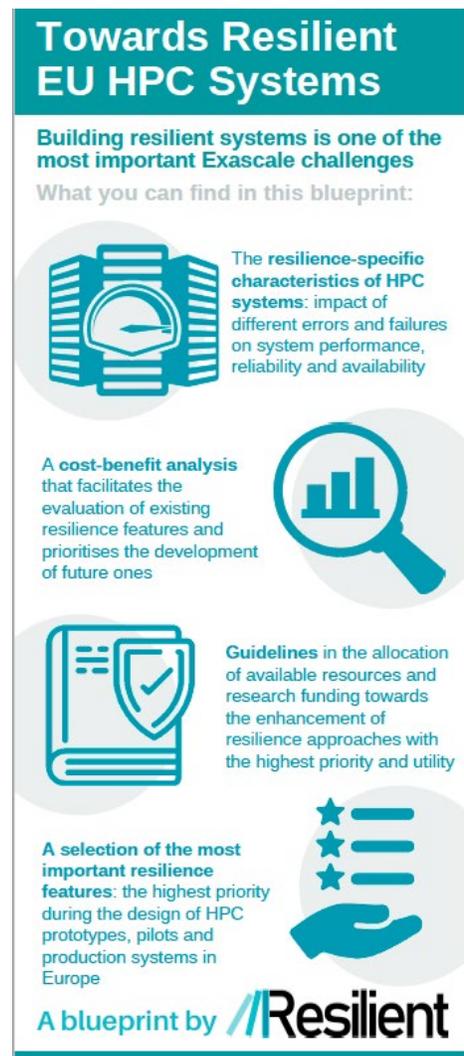# Towards Resilient European HPC systems

HPC system resilience is one of the most important Exascale requirements and challenges. Addressing this challenge requires holistic full-stack solutions including hardware, system software and application enhancements. A selected group of European researchers and HPC companies lead the European HPC resilience initiative, whose aim is foster information exchange and collaboration across the diverse communities affected by resilience concerns. One of the first steps forward is the launch of the document titled Towards Resilient EU HPC Systems: A Blueprint, developed thanks to the collaboration of European research centers, US national laboratories, HPC-related companies and various European HPC projects. The objective of this publication is to spearhead a Europe-wide discussion on HPC system resilience and to help the European HPC community define best practices.

Resilience is widely recognized as a critical challenge for high performance computing (HPC) systems, as a result of the increasing complexity, both at the level of individual hardware and software components and also in terms of subsystems and complete heterogeneous system configurations. Resilience for HPC systems encompasses a wide spectrum of fundamental and applied research and development, including theoretical foundations, fault detection and prediction, monitoring and control, end-to-end data integrity, enabling infrastructure, and resilient computational algorithms. Moreover, facility operations and cost management concerns need also to be weighed in, in the context of a systematic risk management framework.

The document Towards Resilient EU HPC Systems: A Blueprint analyses a wide range of state-of-the-art resilience mechanisms and recommends the most effective approaches to employ in large-scale HPC systems. These guidelines are useful in the allocation of available resources, as well as guiding researchers and research funding initiatives towards the enhancement of resilience approaches with the highest priority and utility. Although this work is focused on the needs of next generation HPC prototypes, pilots and production systems in Europe, the principles and evaluations are also applicable globally. The current version of the blueprint covers individual HPC nodes, encompassing CPU, memory, intra-

node interconnects and emerging FPGA-based hardware accelerators. With community support and feedback, the new document will be extended to cover GPUs, vector accelerators, interconnect networks and storage.

The need for resilience features has been analyzed based on three principles:

- The resilience features implemented in HPC systems should ensure that the failure rate of the system is below an acceptable threshold, representative of the technology, system size and target application.
- Given the high cost incurred by uncorrected error propagation, if hardware errors occur frequently they should be detected and corrected at low overhead, which is likely only possible in hardware.
- Overheating is one of the main causes of unreliable device behavior. Production HPC systems should prevent overheating while balancing power/energy and performance.

The blueprint is already used to define and evaluate resilience features of the Testbed systems that will be the outcome of the [EuroEXA project.](#) Other European projects involved in the European HPC resilience initiative will follow this practice.

Download Blueprint (PDF)

Video: https://www.youtube.com/watch?v=wNFTlVMneuw

About the European HPC resilience initiative

The recently launched European HPC resilience initiative plans to spearhead a Europe-wide discussion on HPC system resilience. It brings together experts from academia and industry covering the broad spectrum of computing systems technologies to further research and deploy HPC resilience. The main objective of this initiative is to foster information exchange and collaboration across the diverse communities affected by resilience concerns. More information: https://resilienthpc.eu/