Workshop Outbrief

Workshop on Prototyping and Deploying Experimental Software Defined Exchanges (SDXs)

Washington DC - June 5-6, 2014

DRAFT

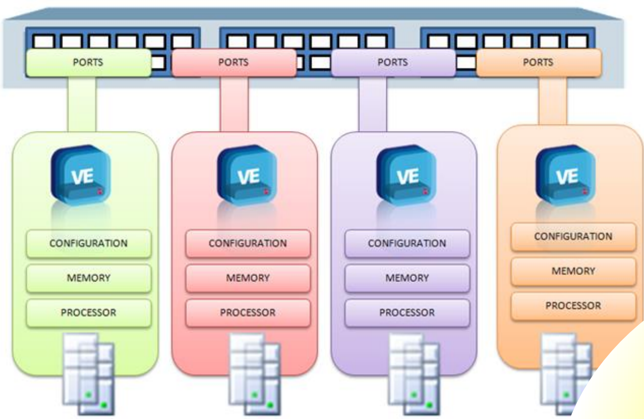
Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

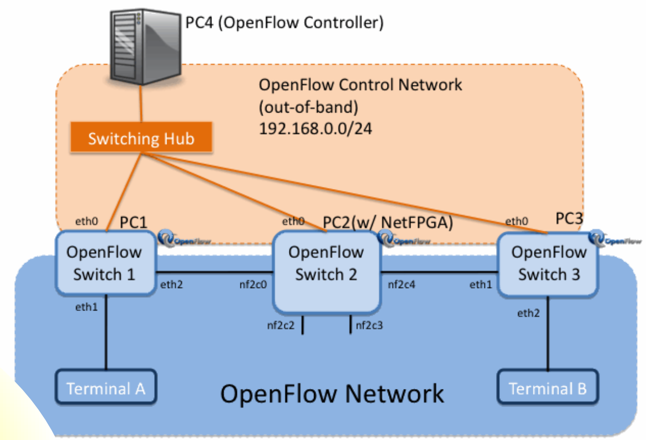
Glossary

- **Peers**: Entities/domains that decide they want to communicate for mutual benefit
- **Resources** include real or virtual computers / storage / networks / instruments / things, or one or more composed sets of higher-level services (middleware, etc.).
- **SDX** (Software Defined Exchange): A real or virtual "meet me" point, where peers meet to communicate, each with its own policies
- **SRP** (Shared Resource Provider): A facility with resources to share. Standard APIs let others access/use the offered resources.
- **SDS** (Software Defined Slice): Collection of resources from SRPs, possibly in different domains, with connecting networks and SDXs
- **Software Defined Infrastructure (SDI):** The collection of SRP provided resources plus networks and SDXs that users / applications can utilize to build end-to-end multi-domain SDSs
- **AAA**: Interoperable Authentication, Authorization, and Accounting software that may be employed when building SDSs

The Internet is Rapidly Evolving

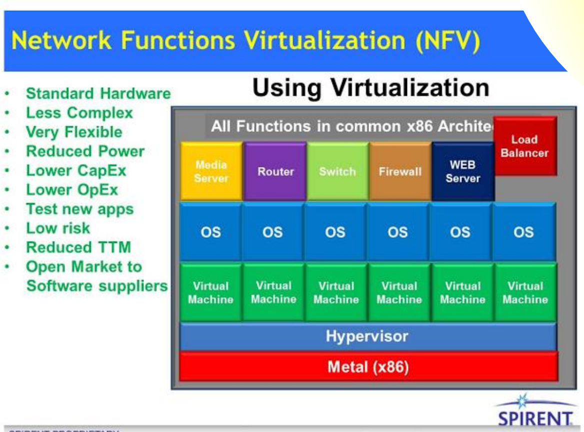


Multi-tenant Datacenters

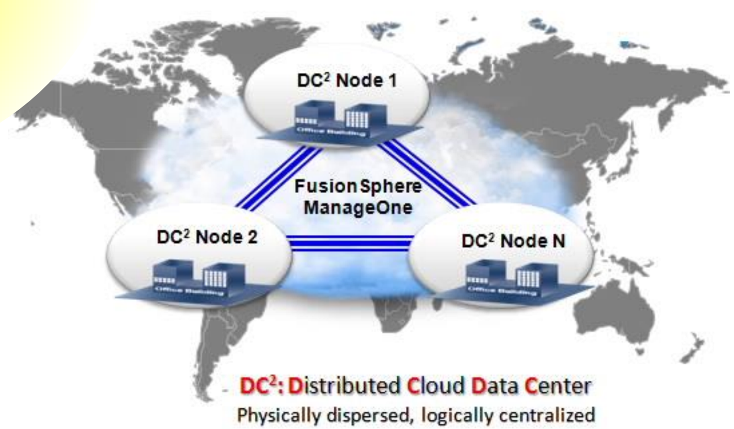


Software Defined Networks

Major trends are converging



Network Functions Virtualization (NFV)



Distributed Datacenters

Exchange Points are (again!) Critically Important

- FIX
- CIX
- NAP
- NGIX
  - MAN LAN
- GOLE

*Interdomain **exchange points** are critical elements in effecting any large-scale network transformation*

- *Technical*
- *Economic*
- *Political*

Next sequel may have started with Nick Feamster's talk at the NITRD SDN workshop (NSF, Dec 2013)

Courtesy Steve Corbato

Software Defined Exchanges (SDXs)

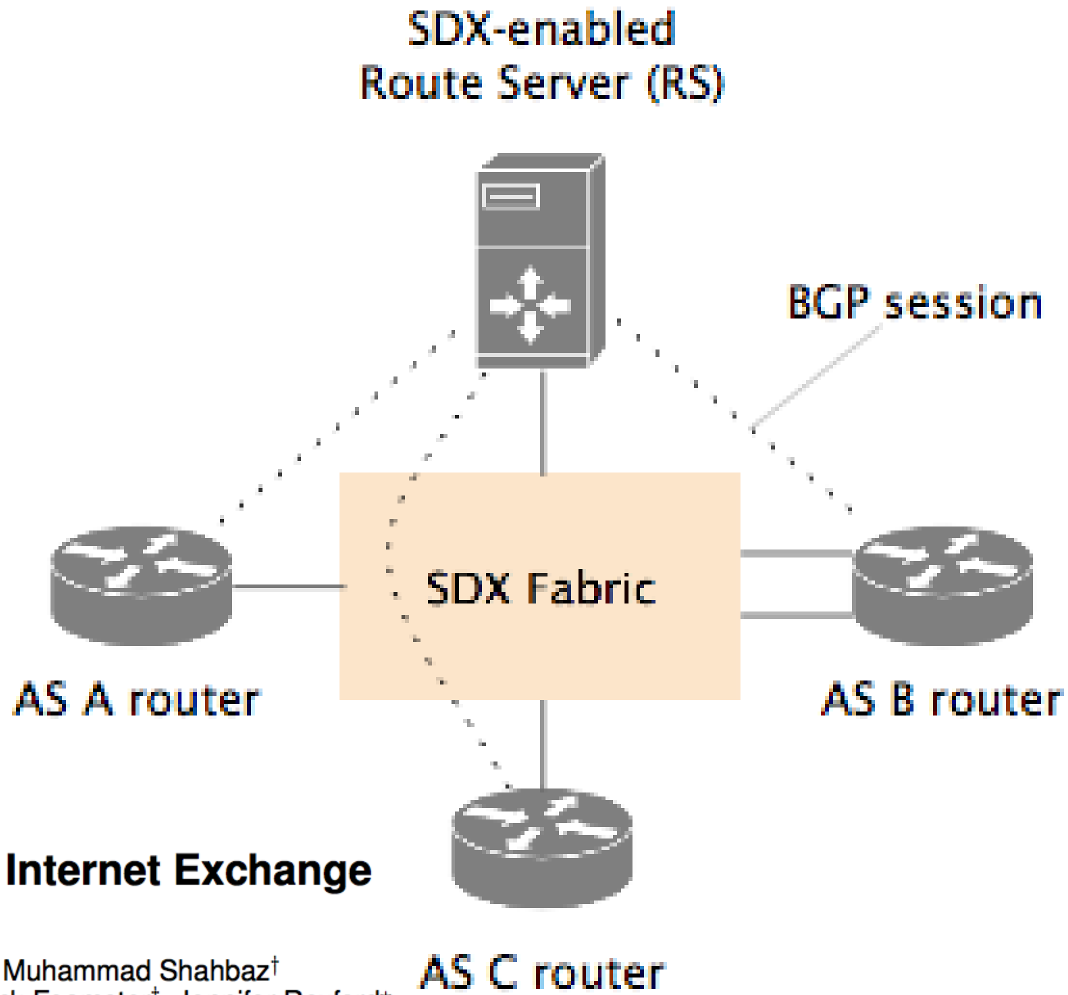
- "Near-term" SDX – pure connectivity/ROUTING
  - Layer 3 (IP) – e.g., connect AS's
  - Layer 2 (Ethernet) – e.g., multi-domain circuits
  - SDN – connect SDN islands
- "Advanced" SDX – with compute/storage
  - Connect SDI islands
  - Compute / storage / network / instruments
  - GENI as an early instance

