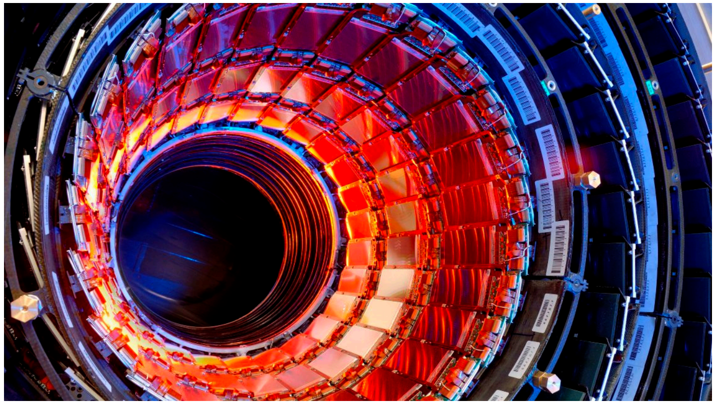
Northeastern

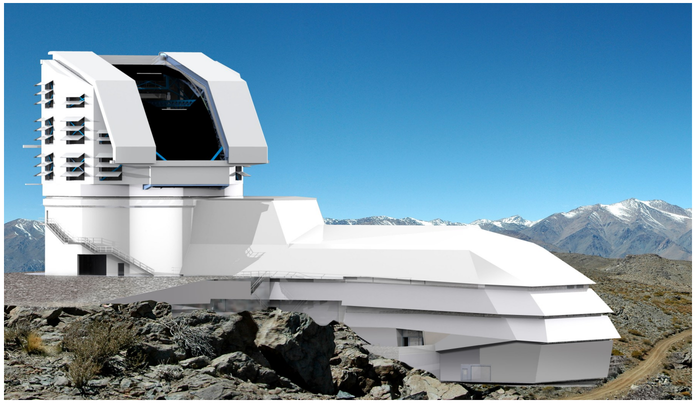
Data-Centric Ecosystems for Large-Scale Data-Intensive Science

Edmund Yeh

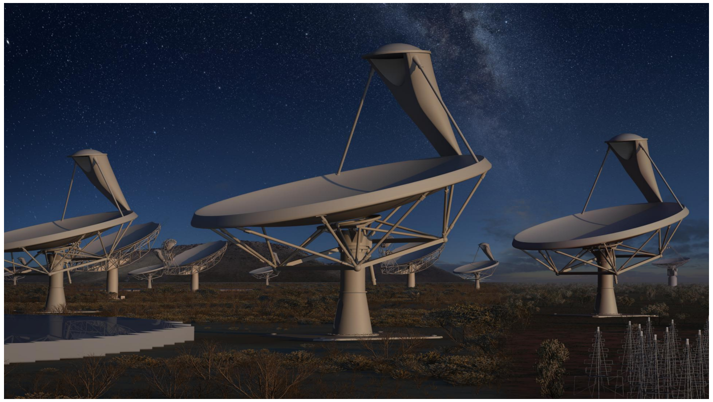
Data-Intensive Science



LHC: Large Hadron Collider



LSST: Large Synoptic Survey Telescope



SKA: Square Kilometer Array

Genomics

Challenges of Data-Intensive Science

- Data-intensive applications face similar set of problems:
  - LHC high energy physics, LSST, LIGO, genomics

- System challenges:
  - Indexing, security, storage, distribution, analysis, learning
  - Coordinated use of computing, storage, network resources

- Today: domain experts are dealing with these systems problems
  - Incremental solutions; developed in isolation; replicated efforts

- Gap between application needs and existing networks/systems
  - Current computer networks/systems focus on addresses, processes, servers, connections
  - Consequently existing security solutions focus on securing data containers and delivery pipes
  - Applications care about data

Data-Centric Networking

- Data-centric approach to system and network design

- Providing system support through the whole data lifecycle:
  - Data production: naming, securing data directly
  - Delivering data using names enables scalable data retrieval
    - in-network caching
    - automated joint caching and forwarding
    - multicast delivery

- Common framework to support different application domains

- Named Data Networking (NDN): http://named-data.net

NSF SANDIE Project

- NSF CC* SANDIE: SDN Assisted NDN for Data Intensive Experiments ($1M grant, started July 2017)

- Team:
  - Northeastern (PI: E. Yeh), Caltech (Co-PI: H. Newman), Colorado State (Co-PI: C. Papadopoulos/C. Partridge)
  - In partnership with other LHC sites and NDN project team

