



SIE

SISTEMAS INFORMÁTICOS EUROPEOS

HPC
CLUSTER
WORKSTATION
STORAGE
BIG DATA
DEEP LEARNING
AI

20
23

INDEX

Our history _____	2
Services _____	2
Our clients _____	3
Workstation _____	5
Shared memory solutions _____	6
Distributed memory solutions _____	7
Cooling _____	7
Ladon Os _____	9
Storage solutions _____	11
NFS & SAMBA solutions _____	11
CEPH server _____	11
Paralell storage solutions _____	12
Deep Learning & Machine Learning _____	13
GPUs _____	13
Nvidia Grace Hopper _____	15
AMD FPGA _____	15
ARM _____	16
IoT & 5G comunications _____	16
Connectivity _____	17
Ethernet _____	17
Low-latency networks _____	18
Artificial Intelligence _____	18
Big Data _____	20
Nvidia Software _____	21
Partners _____	22



OUR HISTORY

Sistemas Informáticos Europeos is a Spanish company with 33 years of existence and 23 years of experience in the HPC field. We provide solutions in HPC, Big Data, Storage, Deep Learning, Machine Learning, and Artificial Intelligence, covering the needs of small workstations to medium-sized clusters, with up to 368 nodes.

We offer our expertise in the field of scientific computing, spanning over 23 years, through our middleware environment called LadonOS. LadonOS is a collection of open-source tools customized for each client, optimizing hardware performance.

Depending on the client's needs, LadonOS can adapt to connectivity environments ranging from gigabit Ethernet to 200 Gb/s and Infiniband from 56 to 400 Gb/s (and in the future, 800 Gb/s), taking into consideration the project's size and client requirements.



1999 Instituto Blas Cabrera

SERVICES



HPC systems consultancy

Since 1999, SIE advice researchers about the use of intensive calculus systems.



Setting-up of the Ladon OS solution in multi-brand systems

Thanks to our experience, we can offer the solution in different systems.



Networks: 10G, 40G, 56G, 100G, 200G & 400G in Ethernet and Infiniband NDR

SIE complete his offer with switch and cards from the market leaders such as Nvidia-Mellanox, DLink, H3C, Broadcom etc...



Own software-hardware configuration SIE Ladón for scientific calculus

SIE offers the best client experience integrating all the system with its own equipment.



GPU y FPGA solutions for paralell calculus

The GPUs & FPGAs are the last HPC technology, they can accelerate programmes 50 times more than a CPU.



24/7 customer service

You can ask to our support team about any incidence in our web.



Suite HPC Ladon OS, based in Rocky Linux 8.8 y 9.X

100% compatible with Red Hat, it offers an opensource software puzzle customizable in every installation.



HADOOP systems Big Data, Deep Learning and Machine Learning solutions

The Apache Hadoop project develops open-source software for reliable, scalable, distributed computing. His software library is a framework that allows for the distributed processing of large data sets across thousands of nodes and petabytes using simple programming models



On the cloud service

We offer HPC services On the cloud for work peaks of private companies and institutions.



Storage solutions

SIE offers complete software and hardware solutions such as FreeNAS, CEPH, for NAS solutions and global disk export with BeeGFS.



Custom-made programming service

We offer a custom-made programming service for deprtaments that need to parallelize or migrate his CPU code to CUDA. Our qualified engineers can make this for you.



OUR CLIENTS

CSIC



PUBLIC CENTERS



PUBLIC UNIVERSITIES



HEALTHCARE



PRIVATE ENTITIES AND INTERNATIONAL PROJECTS



R&D AGREEMENTS



SIE appreciates the trust of hundreds of customers in our SIE Ladon systems and their contributions to enhance and make them more efficient in the scientific computing environment.



The workstations are the response to the necessity of making calculus that last more than 8 hours or to make multiple parallel calculus.

PC's are engineered for 8x5 work (8 hours a day, 5 days a week). However, for scenarios requiring 24x7x365 operation (24 hours a day, 7 days a week, 365 days a year), we need to rely on Workstations that can stay switched on continuously.

SIE Ladon Workstations are equipped with registered ECC memory, with error correction, and an additional chip, preventing calculations to hang.

These stations also offer a noise level of less than 28 dB, a sound that does not disturb the human ear, making them suitable for placement in offices and research departments.

SIE Workstations are available from two manufacturers:



GIGABYTE

Gigabyte workstation W771-Z00



INTEL

We have the new Sapphire Rapids processors. It is a new generation of Intel Xeon Scalable processors designed to enhance performance in artificial intelligence applications and containerized microservices.

It features a new performance core and integrated accelerators, improving performance across a wide range of data center use cases.

Additionally, it includes the new performance core, Intel Advanced Matrix Extensions (AMX), to enhance AI inference and training.

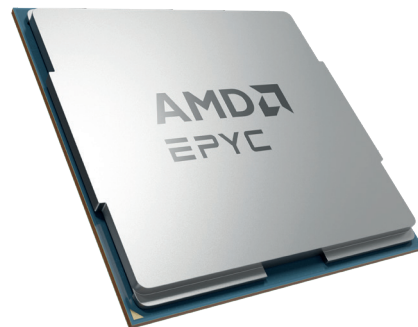
AMD

The new PCI-E 5.0 platforms are now available from AMD with Genoa, Bergamo (9004 family) and Siena (8004 family as low power solution)

It features:

- Ideal performance for pre-calculation and HPC workloads.
- Cutting-edge security features with AMD Infinity Guard offering the advanced capabilities required to help defend against internal and external threats.
- Low power consumption, backed by a broad ecosystem.

Furthermore, it increases the number of cores in the processors: initially up to 96 cores, now up to 128 cores with the new Bergamo series.



“Shared memory solutions vs. distributed memory solutions” These are two different philosophies for addressing the parallelization problem in scientific computing. Both allow for reducing the time required for mathematical processes, which is the base of research using HPC systems. Shared memory is the ideal solution when the researcher can perform calculations with a maximum of 256 cores / 512 threads and 6 TB of RAM. They provide great ease of use since there is no need to use queuing systems, and calculations can be implemented in OpenMP, which is simpler.