Layer 3 SDXs – BGP / Policy

BGP routes for RS

prefix	received
p1	***C***, B
p2	***C***, B
p3	***B***, C
p4	***C***
p5	***A***

elected routes

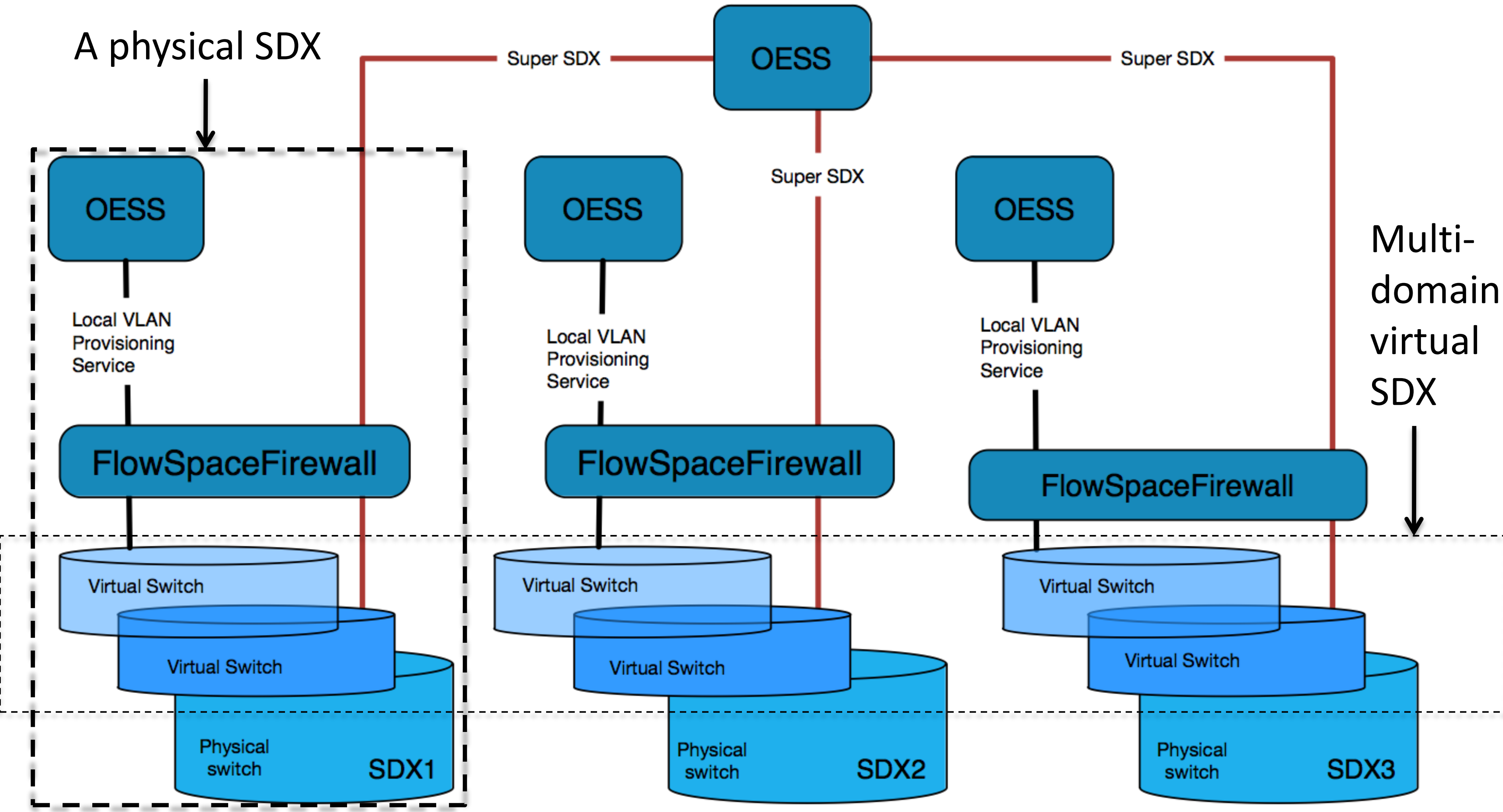


SDX: A Software Defined Internet Exchange

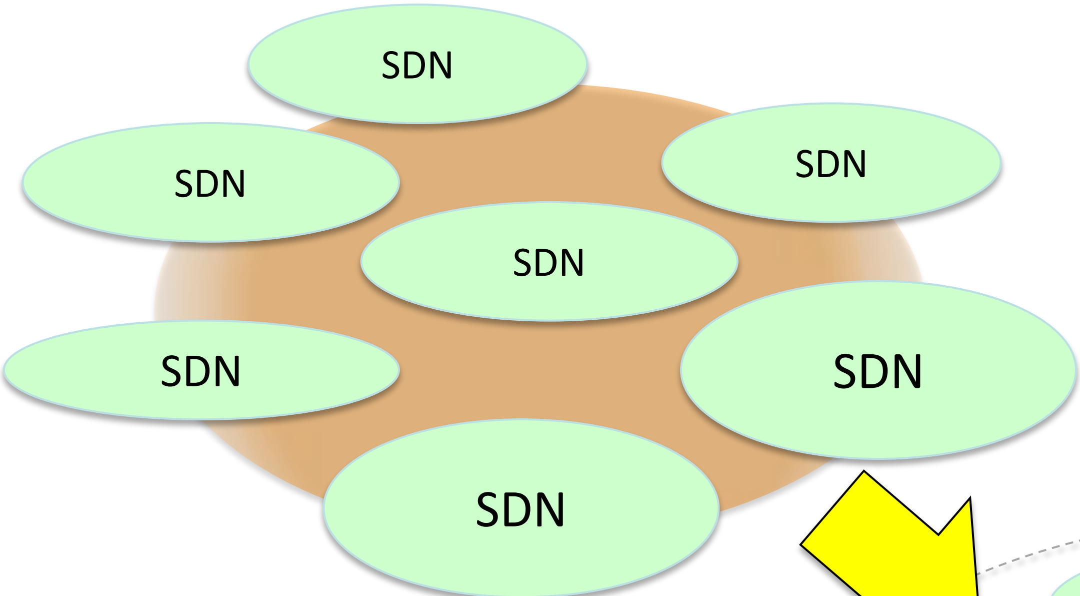
Arpit Gupta†, Laurent Vanbever*, Muhammad Shahbaz†
Sean P. Donovan†, Brandon Schlinker‡, Nick Feamster†, Jennifer Rexford*
Scott Shenker◇, Russ Clark†, Ethan Katz-Bassett‡
†Georgia Tech *Princeton University ◇UC Berkeley ‡Univ. of Southern California

Layer 2 SDXs – Multi-domain Ethernet circuits

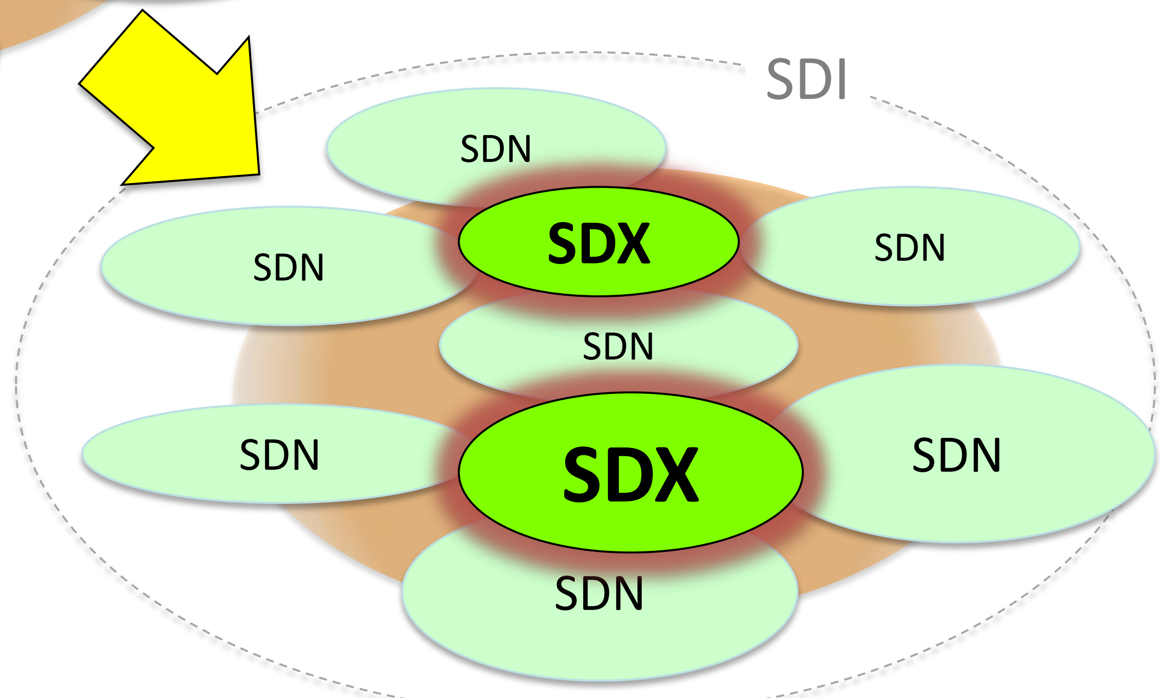
Courtesy Eric Boyd / Internet2



SDN Multi-domain SDXs

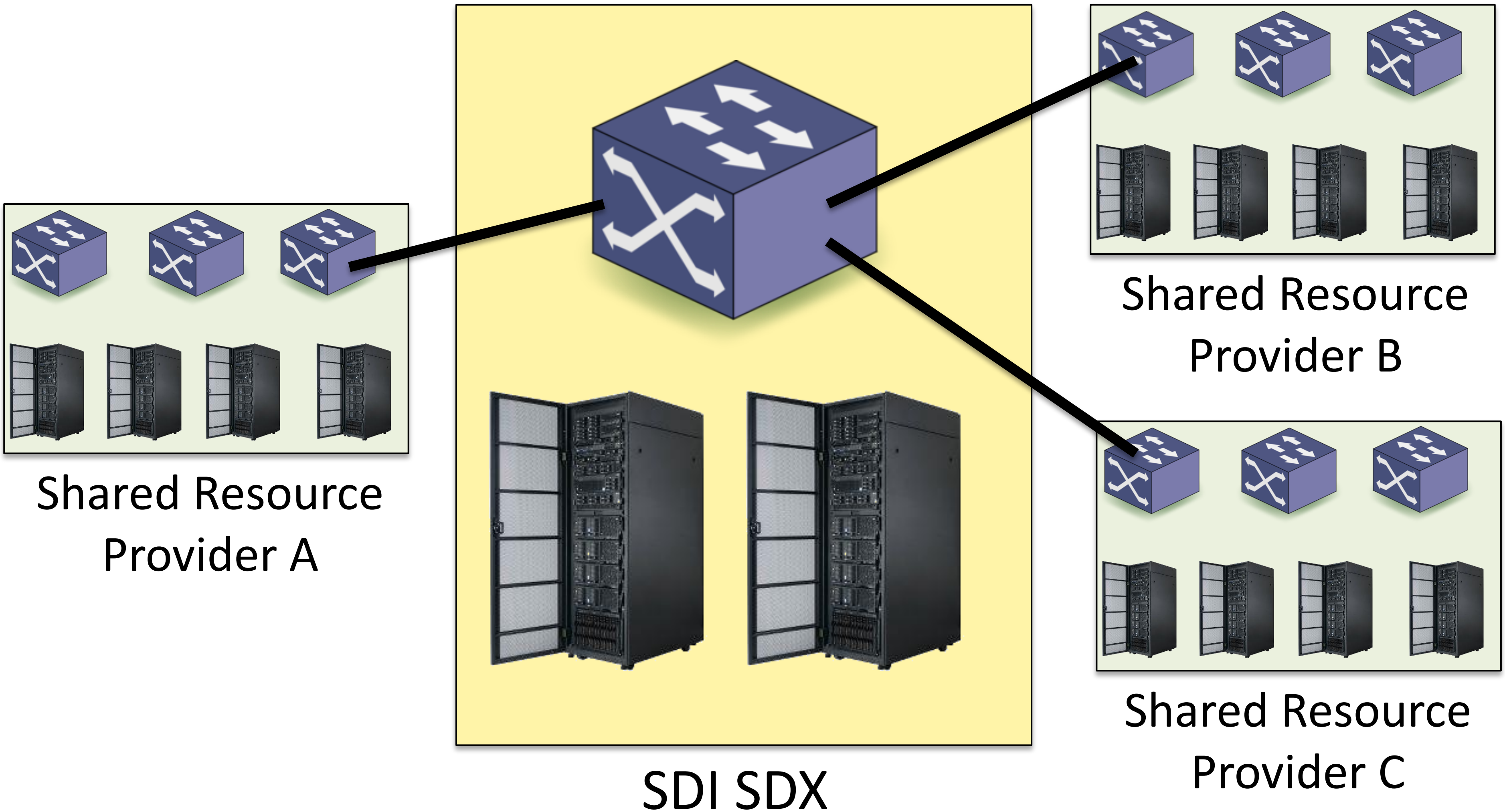


Today: "SDN islands"
GENI slices & VLAN stitching
help point the way



Next Step: Add SDX's
Build a "Rev 0" control plane,
run native next-gen apps
and scientific instruments
spanning multi-domain SDNs

Software Defined Infrastructure (SDI) SDXs



SDX Functionality – Some Ideas

- A place where peering domains come together based on SDX-defined rules of engagement
- Supports establishment of inter-domain connectivity/routing and/or SDSs by applications – supports exchange of information needed to achieve these
- Enforcer of individual domain policies but not itself a policy definer (with exceptions)
- A broker of inter-domain and inter-SRP trust
- A marketplace for services – local and remote
- May be an SRP
- May be virtual or physical
- Peering domains need not be physically connected
- Variation/differentiation among SDX's is useful, but some common services are necessary

What problems are we trying to solve: Why are we building the SDX?

