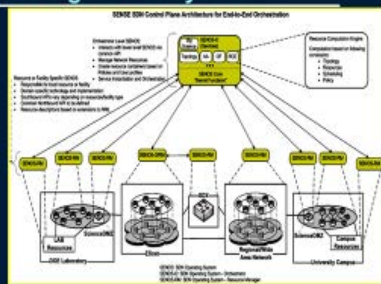

Next Generation Cyberinfrastructures
for LHC and Data Intensive Sciences
SC18
ESnet
CENIC
PACIFIC WAVE
YALE
PACIFIC NORTHWEST GIGAPOP
PRP
BERKELEY LAB
INTERNET2
Fermilab
CERN
13 TeV
LHC Beyond the Higgs Boson
LSST SKA
BioInformatics
Earth Observation
LHC
LSST
SKA
Joint Genome Institute
Gateways to a New Era
Caltech and Partners at SC18
STARLIGHT
+ INDUSTRY
Discovery of the Higgs Boson
July 4, 2012
Science
Physicists Find Elusive Particle Seen as Key to Universe
2013 Nobel Prize
Englert
Higgs
Energy Frontier
Intensity Frontier
Cosmic Frontier
Higgs Boson
Neutrino Mass
Dark Matter
Cosmic Acceleration
Explore the Unknown
A 48 Year Search
Theory : 1964
LHC + Experiments Concept: 1984
Construction: 2001
Operation: 2009
Discovery: 2012
The Economist
A giant leap for science
Advanced Networks Were Essential to the Higgs Discovery and Every Ph.D Thesis of the last 20+ Years
They will be Essential to Future Discoveries, and Every Ph. D Thesis to Come
Entering a new Era of Exploration and Discovery in Data Intensive Sciences
▪We have entered a new era of exploration and discovery
▪In many data intensive fields, from HEP and astrophysics to climate science, genomics, seismology and biomedical research
▪ The largest data- and network-intensive programs from the LHC and HL LHC, to LSST and DESI, LCLS II, the Joint Genome Institute and other emerging areas of growth will face a new round of unprecedented challenges
▪ In global data distribution, processing, access and analysis
▪ In the coordinated use of massive but still limited CPU, storage and network resources.
▪ High-performance networking is a key enabling technology for this research: global science collaborations depend on fast and reliable data transfers and access on regional, national and international scales
Vision: Next Gen Integrated Systems for Exascale Science: a Major Opportunity
Opportunity: Exploit the Synergy among
1. Global operations data and workflow management systems developed by HEP programs, to respond to both steady state and peak demands
▪ Evolving to work with increasingly diverse (HPC) and elastic (Cloud) resources
WLCG
2. Deeply programmable, agile software-defined networks (SDN), emerging as multidomain network operating systems (e.g. SENSE & SDN NGenIA; Multidomain multicontroller SDN)
3. Machine Learning, modeling and game theory: Extract key variables; move to real-time self-optimizing workflows with Reinforcement Learning.
NGenIA-ES Services and Data Flow Diagram
Resources Scheduler
Job Scheduler
Request Manager
WMAgent
Schedd (HTCondor)
CRAB3 backend
Network Resources Scheduler
SENOS
SDN Controller(s) Open Daylight
ALTO
OVS
Data Transfer Scheduler
VO Apps
PhEDEx
ASO
+ ATLAS Big PanDA, LSST, etc.
Exascale Data Transfer Orchestrator
Provisioning
Modeling and Simulation
Machine Learning & Prediction
Monitoring and State Tracking
Leadership and HPC Facilities
US Tier1s
US Tier2s
Cloud
Opportunistic
OSG
SC15-18: SDN Next Generation Terabit/sec Ecosystem for Exascale Science
supercomputing.caltech.edu
SDN-driven flow steering, load balancing, site orchestration Over Terabit/sec Global Networks
SC16+: Consistent Operations with Agile Feedback Major Science Flow Classes Up to High Water Marks
Preview PetaByte Transfers to/ from Site Edges of Exascale Facilities With 100G -1000G DTNs
SC16 SDN-WAN Demonstration End-Points
Caltech, UM, Vanderbilt, UCSD, Dell, 2CRSI, KISTI, StarLight, PRP, FIU, RNP, UNESP, CERN
Pacific Research Platform
900 Gbps Total
Peak of 360 Gbps in the WAN
170G
170G
Global Topology
29 100G NICs; Two 4 X 100G and Two 3 X 100G DTNs; 1.5 Tbps Capability in one Rack; 9 32 X100G Switches
LHC at SC15: Asynchronous Stageout (ASO) with Caltech's SDN Controller
Tbps Ring for SC17: Caltech, Ciena, Scinet, OCC/ StarLight + Many HEP, Network, Vendor Partners at SC16

