GENI Stitching: Under the Hood Or,
How did Heidi & Luisa do that?

Aaron Helsinger
July 22, 2013
www.geni.net

Why is GENI Stitching great?

- A lot happened under the covers in that demo.
- A lot of work goes in to making stitching work

- But experimenters don't have to worry about it!

GENI Stitching is

- Fast

- Instructions are posted so you can do it

- Connect InstaGENI to ExoGENI

- Connect a non GENI host to a GENI VM
  – Across CENIC to a machine at Stanford

Manual Circuits are Hard Work

- ## Custom circuits in GENI are not new
  - ('Slice around the World')
- ## Setting them up meant a lot of engineering work
  - Set up tunnels
  - Find free VLANs
  - Configure Switches
  - Coordinate with VMs
  - Debug circuits
- ## This can take weeks and much grief
- ## And then the circuit often becomes 'stale' and broken

- The demo you saw did it live in minutes
  - Instructions to do it yourself are on the wiki

- Create or find an RSpec that says
  - I want a VM at Utah with 1 interface
  - I want a 2nd VM at the GPO with 1 interface
  - **NEW**: I want a link between those interfaces connecting those aggregates.

- Your tool and the aggregates do the rest.

```
<rspec>
 <node …>
 <node …>
 <link client_id="mylink">
  <component_manager…
  <component_manager…
  <interface_ref …>
  <interface_ref …>
```

How does GENI Stitching Work?

1. Rack Configuration: Long. Done once in advance.

2. Tool & Aggregates make Reservations: Quick, Live, Easy

Creating a Stitchable Network

- Operator / Engineer Setup and Configuration
- Hard work that can take weeks
- Done *ONCE* when a new rack is configured
- Replaces most of the manual circuit work by doing it for multiple VLANs in advance

- Thank you to all the engineers who are working to make this possible!

Rack Stitching Configuration

- Identify a path or paths from the rack to other GENI aggregates.
  - Typically a connection to a national backbone
- Identify the network providers
  - Typically a campus, a regional, and the backbone
- Identify the endpoints and VLAN tags that can be used to connect to the rack

GENI Aggregates — Backbone — Regional — Campus — Rack
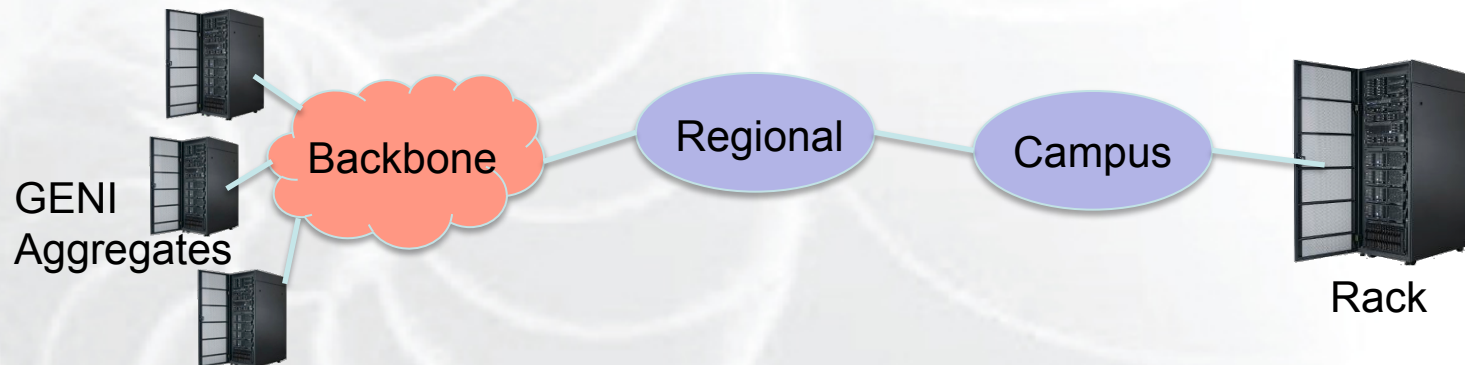
3 Options for controlling circuits at networks:

- Run an Aggregate over Dynamic Circuits
  - We have an OSCARS aggregate. Others possible.

- Run an Aggregate for Static Circuits
  - Network manages the VLANs

- Delegate VLANs to the peer GENI aggregates to manage
  - Set up the VLANs, configure them at the aggregates

- Record this assignment by each intermediate network for easy searching and as the basis for configuring systems
- This work has been prototyped by Tom Lehman and Xi Yang of MAX:
  - A config file captures the information
  - A script generates a wiki page
  - This can be improved and expanded over time.

