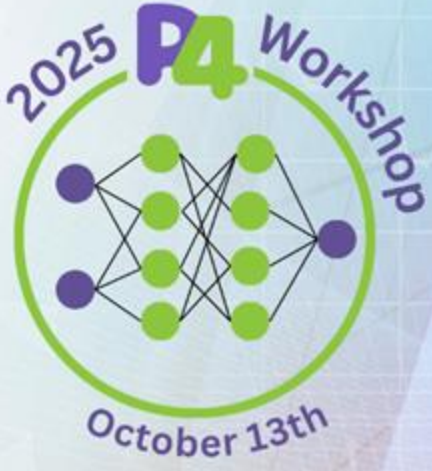


Cross-Federated P4 Research Testbed for Wide-Area Programmable Networking Experiments

Mohammad Firas Sada
University of California, San Diego
San Diego Supercomputer Center

SAN DIEGO SUPERCOMPUTER CENTER

NRP NATIONAL RESEARCH PLATFORM

THE LINUX FOUNDATION

National Research Platform (NRP)

- The **National Research Platform (NRP)** is a distributed computing infrastructure designed to support advanced research across multiple disciplines.
- **Advanced Hardware**: GPUs, FPGAs, SmartNICs…
- **Research Focus**: AI/ML, networking, scientific computing, and data-intensive research
- A partnership of more than **80** institutions.
- Led by researchers at **UC San Diego, University of Nebraska-Lincoln, and MGHPCC.**

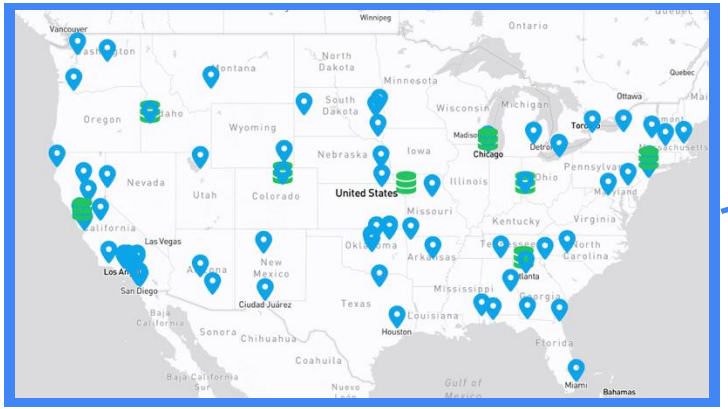
NRP NATIONAL RESEARCH PLATFORM

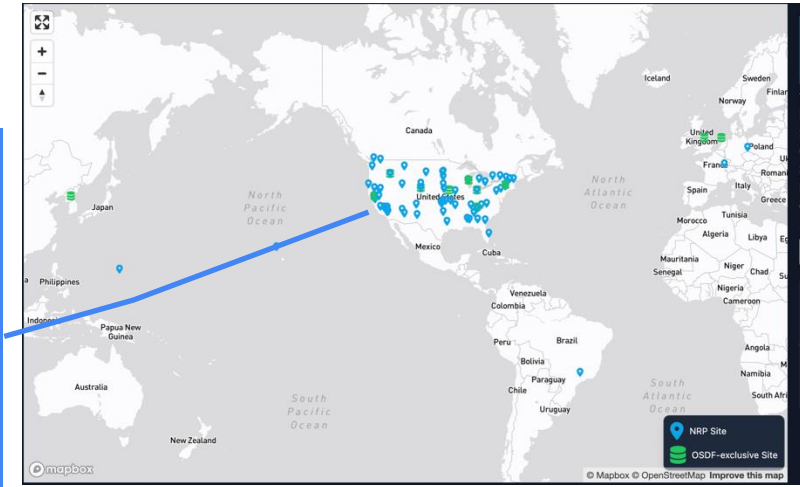
SAN DIEGO SUPERCOMPUTER CENTER

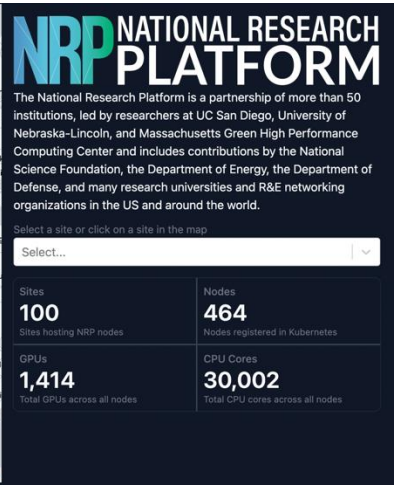
THE LINUX FOUNDATION

National Research Platform (NRP)

- **Open Access**: Available to researchers nationwide.
- **Access Inequality:** Widening gap between institutions with and without CI resources.
- **Architectural Fragmentation:** Post-Moore's Law era complicates domain science adoption.
- NRP operates a primary **Kubernetes** distribution cluster with hardware spanning across **4 continents**, called **Nautilus**







https://dash.nrp-nautilus.io/

NRP operates a primary **Kubernetes** distribution cluster with hardware spanning across **4 continents**, called **Nautilus**

P4

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

Pacific Wave

- High-performance **Internet Exchange** connecting U.S. and international R&E institutions for access to large-scale data and instruments.
- **IGROK Nodes Deployment: Seven** 1U servers with **BlueField-2** DPUs (dual 100GbE), enabling packet processing, AI, and HPC applications.
- **CENIC** made the nodes available on **Nautilus** for enhanced global collaboration and real-time data processing.

