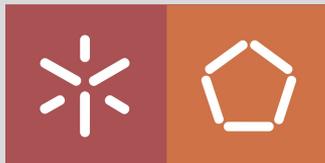




# Computação Avançada

2024/2025



**Universidade do Minho**

Escola de Engenharia

Departamento de Informática

# Para que serve?

# Computação avançada (ou, de elevado desempenho)

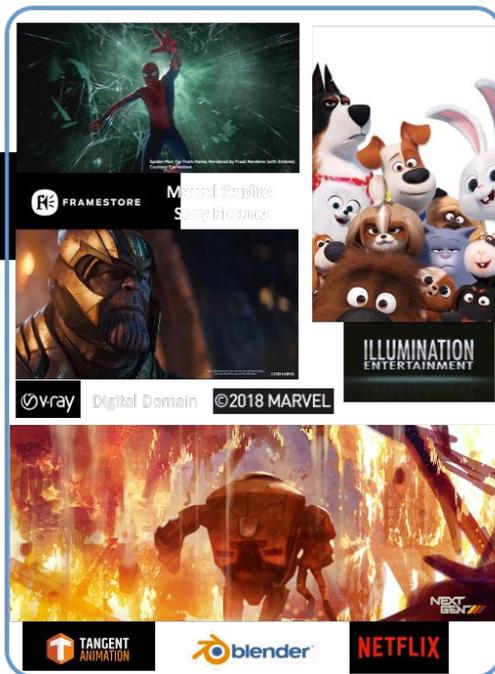
Quais as razões para estudar e utilizar sistemas de computação avançada?

Algumas razões apontadas pela UE, centros HPC (p.ex., TACC, AIST, MACC) e indústria (p.ex., Nvidia, Intel, Google, Amazon)

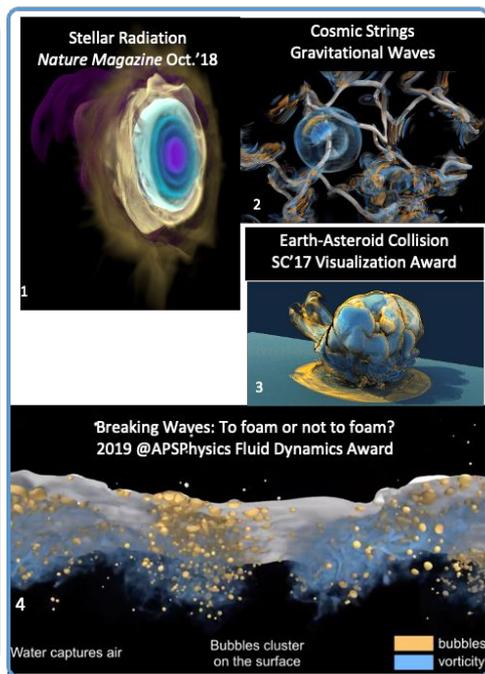
- **Velocidade e Eficiência:** permitem que engenheiros e cientistas abordem problemas complexos e de grandes dimensões
- **Descoberta Científica:** são fundamentais para várias disciplinas da ciência (física, biologia, química), e para descobrir novos tratamentos e medicamentos para a saúde
- **Desenho e otimização de produtos:** amplamente utilizados pela indústria (p.ex., aeroespacial, automóvel, energia, cibersegurança) para otimizar produtos
- **Análise de dados e Inteligência Artificial (IA):** têm servido de suporte fundamental para a análise e construção (treino) de modelos de IA