- Increasing complexity of global multi-domain computing / communications environments
- New paradigms such as SDN, SDI, (distributed) clouds, virtualization/slicing, big data
- Applications that need to utilize resources in multiple domains
  - Science requirements from different discipline groups – workflow, requirements, and network needs for these
  - For R&E, end to end is important – e.g., “hand fashioned circuits for astronomy”
- Declarative control of inter-domain path end to end – BGP doesn’t do this
- Potential to leverage awareness of multi-domain network conditions
- Security included in the infrastructure
- Rich environment for researchers to control their experimental infrastructure needs

Outline

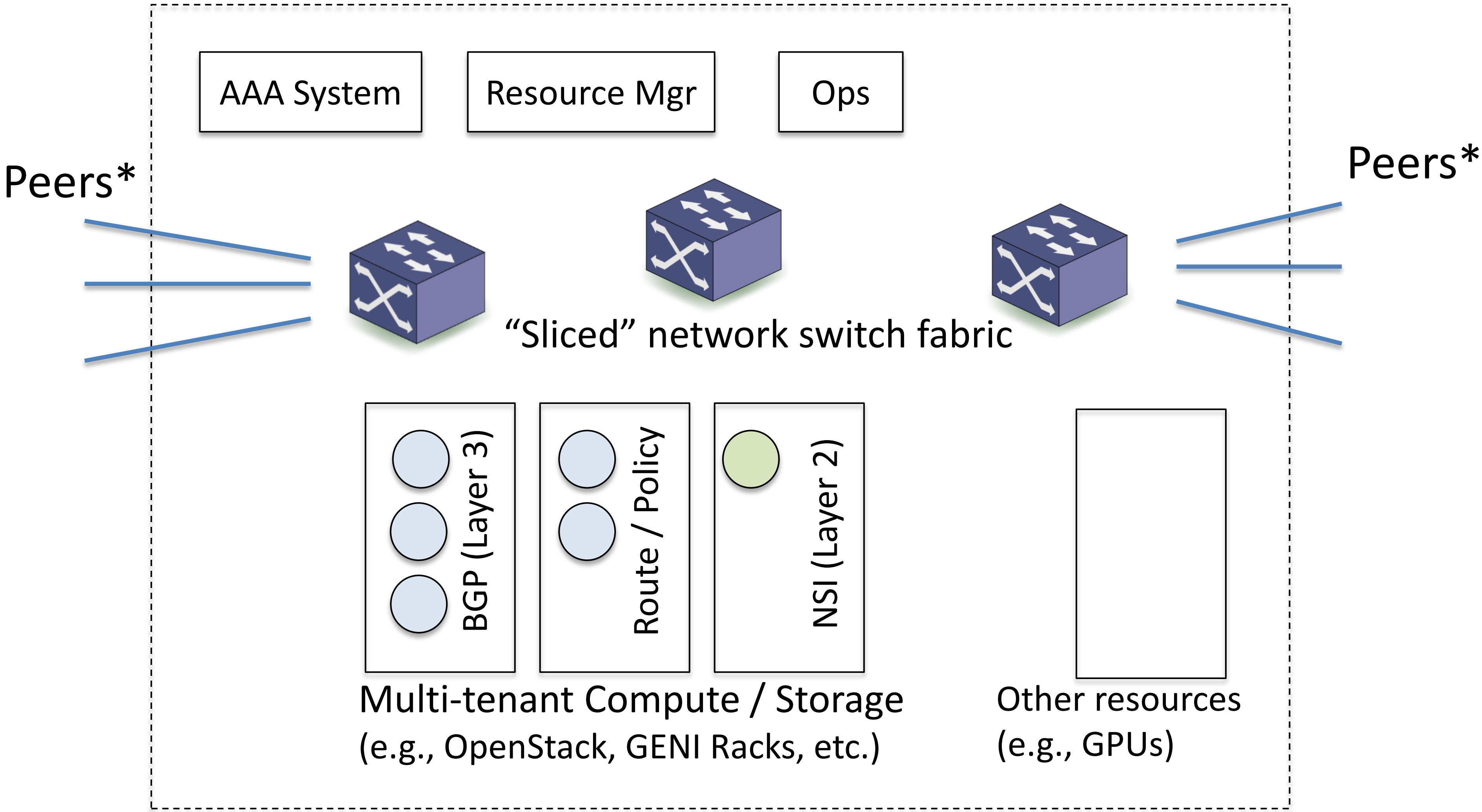
- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

All SDXs Require Local Compute/Storage

- Layer 3 SDX needs route servers, policy servers, OpenFlow controller, and (depending on the implementation) BGP speakers
- Layer 2 SDX needs NSI server, performance monitoring tools, and perhaps topology server
- SDN SDX needs an OpenFlow controller, policy database, and performance monitoring tools
- SDI SDX has similar needs, and also makes some resources available for use within multi-domain slices

Notional SDX Schematic

Note that the SDX's equipment may be physical, virtual, or a mix



* Examples: Campuses, R&E nets, ISPs, Layer 2 networks, HPC, AWS, new SRPs, etc.

What are we building on?

Existing infrastructure and people:

- GENI - multi-domain control frameworks, racks, slicing, clearinghouse, AAA
- Experimental SDXs at MAX, SOX, StarLight
- Internet2 – AL2S, OESS, FlowSpace Firewall
- ESnet – SDN technology in support of data-intensive science
- Existing exchange points, regionals, and backbones
- CC-NIE funded Science DMZ projects
- SDX software: Sigcomm2014, the “original SDX”
- NSI – GLIF
- DYNES
- NSF Cloud Solicitation

Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

Overarching Research Areas

- Security
  - Multi-domain now applies to compute / storage / network / sensors / ….
  - How do SDXs function as security/trust points?
  - What new security techniques can be enabled by SDXs?
- Support for large-scale distributed infrastructure to support science
  - Computational science and experimental & observational science
  - Relation to Grid technologies / futures
- System aspects
  - File system / storage (new forms of virtualized storage architectures)
  - SDX architecture itself
- Multi-domain aspects
  - Control of multi-domain SDNs and infrastructure
  - Multi-domain controllers and inter-domain protocols
  - Federation encodings & formal reasoning
  - Static vs. on-demand federations

Overarching Research Areas (continued)

- Software defined infrastructure
  - Can we build a virtual distributed data center using SDXs?
  - Can (say) a new LHC type project create a template for SDI, try it out in prototypes / emulation, before using
  - "Push button" construction of infrastructure, including stacks
  - Relationship to Internet of Things / Wireless / Big Data
  - Relation to meta-cloud environment
  - How do we reconcile virtual infrastructure with "green"
  - Interactions of measurement / reasoning / etc on the basis of formal models
- Resources
  - Resource description / reasoning
  - Resource reachability protocol
  - Orchestration strategies
  - Virtualization strategies
- Protocols
  - What is the protocol to be used between different controllers/SDXs?
  - EastWest languages
- Policy definition / enforcement in multi-domain / global policy
- Potential economic models

SDN Specific Issues

- SDN and non-SDN traffic in the same SDX
- Can an SDX implement SDN policies on non-SDN traffic?
  - For the data plane, there is no "SDN" traffic. It's the control that determines features.
  - However some apps may require SDN control in order to function?
- Some apps depend on agreements with the recipients
- Recognize ultrahigh flow, bypass SDX to another around the Internet
- Network restrictions – privacy issues on circuits

Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

What are we trying to do?

- Explore new paradigms for inter-domain routing and resource identification/allocation/utilization
- Redefine what peering means. The current model for how peering works in the internet is unsatisfying and not meeting our needs! SDN and SDI should be useful for doing this
- Understand how the new paradigms can support applications that are not well served on today’s Internet and identify new classes of applications that can be supported

Next Steps (1)

- Outline initial guidelines for each of the architectures for each of the SDX types
- Identify multiple SDXs and multiple SRPs to participate in prototyping and experimentation projects
- Each SDX owner implements a Rev 0 prototype based on initial architectural guidelines for a particular SDX type and, if applicable, cooperating SRPs. GENI racks may be used to stage these into use – even internationally
- Each SDX and/or SRP owner brings specific use cases/applications that they will use to demonstrate how this infrastructure provides a benefit
- And iterate multiple times on the above steps

Next Steps (2)

- Engage stakeholder communities
  - SDX developers
  - Infrastructure providers
  - Researchers
  - Industry
  - User communities
  - Application developers
  - International
  - Federal agencies
- Focused workshops
- Presentations at professional meetings
- Collaborative projects
- Coordination and collaboration are essential

Call to Action

- It is essential to encourage research in this area
- It is essential to make sure there is a continuing critical mass of people talking to each other in these areas, as cross-fertilization is very important
- It is essential to continue to build the community, and figure out the relationships between this community and others
- Active engagement, support and funding by U.S. government agencies are essential to the successful prototyping and deployment of experimental SDXs

Additional Material

AGENDA

Thursday, June 5, 2014

Goal for day: Create first cut at overall Near Term and Advanced SDX architectures and motivating apps for both.

Welcome – Lyles, Landweber

Introductory presentations (10 minutes each) – Elliott

Lehman, Mambretti, Vanbever

Benson, Boyd/Vietzke, Monga, Ricart

Complete intro presentations, charge for/transition to break-outs

1. (Hutchins) Near-term SDX (Layers 2/3, SDN) – architecture, security, exemplar apps, timelines
2. (Elliott) Advanced SDX (SDI) – architecture, security, exemplar apps, timelines

Presentations from break-out groups, with discussion

Break-out groups resume, to refine / harmonize their work

Friday, June 6, 2014

Goal for day: Rough plans for infrastructure builds, deployment, software prototyping.

Reprise: show current draft of break-out group slides

Near-term and Advanced break-out groups resume

1. Infrastructure for SDX versions
2. Software prototyping strategies
3. Interoperability vs. multiple designs
4. Timelines

Presentations from break-out groups, with discussion

Call for volunteers / wrap-up

Attendees: Workshop on Prototyping and Deploying Experimental SDXs

JW Marriott Hotel, Washington DC, June 5-6, 2014

Jeannie Albrecht Williams College, jeannie@cs.williams.edu
Celeste Anderson USC, celestea@usc.edu
Theophilus Benson Duke , tbenson@cs.duke.edu
Mark Berman GENI , mberman@bbn.com
Eric Boyd Internet2, eboyd@internet2.edu
Richard Carlson DOE, Richard.Carlson@science.doe.gov
Russ Clark , Georgia Tech, Russ.Clark@gatech.edu
Steve Corbato Utah, Steven Corbato steve.corbato@utah.edu
Brian Court CENIC, bac@cenic.org
Vince Dattoria DOE, vince.dattoria@science.doe.gov
Chip Elliott GPO, celliott@bbn.com
Chin Guok ESNET, chin@es.net
Ron Hutchins Georgia Tech, ron.hutchins@oit.gatech.edu
Julio Ibarra FIU, julio@fiu.edu
Larry Landweber GPO, Larry.landweber@gmail.com
Tom Lehman MAX, tlehman@umd.edu
Mark Luker NCO, luker@nitrd.gov
Bryan Lyles NSF, jlyles@nsf.gov
Joe Mambretti Northwestern , j-mambretti@northwestern.edu
Ron Milford Indiana, rmilford@iu.edu
Grant Miller NITRD, miller@nitrd.gov
Inder Monga ESNET, imonga@es.net
Anita; Nikolich NSF, ANIKOLIC@nsf.gov
Glenn Ricart US IGNITE, glenn.ricart@us-ignite.org
Steve Schwab USC/ISI, schwab@isi.edu
Laurent Vanbever Princeton, vanbever@CS.Princeton.EDU
Rob Vietzke Internet2, rvietzke@internet2.edu
KC Wang Clemson , kwang@clemson.edu
Boyd Wilson Clemson, ubnice@gmail.com
Mike Zink Massacheusetts, zink@ecs.umass.edu

Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

SDN/SDX Program
Research Objectives

SDX Desired Capabilities

- Pervasive Instrumentation and Monitoring (PerfSONAR, etc.)
- Pervasive Modeling of all SDX infrastructure and SDX applications (e.g. along the lines of Tom Lehman's NDL-based examples)
- Automatic support for determining "Application Closures", (i.e. the set of all software, configurations and resources need to re-build that Application)

… in service to Advanced Functionality including:

Automated and Ultimately Reasoning Based Solutions for

- Optimization of Applications, single SDX and multiple SDX
  - resource utilization and performance
- Resource Quotas and Intrusion Detection Systems at the SDX level for
  - Individual Applications, single SDX, and multiple SDX eco-system
  - Aspects addressing: policy vocabulary, policy definition, policy exchange, resource quota enforcement
- Robustness, resilience, security, and portability or reconfigurability of Applications

Comprehensive Analysis of the Impact of Cyber Infrastructure Technology Adoption

- Total Cost of Ownership / Long-term Assessment and Re-evaluation of Analytics
  - Is the current SDX infrastructure providing the best possible service?
  - Subject to the capital and operational costs of computation, storage, networking, and instruments
  - Adopt Business analytics and Return-on-investment analysis used by Internet-scale Cloud Providers

Trust, Authentication, Resource Use Authorization, Auditing/Validation

- "If I can't pass my policies on to another network and trust that they will be implemented…"
- "If I can't verify that the policy given to me is from someone I trust…"
- "If I can't prevent someone from using up all my resources…"
- "I want to push security back to the ingress point to be effective…"
- "I want a researcher sitting on campus to be able to… in an automated seamless fashion…"
- "If the 'incentive' isn't set right…" – SDX marketplace to create good behavior between… SDX supports a mechanism… "proof by induction"
- "I want the networks connecting to the SDX to be able to express their capabilities in a much richer manner and trust what I get."
- "How do I measure or know that what I'm requesting is getting implemented?" – is the stuff still available, for how long, at what granularity? Not insurance but a guarantee (SLA)
- "Provide a better way to see what I'm sending and what I'm receiving, exact distribution of HTTP traffic"
- Unique logging elements for SDX vs IX? – SIEM? Privacy assertions

Policy Domain

- BGP exchanges policy (reachability policy)
  - Policy in the local level
  - Route server – reduce the complexity – clearinghouse or brokerage service for SDX – extended services without BGP constraints
- Chaining policy across multiple domains
- “god box” – we don’t all have to have the same “god” -
- What is the protocol to be used between different controllers? Inter-domain
- Exchange topology? Connectivity vs reachability vs. capability
- Application specific policy – VOIP?

