Eric Boyd, Internet2
Luke Fowler, Indiana University
Luisa Nevers, GPO
Matt Zekauskas, Internet2

INTERNET2 AND GENI: ENABLING INNOVATION THROUGH NETWORK VIRTUALIZATION

Unleashing new waves of global discovery, together.

Advanced Layer2 Service Deployment

Legend:
- AL2S Node
- IP Node
- AL2S-IP Interconnect
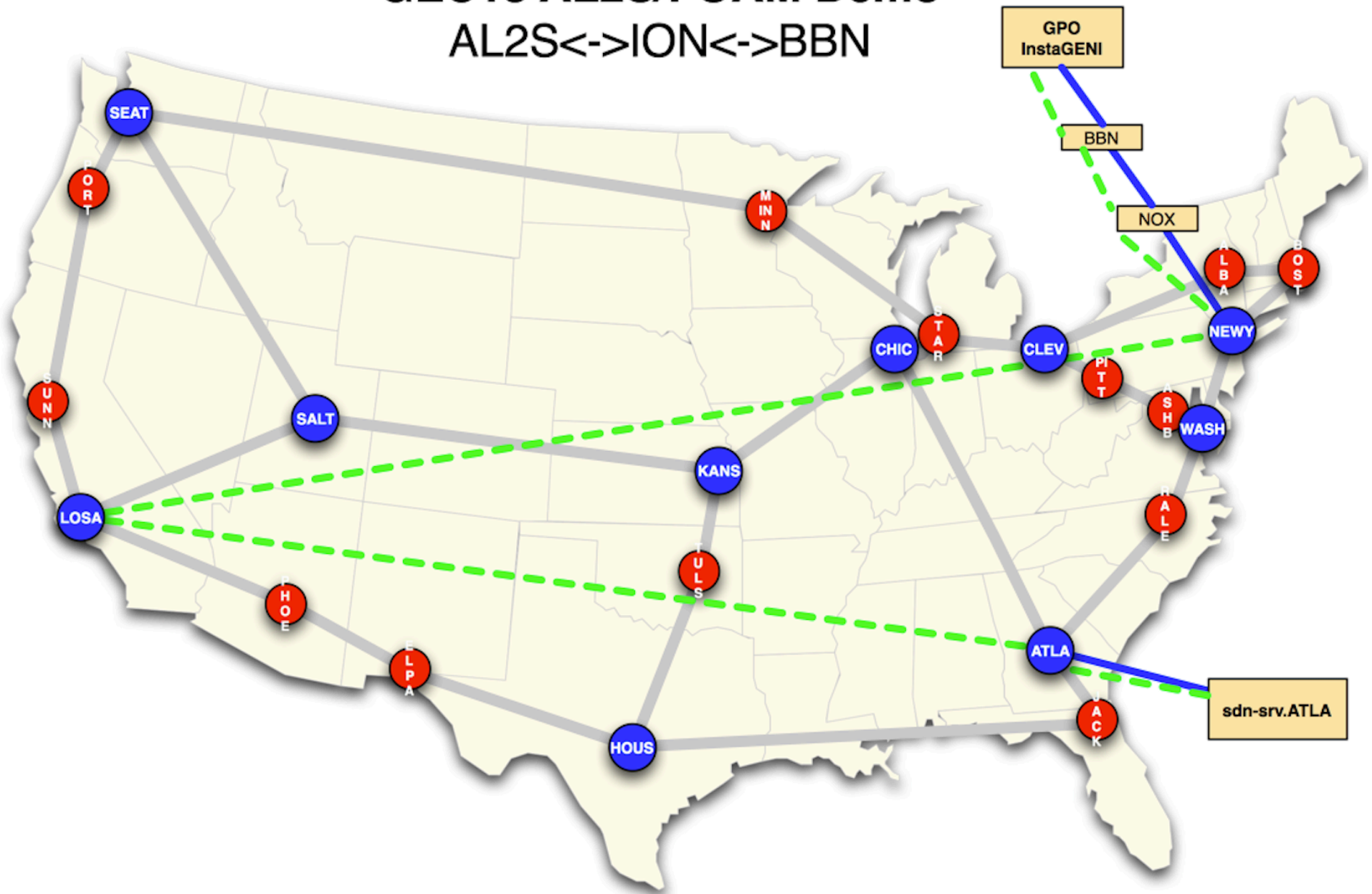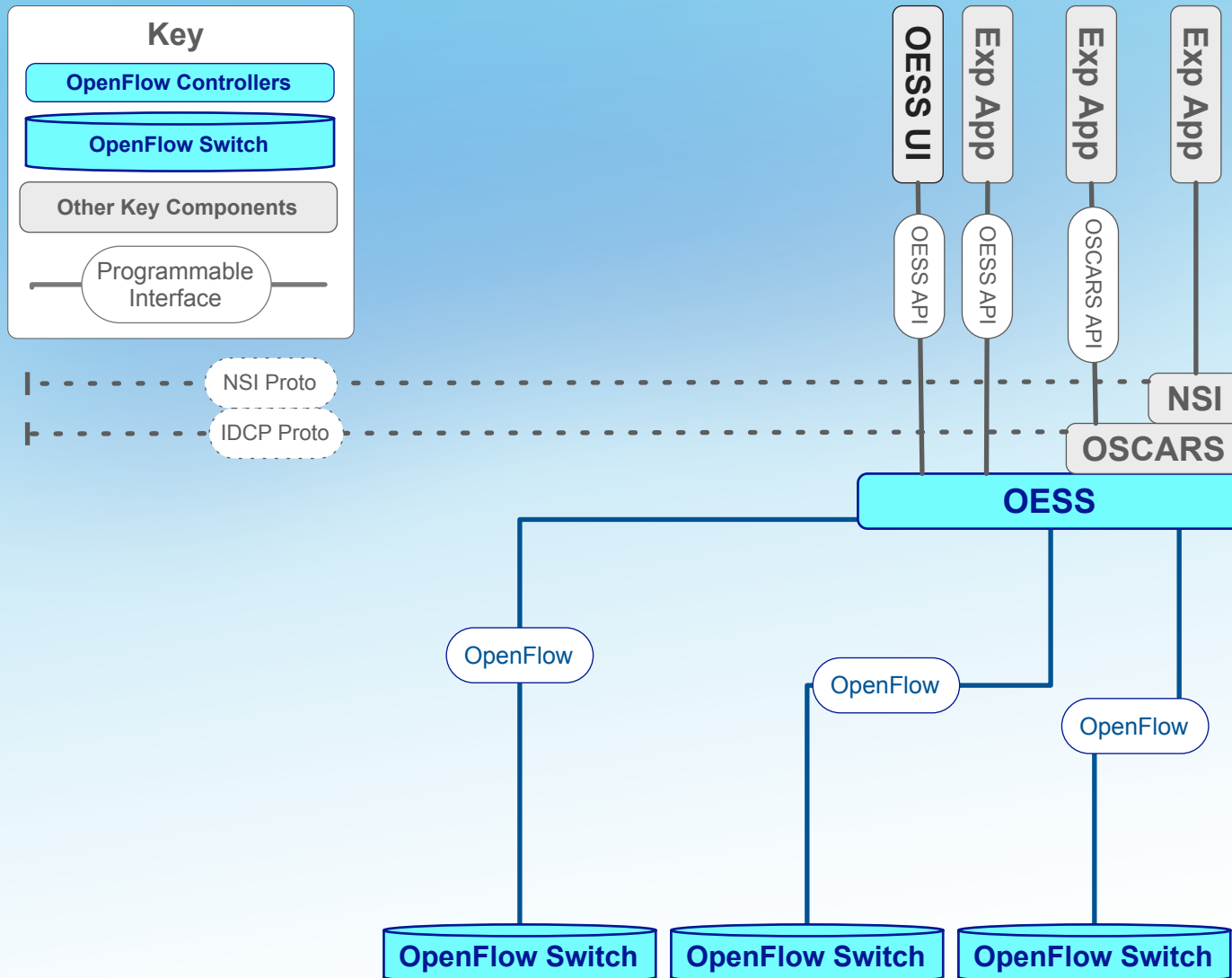- 100G AL2S Backbone

AL2S Software Stack Timeline

- Stitching - Software / Service Deployment Roll-out
  - ION (MAX) Aggregate Manager in Production (GEC17)
  - AL2S FOAM Aggregate Manager in Production (GEC18)
  - Stitching (GEC19)
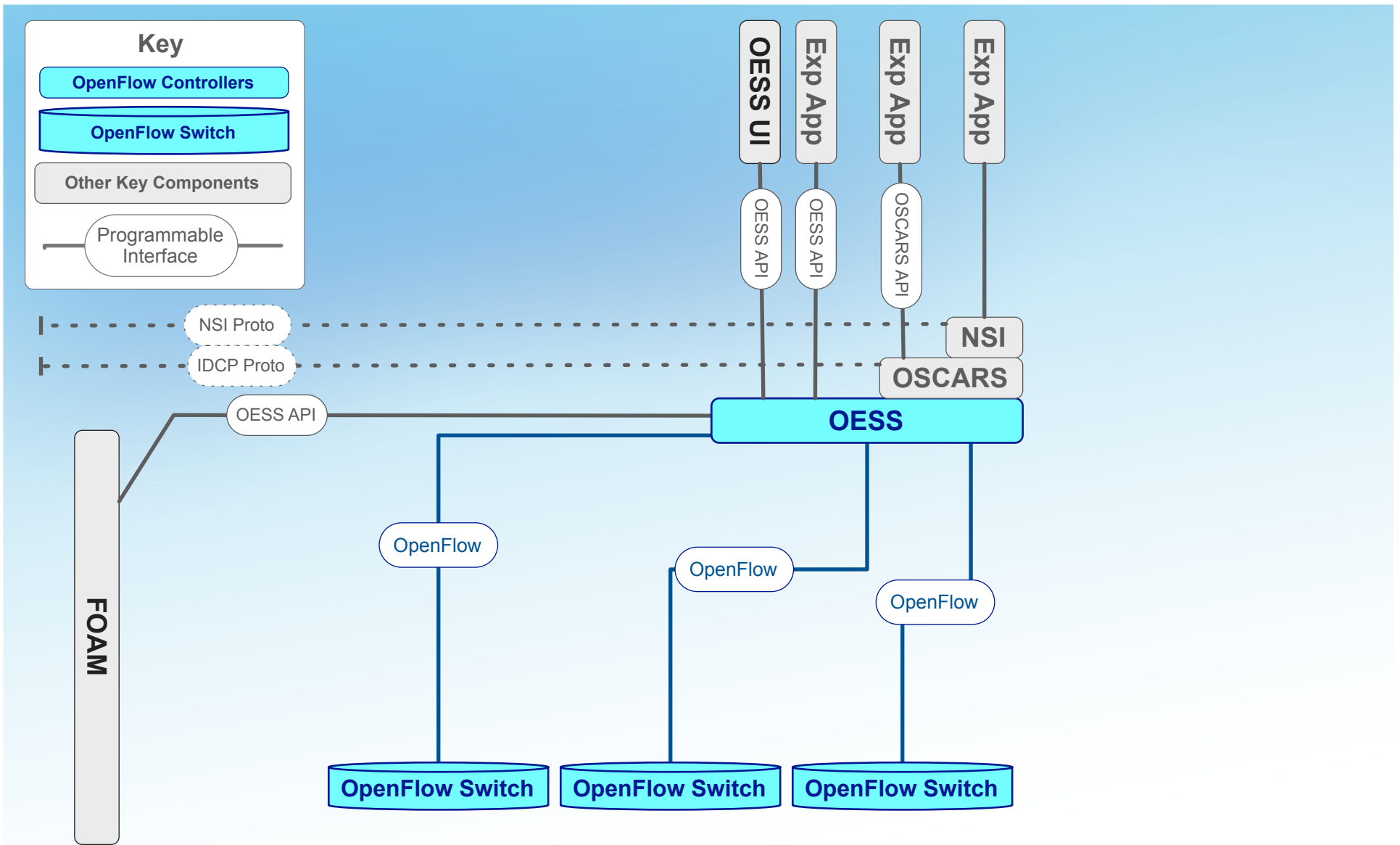
GEC18 AL2S/FOAM Demo
AL2S<->ION<->BBN

