



Strawman

GENI Use Cases

Global Environment for Network Innovations

The GENI Project Office (GPO)

www.geni.net

Clearing house for all GENI news and documents



Overview

- Introduction
 - GENI Experiment A
 - Assumptions in this Use Case
- Strawman Use Case #1
 - Resource Discovery
 - Sliver Creation
 - Experiment Set-Up
- Mini Use Case #2: Emergency Slice Suspension
- Mini Use Case #3: User Opt-In
- Suggestions for future use-cases
- Experiment Details (backup)

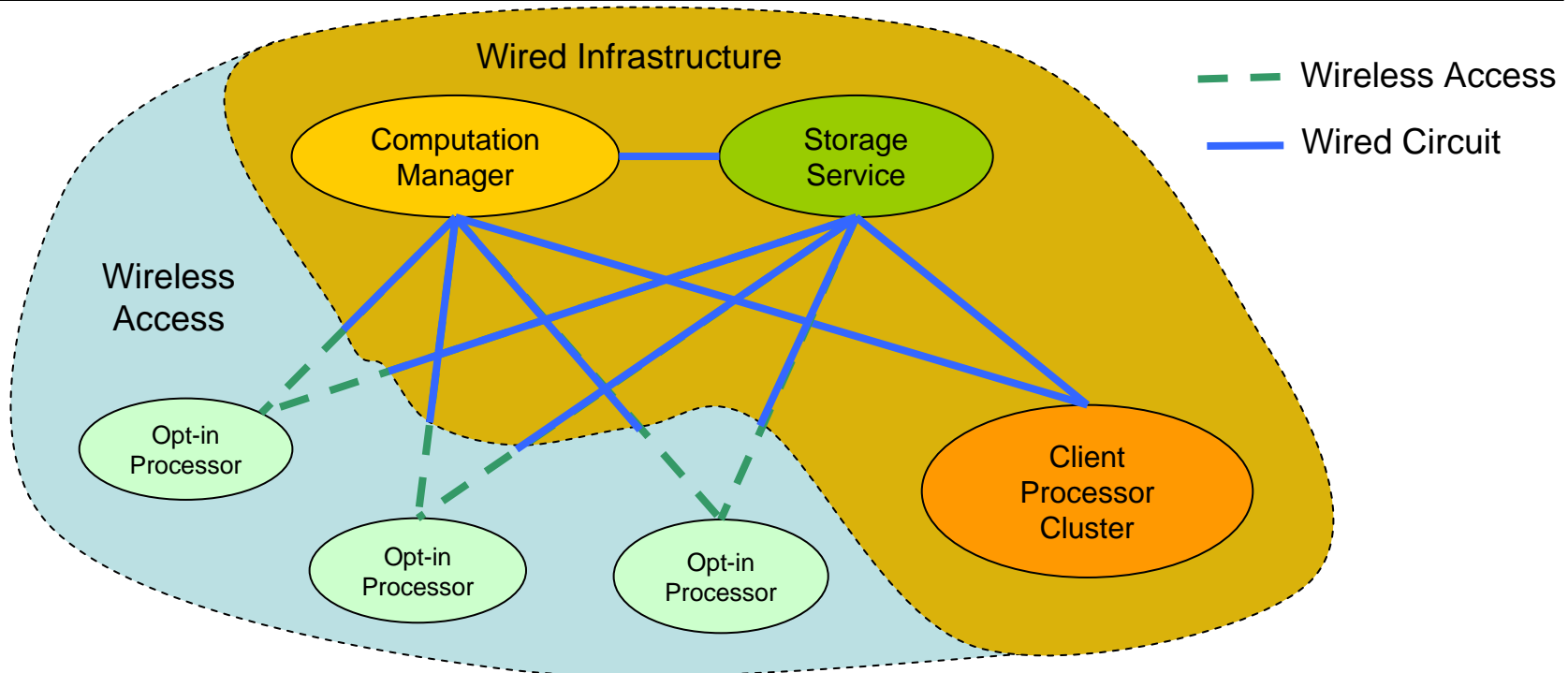


Introduction

- This is a strawman use case; it is intended to be:
 - One possible walk-through of the basic steps needed to establish an experiment
 - A tool for discussing the functions and interfaces GENI should support
 - A template for exploring different experiments and facility designs
 - [Mostly] consistent with the GENI architecture
 - Actions described in GENI architecture documents are shown in italics.
- What this use case is not:
 - Complete
 - Optimal
 - Definitive
 - Final



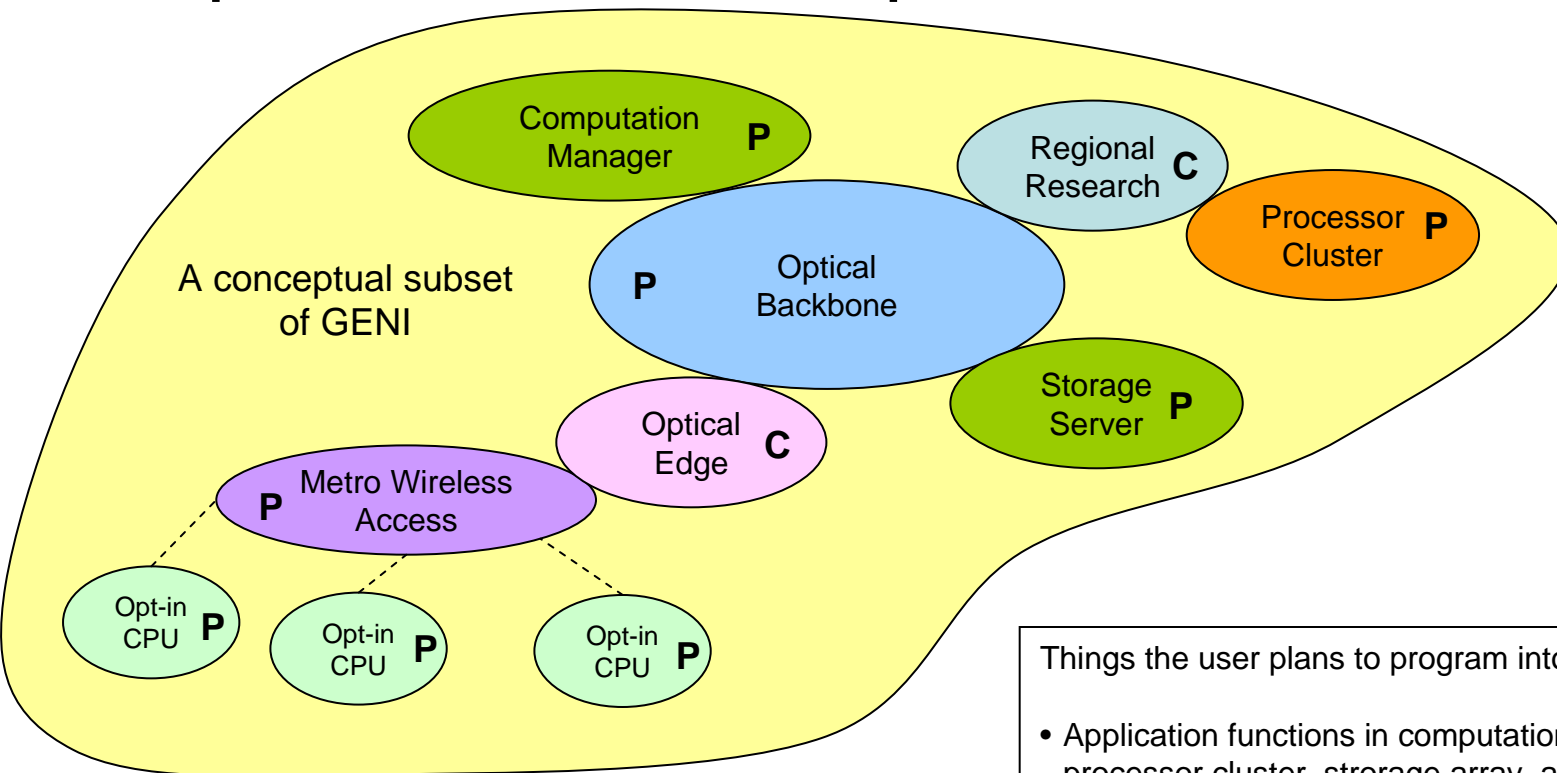
Experiment A Description



A distributed computing environment managed by a computation manager coordinating processes between a geographically disperse set of client processors and storage service

- Involves end-to-end connections spanning multiple heterogeneous network technologies, management and ownership domains
- GbE coordination channel from computation manager to storage service, client processor cluster and wireless access point (not shown)
- 10GbE optical point-multipoint data-file transfer from storage service to computation manager, client processor cluster and wireless access point