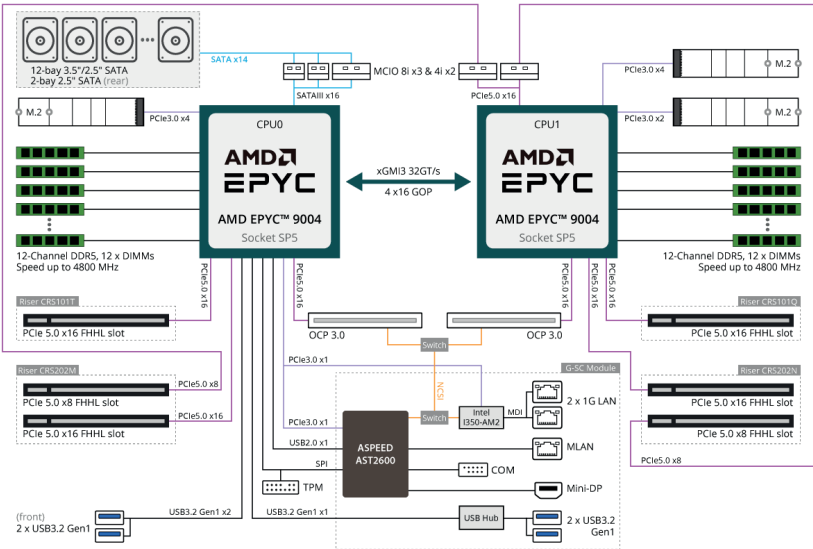
They reduce costs for items such as Ethernet switches, low-latency networks, racks, UPS capacity, and space. Scientists can use both Linux systems (CentOS, Suse, Ubuntu, etc.) and Windows 11 Pro for Workstations (valid for servers), ensuring maximum compatibility with any scientific computing application. Like in the rest of our systems, SIE offers turnkey solutions to start using them from day one.

Our shared memory systems are based on:

- Intel Sapphire and AMD Bergamo solutions in the high-end range. AMD Genoa and AMD Ryzen solutions in the mid-range and low-range, as well as the Genoa F range for users requiring high clock speeds.
- The AMD 9004 processors, which offer the highest core performance, and the latest-generation DDR4 memory solutions. They can be combined with GPU and FPGA solutions to create a high-capacity system in a single machine.

AMD Genoa & Bergamo solutions

Genoa solutions offer very competitive costs, with 256 cores / 5nm lower clock speed, but they are appealing in environments where a lot of memory and low investment and consumption costs are needed.



In recent times, their sales have increased, as they are highly optimal for genetic sequencing solutions, where a significant amount of memory integrated into a single system and substantial disk storage capacity is required.

On the other hand, Bergamo, AMD’s latest offering, is designed for the growing number of horizontal scaling workloads.



DISTRIBUTED MEMORY SOLUTIONS

The distributed memory solutions gives significant developing flexibility, costs reduction and future investment protection.

In these cases, nodes use only 1U, allowing for more than 10,000 cores to be incorporated into a single rack, providing space savings and consumption efficiency with redundant power sources of Platinum and Titanium energy efficiency levels (96% utilization).

We also offer GPU solutions with a density of 8 GPUs in 2U.

The new IPMI 3.3 gigabit implementation for remote management and high security remains compatible with Nagios-type consoles and the new Redfish standard. Furthermore, the new PCI-E 5.0 platforms introduce Dual ROM architecture, protecting systems against attacks.

Connectivity options range from Gigabit to 100G, combined with Mellanox Infiniband networks of up to 200 Gb/s and, in the future, 400 Gb/s and 1 Pb/s.

The nodes can be of mixed environments, incorporating 64-bit Intel Xeon processors, 64-bit AMD Rome processors, nodes with NVIDIA GPUs, and nodes with AMD XILINX and new FPGA models.

Sistemas Informáticos Europeos has integrated common global disk export repositories, NFS systems for storage with Samba services, automatic node shutdown based on SLURM, GPU virtualization based on RCUDA, and more.

The flexibility of these systems allows for starting with basic configurations of one server and four nodes and scaling up to solutions with 384 nodes.

These systems provide graphical administration tools, remote monitoring (GSM), and SLURM queue systems.



Extremadura University

COOLING

Heat neutral emission cluster

One of the major challenges in supercomputing is heat dissipation. As an example, SIE Ladon systems can generate up to 20,000 to 30,000 watts of thermal dissipation per rack, which is significantly higher than what typical data center environments can handle.

Power consumption is another significant limitation of HPC (High-Performance Computing) systems. Despite their integrated automatic node shutdown feature for idle nodes, these systems tend to operate at nearly 100% capacity, 24 hours a day, 365 days a year.

To address this problem, SIE has developed Gas Cooling Backplanes in collaboration with Focus Cooling. Unlike water cooling backplanes used for years, they avoid issues such as maintenance and the risk of leaks. Additionally, they don't require a minimum number of full racks in the installation to be profitable.

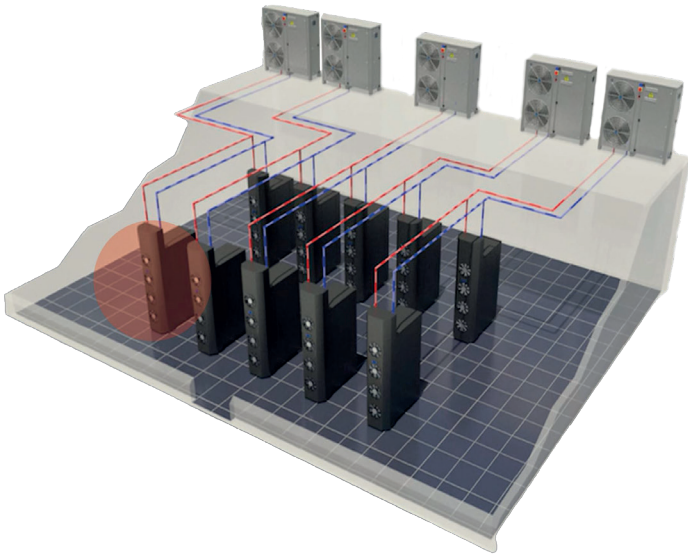


Focus Cooling backplane

This gas is innocuous and doesn't have the problems associated with water and it doesn't affect electronic components. Another advantage is scalability; you can start with a 10 KW system with just one rack and scale up to a 55 KW condenser. They don't require hot and cold aisles or the installation of raised floors.

