

Assoc. Dir. DDIT IES Cloud Engineering

Job ID
REQ-10059805

12月 10, 2025

India

摘要

Responsible for designing, building, and managing a cutting-edge AI and Generative AI infrastructure based on NVIDIA SuperPOD NV72 system, tailored for pharmaceutical business use cases. The platform will enable Biomedical Research Scientists and other business users to accelerate early molecule development and research activities by providing robust, scalable, and secure GPU computing resources.

About the Role

Major Accountabilities:

- Architect and Design: Lead the design and architecture of an NVIDIA SuperPOD-based AI infrastructure platform supporting Generative AI workloads and advanced analytics for pharma use cases like BioNeMo, AlphaFold, ESMFold, OpenFold, ProtGPT2, and NVIDIA

Clara suite.

- Platform Development: Implement ML/Ops solutions (Run:AI) on Kubernetes clusters optimized for NVIDIA GPUs.
- Data Management: Design and implement high-performance data pipelines for large-scale genomics and chemical compound datasets.
- Security and Compliance: Ensure robust security measures and compliance for HPC and multi-cloud environments.
- Performance Optimization: Optimize GPU cluster performance, networking, and storage for cost-efficiency and scalability.
- Innovation: Stay updated with NVIDIA AI infrastructure advancements and HPC trends.

Technical Expertise:

- Expertise in deploying and managing GBX00 GPU-based clusters.
- 8+ years of experience in GPU-based AI infrastructure and HPC systems.
- Understanding of advanced interconnect technologies for GB-series GPUs.
- Performance tuning for multi-node GBX00 workloads using NCCL, CUDA NVLink, NVSwitch, Storage and Inband High-Speed Ethernet Fabric, RDMA tuning, QoS policies, Out of Band Management.
- Redundant power and cooling systems for HPC reliability.
- Cluster Management: NVIDIA Base Command Manager, Slurm, Kubernetes for GPU scheduling.
- Firmware & Driver Management: CUDA, NCCL, InfiniBand drivers, GPU firmware updates.
- EFA, NVLink and InfiniBand switches for ultra-low latency GPU cluster communication.
- Separate Ethernet-based management network for orchestration and monitoring.
- Parallel File Systems: Spectrum Scale (GPFS) or Lustre for high-performance distributed storage.
- Multi-petabyte capacity with NVMe SSD tiers for scratch space and HDD tiers for archival.
- Integration with object storage for AI datasets.
- Monitoring & Troubleshooting: DCGM, Prometheus, Grafana for telemetry and health checks.
- Security & Compliance: RBAC, encryption, secure multi-tenant configurations.
- AI/ML Workflow optimization, troubleshooting and job scheduling

Why consider Novartis?

Our purpose is to reimagine medicine to improve and extend people's lives and our vision is to become the most valued and trusted medicines company in the world. How can we achieve this? With our people. It is our associates that drive us each day to reach our ambitions. Be a part of this mission and join us!

Learn more here:

<https://www.novartis.com/about/strategy/people-and-culture>

Commitment to Diversity and Inclusion:

Novartis is committed to building an outstanding, inclusive work environment and diverse teams' representative of the patients and communities we serve.

Join our Novartis Network: If this role is not suitable to your experience or career goals but you wish to stay connected to hear more about Novartis and our career opportunities, join the Novartis Network here:

<https://talentnetwork.novartis.com/network>

Why Novartis: Helping people with disease and their families takes more than innovative science. It takes a community of smart, passionate people like you. Collaborating, supporting and inspiring each other. Combining to achieve breakthroughs that change patients' lives. Ready to create a brighter future together? <https://www.novartis.com/about/strategy/people-and-culture>

Benefits and Rewards: Read our handbook to learn about all the ways we'll help you thrive personally and professionally: <https://www.novartis.com/careers/benefits-rewards>

部门

Operations

Business Unit

Information Technology

地点

India

站点

Hyderabad (Office)

Company / Legal Entity

IN10 (FCRS = IN010) Novartis Healthcare Private Limited

Alternative Location 1

Barcelona Gran Vía, Spain

Alternative Location 2

Prague, Czech Republic

Functional Area
Technology Transformation

Job Type
Full time

Employment Type
Regular

Shift Work
No

Job ID
REQ-10059805

Assoc. Dir. DDIT IES Cloud Engineering

[Apply to Job](#)



Job ID
REQ-10059805

Assoc. Dir. DDIT IES Cloud Engineering

[Apply to Job](#)

Source URL:

<https://www.novartis.com.cn/careers/career-search/job/details/req-10059805-assoc-dir-ddit-ies-cloud-engineering>

List of links present in page

1. <https://www.novartis.com/about/strategy/people-and-culture>
2. <https://www.novartis.com/careers/benefits-rewards>
3. <https://novartis.wd3.myworkdayjobs.com/en-US/NovartisCareers/job/Hyderabad-Office/Assoc-Dir-DDIT-IES-Cloud-EngineeringREQ-10059805>
4. <https://novartis.wd3.myworkdayjobs.com/en-US/NovartisCareers/job/Hyderabad-Office/Assoc-Dir-DDIT-IES-Cloud-EngineeringREQ-10059805>