Experiment A's Components



P – GENI resources programmed and configured

C – GENI resources configured only

Things the user plans to program into his slice:

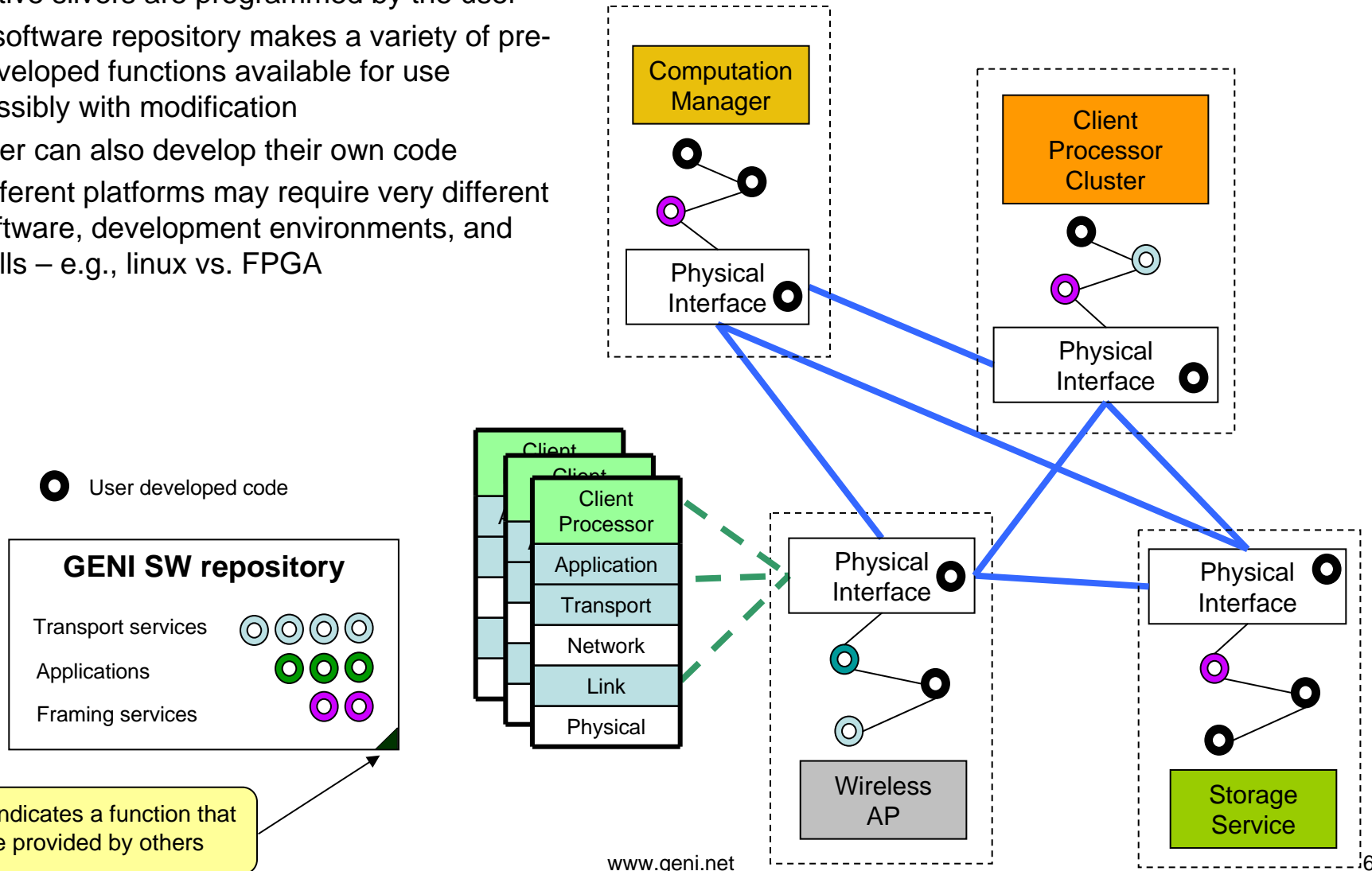
- Application functions in computation manager, processor cluster, storage array, and opt-in CPU's
- Transport functions in computation manager, storage server, and opt-in CPU's
- Link functions in wireless access point and opt-in CPU's
- Phy functions in optical backbone

- The GENI substrate is formed from a diverse set of participating networks
 - Ovals denote a network domain and management/ownership domain (an aggregate)
- User Opt-In enables individual CPU's to be utilized as additional component resources



Active Slivers Draw on a Software Repository

- Active slivers are programmed by the user
- A software repository makes a variety of pre-developed functions available for use possibly with modification
- User can also develop their own code
- Different platforms may require very different software, development environments, and skills – e.g., linux vs. FPGA





Use Case Assumptions

- The researcher and all components...
 - Have established their identities
 - Have registered NSF GENI Clearinghouse
- The researcher has acquired authorization credentials permitting slice creation
- Slice creation has been completed prior to this use case.
 - I.e., a slice has been created for an organization in which the user belongs
- This experiment takes place under the auspice of a single Clearinghouse
 - Uses multiple Management Authorities
 - Use of multiple Clearinghouses is not addressed (but is not excluded)



Use Case Assumptions

- Control & management plane communication in this use case (e.g., for discovery, reservation, software upload, initial debug, and O&M) take place over an IP network.
- No assumptions are made regarding whether reservations are fixed or best-effort; different components could choose independently
- Assumes a sufficiently sophisticated researcher who can work directly with RSpecs and can bring up a distributed experiment on his own. An alternative would be to use helper services (which might be hosted by the Clearinghouse) that could handle sliver-coordination and/or translation from RSpecs to a more user-friendly syntax.
 - Also assuming the researcher handles the operations of his own experiment.
- Pre- and post-staging phases are not addressed but would be useful to allow, e.g., reduced capacity reservations for pre-experiment debugging and trials and post-experiment forensics.



Strawman GENI Use Case #1

From Resource Discovery To
Experiment Setup

Resource Discovery (1 of 2)

INPUTS:

I need the following **edge nodes**:

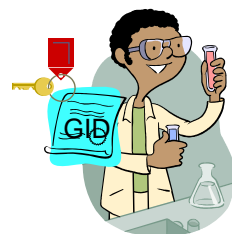
- Two CPU clusters
- One Storage array
- Three access points on a wireless network

I need the following **connections**

- A dedicated GbE service between the two CPU clusters
- A dedicated GbE service between a CPU cluster and the wireless network
- A 10GbE dedicated optical multicast circuit from the storage array to the CPU clusters and wireless network
 - ULH Optical circuit with user defined Ethernet frame encapsulation including strong FEC
- All connections must be over a national back-bone (long-haul scale)

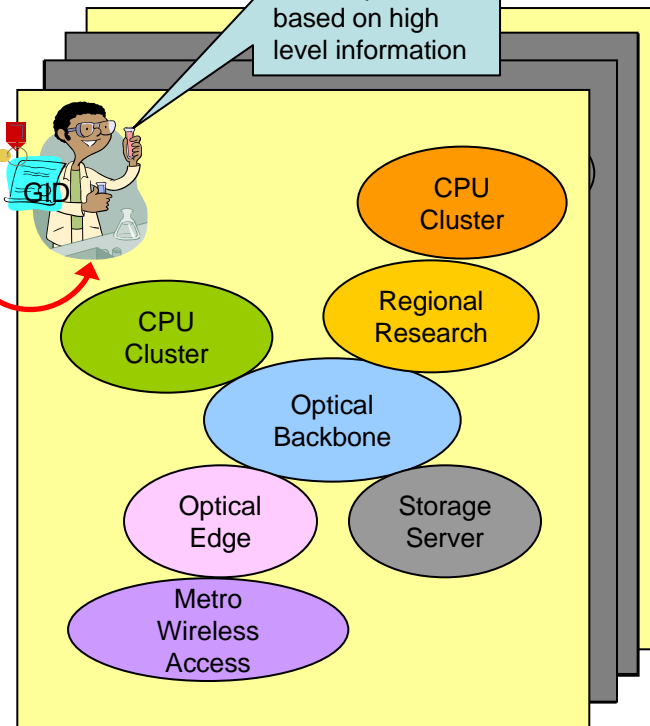
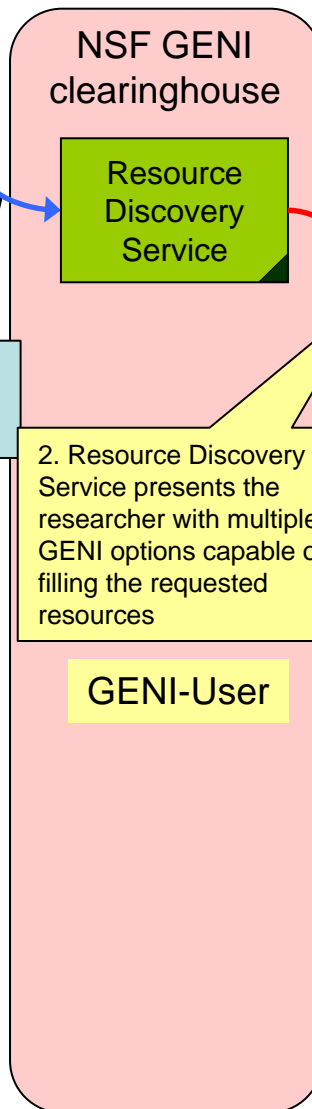
I need the following **measurements**