The **FABRIC Testbed**

- **FABRIC** = **F**ABRIC is **A**daptive Programma**B**le **R**esearch **I**nfrastructure for **C**omputer Science and Science Applications
- International infrastructure for research at scale
- Supports networking, cybersecurity, HPC, VR, ML, 5G, IoT, and more.
- 29 U.S. sites + 4 global sites (Asia, Europe, South America)
- High-speed, programmable network & compute environment
- https://whatisfabric.net

P4 NRP NATIONAL RESEARCH PLATFORM SAN DIEGO SUPERCOMPUTER CENTER THE LINUX FOUNDATION

Ampath

- A high-performance international research **connection point** facilitating **U.S.-Latin America & Caribbean R&E collaboration.**
- Developed by Florida International University (FIU),
- NRP & AMPATH: Deployment of SN1000 Xilinx Alveo FPGA on AMPATH, enabling geographically extended **FPGA experiments** between the U.S., Latin America, and the Caribbean.
- https://ampath.net/

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

The AutoGole/SENSE Topology

ESnet SENSE

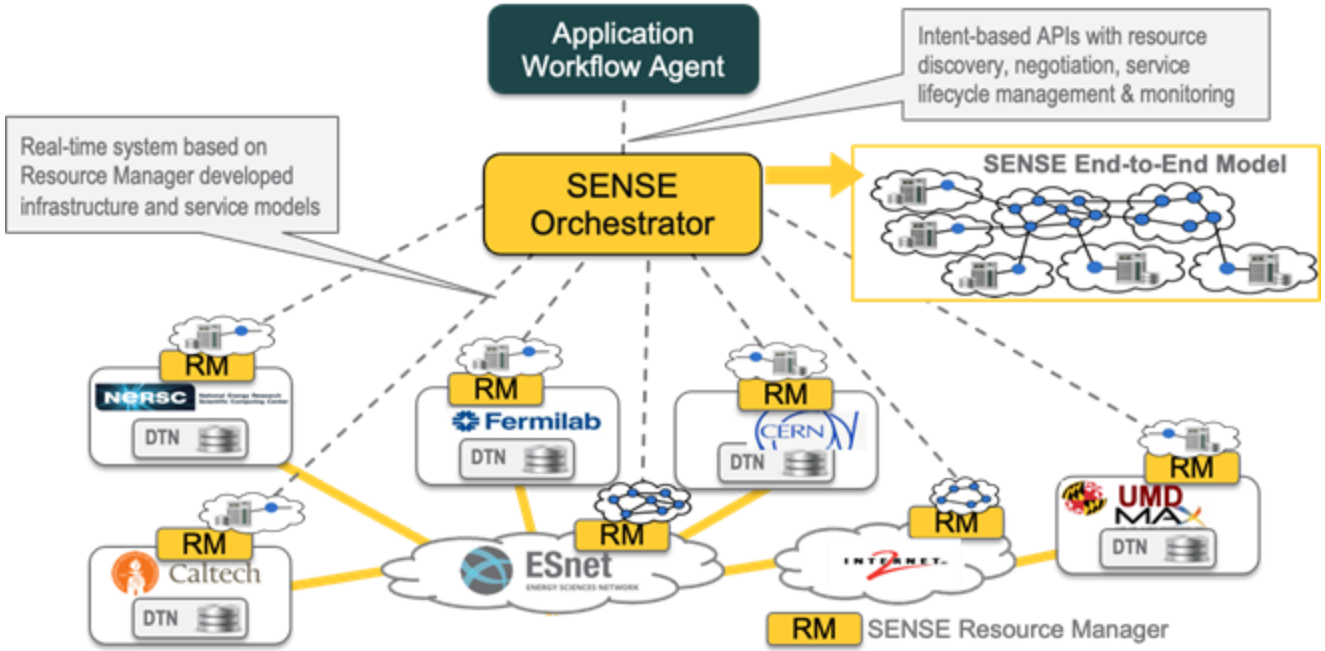
- ESnet (**Energy Sciences Network**) is a **high-performance, high-capacity computer network** funded by the U.S. Department of Energy's (DOE) Office of Science.
- Software-Defined Network for End-to-End Networked Science at Exascale
- Network management system designed to provide tailored and reliable network services for scientific applications.
- Offers scalable approach to network management.
- Enables customized end-to-end service provisioning (L2 and L3) across complex and distributed infrastructures.

