GIMI Tutorial

GIMI Team

University of Massachusetts Amherst

GEC 14, Boston, July 9th

Overview

- Introduction
- Goals
- OML/OMF on ExoGENI
- Visualization
- IRODS
- IREEL

Introduction

- Develop and deploy the GIMI instrumentation and measurement framework
- Capable of supporting needs of GENI experimenters and GENI infrastructure operators
- Uses the ORBIT Measurement Library (OML) and integrated Rule Oriented Data System (iRODS) as its basis
- Will provide libraries to instrument resources, to filter and process measurement flows, and to consume measurement flows
- It will use the iRODS data grid for archiving and further processing
- Will include access control based on accepted GENI policy and authorization mechanisms

Goals

- Provide easy-to-use I&M services for experimenters:
  - deploying slices in GENI on selected types of servers, VMs and racks, WiMAX installation, and sensor networks.
  - interconnected by various types of network paths
- Provide comprehensive infrastructure measurement services for infrastructure operators:
  - deploying measurement slices on selected types of servers, VMs and racks, and WiMAX installations
  - interconnected by a various types of network paths In addition
  - slices established by other infrastructure operators, or by experimenters, can be authorized to gather data from the measurement slices

Goals

- This project will build and operate two persistent services:
  - the GENI Measurement Data Archive Service (i.e., iRODS) and
  - the GENI Experimenter Portal Service
- Collaborate with existing I&M Spiral 2 projects and GEMINI

Initial Focus

- Start with OML, and provides easy-to-use tools for GENI environment
- I&M use cases for experiments:
  - Default OML Client in nodes that gathers passive measurements like INSTOOLS
  - Provides basic node-to-node connectivity tests using ping and Iperf
- Spiral 4 : ORCA servers, VMs and ExoGENI (RENCI) racks, and WiMAX
- Spiral 5: protoGENI/Emulab servers/VMs and InstaGENI (HP) racks

Today's Tutorial

- Based on preceding OMF/OML and ExoGENI tutorials.

- So far:
  - Basics on Orbit Measurement Library and Framework
  - Basics on ExoGENI

- We will build on topologies you have been setting up during ExoGENI tutorial.

Today's Tutorial

- Overall goal is to teach you how to instrument and measure experiments on ExoGENI racks.

- Familiarize you with:
  - OML/OMF on ExoGENI racks and pre-build image
  - Integrated Rule Oriented Data System (iRODS)
  - Internet Remote Emulation Experiment Laboratory (IREEL)

Trac Page

- The accompanying trac page for this tutorial can be found at the following link:

  http://groups.geni.net/geni/wiki/GIMIv1.0Tutorial

Note: The Trac page includes a lot of additional information that is not needed for this tutorial. We added this information to allow experimenters to repeat the tutorial by Themselves at some later point in time.

Common Errors!!

- Be careful cutting and pasting
  - **Doesn't work from Trac!**
  - Sometimes even txt file inserts hidden characters
  - Some lines require edits; don't cut and paste the newline
- Make terminal windows big
  - Sometimes terminal gets corrupted when pasting a line that runs to the next line (especially when backspacing)
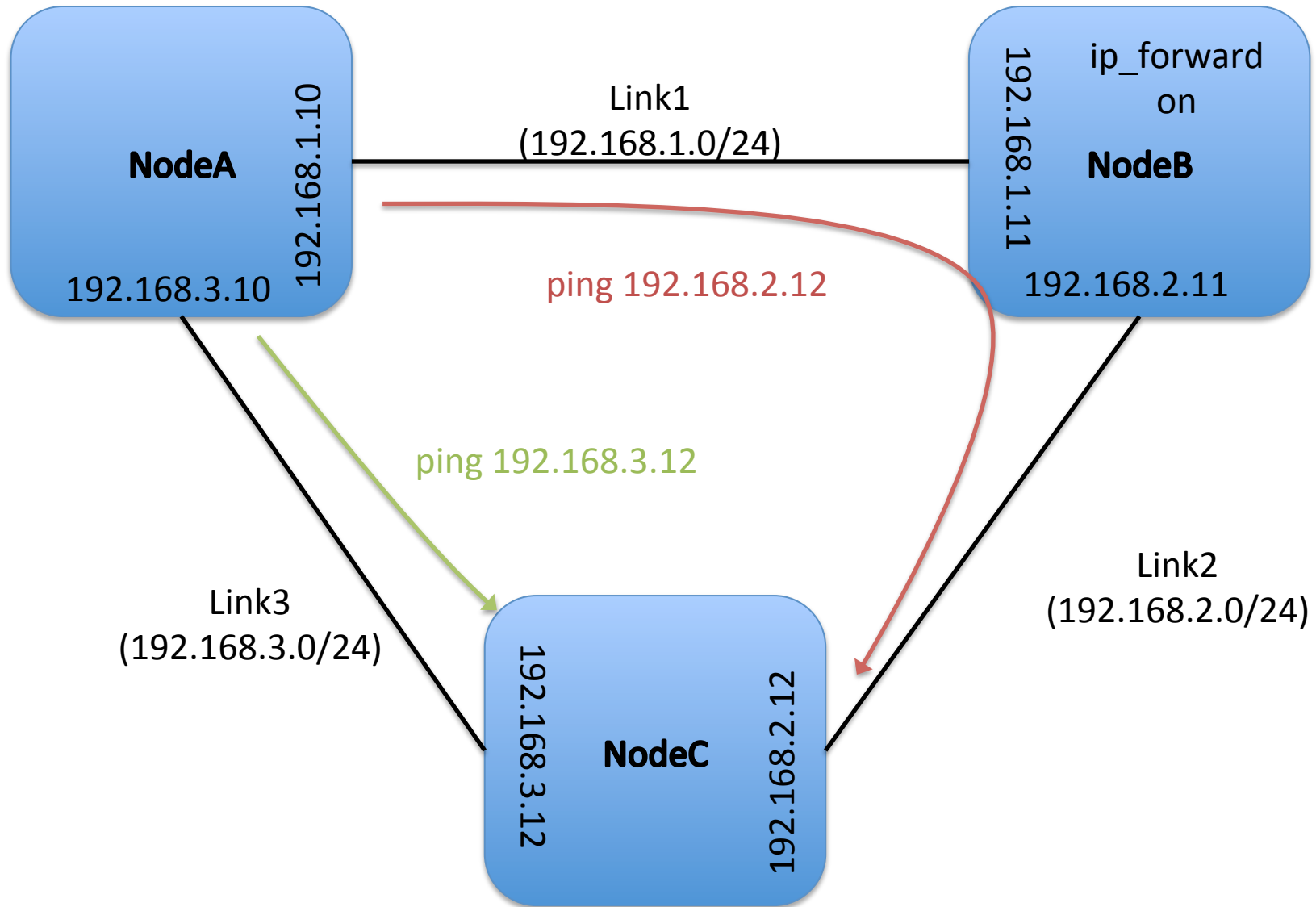
Cut and paste inside the VM:

