Structured Overlay Networks for a New Generation of Internet Services

Amy Babay[1], Claudiu Danilov[4], John Lane[2],
Michal Miskin-Amir[2,3], Daniel Obenshain[1], John Schultz[2,3],
Jonathan Stanton[2,3], Thomas Tantillo[1], and Yair Amir[1,2,3]

[1]Johns Hopkins University, [2]LTN Global Communications,
[3]Spread Concepts LLC, [4]Boeing Research & Technology

JOHNS HOPKINS UNIVERSITY

Distributed Systems and Networks Lab
www.dsn.jhu.edu

The Internet Revolution

A single, multi-purpose, IP-based network

- The art of design - end-to-end principle:
  - Keep it simple in the middle and smart at the edge
    - Simple middle: best effort packet switching, routing (intranet, internet)
    - Smart edge: end-to-end reliability, naming
  - Allowed dramatic success in adaptability and scaling over the past five decades
- Critical to society and standardized
  - The basic services are not likely to change

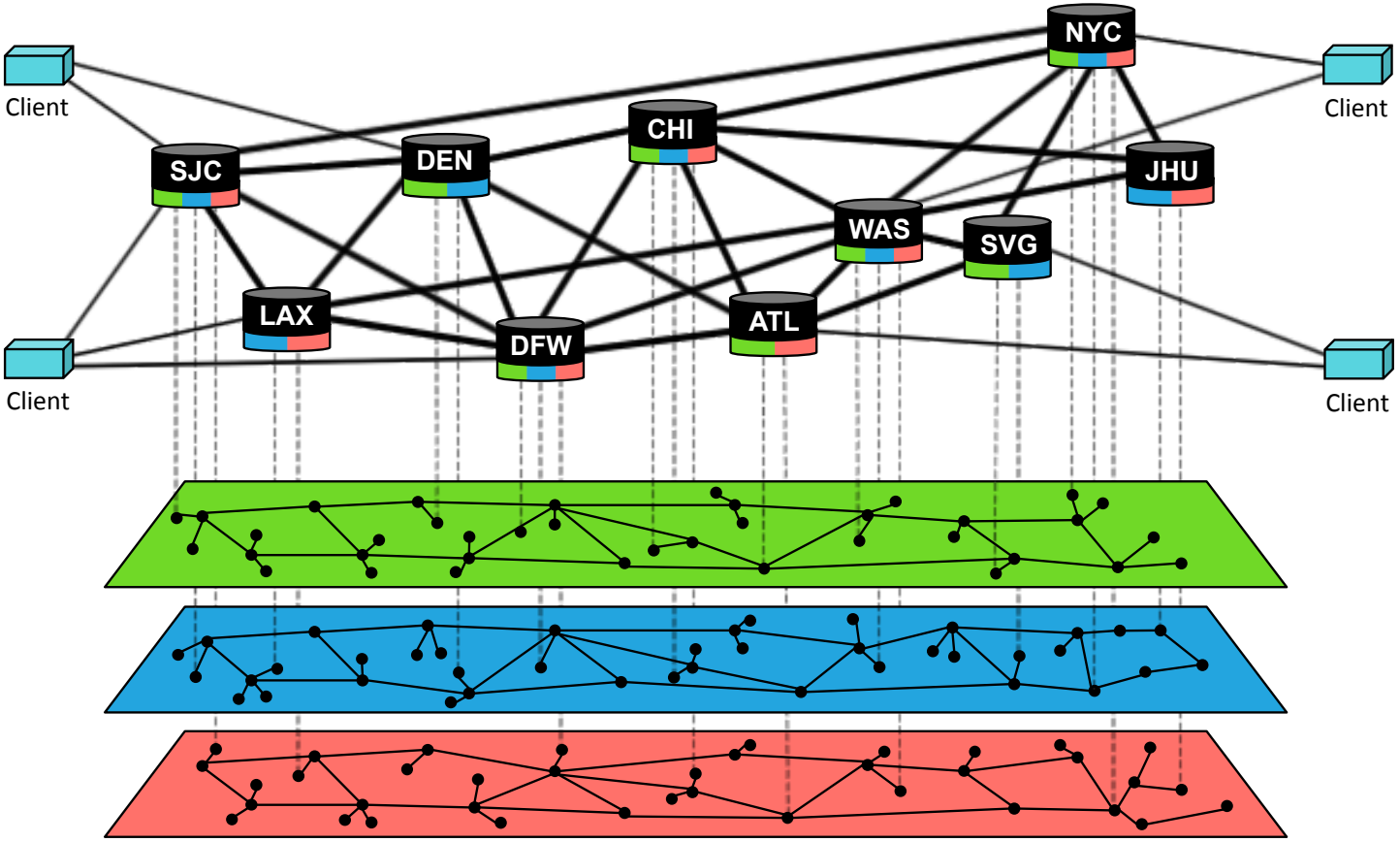
A New Generation of Internet Applications

