

Omnia open-source software

Run AI, HPC and data analytics on the same systems, with greater flexibility.

Table of Contents

Build and manage complex environments, and run them anywhere.	2
Dell Technologies has what you need.	3
Omnia use cases	4
Why Omnia?	4
Deploy faster.	4
Flex with demand.	5
Get instant gratification.	5
Customer success stories	5
Technical specifications	6
Omnia software	6
Support matrix: Software and hardware requirements	6
Services and financing	8
Why choose Dell Technologies	8
Customer Solution Centers	9
AI Experience Zones	9
HPC & AI Innovation Lab	9
HPC & AI Centers of Excellence	9
Proven results	9
Take the next step, today.	10

Build and manage complex AI, HPC and data analytics environments, and run them anywhere.

The convergence of artificial intelligence (AI), high performance computing (HPC) and data analytics is being driven by a proliferation of advanced computing workflows that combine different techniques to solve complex problems. For example, AI and data analytics can augment traditional HPC workloads to speed scientific discovery and innovation. At the same time, data scientists and researchers are developing new processes for solving problems at massive scale that require HPC systems.

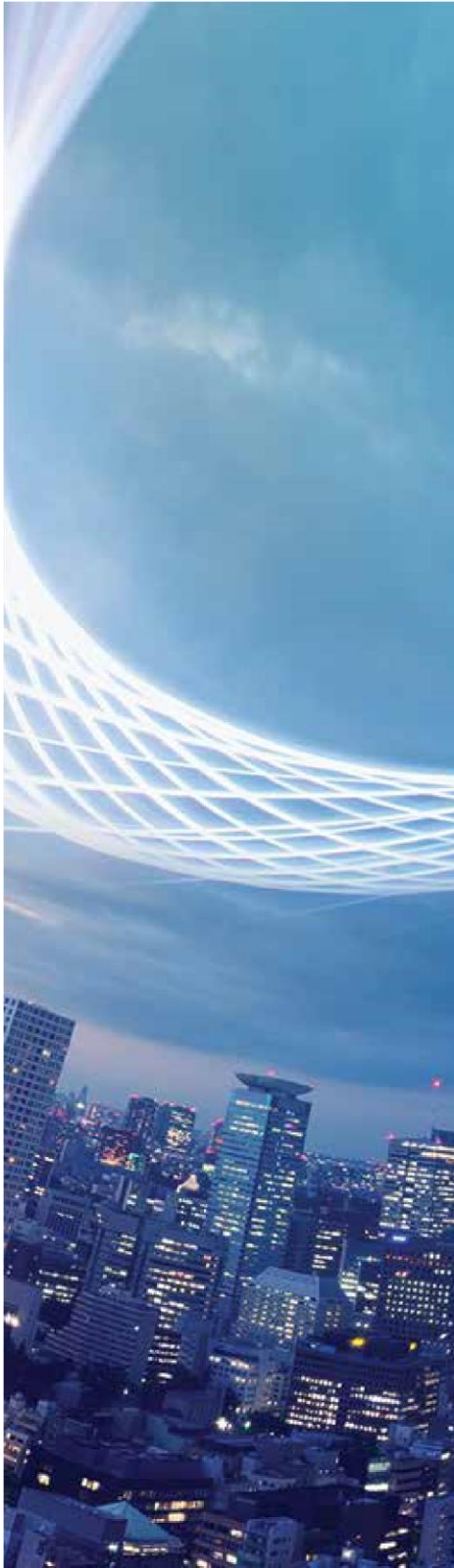
While this convergence is accelerating discovery and innovation, it's also putting pressure on IT to support ever more complex environments. IT teams are being asked to complete manual configurations and reconfigurations of servers, storage and networking as they move nodes between clusters to provide the resources required for shifting workload demands.

The Omnia software stack helps speed and simplify the process of deploying and managing environments for mixed workloads. It abstracts away the manual steps that can slow provisioning and lead to configuration errors, automating the deployment of Slurm® and/or Kubernetes® workload management software along with libraries, frameworks, operators, services, platforms and applications.