GPO InstaGENI

BBN

NOX

sdn-srv.ATLA

AL2S Software Stack   GEC17 Configuration

AL2S Software Stack   GEC18 Configuration

On–demand Secure Circuits and Advance Reservation System

https 🔒 ion.net.internet2.edu:8443/OSCARS/

OSCARS software developed by ESnet, in collaboration with Internet2 and IsI East

October 27, 2013 19:36

<div style="border: 2px solid green;">ACTIVE</div>

| Reservations | Reservation Details | Create Reservation | User Profile | Login/Logout |

NEW GRI [   ] [QUERY]

[REFRESH] [MODIFY] [CANCEL] [CLONE]

GRI	ion.internet2.edu-45641
Status	ACTIVE
User	geni_user
Description	urn:publicid:IDN+ch.geni.net:internet2test+slice+i2gec18 (link-ld
Start date	10/26/2013
Start time	19:27
End date	10/27/2013
End time	19:27
Created time	2013/10/26 19:27
Bandwidth (Mbps)	10
Source	urn:ogf:network:domain=ion.internet2.edu:node=rtr.losa:port=et-10/0/0:link=al2s
Destination	urn:ogf:network:domain=ion.internet2.edu:node=rtr.newy:port=ae0:link=bbn

	VLAN	Hop
	3951	urn:ogf:network:domain=ion.internet2.edu:node=rtr.losa:port=et-10/0/0:link=al2s
	n/a*	urn:ogf:network:domain=ion.internet2.edu:node=rtr.losa:port=et-10/0/0:link=et-10/0/0.201
Path	n/a*	urn:ogf:network:domain=ion.internet2.edu:node=rtr.wash:port=et-9/0/0:link=et-9/0/0.201
	n/a*	urn:ogf:network:domain=ion.internet2.edu:node=rtr.wash:port=et-3/0/0:link=et-3/0/0.202
	n/a*	urn:ogf:network:domain=ion.internet2.edu:node=rtr.newy:port=et-5/0/0:link=et-5/0/0.202
	3747	urn:ogf:network:domain=ion.internet2.edu:node=rtr.newy:port=ae0:link=bbn

Source VLAN	3951
Tagged	true

```
luke — luke@skip:~ — luke@skip:~ — ssh — 121×22

[  3] 7980.0-7982.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7982.0-7984.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7984.0-7986.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7986.0-7988.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7988.0-7990.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7990.0-7992.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7992.0-7994.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7994.0-7996.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7996.0-7998.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 7998.0-8000.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8000.0-8002.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8002.0-8004.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8004.0-8006.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8006.0-8008.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8008.0-8010.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8010.0-8012.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8012.0-8014.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8014.0-8016.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8016.0-8018.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8018.0-8020.0 sec   11.9 MBytes   50.0 Mbits/sec
[  3] 8020.0-8022.0 sec   11.9 MBytes   50.0 Mbits/sec
```

```
[luke@skip ~]$ ssh foam-oess-stage
luke@foam-oess-stage's password:

root@foam-oess-stage:/home/luke# su - gec18

gec18@foam-oess-stage:~$ stitcher.py createsliver i2gec18 rspecs/stitch-gpo-losa --fixedEndpoint
21:16:39 INFO     stitcher: Loading agg_nick_cache file '/home/gec18/.gcf/agg_nick_cache'
21:16:39 INFO     stitcher: Loading config file /home/gec18/.gcf/omni_config
21:16:39 INFO     stitcher: Using control framework portal
21:16:39 INFO     stitcher: Checking that slice i2gec18 is valid...
21:16:41 INFO     stitcher: Slice urn:publicid:IDN+ch.geni.net:internet2test+slice+i2gec18 expires on 2013-11-02
16:46:44 UTC
21:16:41 INFO     stitcher: <Aggregate urn:publicid:IDN+instageni.gpolab.bbn.com+authority+cm> speaks AM API v3,
but sticking with v2
21:16:41 INFO     stitcher: Stitched reservation will include resources from these aggregates:
21:16:41 INFO     stitcher:     <Aggregate urn:publicid:IDN+instageni.gpolab.bbn.com+authority+cm>
21:16:41 INFO     stitcher:     <Aggregate urn:publicid:IDN+ion.internet2.edu+authority+am>
21:16:41 INFO     stitcher: Adding fake Node endpoint
21:16:41 INFO     stitcher: Adding fake iref endpoint on link <DOM Element: link at 0x1ee00e0>
21:16:41 INFO     stitch.Aggregate: Writing to '/tmp/i2gec18-createsliver-request-11-instageni-gpolab-bbn-com.xml'
21:16:41 INFO     stitch.Aggregate:
       Stitcher doing createsliver at https://boss.instageni.gpolab.bbn.com:12369/protogeni/xmlrpc/am
21:16:55 INFO     stitch.Aggregate: Allocation at <Aggregate urn:publicid:IDN+instageni.gpolab.bbn.com+authority
+cm> complete.
21:16:55 INFO     stitch.Aggregate: Writing to '/tmp/i2gec18-createsliver-request-11-ion-internet2-edu.xml'
21:16:55 INFO     stitch.Aggregate:
       Stitcher doing createsliver at http://geni-am.net.internet2.edu:12346
21:17:38 INFO     stitch.Aggregate: DCN AM <Aggregate urn:publicid:IDN+ion.internet2.edu+authority+am>: must wait
for status ready....
21:17:38 INFO     stitch.Aggregate: Pause 30 seconds to let circuit become ready...
21:18:15 INFO     stitch.Aggregate: DCN circuit 45591 is ready
21:18:21 INFO     stitch.Aggregate: Allocation at <Aggregate urn:publicid:IDN+ion.internet2.edu+authority+am>
```

INTERNET2

OS³E
The Open Science, Scholarship & Services Exchange

Feedback Admin

Workgroups > Home > **Circuit Details**

Workgroup: GENI-FOAM

Summary

Description
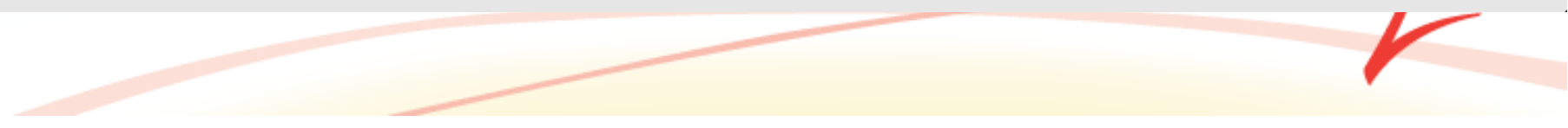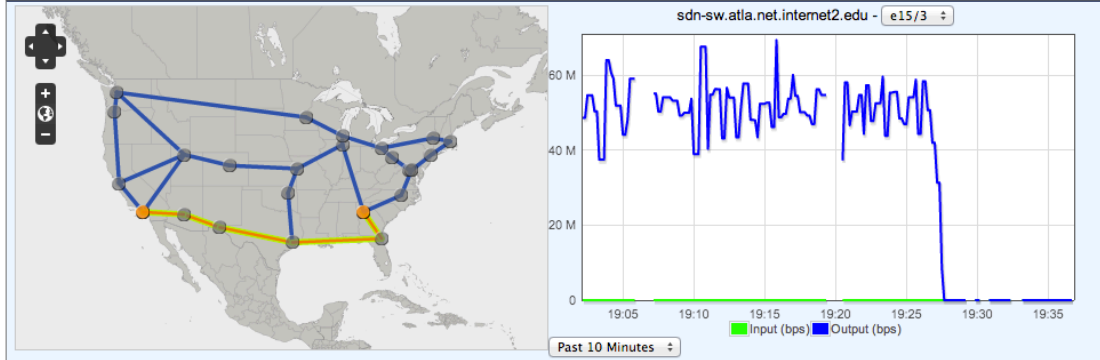urn:publicid:IDN+ch.geni.net:internet2-test+slice+gec18:b0d71581-38c8-4071-bc47-56719e5059fc

Type
Local

Restore To Primary
Off

Status
active

Endpoints

Interface	Interface Description	VLAN
sdn-sw.losa.net.internet2.edu - e1/1	rtr.losa, et-10/0/0	3951
sdn-sw.atla.net.internet2.edu - e15/3	sdn-srv.atla, em2	3951

Edit Circuit

Remove Circuit

Force Reprovision

| **Utilization** | History | Scheduled Events | Circuit Layout | Circuit Layout Raw |

sdn-sw.atla.net.internet2.edu - e15/3

60 M

40 M

20 M

0

19:05   19:10   19:15   19:20   19:25   19:30   19:35

Input (bps)   Output (bps)

Past 10 Minutes

GlobalNOC
Global Research Network Operations Center

Internet2 Network

Stanford – Global Research NOC – Indiana University – Internet2 – OESS – About

Aggregate Manager over AL2S - Testing

- The GENI OESS Test Plan outlines the test goals for the effort and can be found at http://groups.geni.net/geni/wiki/GENIOESSTopologiesTestPlan

- Progress has been made for Functional Tests and for some Cross Connect Tests. Detailed logs and bugs found for tested areas are at http://groups.geni.net/geni/wiki/GENIOESSTopologiesTestStatus

INTERNET2

Functional Testing

- Functional tests have used the GENI AM API V2 interface to successfully
  - List, request, get status and renew resources.
  - Verify ability to request AL2S to AL2S endpoint with:
    - Primary Path and Backup Path
    - Primary Path only
    - No path (endpoints co-locates)
  - Verified recovery from network disruptions:
    - with backup paths
    - without backup path
- Longevity tests had connections with ping traffic for over 24 hours between the following sites:
  - Atlanta to Los Angeles
  - New York to Los Angeles
  - Washington to New York
  - Atlanta to New York
- Verified that OESS Web interface accurately captured traffic load and interruptions for all tests in the test Work group.

