DYNAMO

Funded by the National Science Foundation Grant #1826997

Toward a Dynamic Network-centric Distributed Cloud Platform for Scientific Workflows: A Case Study for Adaptive Weather Sensing

Eric Lyons, George Papadimitriou, Cong Wang, Komal Thareja, Paul Ruth,
J. J. Villalobos, Ivan Rodero, Ewa Deelman, Michael Zink, Anirban Mandal

renci

USC Viterbi School of Engineering
Information Sciences Institute

UMass Amherst

RUTGERS

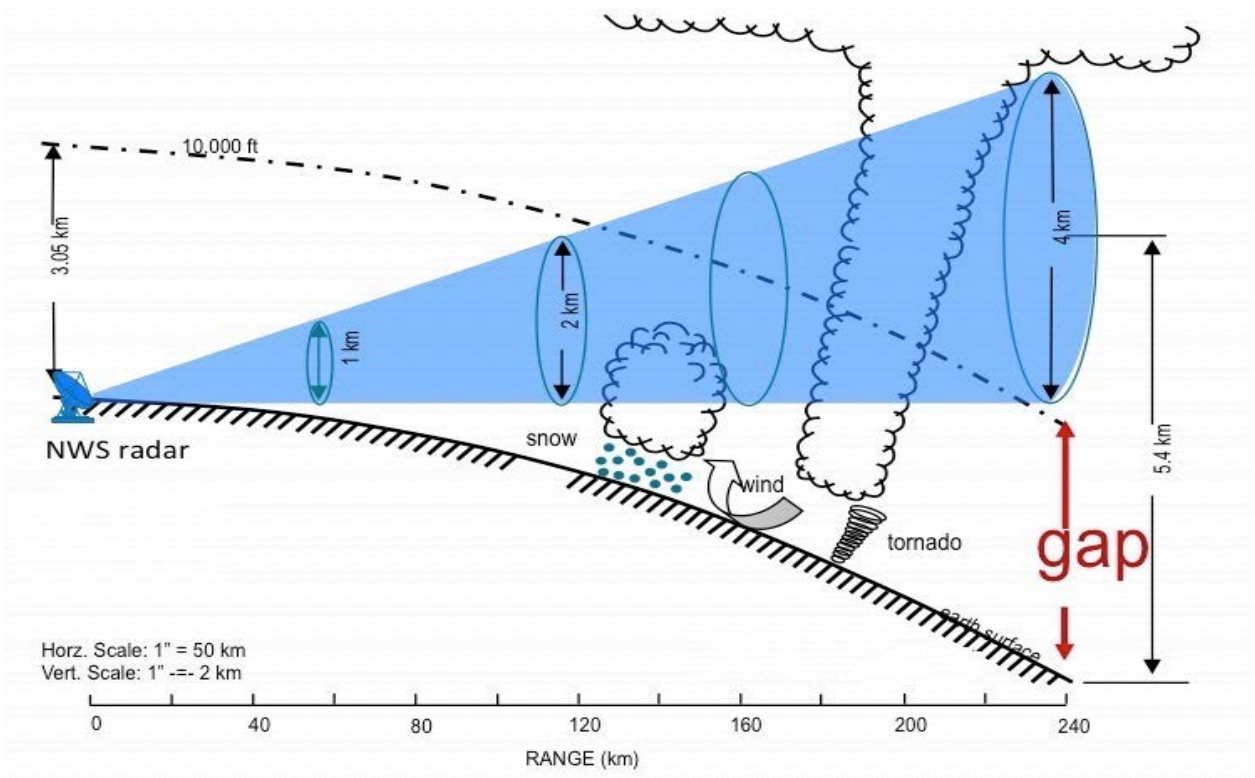
Characteristics of Today's Computational Science

- Data transfer and compute intensive
- Complex workflows
- Distributed data repositories
- Highly distributed compute locations
- Major challenge: Integration of cyber infrastructure to science workflows

DyNamo: Multi-Cloud Platform for Scientific Workflows

- Develop novel algorithms, policies, and mechanisms to offer optimized, adaptive, data flows across different kinds of cyberinfrastructures
- Network-centric platform to bridge the gap between science workflows and high performance network services
- Network-aware workflow scheduling algorithms, predictions and ensemble management – Network-aware Pegasus Workflow Management System (WMS)
- DyNamo case study in this talk
  - Collaborative Adaptive Sensing of the Atmosphere (CASA)

CASA: Collaborative Adaptive Sensing of the Atmosphere



- Traditional Next Generation Weather Radars (NEXRAD)
  - High power, long range
  - Limited ability to observe the lower part of the atmosphere because of the Earth's curvature
- CASA
  - Network of short range Doppler radars
  - Adjustable sensing modes in response to quick weather changes
  - Suitable for near-ground weather events: tornado, hail, high winds

renci

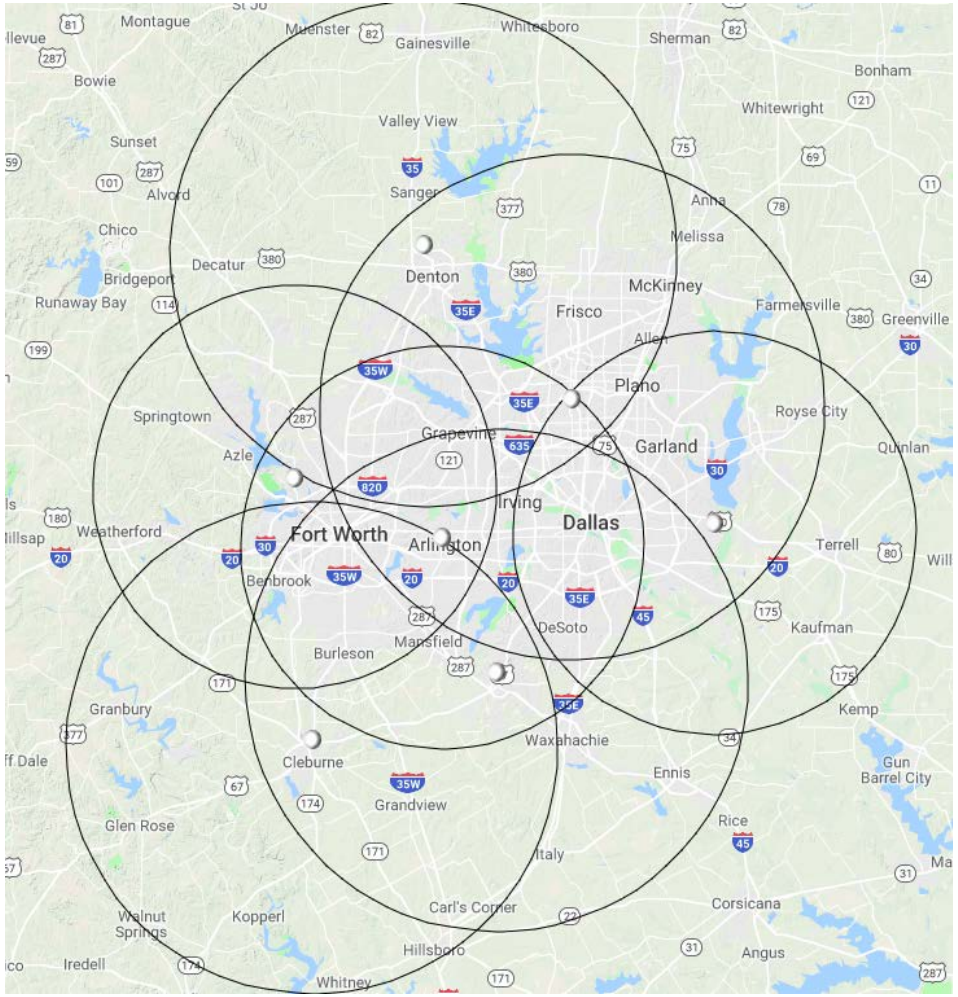
USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