- Communication patterns
  - From point-to-point – to point-to-multipoint – to many-to-many
- High performance reliability
  - "Faster than real-time" file transfers
- Low latency interactivity
  - 100ms for VoIP
  - 80ms for interactive games
  - 65ms for remote manipulation, remote robotic surgery
- End-to-end dependability (availability, reliability)
  - From "e-mail" dependability – to "phone service" dependability – to "TV service" dependability – to "remote surgery" dependability
- System resiliency, security, and access control
  - From E-mail fault tolerance – to financial transaction security – to critical infrastructure (SCADA) intrusion tolerance

The Structured Overlay Vision

- Key idea: puts processing and context into the middle of the network, providing more flexibility and control
  - At overlay level
  - Underlying network maintains the end-to-end principle
- Three structured overlay principles:
  - Resilient network architecture
  - Overlay node software architecture with global state and unlimited programmability
  - Flow-based processing

Resilient Network Architecture

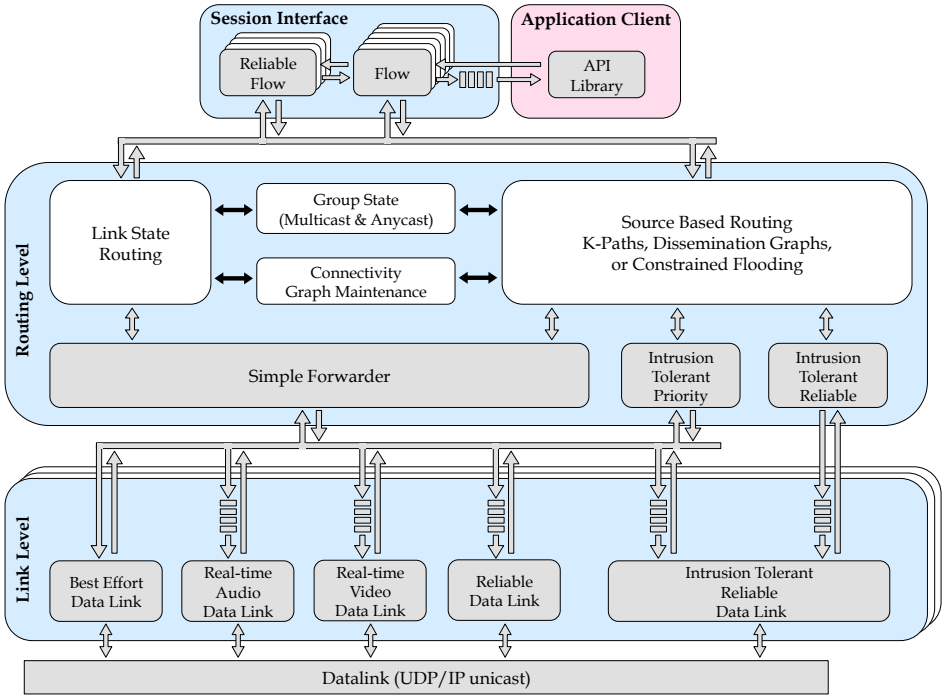


U.S. portion of a resilient structured overlay network with overlay nodes located in strategic datacenters

Overlay Node Software Architecture

- Structured overlay messaging system
  - Running overlay software routers (daemons) on top of UDP as user-level internet applications
  - Using commodity servers in strategic datacenters
- Easy-to-use programming platform
  - API similar to the socket API
  - Additional, seamless API through packet interception
- Deployable
  - Vision partially realized by the Spines messaging system (www.spines.org) and its derivatives