# Mais alguns exemplos de aplicação

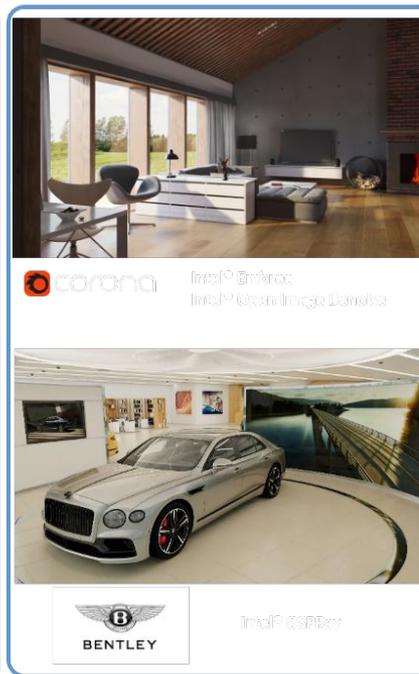
## Indústria Cinematográfica



## Ciência



## Produtos 3D e Arquitectura



## Jogos e Realidade Virtual



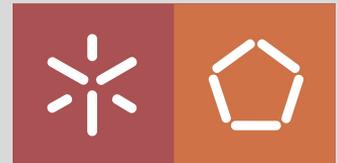
# Desafios

- Os requisitos de armazenamento e processamento de dados têm **crescido significativamente** para dar suporte à nova geração de aplicações
- O que tem levado a infraestruturas de computação avançada cada vez maiores e mais complexas
  - **Deucalion: 1 632** servidores com arquitetura ARM (e mais uma centena de servidores)
  - **Vista: 600** servidores com NVIDIA GH200 Grace Hopper Superchip (35 000\$ por chip)
  - **ABCI 3.0:** novo sistema com **6 128 NVIDIA H200** (~30 000\$ por GPU)
  - **Fugaku: 158 972** servidores ARM

# Visão

- Infelizmente (ou **felizmente**), não basta aumentar e melhorar a infraestrutura (**hardware**) para ter aplicações mais eficientes
  - É preciso saber **implementar/otimizar aplicações e sistemas** que tirem partido do hardware e da grande quantidade de recursos computacionais e de armazenamento disponíveis!
- Este é um perfil de **mãos na massa**, onde vamos estudar, implementar, otimizar e experimentar sistemas e tecnologias usadas na prática (p.ex., Apache Spark, PyTorch, FUSE)
  - Colaborando com investigadores de centros HPC (MACC, TACC, AIST)
  - Adquirindo conceitos transversais e aplicáveis na indústria e em outros ambientes (p.ex., computação em nuvem)

# O que engloba?



**Universidade do Minho**  
Escola de Engenharia  
Departamento de Informática

# Computação Heterogénea de Elevado Desempenho

Confere conhecimentos sobre as características e organização de **servidores heterogéneos** (CPUs + aceleradores) para **programar eficientemente** estes dispositivos.

**Foco na otimização de código para arquiteturas HPC e análise de desempenho**

- Arquiteturas de servidores heterogéneos e supercomputadores
  - Topologias e tecnologias de interligação multi-servidor em escala
- Paralelização híbrida em servidores heterogéneos
  - Paralelismo multi-servidor escalável
  - Escalonamento multi-nível de cargas de trabalho
- Estruturas de dados e técnicas de decomposição de domínios
- Avaliação de desempenho de aplicações de larga escala
  - Utilização de *profilers*

# Análise de Dados de Grandes Dimensões

Pretende dar a conhecer as etapas e desafios da análise de grandes volumes de dados para **desenvolver aplicações centradas em dados mais eficientes.**

**Foco no desenvolvimento de aplicações de Operational Data Analytics (ODA)**

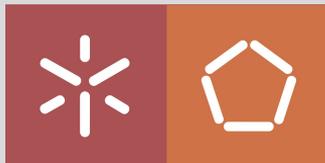
- Pipelines de análise de dados estruturados e semi-estruturados (single e multi-node)
- Modelos de distribuição em processamento distribuído
  - Paralelismo dos dados
  - Paralelismo dos modelos
  - Paralelismo da pipeline
- Processamento em batch, streaming e baseado em modelos de ML/DL (Apache Spark, PyTorch)
- Avaliação, monitorização e otimização de pipelines de processamento distribuído
  - Escalabilidade
  - Coerência
  - Distribuição
  - Expressividade dos dados

# Sistemas de Armazenamento de Dados

Pretende conferir a capacidade de **desenvolver e otimizar soluções para o armazenamento e acesso a dados** com garantias de alto desempenho, escalabilidade e confiabilidade.