NRP NATIONAL RESEARCH PLATFORM

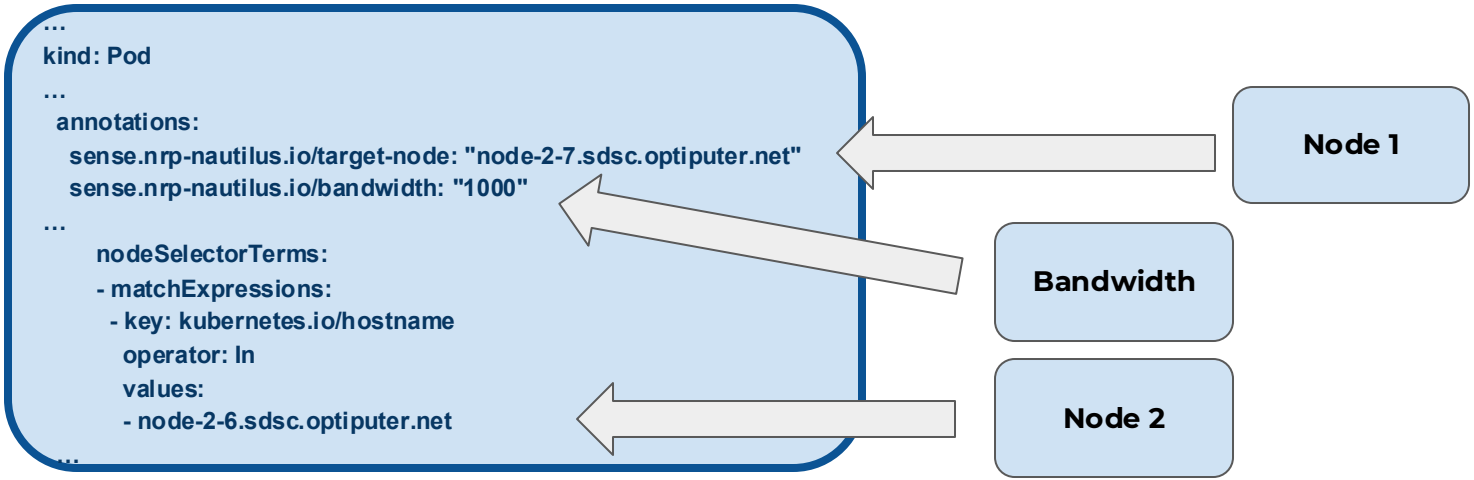
SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

ESnet SENSE


Application Workflow Agent
Intent-based APIs with resource discovery, negotiation, service lifecycle management & monitoring
Real-time system based on Resource Manager developed infrastructure and service models
SENSE Orchestrator
SENSE End-to-End Model
RM
NERSC
DTN
RM
Fermilab
DTN
RM
CERN
DTN
RM
UMD
MAX
DTN
RM
Caltech
DTN
RM
ESnet
RM
INTERNET
RM
SENSE Resource Manager