They can be maintained as cooling systems, similar to the central air conditioning units of the institution, and their scalability is straightforward and cost-effective. As a single data point, although it may vary depending on the project size, they only represent a cost of 1 euro per watt (W).

On the other hand, the efficiency of a conventional air system, due to the recirculation of hot air, is approximately 1-1.8 (for every 1,000 watts of heat generated, 800 watts are needed to cool it). In contrast, these systems offer an efficiency of 1-1.2 (for every 1,000 watts, only 200 watts are needed to cool it). There is no recirculation of hot air, and as it cools attached to the back of the rack, the δT° (temperature difference) is always positive.



SIE has several systems installed at IRB (Barcelona Science Park), the University of Barcelona, the University of Extremadura, DIPIC, the University Rey Juan Carlos, and more.

Inmersion cooling

Recent increases in electricity rates, coupled with projections of higher consumption between 2023 and 2025, make the implementation of new oil-immersed data centers essential.

We introduce a pioneering hosting service based on this innovative technology, unique in the entire Iberian region, which also complies with European safety regulations and allows for densities of up to 100 KW per rack.

This technique, aimed at cooling IT equipment, consists in immersing its components in a thermally conductive and dielectric liquid. With this approach, servers are kept cool, and heat is directly dissipated into the liquid, offering significant environmental and long-term energy-saving advantages.

In collaboration with Submer, we provide consulting and installation of these innovative systems. SIE offers Gigabyte equipment manufactured and certified with these systems, saving on the power consumption of machine fans. This allows us to achieve a PUE of 1.02.



Submer SmartPod



LadonOS is based on CentOS and Rocky Linux distributions, which are open-source alternatives to Red Hat. By using this distribution, the system achieves a perfect balance between reliability, security, and efficiency. It's optimized to provide an environment of total stability on Rocky Linux 8 and 9 or Red Hat 8 and 9 for those environments seeking official support.



Differentiating factor

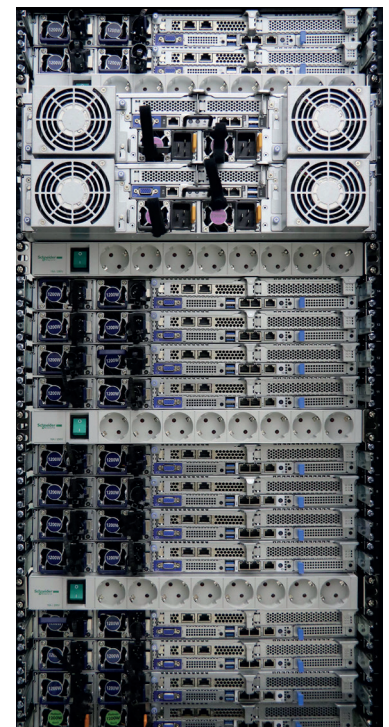
- A completely OpenSource environment, creating an ecosystem of applications.
- Each installation is customized to the final project down to the last detail, optionally offering training for both administrators and users on technical, security, and management aspects.
- Years of updates at no additional cost.
- Base, Epel, and Stream repositories are provided as needed by the project, improving libraries and programs.
- At any time, growth is possible by adding new nodes or upgrading the existing environment, whether in terms of computing capacity, GPUs, FPGAs, or communications.
- NDB support is available during the warranty period.

First-hand consultancy for both everyday tasks and more demanding ones, including:

- System security through user and group control.
- Resource optimization using SLURM-based Queue Managers with automatic node shutdown.
- Software and libraries installation.
- Virtualization in containers or management through modules.
- Backup and duplication solutions.
- Remote and telephone support, with technicians having over 10 years of experience.

Fully customizable hardware down to the last detail:

- Hard drive partitioning.
- Network environment.
- Redundancy and security environments.
- Node and environment partitions.
- Auditing always done with the client to tailor it to real needs and even the smallest detail.
- 100% compatible with various hardware manufacturers.
- Turnkey system.



Universidad de Barcelona

The Ladon OS day by day

LadonOS is designed for 24x7 production environments, and as such, we have a set of control, management, and monitoring tools for component oversight. Installation is carried out using Kickstart, allowing for both initial installation and the rapid restoration of a functional node within a matter of minutes.

The server is highly secure, featuring a web gateway to the rest of the nodes for patch and update deployment, as well as the integration of a firewall and various tools to thwart third-party attacks. System management is performed through IPMI, enabling power control, hardware monitoring, proactive maintenance tasks, and disaster recovery if needed.

Additionally, we have integrated an LDAP-based NIS domain with the following advantages:

- User and group management
- Web-based environment for service management
- Integration of DNS services for hosts
- Integration with Active Directory

Redundancy

LadonOS allows the installation of multiple primary service servers in High Availability (HA), which will provide us with unlimited security in production environments, even if one of the main servers encounters an error. This includes virtualization of various services, such as Proxmox.

In the unlikely event of an error, the primary server transition process is transparent to the user.

Ladon OS: Applications

We can always customize the software and it's not limited to OpenSource. In addition, all kinds of licensed software can be installed on this OS due to the compatibility of the RHEL 8 and 9 core. Furthermore, we provide the system with Docker or Kubernetes containers for compatibility with various software versions.