Performance Level for an Exchange

- Robustness
- Bit Error Rate (BER)
- “If you meet x traffic parameters as measured by this mechanism,… “
- Set up performance across the exchange – if traffic meets pre-arranged agreement.
- Synthesize measurement techniques - “start perf monitoring – stop and checksum ” – are you getting the service you bought (resource accountability)?

Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

Digital Watershed (Clemson)

Digital Watershed - a NSF MRI project at Clemson is developing a complete river sensing, processing, storage, and visualization platform. SDX can be used to facilitate data distribution for the platform.

Astronomy (Clemson)

Astronomy research at regional institute (College of Charleston). SDX can provide compute and network resources that are otherwise hard to acquire.

Genomics Workflow (Clemson)

- SDN-enhanced data transport for multi-campus genomics workflow
- access to seamless data transfer nodes and compute nodes for data processing (data reduction or transformation) in on-campus or regional SDXs. There is no uniform orchestrated solution today - all current solutions require researchers/network engineers on each campus to manually configure data paths and allocate compute resources.

Severe Weather Prediction & Warning (UMass)

a. What is the application?
   CPS for severe weather prediction & warning
b. How will they use the SDX?
   Get a slice that consisst of several SDN domains, computation, storage.
c. What will improve from state of the art today? Why is SDX a better or preferred approach?
   - Computation and storage at "safe" locations
   - Load balancing
   - Alternative routes (circumvent ones that are impacted)

Emergency Management (UMass)

a. What is the application?
   Emergency management. E.g., a passenger train wreck

b. How will they use the SDX?
   Get a slice that consists of several SDN domains, computation, storage, and LTE network.

c. What will improve from state of the art today? Why is SDX a better or preferred approach?
   – Send data multipath and just use the ones that reaches destination first.
   – Get computation and storage that is close to site.

Astronomy (MAX)

a) Brief Description
Astronomy projects such as Pan-STARRS1 and Large Synoptic Survey Telescope projects: two data-intensive research programs in the area of time-domain astrophysics. Both of these applications need to move data from telescopes to their computation environment.

b) How they can use SDX
If there were an infrastructure of distributed SDXs, with compute and storage resources, their workflows could be modified to utilize SDX compute/storage resources which are closer to the data source. This will allow them to more easily move the science data to their applications, and possibly allow them to develop additional workflow features such as increased realtime processing and interaction with data source.

c) What will improve from state of the art today? Why is SDX a better or preferred approach?
This will allow them to more easily move the science data to their applications. By moving their application to a point closer to the data, they benefit from a lower bandwidth-delay product, and no longer have to work out the issues with campus firewalls and other infrastructure slowing down their data transfers. Running their applications in a well engineered SDX, closer to the data source may also allow them to develop additional workflow features incorporating increased realtime processing and interaction with data sources.

Climate Science (MAX)

a) Brief Description
Climate science applications such as the Global Land Cover Facility need to move large amounts of data from sources repositories to their local environment where they utilize local computational facilities to produce value added products. These value added products are then distributed to their users.

b) How they can use SDX
This application could utilize a distributed SDNx to assist with data transfers and computation workflows. Data transfers could be accelerated by breaking large bandwidth-delay TCP sessions into two small bandwidth-delay segments at the ends, and an SDX to SDX transfer in the middle, possibly utilizing a reserved bandwidth path or an alternative data transfer protocol. In addition, SDXs which have high performance connections to compute resources would allow new workflow methods to be evaluated based on just-in-time data movement and utilization of external compute resources to develop the processed data products.

c) What will improve from state of the art today? Why is SDX a better or preferred approach?
Providing an infrastructure to allow them to experiment with an infrastructure which facilitates data movement and provides better access to a variety of compute resources will allow them to more easily experiment and innovate with their workflows and produced products.

Large Scale Simulation (MAX)

a) Brief Description
Large scale agent-based simulation models are used to model a variety of environments and can quickly scale in computational requirements and data generated. Agents typically operate as finite state machines, and as their instruction sets grow, so do the computational requirements and data generated. The types of simulations are often limited by the local computation resources, because moving the large amounts of data out of the source computation environment is problematic.

b) How they can use SDX
A distributed SDX infrastructure with rich connectivity to a variety of compute resource providers, would provide an opportunity for these types of simulation applications to dynamically expand their simulation runs into other computation environments and retrieve the generated data.

c) What will improve from state of the art today? Why is SDX a better or preferred approach?
A distributed SDX could facilitate data movement from computation facilities to long term storage resources. In addition, SDX computation facilities could be utilized explore new workflows based on in-line realtime processing of simulation data, or dynamically bursting into external compute environments as simulation scaling requires.

Train Emergency Event (NSF)

- A freight train with 100 cars of toxic material derails, catches fire and causes a serious emergency across a wide area. There are multiple federal state, local and volunteer organizations involved in the response. To coordinate the emergency a "slice" is set up containing cyber assets such as coordination resources, HPC for predicting the path of the toxic plume, and databases of public facilities, residence information and the status of transportation resources.
- There is insufficient time to issue identity credentials specific to this emergency to the the responders. Thus you have to use the credentials already in place. The slice provides a context for allowing the appropriate emergency response focused use of existing credentials.

Radio Astronomy (StarLight)

Radio astronomy research using real time data flows at multiple sites

2. How will they use the SDX?

Programming of high performance streams among multiple sites in real time

3. What will improve from state of the art today? Why is SDX a better or preferred approach?

Enhanced programming of individual flows, flow integration, and real time correlation

Computational Genomics (StarLight)

2. How will they use the SDX?

Programming of high performance, high capacity streams among multiple sites

3. What will improve from state of the art today? Why is SDX a better or preferred approach?

Enhanced programming of individual flows and flow integration

High Performance Digital Media (StarLight)

High Performance Digital Media Network (HPDMnet): An international, programmable L2 network supporting large scale media streams, including 4k, 8k, and Ultra resolution 3D and VR

2. How will they use the SDX?

Programming of high quality low-latency streams combined into complex flows, including sync'ed audio

3. What will improve from state of the art today? Why is SDX a better or preferred approach?

Enhanced programming of individual flows and flow integration

Large Hadron Collider (StarLight)

High energy physics, Large Hadron Collider networking supporting large scale streams for physics research

2. How will they use the SDX?

Programming of high quality low-latency streams combined into complex flows, among the Tier 0 instrument at CERN and multiple compute and storage centers around the world

3. What will improve from state of the art today? Why is SDX a better or preferred approach?

Enhanced programming of individual flows and flow integration

Genomic transfer for personalized medicine (US Ignite)

Fine-grained differential disease diagnosis and personalized treatment programs often are possible when the patient's genome is sequenced and available for analysis. Sequencing the individual genome is now possible for several hundred dollars, but protecting personal health information "in flight" is extremely important. SDN can provide some kinds of security that arise from flow isolation including precluding man-in-the-middle attacks, spoofing, DNS poisoning attacks, and with performance isolation, denial of service attacks.

The SDX will be used to stitch together protected SDN virtual networking paths to guarantee the SDN flow isolation properties are maintained across administrative domains.

Although encryption is used in genome transport, even today, the key management for this encryption is often weak (e.g. re-used keys) for the convenience of practitioners. Additional safeguards are often aimed at casual crackers and won't stand up long against a determined and serious attack. SDN virtual network paths will preclude some of the most popular points on the attack surface. This application will test and stress big data, privacy, and security issues.

Optional additional SDX service: reducing transmission time by using engineered parallel paths through the multiple networks involved.

Video collaboration (US ignite)

Multi-domain multi-way live high-quality video collaboration.

SDN can reduce the latency in high-quality multi-way live videoconferencing through two mechanisms: (a) Replace MCU with OpenFlow packet duplication rules, and (b) Use SDN paths engineered for low latency and ultra-low jitter.

SDXes will be needed to provide this service on a cross-domain basis. Challenges will include finding low-latency total length paths between multiply-connected domains, providing in-network packet duplication (possibly a service of the SDX), and managing performance to provide ultra-low jitter.

The state of the art today is to "cover up" the long paths, unnecessary hop counts, dropped packets, and MCU delays by adding hundreds of milliseconds of jitter buffer and varying the picture quality. With an SDX engineered to provide the services described in 1a, participants should have a much lower latency (more natural) broadcast-like-quality video conference. This application will test and stress the ability to engineer low-latency paths, providing ultra-low jitter through an SDX, and providing a high-efficiency replacement for IP multicast using OpenFlow.

Optional additional SDX service: transcoding the video for lower-quality devices such as smartphones.

Synchrophasor Real-time Analysis (US Ignite)

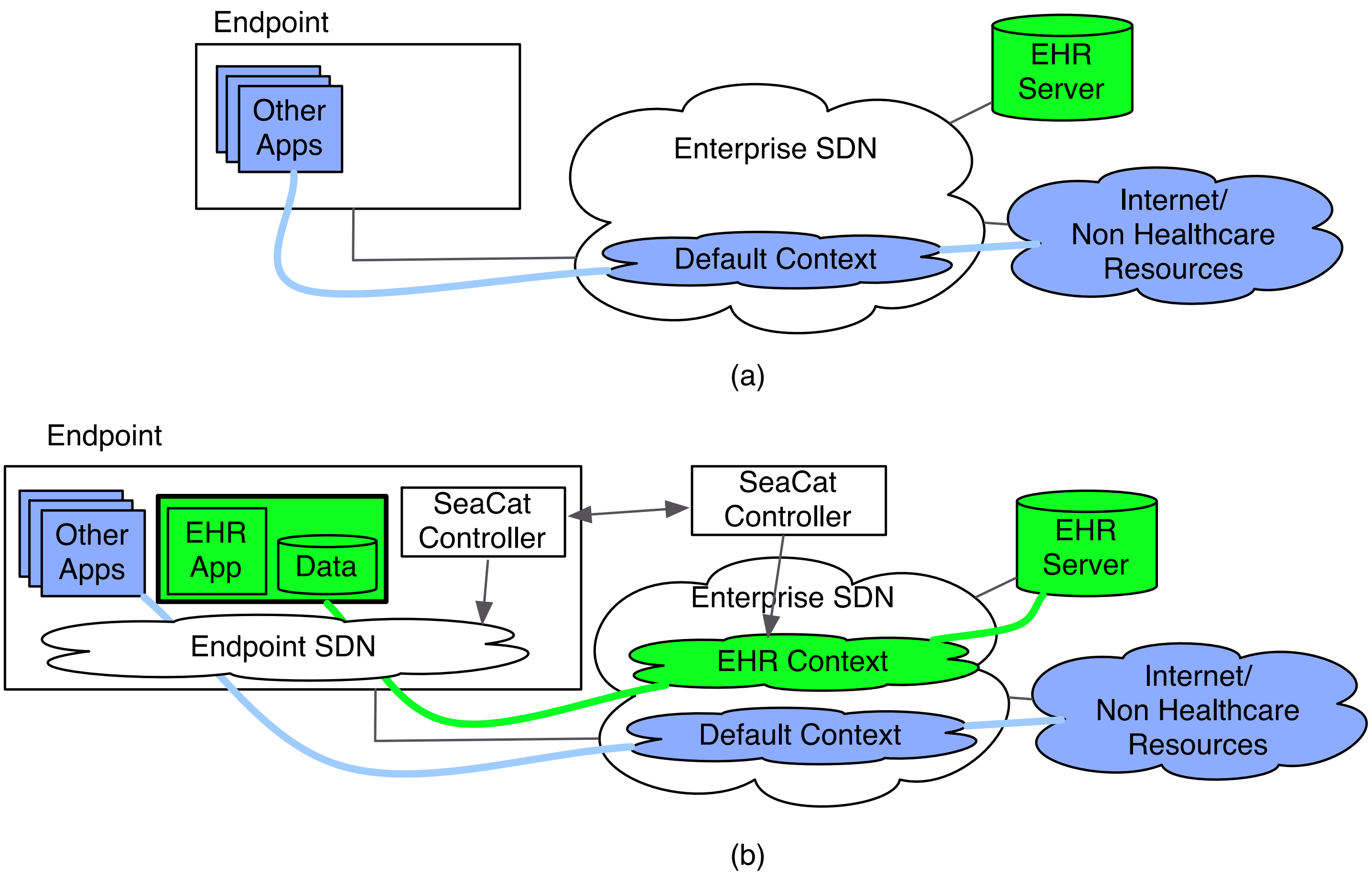
Synchrophasors measure the reactive component of powerline transmission, which is an accurate indication of pending disconnects in the power grid. If acted on quickly enough, the cause of the reactive component can be mitigated or alleviated preventing power outages. Extremely low latency between synchrophasor readings and analysis is needed on a distributed basis.

