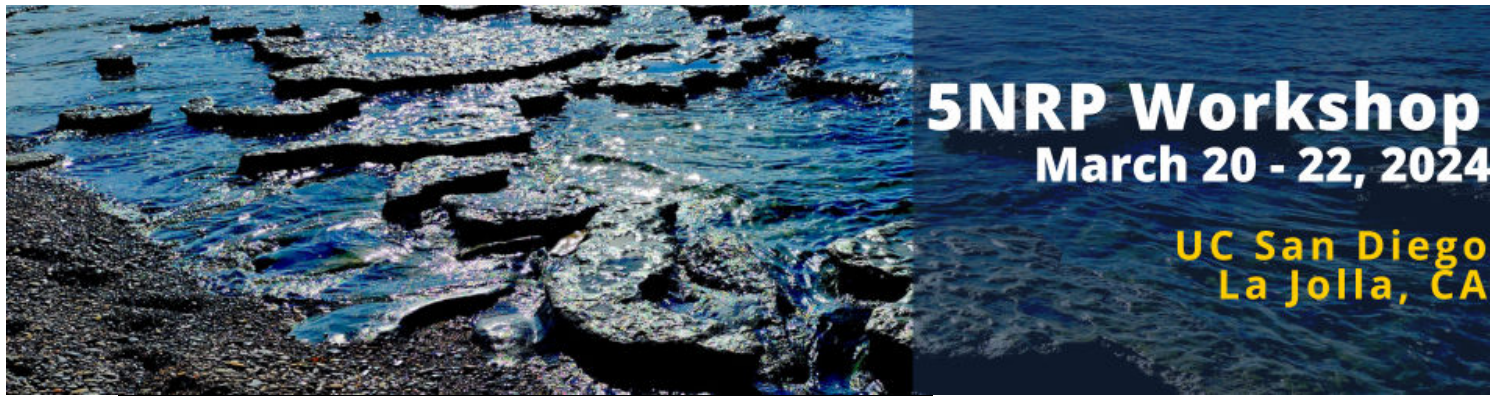




Fifth National Research Platform (5NRP) Workshop March 19 - 22, 2024 Co-located with the FABRIC Community Workshop - KNIT8 Full agenda HERE https://docs.google.com/spreadsheets/u/1/d/1wBgoQuTC8_kQqnHEgZLYGhJIBis_8sbEOCus6nh1bkY/preview?utm_source=hs_email&utm_medium=email&utm_content=2&_hsmt=2	
Live streaming available: Wednesday, 3/20 - https://youtube.com/live/JFrrP5gyjJU?feature=share (https://youtube.com/live/JFrrP5gyjJU?feature=share) Thursday, 3/21 - https://youtube.com/live/5gVlb3QNdMc?feature=share (https://youtube.com/live/5gVlb3QNdMc?feature=share) Friday, 3/22 - https://youtube.com/live/HrS00QFWmx4?feature=share (https://youtube.com/live/HrS00QFWmx4?feature=share)	
TUESDAY, March 19, 2024	
5NRP Tutorial Track NRP Sessions: Yellow FABRIC & NRP Joint Sessions: Red	
Pacific Time	Location: SDSC Conference Room 408
1:00 - 2:00 PM	Kubernetes for Scientists: Examples drawn from AI <i>Tutorial Leaders: Mahdihar Tatini, Dima Mishin, Igor Stiliog, UC San Diego</i> This tutorial will provide a Kubernetes architectural overview, an overview of typical job and workflow submission procedures, and examples provided regarding the various options available to enable optimal use of GPU, CPU, and storage resources for AI use cases. Theoretical information will be paired with hands-on sessions operating on the Prototype National Research Platform (PNRP) production Kubernetes cluster which features a variety of compute and storage resources.
2:00 - 3:00 PM	How to run AI/ML computations on SDSC's Voyager <i>Tutorial Leaders: Paul Rodriguez, Mahidhar Tatini, UC San Diego</i> In this tutorial we will provide information on the Voyager system architecture with details on the Habana processing units (HPUs), provide information on containerized software stacks, file systems, examples using Kubernetes, and overall guidelines. Also, we will discuss options and considerations for scaling training across multiple processors and multiple nodes, including brief introduction to parallelization options, like DeepSpeed and other useful tools.