**Foco em sistemas de armazenamento e na otimização de aplicações centradas em dados**

- Interfaces, dispositivos e tecnologias de armazenamento emergentes
  - Bloco, ficheiro, chave-valor
  - NVMe SSD, memória persistente, CXL, SPDK
- Desenho e arquitecturas de soluções modernas de armazenamento de dados
  - Locais (ext4, RocksDB)
  - Remotas (iSCSI, NFS)
  - Distribuídas (CEPH, Lustre)
- Otimizações de armazenamento locais e distribuídas
  - Desempenho, escalabilidade, confiabilidade
- Avaliação e monitorização de sistemas de armazenamento e aplicações



**Universidade do Minho**

Escola de Engenharia

Departamento de Informática

# O que dizem?

# Testemunhos



As the Director of HPC at TACC, there are a few **qualities I look for in new hires** to support the research of our users. Primarily, we look for applicants with **parallel programming** experience for **multi-node** and **multi-threaded** execution using MPI and OpenMP to support scientific applications. Also, with the **increasing importance of Machine Learning**, we're interviewing candidates with experience managing **large amounts of data on parallel file systems**. These skills are invaluable for our mission.

**John Cazes**  
*Director of HPC*  
TACC

**AIST** is a national research institute in Japan that **performs research to help advance Japanese industry**, especially in the area of **Artificial Intelligence (AI)**. However, it is clear that AI and associated technologies are increasing their global reach and impact. Therefore, it is critical that students develop skill sets in the area of **distributed parallel file systems, storage I/O, and data management**, which will better prepare them for entry into the global workforce.

**Jason Haga**  
*Chief Senior Research Scientist*  
AIST

# Testemunhos



As the Director of HPC at TACC, there are a few **qualities I look for in new hires** to support the research of our users. Primarily, we look for applicants with **parallel programming** experience for **multi-node** and **multi-threaded** execution using MPI and OpenMP to support scientific applications. Also, with the increasing importance of **Machine Learning**, we're interviewing candidates with experience managing **large amounts of data on parallel file systems**. These skills are invaluable for our mission.

**John Cazes**  
*Director of HPC*  
TACC

**AIST** is a national research institute in Japan that **performs research to help advance Japanese industry**, especially in the area of **Artificial Intelligence (AI)**. However, it is clear that AI and associated technologies are increasing their global reach and impact. Therefore, it is critical that students develop skill sets in the area of **distributed parallel file systems, storage I/O, and data management**, which will better prepare them for entry into the global workforce.

**Jason Haga**  
*Chief Senior Research Scientist*  
AIST

**Oportunidade de colaborar/estagiar com centros de computação avançada nacionais (MACC) e internacionais (TACC, AIST, NARlabs, ...)**

# Testemunhos



Feedzai's core scoring systems handle over **60 billion fraud decisions annually**, delivering high-throughput and millisecond-level latency. The Data Reliability team develops **data-intensive architectures** to support the computation of fraud profiles, including **parallel computing, prefetching, sharding, batching** and **event streaming**, where **data locality, availability** and **consistency** are critical requirements.

**João Azevedo**  
*Data Reliability Engineer*  
*Feedzai*

Atualmente, sou responsável por integrar soluções da Elastic na plataforma SaaS da empresa. Conceitos sobre **processamento distribuído de dados, caching e prefetching**, entre outros, são fundamentais para compreender como a nossa **base de dados Elasticsearch** funciona e como a integrar numa plataforma (por exemplo para escolher as melhores configurações e hardware).