SDX will be used to aggregate local synchrophasor flows using local processing as well as very low latency and ultra-low jitter transmission to the next level synchrophasor aggregation processing point.

Today's shared networks are not fast enough or low-enough latency to permit wide-area synchrophasor outage protection.

Collaborators: K. Van der Merwe, T. Cheatham, and J. Breen, University of Utah

Health Care Application Extensible to an SDX (Utah)



Using an SDN and sliced Shared Resource Providers (SRPs - e.g., ACIREF centers) for securely handling protected/sensitive information such as Electronic Health Records (EHR)

Outline

- SDX Background & Context
- SDX Conceptual Architecture
- Research Topics
- Next Steps / Call to Action
- Exemplar Applications
- Imagining the Future

“Imagining” by Steve Corbató, U. of Utah

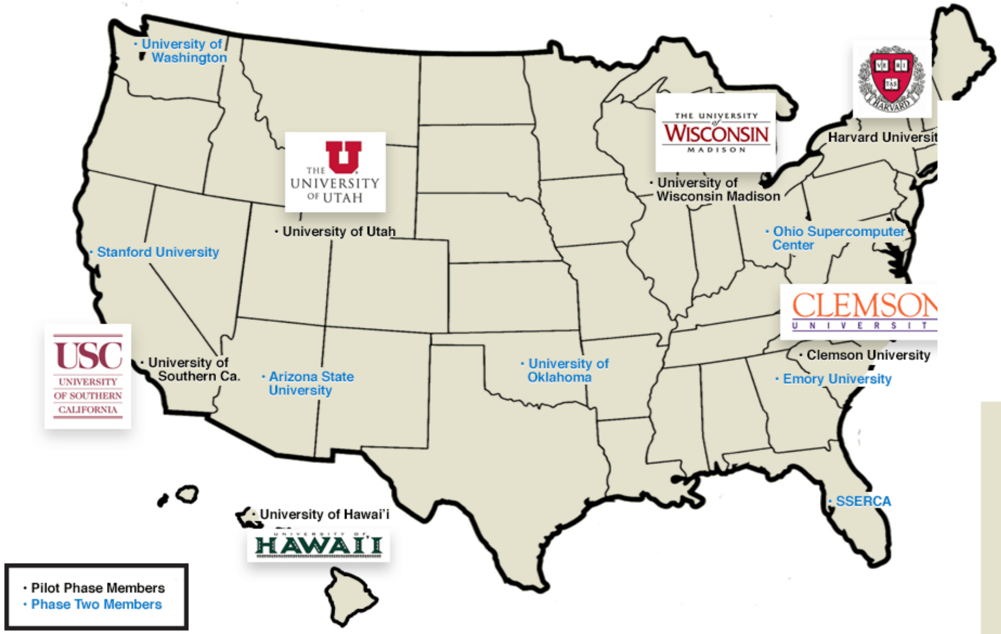
Leveraging the ACI-REF Collaboration ('Condo of Condos') for a national SDX & SRP fabric

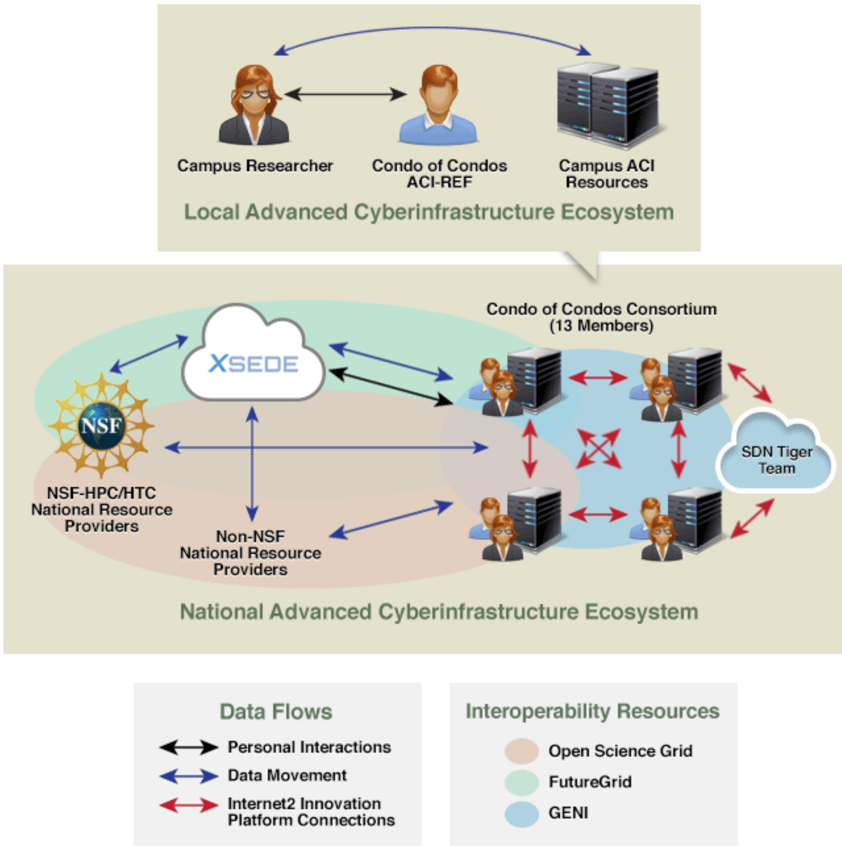
Steve Corbató
Interim CIO, University of Utah and Utah System of Higher Education
Interim Director, Center for High Performance Computing (CHPC)

GENI SDX Workshop – Washington DC – June 6, 2014

ACI-REF project: Condo of Condos

NSF #CNS-1338155 – PI: Jim Bottum, Clemson





ACI-REF campus research computing centers are federating to share resources – both human and computational

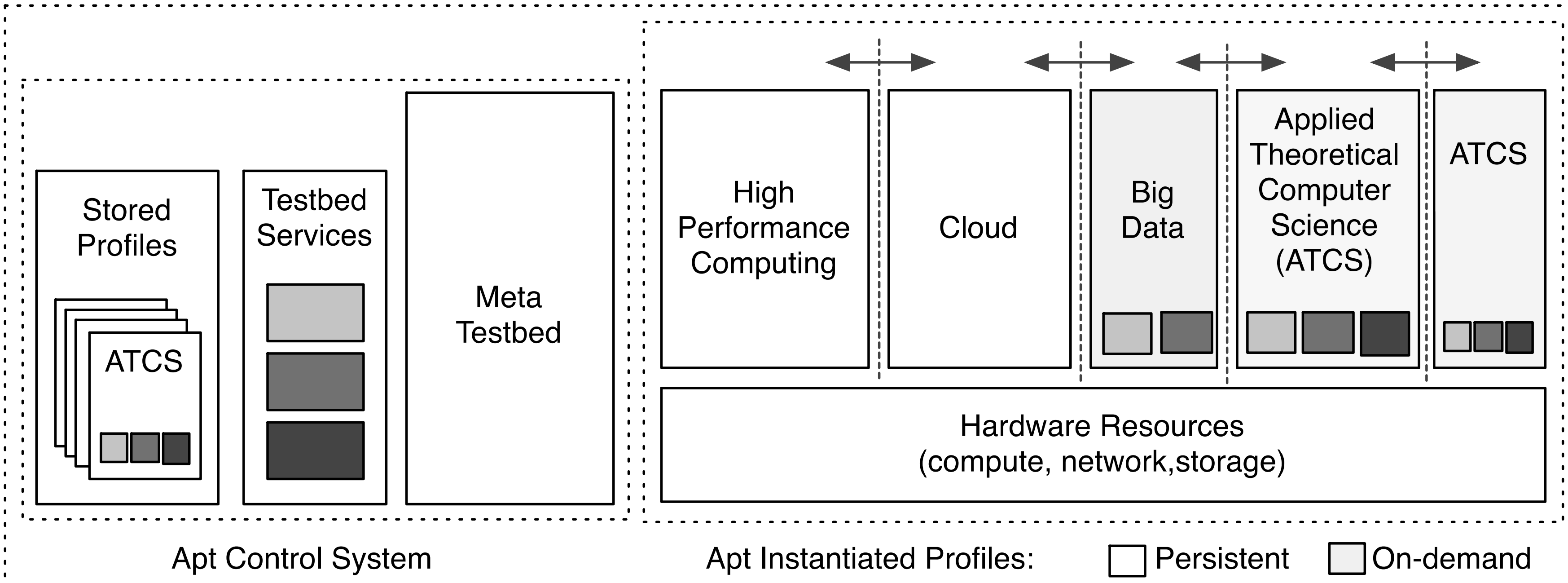
Also all connect to the Internet2 Layer 2/3 Network at 100G

Building a nationally distributed SDX with federated ACI-REF SRPs

- ACI-REF campus research computing centers have advanced networking expertise and strong interest in deploying SDN internally
- These centers already serve as Shared Resource Providers (SRPs) that can attach to *local SDXes* leveraging data center/cloud resources
- Slices of these SDXes can be stitched together to create a national ACI-REF SDX fabric
- Potential applications: genomics, secure computing with patient health information, weather forecasting, collaborative Big Data projects
- Novel use of SDN/SDX: this allows for greater degree of control, granularity, and traffic isolation than possible in the current Internet

Apt: A Novel Cluster for Research

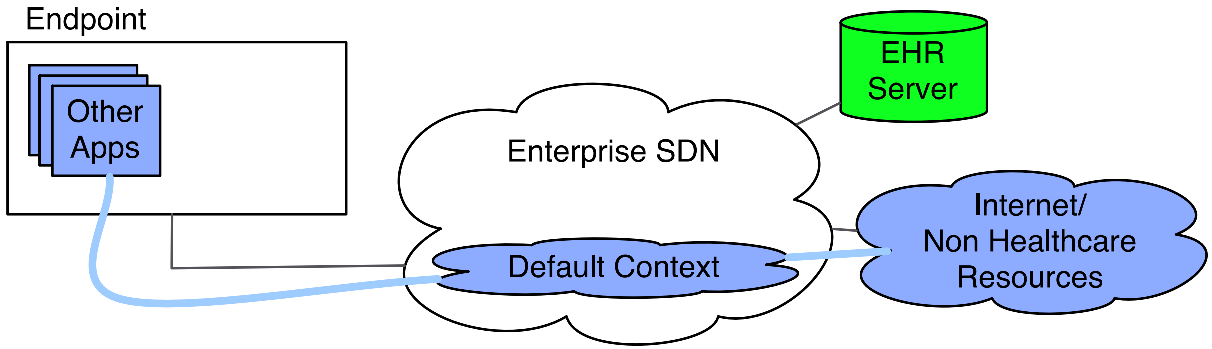
NSF #CNS-1338155 – PI: Ricci – Co-PIs: Corbató, Eide, Facelli, Van der Merwe