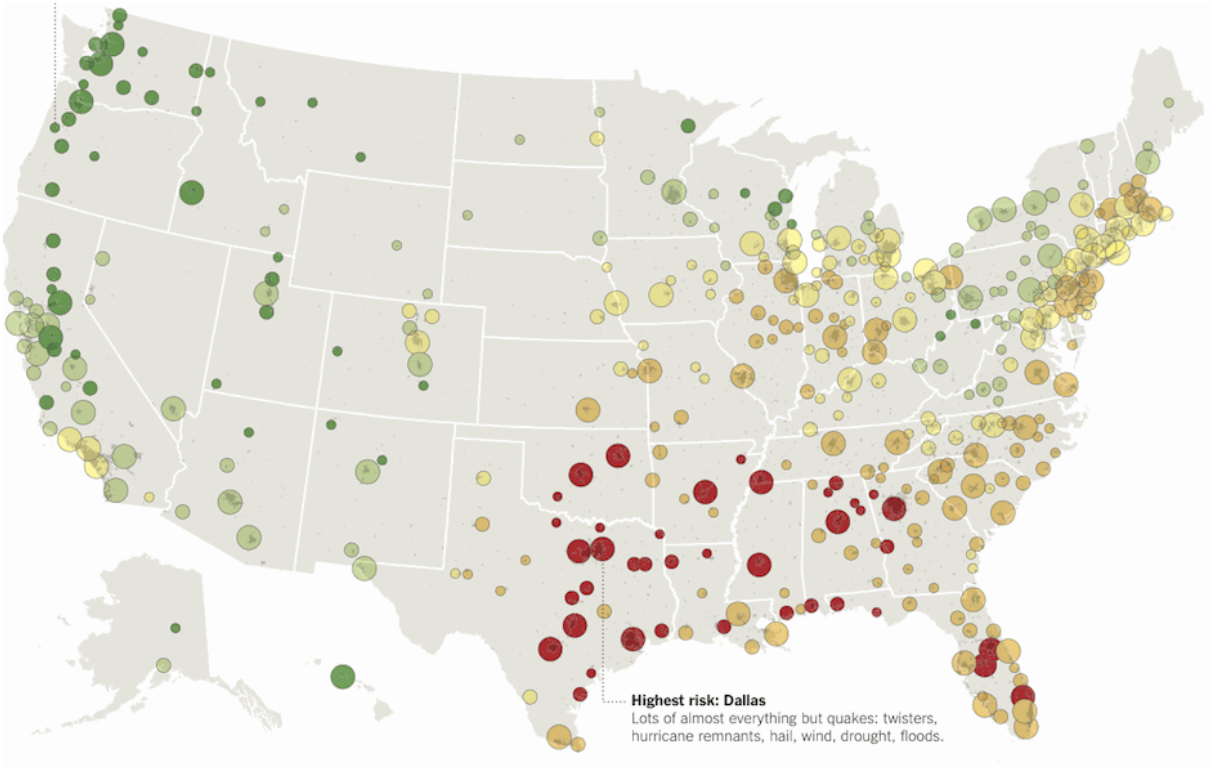
Rutgers

DYNAMO

CASA DFW Network



- > 7M people, >100K businesses, >1500 Corporate HQs



renci

USC Viterbi School of Engineering Information Sciences Institute

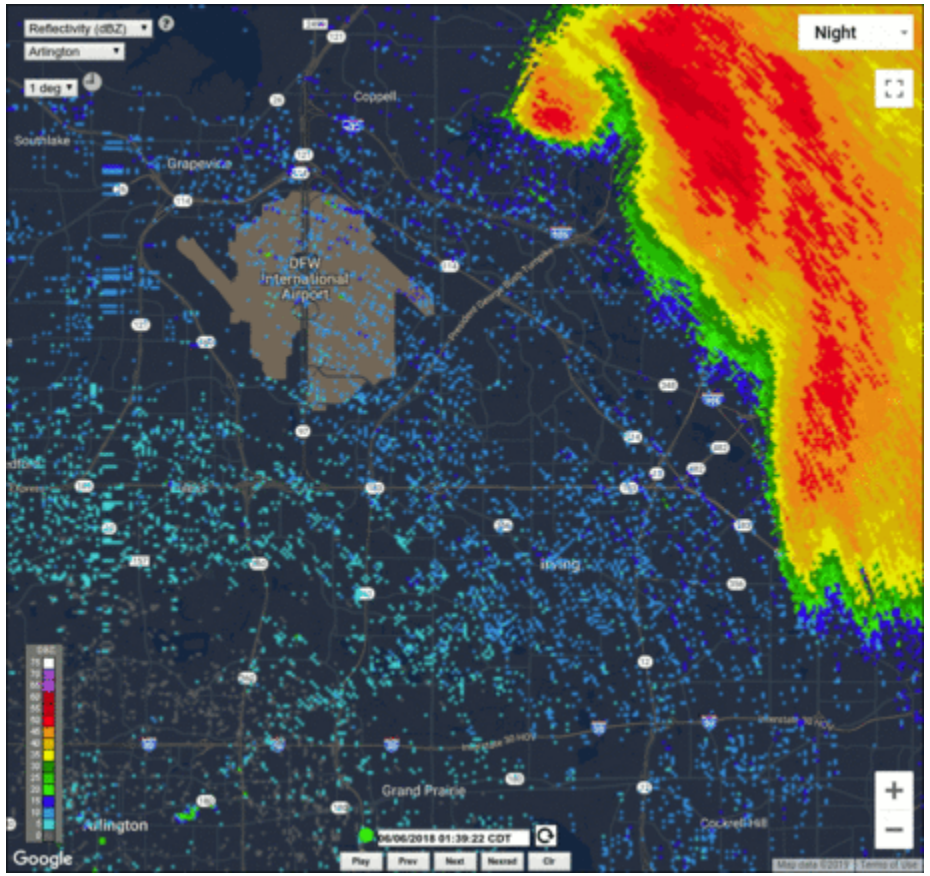
UMass Amherst

RUTGERS

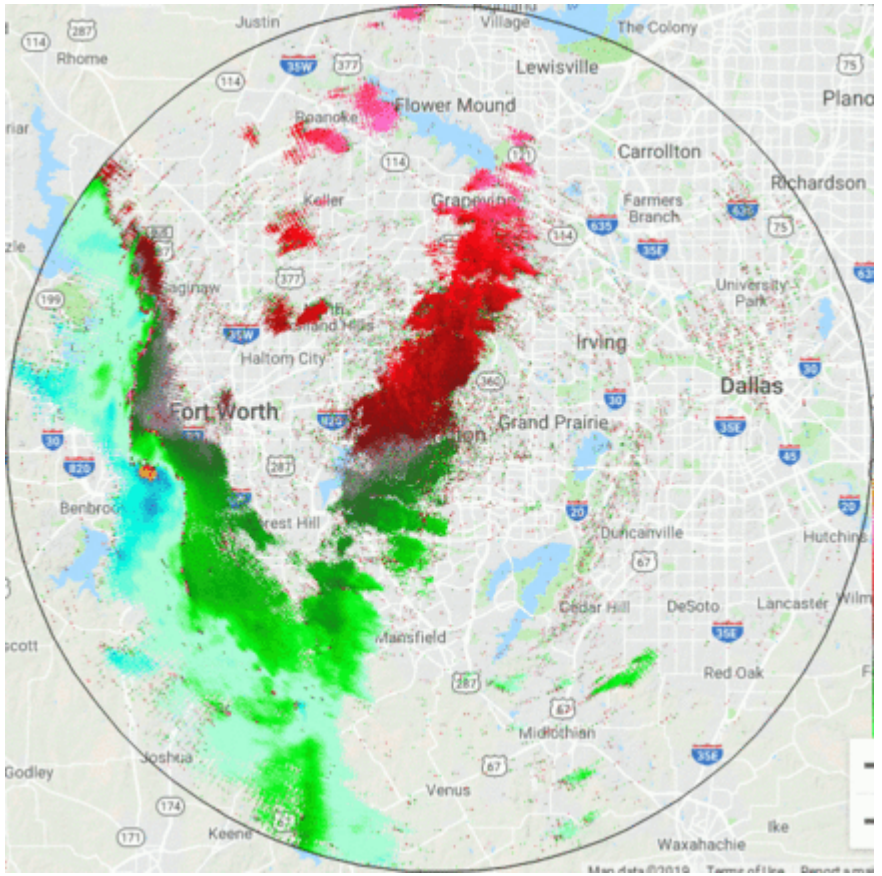
CASA Data Rates

- ~100 Mbps per radar raw data, processed locally
- ~10 Mbps per radar “moment” data, transferred across network
- ~1 Mbps gridded product data
- Transferred to DFW Radar Operations Center at NOAA SRH
- Transferred to Univ. Of North Texas for DYNAMO ingest