SENSE: SDN for End-to-end Networked Science at the Exascale

ESnet Caltech Fermilab LBNL/NERSC Argonne Maryland



- ❒ Mission Goals:
  - ❒ Improve end-to-end performance of science workflows
  - ❒ Enabling new paradigms: e.g. creating dynamic distributed 'Superfacilities'.
- ❒ Comprehensive Approach: An end-to-end SDN Operating System, with:
- ❒ Intent-based interfaces, providing intuitive access to intelligent SDN services
- ❒ Auto-provisioning of network devices and Data Transfer Nodes
- ❒ Network measurement, analytics and feedback to build resilience
- ❒ Longer Term: Ai Optimized policy-guided E2E Orchestration

SENSE: SDN for End-to-end Networked Science at the Exascale

Vision and Objectives

A new paradigm for Application to Network Interactions

- **Intent Based** – Abstract requests and **questions** in the context of the application objectives.
- **Interactive** – what is possible? what is recommended? let's negotiate.
- **Real-time** – resource availability, provisioning options, service status, troubleshooting.
- **End-to-End** – multi-domain networks, end sites, and the network stack inside the end systems.
- **Full Service Lifecycle Interactions** – continuous conversation between application and network for the service duration.

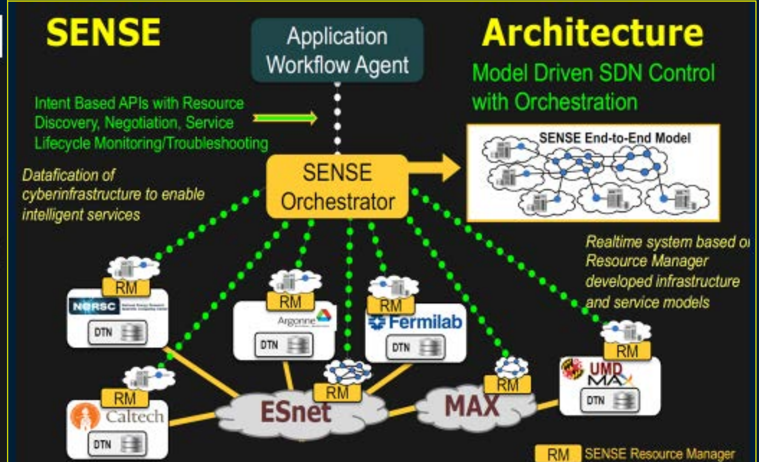
SENSE: SDN for End-to-end Networked Science at the Exascale

SENSE Solution Approach

End-to-End model-based distributed resource reasoning and intelligent service orchestration

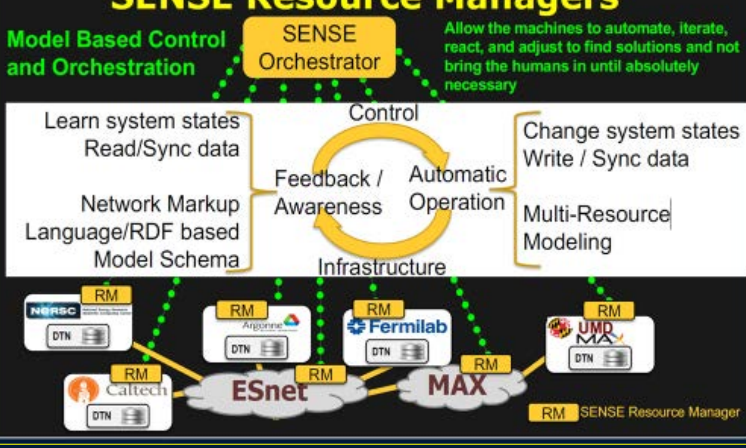
- Hierarchical service resource architecture
- Unified network and end-site resource modeling and computation
- Model based realtime control
- Application driven orchestration workflow
- End-to-end network data collection and analytics integration