- CPU utilization
- RF spectrum analysis
- Optical backbone BER (pre-FEC and post-FEC) and OSNR
- Packet Error Ratio



1. Researcher queries Discovery Service for candidate resources

User-GENI



OUTPUTS: Candidate solution descriptions

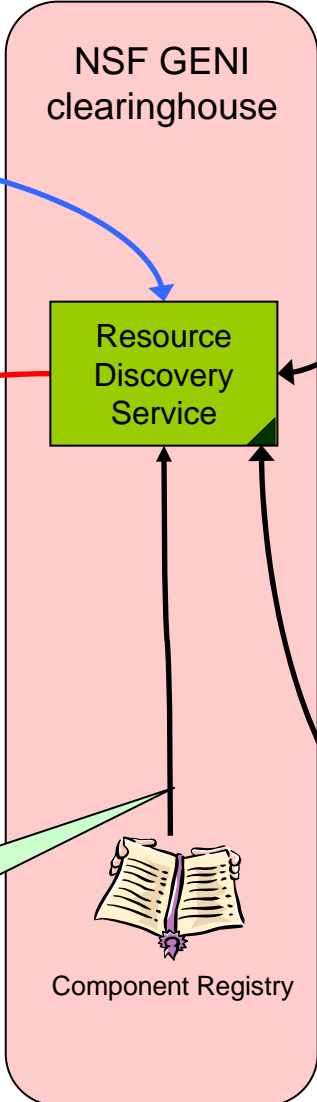
- A high level view of GENI resources potentially satisfying researcher needs
- A list of networks
- Components (nodes, connections, services)
 - A CPU cluster located at XYZ
 - A GbE connection between A:B
 - An optical multicast topology in the optical backbone network
 - GENI Policies

Resource Discovery (2 of 2)

I need detailed specifications to make my final choice of candidate solutions



1. Researcher queries Discovery Service for detailed capability & availability.

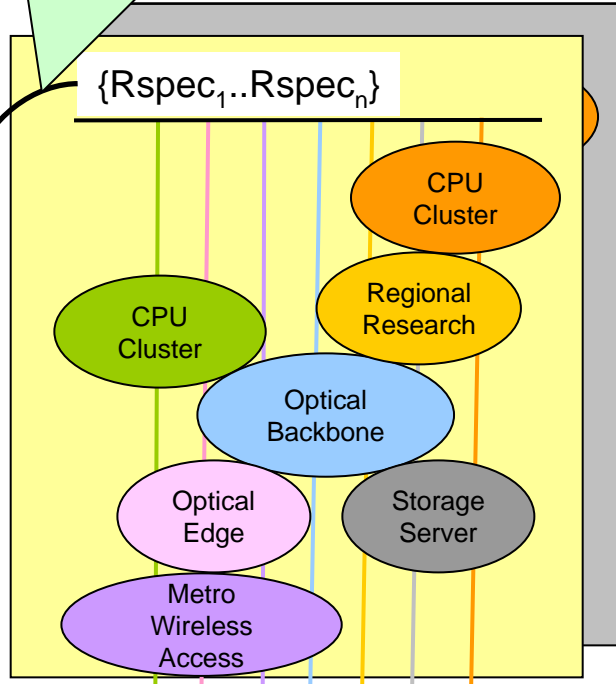


2. Resource services provide user with lists of capabilities and their availability. Registry provides URI's (one per aggregate) for reservation of resources within selected option.

0a. Discovery service(s) subscribe to notification of new components from component registry.

3. Researcher selects desired GENI resource option (i.e, he chooses a specific set of R-specs) based on specific capabilities and availability
- CPU count
 - Storage capacity
 - Optical reach
 - Wireless access location
 - Scheduled downtime
 - Dates of interest

0b. Discovery service(s) subscribe to capability (RSpec) and schedule advertisements from Component Managers.

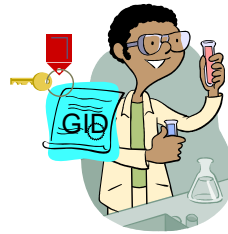


$\{Coarse-grain scheduling_{1..n}\}$

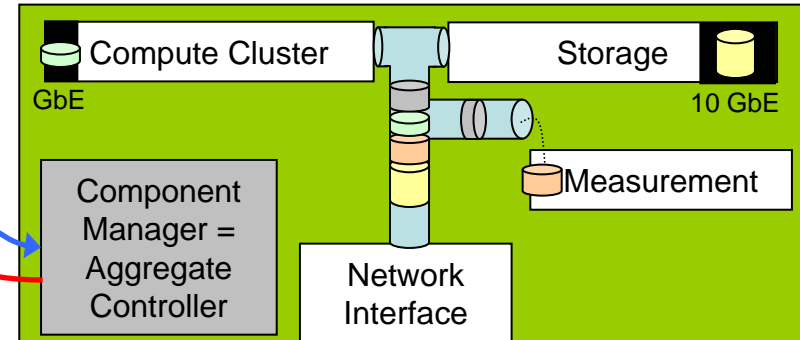
Sliver Creation: Computation Resources (1 of 3)

I want to reserve all the pieces to build my experiment, I will start with a CPU cluster, reserving the following

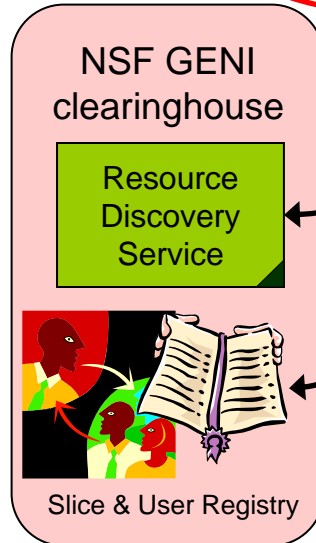
- 300 Processors
- 10 TB Local storage (multiple disks)
- 1 GbE and 10 GbE link capacity from network interface to CPU's and 10TB drive
- CPU utilization measurement w/ data transfer rates



1. Researcher submits credentials and request to component manager (CM) for resources, that includes a reservation time



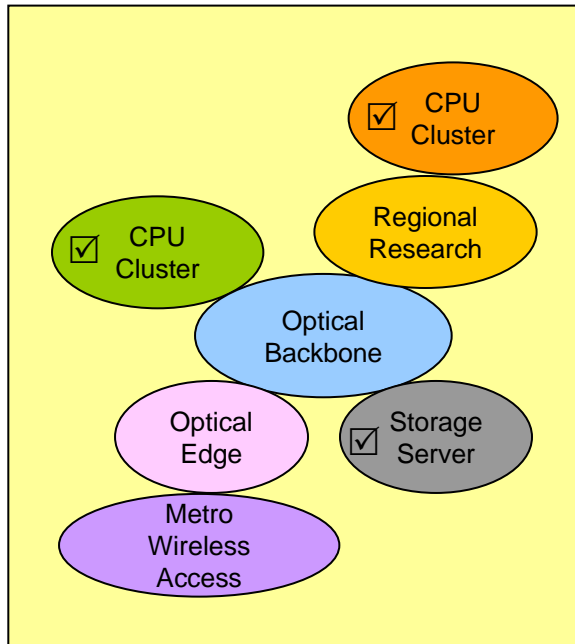
2. CM checks policies against credentials and accepts reservation by returning signed RSpec to User (called a ticket).



3a. CM sends schedule update with reservation information (resources and dates)

3b. CM sends copy of ticket to Slice Registry (who tracks resources in each slice).

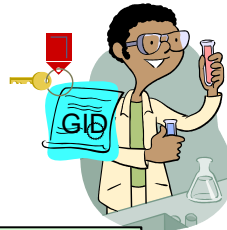
A similar set of actions are performed for the other CPU cluster and storage array



At this stage, the researcher has a right to use specific resources (i.e., establish slivers) from several of the aggregates. However, these resources are not active and have not been composed into a coherent experiment.