3:00 - 4:00 PM	How to run AI/ML computations on SDSC's Expanse <i>Tutorial Leaders: Paul Rodriguez, Mahidhar Tatini, UC San Diego</i> In this tutorial we will provide information on the Expanse system architecture, with details on the available GPU resources and scheduling, using containerized and conda based software stacks, and examples of batch and interactive use of TensorFlow, PyTorch. Attendees will be able to use the Expanse portal to run Jupyter notebooks with AI/ML examples. Additional information will be provided on running multi-node TensorFlow workloads on Expanse.
4:00 - 5:00 PM	Writing a successful CC* Grant <i>Tutorial Lead: Jerry Sheehan, Salk Institute and Jason Zurawski, ESnet</i> This tutorial will explore best practices for developing NSF Campus Cyberinfrastructure proposals. Topics addressed will include the 2024 CC* Solicitation, Identifying science drivers, Best Practices for developing your proposal technology plan, Strategies to meet your 20% Resource Sharing requirements, and Tips to tell your Broader Impact story. At the conclusion of the tutorial, NSF Program Director from the Office of Advanced Cyberinfrastructure Kevin Thompson, will participate in a question-and-answer session with participants.
5:00 - 7:00 PM	FABRIC Demo/Poster Night Location: SDSC Auditorium
WEDNESDAY, March 20, 2024	
NRP Track NRP Sessions: Yellow FABRIC & NRP Joint Sessions: Red	
Pacific Time	Location: Qualcomm Institute
8:00 - 9:00 AM	Breakfast Location: QI Courtyard
9:00 - 9:45 AM	[Keynote] 5NRP / NAIRR Pilot Location: QI Atkinson Auditorium <i>Dr. Katie Antypas, NSF OAC</i>
9:45 - 10:15 AM	Q&A with Dr. Antypas
10:15 - 10:30 AM	Break
10:30 - 11:00 AM	NRP, from AI to Networking Research to Education Location: QI Atkinson Auditorium <i>Frank Wuerthwein, UC San Diego</i> We will present the vision and status of NRP with particular focus on how NRP addresses three challenges. - The gap between those who have and those who can't afford is getting wider. - Education increasingly requires significant cyberinfrastructure - The end of Moore's Law is leading to a proliferation of "architectures", increasing the complexity of computing to the point where domain science adoption of innovations is at risk We will look at these challenges with AI as a lense. Thought these challenges are far more general.
Using the NRP for Large Scale AI/ML Research Computing Part 1:	

11:00 - 11:30 AM	Bringing AI to wireless sensing, communication, networking with NRP <i>Dinesh Bharadia, UC San Diego</i> Wireless radios can sense and communicate robustly in non-line of sight and inclement weather, enabling the entire era of mobile and AR/VR computing and interactions. This talk will present extensive data sets, deep learning models, and digital twins developed to enable wireless radios to provide sensing for autonomous cars, robotics, AR/VR sensing and content delivery, and 6G. In the era of generative AI, limited to only meaningfully processing large-scale text, image, and audio data due to abundant datasets, it can't penetrate these new forms of sensing; the lab's broader goal is to enable that in the next decade.
11:30 - 12:00 PM	Knowledge Graph-based Industrial AI for Future Manufacturing <i>Bingbing Li, CSU Northridge</i> The research project advances industrial AI research in multi-domain knowledge representation learning, differentiable reasoning, and promote use-inspired AI research to accelerate AI-powered innovation and best practices for future manufacturing. The industrial AI research will build upon Multi-domain Knowledge Graph (KG) under four thrusts: 1) KG initialization with supervised initial KG construction with documents, data, and models; 2) model and KG learning such as self-supervised learning from domain data; 3) KG reasoning and learning such as combining symbolic reasoning with representation learning for KG reasoning; and 4) KG validation such as planning, data acquisition, and model validation.