INTERNET

ION to AL2S Connection Testing

- Once a functional baseline was established testing moved to ION and AL2S OESS Cross connects.
- Used the the GENI AM API V2 interface to request
    - AL2S test endpoints and primary paths from the OESS FOAM aggregate
    - ION endpoints from the GENI Stitching Computation service via Internet2 ION aggregate.
- For this test a set of 10 cross connects were configured for the GENI Stitching Computation Service:

- 1. sdn-sw.losa e1/1 <-> rtr.losa:port=et-10/0/0
    2. sdn-sw.atla e15/1 <-> rtr.atla:port=xe-0/3/0
    3. sdn-sw.chic e3/1 <-> rtr.chic:port=et-10/0/0
    4. sdn-sw.clev e5/1 <-> rtr.clev:port=et-5/0/0
    5. sdn-sw.hous e15/3 <-> rtr.hous:port=xe-0/1/3
    6. sdn-sw.kans e15/1 <-> rtr.kans:port=xe-0/0/3
    7. sdn-sw.newy32aoa e3/2 <-> rtr.newy:port=et-5/0/0
    8. sdn-sw.salt e15/1 <-> rtr.salt:port=xe-0/1/1
    9. sdn-sw.seat e-2/0/0.0 <-> rtr.seat:port=et-5/0/0
    10. sdn-sw.wash e5/2 <-> rtr.wash:port=et-9/0/0

- All ten were successfully used to exchange traffic with the same two host endpoints (GPO InstaGENI VM and AL2S Atlanta test endpoint host) but changing the path between them to use each of the ten above. Traffic was ping only.or all tests in the test Work group.

INTERNET2

Performance Measurement

- Some performance information was captured for:
  - AL2S to AL2S connections
  - ION to AL2S cross connections
- Summary results can be found at:
  - http://groups.geni.net/geni/wiki/GENIOESSTopologiesPerformance .
- Note: As can be seen from the summary, the AL2S to AL2S connections perform well while the same cannot be said for the ION to AL2S connections.
- May be a tuning issue or testing artifact.
- Need to investigate the traffic generation variability for the ION test endpoint, after GEC 18.

INTERNET2

Unresolved Issues

- The most critical issues found are in the area of the OESS aggregate status.
- The OESS Servers does not report failures that occur after the request is submitted.
- There is no visible user notification if a VLAN is not provisioned (ticket 1141) .
- In addition to the sliver status not being updated for failures; when a sliver is ready, the sliver status remains in "configuring" state (ticket 1140).
- There is one critical bug for resource management, the aggregate does not delete slivers when they expire(ticket 1142).
- The full list list of unresolved issues can be found at: http://groups.geni.net/geni/query?status=new&status=assigned&status=reopened&keywords=~oess&order=priority

INTERNET2

This is what we have been able to say for about a year:
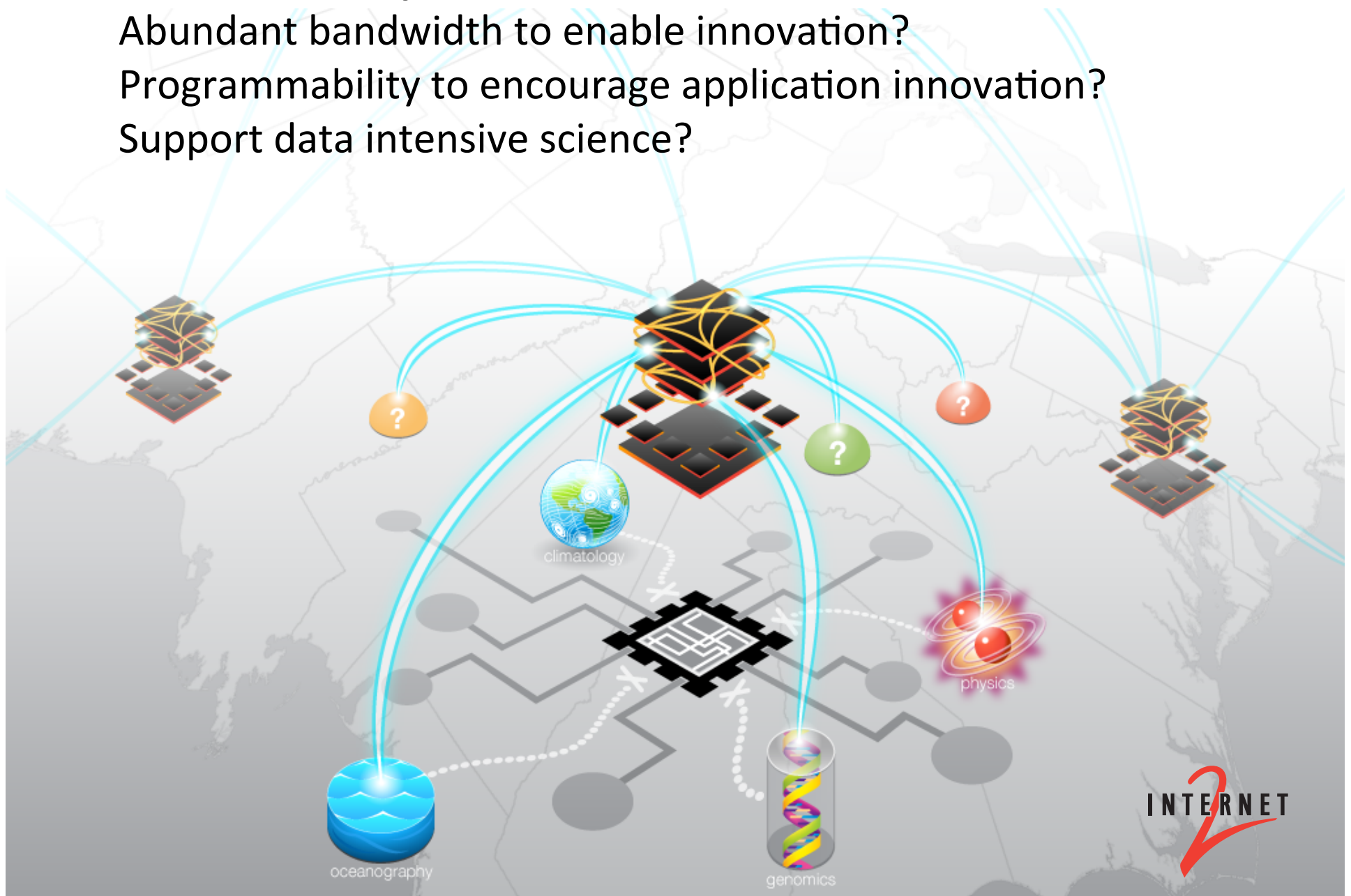The **100G** testbed of innovation for tomorrow's Internet is available nationwide, right now.
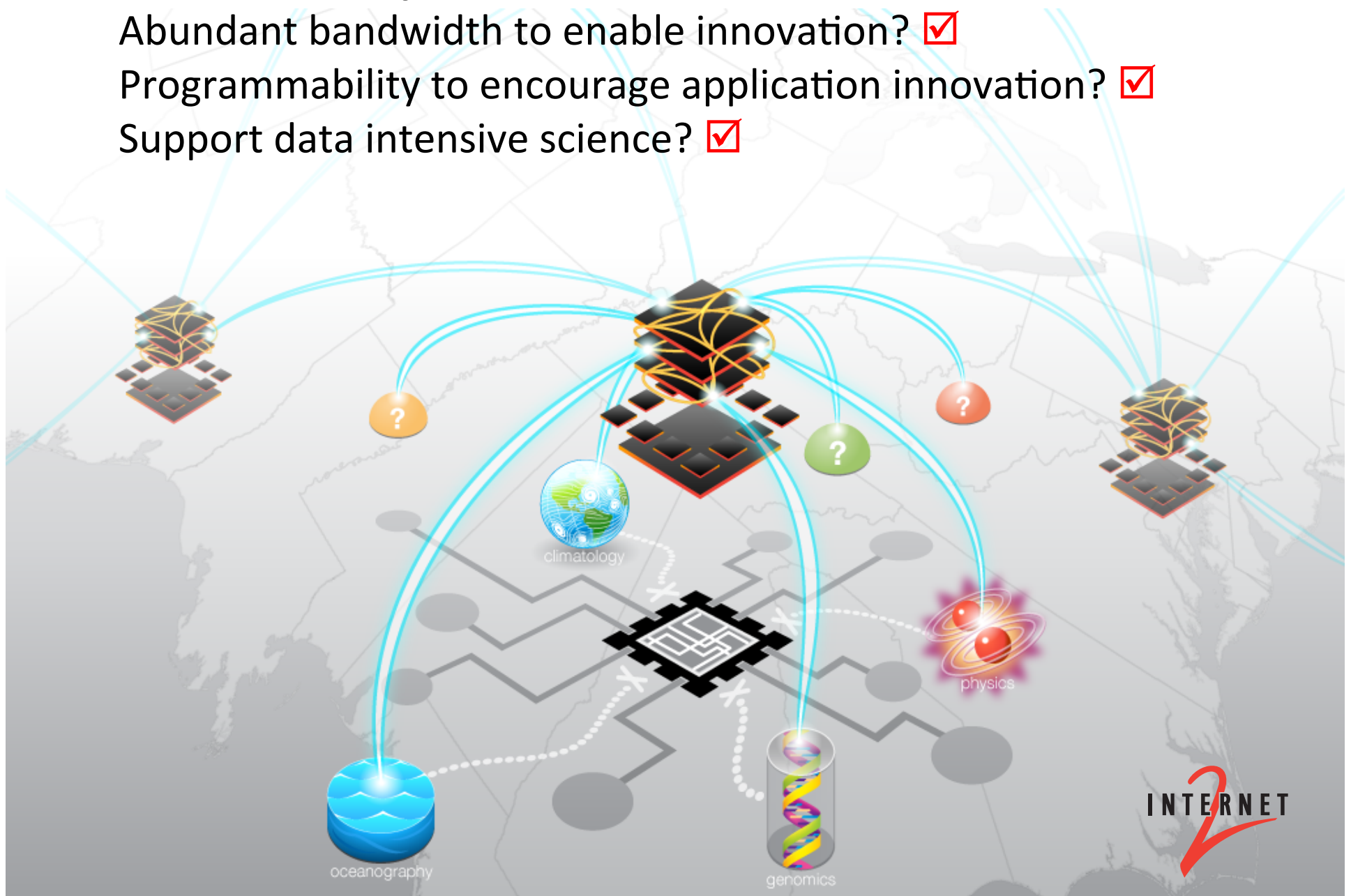
Does this create a platform for innovation?
   Abundant bandwidth to enable innovation?
   Programmability to encourage application innovation?
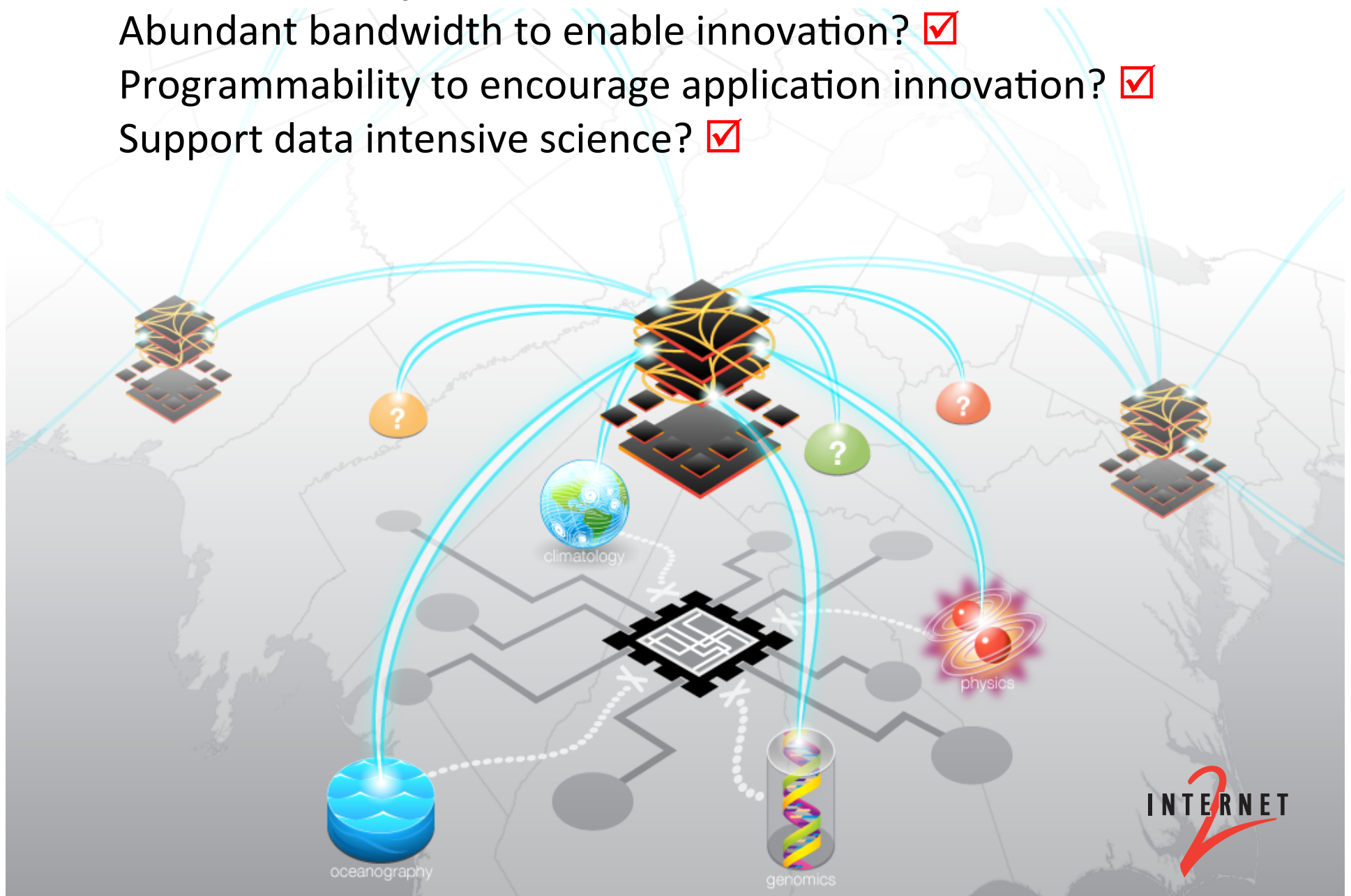   Support data intensive science?