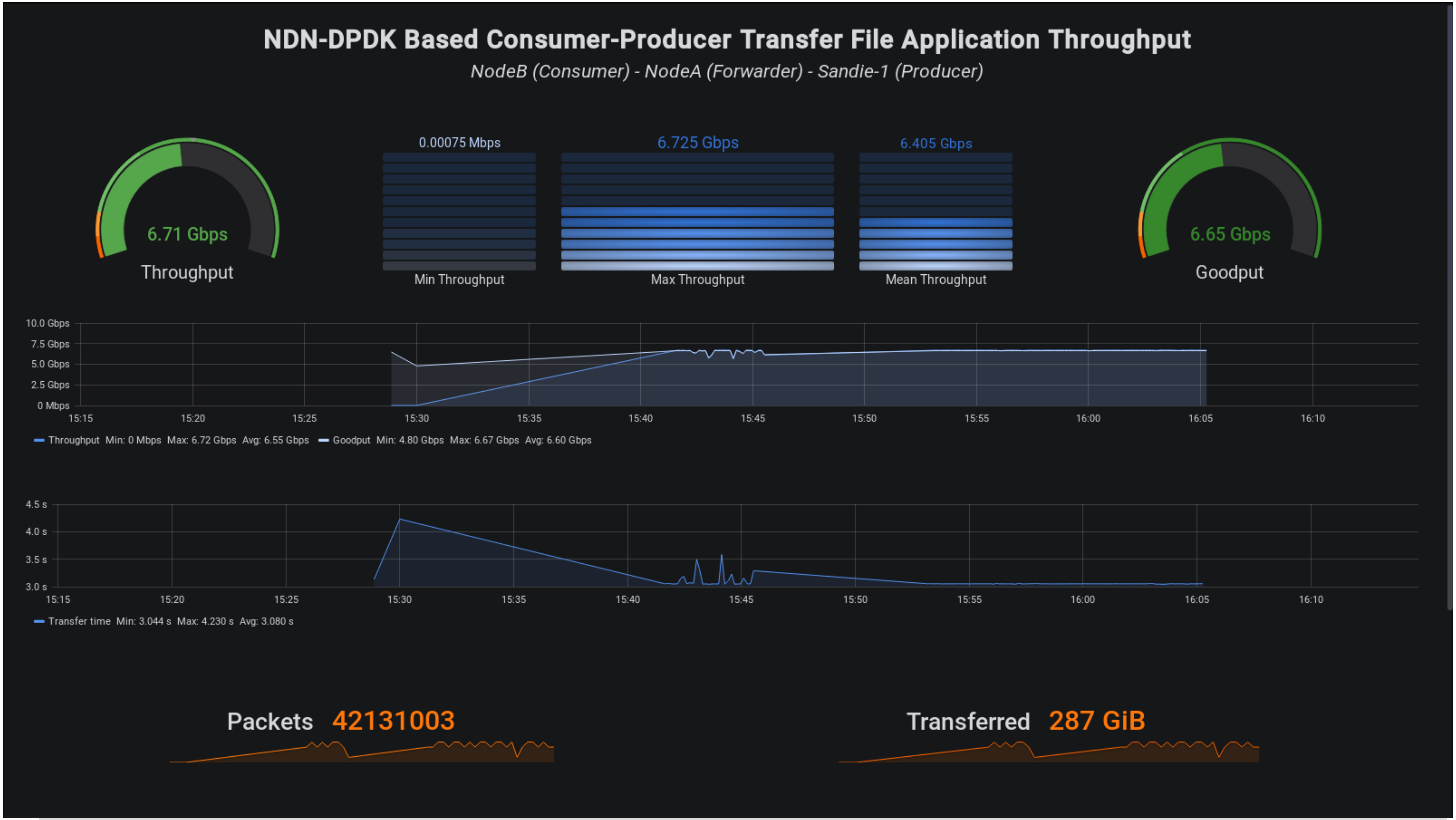
- Approach: Use SDN-assisted NDN to redesign LHC HEP network; optimize workflow

- Solutions:
  - Development of NDN naming scheme for fast access and efficient communication in HEP and other fields -
  - Deployment of NDN edge caches with SSDs at multiple sites
  - Simultaneously optimization of caching ("hot" datasets) and forwarding

SANDIE Results

- Feasibility of NDN-based data distribution system for LHC first demonstrated at SC 18 (Dallas)
- SANDIE demo at SC 19 (Denver) showed greatly improved throughput and delay performance by leveraging:
  - VIP jointly optimized caching and forwarding algorithms developed by Northeastern (Yeh et al. 2014)
  - High-speed NDN-DPDK forwarder developed by NIST (Shi et al. 2019)
  - NDN-DPDK-based consumer and producer developed by Caltech (Iordache, et al. 2019)
- Demo over SANDIE transcontinental layer-2 demo testbed (SC19 – Caltech – Northeastern – Colorado State)
- Achieved over 6.7 Gbps throughput (single thread) between NDN-DPDK-based consumer and producer
- Optimized caching and forwarding algorithms decreased download times by a factor of 10.