Single radar Reflectivity

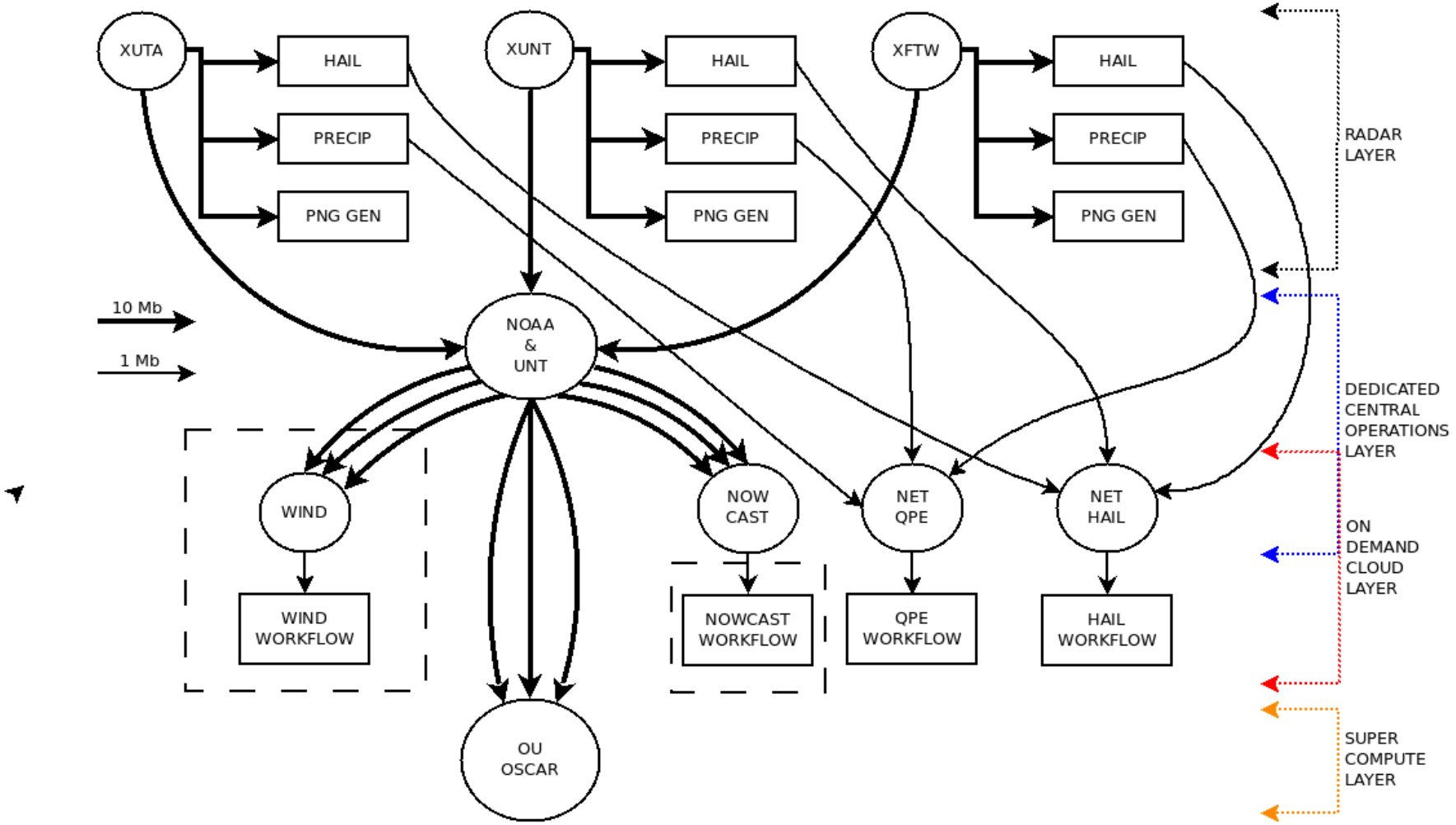


Single radar Velocity



DYNAMO

CASA Workflows



renci

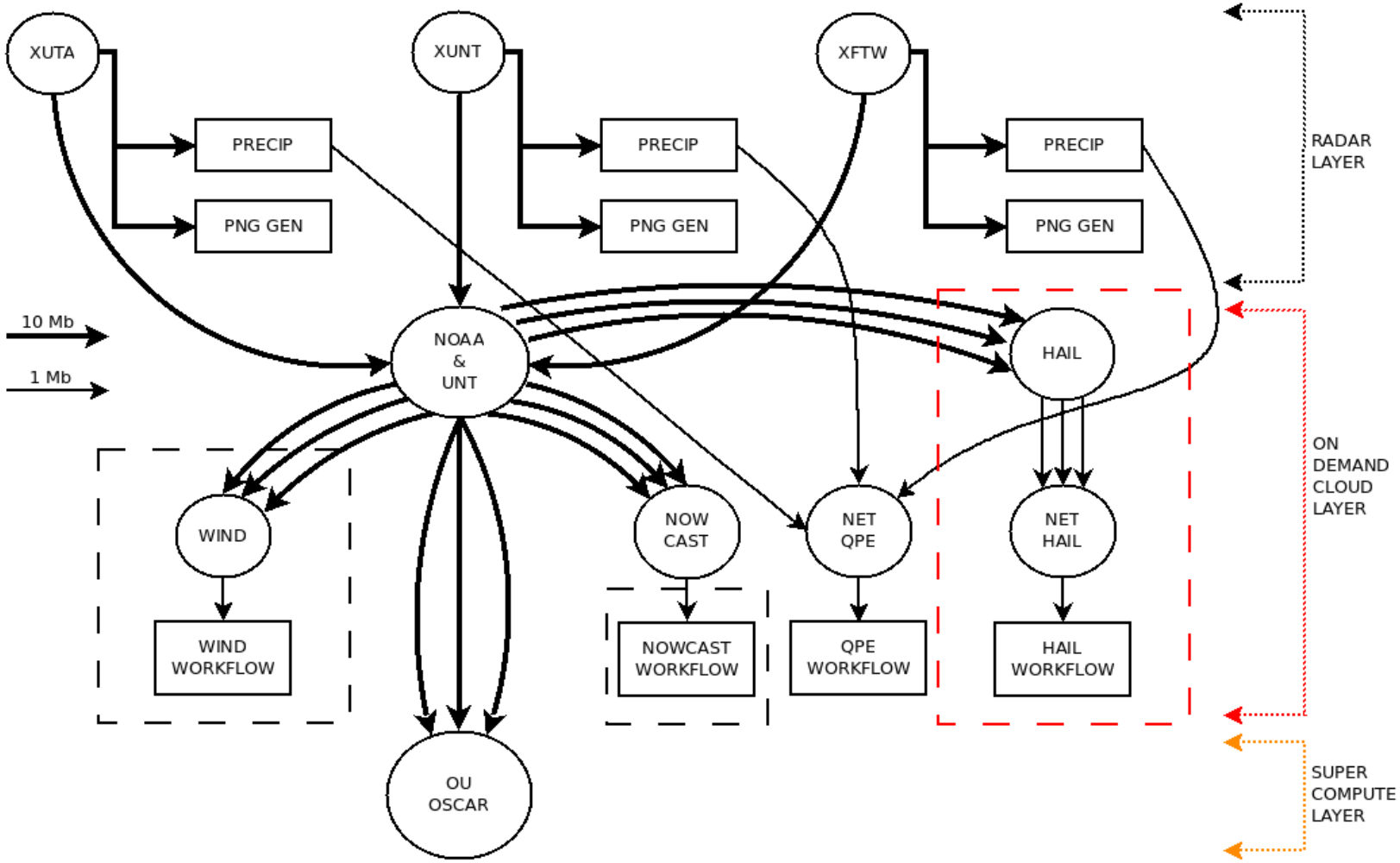
USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

RUTGERS

DYNAMO

CASA Workflows



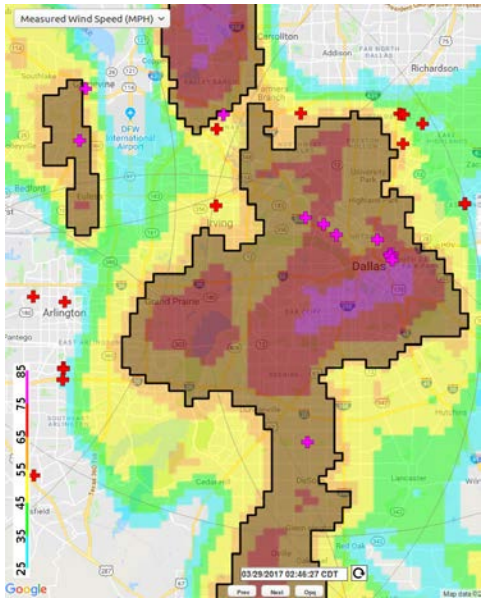
renci

USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

RUTGERS

CASA Wind Workflow

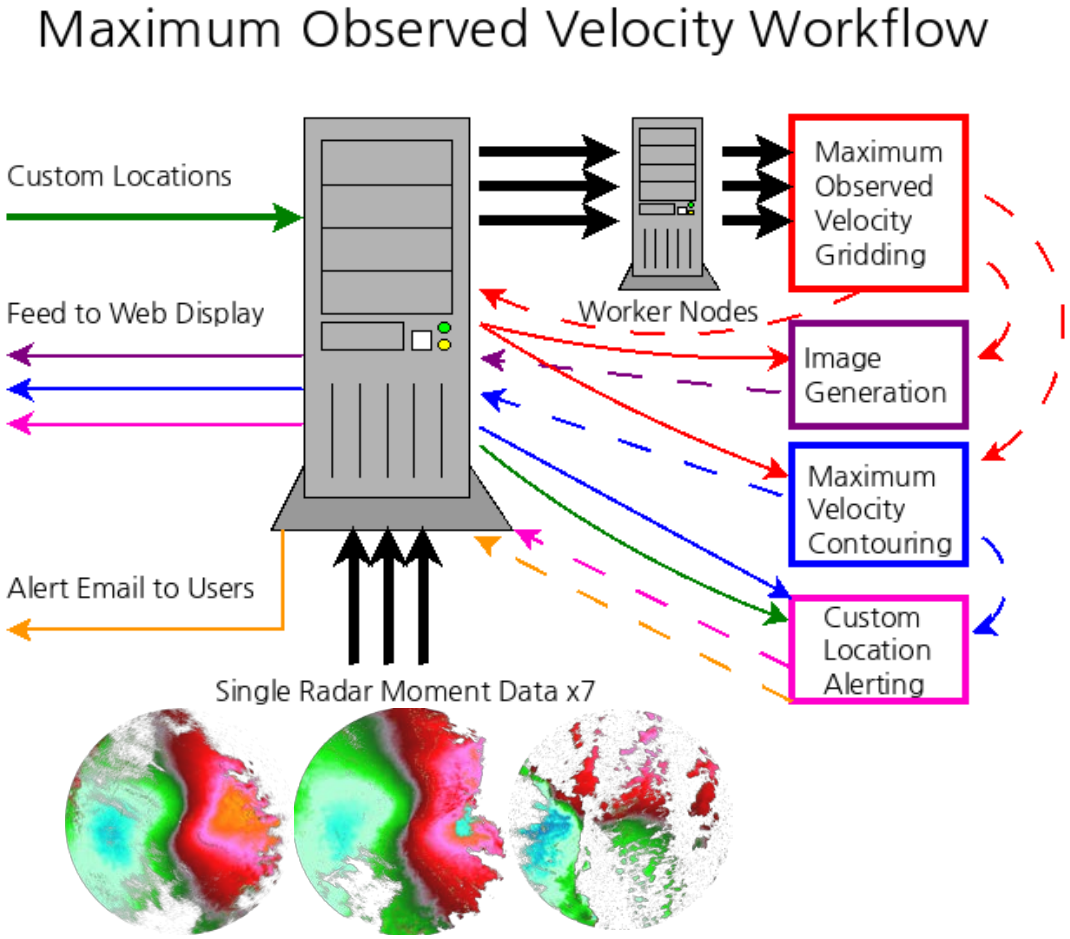


CASA automated notification for BAYLOR UNIVERSITY MEDICAL CENTER

noreply@papajim.eu
Mon 11/5, 2:19 PM
You

siteLocation: BAYLOR UNIVERSITY MEDICAL CENTER
alert type: WINDS_CASA mag: 58MG
timestamp: 2017-03-29T07:55:00Z

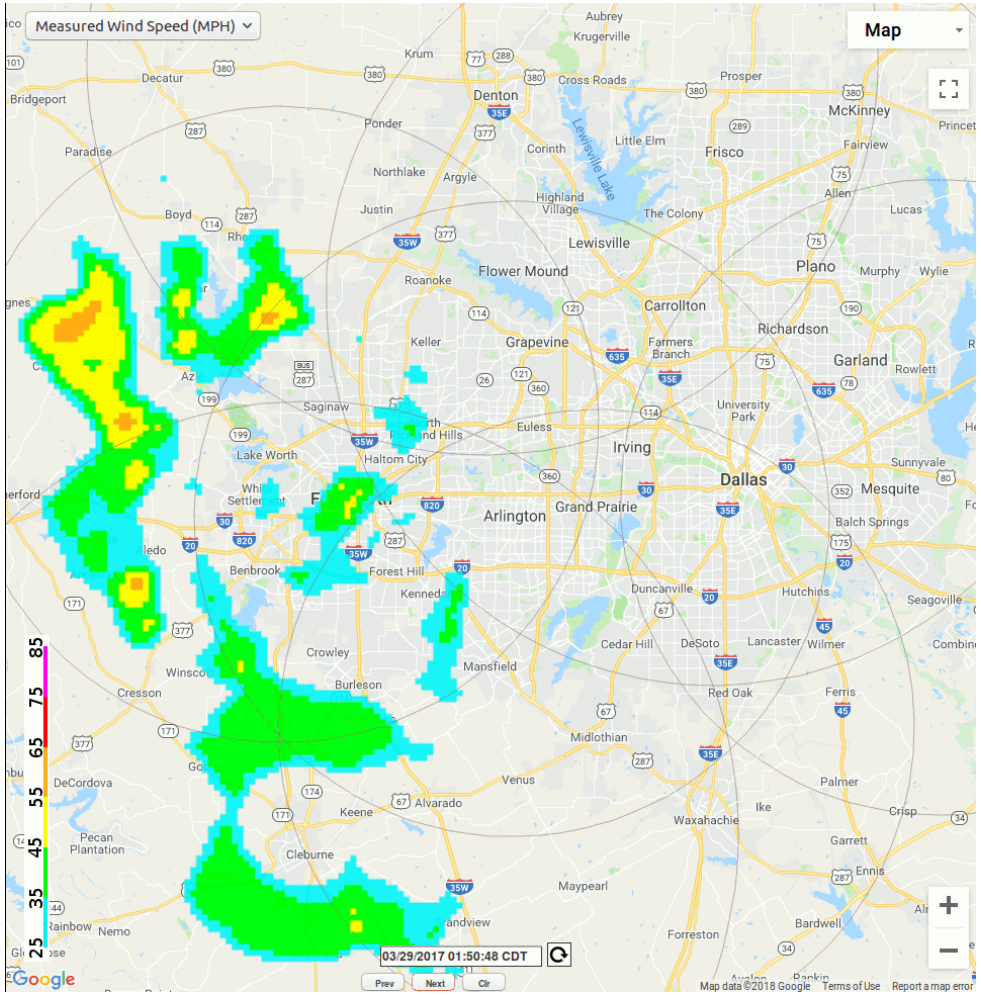
To be removed from this mailing list, please send email to elyons@engin.umass.edu.