- Firefox/Flukes:
  - Cut = Ctrl-X
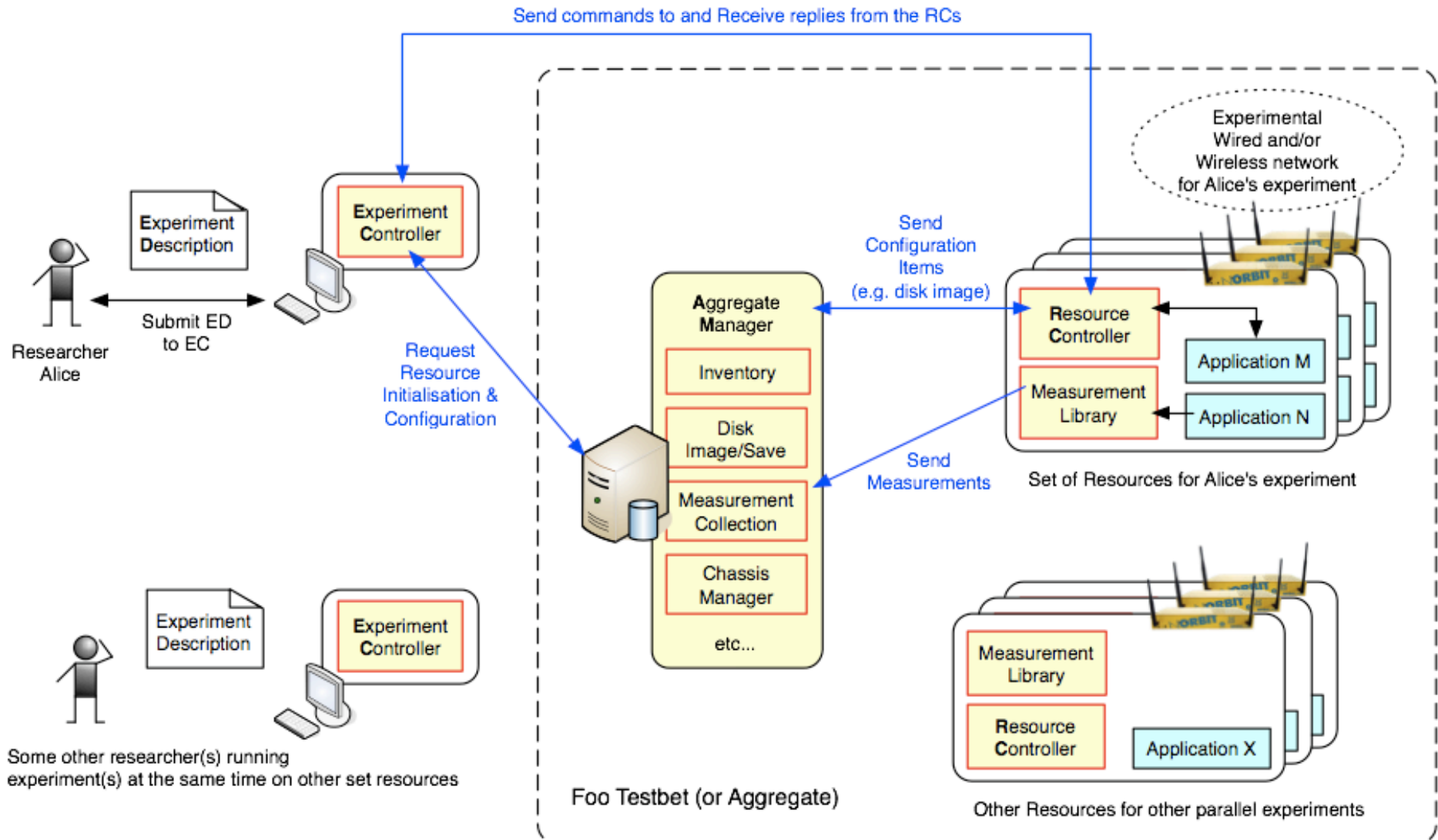 - Copy = Ctrl-C
  - Paste = Ctrl-V