Queue managing: SLURM

- Scalability and performance
- Free and OpenSource as our philosophy
- Compatibility with a wide range of programming languages
- Utilizes 100% of resources
- Customizable jobs on demand
- Professional support offered by SchedMD
- Integrated node shutdown system for reduced energy costs and greater ecological benefits.



STORAGE SOLUTIONS



SIE offers a wide range of solutions, from simple needs to secure repositories for Big Data or Parallel Storage, for all kinds of applications in various fields such as:

- Bioinformatics
- Engineering
- Data analysis
- Backup

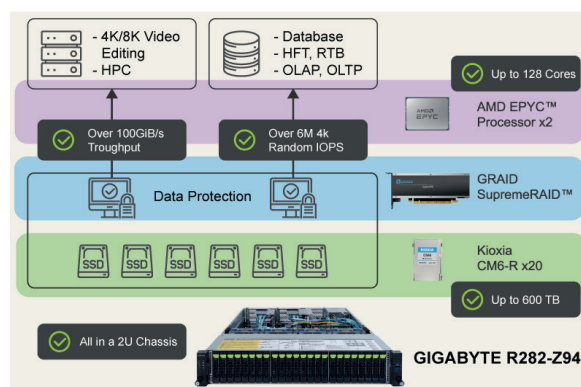
NFS & SAMBA SOLUTIONS

NFS

Based on the native Linux protocol NFS (Network File System), which is a distributed file system optimized for a network environment, allowing us to handle files as if they were local.

It follows the client-server model, where one system provides access to its storage resources (files and directories) over the network to other systems acting as clients.

The inclusion of GRAID supremeRAID controllers allows us to offer solutions with up to 24 NVME disks and 16 million IOPS.



SAMBA

The SAMBA protocol is a free software implementation that enables interoperability between Windows and Unix/Linux-based operating systems.

With SAMBA, it is possible to share files, printers, and other resources between these systems, allowing computers with different operating systems to communicate and collaborate on a network.

CEPH SERVER

Ceph is an open source storage system designed to handle a big amount of data, offering high availability, error tolerance and performance.

This servers plays an important role in the administration and distribution of the cluster's data.

Each server have an specific part in the storage process, redundancy and data recovery in order to create a reliable and scalable system that can be adapted to the storage necessities of big amounts of data.



BeeGFS

For parallel computing environments and distributed file systems, the BeeGFS solution is used. BeeGFS is a global disk export solution that allows all nodes to write simultaneously to a single centralized repository.

BeeGFS solutions can utilize 1 Gb LAN, the latest 100GBaseT, or even Low-Latency Infiniband networks in any of their modes (EDR, HDR up to 400GB). You could even start with Ethernet and later transition to Infiniband, making use of the entire infrastructure with minimal downtime.

BeeGFS services include:

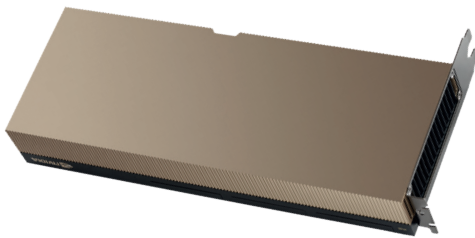
- **Distributed Storage:** BeeGFS divides data into blocks and distributes them across multiple nodes in a cluster. This enables parallel data access, significantly improving read and write performance, especially for applications dealing with large data volumes.
- **High Availability:** BeeGFS is designed to be highly reliable. Data is automatically replicated on different nodes to ensure availability in the event of hardware or network failures.
- **Scalable Performance:** BeeGFS can horizontally scale to handle increasingly larger workloads. As more nodes are added to the cluster, the overall file system performance increases proportionally.
- **Simplified Management:** BeeGFS provides management and monitoring tools that allow administrators to monitor system status, adjust configurations, and efficiently perform maintenance tasks.
- **Low Latency:** BeeGFS is Infiniband-native to minimize latency, making it suitable for applications requiring fast data access, such as simulation, scientific data analysis, and rendering.
- **Multi-Protocol Compatibility:** BeeGFS offers support for various access protocols, including POSIX (Unix/Linux standard file system), NFS (Network File System), and SMB (Server Message Block), enabling existing applications to work with the file system.
- **Caching and Tiering:** BeeGFS allows for storage layer configurations, including the use of high-speed storage systems (such as SSD SAS or SSD NVME) for caching (metadata) and slower but higher-capacity data storage layers (like HDDs) for long-term storage.
- **Integration into the HPC Ecosystem:** BeeGFS integrates well with other typical components of HPC clusters, such as job management systems and tape storage systems, making it easier to build and manage high-performance computing environments.



Just like 8 years ago, out of every 10 systems that were commercialized, only 1 incorporated GPUs, and just 2 years ago, it was 50% of them. Currently, SIE supplies more systems with GPUs-based coprocessors than those without them.

The approach to entering the world of GPUs requires certain prerequisites:

1. The calculation must be highly parallelizable.
2. It should not require a significant amount of memory, although this limitation has partly dissipated with cards having 96 GB of memory and systems with 8 GPUs interconnected by NVLINK.
3. It should be migratable to the CUDA environment. More and more languages like C++ , Python, and proprietary applications like Matlab have libraries that facilitate this migration. Open Source environments like OpenACC and Rapids have enabled more scientists to transition their applications from x86 architectures to CUDA relatively easily.



GPUs

The fact that there has been upward compatibility from CUDA 1 to CUDA 12 has given the computational environment confidence that what has been developed previously will continue to be operational in future versions, guaranteeing future investments.

Currently, Sistemas Informáticos Europeos offers a wide range of solutions that we can group from smaller to larger:

- SIE Ladon GPU Workstation: Systems for office or low-noise laboratory environments that do not require a rack.
- SIE Ladon GPU Basic: This is a system designed for a standard 19" rack in 2U, which combines the use of GPUs with high storage capacity, allowing it to be a standalone system (integratable in the future into a LadonOS cluster). In terms of GPUs, it can support up to 4 cards from the RTX, Compute, or Visualization series, and in terms of storage, it can go up to 200 TB based on SAS Enterprise disks or SATA SSDs, providing RAID 5 or 6 security.