- Ingests compressed radar moment data x 7 radars
- Ingests GIS list of infrastructure
- gunzips
- Combines into grid of maximum observed wind speed
- Makes png
- Contours with velocity thresholds
- Compares contours with infrastructure and sends alert emails

DYNAMO

CASA Wind Workflow



- Ingests compressed radar moment data x 7 radars
- Ingests GIS list of infrastructure
- gunzips
- Combines into grid of maximum observed wind speed
- Makes png
- Contours with velocity thresholds
- Compares contours with infrastructure and sends alert emails

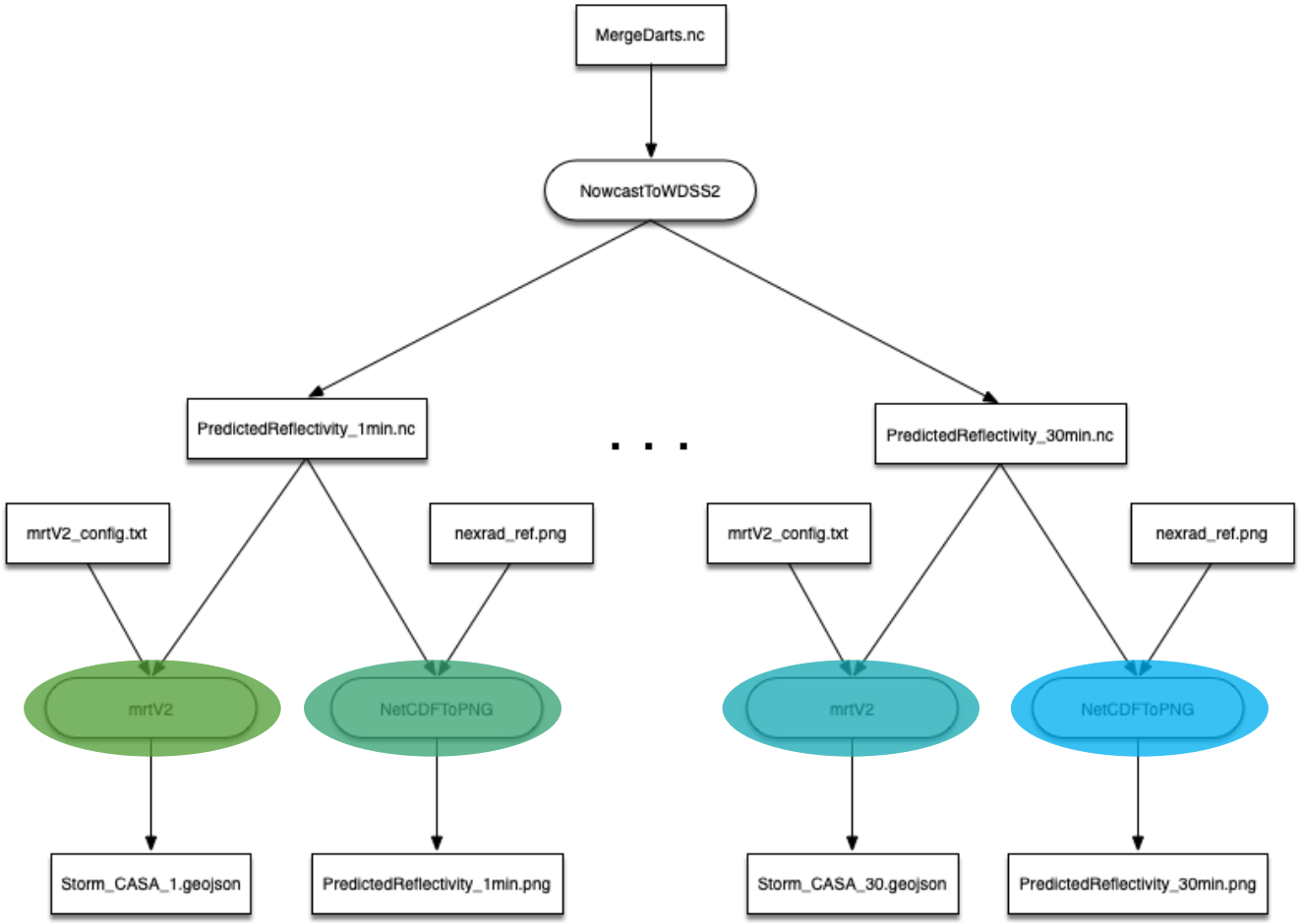
renci

USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

RUTGERS

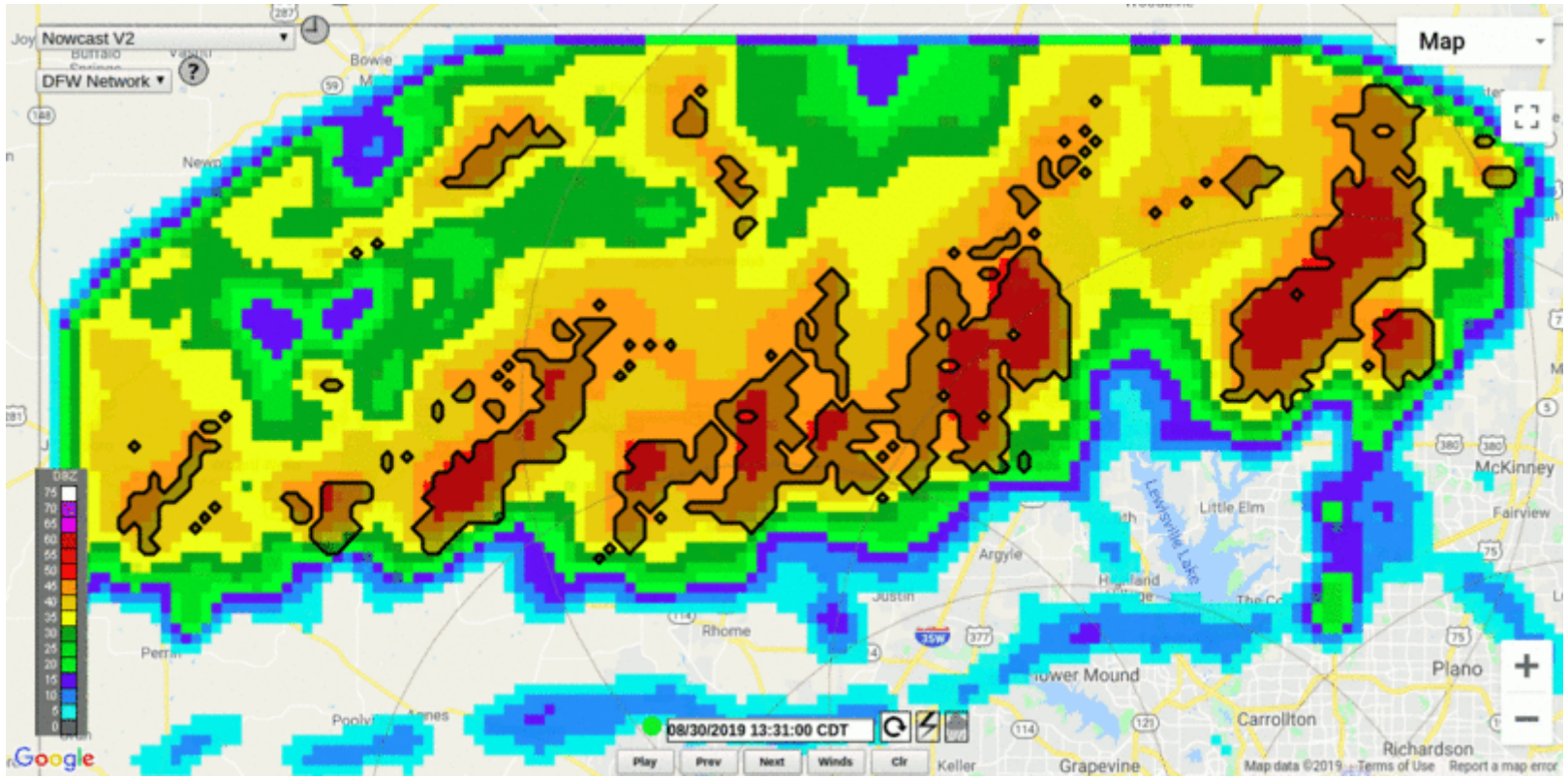
Pegasus Workflow – Nowcast



- Each Nowcast workflow has 63 compute tasks.
- Pegasus supports automated clustering of the compute tasks.
  - Tasks that exist on the same level and use the same executables can be clustered together
  - Custom clustering based on labels is also supported
- All tasks are executed within a Docker container
  - Consistent environment across execution sites

DYNAMO

Pegasus Workflow – Nowcast



- Composites single radar reflectivity data
- Ingests balloon sounding data
- 1 minute forecasts out to 30 minutes
- 31 grids/minute
- Creates pngs
- Contour multiple thresholds