- Terminal:
  - Cut = Shift+Ctrl-X
  - Copy = Shift + Ctrl-C
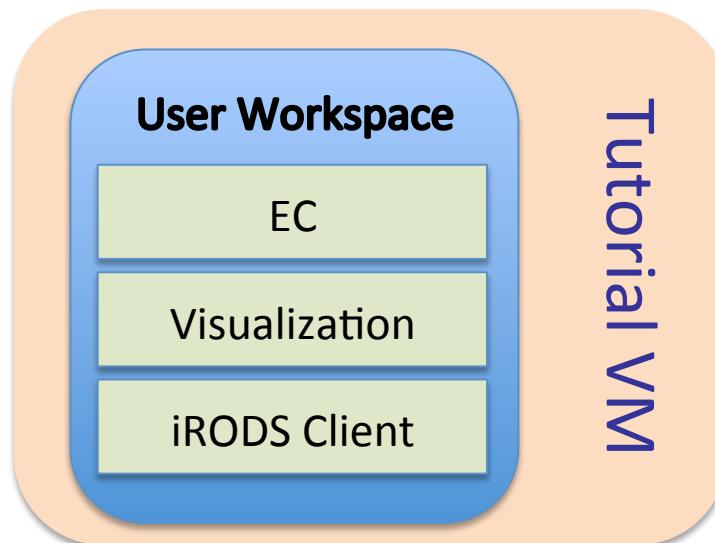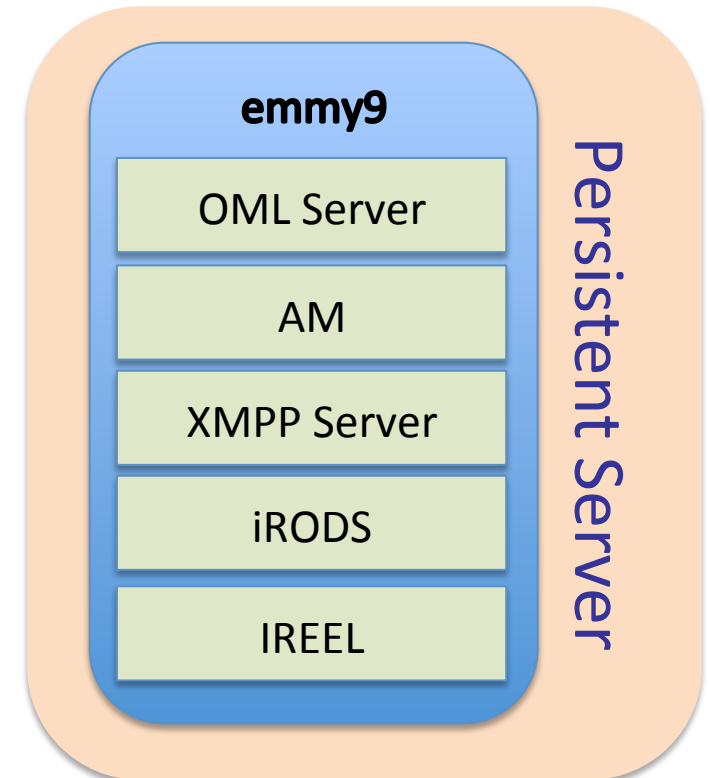 - Paste = Shift + Ctrl-V
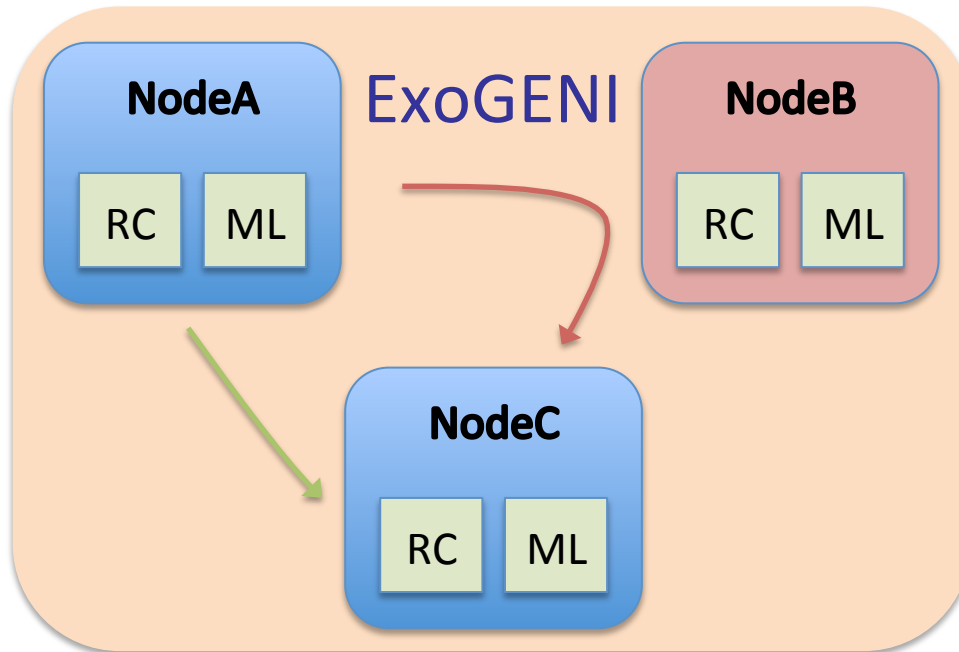
OML/OMF on ExoGENI (1)

- First part of tutorial:

