Optimization of Rucio-SENSE DMM

Camille Sicat, with mentorship by Aashay Arora and Diego Davila

UC San Diego

Premise

- Current experiments, such as CMS, generate as much as 20 PB/year
- Future planned experiments, such as the HL-LHC, will generate 30x more data
- To replicate this volume of data to storage sites, need >=100 Gbps speeds
  - Estimated 400 Gbps between sites in the U.S.
- Current network model, networks dominated by large transfers
- Need more accountable network usage

Rucio-SENSE DMM

- SENSE (**S**oftware-Defined Network for **E**nd-to-end **N**etworked **S**cience at the **E**xascale) allows for customizable multi-domain orchestration
  - For use on experiments with individual domain science workflows + requirements
  - Pushes QoS and routing rules
- Rucio - CERN's data management software
  - Provides scalable data storage, transfer, replication, etc. across different physical locations
  - Allows for individualized tagging/tracking of data

A High-Level Description of DMM Performance

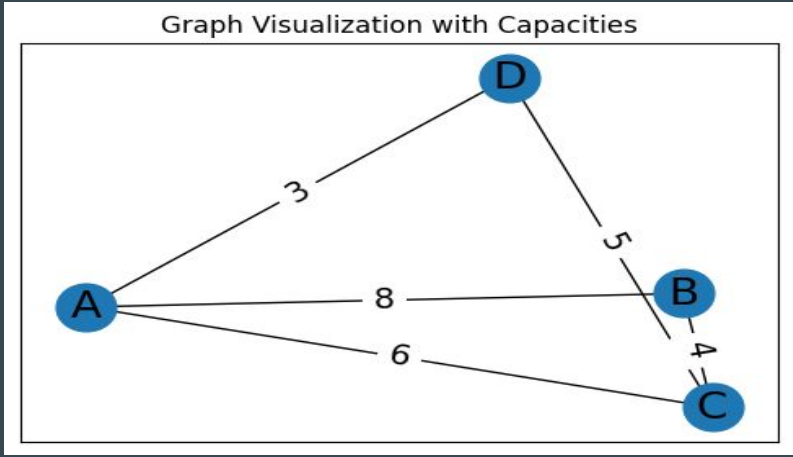
- The Data Movement Manager (DMM) is the interface between SENSE and Rucio
    - From Rucio: source/destination RSE names, number of bytes, priority
    - From SENSE: Constructs P2P VLANs for each set of endpoints
    - Monitors status + performance of all dataflows
- Based on transfer metadata, DMM calculates and provisions network bandwidth for each rule
    - Constantly assesses current bandwidth usage and re-allocates depending on rule priority

Project Outcomes: Monitoring

- Network accountability means being able to track transfer rates and points of failure
- Implementation of DMM's monitoring system through the monit daemon
  - Online monitoring: live tracking of throughput for each rule every 10s using metrics from Prometheus
    - Acts as a live performance report for transfer rates, points of failure, etc.
  - Offline monitoring: correlating data from Prometheus and FTS to analyze completed rules
    - Allows for comparison of performance across different transfers

Project Outcomes

- Attempted optimization of network bandwidth allocation algorithm
  - Tested implementation of weighted fair queuing, linear programming, and max_min_fairness algorithms
  - Requires further research
- Time constraints did not permit optimization of DMM's interactions with Rucio



UC San Diego

Accomplishments

- Developed strong skills in GitHub, Python, SQL, HTML, and graphing algorithms
- Gained hands-on experience on integrating Docker, Kubernetes, Prometheus, FTS, and Flask for project development and deployment
- Emphasized the importance of collaboration and time management

Summary

- Future experiments will generate large amounts of data
- Fellowship was focused on making the Rucio-SENSE DMM capable of holding networks more accountable for data transfers with large amounts of data
- Successfully implemented throughput monitoring for this goal, with need for further study of optimization methods

Thank you!