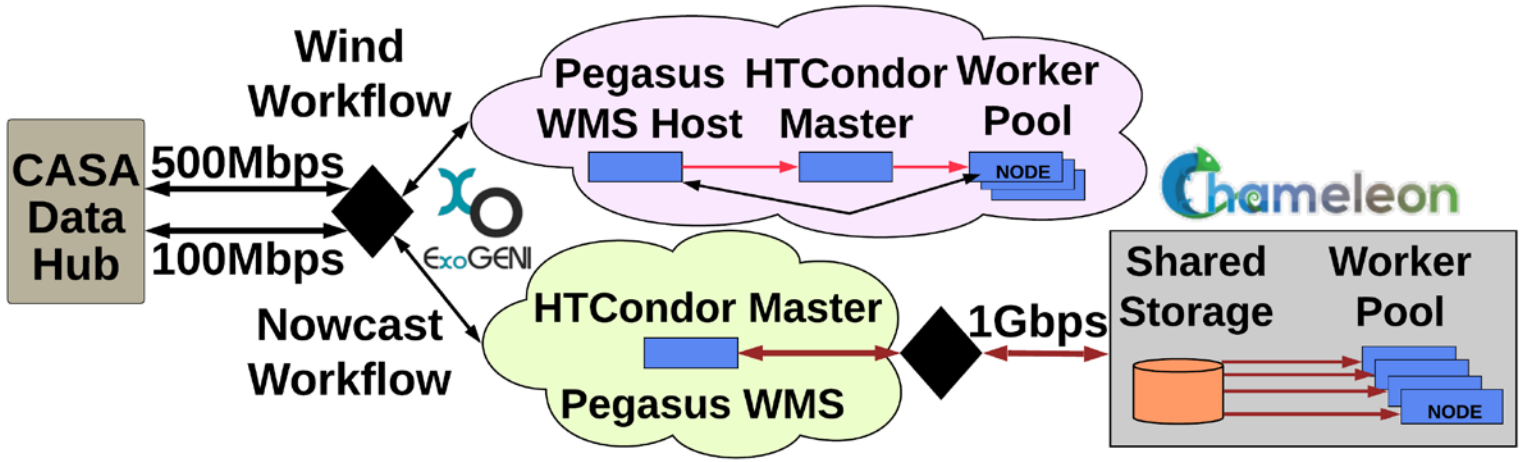
renci

USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

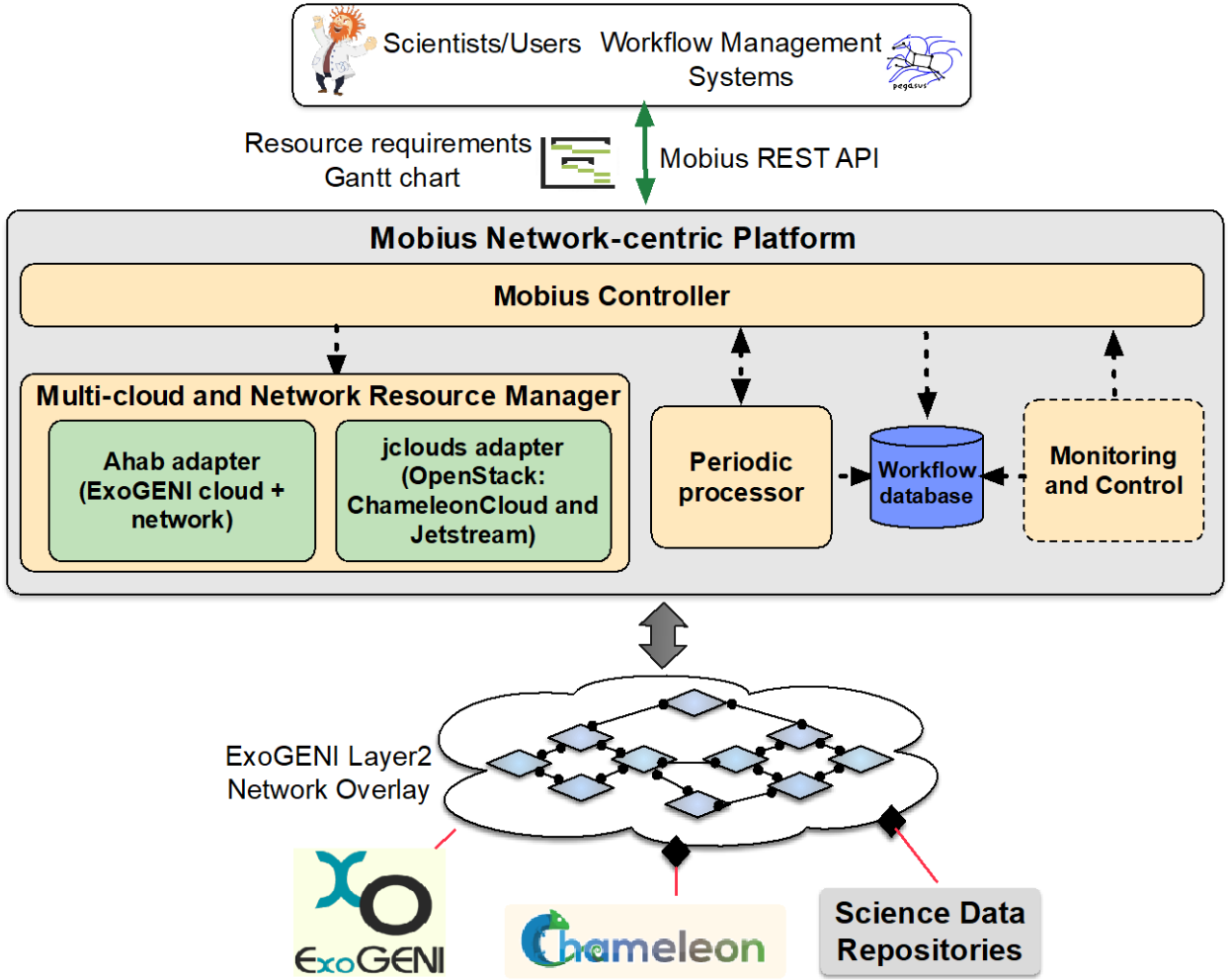
RUTGERS

DyNamo Platform Overview



- High speed data movement via ExoGENI's dedicated layer-2 overlay networks
- Compute and storage resources on both ExoGENI and Chameleon clouds
- Dynamic resource provisioning on ExoGENI and Chameleon clouds
- Workflow instrument with Pegasus WMS and HTCondor

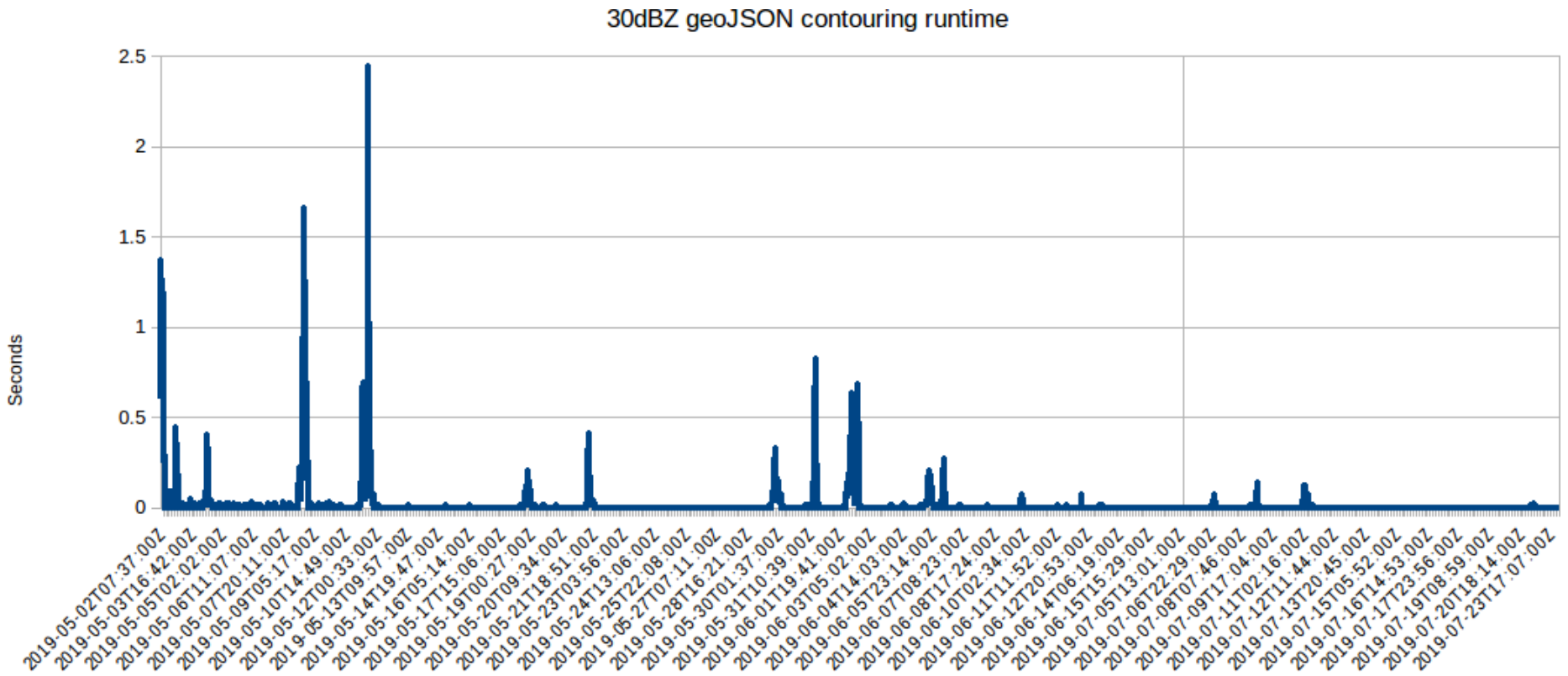
Workflow Provisioning and Orchestration with DyNamo



- Resource requirements are generated using Gantt chart
- Mobius network-centric platform
  - Multi-cloud provision compute and network resources
  - Periodic processor
  - Resource monitoring and control

DYNAMO

The Need for Elasticity



- 90K runs over 2.5 mos
- 1 min, 1 threshold
- Largely near zero runtime, punctuated by notable spikes when widespread weather occurs
- Strongly argues for scalability!

renci

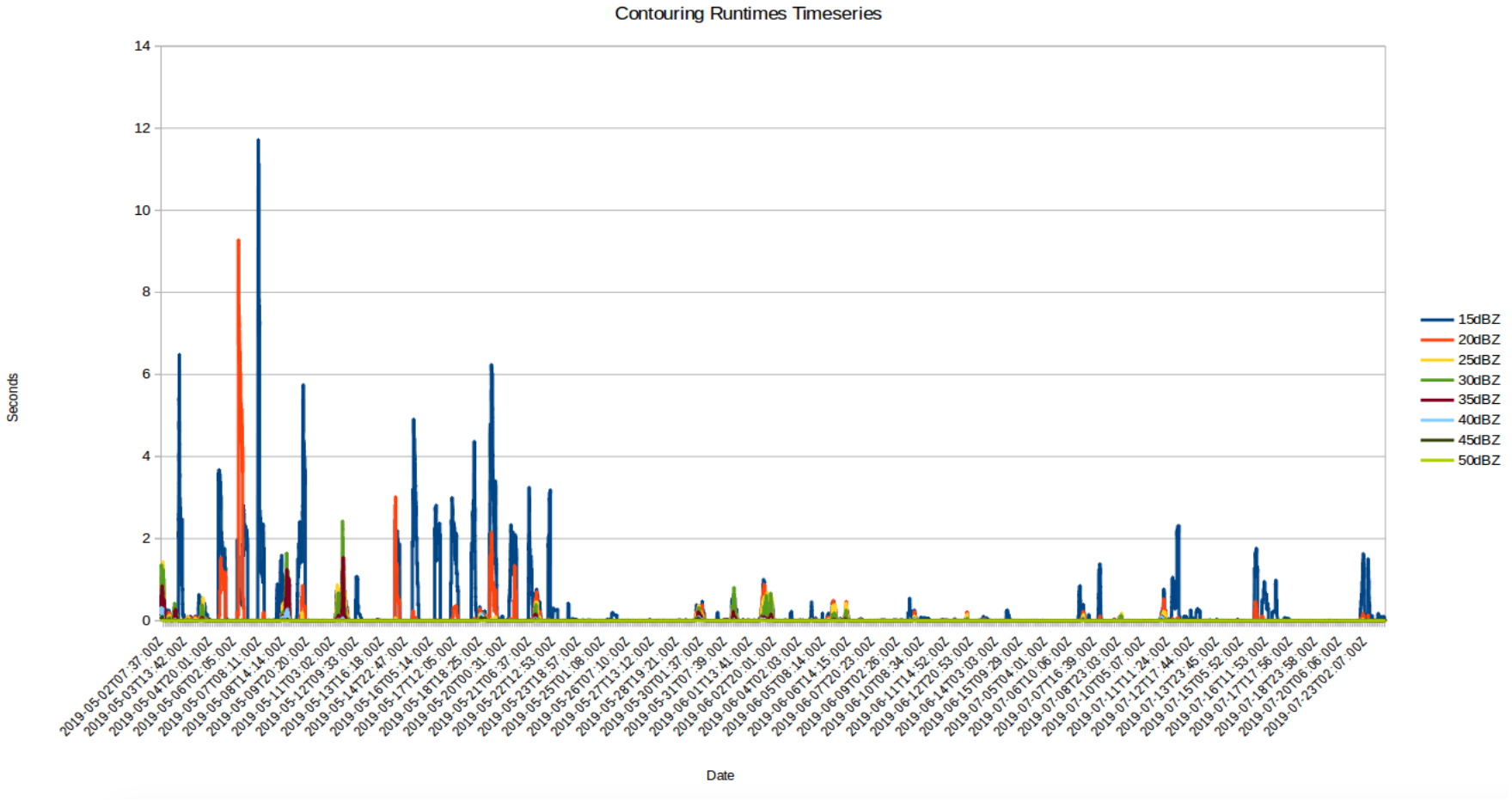
USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