This system features 12 front bays for 3.5"/2.5" drives and two rear bays for 2.5" drives. It is equipped with Broadcom/LSI 3108 controllers with 16 ports, allowing the disks to be segmented into 4 groups. This configuration enables the disk throughput to be accessible from the processor calculations without bottlenecks. Even as the most basic system, it comes standard with redundant power supplies with Platinum efficiency.

SIE Ladon Multi GPU: In this type of node, we can provide solutions with 8 GPUs in 2U or up to 10 GPUs in 4U. Card connectivity is achieved through PCIe 3.0 or 4.0 buses, depending on whether it's based on Xeon Refresh or AMD Rome processors, respectively.



These nodes are ideal as standalone units or integrated as supported GPU nodes within clusters using the LadonOS system, with either Ethernet or InfiniBand RDMA connectivity, depending on the customer's needs.

SIE Ladon LDX-8:

This is the most advanced solution, available in LDX-4 (with 4 GPUs) or LDX-8 (with 8 GPUs) models, all interconnected through NVLINK interfaces. Due to their special features, these systems come pre-configured with all NVIDIA H100 GPUs and, depending on the usage, HPC or Deep Learning libraries such as Keras, TensorFlow, etc. They can be based on Ubuntu, CentOS, or Red Hat operating systems, with LadonOS middleware.



SIE Ladon GPU Workstation:

These are systems designed for office or low-noise laboratory environments that do not require a rack. They are marketed as turnkey closed solutions, exclusively through our partner, Informática El Corte Inglés, due to their special characteristics and advanced technology. Support is available in both English and Spanish, with immediate remote or telephone assistance and 9x5 NBD (Next Business Day) incident resolution.

In addition to the standard 3-year on-site warranty for all SIE Ladon systems, they include installation at the customer's site. Optionally, training, consultancy, and the development of tools for structural biology and drug discovery projects, focused on vertical sectors of the industry, can also be provided with them.

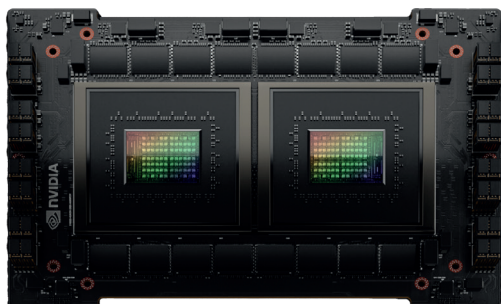


NVIDIA GRACE HOPPER

It's a groundbreaking accelerated CPU designed from the ground up for AI and high-performance computing (HPC) applications at a large scale. The superchip will offer up to 10 times more performance for applications handling terabytes of data, enabling scientists and researchers to achieve unprecedented solutions for the world's most complex problems.

The Grace Hopper Superchip combines the Grace and Hopper architectures using NVIDIA NVLink-C2C to provide a coherent CPU-GPU shared memory model for accelerated HPC and AI applications. The solution is implemented by Gigabyte.

- CPU and GPU designed for large-scale AI and HPC
- New coherent interface of 900 gigabytes per second (GB/s), 7 times faster than PCIe Gen5.
- Powers accelerated computing and generative AI with GPU HBM3 and HBM3e memory
- Runs all NVIDIA software stacks and platforms, such as NVIDIA HPC SDK, NVIDIA AI, and NVIDIA Omniverse. It's the first solution where memory is coherent between CPU and GPU.



AMD FPGA

AMD offers a comprehensive portfolio of multiple nodes to address requirements across a wide range of applications.

They provide a low-cost, small-sized Field-Programmable Gate Array (FPGA) array to take software-defined technology to the next level. AMD FPGA offers system integration while optimizing performance per watt.

AMD's acquisition of Xilinx brings new technology to the field of FPGAs, leveraging this investment for future developments, including enhancing FPGAs focused on Artificial Intelligence.



Whether you are designing a high-performance, cutting-edge network application or seeking a cost-effective, small-sized FPGA, FPGAs and 3D Integrated Circuits (3D ICs) in the Artix®, Kintex®, and Virtex® families facilitate system integration while optimizing performance per watt.

AMD XILINX Alveo™ data center accelerator cards are designed to meet the ever-changing needs of modern data centers, providing up to 90 times the performance of CPUs for key workloads, including machine learning inference, video transcoding, and database search and analysis.



Built on the AMD 16nm UltraScale™ architecture, Alveo accelerator cards are designed to accommodate changing acceleration requirements and algorithm standards. They can accelerate any workload without hardware changes and reduce the total cost of ownership.

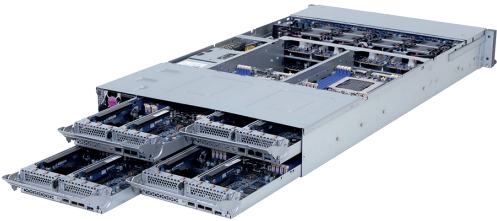
ARM



The latest developments based on CentOS and Ubuntu now have a structure for use on ARM architecture platforms. LadonOS 8 includes a version for this technology.

The CentOS ARM distribution is compatible with ARM systems like AmpereOne (64-bit ARM) in all its variations. The new 2-socket ARM Ampere Altra platforms bring new levels of power. These new platforms allow for double twin ARM-based machines with 10G connectivity and DDR4 memory. Furthermore, these new platforms have 6 Gb/s SATA and 12 Gb/s SAS connectivity and feature a “mezzanine” slot for Infiniband and other possible connections.

With the new ARM Cavium generation, we can tackle high-powered and high-density HPC solutions, offering much lower power consumption (from 310 W to 190 W) than Intel platforms and at a more cost-effective price. Thanks to the efforts of the Open Source community, we have other publicly available tools like Python, PHP, KVM, etc., that complete the ecosystem.