The K8s SENSE Operator



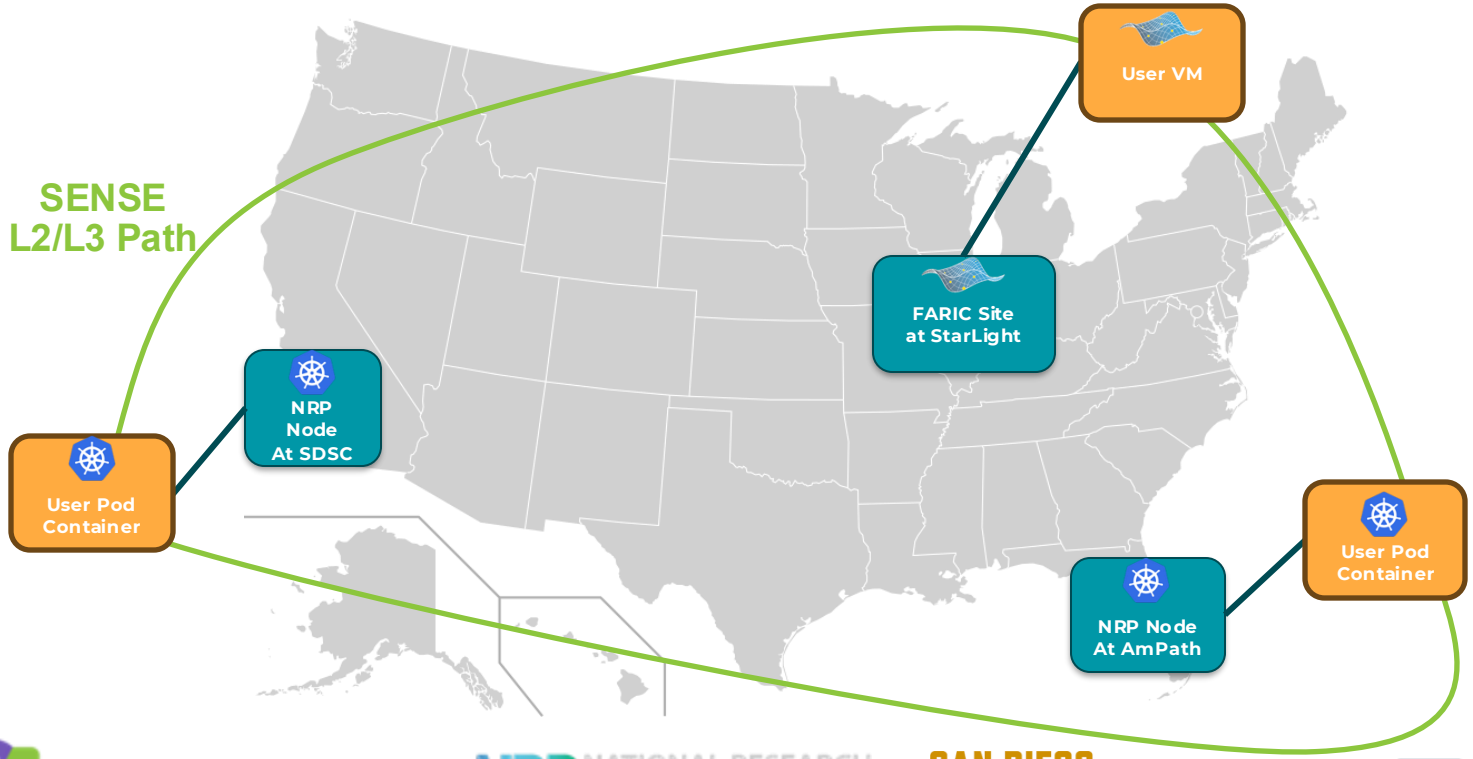
Start k8s SensePath CRD → Start SENSE Path → Update Status in SensePath → Create Multus NADs → Patch annotated pods → Delete path in 7 days unless extended

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

The AutoGole/SENSE Topology



NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

Hardware on NRP

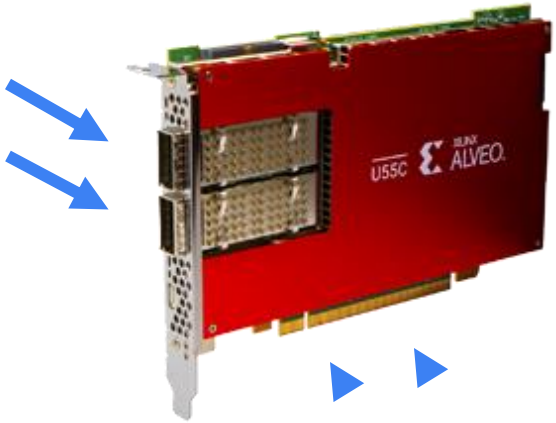
- **ARM + x86_64 Nodes**
- **~1400 GPUs with various models**
  - Good for training most of the ML models
  - 4-8 GPUs / node: can request all of them
  - 550 GB RAM

~600 RTX/GTX “Gaming” GPUs

- **8 Qualcomm Cloud AI 100 Ultra Inference & Fine-tuning Cards:**
  - 1 card = 4 SoCs (System-on-Chip)
  - Each SoC capable of running 25B Param LLM with out-of-the-box config
  - Provisioned via k8s device plugin

P4 Hardware on NRP

- BlueField-2 DPUs:
  - 7 DPUs at Different sites: San Diego, Guam, Chicago..
  - 2x100Gbps: ARM CPU + ConnectX-6 Dx
  - Native **P4* / DOCA support + DPDK**
- Alveo U55C SmartNICs*:
  - 32 FPGAs
  - 2x100Gbps: FPGA + 16GB HBM Memory
  - Uses ESnet SmartNIC framework (AMD OpenNIC Shell P4) **+ DPDK**
  - At SDSC
- Tofino2 Switches



P4 Hardware on FABRIC

- Alveo U280 SmartNICs*:
  - 28 FPGAs
  - 2x100Gbps: FPGA + 2x 16 DDR4 Memory
  - Uses ESnet SmartNIC framework (AMD OpenNIC Shell P4) **+ DPDK**
  - 1 Card at each site
- Dedicated Tofino2 Switches
- Bluefield3 DPUs **+ DPDK**