LINK#1					
InterfaceA				**InterfaceZ**	
Description	**ComponentId**		**AggregateManagerURL**	**Description**	**ComponentId**
BBN GPO-IG to Internet2 ION via NOX	urn:publicid:IDN+instageni.gpolab.bbn.com+interface+procurve2:5.24		https://www.instageni.gpolab.bbn.com:12369 /protogeni/xmlrpc/am/2.0	Internet2 ION to BBN GPO-IG via NOX	urn:publicid:IDN+ion.internet2.edu+interface+rtr.newy:ae0:bbn
Intermediate Networks					
urn:publicid:IDN+nox.org					
Capacity	**MaximumReservableCapacity**		**MinimumReservableCapacity**	**Granularity**	**AvailableVlanRange**
1000Mbps	1000Mbps		1Mbps	n/a	3747-3749

- Configure the VLAN ranges
  - Networks that provide **dynamic** circuit services (OSCARS) can run a GENI Aggregate Manager
    - Internet2 runs such an aggregate over ION.
    - Such networks configure their aggregate manager and their OSCARS interfaces.
  - Networks that provide **static** circuits will provision all of the promised VLANs between fixed switch/port endpoints
    - Relying on the GENI aggregates at the endpoints to control use of the VLANs.

GENI
AMs

ION

NOX

BBN

IG GPO

- ION is an Internet2 service that uses OSCARS to do dynamic circuits and does VLAN translation

- Tom Lehman and Xi Yang of MAX wrote an SFA based aggregate manager that translates GENI calls to OSCARS calls

- Internet2 operates this aggregate
  - *Other OSCARS networks (regionals?) could use it too*

- This is powerful:
  - GENI experimenters can connect arbitrary GENI ION endpoints when they need them.
  - This enables stitching to non GENI resources (i.e. through CENIC to a node in Stanford as in the demo).

Thank you MAX and Internet2!

1. Rack Configuration: Long. Done once in advance.

2. Tool & Aggregates make Reservations: Quick, Live, Easy

Experimenter View: Creating a Circuit

- GENI Stitching means experimenters don't have to worry about all that work. It's already done for you.

- The demo showed what an experimenter sees and does.
  - Quick, Simple, and Documented

- What happens under the covers?
  - Remember: Experimenter just says they want a link; tools and aggregates do the rest. It is easy and getting easier.
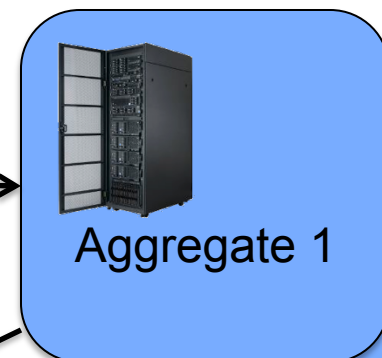
Experimenter View: Creating a Circuit
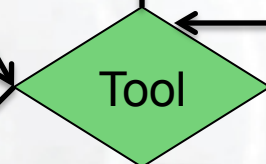
2. Send Path Request to Stitching Computation Service (SCS)
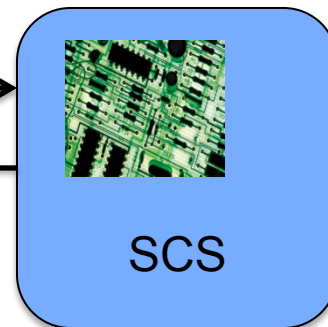
1. Simple Request

```
<link client_id="mylink">
 <component_manager…
 <component_manager…
```