Sliver Creation: Aggregate Interconnection (2 of 3)

Now that my edge processor resources are reserved, I need to establish data link connections across the multiple network domains. But, how do I set-up connections between two aggregates?

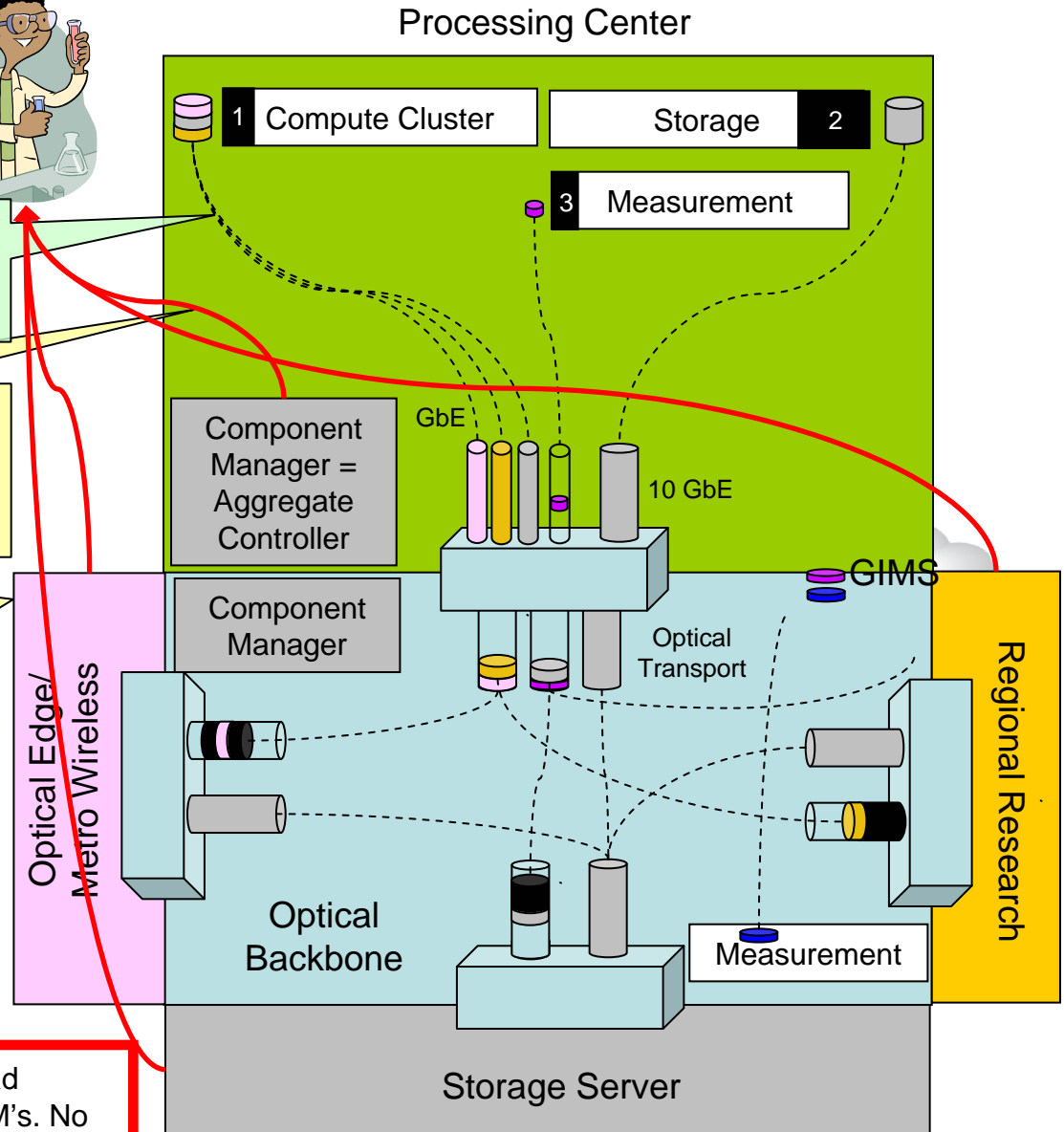


The reservation specifies network bandwidth between the 3 components and Optical Backbone. The Processing Center CM handles establishing requested internal network configuration. Connectivity to OB already exists.

When the CM issues a ticket, the signed Rspec will contain **interconnection parameters** of each network reservation, e.g., port number, VLAN ID, source Ethernet address, so the OB CM can associate network resources in PC to those in OB.

This is repeated for the Regional Research, Storage Server, Optical Edge (and Metro Wireless) networks.

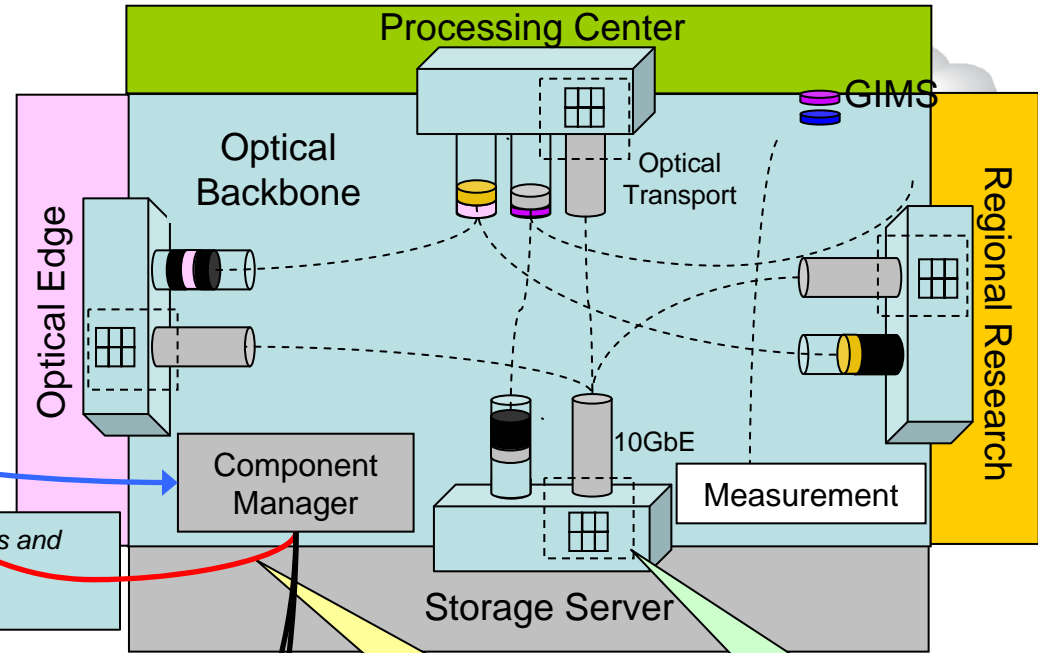
At this stage, the researcher has obtained payload mapping information to each of the aggregate CM's. No connections have been established



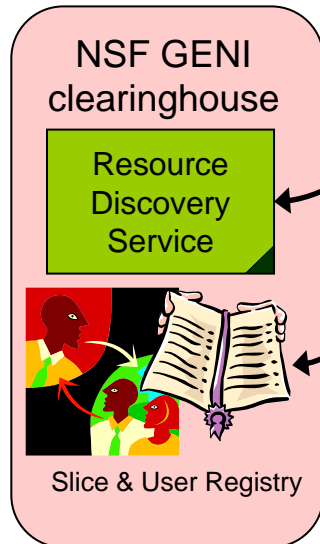
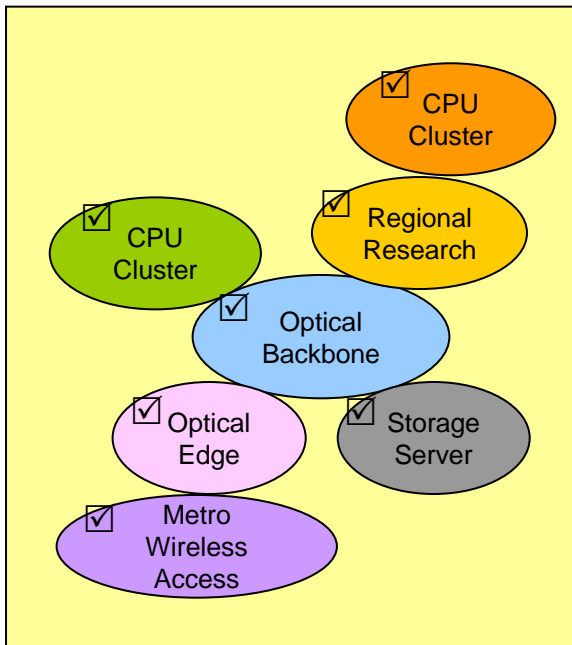
Sliver Creation: Networking Resources (3 of 3)