  – Instrument topology created in preceding ExoGENI tutorial

  – Perform Iperf and ping (if time permits) measurement
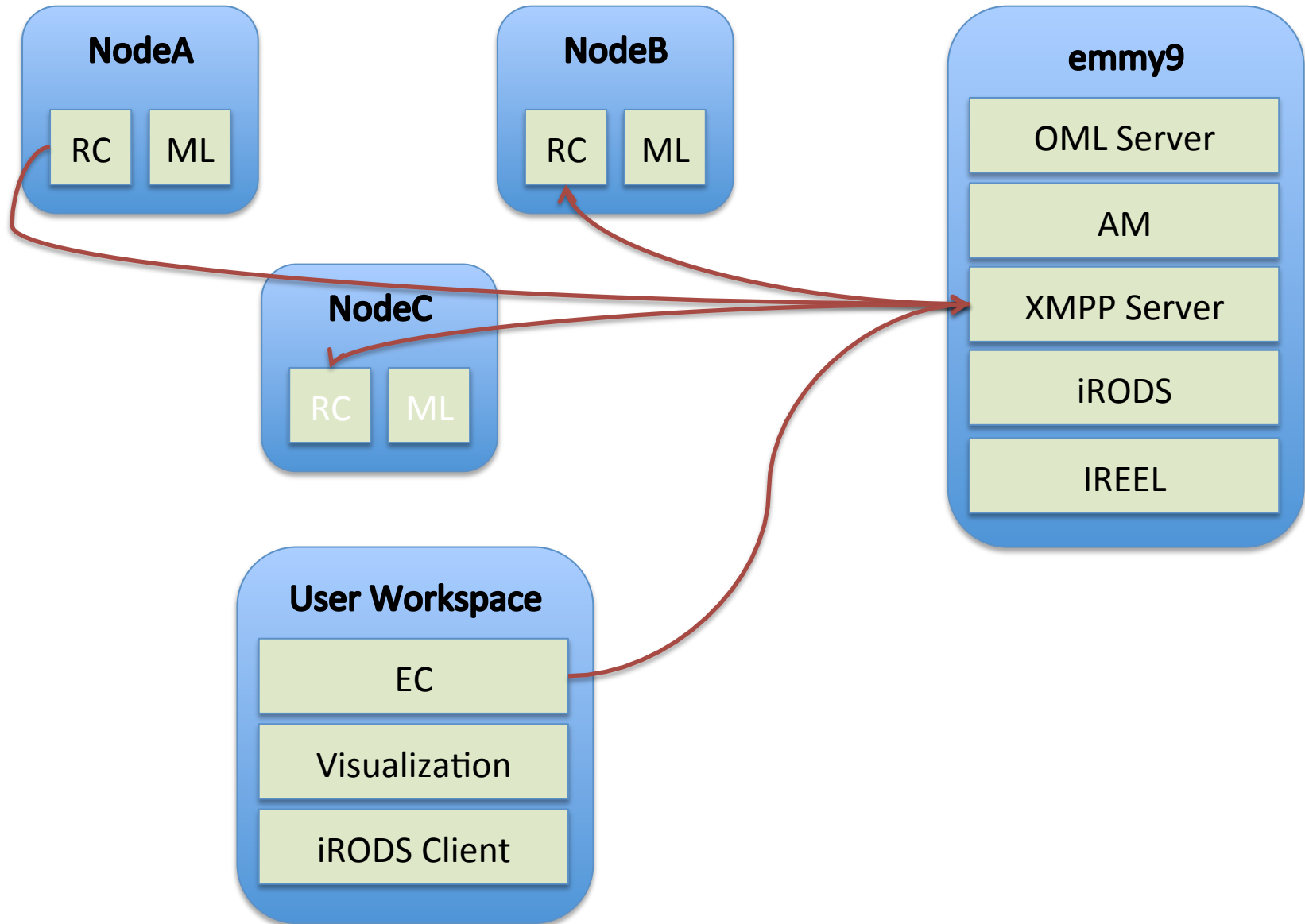
  – Analyze data from measurement