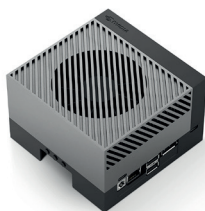


IoT & 5G COMMUNICATIONS

The introduction of 5G networks will enable communication between IoT devices. These devices will be connected and capable of exchanging data at high speeds, facilitating real-time decision-making. With the support of Gigabyte and Nvidia, SIE has the complete ecosystem required for this:

1. Low-power IoT devices with a significant number of network ports and RS-232 communication capabilities, allowing them to locally filter relevant data and send it to the cloud. In this field, collaborations with universities such as the Polytechnic University of Valencia, the Catholic University of Murcia, GlobalIoT, Nutricontrol, and Vielca enable us to offer solutions for greenhouse and plantation control and monitoring. The range of professional Intel drones allows us to create swarms of them to monitor extensive crop areas and supervise pest outbreaks. We have solutions such as the Gigabyte Brix and the Nvidia Jetson AGX Orin developer kit.





2. Telco equipment, designed for real-time 5G signal processing, offers solutions with up to 10 GPUs, allowing for the incorporation of RTX or Tesla cards based on the customer's needs, or Intel FPGAs. In the new 5G networks, signal processing needs to be faster, and priorities are established based on layer 7, at the application level. There is also a new low power solution based on AMD 8004PN serie (Siena).

Much of the intelligence of routers is transferred to servers. The solutions provided by SIE are well-suited for these environments and equipped with coprocessors, offering lower response times and latency.

3. Lastly, data must be analyzed in Machine Learning and Deep Learning systems, which require ample storage capacity and efficient searching within unstructured information. Big Data systems address this issue.

CONNECTIVITY

Clusters, like any other computer system, have seen improvements in their connectivity. These systems have their own internal networks for two reasons:

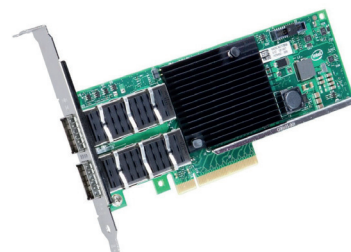
1. To prevent network traffic from the institution or company from affecting and potentially overwhelming the internal management of the system. Additionally, they use specific low-latency networks that are not relevant in general-use environments.
2. Particularly in Ethernet networks, to enable the implementation of software firewalls between independent physical cards, enhancing security and preventing attacks.

ETHERNET

Ethernet networks have evolved significantly in the last 20 years. In terms of IPMI connectivity, we have transitioned from a standard of 100 Gb/s to gigabit, allowing remote management systems to be much more graphical and have vastly improved functionality.

For data communication, we are witnessing an unstoppable evolution. Storage within the cluster and Big Data systems are demanding exponential growth in bandwidth. Gigabit networks are practically disappearing in favor of 10GbaseT Ethernet, which is compatible with the previous standard and is now standard on most servers.

Networks at 25 Gb/s, 50 Gb/s, and 100 Gb/s are becoming common as backbones and in critical systems that handle up to petabytes of data.



LOW-LATENCY NETWORKS

Today, InfiniBand is the IEEE standard that has become the preferred choice for low-latency connectivity for several reasons:

- 1. Since it is commonly used within clusters, copper-based connections make its deployment and configuration much easier while also saving on infrastructure costs.
- 2. From the DDR versions at 20 Gb/s, QDR at 40 Gb/s, FDR at 56 Gb/s, EDR at 100 Gb/s, and the HDR at 200 Gb/s, the new NDR at 400 Gb/s up to what we will see in the coming years at 800 Gb/s, they offer the advantages of upward compatibility, similar to Ethernet.
- 3. Although the leading manufacturer in this segment is Mellanox, as it is a market standard, it ensures long-term continuity and full compatibility. Both Open MPI and Intel MPI have Open Fabric versions.



INFINIBAND PACK



ETHERNET PACK

It is a protocol with much lower latency and a smaller header and trailer than Ethernet. Unlike Ethernet, it does not have as many options for rerouting and error correction, but it prioritizes its lightweight nature and speed in handling frames and managing collision semaphores.

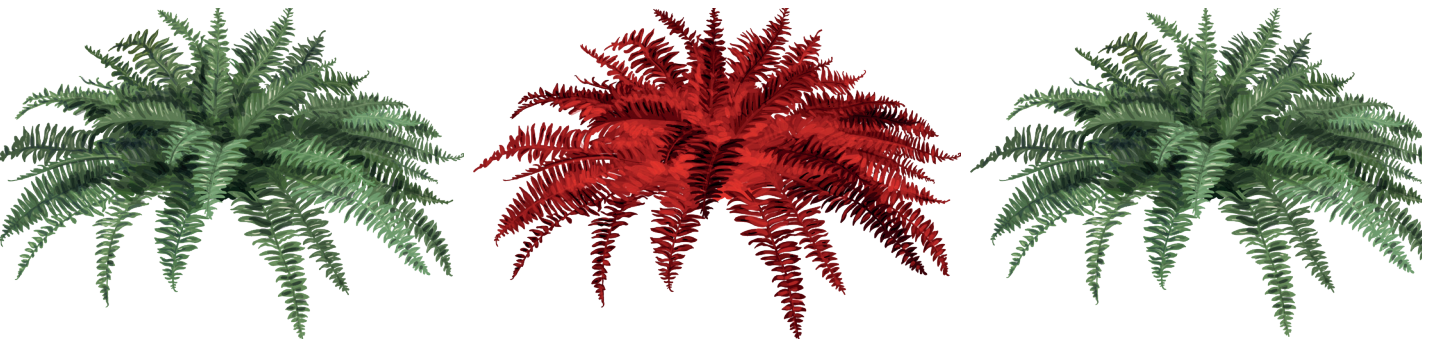
ARTIFICIAL INTELLIGENCE

The evolution of computer systems has allowed theories like those of Alan Turing to become a reality. Computers now achieve accuracy rates of 99.9% or even higher, superior even for humans.