For advanced computing applications such as simulation, high-throughput computing (HTC), machine learning (ML) and deep learning (DL), and data analytics, Omnia gives IT the flexibility to run these workloads in the same environment, with a single interface for cluster provisioning and deployment, with easy-to-use point-and-click templates.

With Omnia, you can compose a unified architecture with multi-purpose, balanced nodes to support multiple workloads, and quickly re-compose resources to meet demands both now and in the future. Omnia is an open source community project started in the Dell Technologies HPC & AI Innovation Lab and you're invited to download and participate on [GitHub](#).





Dell Technologies has what you need.

Expertise and guidance

Technology is evolving quickly, so your team may not have time to design, deploy and manage solution stacks optimized for multiple workloads. While advanced AI, HPC and data analytics computing might seem like the latest IT trend, Dell Technologies has been a leader in the advanced computing space for over a decade, with proven products, solutions and expertise. Dell Technologies has a team of AI, HPC and data analytics experts dedicated to staying on the cutting edge, testing new technologies and tuning solutions to your applications to help you keep pace with this constantly evolving landscape.

Dell Validated Designs

Validated Designs are workload-optimized rack-level systems with servers, software, networking, storage and services to scale faster with the confidence of an engineering-tested solution while saving valuable time and resources.

- Validated Designs for HPC are scalable systems tested and tuned for specific vertical market applications such as life sciences, digital manufacturing and research.
- Validated Designs for AI help make AI simpler with designs enabling you to get faster, deeper insights delivered with proven AI expertise.
- Validated Designs for Analytics speed time to insight with architectures, integrated systems and services optimized for big data analytics.
- Validated Designs for HPC Storage make it easier to unlock the value of your data with scalable solutions for NFS, PixStor™ and/or BeeGFS® storage.

Solutions customized for your environment

Dell Technologies uniquely provides an extensive portfolio of technologies to deliver the advanced computing solutions that underpin successful AI, HPC and data analytics implementations. With years of experience and an ecosystem of curated technology and service partners, Dell Technologies provides innovative solutions, workstations, servers, networking, storage and services that reduce complexity and enable you to capitalize on a universe of data.

Omnia use cases

Omnia lends itself to a number of use cases, including:

- AI clusters for training of large-scale generative models such as LLM or image generation
- AI inferencing for serving end user access of multiple instances of pre-trained models
- HPC clusters for tightly-coupled and loosely-coupled parallel computation
- General-purpose accelerated server clusters with RDMA-enabled fabrics
- High performance data analytics (HPDA) clusters for large-scale distributed data analysis
- Multi-user HPC/AI/HPDA systems
- Virtualized HPC/AI/HPDA clusters
- Hybrid HPC/AI/HPDA clusters
- Edge deployments for AI inference.

Why Omnia

Omnia is open-source software for deploying and managing high performance clusters for HPC, AI, and data analytics workloads. For Dell PowerEdge servers running RPM or DEB-based Linux® OS images, Omnia installs Kubernetes and/or slurm for managing jobs and enables installation of many other packages and services for running diverse workloads on the same converged solution. Developers are continually extending Omnia to speed deployment of new infrastructure into resource pools that can easily be allocated and re-allocated to different workloads. Omnia can make it faster and easier for IT to provide the right tools for the right job on the right infrastructure at the right time.

Deploy faster.

When HPC teams are asked to run AI and data analytics jobs within the same infrastructure to help save time and resources, reconfiguring can be a manual and time-consuming process. Omnia automates the deployment of high-performance clusters for AI, HPC, and data analytics workloads to create a single pool of flexible resources. It imprints a software solution onto each server based on the use case — for example, HPC simulations, neural networks for AI, or in-memory graph processing for data analytics — to reduce deployment time from weeks to hours.

Flex with demand.

When provisioning and deployment are tied together in an image-based architecture, teams can't quickly pivot, or flex to meet specific workload needs while taking advantage of new and diverse technologies. Simulation and modeling are compute-intensive with jobs submitted by a batch scheduler, taking hours or days to run. Data ingest requires very high bandwidth (GB/s) performance at scale to sustain data rates. AI training requires high throughput (IO/s) and low latency for continuous and repetitive computational analysis of the data.