SCS

3. Get Expanded Request

Tool

4. Send Request to Aggregate 1

Aggregate 1

5. Get Manifest

Aggregate 2

```
<link client_id="mylink">
 <component_manager…
 <component_manager…
…
<stitching>
 <path id="mylink">
  <hop id="1">
   <link id="switch1:port1"
…
    <vlan…>3747</vlan…>
```
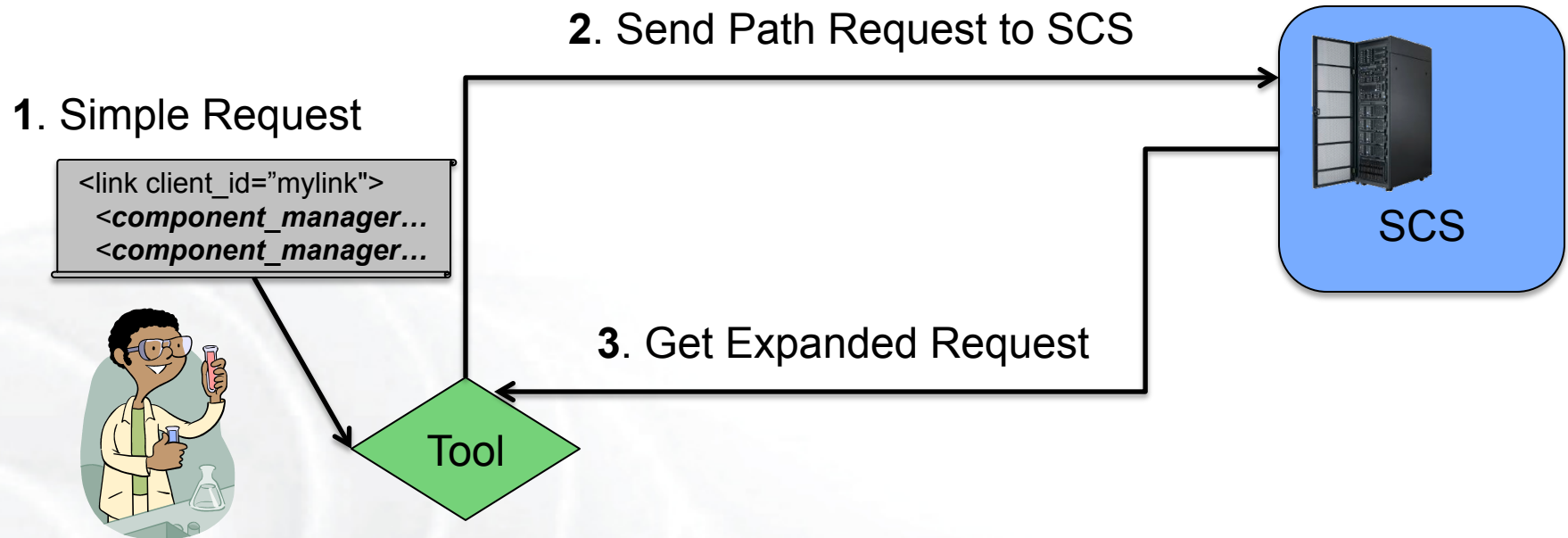
6. Repeat for Other Aggregates

7. Manifest Back

- Design your Experiment to include
  - 2+ nodes in different places (could be IG, EG, even a non-aggregate node as in the Stanford/CENIC demo)
  - A link between them
- Tool helps you create an RSpec with a **new link**:
  - From MyNode: Interface to MyOtherNode: Interface2
  - Connecting ThisFirstAggregate to ThatOtherAggregate
- One RSpec for whole slice

- Done.

The rest is done for you.

```
<node …>
<node …>
<link client_id="mylink">
 <component_manager…
 <component_manager…
 <interface_ref …>
 <interface_ref …>
```

- ## No central GENI stitching authority
  - Stitching is just a set of resource reservations at multiple aggregates.

- ## Any tool can do this.
  - The demo used an Omni script called 'stitcher.py'.
  1. Expand your request to find a path for your circuit
  2. Generate a request RSpec for each aggregate and make the reservations
  3. Check if any dynamic circuits were successfully created
  4. Report back a combined summary of what you have at all the aggregates

Experimenter View: Find a Path

2. Send Path Request to SCS

1. Simple Request

```
<link client_id="mylink">
  <component_manager…
  <component_manager…
```

SCS

3. Get Expanded Request

Tool

- Connecting 2 aggregates means finding a path
  - May include intermediate aggregates, like ION
  - Must check available VLANs
- Find connections in Aggregate advertisements

- Tool computes possible paths and VLANs
- Order of reservations impacts chance of success
  - Some aggregates must pick the VLAN
  - VLAN translating aggregates can go last
- Your tool does all this for you