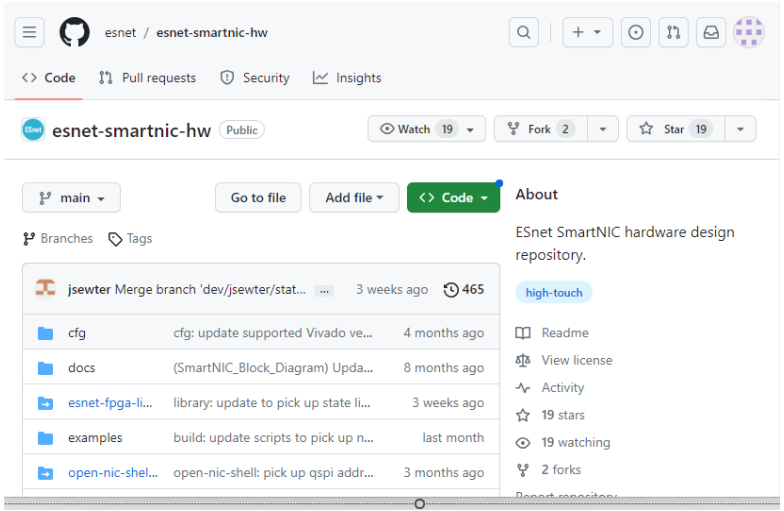
NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

ESnet P4 SmartNIC

- ESnet SmartNIC framework provides an entire workflow for AMD/Xilinx Alveo FPGAs.
- It is open-source (on github).
- It seamlessly integrates Xilinx tools along with various tools like DPDK to provide an easy way of programming Alveo FPGAs as SmartNICs.
- Various debugging, testing and simulating tools.
- Containerized environment that makes it as easy as plug-and-play for P4 on FPGAs.

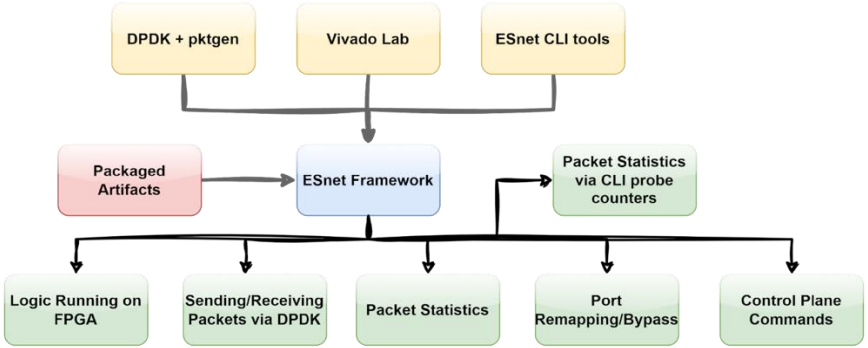


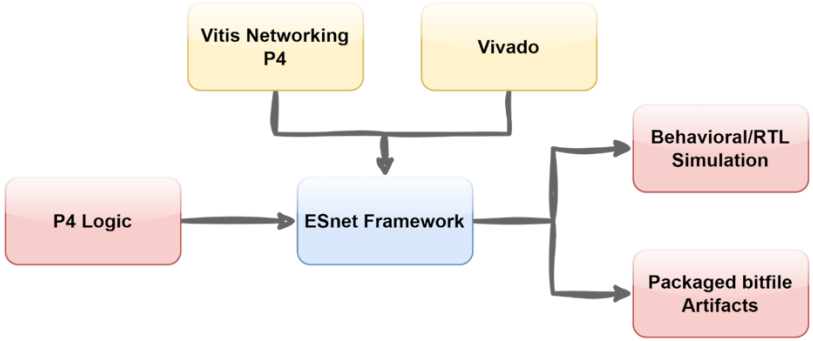
NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