Omnia can compose the solution stack to support a variety of workload demands for technologies from multiple vendors with an infrastructure-as-code approach. It supports multiple user and workload types as well, enabling you to compose and recompose resource pools. It uses repeatable, simplified workflows that enable you to build, scale and manage complex environments, based on component modules, profiles and roles, that can run anywhere.

Get instant gratification.

Workload-specialized systems are often siloed with a diverse set of hardware and software combinations accumulated over time. While it may have been the best option then, proprietary or closed software can limit choice of applications, developer platforms, desired libraries, middleware, operators and back-end services.

Omnia is open source, so you can shape it to meet your needs in an instant and in the future. Dell Technologies integrates the latest open-source tools and invites you to participate in the community. The collective power of a talented community working in concert not only delivers more ideas, but also speeds development and troubleshooting — all with support available from Dell Technologies.



You don't need to become an expert when using Omnia. Omnia abstracts away the manual steps that can lead to configuration mistakes.

Customer success stories

Arizona State University: [Accelerating scientific research with HPC](#)

Flexible resources

Testing Omnia for cloud environments

Advance science

Dedicated to giving minutes back to science

Simplified

Workload migration

Read more [customer stories](#)

Telemetry and Cluster Visualization

- Omnia fetches key performance indicators from iDRAC (power, temperatures) and at the OS level (CPU/GPU/RAM utilization) in the cluster
- Omnia also supports fetching health and job data from Kubernetes managers and Slurm controllers
- The telemetry data is plotted on [Grafana](#) to provide better visualization capabilities.
- Four visualization plugins are supported to provide and analyze iDRAC and Slurm data:
 - Parallel Coordinate
 - Spiral
 - Sankey
 - Stream-net (aka. Power Map)

Technical specifications

Omnia software

Omnia is a collection of Ansible playbooks for automating OS provisioning, deployment of open source Kubernetes and Slurm services; installation of hardware drivers, optimization libraries, and machine learning frameworks/platforms; and serving pre-trained AI inferencing models from the internet or running of research or commercial HPC applications Kubernetes. This is accomplished by installing software from a variety of sources, including Standard Rocky® and [ELRepo](#) repositories, Helm® repositories, and Source code compilation. Omnia can be downloaded from GitHub® at: <https://github.com/dell/omnia>

Support matrix: Software and hardware requirements

These options are continually expanding. Please check github for the latest information.

Dell EMC PowerEdge Intel Servers 14G	C4140, C6420, R240, R340, R440, R540, R640, R740, R740xd, R740xd2, R840, R940, R940xa
Dell EMC PowerEdge Intel Servers 15G	C6520, R650, R750, R750xa
Dell EMC PowerEdge Intel Servers 16G	R660, R760, R760xa, R760xd2, C6620, XE8640, XE9640, XE9680
Dell EMC PowerEdge AMD Servers 14G	R6415, R7415, R7425
Dell EMC PowerEdge AMD Servers 15G	C6525, R6515, R6525, R7525, XE8545
Dell EMC PowerEdge AMD Servers 16G	R6625, R7625, R6615, R7615
Nvidia Accelerators	A10, A100, A30, H100, L40, T4
AMD Accelerators	Mi100, Mi200, Mi210, Mi300x
Intel Accelerators	Gaudi 3
Operating system (Control Plane)	RHEL 8.8 Full Rocky 8.8 Full Ubuntu 22.04.5 Server
Operating system deployed by Omnia on bare-metal servers	RHEL 8.8 Minimal Rocky 8.8 Minimal Ubuntu 22.04.5 Server
xCAT	2.16.5
Slurm Workload Manager	20.11.9
Kubernetes Controllers on Control Plane	1.29.5
Kubernetes Controllers on Head Node and Compute	1.29.5
CSI PowerScale Driver	2.11.0
vLLM	v0.2.4 (AMD), v0.4.0 (Nvidia) v0.5.3.post1+Gaudi-1.18 (Gaudi)
KServe	v0.13.0
Jupyterhub	3.2.0
Pytorch	latest (AMD, CPU), 23.12-py3 (Nvidia), 2.4.0 (Gaudi)
TensorFlow	latest (AMD, CPU), 23.12-tf2-py3 (Nvidia)
Kubeflow	v1.9.1
DeepSpeed for PyTorch	v2Beta1 (Gaudi)
Prometheus	2.23
Helm	3.14.2
Grafana	8.3.2
FreeIPA	4.9.11
LDAP ToolBox	2.5.16
Dell EMC PowerVault Storage	PowerVault ME4084, ME4024, ME4012, ME5012, ME5024, ME5084 Storage Arrays
Dell EMC PowerScale Storage	PowerScale F710, F600, H7000, H5600
Dell EMC Networking Switches	PowerSwitch S3048-ON, S5232F-ON, S5232F-ON, S4148T-ON PowerSwitch N3248TE-ON PowerSwitch Z9332F-ON, Z9264F-ON, Z9264F-ON, Z9664F-ON, Z9432F-ON, Z9864F-ON
Nvidia InfiniBand Switches	NVIDIA SB7890-* EDR Infiniband Switch NVIDIA QM8700-* Quantum HDR InfiniBand