Stitching Computation Service
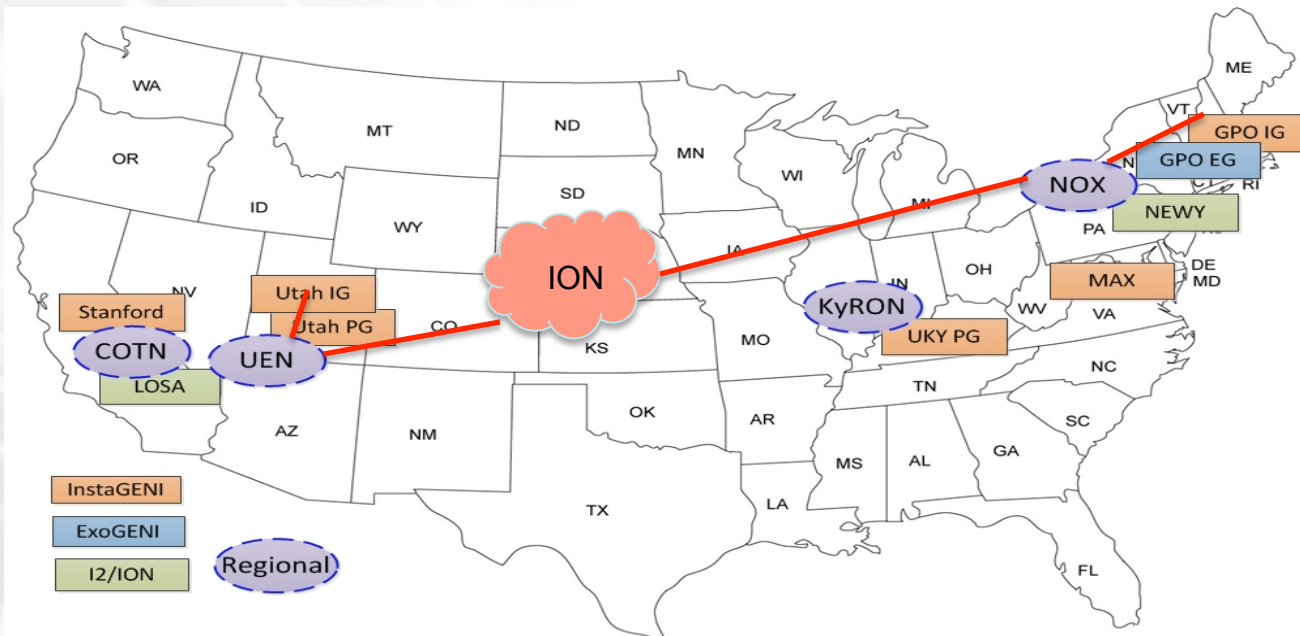
- That is hard, lots of heuristics
- Stitching Computation Service (SCS) for path and workflow computation
  - Includes many heuristics to optimize path, chance of success
  - Allows excluding particular connection points, VLANs
- Tom Lehman and Xi Yang wrote this optional service
  - Other tools may use different heuristics
- Stitcher uses the SCS

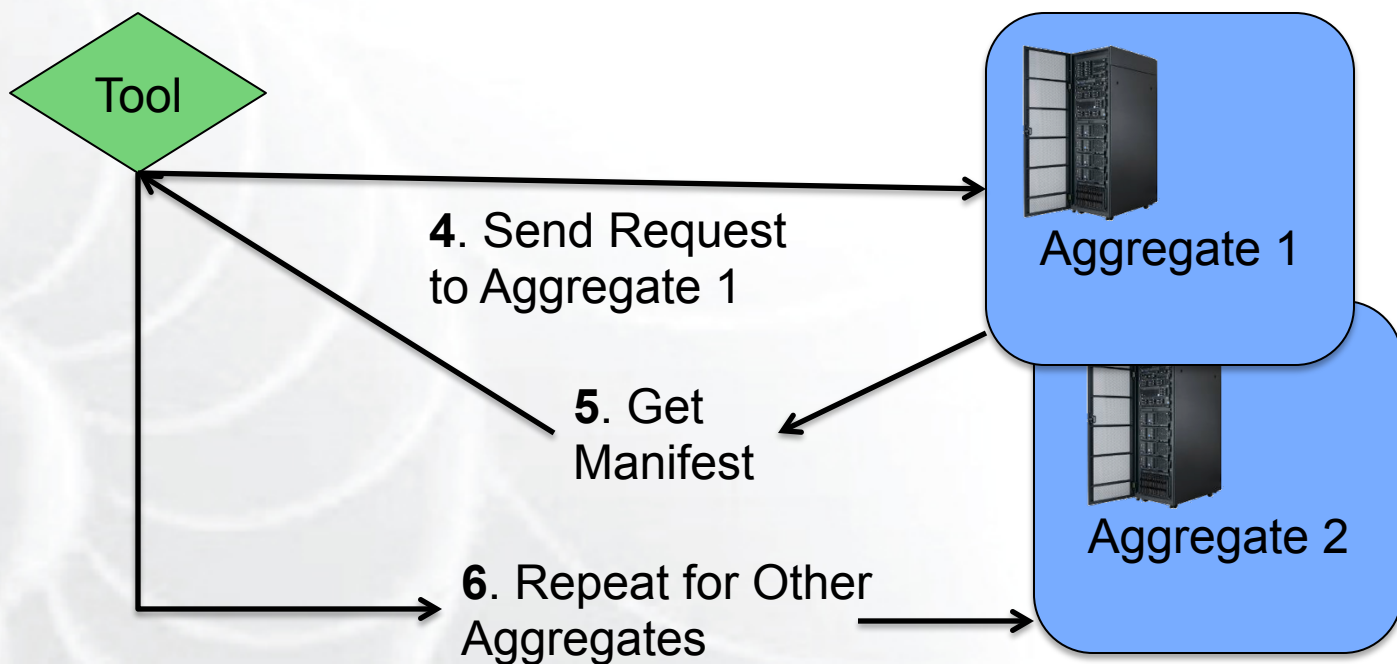 http://geni.maxgigapop.net/twiki/bin/view/GENI/NetworkStitchingAPI