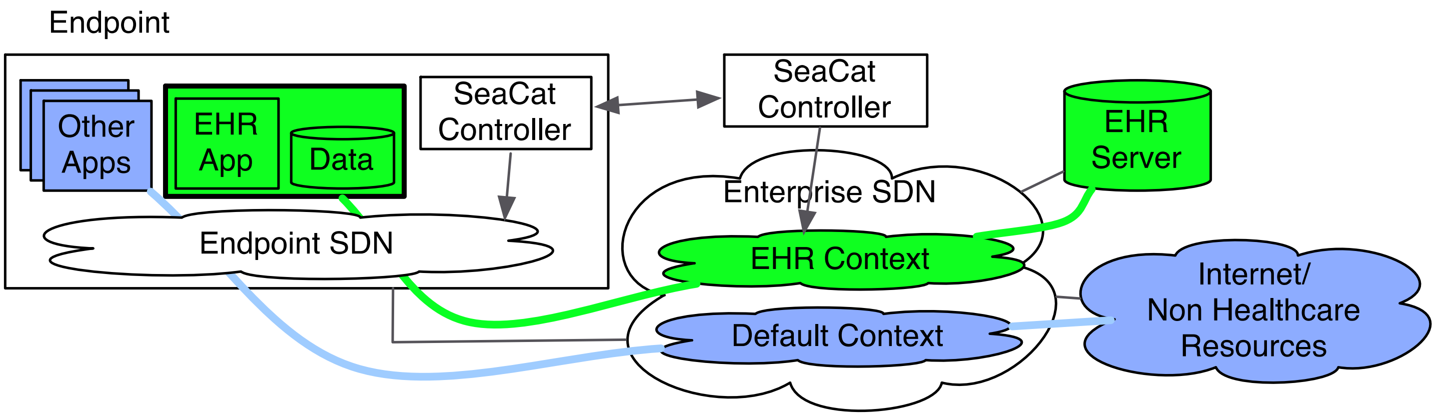
A potential testbed platform for an SDX-attached Shared Resource Provider (SRP)

Collaborators: K. Van der Merwe, T. Cheatham, and J. Breen, University of Utah

Health Care Application Extensible to an SDX



(a)



(b)

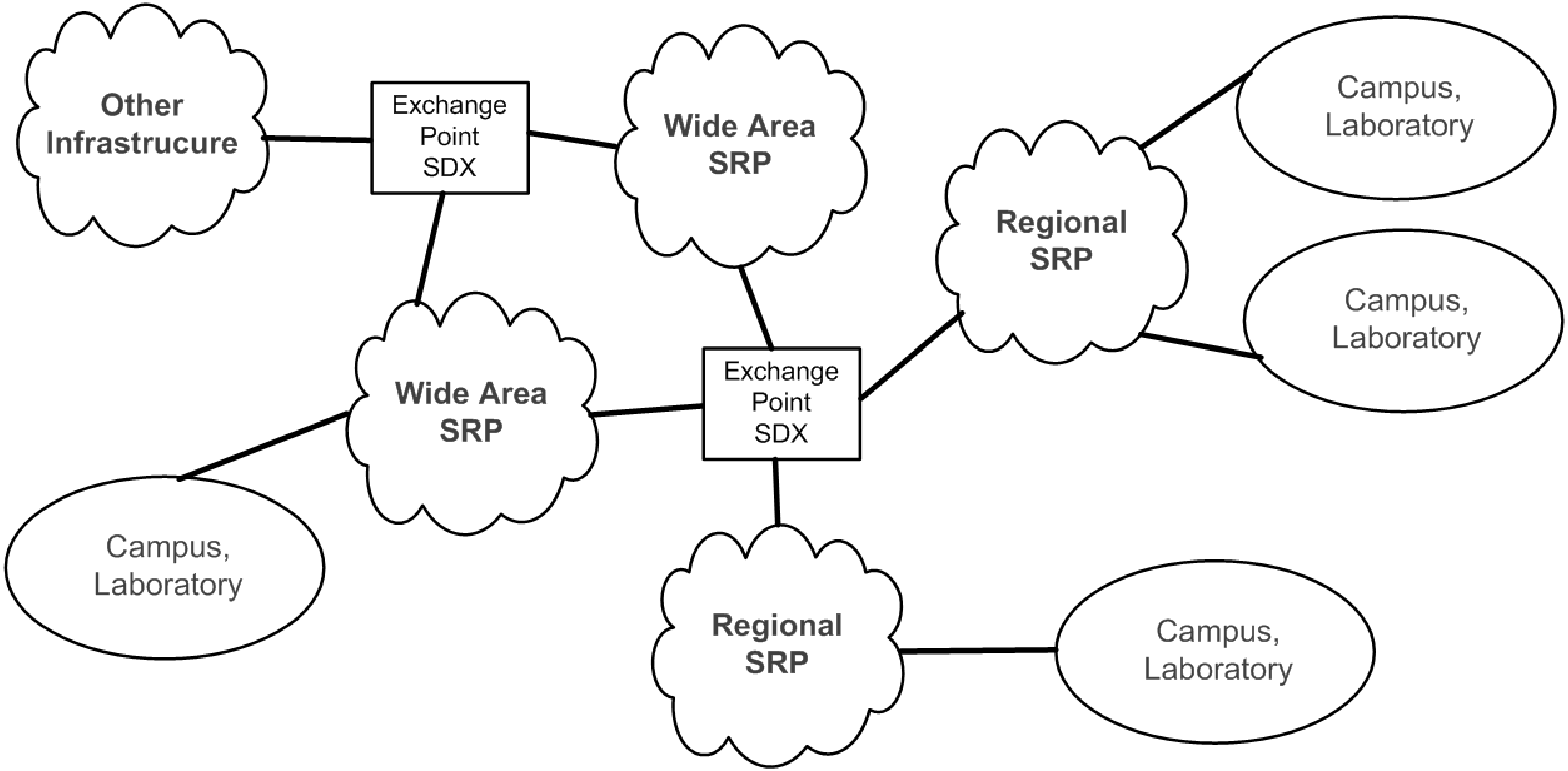
Using an SDN and sliced Shared Resource Providers (SRPs - e.g., ACIREF centers) for securely handling protected/sensitive information such as Electronic Health Records (EHR)

“Imagining” by Tom Lehman, MAX

a. "I can imagine"
A national scale deployment of SDX and SRP which allow prototyping and experimentation in both the architecture and use case dimensions.

Architecture investigation areas include:
-design and control of an individual SDX and/or SRP
-resource advertisement requirements and mechanisms
-multi-domain SDX interoperation and service provisioning
-SDX policy and AAA

Use case investigation areas include:
-identify multiple applications
-design specific modifications to their workflows which utilize the unique properties of the SDX/SRP prototype infrastructure


Other Infrastrucure
Exchange Point SDX
Wide Area SRP
Regional SRP
Campus, Laboratory
Campus, Laboratory
Wide Area SRP
Exchange Point SDX
Campus, Laboratory
Regional SRP
Campus, Laboratory

c. What needs to get done / steps

- identify 3 or more SDXs and multiple connecting SRPs to participate in prototyping and experimentation project
- outline initial guidelines for each of the architecture investigation areas
- ask each SDX and/or SRP owner to design a prototype based on initial architectural guidelines
- the above step should result in identification of specific software and features sets to be utilized at each SDX/SRP.
- ask each SDX and/or SRP owner to bring one specific use case that they will demonstrate how this infrastructure provides a benefit
- iterate multiple times on the above steps

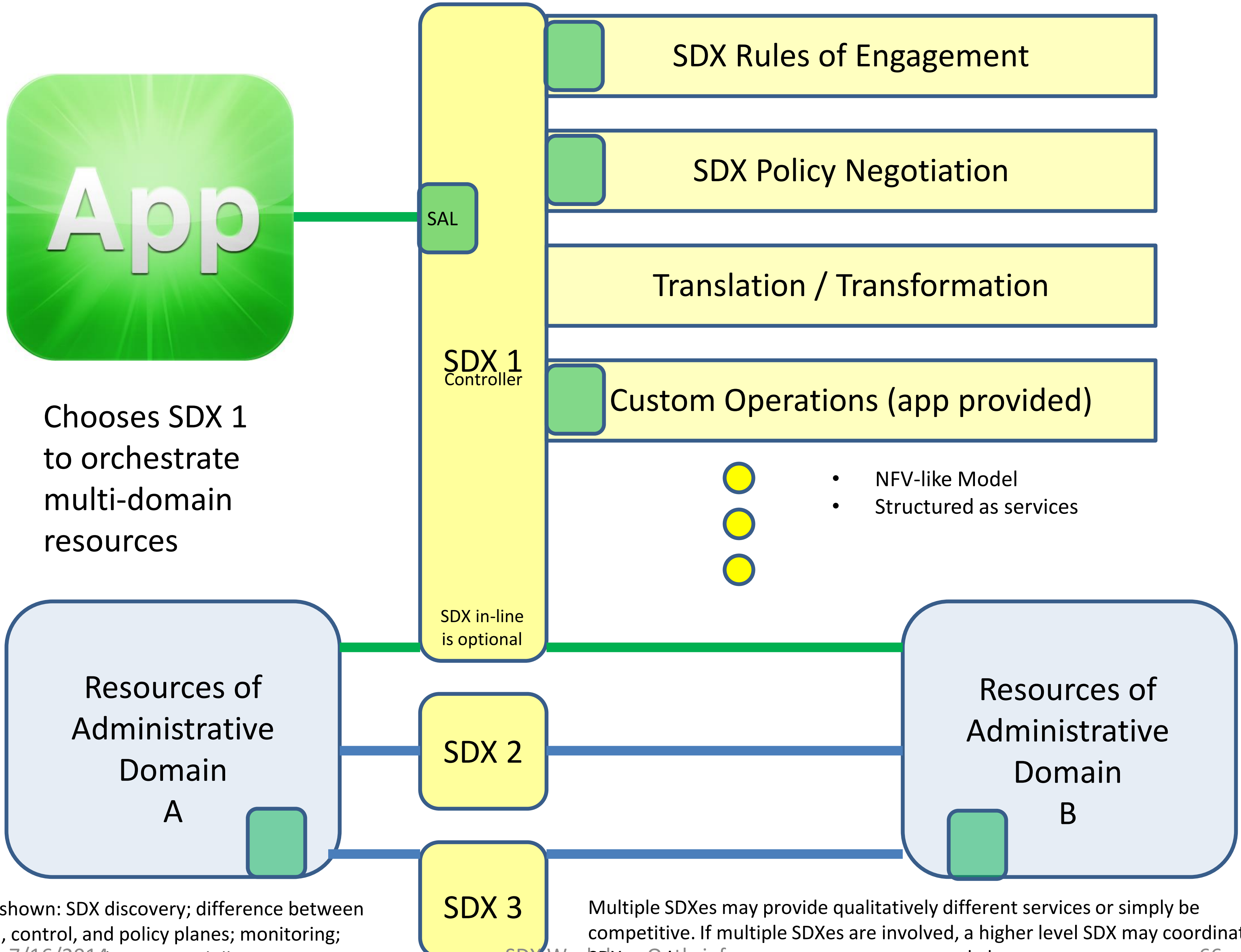
d. how could we build ‘N’ of these around the US ?
Find some specific SDX and SRP owners who are willing to participate in something like the architectural diagram. Possibilities include:

- Internet2 AL2S
- ESnet (Testbed or production network)
- WIX
- MAX
- MANLAN
- Starlight
- SOX
- Clemson
- UMD CLPK
- others

“Imagining” by Glenn Ricart, US Ignite

SDX

- Software-defined services Exchange
- Application-specific domain synthesis
- Services Orchestrator



Not shown: SDX discovery; difference between data, control, and policy planes; monitoring; resources usage limitations; billing; trust; etc.

Multiple SDXes may provide qualitatively different services or simply be competitive. If multiple SDXes are involved, a higher level SDX may coordinate SDX may issue requests to a separate control plane.

Glenn Ricart

Next Two Years

- Competitively commission 4-7 SDXes with supporting applications and resource partners (see next slide)
- Build an SDX community of researchers, prototype operators, users, industry partners
- Support research and prototypes of inter-domain
  - Resource negotiation
  - Trust
  - Applications
  - Etc.