Overlay Node Software Architecture

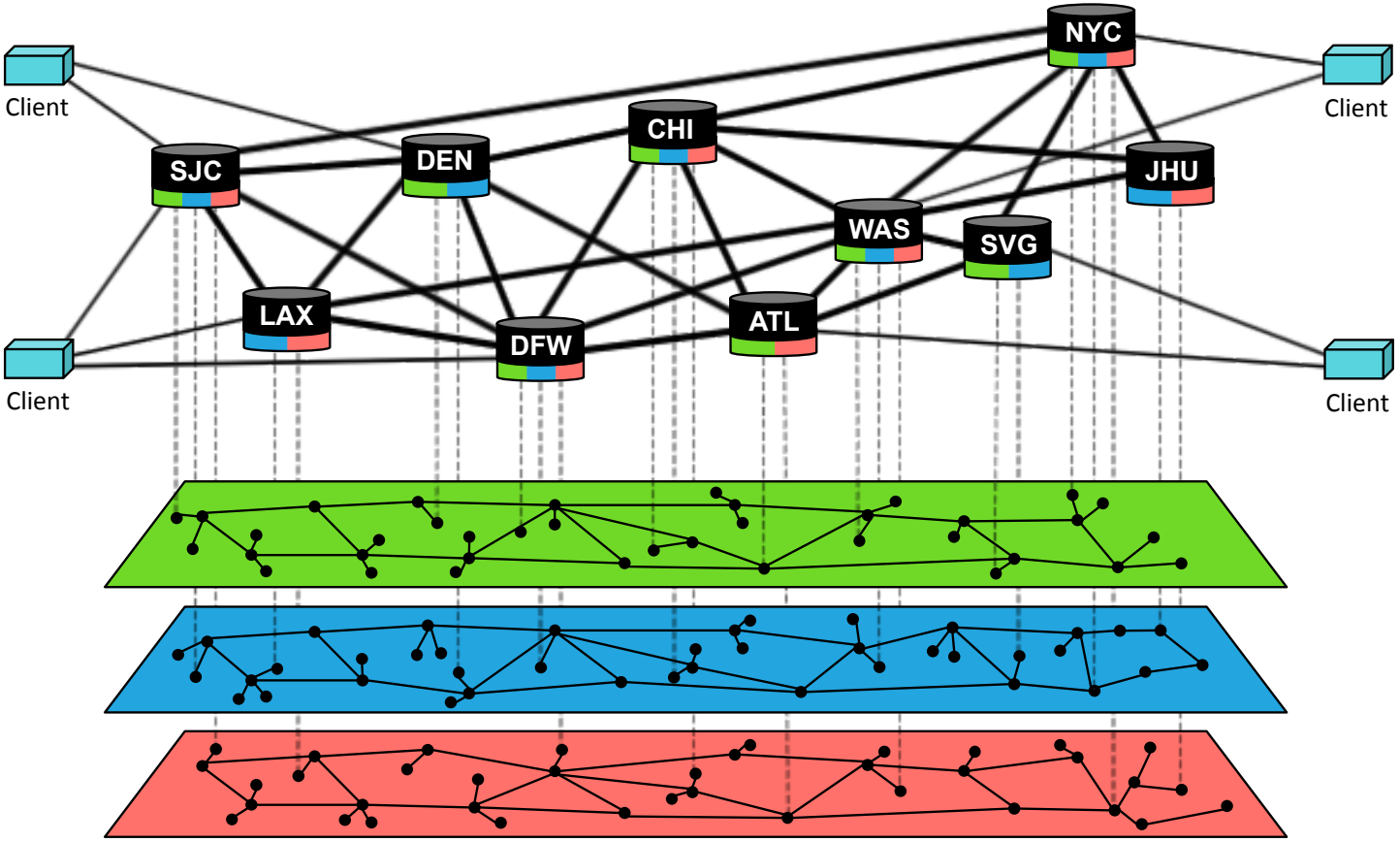


- Global State
  - Possible due to the relatively small number of nodes (e.g. a few tens)
- Unlimited programmability
  - General purpose computers (or clusters) in datacenters
  - Flexible and extensible architecture

Flow-Based Processing

- Leverages flow-specific context
  - Hop-by-hop recovery
  - De-duplication of retransmitted or redundantly transmitted packets in the middle of the network
  - Enhanced resiliency through flow-based fairness
- Allows different services to be selected for different application flows

Structured Overlay Framework



Feasible through a service provider paradigm (just like cloud computing)