INTERNET2

Does this create a platform for innovation?

Abundant bandwidth to enable innovation? ☑

Programmability to encourage application innovation? ☑

Support data intensive science? ☑

Does this create a platform for innovation?

Abundant bandwidth to enable innovation? ☑

Programmability to encourage application innovation? ☑

Support data intensive science? ☑

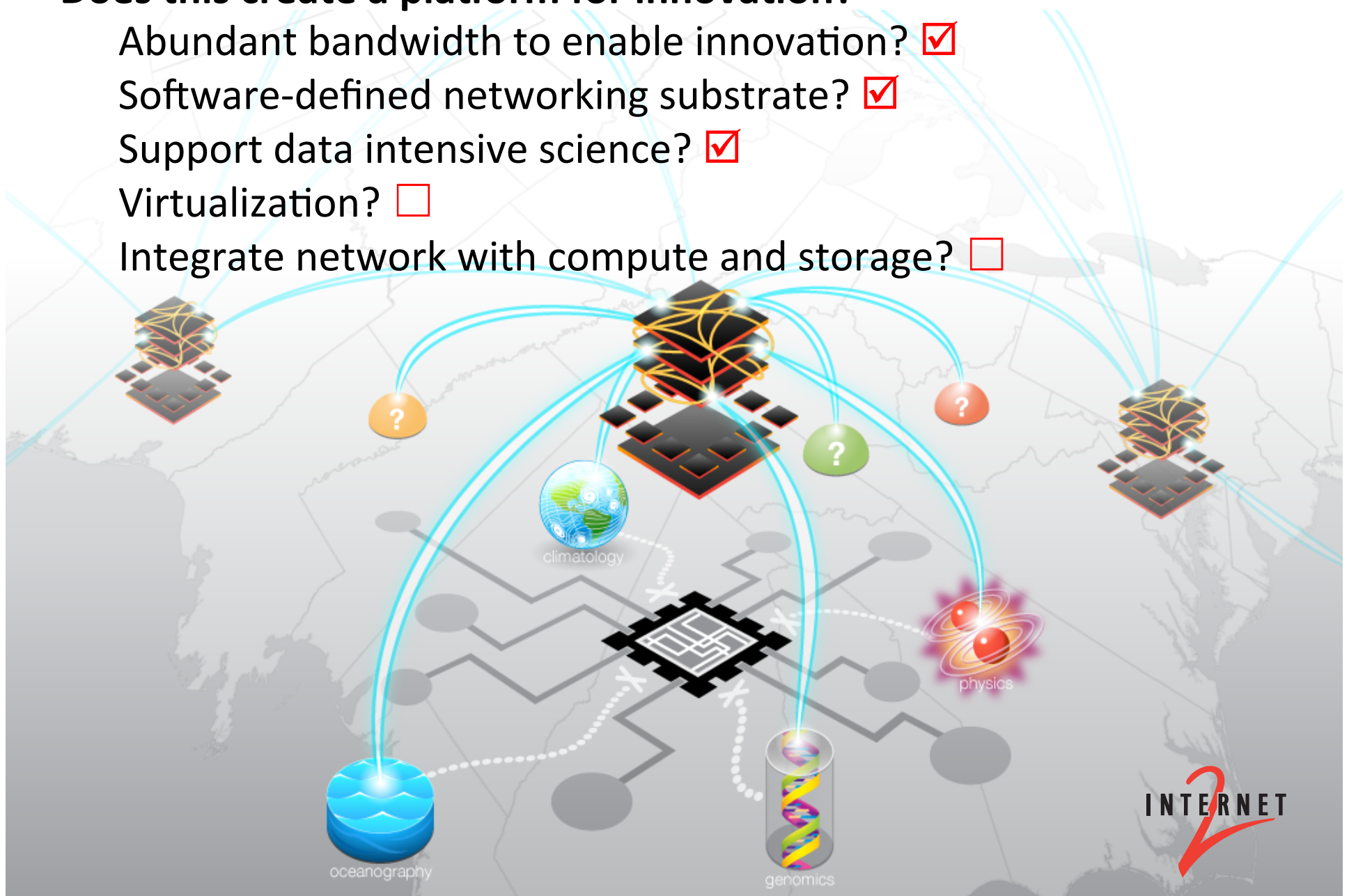**Does this create a platform for innovation?**

Abundant bandwidth to enable innovation? ☑

Software-defined networking substrate? ☑

Support data intensive science? ☑

Virtualization? ☐

Integrate network with compute and storage? ☐

Does this create a platform for innovation?