Now that aggregate interconnection mapping is established, I want to reserve the following resources:

- Optical multipoint topology from storage server to PC,OE, RR network interfaces
- FPGA framers on linecards mapped to 10GbE payload at each aggregate interface
- 1 GbE link capacity between Processing center and other aggregate interfaces
- BER and OSNR measurements on all links
- Measurement data transfer rates



1. Researcher submits credentials, aggregate interconnection maps and request to component manager (CM) for resources, that includes a reservation time



2. CM checks policies against credentials and accepts reservation by returning signed RSpec to User

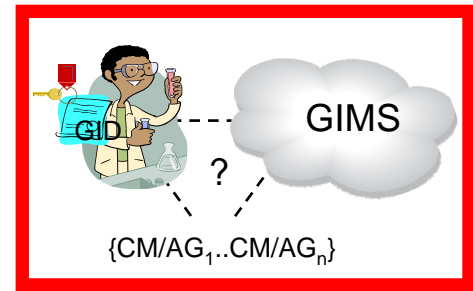
3a. CM maps payloads to interfaces and provides information in ticket returned to user

3b. CM sends status update with reservation information

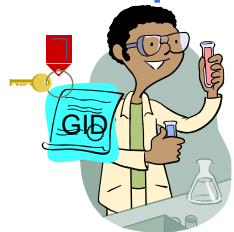
3c. CM sends copy of ticket to Slice Registry

At this stage, a complete slice reservation exists. However, until that reservation is exercised (i.e., tickets redeemed), active slivers cannot be programmed.

Experiment Set-Up (1 of 2)



I have tickets for my resources and all aggregate interconnection maps. It is now time to set-up my experiment



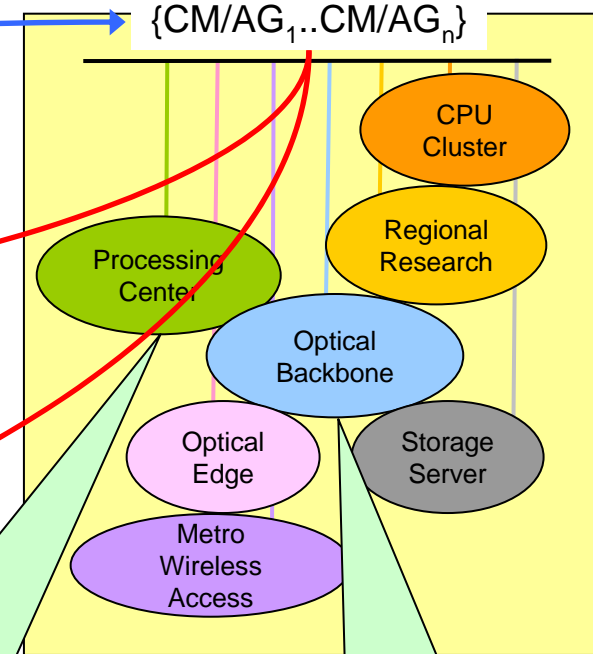
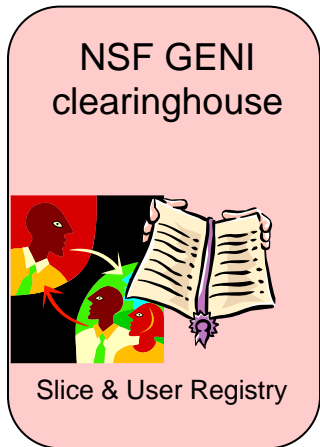
1. Researcher submits tickets to the CM/AG's

3. CM/AG's issue identification of specific resources allocated to slice & coordinates for uploading code & configurations.

4. Copy of allocated resource identification sent to Slice registry and can be used when identifying misbehaving slices

2a. Processing center CM/AG's execute reservations:
300 CPU's interconnected forming an isolated cluster; Multiple disks connected to provide an 10TB composite capacity; Payloads established over links from cluster and 10TB storage to designated network interface; Measurement links to GIMS established

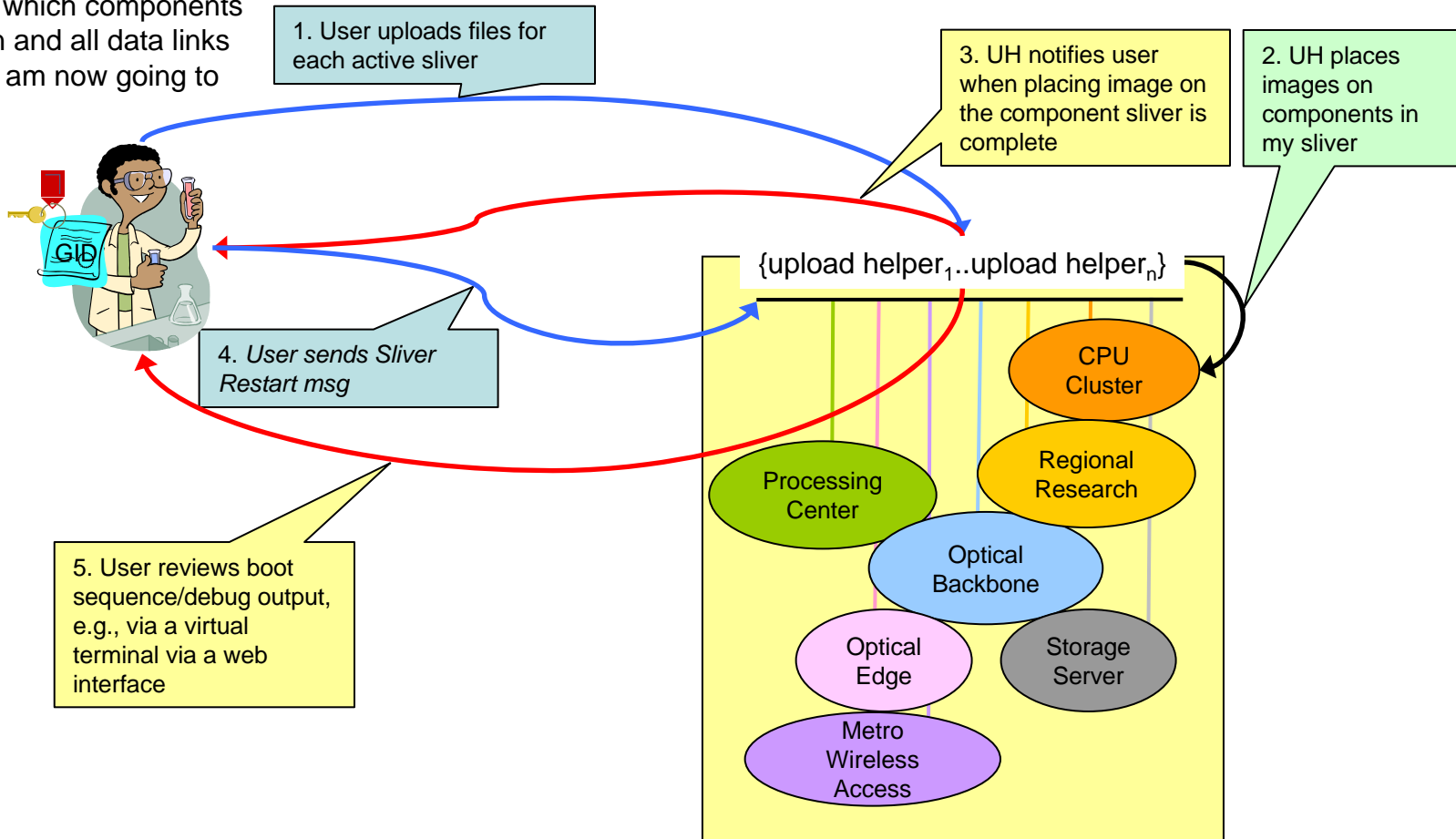
2b. Optical Backbone CM/AG's execute reservations:
Optical multipoint topology established; Payloads mapped and links established between aggregate interfaces (PC,RR,OE,SS); Measurement links to GIMS established



Experiment Set-Up (2 of 2)

All my reservations have been accepted, I know which components I have been given and all data links are established. I am now going to load my software