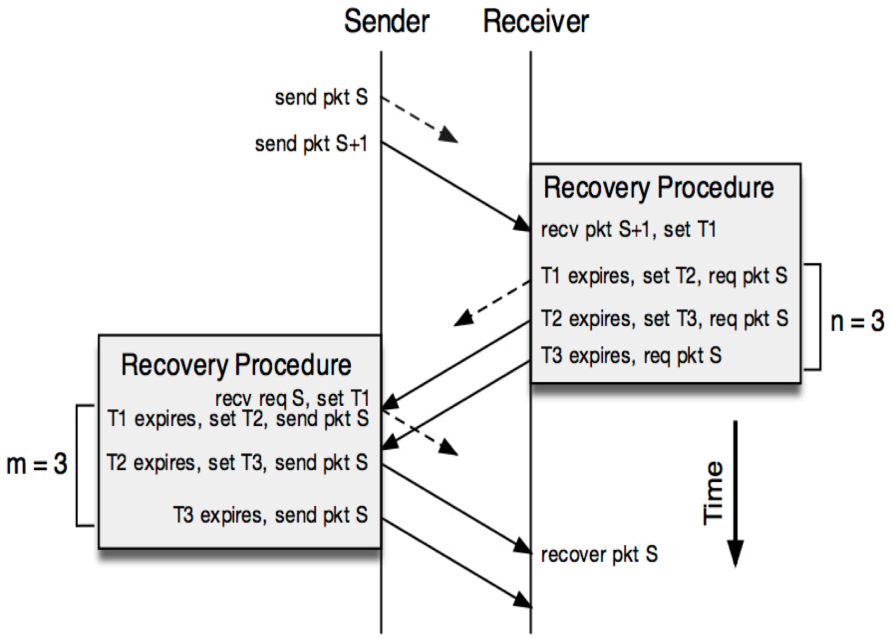
Example Applications

- Streaming (+ interactivity)
  - Broadcast-quality video transport
  - Live broadcast-quality video transport
  - Real-time remote manipulation
- Global Monitoring and Control
  - Resilient monitoring and control
  - Intrusion-tolerant monitoring and control
  - Monitoring and control of critical infrastructure

Live Broadcast-Quality Video Transport

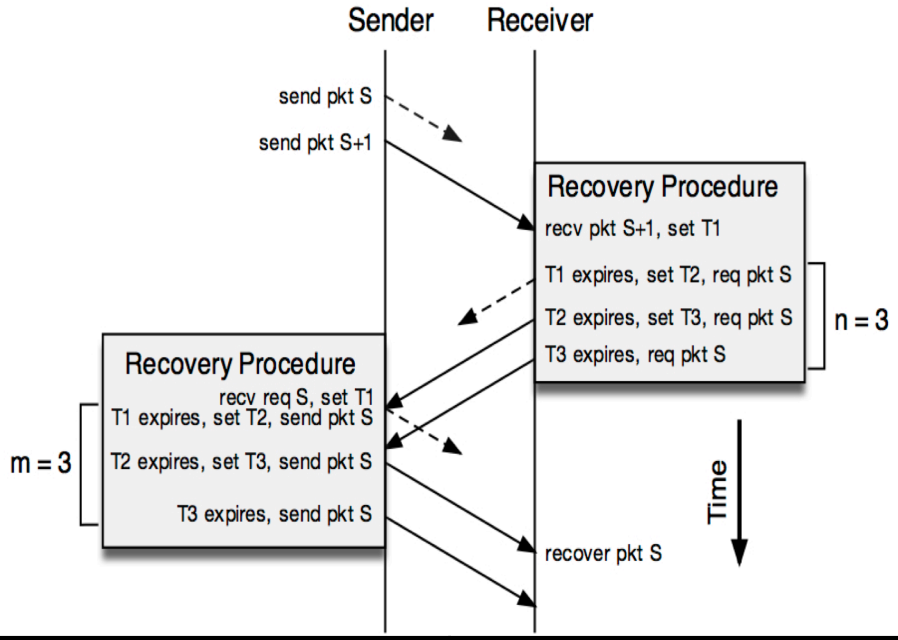
- Requires high availability, multicast, reliability (99.999%), and timeliness (~200ms one way)

Live Broadcast-Quality Video Transport



NM-strikes overlay link protocol: guaranteed timeliness, “almost reliable” delivery

Live Broadcast-Quality Video Transport



Network packet loss on one link (assuming 66% burstiness)	Loss experienced by flows on the LTN Network
2%	< 0.0003%
5%	< 0.003%
10%	< 0.03%

