Expand: High Performance Storage System for HPC and Big Data Environments" (TED2021-131798B-I00)

High Performance Storage Sytems for HPC and Big Data (Expand)

D 4.2
Final report of the project and conclusions

Universidad Carlos III de Madrid

June, 2025

CONTENTS

1. MAIN RESULTS 1
1.1. Project website and deliverables 1
1.2. Publications 1
1.2.1. Publications in journals 1
1.2.2. International conferences 2
1.2.3. National conferences 2
1.3. Theses under development. 3
1.4. Results obtained in the IO500 benchmark 3

1. MAIN RESULTS

This report presents the main results obtained with the development of the project:

- Project website and deliverables.
- Publications.
- Theses under development.
- Results obtained in the IO500 benchmark.

1.1. Project website and deliverables

The website with the project information is available at this address:

`https://expand-arcos.github.io`

The Expand source code is available at this address:

`https://github.com/xpn-arcos/xpn`

1.2. Publications

1.2.1. Publications in journals

- Dario Muñoz-Muñoz, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, Jesus Carretero, "Malleability and fault tolerance in ad-hoc parallel file systems", *Cluster Computing* [Q1, Computer science, theory & methods], June 2025
  doi: 10.1007/s10586-025-05575-8.
- Elias Del-Pozo-Puñal, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, "Hierarchical and distributed data storage for Computing Continuum", *Future Generation Computer Systems* [Q1, Computer science, theory & methods], vol. 174, June 2025
  doi: 10.1016/j.future.2025.107931.

Submitted under review

- Diego Camarmas-Alonso, Felix Garcia-Carballeira, Alejandro Calderon-Mateos, Jesus Carretero, "Improving I/O performance in HPC environments using the Expand Ad-Hoc file system", International Journal of High Performance Computing Applications [Q3, Computer science, hardware & architecture]. Under second review.

1.2.2. International conferences

- Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, Jesus Carretero, "A new Ad-Hoc parallel file system for HPC environments based on the Expand parallel file system", *22nd International Symposium on Parallel and Distributed Computing (ISPDC)* [CORE C], pp. 69-76, July 10-12 2023, Bucharest, Romania
  doi: 10.1109/ISPDC59212.2023.00015.

- Dario Muñoz-Muñoz, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, Jesus Carretero, "Fault Tolerant in the Expand Ad-Hoc Parallel File System", *Euro-Par 2024 (CORE B). August 26-30 2024, Madrid, Spain*
  doi: 10.1007/978-3-031-90200-0_26.

- Dario Muñoz-Muñoz, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, Jesus Carretero, "Malleability in the Expand Ad-Hoc parallel file system", *3rd EuroHPC Workshop on Dynamic Resources in HPC. Euro-Par 2024* [CORE B] August 26-30 2024, Madrid, Spain
  doi: 10.1007/978-3-031-90200-0_26.

1.2.3. National conferences

- Diego Camarmas-Alonso, Felix Garcia-Carballeira, Alejandro Calderon-Mateos, Darío Muñoz-Muñoz, Jesus Carretero, "Evaluación del sistema de ficheros Expand Ad-Hoc con aplicaciones de uso intensivo de datos", *XXXV Jornadas de Paralelismo (JP25)*, June 25-27 2025, Sevilla.

- Elías Del-Pozo-Puñal, Félix García-Carballeira, {Diego Camarmas-Alonso, Alejandro Calderon-Mateos, "Sistema de almacenamiento para computing continuum: aplicación a sistemas de información ferroviaria", *XXXV Jornadas de Paralelismo (JP25)*, June 25-27 2025, Sevilla.

- Gabriel Sotodosos-Morales, Félix García-Carballeira, Diego Camarmas-Alonso, Alejandro Calderón-Mateos, Darío Muñoz-Muñoz, Jesús Carretero, "Optimización de entornos de Big Data Analytics mediante sistemas de ficheros paralelos Ad-hoc", *XXXV Jornadas de Paralelismo (JP25)*, June 25-27 2024, Sevilla.

- Dario Muñoz-Muñoz, Diego Camarmas-Alonso, Felix Garcia-Carballeira, Alejandro Calderon-Mateos, Jesus Carretero, "Tolerancia a fallos en el sistema de ficheros paralelo Expand Ad-Hoc", *XXXIV Jornadas de Paralelismo (JP24)*, pp. 271-279, June 17-19 2024, A Coruña
  doi: 10.5281/zenodo.12743582.

- Elias Del-Pozo-Puñal, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, "Evaluación del rendimiento de un sistema de ficheros para sistemas IoT", *XXXIV Jornadas de Paralelismo (JP24)*, pp. 289-298, June 17-19 2024, A Coruña
  doi: 10.5281/zenodo.12094741.

- Diego Camarmas-Alonso, Felix Garcia-Carballeira, Alejandro Calderon-Mateos, Jesus Carretero, "Evaluación de rendimiento del sistema de ficheros paralelo Expand Ad-Hoc en MareNostrum 4", *XXXIII Jornadas de Paralelismo (JP23)*, pp. 397-404, September 20-22 2023, Ciudad Real
  doi: 10.5281/zenodo.8378956.

- Elias Del-Pozo-Puñal, Felix Garcia-Carballeira, Diego Camarmas-Alonso, Alejandro Calderon-Mateos, "Sistema de Ficheros Distribuido para IoT basado en Expand", *XXXIII Jornadas de Paralelismo (JP23)*, pp. 559-567, September 20-22 2023, Ciudad Real
  doi: 10.5281/zenodo.10706248.

1.3. Theses under development

The results obtained in the project have given rise to the following doctoral theses:

- Ph.D.: Elías del Pozo
  Advisors: Félix García Carballeira and Alejandro Calderón Mateos
  Title: Data management techniques in High-End and Edge Computing systems
  Universidad Carlos III de Madrid
  Expected Reading Date: October 2025
- Ph.D.: Diego Camarmas
  Advisors: Félix García Carballeira y Jesús Carretero Pérez
  Title: Gestión dinámica de datos en sistemas distribuidos de gran escala.
  Universidad Carlos III de Madrid
  Expected Reading Date: October 2025
- Ph.D.: Darío Muñoz Muñoz
  Advisor: Félix García Carballeira
  Title: New techniques to improve performance and usability of Ad-hoc file systems in HPC
  Universidad Carlos III de Madrid
  Expected Reading Date: 2027

1.4. Results obtained in the IO500 benchmark

Expand has been included in the IO500 list (**https://io500.org**), where positions are as follows:

- MareNostrum 4 supercomputer evaluation (SC23):
  - 69th position out of 101 in the 10 Node Research list.
  - 184th position out of 236 in the Full list.
- Leonardo supercomputer evaluation (SC24):
  - 47th position out of 113 in the 10 Node Research list.
  - 114th position out of 268 in the Full list.
- C3-UC3M supercomputer evaluation (ISC25):
  - 78th position out of 118 in the 10 Node Research list.
  - 205th position out of 284 in the Full list.