Solution overview

Switch NVIDIA QM9700-* NDR Infiniband Switch

Solution overview

Security

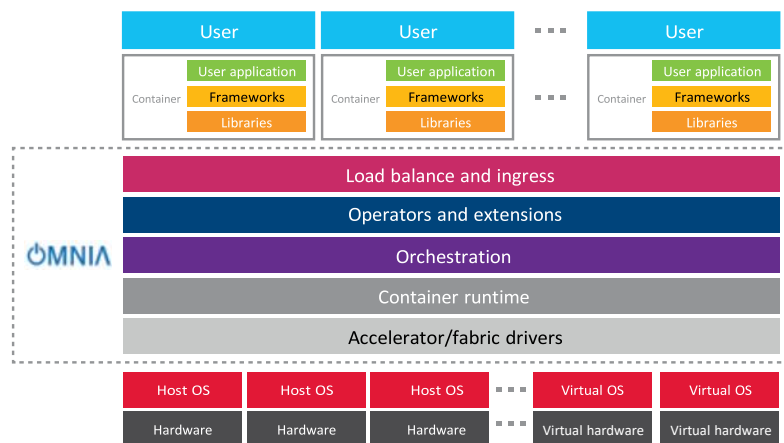
- Omnia adheres to most requirements of NIST 800-53 and NIST 800-171 guidelines on the control plane and login node.
- Omnia supports both FreeIPA, OpenLDAP, external LDAP replication to provide authentication and authorization.
- Email alerts have been added in case of login failures.
- Administrator can restrict users or hosts from accessing the control plane and login node over SSH.
- Malicious or unwanted network software access can be restricted by the administrator.
- Admins can restrict the idle time allowed in an SSH session.
- Security on audit log access is provided.
- Program execution on the control plane and login node is logged using snoopy tool.
- User activity on the control plane and login node is monitored using psacct/account tools installed by Omnia.

“HPC systems deployment can be difficult, and the addition of AI and data analytics makes it even more complicated. We created Omnia to make it easier, incorporating expertise from Dell's HPC & AI Solutions engineers, our HPC & AI Centers of Excellence, and across the HPC Community. And as the HPC landscape changes, whether that be on premises or in the cloud, Omnia will continue to grow and evolve to meet our customers' and the Community's needs.”

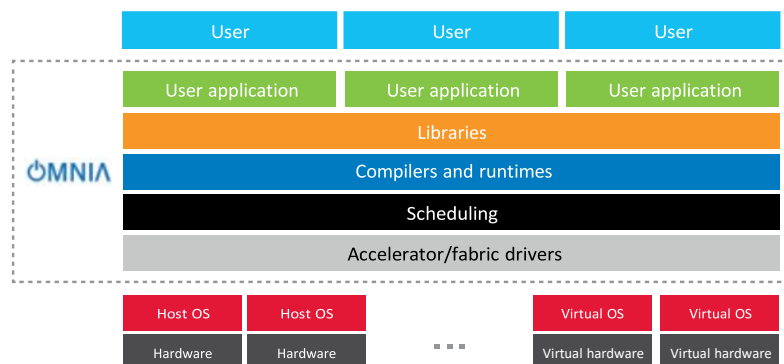
— John Lockman,
Technologist, HPC & AI

Omnia can install Kubernetes, Slurm or both, along with additional drivers, services, libraries, and user applications.