Rutgers

DYNAMO

The Need for Elasticity



- 90K runs over 2.5 mos
- 1 min, 1 threshold
- Largely near zero runtime, punctuated by notable spikes when widespread weather occurs
- Strongly argues for scalability!

renci

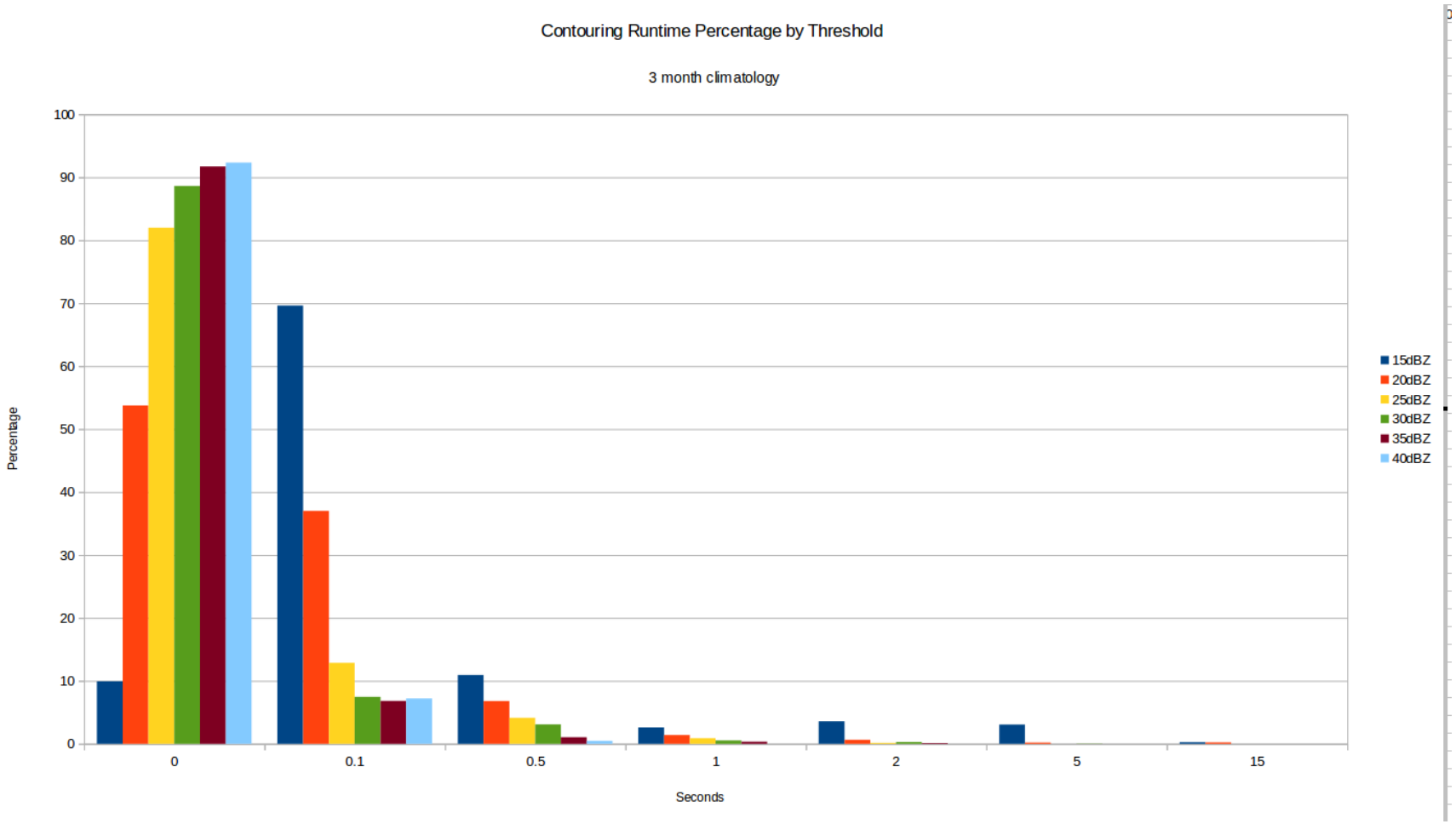
USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

Rutgers

DYNAMO

The Need for Elasticity



- 90K runs over 2.5 mos
- 1 min, 1 threshold
- Largely near zero runtime, punctuated by notable spikes when widespread weather occurs
- Strongly argues for scalability!

renci

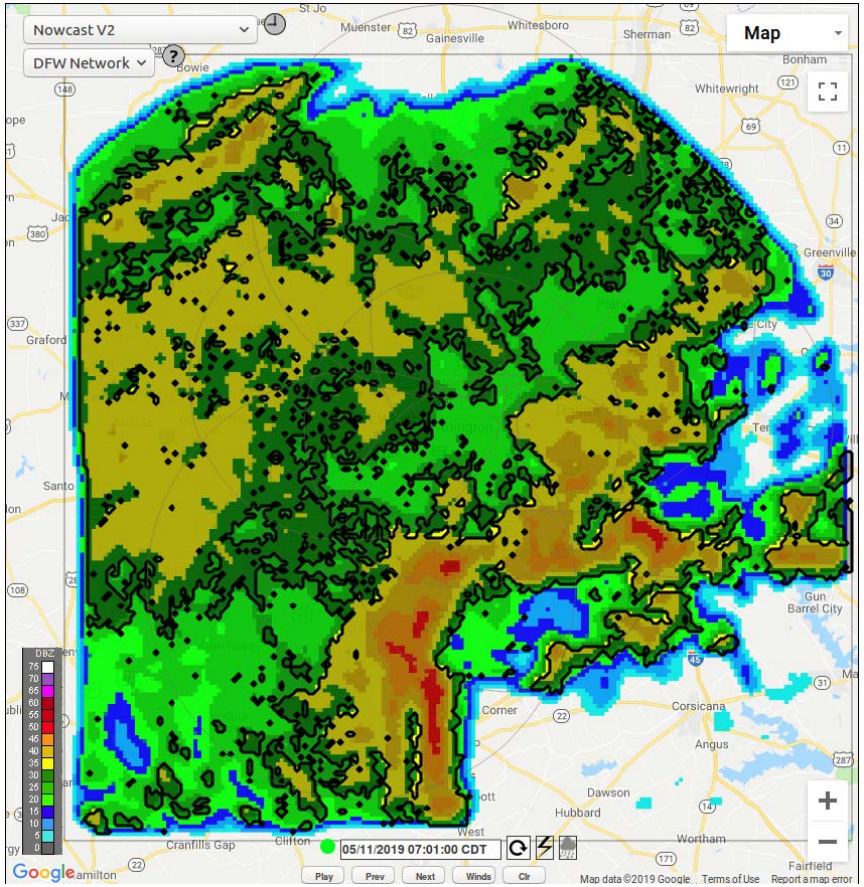
USC Viterbi School of Engineering Information Sciences Institute

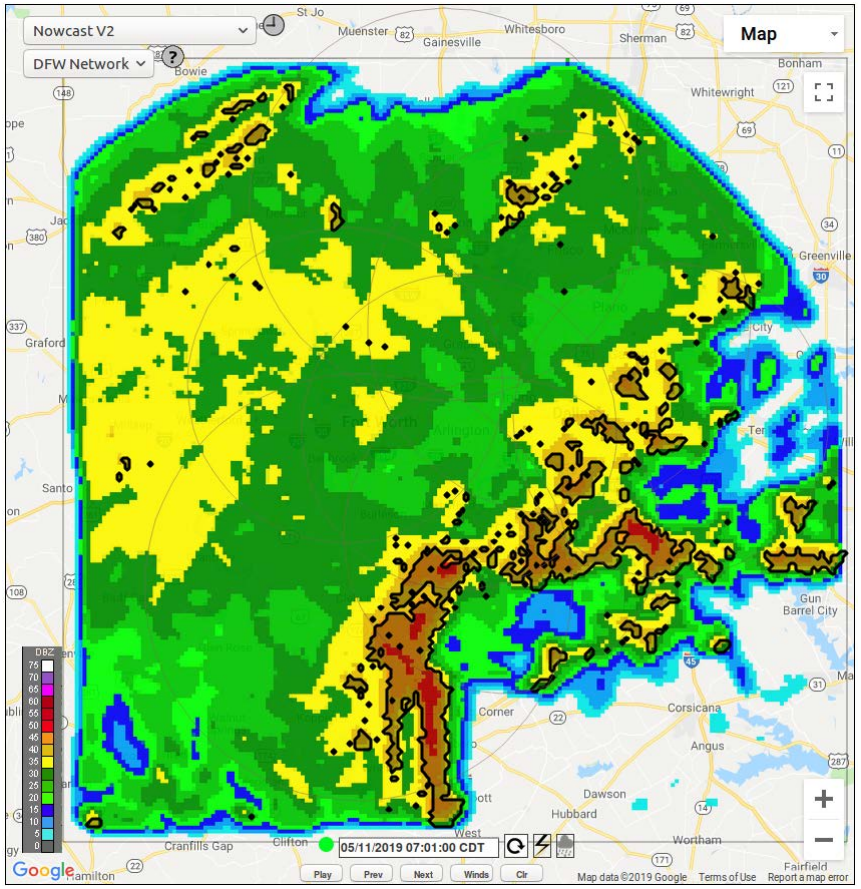
UMass Amherst

Rutgers

DYNAMO

The Need for Elasticity





- 90K runs over 2.5 mos
- 1 min, 1 threshold
- Largely near zero runtime, punctuated by notable spikes when widespread weather occurs
- Strongly argues for scalability!