ESnet P4 SmartNIC





https://docs.nrp.ai/userdocs/fpgas/esnet/

Starting a P4 Testbed

```
-> % ./create_multus_nad.sh
=====================================
🚀 Multus NAD Creation Script
=====================================

📋 Configuration:
  • Namespace: mfsada
  • NAD Name: macvlan-p4-vlan3111
  • VLAN ID: 3111
  • IP Range: 192.168.100.0/24
  • Gateway: 192.168.100.1
  • IP Start: 192.168.100.20
  • IP End: 192.168.100.50

🔍 Checking kubectl availability...
✅ kubectl command found

🔍 Checking Multus CNI installation...
✅ Multus CNI found

🔧 Creating namespace mfsada...
namespace/mfsada unchanged
✅ Namespace mfsada ready

🔧 Creating NetworkAttachmentDefinition YAML...
✅ NetworkAttachmentDefinition YAML created: macvlan-nad.yaml

📄 YAML Content:
==================
apiVersion: "k8s.cni.cncf.io/v1"
kind: NetworkAttachmentDefinition
metadata:
  name: macvlan-p4-vlan3111
  namespace: mfsada
  annotations:
    k8s.v1.cni.cncf.io/resourceName: macvlan-p4
spec:
  config: |
    {
      "cniVersion": "0.3.1",
      "name": "macvlan-p4-vlan3111",
      "type": "macvlan",
      "master": "eth0",
      "mode": "bridge",
      "vlan": 3111,
      "ipam": {
        "type": "host-local",
        "subnet": "192.168.100.0/24",
        "gateway": "192.168.100.1",
        "rangeStart": "192.168.100.20",
        "rangeEnd": "192.168.100.50"
      }
    }
==================

Do you want to apply this NetworkAttachmentDefinition? (y/n):
```

```
-> % ./generate_pod_yaml.sh
=====================================
🚀 Generate Pod YAML Script
=====================================

📋 Configuration:
  • Namespace: mfsada
  • NAD Name: macvlan-p4-vlan3111
  • Node 1: node-2-6.sdsc.optiputer.net
  • Node 2: node-2-7.sdsc.optiputer.net
  • Node 3: node-2-8.sdsc.optiputer.net

🔧 Generating pod YAML files in pod_yaml/ directory...

📄 Creating P4 Switch Pod (Node 1) YAML...
✅ p4-switch-1.yaml created

📄 Creating P4 Switch Pod (Node 2) YAML...
✅ p4-switch-2.yaml created

📄 Creating P4 Controller Pod YAML...
✅ p4-controller.yaml created

📄 Creating Traffic Generator Pod YAML...
✅ traffic-generator.yaml created

📄 Creating P4 Monitor Pod YAML...
✅ p4-monitor.yaml created

📄 Creating P4 Test Client Pod YAML...
✅ p4-test-client.yaml created

📄 Creating P4 Test Server Pod YAML...
✅ p4-test-server.yaml created

📄 Creating Apply All Pods Script...
✅ apply_all_pods.sh created

📄 Creating Delete All Pods Script...
✅ delete_all_pods.sh created

📄 Creating README...
✅ README.md created

✅ All pod YAML files generated successfully!

📁 Generated files in pod_yaml/ directory:
  • p4-switch-1.yaml - P4 Switch Pod 1
  • p4-switch-2.yaml - P4 Switch Pod 2
  • p4-controller.yaml - P4 Controller Pod
  • traffic-generator.yaml - Traffic Generator Pod
  • p4-monitor.yaml - P4 Monitor Pod
  • p4-test-client.yaml - P4 Test Client Pod
  • p4-test-server.yaml - P4 Test Server Pod
  • apply_all_pods.sh - Script to apply all pods
  • delete_all_pods.sh - Script to delete all pods
  • README.md - Documentation and usage
```

Per-Packet Telemetry

Every packet is stripped of its payload and encapsulated with a 64-bit FPGA timestamp, and sent to “workers”.

Based on ESnet’s High Touch model:

- Packet truncation at line rate from Tofino switches
- FPGA accelerated data reduction that can process up to 300Mpps
- FPGA 1ns accuracy time stamping
- Kafka based 24/7 streaming central database
- PCAP capture of any subset of flows

Some use cases:

- Latency Analysis: End-to-end path latency measured via FPGA timestamps (e.g., PacWave-to-FABRIC links).
- Traffic Steering: SRv6 segments route telemetry packets to collector nodes (e.g., CIENA site at exhibitions).
- Trend Analysis and Anomaly Investigation

FPGA Telemetry with SRv6 Encapsulation

Segment Routing (SR):

A network routing technique that uses segments (predefined paths) to guide packets through the network, reducing the need for complex stateful routing tables.

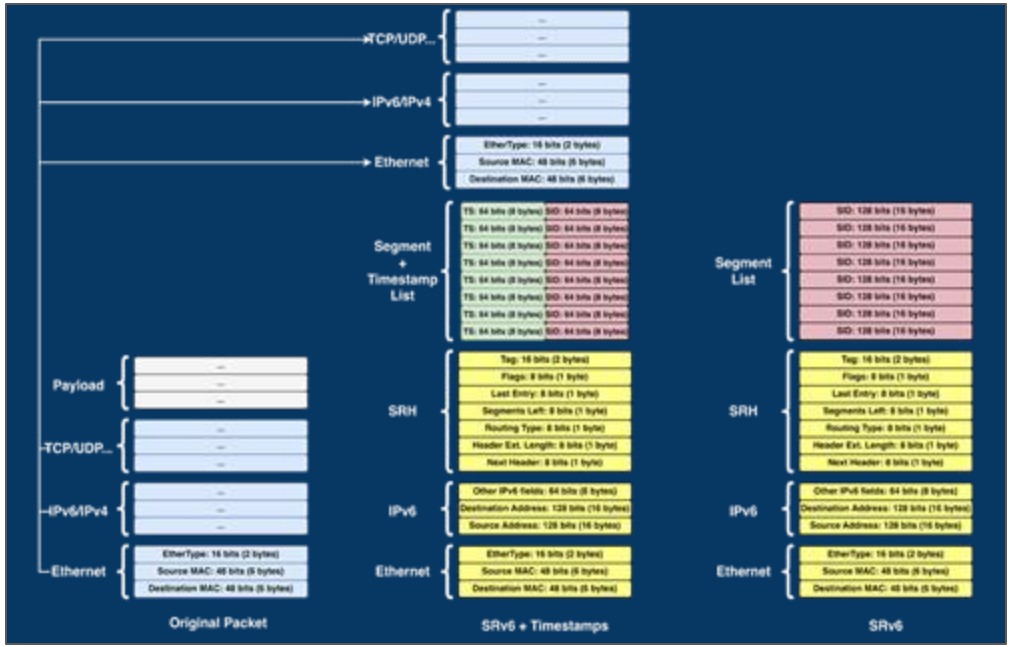
SRv6 (Segment Routing IPv6):