Backup

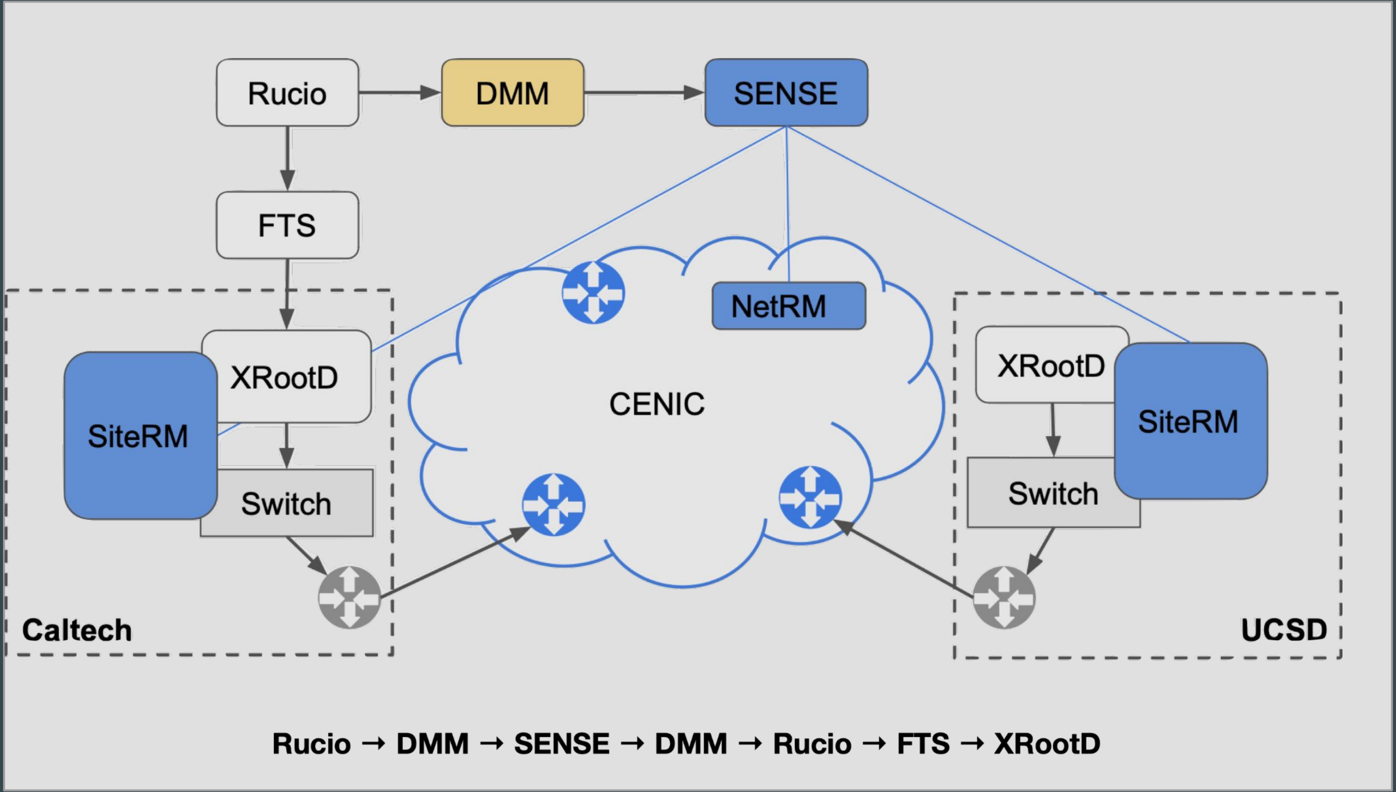
	# of collissions	# of events simulated	RAW event size [MB]	AOD event size [MB]	Total per year [PB]
Run 2	9 Billion	22 Billion	0.9	0.35	~20
HL-LHC	56 Billion	64 Billion	6.5	2	~600

The beams get "brighter" by x6
Data taking rate goes up by x6
Simulations go up by x3

Primary Data volume per year goes up by x30

	RAW	AOD	MINI	NANO
Run 2	0.9 MB/event	0.35 MB/event	0.035 MB/event	0.001MB/event
	8 PB/year	16 PB/year	1 PB/year	0.031 PB/year
HL-LHC	6.5 MB/event	2.0 MB/event	0.250 MB/event	0.002 MB/event
	364 PB/year	240 PB/year	30 PB/year	0.24 PB/year

Detailed Model



UC San Diego

Transfer Performance

UC San Diego