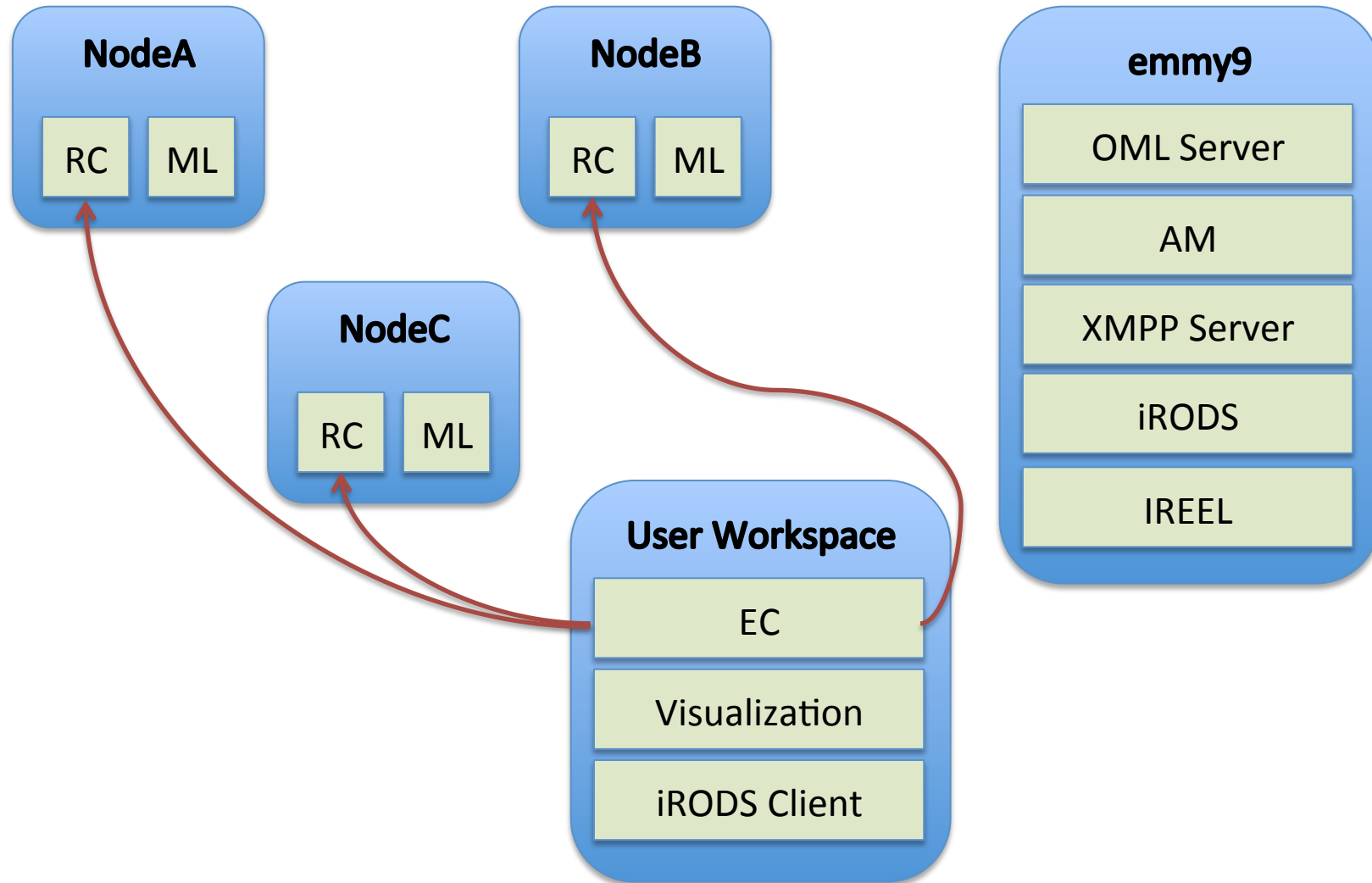# Experiment Topology

GIMI

NodeA
192.168.1.10
192.168.3.10

Link1
(192.168.1.0/24)

ip_forward on
NodeB
192.168.1.11
192.168.2.11

ping 192.168.2.12

ping 192.168.3.12

Link3
(192.168.3.0/24)

Link2
(192.168.2.0/24)

NodeC
192.168.3.12
192.168.2.12

13

OMF

Send commands to and Receive replies from the RCs

Experiment Description

Researcher Alice

Submit ED to EC

Experiment Controller

Request Resource Initialisation & Configuration

Experiment Description

Experiment Controller

Some other researcher(s) running experiment(s) at the same time on other set resources

Aggregate Manager

Inventory

Disk Image/Save

Measurement Collection

Chassis Manager

etc...

Foo Testbet (or Aggregate)

Send Configuration Items (e.g. disk image)

Experimental Wired and/or Wireless network for Alice's experiment

Resource Controller

Measurement Library

Application M

Application N

Set of Resources for Alice's experiment

Send Measurements

Measurement Library

Resource Controller

Application X

Other Resources for other parallel experiments
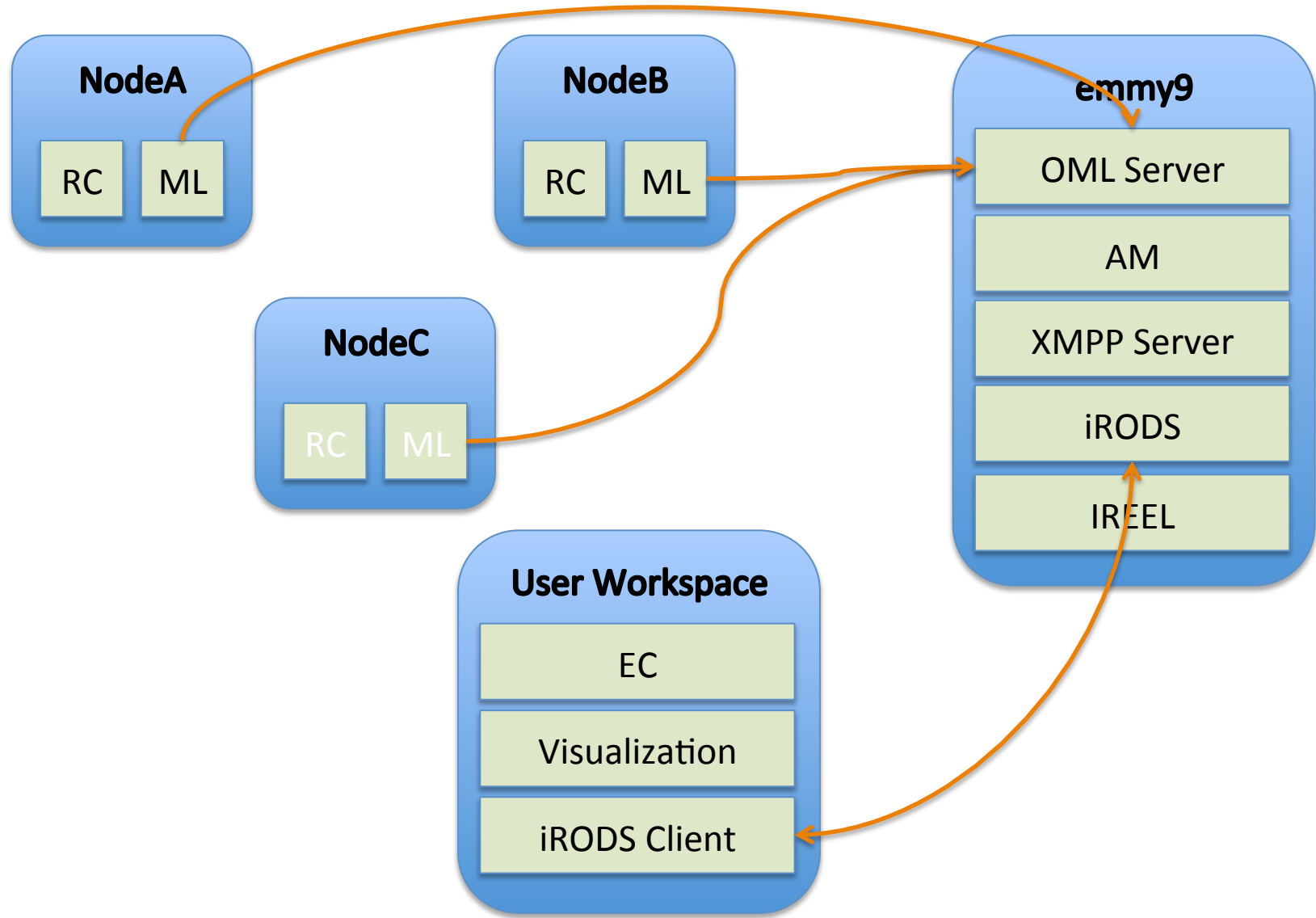
GIMI Measurement Environment