How to do it

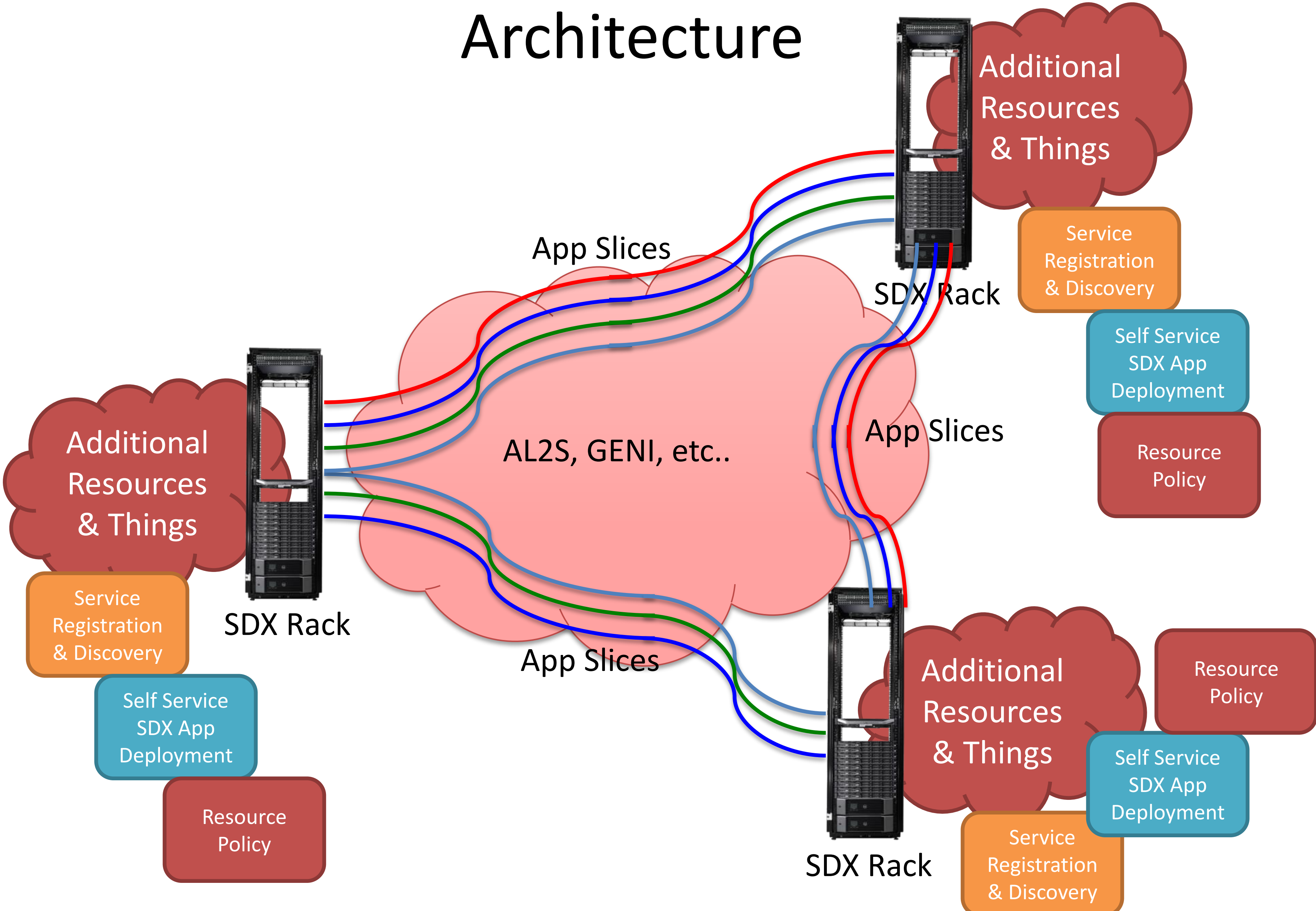
- Commission 4-7 prototype SDXes
  - Each SDX design should be based on supporting one or a few (classes of) inter-domain applications, which should also be developed or adapted to the SDX model as part of the commission
  - Seek a variety of applications and therefore designs
  - Consider existing building blocks GENI, OpenDaylight, OSCARS, Cyan, Overture, HP NFVO, ONOS, OpenContrail
  - Each SDX should be supported by two or more administrative domains that the SDX interconnects
  - Specific areas to be considered:
    - Performance (avoid any degradation)
    - Optimization (as appropriate to application(s))
    - Trust, identity, authorization (as appropriate to application(s))
    - Security (as appropriate to application(s))

“Imagining” by Boyd Wilson & KC Wang, Clemson

Imagine …

- SDX Racks (Short Term)
  - GENI Racks + base SDX software
  - Service registration and discovery software (e.g., PerfSonar style discovery)
  - Authentication software
  - Plug-in services
    - data transfer node, high throughput storage
    - Application specific services (e.g., agents installed and launched on-the-fly by applications)
  - Interfaces for …
    - applications to inquire and specify resources
    - Policy specification for resources
- Marketplace of SDX services (Long Term)
  - Self-service, on-demand services instantiated by applications
    - Push button deployment
    - Application registration & discovery (including authentication and …)

Architecture



What needs to get done

- Deployment
  - Software installation on GENI racks
  - Network configuration across campuses, regional, national
- Bring up specific pilot applications
  - Can be partly manual
  - Identify resource and attributes
  - Identify policies (with network operators)
  - Multi-domain provisioning and policy enforcement
- Integrate core services
  - Service registration & discovery
  - AAA
- Self Service Interface to push button deploy services into SDX
- Analyze trust/security issues at various levels of SDX Hierarchy

How To Build 2-3yrs

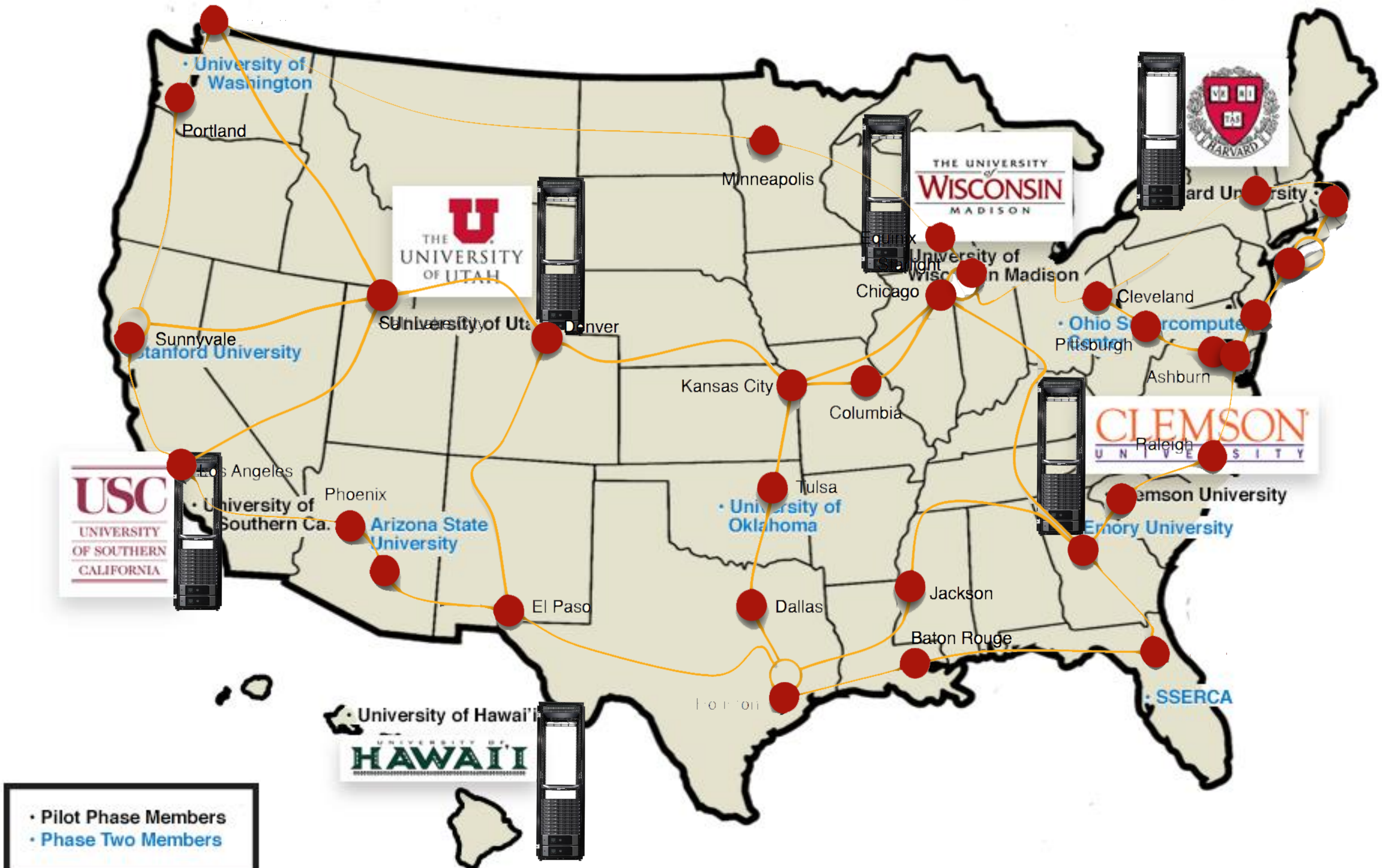
Yr 0-1.5

- Deploy a few SDX Racks across domains
- Integrate proof-of-concept software stack
- Prototype pilot applications and SDX services
- Identify straw man policy templates

Yr 1.5-3

- Second spiral
- Consensus/interoperation/refinement of solution(s) of
  - authorization of resources, QOS, Multi-Domain, …
- Add SDX Racks

Example ACI-REF SDX Racks



App – Digital Watershed

- Aquatic Sensors in streams, rivers, lakes measuring all kinds of useful things. Can model the progression of pollutants.
- With SDX a large number of bodies of water could to be sampled with a large number of sensors, compute, storage and network could be computed more real time closer to where the data is collected vs. slower compile and batch run.

G. W. Eidson,[1] S. T. Esswein,[2] J. B. Gemmill,[3,4] J. O. Hallstrom,[4] T. R. Howard,[3] J. K. Lawrence,[5] C. J. Post,[2] C. B. Sawyer,[6] K.-C. Wang,[5] and D. L. White[3]

App – Genomics

- Dr. Feltus

Dr. Feltus has worked with the CITI group at Clemson to optimize older Genomics codes and reduced tens of hours of processing down to Minutes.   But with this improvement, there is still a need for SDX to provide the ability to bring the computation nearer to the larger data stores.

Gas Cloud meets Black Hole…

P. Chris Fragile at the College of Charleston

predict the fate of a newly discovered gas cloud known as G2 which will soon (beginning June 2013) will be devoured by Sgr A*, the supermassive black hole that resides at the center of the Milky Way galaxy, over 50,000 hrs of compute time was needed for the study.

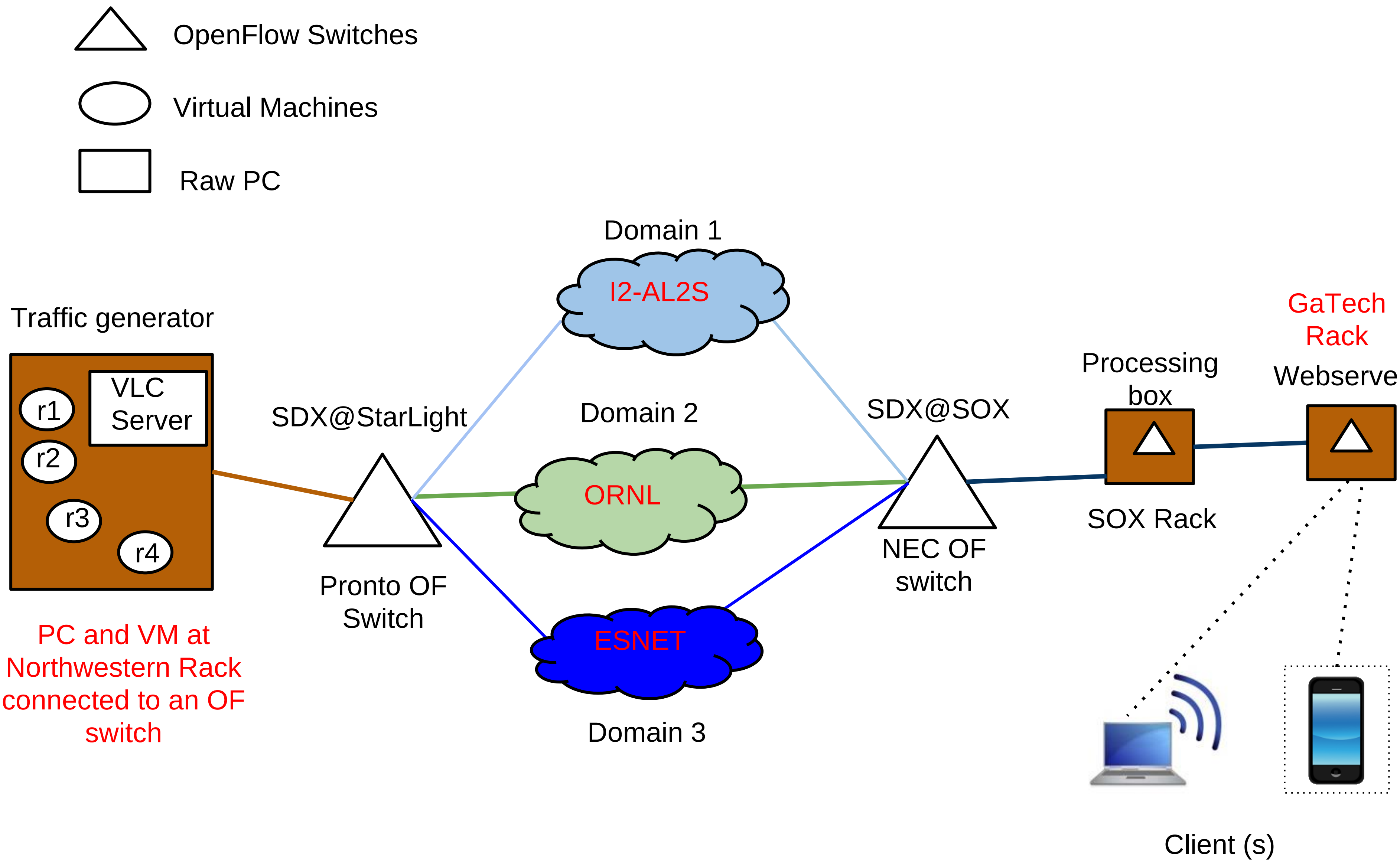
SDX could make studies like this easier, as College of Charleston does not have resources, the SDX could make it easier to find the resources needed and the self service computing/network/storage infrastructure can help big researchers at smaller Universities and Colleges. In addition the computation can be done more closely to the data.