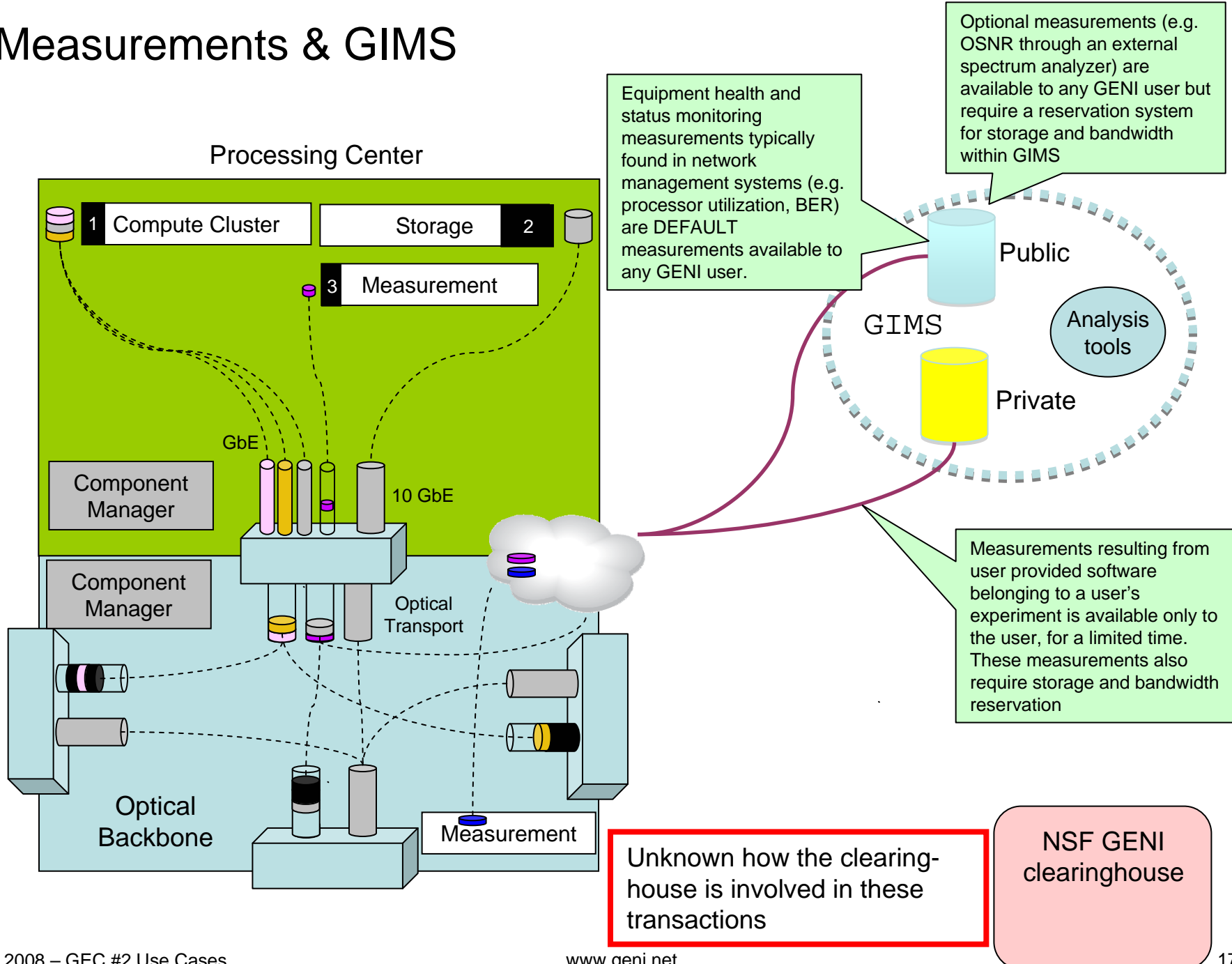
The earlier step of ticket redemption provides user with URLs for file upload, sliver control, debug output



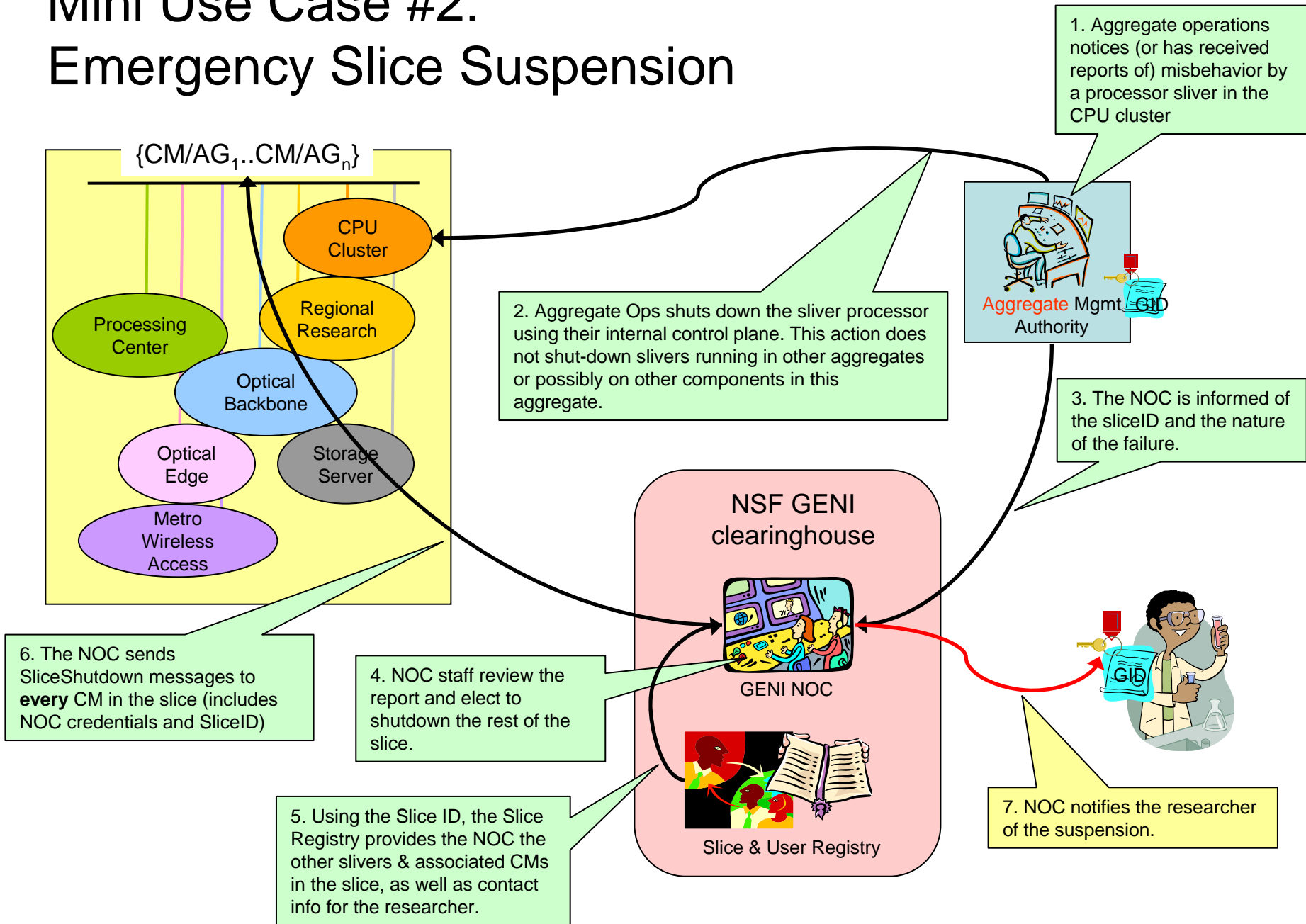
NSF GENI clearinghouse

The clearinghouse is not involved in these transactions

Measurements & GIMS



Mini Use Case #2: Emergency Slice Suspension

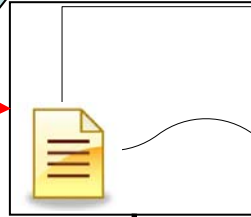


Mini Use Case #3: User Opt-In

I want to allow others to join my experiment



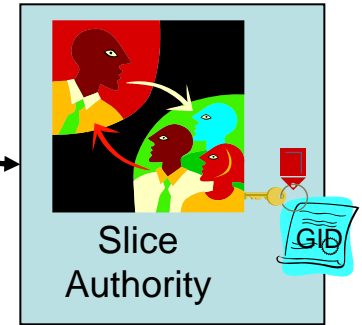
1. Experimenter posts web page with information about experiment



NSF GENI clearinghouse

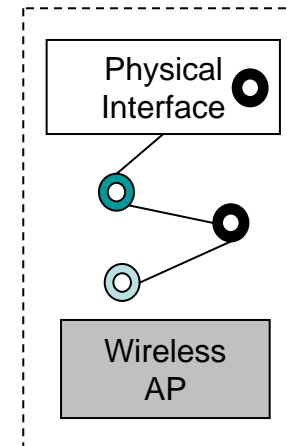
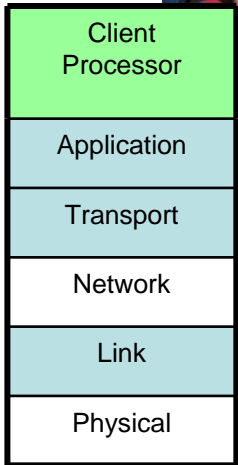
5. Slice authority registers identity (if needed) and coordinates of opt-in user (for shutdown, legal indemnity).

