

San Diego Supercomputer Center Welcomes ‘Gordon’ Supercomputer as a Research Powerhouse

Unique Flash-based System to Help Solve Large Data-Intensive Science Challenges



SDSC's Gordon Supercomputer. Photo by Erik Jepsen/UC San Diego Publications

When it officially comes online in early January, *Gordon*, a unique new supercomputer at the San Diego Supercomputer Center (SDSC), will help researchers tackle the most vexing data-intensive challenges, from mapping genomes for personalized medicine to rapidly calculating thousands of “what-if” scenarios affecting everything from traffic patterns to climate change.

Gordon by the Numbers

“*Gordon* is an extremely important resource because it is dedicated to solving critical science and societal problems currently overwhelmed by the vast amount of data generated by the digital devices of our era,” said UC San Diego Chancellor Marye Anne Fox at the formal launch of SDSC’s newest supercomputer, held Dec. 5 at SDSC. “Just like SDSC, *Gordon* will be a central point of collaboration. Today’s global challenges are increasingly interdisciplinary in nature, and require an interdisciplinary approach to solve.”

San Diego Mayor Jerry Sanders followed the chancellor’s remarks, quipping that “We expect this kind of thing out of UCSD once a week,” and called the launch of *Gordon* “yet another example” of why the university is one of the nation’s top research schools.

“The era of data-intensive supercomputing begins with *Gordon*,” declared SDSC Director Michael Norman, who along with SDSC Associate Director Allan Snively convinced the National Science Foundation (NSF) about three years ago to invest \$20 million in a completely new kind of supercomputer.

“Every year, we double the amount of information being generated, and we now are being overwhelmed by the data we are able to produce with our own computers,” said Norman. “So it stands to reason that we needed a new kind of computer.”

Popcorn, anyone?

So what’s unique about *Gordon*? It is the first supercomputer to employ massive amounts of flash-based memory – which is common in smaller devices such as laptops or cell phones – to help speed solutions now hamstrung by slower spinning disk memory. Think of *Gordon* as the world’s largest thumb drive, but with the capability to ingest about 220 movies per second from

Gordon is capable of handling massive data bases while providing up to 100 times faster speeds when compared to hard drive disk systems for some queries. Here’s a look at *Gordon*’s calculating “horsepower.”

- **4 petabytes of disk storage.** One petabyte equals a quadrillion (1,000 trillion) bytes of information. It would take 1,900 years to listen to a petabyte’s worth of songs – if you had a large enough MP3 player!

- **64 terabytes of random access memory.** Just one terabyte equals all the information printed on paper made from 50,000 trees. *Gordon*’s 64 terabytes of RAM is more than 16,000 times the memory of a standard MacBook Pro.

- **280+ teraflops of compute power.** Supercomputers are typically measured in “FLOP/s” or Floating point Operations per Second, or the ability to do mathematic calculations at peak speed. A teraflop is one trillion floating point operations. A supercomputer capable of operating at one TFlop/s is about 150 times faster than having every person in the world performing one operation per second on a hand calculator.

- **300 terabytes of flash memory.** *Gordon*’s 300 terabytes of flash memory is 4800 times greater than the storage capacity of a top-of-the-line 64GB iPad 2.

- **100,000 movies:** That’s the entire Netflix movie catalog, which *Gordon* could ingest at the rate of 220 movies per second – and still have room for 200,000 more movies!

- **20,000,000 dollars.** *Gordon* is the result of a \$20 million, five-year grant from the National Science Foundation to develop a new kind of data-intensive supercomputer.

- **36,000,000 IOPS.** IOPS means input/output operations per second, an important measure for data-intensive computing since it indicates the ability of a storage system to perform I/O operations on randomly organized data – critical

Netflix, or consume the entire catalog of about 100,000 Netflix movies – while still having room for another 200,000 titles. That’s a lot of popcorn.

On a more scientific note, *Gordon* will have the ability to hold 100,000 entire human genomes in its flash memory system. “I don’t think there are 100,000 human genomes yet, which gives you a sense how deep we are going with this new technology,” said Norman.

Added Nicholas Schork, professor at The Scripps

Research Institute and the Scripps Translational Science Institute, and an expert in genomic data who plans to use *Gordon* as a resource: “We need computational advances such as *Gordon*, and I am glad to see the confluence of genomics and computational science.”

While *Gordon* will serve UC San Diego and UC researchers as well as be available for use by industry and government agencies, it will also serve as a national computing resource as part of the NSF’s new XSEDE (Extreme Science and Engineering Discovery Environment) program, a nationwide partnership comprising 16 supercomputers and high-end visualization and data analysis resources. The five-year, \$121-million NSF project expands on the NSF’s TeraGrid project, in which more than 10,000 scientists used the TeraGrid for thousands of research projects, at no cost to the researchers.

Superfast... and capable

Along with its innovative “supernodes” that create shared memory systems, *Gordon* will have the capability to process large data-intensive problems about 10 times faster than other supercomputers. In fact, at launch *Gordon* will be the most powerful supercomputer in all of Southern California in terms of speed and capability, and was recently ranked as among the top 50 fastest supercomputers in the world in terms of speed or doing pure math (see top500.org).

More importantly, *Gordon* is now at the top in terms of how many I/O (input/output) operations it can do per second, which is a critical measure for doing data-intensive computing, such as sifting through huge datasets just to find a sliver of meaningful information. In recent validation tests, *Gordon* achieved an unprecedented 36 million IOPS (input/output operations per second), making it the most powerful supercomputer ever commissioned by the NSF for doing I/O and breaking the previous (2010) record of only 4.2 million IOPS.

The XSEDE resource allocations committee, made up of computational science experts from across the country, has just reviewed the first allocation requests for using *Gordon*. “I view *Gordon* as a new kind of vessel, a ship that will take us on new voyages to makes new discoveries in new areas of science,” said Norman.

for database and data mining applications. *Gordon* now holds the record for IOPS capability, outpacing the previous record holder by more than eight times.

Now you know: SDSC’s newest supercomputer is named after the 1930s science fiction comic books and 1950s television series called “Flash Gordon.”



Mayor Jerry Sanders, UCSD Chancellor Marye Anne Fox, Director Michael Norman at the Gordon Launch event, photo: Alan Decker.

Those voyages mean *Gordon* can serve as a compute resource to help map the universe with the most detailed observations ever made, compare thousands of genomes to help find cures to diseases – or even detect predispositions to certain diseases – or be used to produce ultra-detailed simulations of earthquakes to help mitigate building damage, aid rescue efforts, and reduce potential injuries and loss of life. Other possible applications include helping sociologists detect potentially significant changes and patterns in population, as well as helping economists spot how such things as computerized trading affect world financial markets.

“*Gordon’s* extraordinary speed makes it possible for researchers to tackle questions they couldn't address before, simply because they didn't have enough computing power,” added Norman.

Keep up with campus news by subscribing to *This Week @ UC San Diego*