12:00 - 1:30 PM	Lunch <i>Location: QI Courtyard</i>
Pacific Time	Location: Qualcomm Institute
Using the NRP for Large Scale Computing - Part 2	
1:30-2:00 PM	Learning Humanoid Robots <i>Xiaolong Wang, UC San Diego</i> Having a humanoid robot operating like a human has been a long-standing goal in robotics. The applications go much beyond moving boxes in the factory. The humanoid robot provides a generalized purpose platform to conduct diverse tasks we do in our daily lives. In this talk, we study learning-based approaches for both the mobility and manipulation skills of the humanoid robot, with the goal of generalization to diverse tasks, objects and scenes. With Reinforcement Learning and Sim2real transfer, we will first discuss how to perform whole-body control in humanoids to generate rich, diverse, and expressive motions. We then discuss what will be the foundation model for the robot to perceive the visual world and perform mobile manipulation in the wild. Finally, we will dive into the humanoid's manipulator, the hands, and study tactile sensing for flexible manipulations.
2:00 - 2:30PM	Machine Learning on Secure Human Time Series <i>Benjamin Smarr, UC San Diego</i> The NRP provides extensible resources without the associated costs of commercial platforms, suggesting a role for the NRP in enabling health algorithms run in at risk communities without the need to charge participants. The NRP is not designed to handle sensitive medical information however. During the pandemic across 2020 and 2021 we used the NRP to host daily covid screening algorithms for health care front line workers using wearable device data without compromising security. Here we discuss hashing, data segregation, and collaborative efforts that made this possible within the NRP.
2:30 - 3:00 PM	Neural Radiance Fields for View Synthesis <i>Ravi Ramamoorthi, UC San Diego</i> Many applications like virtual reality, e-commerce, and three-dimensional immersive digital photography require the ability to take a few images of a scene of interest, and be able to walk around the object or scene, synthesizing views from new directions. This problem is known as view synthesis in computer vision and graphics, and a successful solution promises to revolutionize image capture, from 2D flat images to immersive 3D experiences. Historically, methods focused on explicit or implicit 3D reconstruction which was prone to errors. We have introduced a pioneering new approach in terms of volumetric neural radiance fields (NeRFs) to represent the 3D information in the scene, leading to unparalleled levels of realism; moreover the NeRF representation can be a core 3D component in a wide variety of applications. This paper briefly discusses neural radiance fields and their subsequent impact and newer developments for view synthesis and related problems. Solving for a NeRF involves large-scale optimization, and in some cases pre-training across multiple scenes for generalization, which benefits greatly from the computational support provided by the national research partnership, and this computational infrastructure has been critical to many of our newest advances.
3:00 - 3:30 PM	Break
3:30 - 4:00 PM	National Data Platform (NDP) as a Research Resource for AI <i>Ilkay Altintas, UC San Diego</i> Open and equitable access to scientific data is essential to addressing important scientific and societal grand challenges, and to research enterprise more broadly. This session discusses the importance and urgency of open and equitable data access, and introduces the vision and architecture of the National Data Platform, a recently launched project aimed at catalyzing an open, equitable and extensible data and service ecosystem for AI workflows.
4:00 - 4:30 PM	AI for the University as an Enterprise <i>Vince Kellen, UC San Diego CIO</i> Vince Kellen, Chief Information Officer (CIO) of UC San Diego, will explore the transformative impact of artificial intelligence (AI) on higher education institutions. Vince shares insights into how UC San Diego is leveraging AI as an enterprise to enhance various aspects of university operations, academic research, and student experiences.
4:30 - 5:00 PM	NDP LLM-as-a-Service on NRP Tutorial <i>Ilkay Altintas, UC San Diego</i> This tutorial offers an introduction to utilizing open Large Language Models (LLMs) as a service through the National Data Platform (NDP), focusing on their integration into various domain-specific application workflows. Participants will learn how to access and deploy LLM services through practical notebooks on NDP with hands-on examples running on NRP, navigating through setup, customization, and interaction steps. The technical depth of this tutorial session is balanced toward practical applications and accessibility to a broad audience to take advantage of models available on HuggingFace.
5:00 - 7:00 PM	FABRIC/NRP Joint Social <i>Location: QI Atkinson Auditorium</i>
THURSDAY, March 21, 2024	
NRP Track	
NRP Sessions: Yellow FABRIC & NRP Joint Sessions: Red	
Pacific Time	Location: Qualcomm Institute
8:00 - 9:00 AM	Breakfast <i>Location: QI Courtyard</i>
9:00 - 9:45 AM	FABRIC Keynote <i>Location: QI Atkinson Auditorium</i> <i>kc claffy, UC San Diego</i> The goal of the GMI project (Designing a Global Measurement Infrastructure to Improve Internet Security, or GMI3S) is to design a new generation of measurement infrastructure for the Internet, which will support collection, curation, archiving, and expanded sharing of data needed to advance critical scientific research on the security, stability, and resilience of Internet infrastructure.