XMPP Messaging

EC/RC Messaging

Data Collection

OML/OMF on ExoGENI (2)

- Please bring up the VirtualBox image (GEC14_Tutorials_Final.ova).

- Log into VM with the *gec14user* and use PW *gec14user*.

- Open a Firefox web browser

- Open a terminal window.

OML/OMF on ExoGENI (3)

- Please use the topology you created in the ExoGENI tutorial.

- Download GIMI tutorial specific configuration files by issuing the following command in a user workspace terminal:

```
$ cd ~/Tutorials/GIMI/common/
$ git pull
```

- VMs on ExoGENI nodes include the following SW:
  - OMF (AM, RC, EC), OML, OMLified Iperf, Nmetrics, iRODS client,R

OML/OMF on ExoGENI (4)

- Copy the template iRODS configuration file to .irodsEnv with the following command:

```
$ cp ~/.irods/gimiIrodsEnv ~/.irods/.irodsEnv
```

- Open ~/.irods/.irodsEnv with your favorite text editor (we recommend *nano*) and change gimiXX to your assigned username (e.g., gimi04).

OML/OMF on ExoGENI (5)

iRODS personal configuration file.

This file was automatically created during iRODS installation.

Created Thu Feb 16 14:06:27 2012

iRODS server host name:

irodsHost 'emmy8.casa.umass.edu'

iRODS server port number:

irodsPort 1247

Default storage resource name:

irodsDefResource 'iRODSUmass' (remove 1)

Home directory in iRODS:

irodsHome '/geniRenci/home/gimiXX'

Current directory in iRODS:

irodsCwd '/geniRenci/home/gimiXX'

Account name:

irodsUserName 'gimiXX'

Zone:

irodsZone 'geniRenci'

OML/OMF on ExoGENI (6)

- Register with iRODS server by issuing the following command (more details on iRODS will be given shortly):

  ```
  $ iinit
  ```

- You will be prompted for a password. Please type in the password you were provided with on the paper handout!!

OML/OMF on ExoGENI (7)

Upfront:

- OMF controller controls the experiment based on the unique hostname and Experiment name.

- In the case of this tutorial the individual experiment name is: gimiXX-tutorial

- ExoGENI nodes hostnames are always initally set to "debian"

OML/OMF on ExoGENI (8)

Upfront: ExoGENI post boot script sets

- the hostnames to the unique node names: E.g., gimiXX-tutorial-nodeA, gimiXX-tutorial-nodeB, gimiXX-tutorial-nodeC

- the experiment name to the slice name of your ExoGENI request, which should be unique: gimiXX-tutorial

Post boot script

```
$ hostname gimi01-tutorial-nodeA

$ apt-get update

$ curl https://pkg.mytestbed.net/ubuntu/oneiric/oml2-
    iperf_2.0.5-1ubuntu5_amd64.deb -o /root/iperf.deb

$ dpkg -i /root/iperf.deb

$ route add -net 192.168.2.0 netmask 255.255.255.0 gw 192.168.1.11

$ curl http://emmy9.casa.umass.edu/pingWrap.rb -o /root/pingWrap.rb

$ chmod +x /root/pingWrap.rb

$ gem install oml4r

$ omf_create_psnode-5.4 emmy9.casa.umass.edu mkslice gimiXX-tutorial
gimiXX- tutorial-nodeA

$ curl http://emmy8.casa.umass.edu/enrolled.patch -o enrolled.patch

$ patch -p1 < /enrolled.patch

$ curl http://emmy8.casa.umass.edu/omf-resctl.yaml -o /etc/omf-
resctl-5.4/omf-    resctl.yaml

$ perl -i.bak -pe "s/\:slice\:/\:slice\: gimi01-tutorial/g" /etc/omf-
    resctl-5.4/omf-resctl.yaml

$ /etc/init.d/omf-resctl-5.4 restart
```

OML/OMF on ExoGENI (9)

Verifying initial setup:

- Once the VMs are brought up on ExoGENI nodes, open three terminals in your VM.

