courses in 200 specializations. For more than a quarter century, Purdue University has been the largest or second largest in undergraduate engineering enrollment in the country. Purdue research exceeds $347.1 million a year using more than 400 research labs. For more information visit purdue.edu.

About Dell Technologies

Dell Technologies (NYSE:DELL) helps organizations and individuals build their digital future and transform how they work, live, and play. The company provides customers with the industry's broadest and most innovative technology and services portfolio for the data era.

About Microsoft Azure

The Azure cloud platform is more than 200 products and cloud services designed to help you bring new solutions to life—to solve today's challenges and create the future. Build, run, and manage applications across multiple clouds, on-premises, and at the edge, with the tools and frameworks of your choice. For more information visit azure.microsoft.com.

About AMD

For over 50 years AMD has driven innovation in high-performance computing, graphics, and visualization technologies—the building blocks for gaming, immersive platforms, and the data center. Hundreds of millions of consumers, leading Fortune 500 businesses, and cutting-edge scientific research facilities around the world rely on AMD technology daily to improve how they live, work, and play. AMD employees around the world are focused on building great products that push the boundaries of what is possible. For more information about how AMD is enabling today and inspiring tomorrow, visit amd.com/EPYC.