9:45 - 9:50 AM	Q&A with kc claffy

9:50 - 10:30 AM	FABRIC Keynote: In Search of a Networking Unicorn: Realizing Closed-Loop ML Pipeline for Networking Location: QI Atkinson Auditorium <i>Arpit Gupta, UC Santa Barbara</i> AI and ML promise to enhance networking solutions but face skepticism regarding their reliability in critical network decision-making due to issues like inadequate training data and a focus on narrow performance metrics. This talk will present a novel closed-loop ML workflow that emphasizes iterative training data collection and model analysis to improve generalization, featuring tools such as Trustee for explainability, netUnicorn for data gathering, and PINOT for data infrastructure, aiming to foster community-wide support for creating production-ready ML artifacts for networking.
10:30 - 10:35 AM	Q&A with Arpit Gupta
10:35 - 10:45 AM	Break
10:45 - 11:45 AM	Keynote Demo - FPGA Location: QI Atkinson Auditorium <i>Nik Sultana, Illinois Institute of Technology</i> The FABRIC testbed provides its users with access to AMD-Xilinx Alveo cards -- programmable, FPGA-based network cards that can support research experiments at scale. This talk+demo-tutorial combo presentation will provide example applications of these cards on FABRIC and provide you with a step-by-step guide to start using them.
11:45 - 12:00 PM	Q&A with Nik Sultana
12:00 - 1:30 PM	Lunch Location: QI Courtyard
12:20 - 1:20 PM	Lunch/National Research Platform International Extensions, Including the Global Research Platform Location: VROOM, Atkinson Hall (next to the Auditorium) <i>Bring lunches from Lobby</i> Agenda A) Introduction to Session (Joe Mambretti) (12:15 – 12:25, 10 min) B) NSF International Research Connections Program (Kevin Thompson) (12:25-12:35, 10 min) C) AutoGOLE/SENSE (Tom Lehman) (12:35-12:45, 10 min) D) Asia Pacific Research Platform (Jeonghoon Moon) (12:45-12:55, 10 min) E) Ampath/AmLight (Julio Ibarra) (12:55-1:05, 10 min) F) CENI International Testbed (Gauravdeep Shami) (1:05-1:15) (10 min) G) GNA-BIS Overview (Harvey Newman) (1:15-1:25, 10 min)
Using the NRP for Large-Scale AI/ML applied to Disciplinary Sciences	
1:30 - 2:00 PM	AI/ML Computing for the Laser Interferometric Gravitational Wave Observatory (LIGO) <i>Phil Harris, MIT</i> Deep learning algorithms have excelled in various domains. Despite this success, few deep-learning models have seen full end-to-end deployment in gravitational-wave searches, both in real-time and on archival data. In particular, there is a lack of standardized software tools for quick implementation and development of novel AI ideas. We address this gap by developing the ML4GW and HERMES libraries. We show how these libraries enhance efficiency and AI model robustness in the context of a broad range of gravitational wave analyses with an emphasis on real-time application, scalability to heterogeneous computing resources, and streamlining the training to deployment cycles for machine learning models. We present this work within the larger effort toward heterogeneous integration of large scale physics experiments and discuss where this work is going. Out of this scope, we present a possible model of next-generation computing for scientific experiments.
2:00 - 2:30 PM	SmokeyNet for Wildfire Smoke Detection <i>Mai H. Nguyen, SDSC</i> Early detection of wildfires is essential to minimize any damage and potential catastrophic losses. We present our work on SmokeyNet, a deep learning approach for detecting smoke plumes from wildfires. Smokeynet can integrate several data sources and can be used for early notification of wildfires, potentially reducing the time to wildfire response.