Abundant bandwidth to enable innovation? ☑
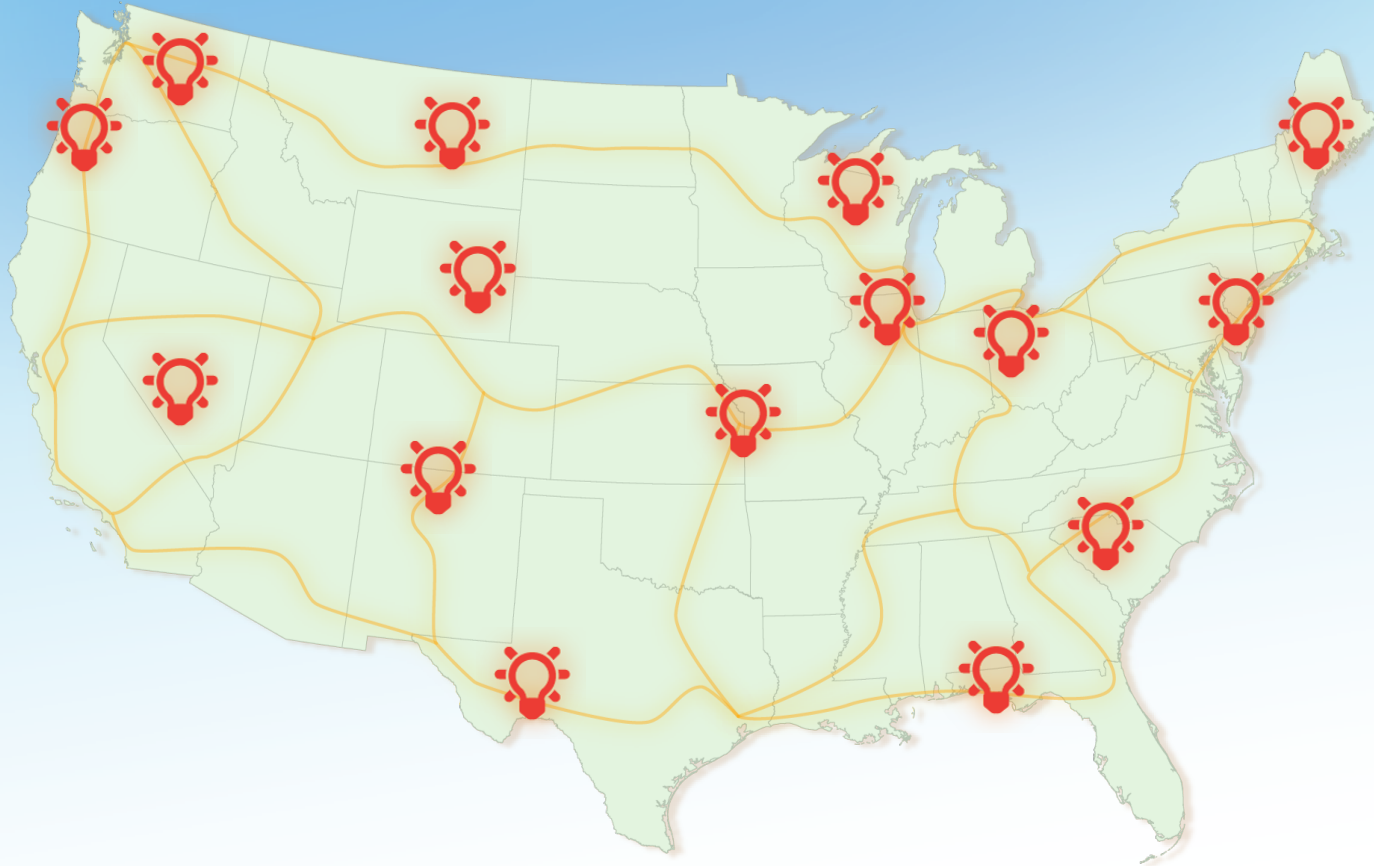
Software-defined networking substrate? ☑

Support data intensive science? ☑

Virtualization? ☐

Integrate network with compute and storage? ☐

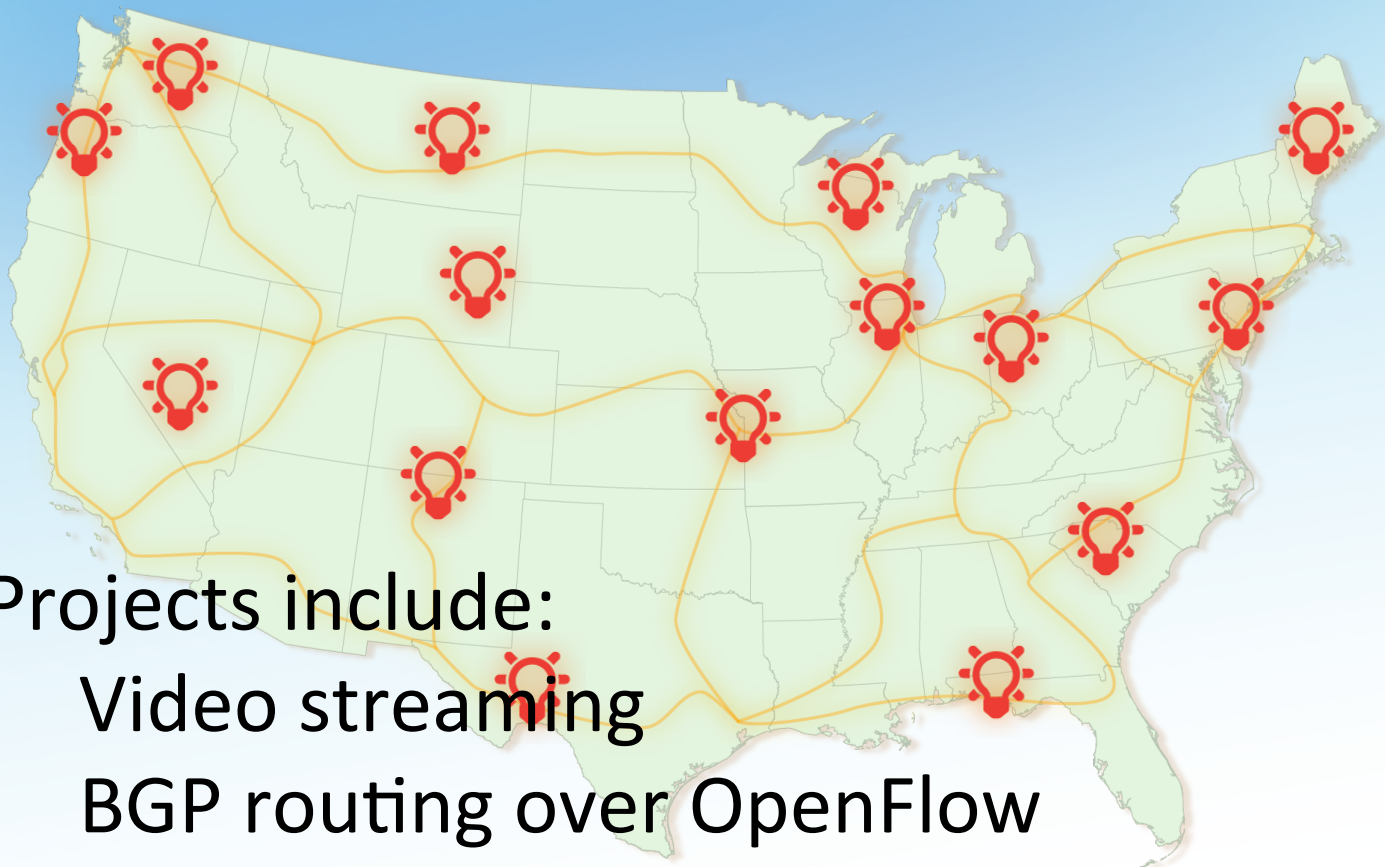# 2013 Internet2 Innovative Application Awards
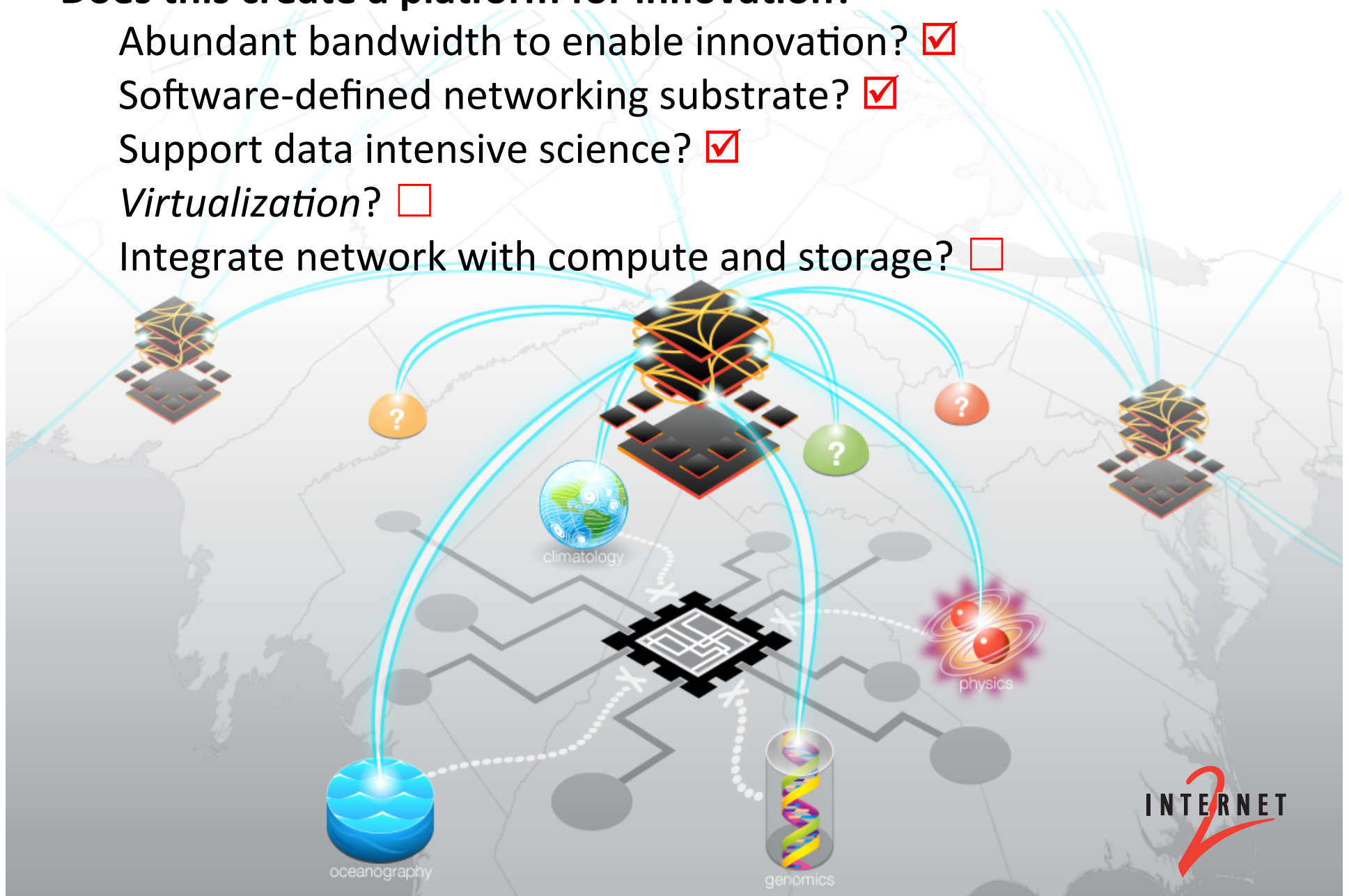
GOLD

JUNIPER
NETWORKS®

SILVER

ciena.