SENSE Architecture

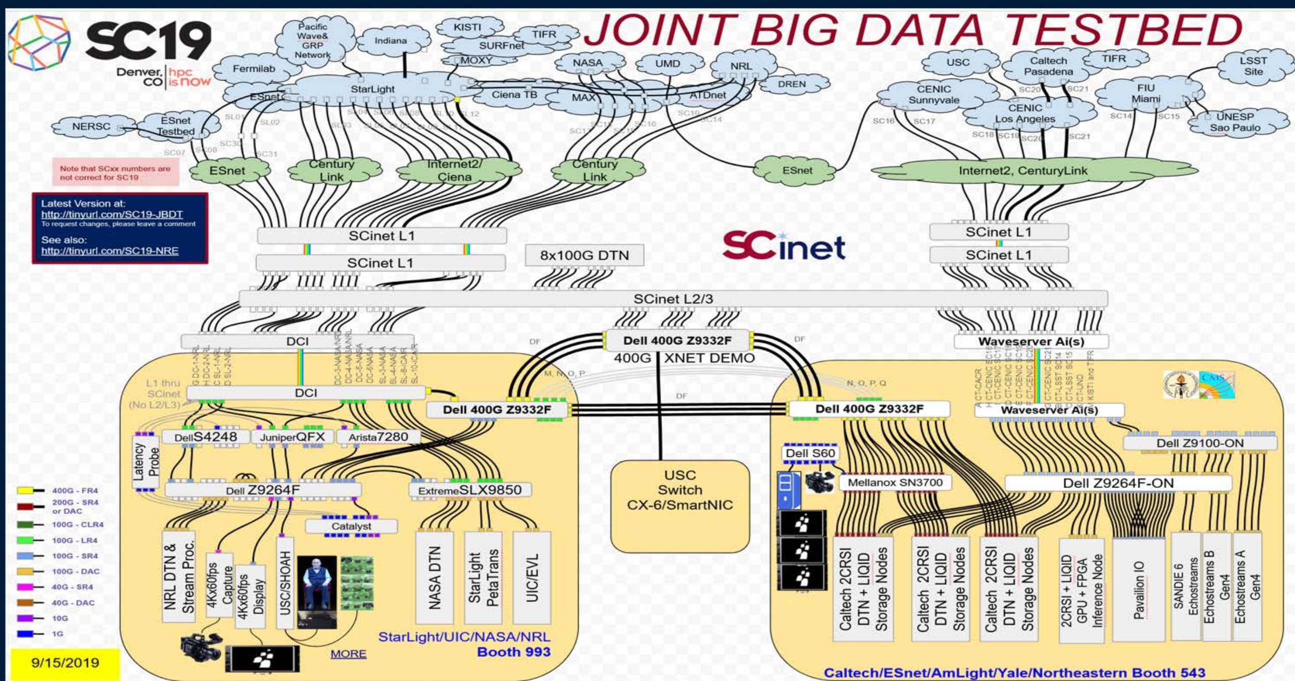


SENSE: SDN for End-to-end Networked Science at the Exascale

SENSE Resource Managers



Caltech and Partners at SC19



Microcosm: Creating the Future of SCinet and of Networks for Science

- LHC/HEP, LSST/Astrophysics; AmLight Express+Protect, SENSE, SANDIE(NDN), SDN NGenIA, Mercator, Carbide Multicontroller SDN Projects
- Ai Presentations: CMS Trigger w/Fast Training and Intereference, Higgs Bosons and Interaction Networks, Quantum Noise Mitigation, Inline Monitoring + Decisions
- DWDM+Waveserver Ais in the Caltech booth: *400G waves, 16 100G clients*
- "Caltech-Starlight-SCinet+USC" Triangle: 400GE Dell + Mellanox 200GE Switches
- ~10 Tbps Server Capacity at the Caltech Booth in 1/2 Rack
  - 3 2CRSI 4-Node + Echostreams AMD Rome (PCIe Gen4) Servers; Pavilion IO NVMeoF; 28 processors, 28 *200GE + ~40 100GE NICs, ~160 SSDs*
  - QSFP56 200GE + DD 400GE Optics; 400G to 2 X 200G Splitters
- Network, Server, Storage Partners: SCinet, Ciena, Mellanox, 2CRSI, Echostreams, Pavilion IO, LIQID, XiLinx
- Science and Network Partners: USC, AmLight, Starlight, CENIC, Pacific Wave, SURFnet, Maryland/MAX, USCD, UERJ, GridUNESP, KISTI/KASI, Michigan, NCSA
- WAN Sites: Caltech, FIU, Maryland, Starlight, PRP, UCSD, Stanford, LBNL, CENIC, Fermilab, AmLight, LSST (Chile), Grid UNESP, UERJ (Rio), SURFnet, KISTI/KASI, CERN, TIFR
- Caltech Booth to WANs: 400G to Caltech + USC campuses; 400G to PRP/TNRP via CENIC (UCSD, Stanford, LBNL, UCSC, et al); 200G to Brazil+Chile via AmLight Express (200G Triangle); 200G to ESnet via Sunnyvale