Omnia Kubernetes stack



Omnia Slurm stack



Solution overview

“ We choose Dell because it’s the best in quality and the best in support. I am not joking. We now have around 600 servers in our data center, including different generations from Dell, and we have statistics that show us that Dell is the best in quality and support.”

— Maurizio Davini
Chief Technology Officer
University of Pisa¹

Red Hat® Ansible® Automation Platform

Ansible is a simple-to-use IT automation engine that transforms repetitive, inefficient tasks into predictable, scalable and simple processes. Ansible automation lets developers spend more time on their work and helps operations more easily support deployment pipelines. Together, these capabilities create a quick, comprehensive and coordinated approach to delivering business value. Learn more about [Red Hat Ansible](#)

Services and financing

Dell Technologies is there every step of the way, linking people, processes and technology to accelerate innovation and enable optimal business outcomes.

- [Consulting](#) helps you create a competitive advantage for your business. Our expert consultants work with companies at all stages of analytics to help you plan, implement, and optimize solutions that enable you to unlock your data capital and support advanced techniques, such as AI and ML.
- [Deployment](#) helps you streamline complexity and bring new IT investments online as quickly as possible. Leverage our 30+ years of experience for efficient and reliable solution deployment to accelerate adoption and return on investment while freeing IT staff for more strategic work.
- [Support](#) driven by AI and DL will change the way you think about support with smart, ground-breaking technology backed by experts to help you maximize productivity, uptime and convenience. Experience more than fast problem resolution—our AI engine proactively detects and prevents issues before they impact performance.
- [Payment Solutions](#) from Dell Financial Services help you maximize your IT budget and get the technology you need today. Our portfolio includes traditional leasing and financing options, as well as advanced flexible consumption products.
- [APEX](#) delivers cloud services for a range of data and workload requirements, enabling you to simplify transformation, adapt to evolving conditions, and stay in control of your data. APEX is based on innovative Dell Technologies infrastructure built with flexibility and performance.
- [Managed Services](#) can help reduce the cost, complexity and risk of managing IT so you can focus your resources on digital innovation and transformation while our experts help optimize your IT operations and investment.
- [Residency Services](#) provide the expertise needed to drive effective IT

Why choose Dell Technologies

We’re committed to advancing HPC, AI and data analytics, and we’ve dedicated a great deal of resources toward that goal.

- Schedule an [executive briefing](#) and collaborate on ways to reach your business goals.
- [Dell Technologies Customer Solution Centers](#) are staffed with computer scientists, engineers and subject matter experts in a variety of disciplines.
- We are committed to [providing you with choice](#). We believe in being open, and we publish our performance results at [hpcatdell.com](#) and on the [InfoHub](#). The [Dell HPC Community](#) regularly hosts speakers from academia and industry to give their perspectives on what is coming next.
- Dell Technologies is the only company in the world with a portfolio that spans from workstations to supercomputers, including servers, networking, storage, software and services.
- Because Dell Technologies offers such a wide selection of solutions, we can act as a trusted advisor without trying to sell you a one-size-fits-all approach to your problem. That range of solutions has also given us the expertise to understand a broad spectrum of challenges and ways to address them.

¹Dell Technologies case study, [Capitalizing on virtualization](#), August 2020

Solution overview

“Our partnership with Dell Technologies allows us to take advantage of the full breadth and depth of their compute, storage, networking and security solutions.”