renci

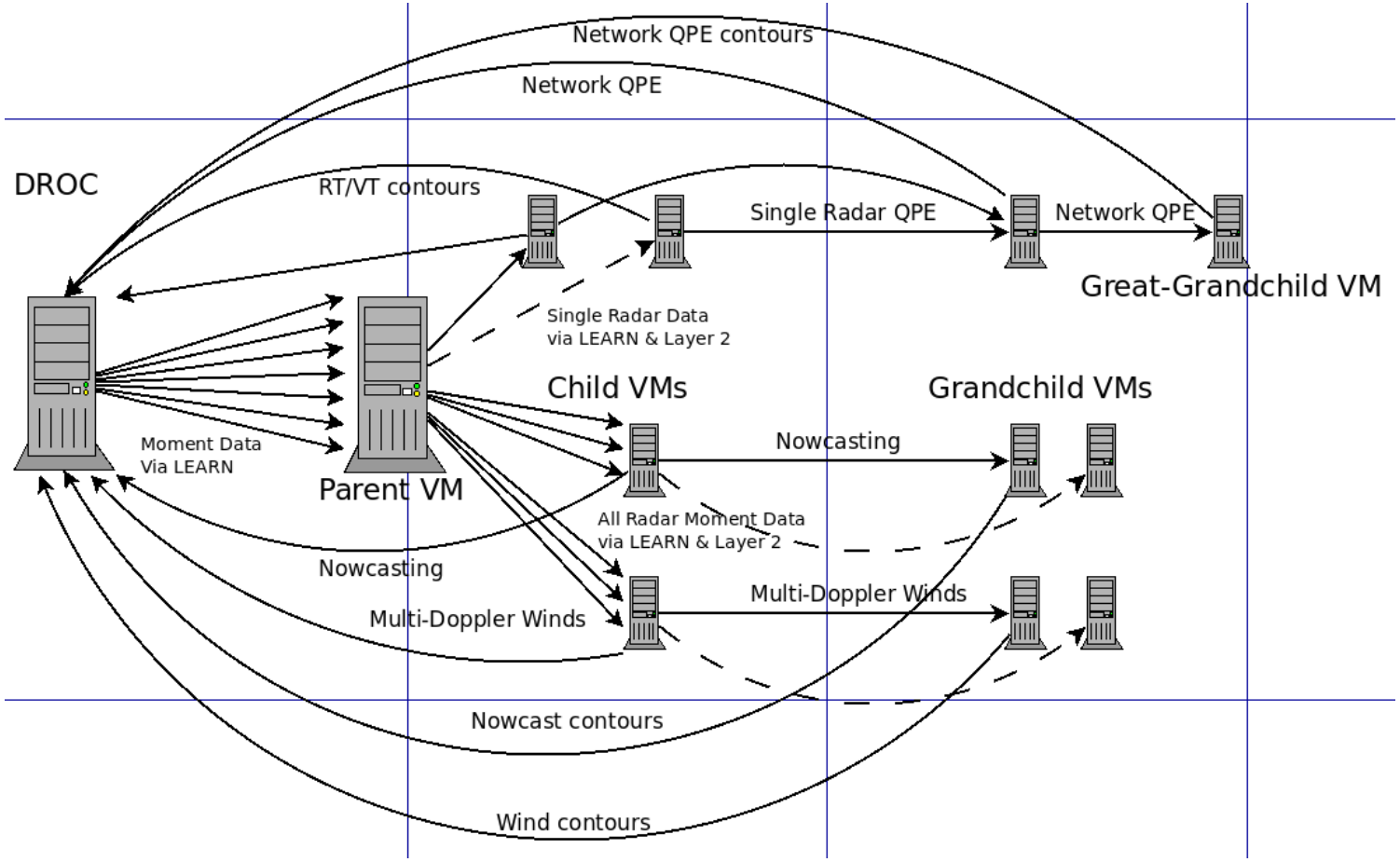
USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

Rutgers

DYNAMO

CASA Workflows Before DyNamo



- Deployed on ExoGENI
- Not scalable with load fluctuations
- Layer 3 only… no layer 2 stitchports
- No WMS to make efficient use of available processing
- Not containerized

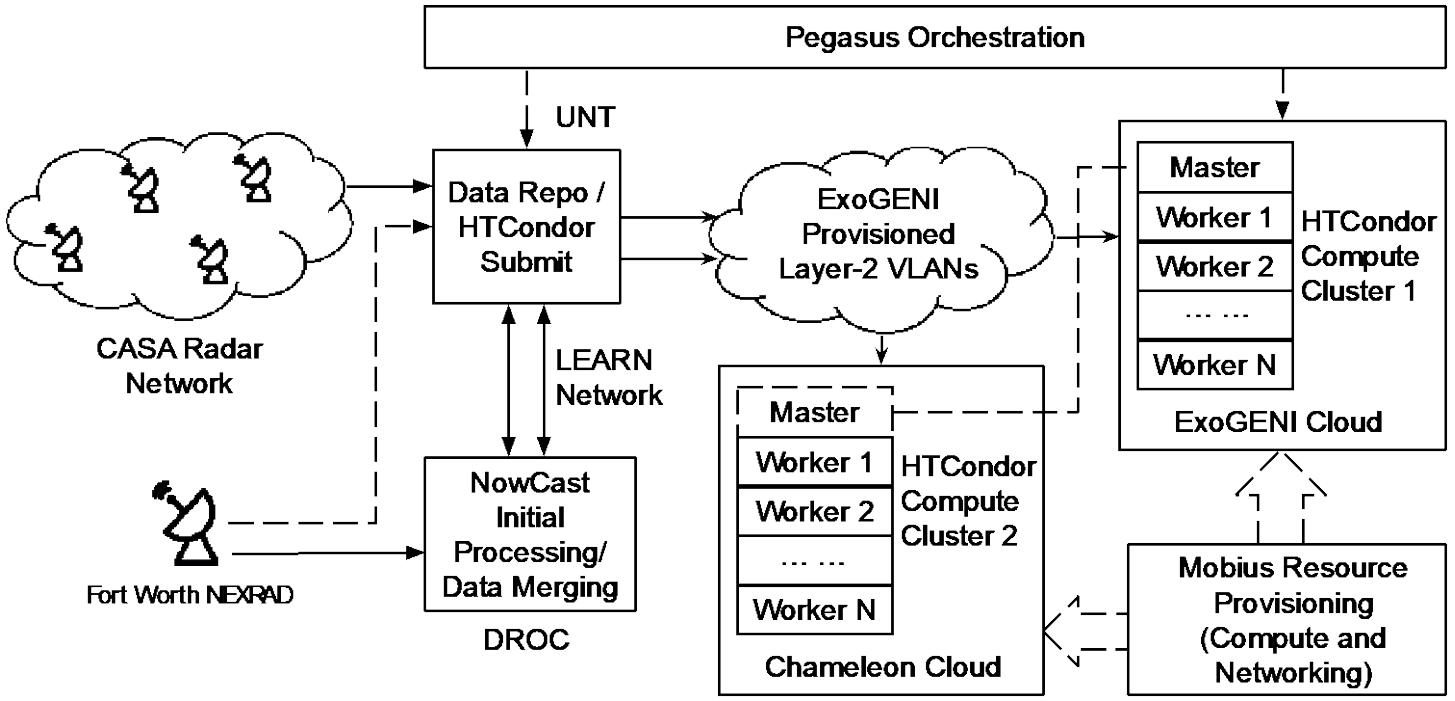
renci

USC Viterbi School of Engineering Information Sciences Institute

UMass Amherst

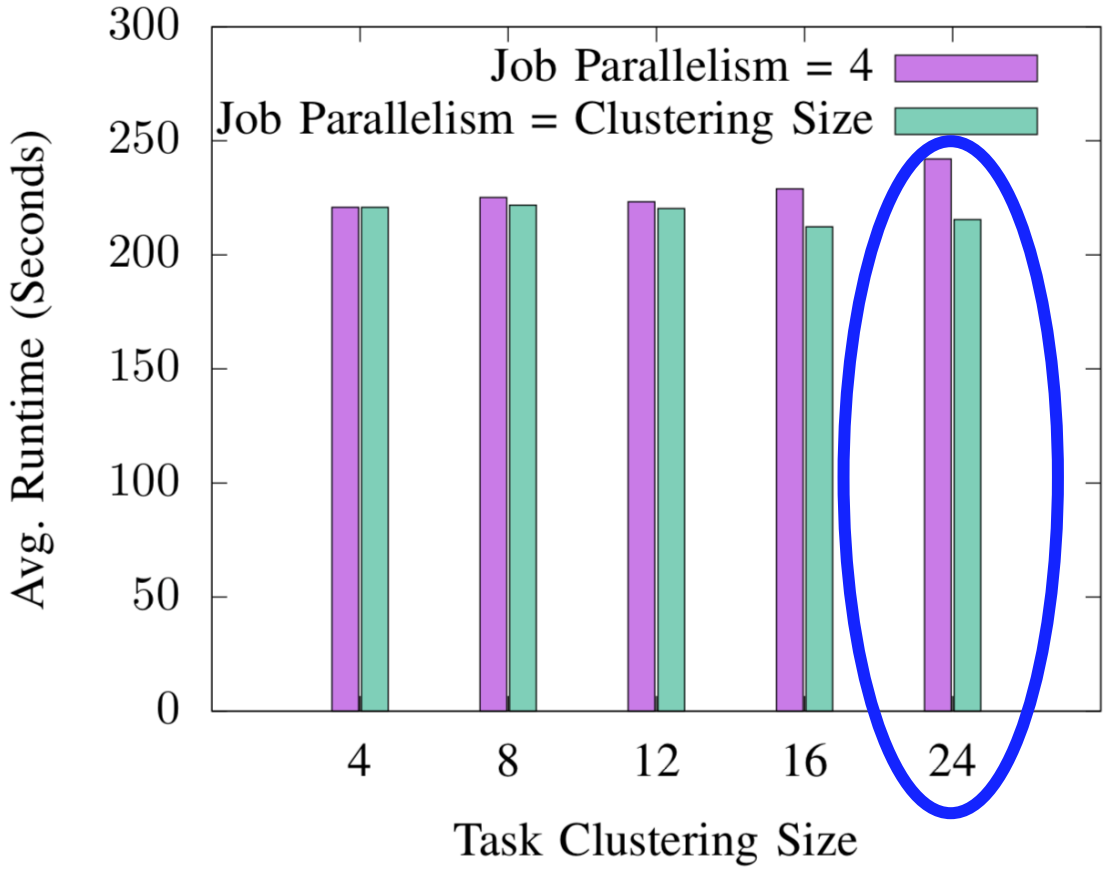
Rutgers

DyNamo Workflow Deployment



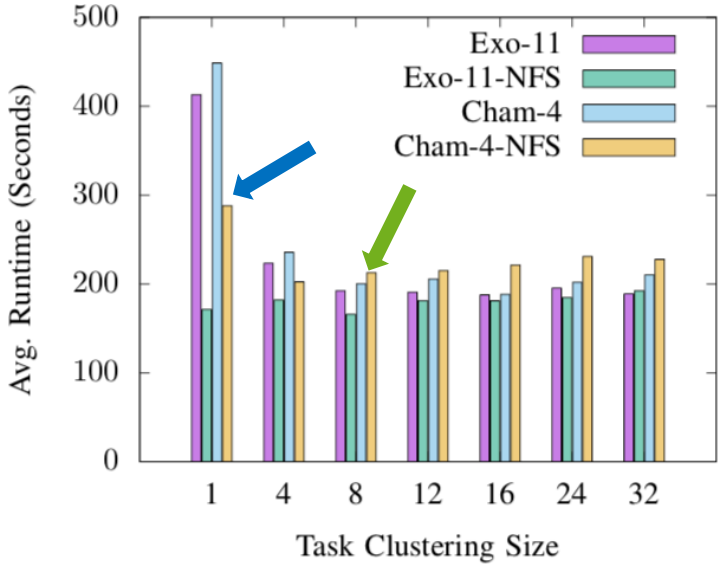
- Prototype
  - Heterogeneous: ExoGENI and Chameleon testbeds
- ExoGENI
  - 11 workers (VMs)
  - 4 cores and 12 GB RAM
  - NFS storage
- Chameleon
  - 4 workers (bare metal)
  - 24 cores and 192 GB RAM
- Connected via 10 Gbps network
- Data repo
  - UNT via L2 stitching port