Now, all these mathematical theories can be put into practice. According to prestigious consulting firms, the change of the decade will require scientists to transition from processing on the scale of Pflops to the scale of Exaflops. Additionally, the new AI-driven decision systems demand that this process must be carried out in real-time.

Let's consider an example. A hyperspectral camera can scan from the infrared to the ultraviolet spectrum, recording all these values. This means that the information captured at a single point goes from having 3 basic colors (Red, Green, Blue with their 256 shades) to having 1,000 values.

For a human to analyze it, these values from the invisible spectrum (infrared and ultraviolet) need to be represented within the visible spectrum that can be perceived by the human eye.



In a simple example, if we take a photograph of 3 ferns, and the one in the center is artificial, with a conventional camera, we would see 3 identical plants that, to our eyes and especially when viewed on a screen, appear to be identical, even though they are not in their composition. Two of them have living plant cells, and the artificial one is made of polymers.

Now, let's apply this to medical imaging: Imagine a patient who undergoes a CT scan and a cancerous tumor is discovered, and for added sensitivity, let's say it's located in the brain. When the surgeon begins the operation, the cancerous area will be where the scan indicated it to be. However, as soon as the cranial bone is removed, the patient's shape will shift, and the surgeon will have to operate blindly.

But if we had an artificial intelligence system that could translate the hyperspectral camera image of the tumor tissue (the anomaly, like the artificial fern) into the visible spectrum in real-time or within seconds, the surgeon, using markers, would know exactly where to make the incision with the scalpel.

The application is clear: going from approximate cutting and operating blindly to knowing precisely where the tumor is can make a significant difference in the patient's future quality of life and the success or failure of cancer removal.

This is just one example of the many fields where Artificial Intelligence can assist in science, and achieving Exaflop-level computing capabilities in real-time represents a significant scientific advancement.

SIE Ladon systems for training (pre-learning of the AI environment) and inference (the use of the system at high speed) allow us to address these and other problems, and they will continue to improve and increase their capacity rapidly.

NVIDIA Tesla cards, being an application of GPUs and graphics processors to high-performance computing systems, are highly effective. The SIE Ladon 32 A4000 and 16 A4000 solutions allow for the deployment of 32 and 16 cards on a 2U and 4U platform, respectively.



These systems provide more than 8 ExaTops and are ideal for performing complex tasks in the field of Artificial Intelligence. Another advantage of this solution, based on the new Tesla T4 cards, is that they consume less than 3,000 watts, making it an environmentally-friendly system.

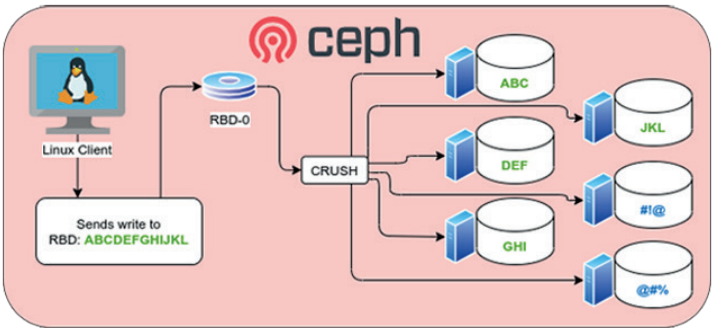
We also offer solutions based on the new Intel Stratix 10 cards, which efficiently address the problem with energy efficiency.

BIG DATA

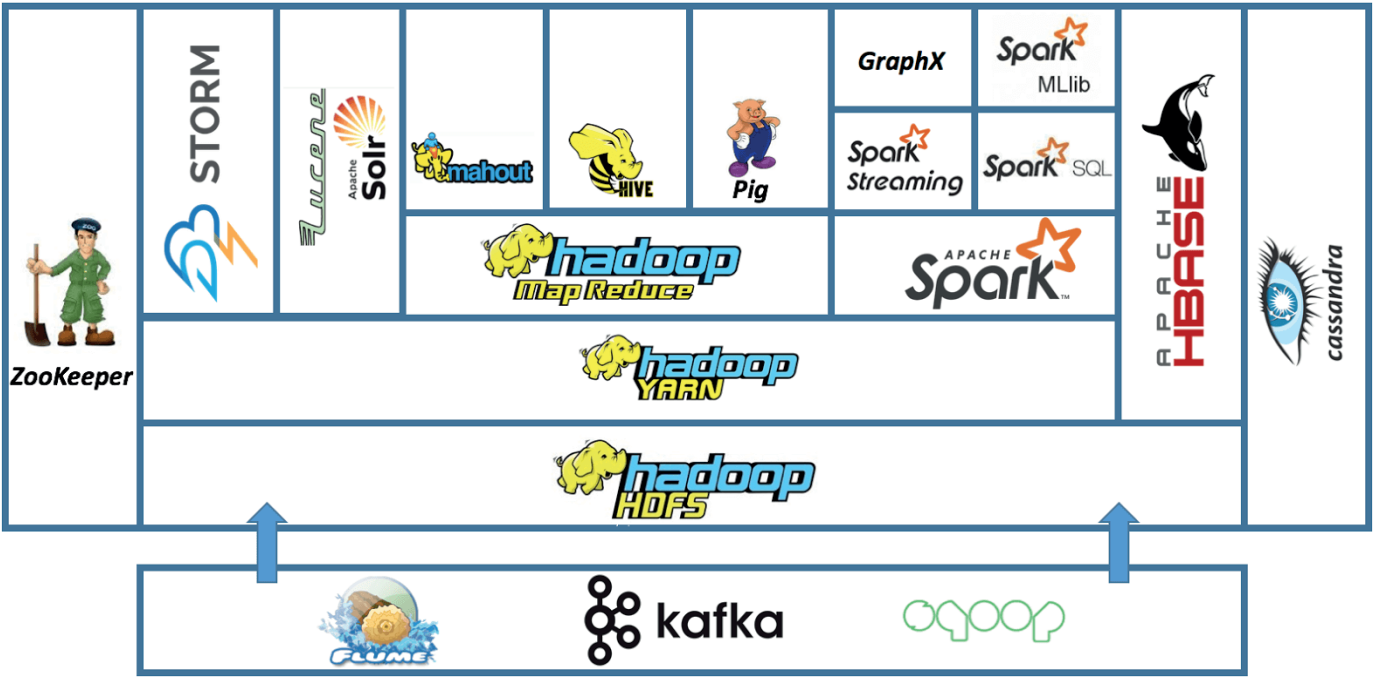
Stratio

Stratio is a Hadoop-based big data environment that is straightforward to configure and delivers excellent performance. All of this is manageable through a user-friendly web interface.