BRONZE

BROCADE

INTERNET 2

2013 Internet2 Innovative Application Awards

GOLD

JUNIPER
NETWORKS

SILVER

ciena

BRONZE

BROCADE

Projects include:
  Video streaming
  BGP routing over OpenFlow
  SDN for Exchange Points
  GENI integration

INTERNET2

- We've had virtualization of storage and servers for quite some time

- How to define Network Virtualization?

- "Virtualization is the core principle in overlays, both allowing nodes to treat an overlay as if it were the native network, and allowing multiple overlays to simultaneously use the same underlying overlay infrastructure." (2004 – Anderson, Peterson, Shenker, Turner)

- **So what does that mean in a practical sense?**
  - **Decouple control plane from data plane**
  - **Enable multiple virtual control planes on a common physical data plane**
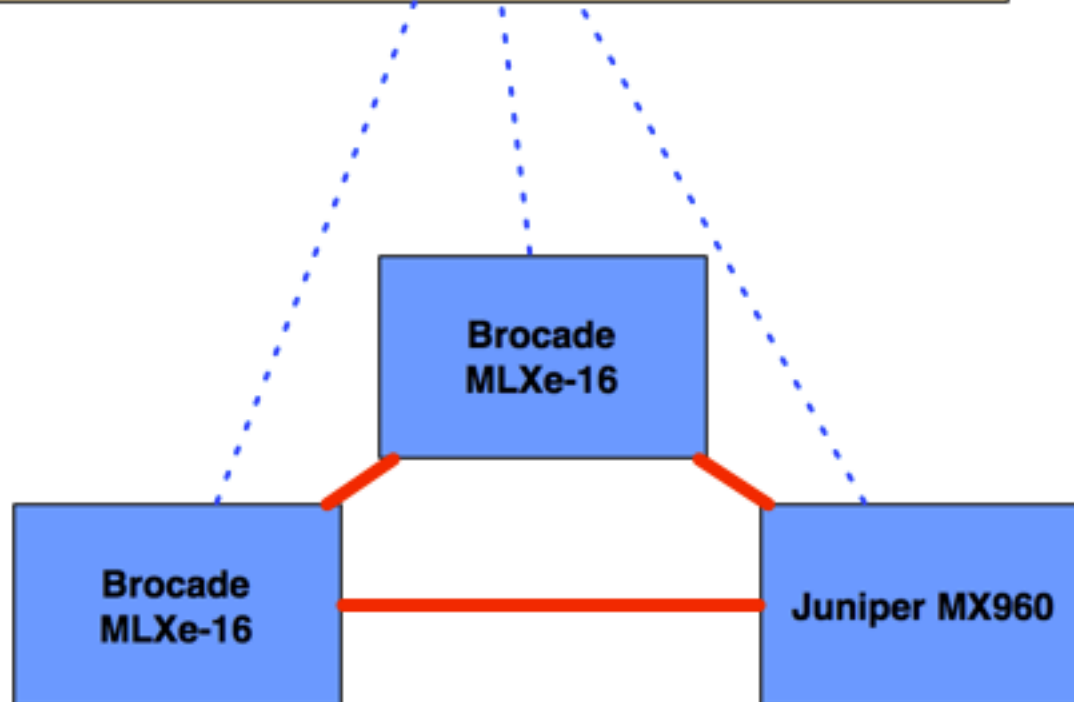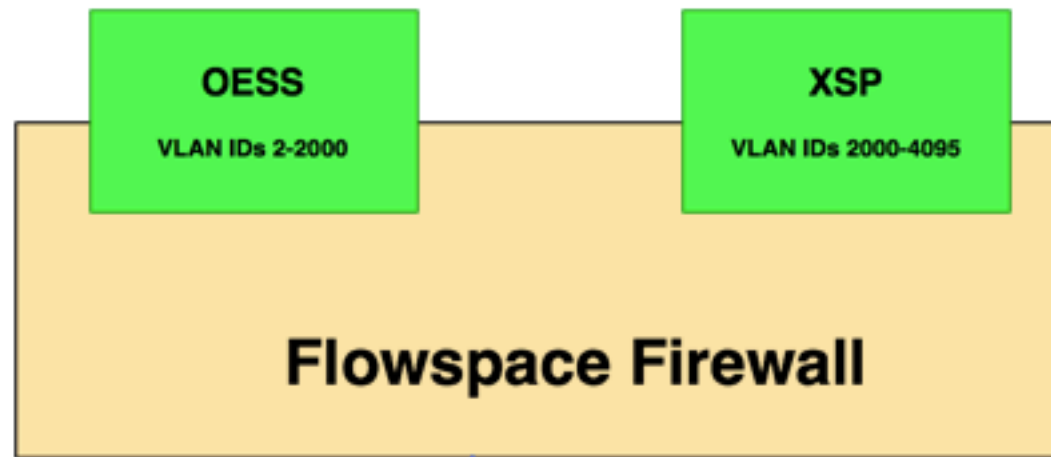
- **Internet2 in partnership with Indiana University has been building / deploying an architecture to support network virtualization**
  - **Provide network multi-tenancy at Layer 2 and Layer 3**
  - **Enforce non-overlapping Layer 2 tag-based flowspace**
    - **Experiment Foo can use VLAN tag range 1-200 (a sliver)**
    - **Experiment Bar can use VLAN tag range 201-400 (a sliver)**

- **How to implement virtualization?**
  - **First we looked at Flowvisor**
  - **Then we realized we needed something slightly different … Flowspace Firewall**

Experimenter Control of Internet2

- Now support Experimenter Control of lab instance of Internet2 SDN Substrate
  - iDREAM GENI
- Aim to support Experimenter Control of Internet2 SDN Substrate

INTERNET2

monitor.incntre.iu.edu/fsfscope/topology/generic/1/

Periscope Testbed Monitoring

Topologies

- fsf

Misc

Transfer Rate

MB/s

0

sw3

pc1

pc2

SDN-Testlab

sw2

sw1

Floodlight | XSP

❌ Delete selected paths 💾 Save topology locations

| GRI	Status	Source	Destination	Start Time	Duration (s)	Misc	
XSP-netPath-1542	DOWN	pc1	pc2	2013-10-27 19:35:21	...	alt	
XSP-netPath-1543	DOWN	pc1	pc2	2013-10-27 19:35:46	...		
XSP-netPath-1544	DOWN	pc1	pc2	2013-10-27 19:36:10	...	alt	
XSP-netPath-1545	DOWN	pc1	pc2	2013-10-27 19:36:35	...		
XSP-netPath-1546	UP	pc1	pc2	2013-10-27 19:37:00	...	alt	

reen Shot 2013–10–27 at 5 09 39 PM

DRAFT AL2S Software Stack Timeline

- Experimental Controller - Software Deployment Roll-out
  - Flowspace Firewall (Late Q4 2013)
  - First Experimental App (Early Q1 2014)
  - First Experimental App over Flowspace Firewall (Late Q1 2014)
- Experimental Controller - Service Roll-out
  - Operational Burn-In of Flowspace Firewall (Q1 2014)
  - GPO Burn-In of Flowspace Firewall (Late Q1 2014)
  - First Testbed-Validated Experimental App (Q2 2014)
  - FOAM-initiated slice for Testbed-Validated Experimental App (Late Q2 2014)
  - FOAM-initiated slice for Experimental App (Possible?)

INTERNET2

DRAFT GENI Experimental Controller Process

- GOAL: Run an experimental controller on an at-scale nationwide production network
- What comes first: Community Conversation
  - Owners of AL2S ports
  - Network researchers (e.g. GENI community)
  - Network operators
- While we're in the early stages: Example Process
  - Contact that NOC and ask for a slice
  - Deploy app on testbed (e.g. NEC-based NDDI, AL2S-like iDREAM) running Flowspace Firewall software stack
  - Experimenter / NOC jointly develop app test plan
  - NOC validates that app is safe
  - FOAM-initiated slice provisioned / experimental controller deployed
  - Run experiment on AL2S