4. User is directed to opt-in page which verifies user agrees to (standardized) terms and conditions of experiment. If approved, user is permitted to join experiment.



2. Opt-in user downloads and installs experimental code

3. System starts-up, searches for and discovers participating access point





Need more use cases on...

- **Other Approaches**
 - “Greatest use of primitives” vs. “bundled services”
- **Discovery**
 - Multiple 3rd party resource discovery services
 - Coordinating reservations across multiple scheduled components
- **Slice Evolution**
 - Slices that grow over time
 - Best-effort pre- and post-staging phase
- **Authorization**
 - What role does the Slice Authority play?
 - How does researcher acquire credentials for slice creation?
 - Delegation of permission to reserve resources (does authority lie in slices or users?)
 - Resource brokerages, sub-dividing & reselling tickets, bargaining (swapping) resources
- **Failsafe**
 - Big Red Switch (e.g., between GENI and Internet)
 - Keep-alives (challenging for disconnected usage)
- **Operations**
 - Handling outages



Seeking Feedback

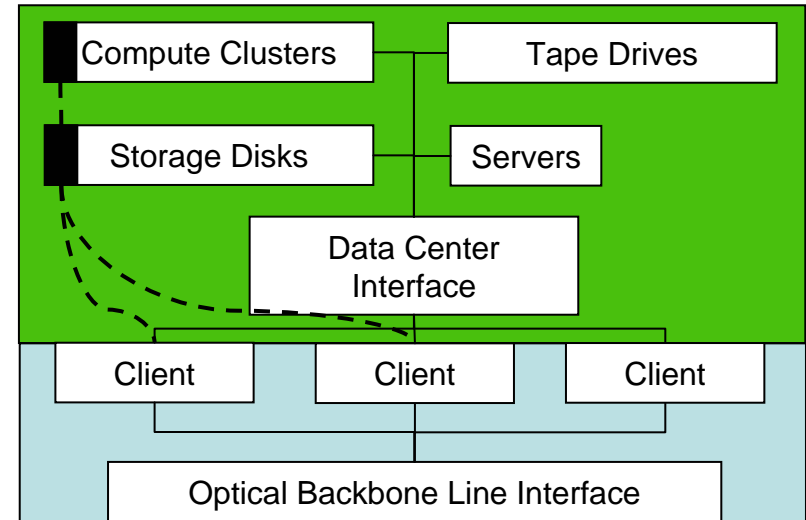
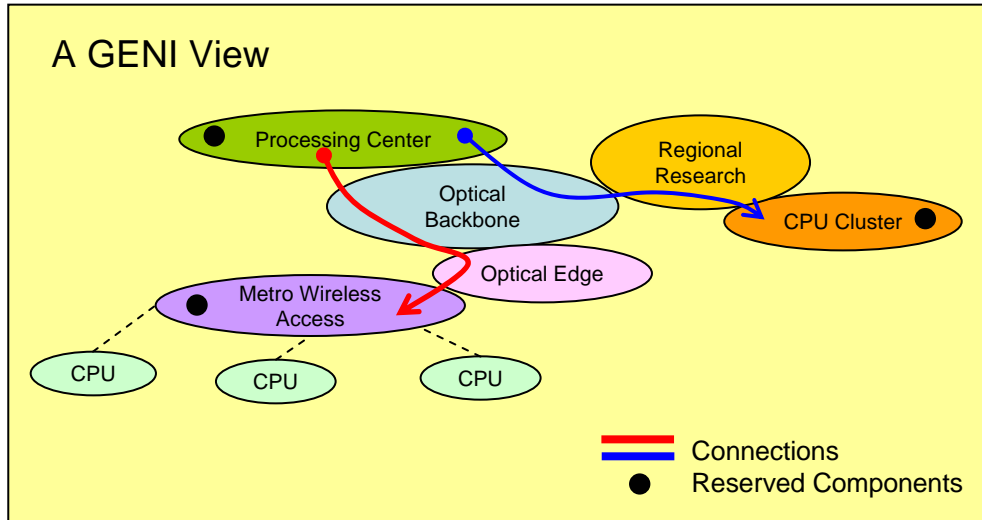
- Where are these use cases incorrect?
- What are the alternatives?
- What are the implications for <wg-group>?



Some Experiment Details

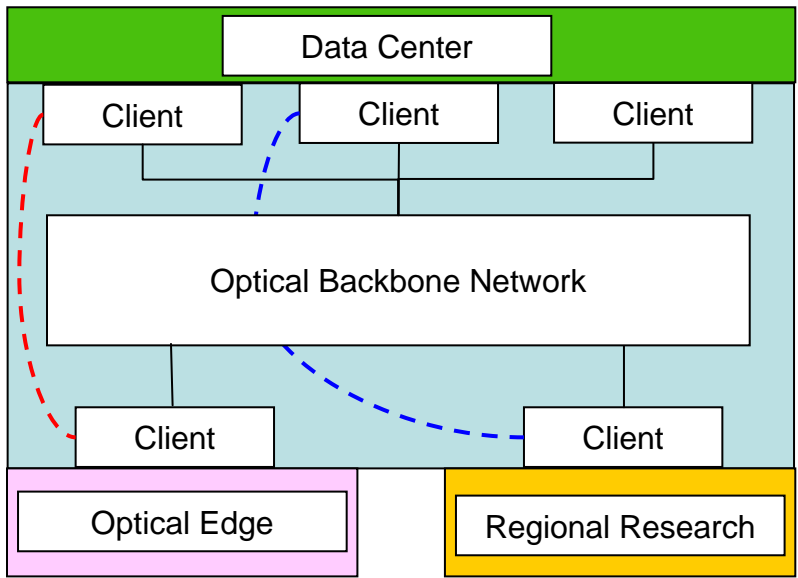
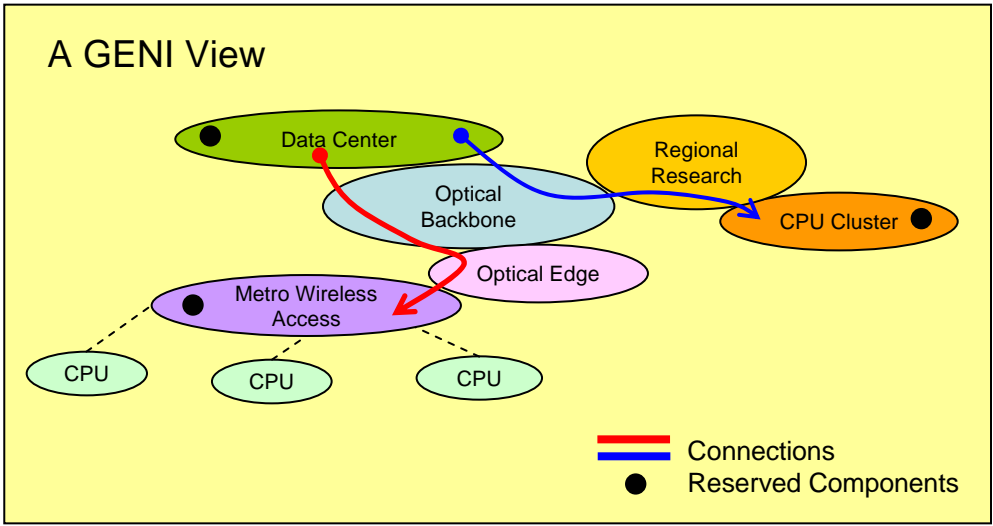
*Not required for understanding the use
case*