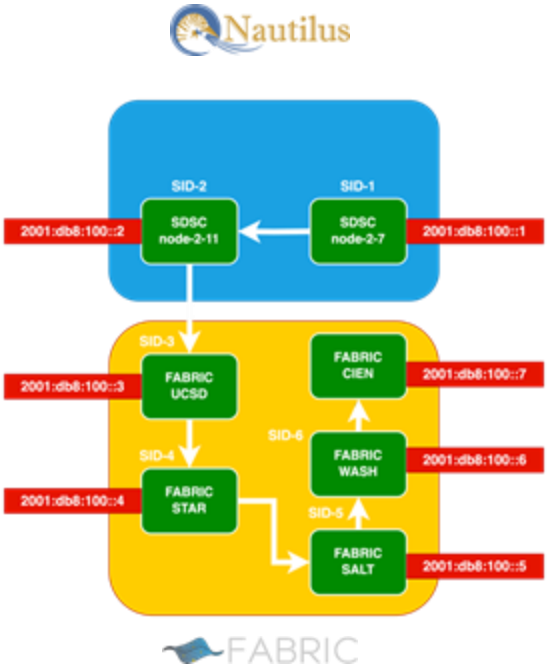
An extension of SR using IPv6 addresses to define routing paths, allowing for more scalable and flexible traffic engineering in modern networks.

Modified SRv6:

- Encapsulated SRv6 Packet
- Segment List: 64-bit Segment ID + 64-bit FPGA timestamp.
- Path: Timestamps preserved end-to-end for latency analysis.
- Original Packet: Retained: Ethernet + IP + Transport headers (payload discarded).
- Payload: Original stripped headers (no application data).

FPGA Telemetry with SRv6 Encapsulation





Real-Time In-Network Machine Learning

“Real-Time In-Network Machine Learning on P4-Programmable FPGA SmartNICs with Fixed-Point Arithmetic and Taylor” (arXiv:2507.00428)

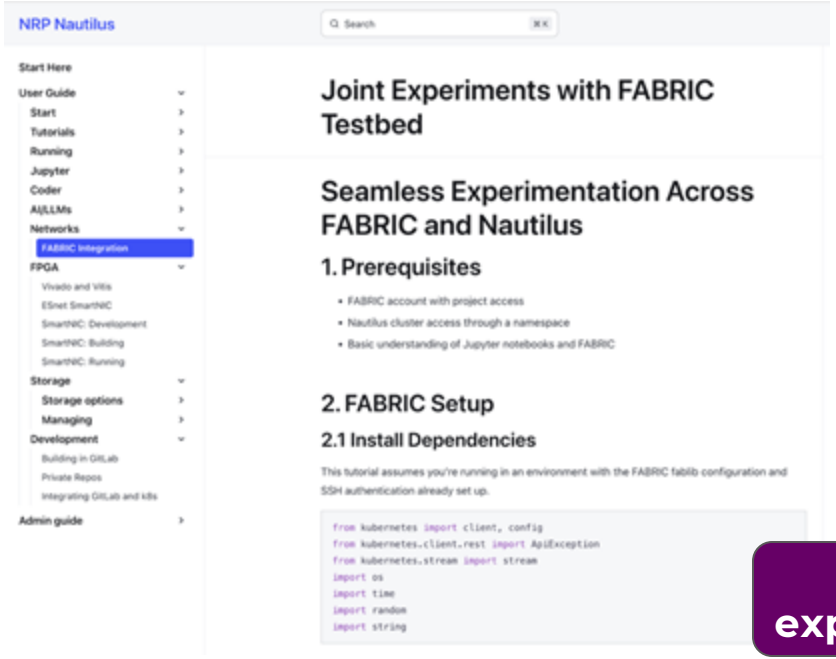
- Fixed-Point Arithmetic + Taylor Series Approximations
  - Convert floating point operations into fixed point
  - Approximate nonlinear functions (e.g. activation) via Taylor expansions in the data plane
- Separation of Data Plane & Control Plane Roles
  - Control plane vs data plane stores model parameters (weights, biases) in tables
  - Data plane handles packet flow, feature extraction, and approximate inference logic
- Packet Encapsulation / Inference Embedding
  - Input features carried in packet headers
  - Inference result embedded in outgoing packet
  - Entire pipeline stays at line rate (minimal extra latency)

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

Experiments



https://docs.nrp.ai

Reproducibility is the experiments' main objective

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION

Thank you!