- Login to each of the ExoGENI nodes.
  - Use Flukes
  - Click on the *Manifest* tab
  - Specify the name of your slice: *gimiXX-tutorial*
  - Click *Query for Manifest*
  - Rightclick on each node and select *Login to node*

27

OML/OMF on ExoGENI (10)

Verifying initial setup:

- Check if the hostnames are set to their node names by typing "hostname" command in each of the three terminals.

- You can verify if the "hostname" and "Experiment name" are set in the RC config file.

- `$ cat /etc/omf-resctl-5.4/omf-resctl.yaml`

OML/OMF on ExoGENI (11)

Verifying initial setup:

- Verify the ":name:" tag to be set to "hostname", and ":slice:" tag to be set to "Slice name".

- If both are set correctly, we are all set with the configuration of the ExoGENI VMs.

OML/OMF on ExoGENI (12)

Step 1: How to start RC on ExoGENI VMs

- Re-Start the RC on each of the nodes:
  ```
  $ /etc/init.d/omf-resctl-5.4 restart
  ```

- Verify if the node is connected to the XMPP server by checking the log:
  ```
  $ cat /var/log/omf-resctl-5.4.log
  ```

OML/OMF on ExoGENI (13)

Step 1: Example log file

```
2012-06-30 23:49:10 DEBUG nodeAgent::OMFPubSubTransport:
Listening on '/OMF/dilip-testing/resources/nodeA' at
'emmy9.casa.umass.edu'


2012-06-30 23:49:10 DEBUG nodeAgent::OMFPubSubTransport:
Listening on '/OMF/dilip-testing' at
'emmy9.casa.umass.edu'
```

Resource Controller

```
# Agent settings
:agent:

  # Name (i.e. unique HRN ID) of this resource
  # Or this could also be passed as a command line parameter "--name"
  # This is either a fully defined string, e.g. "my_resource_name"
  # Or a string for which some values will be replaced by the running RC,
  # currently we support the values: %hostname%, %macaddr%, %fqdn%
  # For example, if you use "some_prefix.%hostname%.some_suffix"
  # Then if your hostname is 'node1', then your RC name will be
  # 'some_prefix.node1.some_suffix'          Node Name
  # %macaddr% is replaced with the MAC address of the control interface
  # %fqdn% is the fully qualified hostname (incl. domain part)
  :name: 'nodeA'
                                      Experiment Name
  # Name (i.e. unique HRN ID) of the slice to which this resource is assigned
  :slice: gec14tutorial

  # shrink the filesystem size before saving an image
  # grow the filesystem to match the disk size after loading an image
  # enabling this may slow down load/save significantly
  :resizefs: false

  # driver to use for Atheros AR5*** cards
  # valid options are madwifi, ath5k
  :ar5xxx_driver: ath5k
```

OML/OMF on ExoGENI (14)

Step 4: Start the experiment .

- You are given an example experiment written in ruby. We start the experiment with OMF controller.

- The experiment description is stored on the tutorial VM under *~/Tutorial/GIMI/common/tcp_iperf.rb*

- Start the Experiment:
  ```
  $ omf-5.4 exec --no-cmc -S gimiXX-gec14
  tcp_iperf.rb -- --source1 gimiXX-tutorial-
  nodeA --sink gimiXX-tutorial-nodeC
  ```

OML/OMF on ExoGENI (15)

Step 4: Results.

- Ignore:

  ERROR NodeHandler: The resource 'nodeC' reports that an error occurred
  ERROR NodeHandler: while running the application 'iperf_app#3'
  ERROR NodeHandler: The error message is 'INFO Net_stream: connecting to host tcp://
  emmy9.casa.umass.edu:3003

- If following message don't appear after approximately 3 minutes, check with us!!

  INFO EXPERIMENT_DONE: Event triggered. Starting the associated tasks. INFO NodeHandler:
  INFO NodeHandler: Shutting down experiment, please wait... INFO NodeHandler: INFO run:
  Experiment gimi04Test070204 finished after 3:02

OML/OMF on ExoGENI (17)

Step 5: Visualization

- The experiment finishes up and the results will automatically be copied to your iRODS home directory. (We will talk about this in a minute.)

Visualization of Data (1)

- Open another terminal in you user workspace.

- Goto "~/Tutorials/GIMI/common" directory

- Run the following command to create the visualization of the experiment carried out. (Please change "gimiXX" in the command to your username.)
  ```
  $ cd ~/Tutorials/GIMI/common
  $ ./tutorial_viz.sh gimiXX-tutorial
  ```

Visualization of Data (2)

- Open the firefox browser and type "127.0.0.1/oml.html" to view the visualization!!

- The visualization script contains a "R" script to generate pdf/jpg based on the sqlite3 measurement database file generated by the OMLified application.

- The script is located in ~/Tutorials/GIMI/common/R_script_viz.r.

Running Experiment for Longer

- If you want to run the experiment for long time, edit the "tcp_iperf.rb" file to change the interval or edit the "tutorial_ec_script.sh" to change the EC command and run the experiment in a loop.