Processing Center: Discovery, Reservation and Configuration



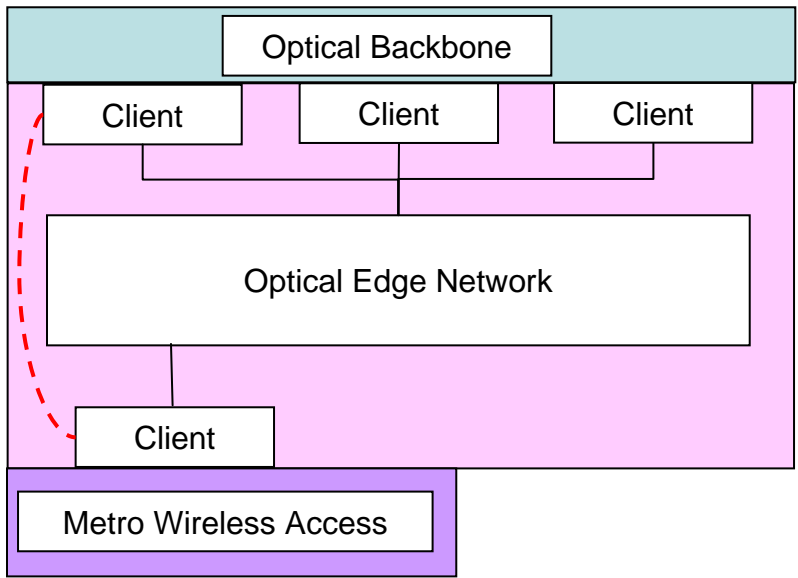
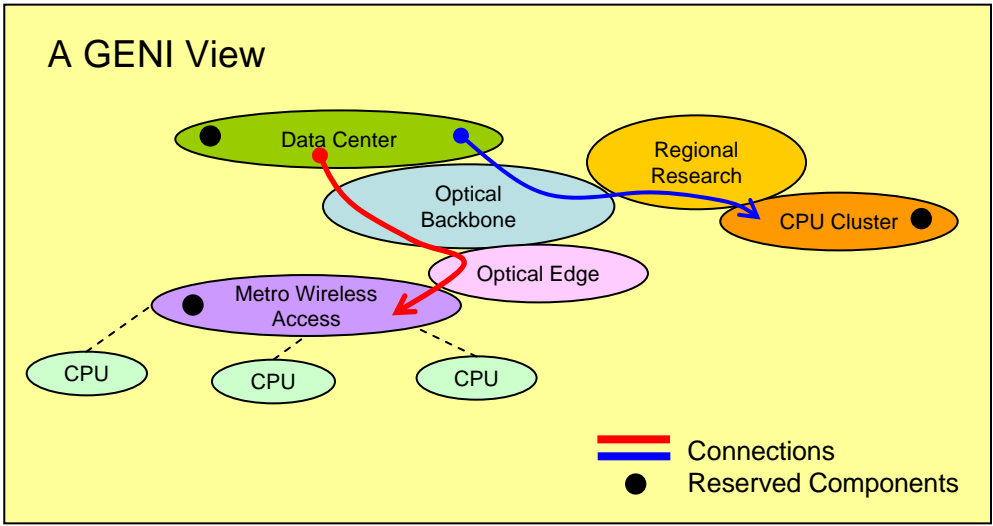
- The Processing Center is a service rich data center offering high performance computing, storage, servers, etc. all tied together with a complex network and connected into multiple client service interfaces on an optical backbone
- Processors, storage capacity, bandwidths and interface links discovered per Processing Center aggregate R-Spec
- Researcher makes request for Processing Center resources (storage capacity, processor cycles, link capacity between internal storage and processor, link service and capacity (e.g. GbE) at client interface with Optical Backbone
- Processing Center **aggregate controller** sets up virtual storage and processors (black boxes) connected with virtual links (dashed lines), and maps this to one or more of the client service interface with the Optical Backbone

Optical Backbone: Discovery, Reservation and Configuration



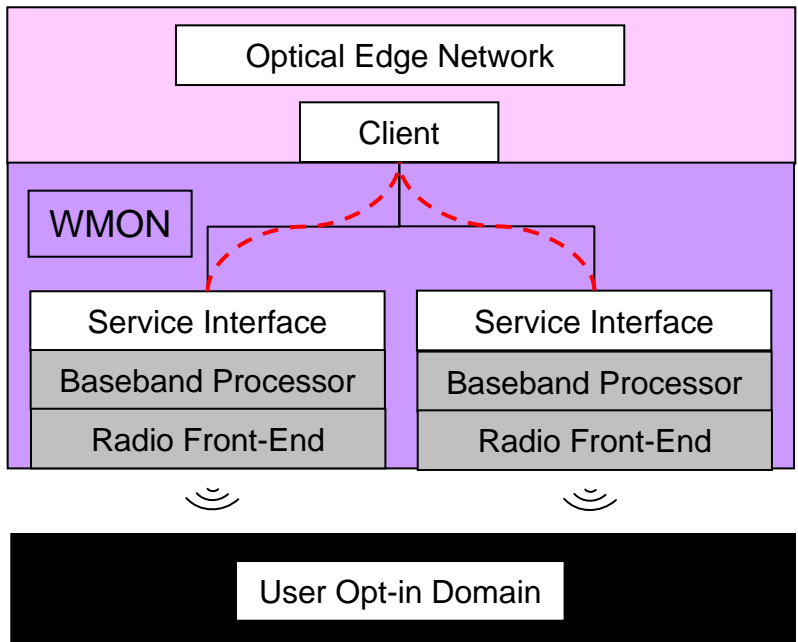
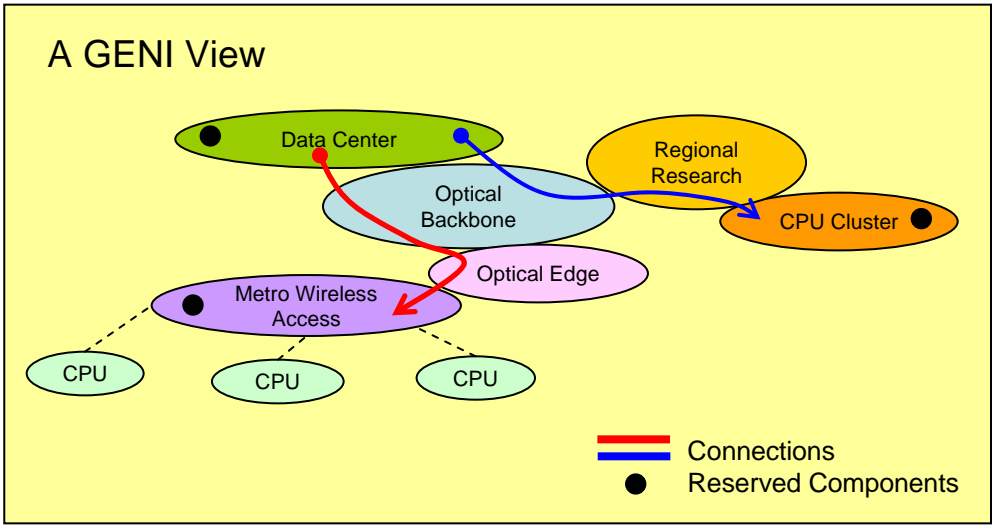
- The Optical Backbone is a national scale fiber optic network offering line and tributary rate services (lambda, SONET, VLAN, packet switched, etc.) with a geographically disperse set of access points and services
- Backbone data services, interfaces, rates, and geographic locations discovered per optical backbone R-spec
- Researcher makes request for Optical Backbone resources, data service between client interfaces with Optical Edge and Regional Research networks.
- Optical Backbone **aggregate controller** sets up virtual links (dashed lines), and maps them to the respective client interfaces with the Optical Edge and Regional Research networks

Optical Edge: Discovery, Reservation and Configuration



- The Optical Edge is a regional scale fiber optic network offering line and tributary rate services (lambda, SONET, VLAN, packet switched, etc.) with a geographically disperse set of access points and services
- Backbone data services, interfaces, rates, and geographic locations discovered per optical edge R-spec
- Researcher makes request for Optical Edge resources, data service between client interfaces with Metro Wireless Access network
- Optical Edge **aggregate controller** sets up virtual link (dashed lines), and maps it to the respective client interfaces with the Metro Wireless Access network

Metro Wireless Access: Discovery, Reservation and Configuration



- Metro Wireless Access is a network offering data services over a geographically disperse set of wireless access points all feeding into a single optical edge client service interface
- Data services, interfaces, rates, geographic locations, wireless access point resources (radio frequency, processors, memory), and wireless monitor all discovered per Metro Wireless Access R-spec
- Researcher makes request for data services, as well as processor, memory and radio resources for a set of wireless access points
- Metro Wireless Access **aggregate controller** sets up virtual link (dashed lines), and maps it to the respective client interfaces with the Optical Edge Network, as well as reserves radio resources (processor, memory, and frequency) for the select wireless access points