INTERNET2
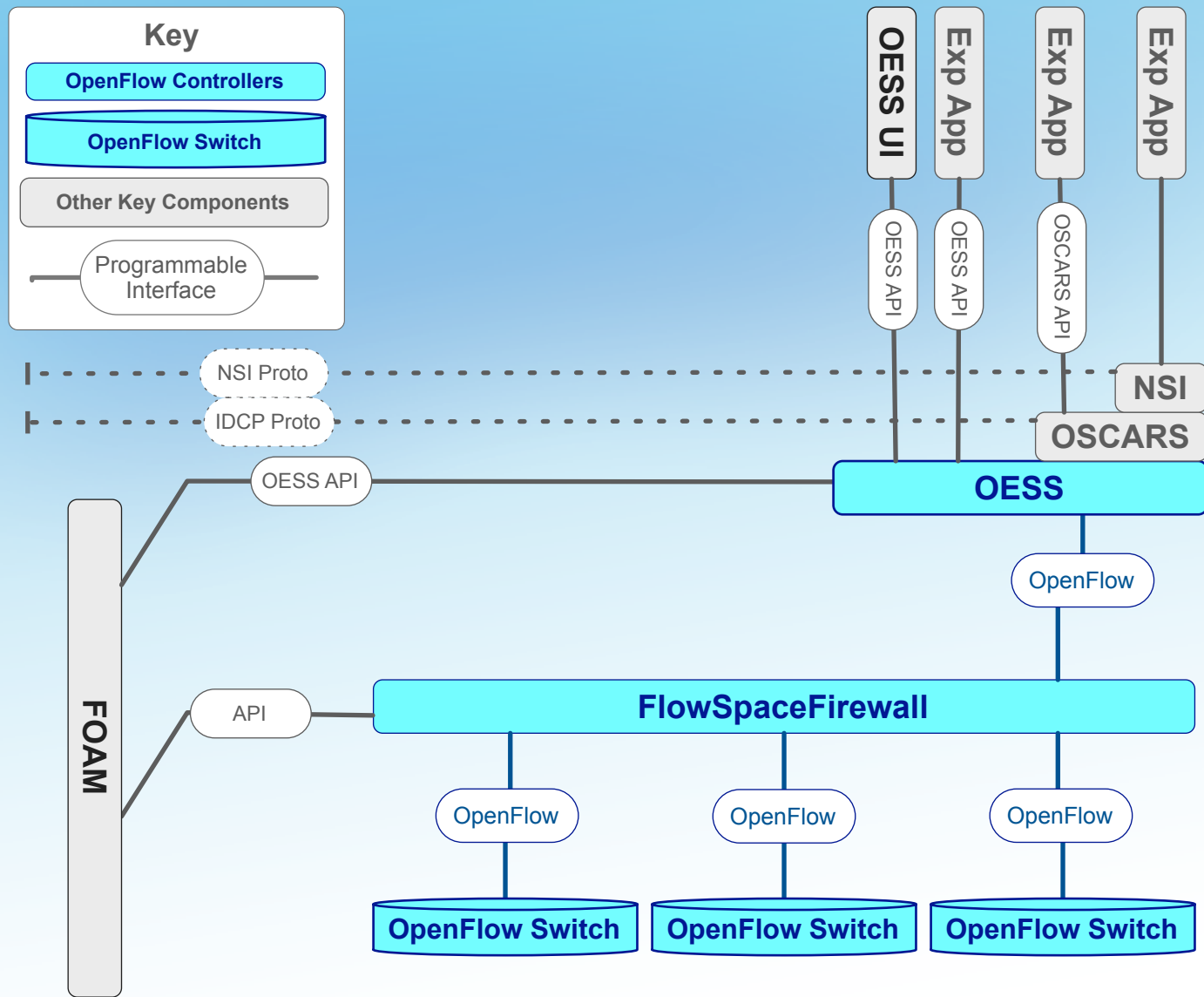
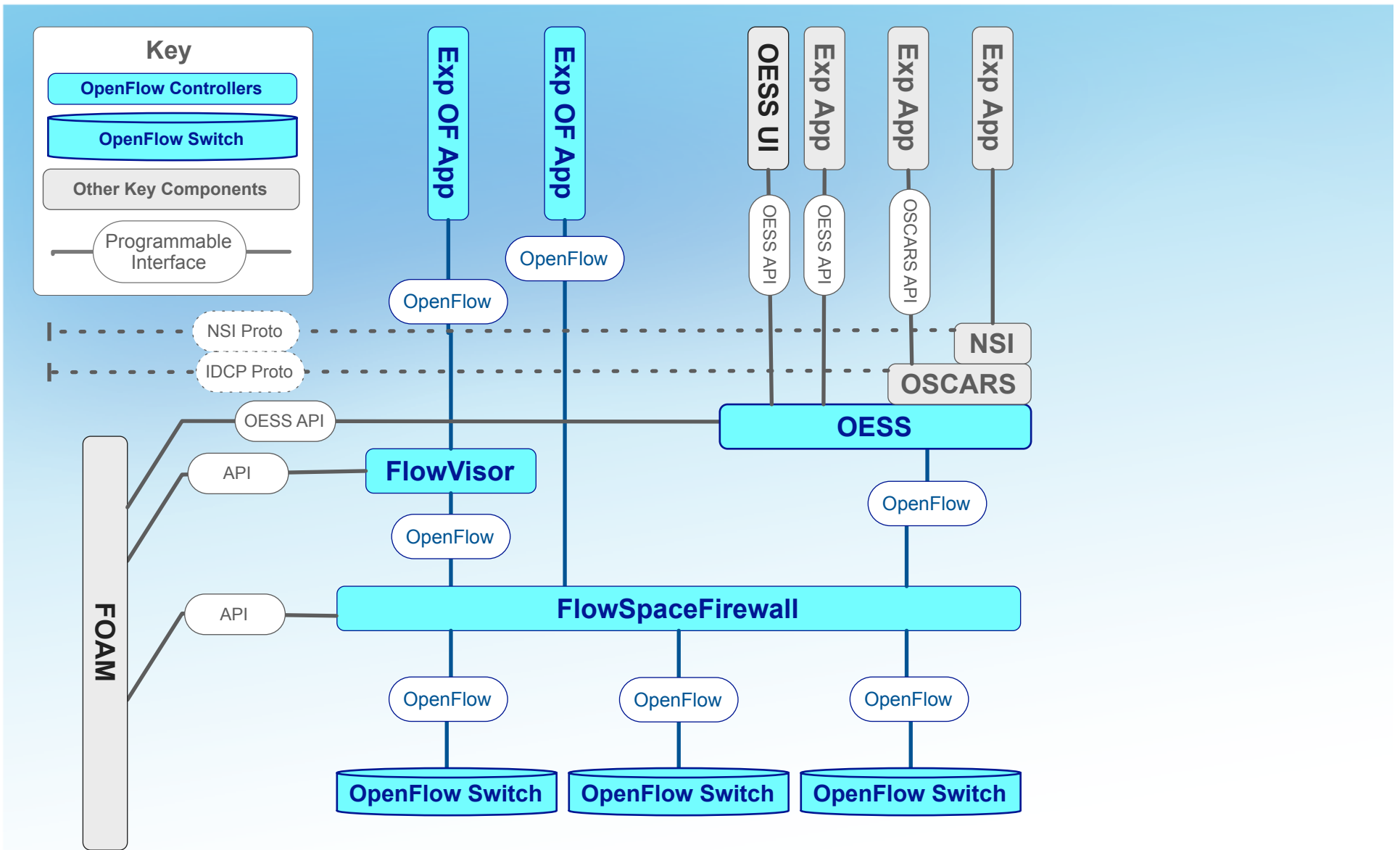AL2S Software Stack   Late Q4 2013?
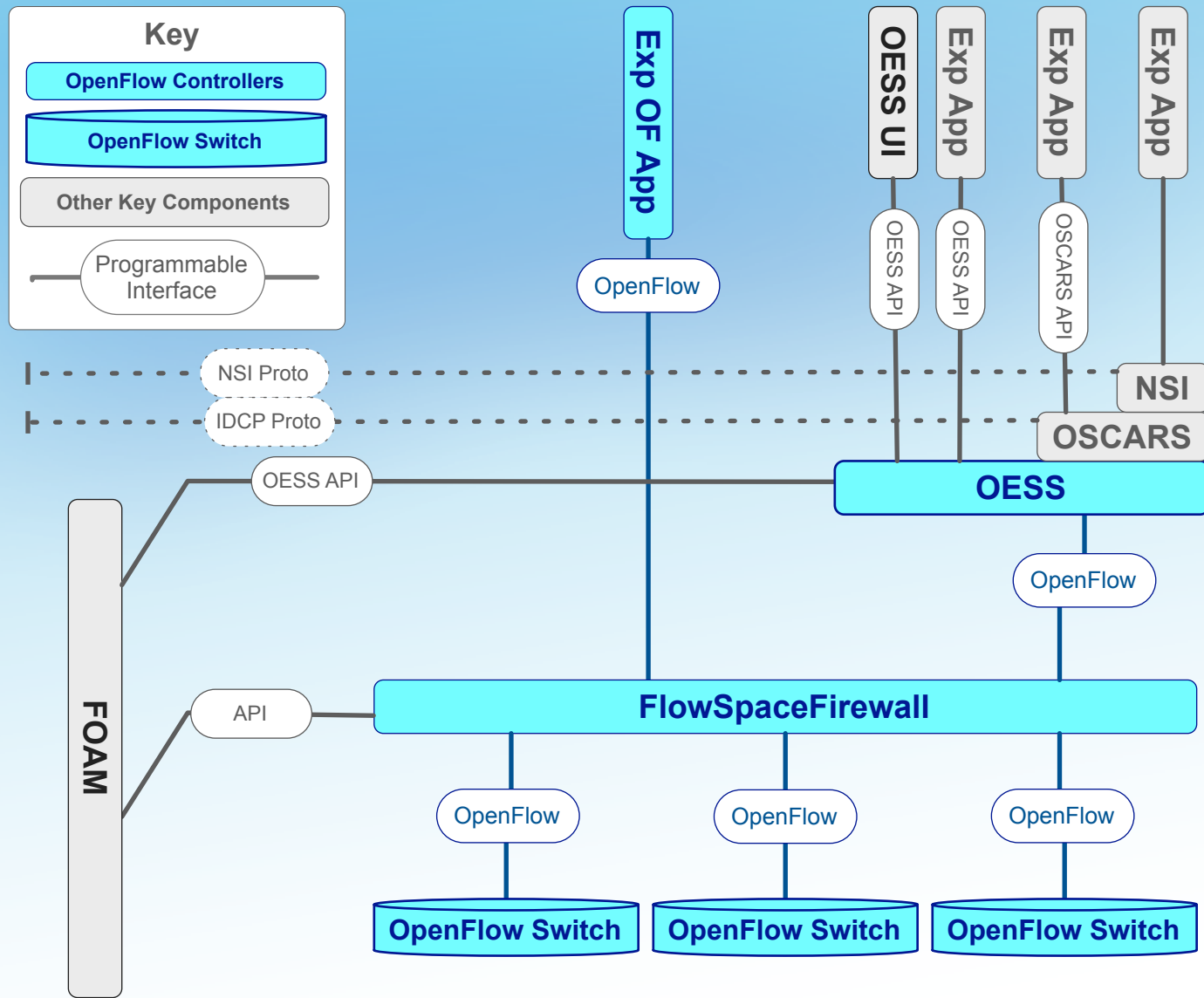
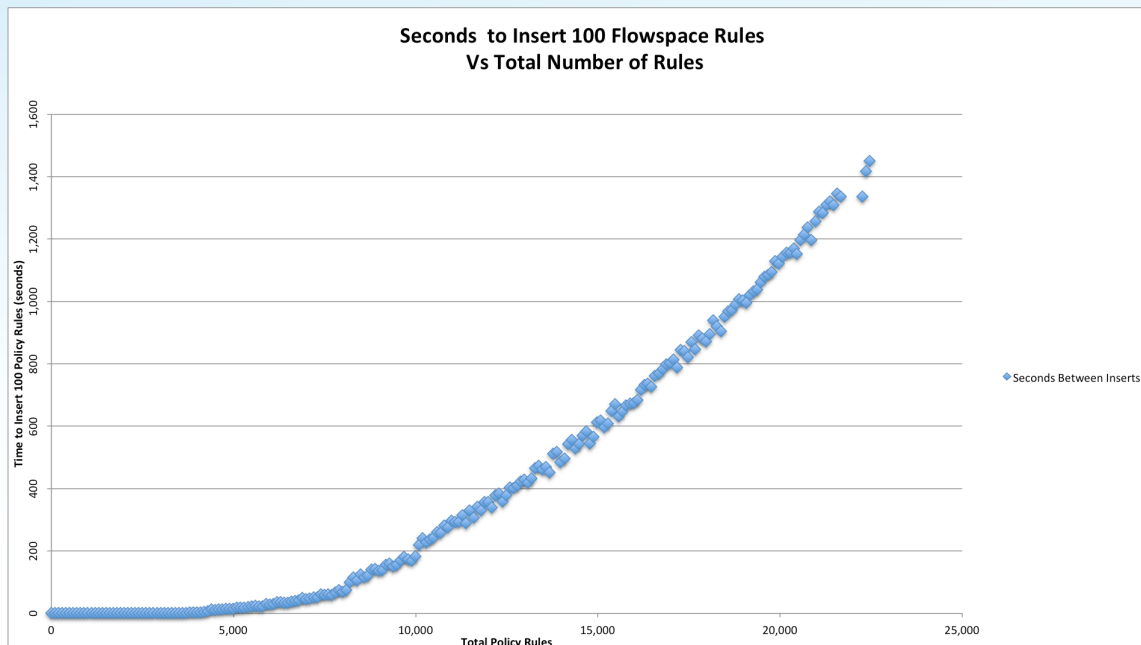AL2S Software Stack   Q1 Early 2014?