**Cláudia Correia**  
*Senior Software Engineer*  
*Elastic*

# Testemunhos



Feedzai's core scoring systems handle over **60 billion fraud decisions annually**, delivering high-throughput and millisecond-level latency. The Data Reliability team develops **data-intensive architectures** to support the computation of fraud profiles, including **parallel computing, prefetching, sharding, batching** and **event streaming**, where **data locality, availability** and **consistency** are critical requirements.

**João Azevedo**  
*Data Reliability Engineer*  
Feedzai

Atualmente, sou responsável por integrar soluções da Elastic na plataforma SaaS da empresa. Conceitos sobre **processamento distribuído de dados, caching e prefetching**, entre outros, são fundamentais para compreender como a nossa **base de dados Elasticsearch** funciona e como a integrar numa plataforma (por exemplo para escolher as melhores configurações e hardware).

**Cláudia Correia**  
*Senior Software Engineer*  
Elastic

**Outros colegas, que trabalharam conosco, estão agora a aplicar estes conceitos em empresas como a RedHat, Critical Techworks, Bosh, Kodly, OLYX ...**

# Testemunhos



O meu trabalho de doutoramento pretende melhorar a **qualidade de serviço dos sistemas de armazenamento em centros HPC**. Para tal, é fundamental compreender o **funcionamento** e a **complexidade** destas **infraestruturas de grande escala**. Nomeadamente, como projetar soluções escaláveis, entender a orquestração dos sistemas de armazenamento e como desenvolver algoritmos eficientes para resolver problemas complexos.

**Mariana Miranda**  
*Aluna de Doutoramento*  
*HASLab*

O meu trabalho de doutoramento pretende desenvolver novas soluções para testar a **resiliência de aplicações que gerem dados críticos**. Para tal são necessários conhecimentos de sistemas operativos, **sistemas de ficheiros** e, sobretudo, no **funcionamento da pilha de armazenamento**. Ainda, é essencial compreender como as **aplicações locais e distribuídas interagem com essa pilha**.

**Maria Ramos**  
*Aluna de Doutoramento*  
*HASLab*

# Testemunhos



O meu trabalho de doutoramento pretende melhorar a **qualidade de serviço dos sistemas de armazenamento em centros HPC**. Para tal, é fundamental compreender o **funcionamento** e a **complexidade** destas **infraestruturas de grande escala**. Nomeadamente, como projetar soluções escaláveis, entender a orquestração dos sistemas de armazenamento e como desenvolver algoritmos eficientes para resolver problemas complexos.

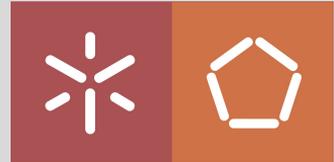
**Mariana Miranda**  
*Aluna de Doutoramento*  
*HASLab*

O meu trabalho de doutoramento pretende desenvolver novas soluções para testar a **resiliência de aplicações que gerem dados críticos**. Para tal são necessários conhecimentos de sistemas operativos, **sistemas de ficheiros** e, sobretudo, no **funcionamento da pilha de armazenamento**. Ainda, é essencial compreender como as **aplicações locais e distribuídas interagem com essa pilha**.

**Maria Ramos**  
*Aluna de Doutoramento*  
*HASLab*

**Oportunidade de trabalhar conosco em projetos nacionais e internacionais! E colaborar com outros investigadores nacionais (UL, UC), e internacionais (MCGill, FIU, UTAustin)**

# Quem?



**Universidade do Minho**  
Escola de Engenharia  
Departamento de Informática

# Equipa



**André Pereira**  
ampereira@di.uminho.pt



**António Sousa**  
als@di.uminho.pt



**Cláudia Brito**  
d13235@di.uminho.pt



**João Paulo**  
jtpaulo@di.uminho.pt



**Ricardo Macedo**  
d12010@di.uminho.pt

# A nossa porta está aberta

Piso 2  
Departamento de Informática

## Laboratórios

2.12

2.13

2.17