2:30 - 3:00 PM	AI/ML for Genomics <i>Aman Patel, Stanford University</i> "Using compute resources like those provided by the NRP, we have trained an atlas of thousands of deep learning models, each of which is designed to learn the DNA syntax underpinning gene regulation in a particular cell type. A key application of our trained models involves predicting the effects of genetic changes - for example, from individuals with genetic diseases or ancient humans living tens of thousands of years ago - in altering gene regulation, thus furthering our knowledge of a variety of relevant diseases and systems. Overall, through deep learning and high-performance computing, we are able to drastically enhance our understanding of biology, disease, and fundamental human genetics."
3:00 - 3:30 PM	Break
3:30 - 4:00 PM	AI/ML Computing Across Diverse Disciplines <i>Larry Smarr, UC San Diego</i> NRP storage and computing resources are being used to investigate how AI/ML algorithms can be applied to a wide range of disciplines. I will give brief examples drawn from the research projects of a number of large NRP users from chemistry, physics, healthcare, drug discovery, robotics and manufacturing.
4:00 - 5:00PM	Expanding the AI/ML Workforce Through Coursework <i>Michael Farley, SDSU, Valerie Polichar, UC San Diego</i> Michael Farley, San Diego State's Chief Technology Research Officer, will share how SDSU has leveraged the NRPs "bring-your-own-resource" model to support instruction with their VERNE cluster. VERNE provides faculty an easy to use resource to teach a variety of courses that require access to advanced cyberinfrastructure. Dr. Polichar will share how UC San Diego created a direct line from research to instruction by embracing vision and technology developed by elite researchers to build UCSD's Data Science/Machine Learning Platform (DS/MLP) for graduate and undergraduate instruction. Now a showcase for modern technology education, the DS/MLP is used by thousands of students every quarter for classes in every division, from Electrical Engineering to Visual Art to Public Health, instilling market-competitive skills in students across a broad range of disciplines.
FRIDAY, March 22, 2024	
NRP Track	
Expanding NRP Usage on Campuses through RONS	
Pacific Time	Location: Qualcomm Institute
7:00 - 8:00 AM	Breakfast Location: QI Courtyard
8:00 - 8:30 AM	The Quilt <i>Jen Leasure, The Quilt</i>
8:30 - 9:00 AM	Great Plains Network <i>Alex Hurt, UMissouri and Derek Weitzel, UNL - GP-Engine</i> This talk will present the latest research and teaching activities being performed on the National Research Platform in the Great Plains Network, which encompasses 8 universities in 6 great-plains states. This presentation will also include the activities and outreach performed as part of the GP-ENGINE project, which has expanded NRP usage throughout the GPN and added several advanced, high-performance GPU nodes as well as facilitated the addition of the first Grace Hopper Super-Chip to the National Research Platform.

9:00 - 9:30 AM	The Technology Infrastructure for Data Exploration (TIDE) and the RON as an Enabling Partner for Digital Equity Chris Bruton, CENIC, Jerry Sheehan, Salk/SDSU Join us for an insightful conference session featuring an adjunct faculty member from SDSU's College of Sciences, a distinguished individual who previously held the position of vice president of information technology and chief information officer. Discover the pivotal role regional optical networks play in tackling equity issues within cyberinfrastructure. Gain valuable insights as they share lessons learned from CSU, shedding light on challenges and research data requirements in this dynamic field.
9:30 - 10:00 AM	StarLight/MREN <i>Joe Manbretti, Northwestern</i> This presentation will describe how RONs can provide services to support large scale computational science communities for data intensive research.
10:00 - 10:30 AM	Break
10:30 - 11:30 AM	Expanding NRP Usage on Your Campus Using Jupyter Notebooks <i>Dung Vu, Youngsu Kim, CSUSB</i> Being a typical R2 teaching institution and a Minority Serving Institution, CSUSB did not have cyber infrastructure to support research. NRP has not only provided computing resources, hands-on expertise, but also technical support and solutions. As the result, using Jupyter hubs on NRP's Kubernetes platform, CSUSB faculty and students can access high performance research computing at their fingertips. In this presentation, we outline the history of how CSUSB has leveraged NRP's computing resources to advance research computing in our campus, emphasizing the integral role of the National Research Platform (NRP). We also demonstrate a few examples of the effective use of Jupyter Hub and Notebooks in supporting our HPC users.
11:30 - 12:00 PM	Toward NAIRR <i>Frank Wuerthwein, UC San Diego</i>
12:00 PM	Adjourn