— David J. Brzozowski Jr
Chief Technology Officer
Medacist²

Customer Solution Centers

Our global network of dedicated [Dell Technologies Customer Solution Centers](#) are trusted environments where world-class IT experts collaborate with you to share best practices, facilitate in-depth discussions of effective business strategies and help your business become more successful and competitive. Dell Technologies Customer Solution Centers reduce the risks associated with new technology investments and can help improve speed of implementation.

AI Experience Zones

Curious about AI and what it can do for your business? Run demos, try proofs of concept and pilot software in Singapore, Seoul, Sydney, Bangalore and other Customer Solution Centers. Dell Technologies experts are available to collaborate and share best practices as you can explore the latest technology, get the information and hands-on experience you need for your advanced computing workloads.

HPC & AI Innovation Lab

The [Dell Technologies HPC & AI Innovation Lab](#) in Austin, Texas, is a flagship innovation center. Housed in a 13,000-square-foot data center, it gives you access to thousands of Dell servers, three powerful HPC clusters, and sophisticated storage and network systems. It's staffed by a dedicated group of computer scientists, engineers and subject matter experts who actively partner and collaborate with customers and other members of the HPC community. The team engineers HPC and AI solutions, tests new and emerging technologies, and shares expertise including performance results and best practices.

HPC & AI Centers of Excellence

As data analytics, HPC and AI converge and the technology evolves, Dell Technologies worldwide [HPC & AI Centers of Excellence](#) provide thought leadership, test new technologies and share best practices. They maintain local industry partnerships and have direct access to Dell Technologies and other technology creators to incorporate your feedback and needs into their roadmaps. Through collaboration, Dell Technologies HPC & AI Centers of Excellence provide a network of resources based on the wide-ranging know-how and experience in the community.

Proven results

Dell Technologies holds leadership positions in some of the biggest and largest-growth categories in the IT infrastructure business, and that means customers can confidently source information technology needs from Dell Technologies.

- #1 in servers³
- #1 in converged and hyperconverged infrastructure (HCI)⁴
- #1 in storage⁵
- #1 cloud IT infrastructure⁶

See [Dell Technologies Key Facts](#).

² Dell Technologies case study, [Medacist Advances Healthcare Analytics with AI running on Dell EMC PowerEdge](#) and PowerScale, January 2021.

³ IDC WW Quarterly Server Tracker, Units & Revenue.

⁴ IDC WW Quarterly WW Quarterly Converged Systems Tracker, Vendor Revenue.

⁵ IDC WW Quarterly Enterprise Storage Systems Tracker.

⁶ IDC WW Quarterly Cloud IT Infrastructure Tracker, Vendor Revenue.

Solution overview

Start working with Omnia.

<https://github.com/dell/omnia>

Learn more.

[HPC/AI Engineering](#)

[Dell.com/HPC](#)

[Dell.com/AI](#)

[Dell.com/Analytics](#)

[InfoHub](#)

Join the HPC Community: [DellHPC.org](#)



Take the next step, today.

Don't wait to reap the benefits of an open-source software solution designed to help you deploy faster, leverage fluid pools of resources, and integrate complete lifecycle management for unified data analytics, AI and HPC clusters. Visit [Omnia on Github](#) and contact your Dell Technologies representative to find out more, today.

Contact us.

To learn more, visit [Dell.com/HPC](#) or [contact](#) your local representative or authorized reseller.

DELLTechnologies

Copyright © 2022 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries.

Slurm® is a registered trademark of SchedMD LLC. Kubernetes®, Helm®, Prometheus®, and OpenHPC™ are trademarks of The Linux Foundation. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. Red Hat®, CentOS®, and Ansible® are registered trademarks of Red Hat, Inc. in the United States and other countries. BeeGFS® is a registered trademark of Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. PixStor™ is a trademark of Arcapix Holdings. VMware® products are covered by one or more patents listed at <http://www.vmware.com/go/patents>. VMware® is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. GitHub® is an exclusive trademark registered in the United States by GitHub, Inc. Other trademarks may be the property of their respective owners. Published in the USA 8/22 Solution overview OMNIA-SWSTK-SO-102

Dell Technologies believes the information in this document is accurate as of its publication date. The information is subject to change without notice.