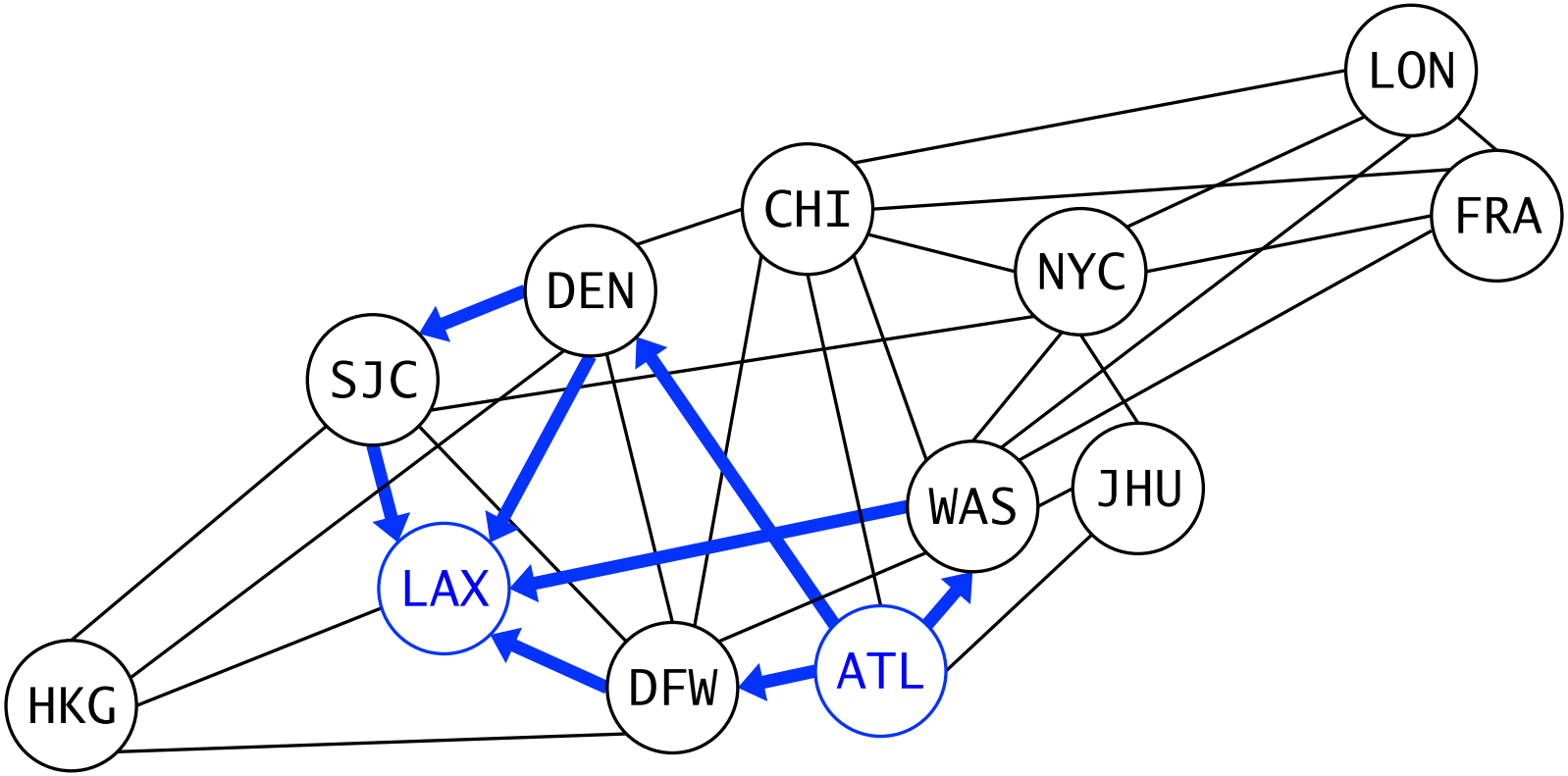
Near Future: Remote Manipulation

65ms **one-way** latency requirement

40ms one-way propagation delay across North America

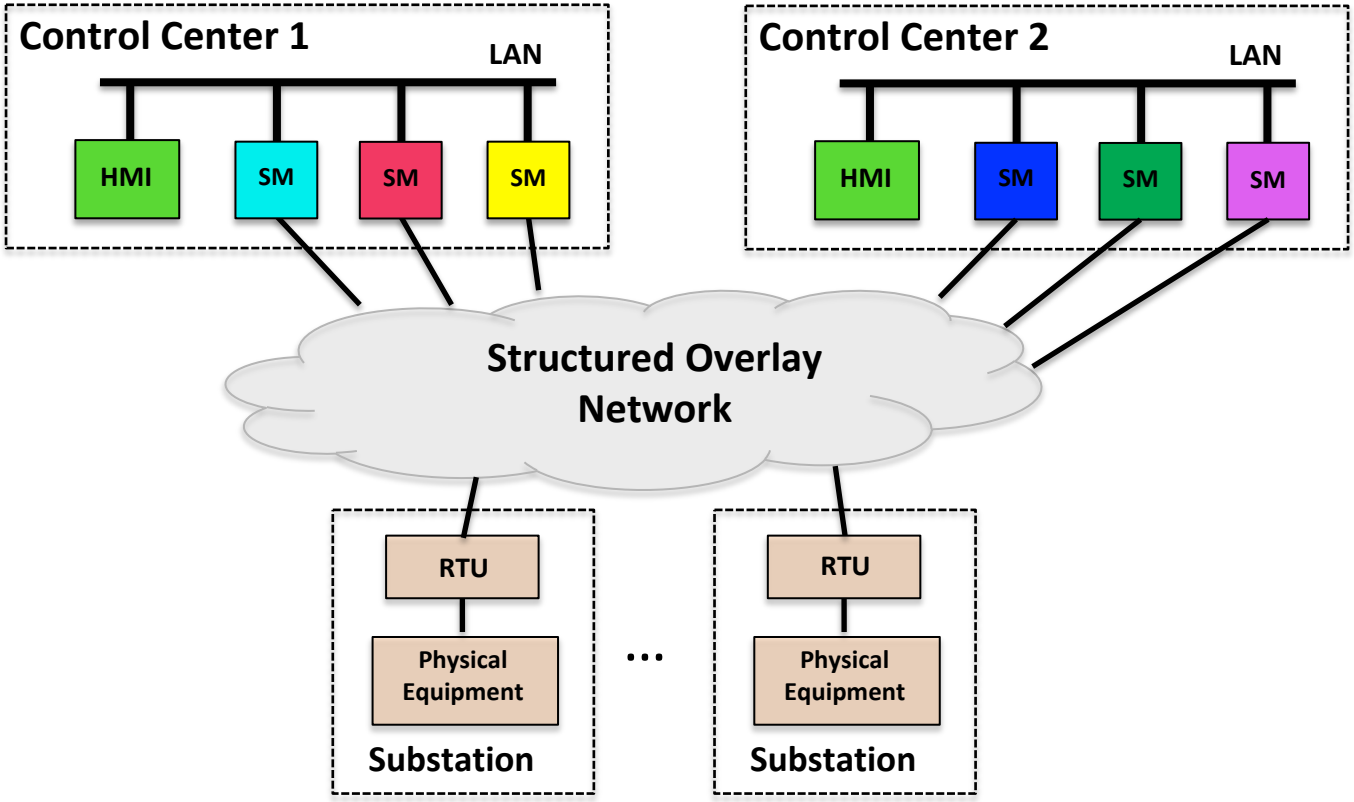
Near Future: Remote Manipulation

- Dissemination graphs with targeted redundancy



Increase redundancy in problematic areas of the network

Near Future: Intrusion-Tolerant SCADA



SCADA for the power grid requires both extreme resilience and guaranteed timeliness (on the order of 100-200ms)

Putting it in Context

- P2P Overlays
  - Generally include a large number of peers and use self-organizing server-less architectures
  - Investment associated with structured overlays offers better performance and resilience
- MPLS
  - Protected virtual circuit capability over single provider IP network
  - Provides bandwidth allocation, traffic class prioritization, multicast
  - Routers provide packet forwarding; cannot support higher-level services that require significant processing and state maintenance
- Software-Defined Networking
  - Offers enhanced network programmability
  - Focuses on separation of control and data planes and improving network management through control-plane innovation

Beyond: Unlimited Potential for a New Generation of Internet Services through Structured Overlays

- Network service that is: authenticated, authorized, admission-controlled, and time-guaranteed, with multicast capabilities
- Seamless support for existing Internet services
  - Obtain core overlay benefits with no application changes through interceptors
- New services
  - Taking advantage of advanced in-network processing capabilities (compound flows)
- For a price…