XRootD with NDN-DPDK: Throughput (1 Thread)



Northeastern

NSF N-DISE Project

- NSF CC* N-DISE: NDN for Data Intensive Science Experiments ($875K grant, 2020-2022)

- Team:
  - Northeastern (PI: E. Yeh), Caltech (Co-PI: H. Newman), UCLA (Co-PIs: L. Zhang and J. Cong), Tennessee Tech (Co-PI: S. Shannigrahi)
  - In partnership with LHC, genomics collaborators and NDN project team

- Challenges:
  - LHC data volume to grow 10x due to High Luminosity LHC (2027): increased data complexity
  - Human genome data, Earth Biogenome (~ exabyte range)
  - Need to use diverse computation, storage, networking resources

- Approach: build data-centric ecosystem to provide agile, integrated, interoperable, scalable, robust and trustworthy solutions for heterogeneous data-intensive domains

N-DISE Goals

- Deploy, commission first prototype production-ready NDN-based petascale data distribution, caching, access, computation system serving major science programs
- LHC high energy physics program as leading target use case. BioGenome, human genome projects, ATLAS, LSST, SKA as future use cases
- Leverage NDN protocols, high throughput forwarding/caching methods, containerization techniques
- Integrated with SDN methods and FPGA acceleration subsystems
- Deliver LHC data over wide area network at throughputs ~ 100 Gbps
- Dramatically decrease download times by using optimizing caching
- Enhanced testbed with high performance NDN data cache servers

N-DISE Research Agenda

- **NDN-DPDK Consumer and Producer**
  - Multi-threaded consumer and producer applications: linear throughput scaling
  - Integrate with NDN-DPDK XRootD plugin
- **Containers for NDN-DPDK and applications**
  - For diverse server equipment and interfaces
  - Docker containers: host guest OS, state restoration, ease upgrading
- **Data integrity and provenance**
  - Data origin authentication
  - Data manifest: compute hashes of data segments, sign hashes

N-DISE Research Agenda

- **Congestion control and retransmission**
- **Caching and Storage**
  - Multi-threaded VIP caching and forwarding
  - Hierarchical cache system (SSDs, Intel Optane, etc.) for handling data volume
- **FPGA acceleration**
  - Programming adaptive, high throughput, low latency, energy efficiency
  - Accelerate hash functions and lookups in NDN-DPDK Name Dispatch Table (NDT), PIT-CS Composite Table (PCCT), and FIB

Conclusions

- Data-intensive science applications require fundamental network/systems solutions to address common needs

- NDN provides data-centric system support through whole data lifecycle -
  - natural fit for LHC, genomics and other data-intensive applications

- High performance NDN: SANDIE showed throughput at 6.7 Gbps over WAN
  - NDN-DPDK forwarder
  - VIP jointly optimized caching and forwarding algorithms

- N-DISE: toward first prototype production-ready NDN system: integration with SDN, FPGA, containerization

- Seeking long-term collaboration with domain science, networking and computer systems communities

References

- C. Iordache, R. Liu, J. Balcas, R. Srivinskas, Y. Wu, C. Fan, S. Shannigrahi, H. Newman, and E. Yeh, “Named Data Networking based File Access for XRootD." Proceedings of the International Conference on Computing in High Energy and Nuclear Physics (CHEP), 2019.

- C. Fan, S. Shannigrahi, S. Dibendetto, C. Olschanowsky, C. Papadopoulos, H. Newman, E. Yeh, J. Vlimant, A. Amin, D. Kcira, I. Legrand, and R. Voicu, “Managing Scientic Data with Named Data Networking." Proceedings of the International Workshop on Network-aware Data Management (NDM), 2015.

- S. Shannigrahi, C. Papadopoulos, E. Yeh, H. Newman, A. Barczyk, R. Liu, A. Sim, I. Monga, J. Vlimant, and J. Wu, “Named Data Networking in Climate Research and HEP Applications." Proceedings of the International Conference on Computing in High Energy and Nuclear Physics (CHEP), 2015.

Northeastern

N-DISE Genomics Pilot

- Genomics use case complements HEP use case
    - Both static and dynamic data sets
    - Centralized (e.g. NCBI) and distributed data publishers
- N-DISE genomics agenda:
    - High-speed data transfer between arbitrary entities (e.g. institutional labs)
    - Porting tools and protocols developed for HEP community into genomics
        - e.g. data discovery, optimized caching algorithm
    - Containerize NDN-forwarder and applications which benefit both HEP and genomics
        - e.g. lessons from deploying kubernetes based genomics workflows can help HEP communities to deploy their workflows to the cloud