

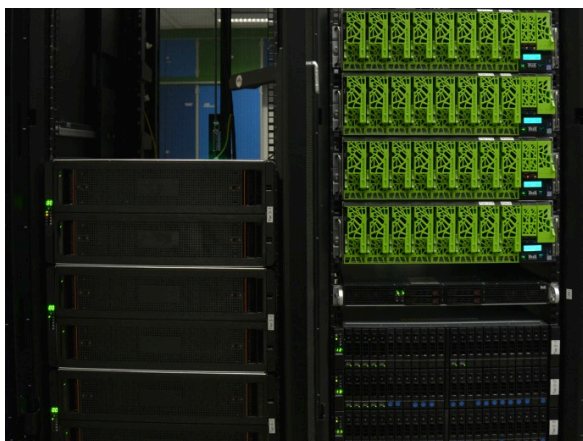


# SAGE

## Results in Brief

### New data storage system takes the bite out of exascale supercomputing

Exascale supercomputers process up to 1 000 terabytes (TB) of data per day, although just moving it to the processor takes hours. An EU innovation removes the bottleneck.



© Dr Sai Narasimhamurthy

Supercomputers are machines with many (these days, thousands) processors that work in parallel to achieve calculation rates far beyond normal computers. The latest generation is known as exascale supercomputers.


Defined as those exceeding one billion billion calculations per second, the newest machines represent a thousandfold speed increase compared to the best of only a decade ago.

Such equipment is used in research fields demanding the ultimate computational power, for example: weather/climate studies, genomics and human brain simulations.

Current data management technologies already struggle with the demands of supercomputers. For instance, a conventional high-performance supercomputer might run a simulation on 8 000 plus processors that produces 25 TB of data every day. Processing the raw data multiplies the amount by two or three. Some applications already have to read hundreds of terabytes. Now, with exascale supercomputers, applications producing petabytes (1 000 TB) will be commonplace.

Computers generally store data in one place and move them to another for analysis or processing. Currently, even with the best available networks, moving terabytes or petabytes of data can take many hours. This represents a serious bottleneck. In addition, the data movement requires tremendous amounts of power, in the range of hundreds of megawatts.

## Removing the bottleneck

The EU-funded [SAGE](#)  project developed a new data storage system able to meet the demands of exascale computing. The innovation minimises the need for data transport.

“Instead of moving data, our system takes the computation to the storage system,” explains project leader Dr Sai Narasimhamurthy. Data can be processed in, or close to, the storage location. Supercomputing applications can drop in analytical modules as needed.

SAGE’s ‘intelligent storage’ system additionally involves optimising data storage. Data may be stored in any of several tiers, including: conventional hard discs, solid state discs and non-volatile memory. Each has certain performance properties. The SAGE system moves data to the tier with the appropriate performance characteristics at the right time. This improves performance.

The combination of both aspects means flexibility and versatility. Applications having various complex data formats can use different types of data management tools. This yields a powerful and extendable application programming interface, which the SAGE team also developed.

## Prototype demonstration

“Our prototype was ‘very small’,” adds Dr Narasimhamurthy, “able to handle less than half a petabyte of data. Also, our software is not yet optimised.” Therefore, it is unrealistic to compare prototype performance against large production clusters. Instead, the team’s main focus was proving that the methods and techniques work; they do, and can easily upscale to larger storage hardware. Reception from the scientific community has been very positive.

Following the successful demonstration, the project will continue as Sage2. The new project will extend the SAGE prototype and explore new ways of using distributed non-volatile memory storage. It will also examine artificial intelligence and ‘deep learning’ applications of exascale supercomputers.

The SAGE system will remove or greatly reduce bottlenecks affecting exascale

supercomputers, allowing the machines to operate closer to full speed. Also, power consumption is ultimately expected to be around 100 times less than that used in current systems.

The worldwide market for high-performance computing storage is expected to reach USD 6 billion in 2021. The market for Big Data analytics and cloud storage will be even bigger. The SAGE projects will be targeting the European components of both markets.

## Keywords

[SAGE](#)

[supercomputers](#)

[exascale](#)

[bottleneck](#)

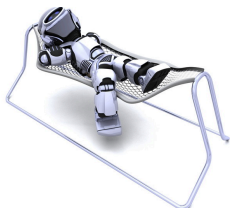
[petabyte](#)

[data storage](#)

[data transport](#)

[intelligent storage](#)

## Discover other articles in the same domain of application



[Dream-like processes could help build more human-centric robots](#)

30 September 2019



[Robot assistants help the elderly stay independent](#)

24 August 2020





## New robots set to transform farming

22 May 2018



## Buildings get smarter with data-driven architecture

21 February 2024



### Project Information

#### SAGE

Grant agreement ID: 671500

[Project website](#) 

#### DOI

[10.3030/671500](https://doi.org/10.3030/671500) 

Project closed

#### EC signature date

29 June 2015

#### Start date

1 September 2015

#### End date

31 August 2018

#### Funded under

EXCELLENT SCIENCE - Future and Emerging Technologies (FET)

#### Total cost

€ 7 882 531,25

#### EU contribution

€ 7 882 531,00

#### Coordinated by

SEAGATE SYSTEMS UK LIMITED

 United Kingdom

**Last update:** 18 February 2019

**Permalink:** <https://cordis.europa.eu/article/id/248967-new-data-storage-system-takes-the-bite-out-of-exascale-supercomputing>

European Union, 2025