AL2S Software Stack   Q1 Late 2014?

FlowVisor Performance Issues

- Does not support VLAN Tag range-based policy
  - need 1 policy for every tag on every port in a flowspace
- ~ 1 million policy rules for the AL2S network
  - 28 switches, 10ports each, 4096 policies per port
- Unable to load this many rules in a acceptable time
  - Non-linear

Seconds to Insert 100 Flowspace Rules
Vs Total Number of Rules

Time to Insert 100 Policy Rules (seconds) — y-axis: 0, 200, 400, 600, 800, 1,000, 1,200, 1,400, 1,600

Total Policy Rules — x-axis: 0, 5,000, 10,000, 15,000, 20,000, 25,000

◆ Seconds Between Inserts

INTERNET2

FlowVisor Usability Issues

- Policy defined using port numbers not names
  - Port numbers on some systems are ephemeral
  - Difficult for humans to parse
- Policy defined using DPID vs symbolic name
  - DPID on some systems is ephemeral
  - Difficult for humans to parse

```
rule 6182:
FlowEntry[dpid=[00:00:00:a0:a5:7a:d7:34],ruleMatch=[OFMatch[in_port=59590,dl_vlan=4092]],actio
nsList=[Slice:nddi=7],id=[7200],priority=[10],]
rule 6183:
FlowEntry[dpid=[00:00:00:a0:a5:7a:d7:34],ruleMatch=[OFMatch[in_port=59590,dl_vlan=4093]],actio
nsList=[Slice:nddi=7],id=[7201],priority=[10],]
rule 6184:
FlowEntry[dpid=[00:00:00:a0:a5:7a:d7:34],ruleMatch=[OFMatch[in_port=59590,dl_vlan=4094]],actio
nsList=[Slice:nddi=7],id=[7202],priority=[10],]
```

INTERNET2

Looking beyond FlowVisor

- FlowVisor was designed to provide Flowspace translation
- Translating VLAN tags requires a 1 to 1 mapping
  - Architectural issues behind this
- For AL2S we are more interested in protection than translation
- We need a firewall to keep an OpenFlow application within its defined slice. Slice isolation is essential.
- After working with OnLab, we came to agreement that a separate application would be the most expedient path to resolve
- We need a FlowSpace Firewall.

INTERNET2

FlowSpace Firewall

- Simple VLAN Tag based flowspace firewall / proxy
- Policy definition and enforcement support range operations
  - < 1,000 policies to support 3 slices using the entire flowspace
- Per slice total rule limits
- Per slice per switch flow modification rate limits *(planned)*
- Built upon FloodLight
- Designed for production use.

Developed by Internet2 with GlobalNOC Software Engineering

FlowSpace Firewall Config Example

```
<flowspace_firewall>
   <switch name="foo"  dpid="5" flush_rules_on_connect="false" />
   <switch name="foo1" dpid="2" flush_rules_on_connect="false" />
   <switch name="foo2" dpid="3" flush_rules_on_connect="false" />
   <switch name="foo3" dpid="4" flush_rules_on_connect="false" />

   <slice name="OESS1">
     <switch name="foo" max_flows="10" flow_rate="1">
       <port name="s5-eth1">
        <range start="1" end="2000"/>
       </port>
       <port name="s5-eth2">
         <range start="1" end="2000" />
       </port>
     </switch>

     <controller ip_address="140.182.45.45" ssl="false" port="6633" />

   </slice>
</flowspace_firewall>
```
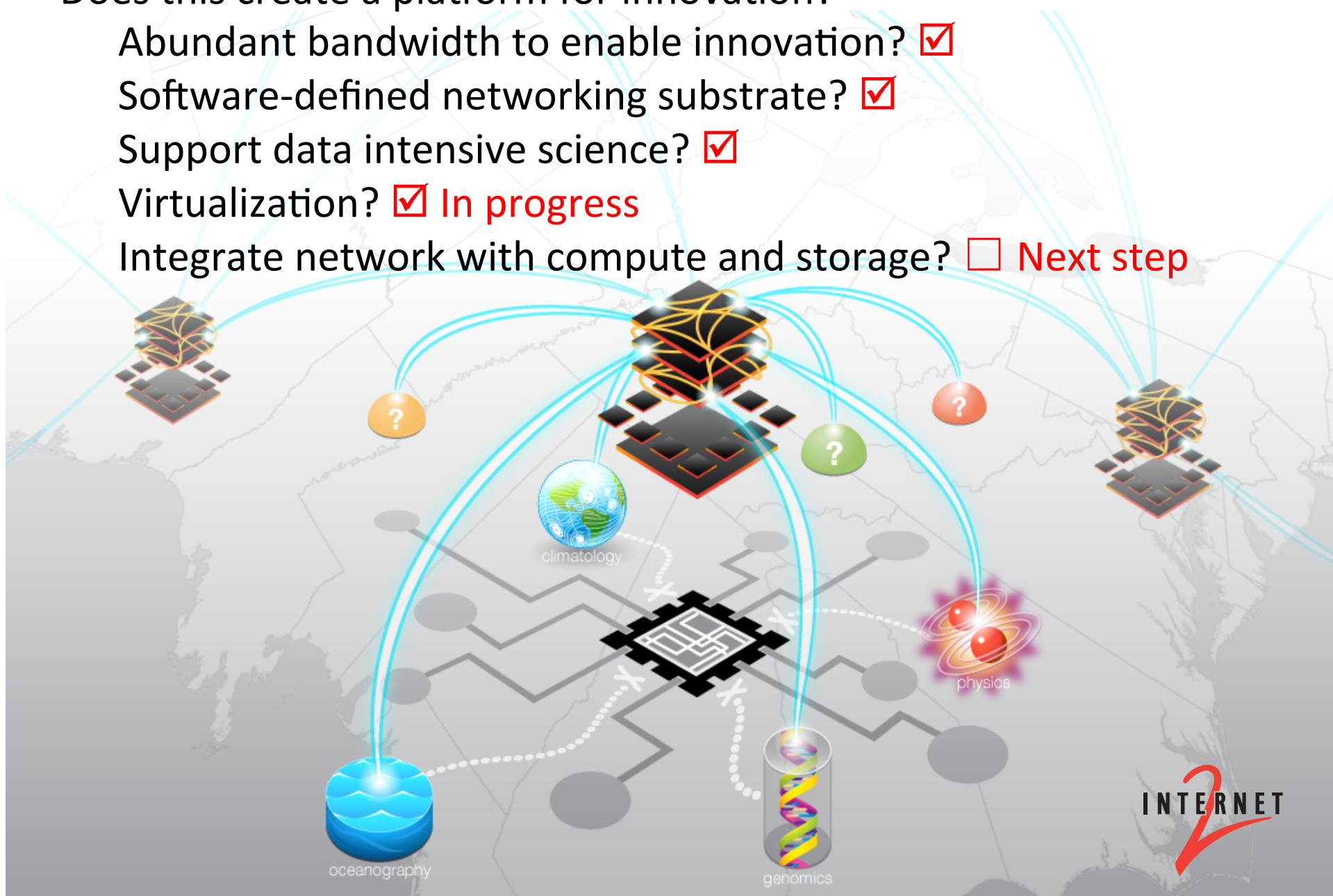
Symbolic names reduce policy churn

limits protect network

Range expression for sanity

INTERNET2

Does this create a platform for innovation?

    Abundant bandwidth to enable innovation? ☑

    Software-defined networking substrate? ☑

    Support data intensive science? ☑

    Virtualization? ☑ In progress

    Integrate network with compute and storage? ☐ Next step

ENABLING INNOVATION THROUGH NETWORK VIRTUALIZATION
(AND INTEGRATION OF COMPUTE AND STORAGE)

Eric Boyd, Internet2

Luke Fowler, Indiana University

Matt Zekauskas, Internet2

Thank you. For more information,
visit http://www.internet2.edu
or e-mail innovation@internet2.edu