Evaluation – Effect of Job Parallelism

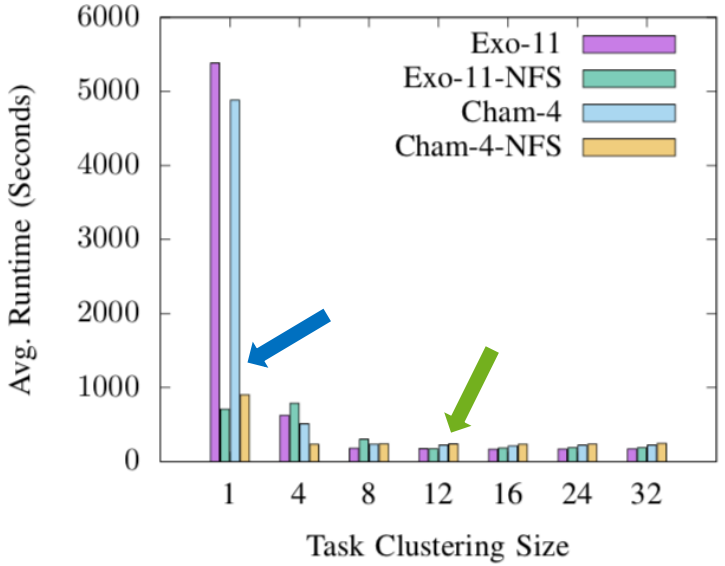


- Run time comparison
  - Parallelism: 4
  - Parallelism: clustering size
- Higher Parallelism gives better performance (less runtime)
- However this gain in performance doesn't justify the additional resource demands.
- In the worst case scenario the increase is ~30seconds.

Evaluation – Nowcast Execution



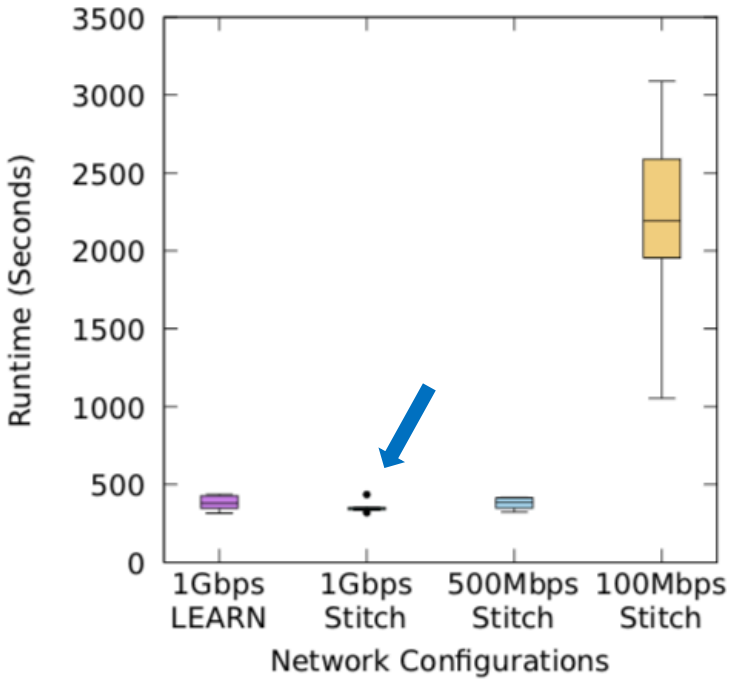
(b) Single Workflow Runs.



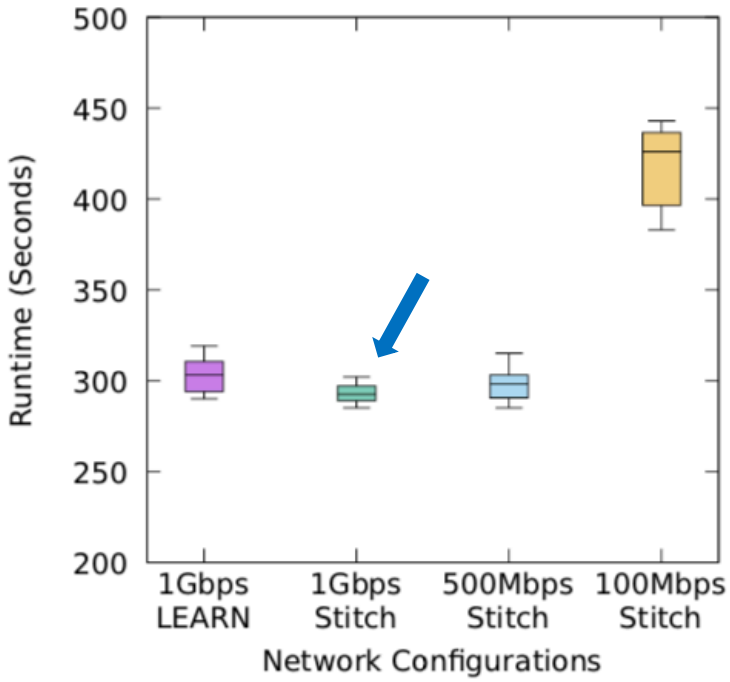
(c) Workflow Ensemble Runs.

- Dedicated Nowcast workflow vs. Nowcast with competing workflows
- NFS Significantly improves execution when cluster size is small
- ExoGENI (11 small workers) tend to be faster than Chameleon (4 large workers)
  - Massive parallelization improves workflow performance (*hint:* IO)

Evaluation – Data Transfer Performance



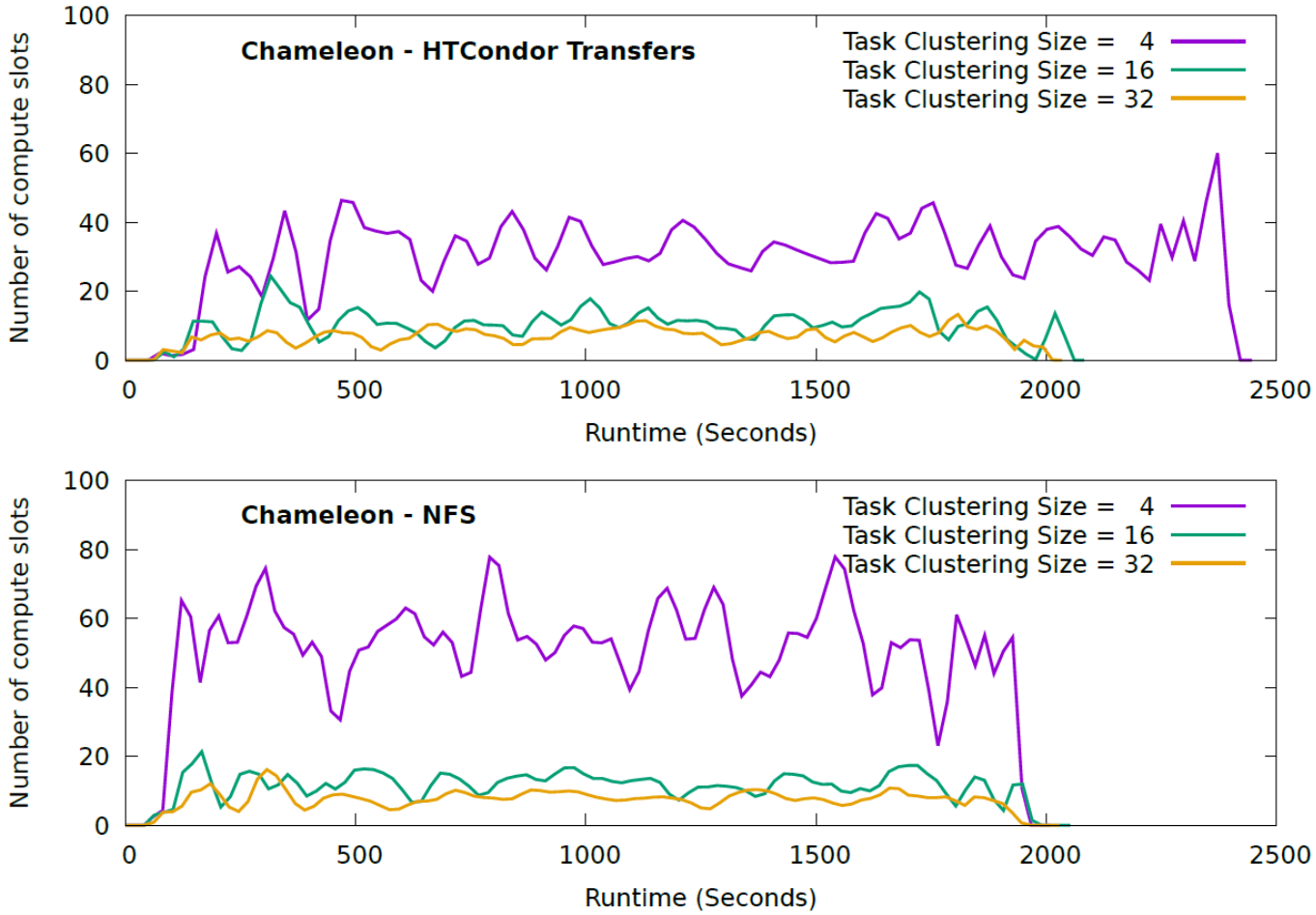
(a) Exogeni - HTCondor Transfers



(b) Exogeni - NFS

- 1 Gbps L2 stitching setup gives best results
- 1 Gbps over LEARN performs similar to 500 Mbps L2 stitching
- 500 Mbps L2 stitching is sufficient for CASA data transfer among facilities

Evaluation – Required Resources by CASA Workflows



- Amount of resources required by compute intensive workflows like Nowcast
- Number of active compute slots for Nowcast
- Chameleon, with HTCondor transfers vs. using NFS
- Clustering of 4 tasks creates high demands (40-80 slots)
- Clustering of 16-32 decreases compute slot demand (<20 slots)

Summary

- DyNamo: a multi-cloud platform with high-performance adaptive computing and networking support for science workflows
- DyNamo enables automation, dynamic infrastructure management
- DyNamo usecase: CASA, that requires on-demand, high-bandwidth paths from CASA central to distributed CI
- Improved CASA capability:
  - Higher data transfer
  - Less workflow runtime
  - Automated resource provisioning and workflow execution

DYNAMO

Funded by the National
Science Foundation
Grant #1826997

Thank you !

Questions ?

renci

USC Viterbi
School of Engineering
Information Sciences Institute

UMass Amherst

RUTGERS