“Imagining” by Mike Zink, UMass

I can imagine

- A series of heterogeneous SDXs available to the research community within the next 12 month
- People working on compelling applications that make use of SDXs
- New SDX placed at “strategically good” locations (Utah, MGHPCC).
- Opt-in from domain scientists (genomic sequencing, physics, …)

Architecture



What needs to get done / steps

- SDXs have to be implemented and made available to the research community.
- Compelling applications that show the benefits of SDXs have to be developed.
- Community engagement (workshops, demos, tutorials)
- Industry??

How could we build 2-3 of these around the US ?

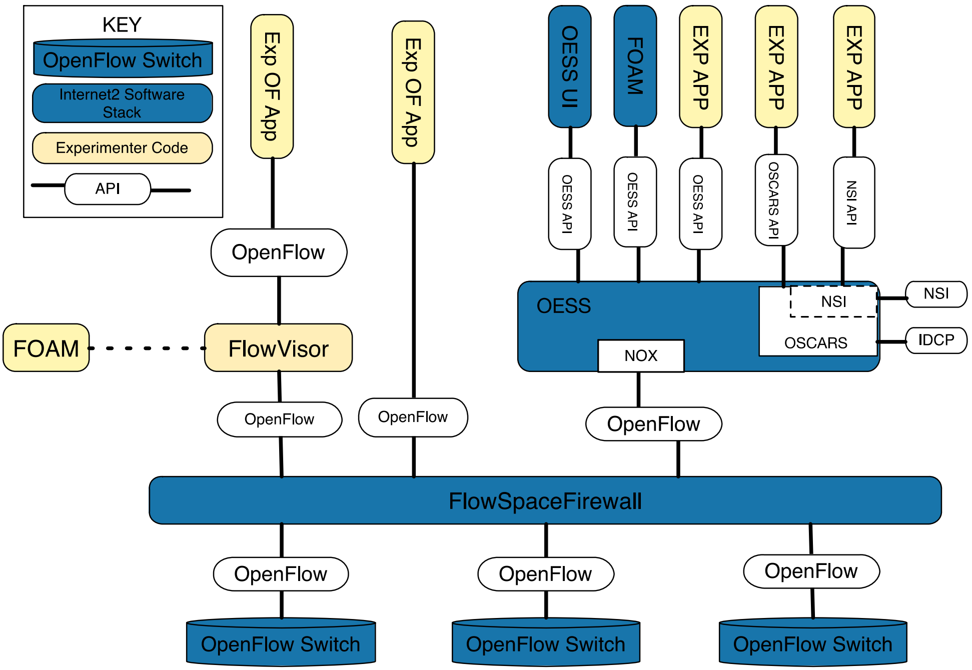
- Looks like some of them are being in the process of being built.
- How can community channel their “desires”?
- Spiral development: Build 2-3 and then build another set (3-5) that build on experiences from first cycle.
- Make them open to the research community (e.g., via GENI control framework).

“Imagining” by Eric Boyd, Internet2

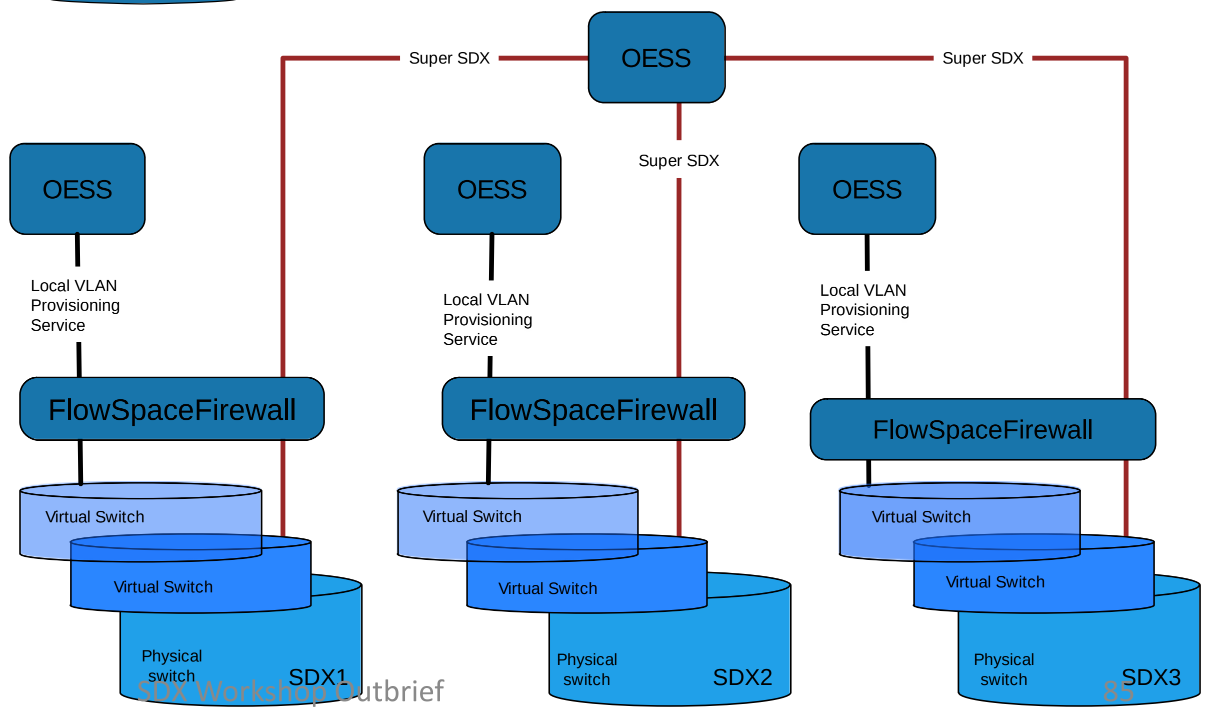
I can imagine …

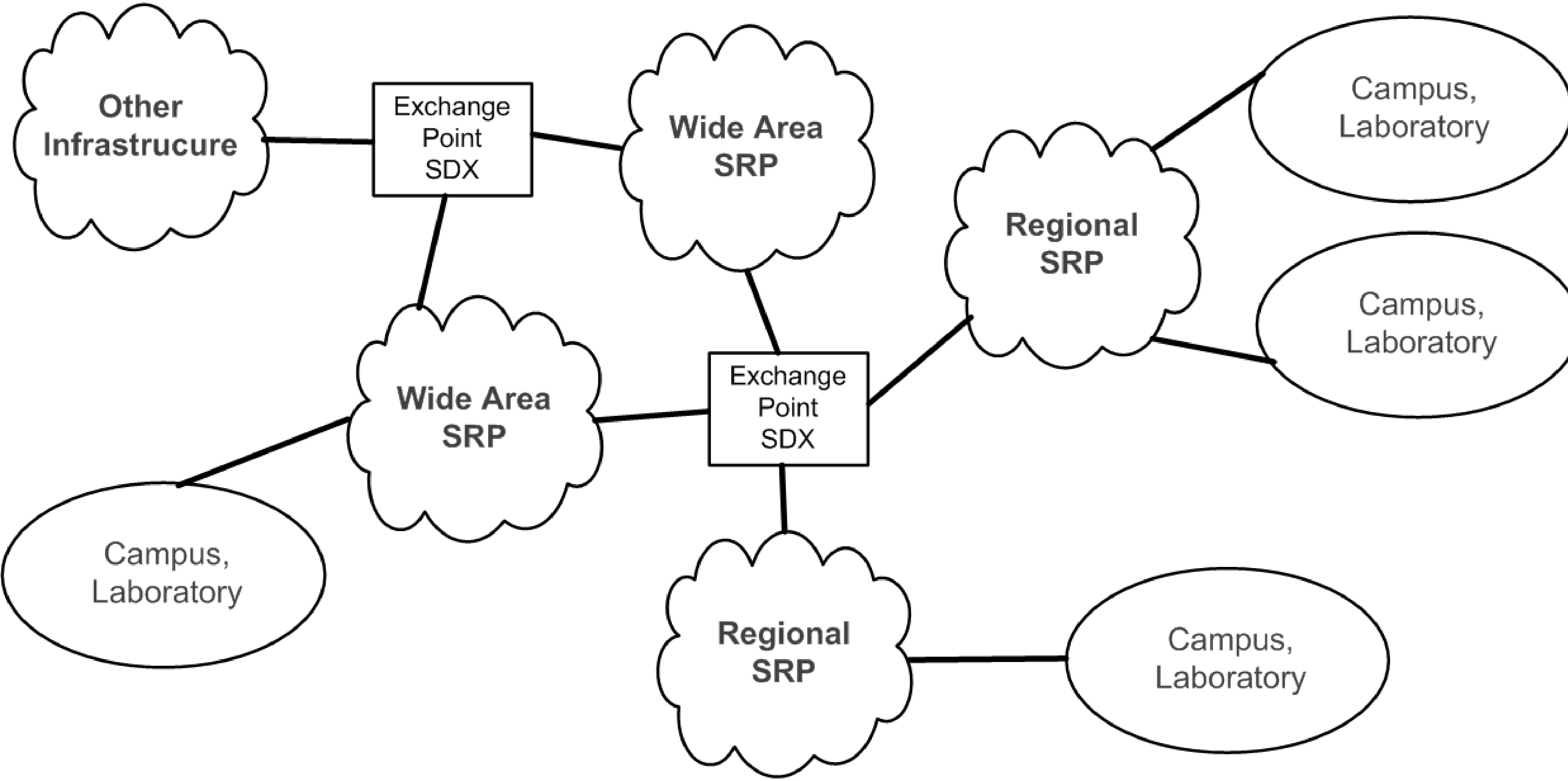
- There are a half dozen R&E SDXs in the US.
- Most such SDXs are also SRPs
  - Network element (big switch)
  - Software stack enabling virtualization
  - Integrated compute and storage
  - Merged into one or more Super-SDXs
- Network virtualization as part of compute and storage virtualization is a normal function
- The focus is on the “distributed supercomputer” …
- … not “independent HPC clusters connected by a network”

Sample SDX Architecture and Deployment



Possible Next Generation Super SDX





Courtesy Tom Lehman, MAX

What Needs to Get Done

- Analysis
  - Identify several reference applications
  - Define the SDX value proposition for each
- Implementation
  - Identify and implement a reference software stack that supports
    - Make network secondary to compute and storage
    - Integrate Network Virtualization
  - Deploy reference software stack at several IXPs (to become SDXs)
  - Define and implement a reference policy framework
  - Need virtual machines to support applications
  - Need testing harness
- Evaluation
  - Demonstrate the applications on the SDX
  - Evaluate the successes and failures of 1st Generation SDXs

How to Proceed?

- NSF to fund 3-4 likely IXP candidates
- Candidates to collaborate, not compete
- Identify a target date (Tech X 2014)
  - Initial implementations
  - Progress Update
- Final target date (Tech X 2015)
  - Final Implementation
  - Report out to community