- Stitching RSpec Extension describes connections, paths, and requested or allocated VLANs

- This is how tools and aggregates talk about stitched circuits

http://www.geni.net/resources/rspec/ext/stitch/

```
<stitching>
 <path id="mylink">
  <hop id="1">
   <link id=switch1:port1">
     ….
      <vlan…>3747</vlan…>
  <hop id="2">….
```

- ## In Aggregate advertisements
  - Local switch ports and VLAN ranges, and the remote switch port they connect to

- ## The SCS adds to your request a stitching extension
  - For each circuit, a series of switch ports / VLAN tags (hops) requested

- ## Manifests:
 - Your request with the VLAN tags assigned for each hop

```
<stitching>
 <path id="mylink">
  <hop id="1">
   <link id=switch1:port1">
    ….
     <vlan…>3747</vlan…>
  <hop id="2">….
```

Tool

4. Send Request to Aggregate 1

5. Get Manifest

6. Repeat for Other Aggregates

Aggregate 1

Aggregate 2

- Circuits and VLANs are resources

- Reserve through GENI AM API 'CreateSliver' calls

- Order of reservations matters
  - Use the computed workflow

- Stitcher used Omni to make the AM API calls

Tool

Aggregate 1

Aggregate 2
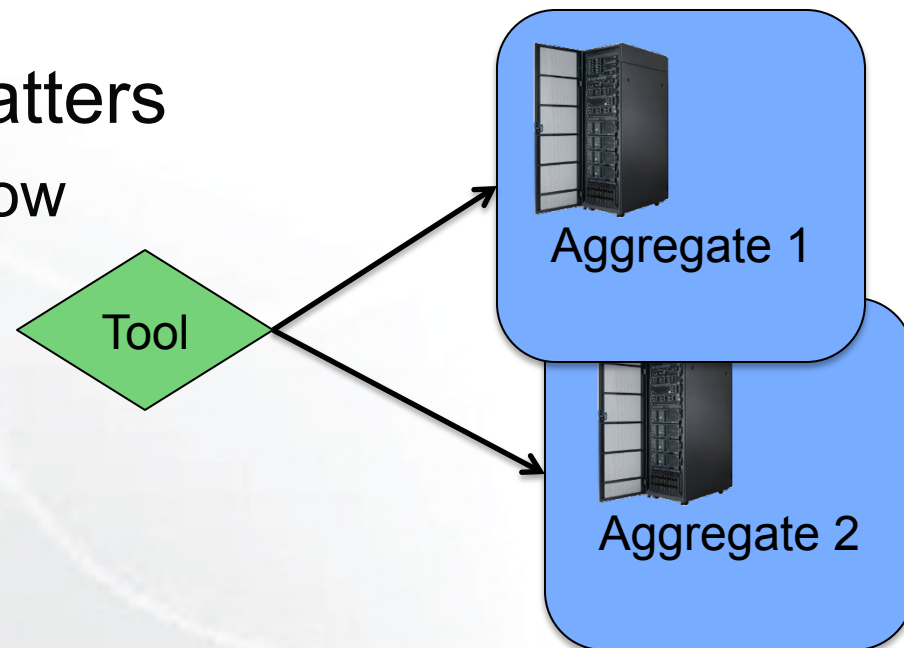
- **Read Request and Stitching Extension**
  - Is (a) VLAN available?
  - Is node available?