- The National Research Platform: https://nrp.ai
- The FABRIC Testbed: https://whatisfabric.net
- Real-time available hardware resources: https://nrp.ai/viz/resources
- Training on NRP: https://nrp.ai/training
- Running SmartNIC experiments: https://nrp.ai/documentation/userdocs/fpgas/esnet_development/
- Community Gitlab: https://gitlab.nrp-nautilus.io/
- Matrix chat & community: https://nrp.ai/contact/
- ESnet SENSE and Multus: https://nrp.ai/documentation/admindocs/cluster/sense-multus/

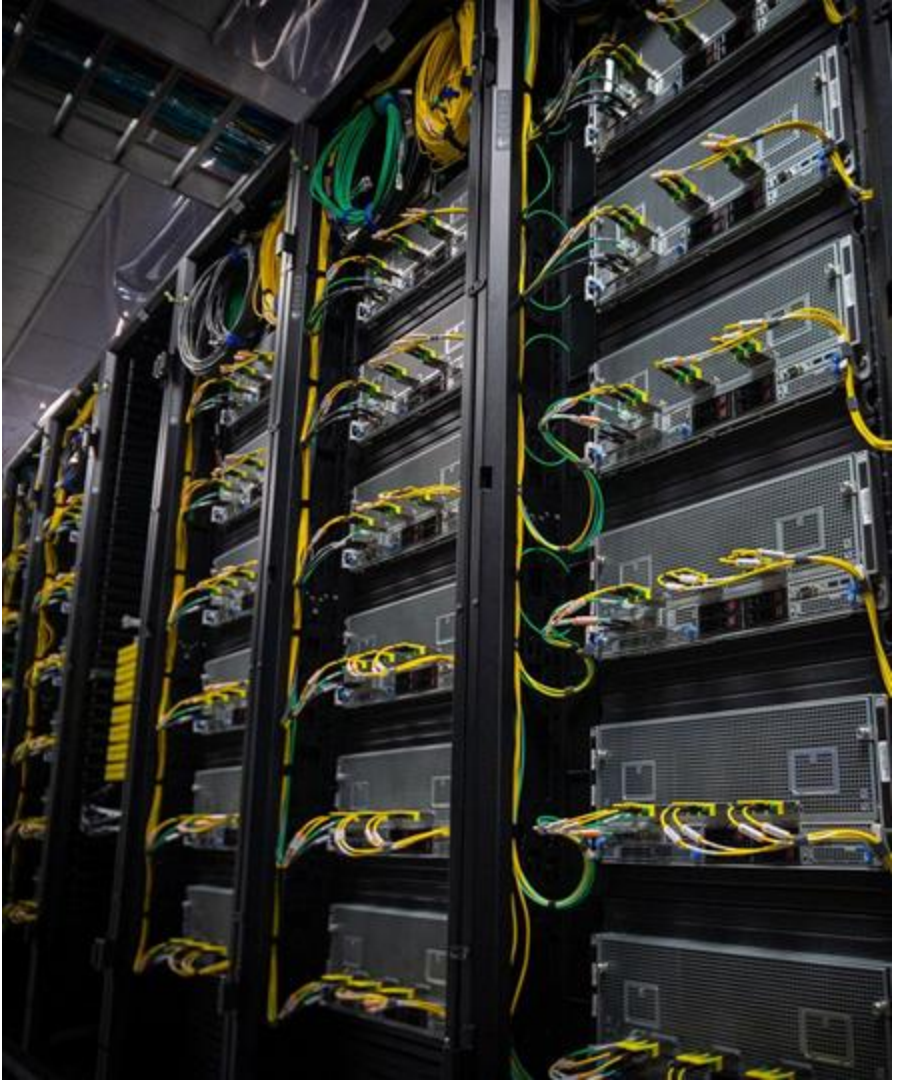
NRP NATIONAL RESEARCH PLATFORM

https://nrp.ai

NRP NATIONAL RESEARCH PLATFORM

SAN DIEGO SUPERCOMPUTER CENTER

THE LINUX FOUNDATION



Helpful links:

- **The National Research Platform: https://nrp.ai/**
- **Link to join: https://portal.nrp-nautilus.io/**
- **Real-time available hardware resources: https:portal.nrp-nautilus.io/resources**
- **Running SmartNIC experiments: https://nrp.ai/documentation/userdocs/fpgas/esnet_development/**
- **Publicly available LLMs: https://nrp.ai/documentation/userdocs/ai/llm-managed/**
- **Community Gitlab: https://gitlab.nrp-nautilus.io/**
- **Matrix chat & community: https://nrp.ai/contact/**
- **ESnet SENSE and Multus: https://nrp.ai/documentation/admindocs/cluster/sense-multus/**

NRP
NATIONAL RESEARCH PLATFORM

https://nrp.ai