Ceph File System is a free distributed file system designed for use with large volumes of data and Big Data. Ceph aims to be distributed without any single points of failure. The data benefits from error-free replication, making it highly fault-tolerant.



Additionally, it allows for the separation of Metadata components, caching, and storage, dynamic data rebalancing, subdirectory snapshots, FUSE support, NFS export, and compatibility with HADOOP.



Nvidia Omniverse

This is a platform and set of tools developed by NVIDIA with the aim of facilitating collaboration and the creation of real-time 3D content.

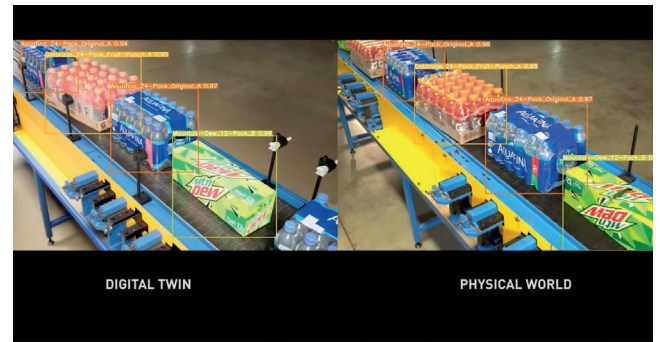
It is designed to connect different teams involved in digital content creation, allowing them to work together more efficiently and effectively in a shared virtual environment.

One of the major innovations is that Omniverse accepts a wide range of file formats as it is based on Universal Scene Description (USD) and provides compatibility between software programs such as Blender, Photoshop, or Unreal Engine, for example.

Another key feature of Omniverse is the creation of real-time collaboration environments. This means that multiple people can collaborate on the same project simultaneously, including features like real-time rendering.

SIE have also worked with workstations like the W771-Z00 that support Nvidia Omniverse, proving to be an ideal combination for tasks such as digital twins or 3D simulations.

We offer training and implementation services for this ecosystem.



Nvidia vGPU

It is a technology developed by NVIDIA that allows for the sharing and virtualization of the graphical processing power of NVIDIA graphics cards in virtualization environments. In essence, vGPU enables multiple virtual machines (VMs) on a server to efficiently share and utilize graphical resources, which is particularly beneficial in enterprise, data center, and cloud computing environments.

With vGPU, graphically intensive applications within virtual machines can benefit from the performance of NVIDIA graphics cards. Additionally, vGPU technology is managed through management software provided by NVIDIA. This enables IT administrators to centrally allocate and manage graphical resources, simplifying the deployment and maintenance of virtualization solutions.

Using vGPU can be more cost-effective compared to providing a separate physical graphics card for each virtual machine. Furthermore, it can achieve greater resource utilization and improved workload distribution.





SIE is currently Elite level partner in the NVIDIA Partner Network Solutions Integration (NPN) for Compute and Preferred for Visualization, focusing on HPC, Deep Learning, and AI environments. This agreement allows us to have certified technicians and sales professionals with high qualifications to undertake GPU-based projects. Furthermore, it enables us to offer special discounts for Education.

In addition to NVIDIA, we also use Mellanox's InfiniBand technology, with EDR, FDR, and HDR solutions of up to 200 Gb/s. They also provide Ethernet solutions of up to 100 Gb/s with very low latency. NVIDIA's recent acquisition of Mellanox will create new connectivity synergies between GPU nodes through RDMA protocol and new implementations.

NVLink and NVSwitch connectivity allow us to harness the power of up to 64 GPUs in complex AI problems.

GIGABYTE™

Gigabyte is known for its production of motherboards and other hardware components. Over the years, Gigabyte has built a strong reputation in the technology industry and has gained recognition for the quality and innovation of its products. They have been involved in the development of advanced technologies and have been an influential presence in the computer hardware components market.

SIE works directly with Gigabyte in Taiwan and the Netherlands (now GigaComputing), which allows us to become the first Spanish manufacturer certified by this Asian giant to integrate their platforms into CPU, GPU, and storage computing solutions. We are the only Gigabyte Server Integrator and Gigabyte Distributor simultaneously for Southern Europe and its one of the most important partners.



AMD provides technologies to accelerate a wide range of data center workloads, from general computing tasks to technical computing, native cloud computing, and accelerated computing. Furthermore, it offers advanced AI acceleration to deliver outstanding performance and efficiency to make the world smarter.

SIE leverages AMD's expertise in supercomputing by acquiring their advanced processors and integrating them into their solutions.



SIE has been collaborating with Inetum for over seventeen years, both in the public and private sectors, as well as on international projects. Together, we have implemented solutions based on SIE Ladón, as well as others on multi-brand platforms, installing the SIE Ladon OS solution for cluster deployment. SIE works with Inetum through the Spanish Dynamic Procurement System (Centralized State Procurement).



Our projects endorse our experience in the field of HPC.
We are an efficient and experienced team with more than 30 years of successful projects.

Many of them are with the most prestigious institutions:

- Vigo University
- Autonomous University of Madrid
- Zaragoza University
- Jaén University
- Polytechnic University of Valencia
- La Laguna University
- Sevilla University



- Rey Juan Carlos University
- Barcelona University
- Aveiro University
- Addis Ababa University
- Biodonostia
- Servicio Murciano de Salud
- Donostia International Physics Center (DIPC)