- **Handle Request**
  - Configure the node
  - Configure the switch to connect the VLAN to the node

- **Book-keeping**
  - Node and VLAN are taken

- **Return manifest**
  - VLAN is reported in stitching extension

Request

Manifest

Set VLAN 3747

Handle Aggregate Response
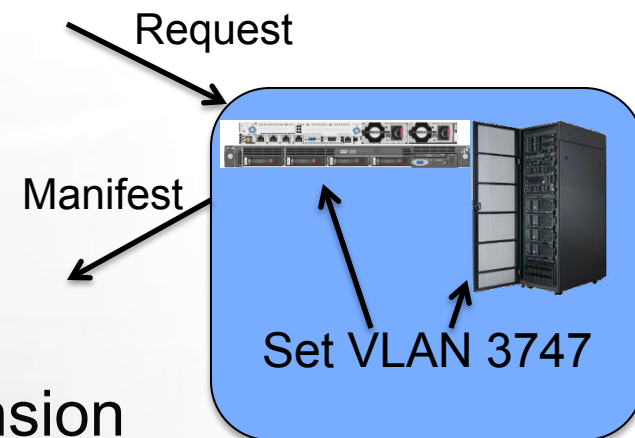
- Tool reads VLAN out of manifest

- VLAN is inserted into request at next aggregate
 - GPO IG picks tag 3747, so request to ION uses 3747

```
<link client_id="mylink">
  <component_manager…
  <component_manager…
…
<stitching>
  <path id="mylink">
    <hop id="switch1:port1">
    <vlan>3747</vlan>
```

- Tool handles many problems
  - VLAN is in use? Tool should try another
  - No VLANs available here? Try a different path
  - Something else, like no node available? Tell the user

- Stitcher checks for and handles these things.

- Dynamic Circuits take Time
  - OSCARS is configuring all your routers
  - This is real work

- While circuits sometimes fail…
- Tool can check and retry if it fails

  - Wait while the tool does the work for you

Tool
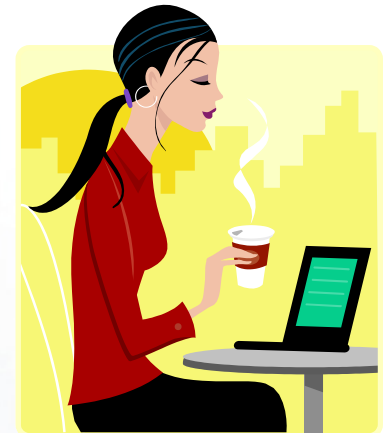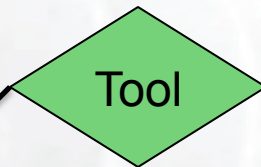
```
<link client_id="mylink">
 <component_manager…
 <component_manager…
…
<stitching>
 <path id="mylink">
  <hop id="1">
   <link id="switch1:port1"
…
    <vlan…>3747</vlan…>
```

7. Manifest Back

- You have your circuit connecting your nodes!

- What do you have?
  - Details aren't usually interesting
    - Which VLAN tag, etc

```
<link client_id="mylink">
  <component_manager…
  <component_manager…
…
<stitching>
  <path id="mylink">
    <hop id="switch1:port1">
      <vlan>3747</vlan>
```

- Stitcher combines the manifests from all your aggregates
  - You may have reservations places you didn't expect, like ION

- This might seem like a lot. But you saw it live in minutes.

- ## Testing Continues
  - Detailed Test Plan
  - Early network and aggregate partners helping
- ## More Sites
  - After testing, all GENI racks
- ## More Providers
  - NLR, regionals, and campuses
- ## More VLANs
  - VLAN ranges are small now – a scarce resource
  - GPO & rack teams will work to expand these
- ## Monitoring
  - VLAN use and circuit availability

- Omni 2.4 (improved stitcher) is coming soon
- Flack stitching support is in testing

- Portal? Other tools?
- Other kinds of circuits
  - VLANs that support OpenFlow controller routing
  - Multipoint?

GENI Stitching is Easy and getting Easier

- GENI Stitching provides experimenters the ability to create
  - Custom circuits
  - On-demand
  - Across all of GENI.
- While this basic capability has existed for years, this is now more broadly available or simpler to use.
- GENI makes custom topologies easy
  - It is getting bigger and easier.