- After you have done editing, execute the bash script in command line by typing
```
$ cd ~/Tutorials/GIMI/common
$ ./tutorial_ec_script.sh
```

iRODS

- Integrated Rule-Oriented Data System that aims at managing distributed massive data
- Open source initiative (+13 year development and ~ $20M NSF funding)
- Collaboration between UNC (DICE), RENCI, UCSD
- Applications:
  - Data grids, Institutional repositories, Libraries, Archives
  - Astronomy, High Energy Physics, Earth, Environment, Genomics…
- Scale: hundreds of millions of files, petabytes of data, tens of federated data grids
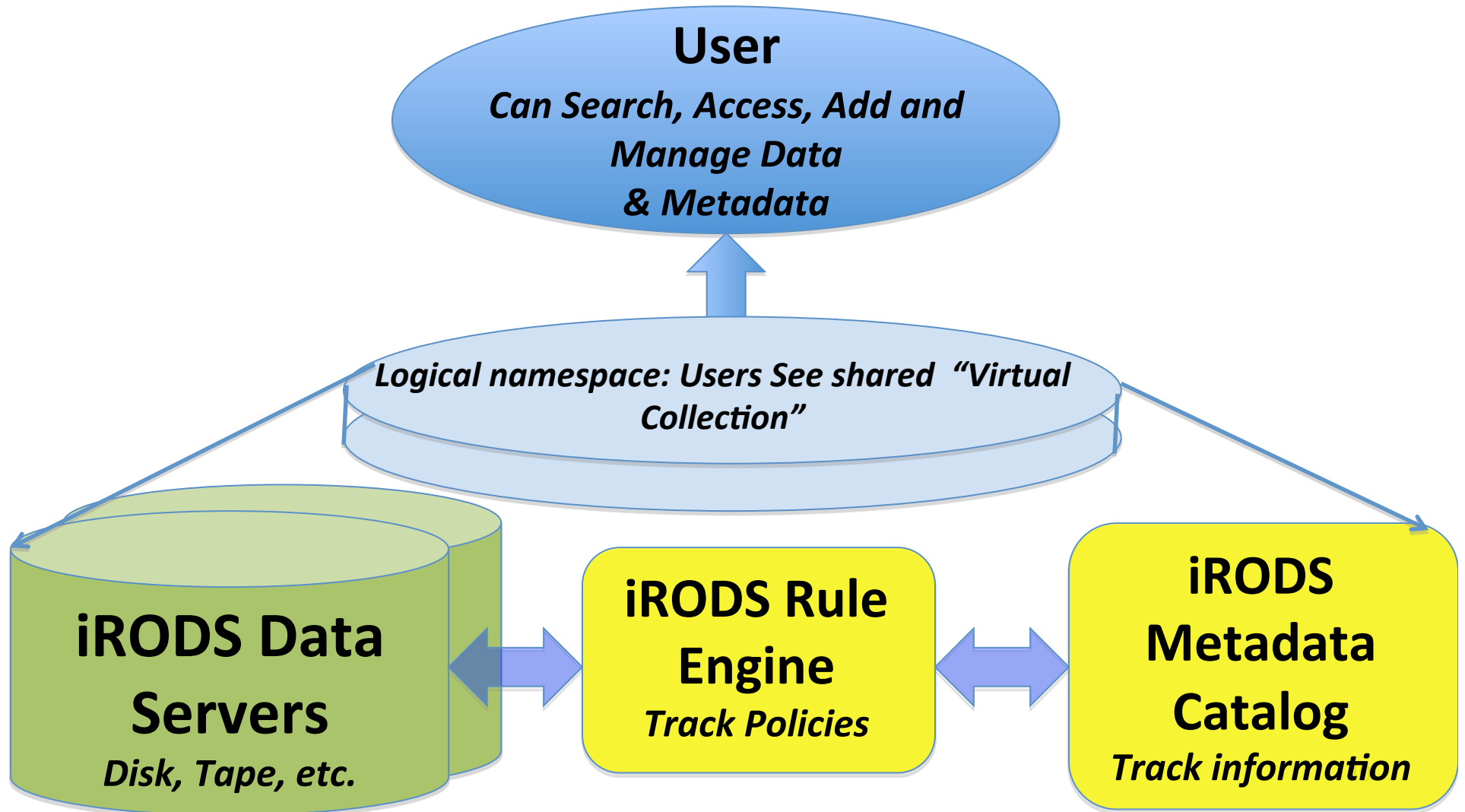
Managing Data - Virtualization

Access Interface

Standard Micro-services

Data Grid

Standard Operations (POSIX,ODBC…)

Data obj, DB obj, Workflow

Map from actions requested by the access method to a standard set of Micro-services.

Map the standard Micro-services to standard operations.

Map the operations to protocol supported by the operating system.

Managing Computation

- Why?
  - May be easier to move computation to data when data size is too large and computation is simple
  - Reduce latency by local processing can be critical

- How?
  - Rule: Action | Condition | MS_1, …, MS_n | RMS_1, …, RMS_n
  - Micro-services: +250 well-defined functions
  - Rules invoked by servers to enforce policies
  - Rules invoked by clients to run workflows on servers

Policies – Actionable Rules

- Retention, disposition, distribution, arrangement
- Authenticity, provenance, description
- Integrity, replication, synchronization
- Deletion, trash cans, versioning
- Archiving, staging, caching
- Authentication, authorization, redaction
- Access, approval, IRB, audit trails, report generation
- Assessment criteria, validation
- Derived data product generation, format parsing
- Federation of independent data grids

Data Life Stages in iRODS' View

Data Collection	Data Grid	Data Processing Pipeline	Digital Library	Reference Collection	Federation
Private	Shared	Analyzed	Published	Preserved	Sustained
Local Policy	Distribution Policy	Service Policy	Description Policy	Representation Policy	Re-purposing Policy
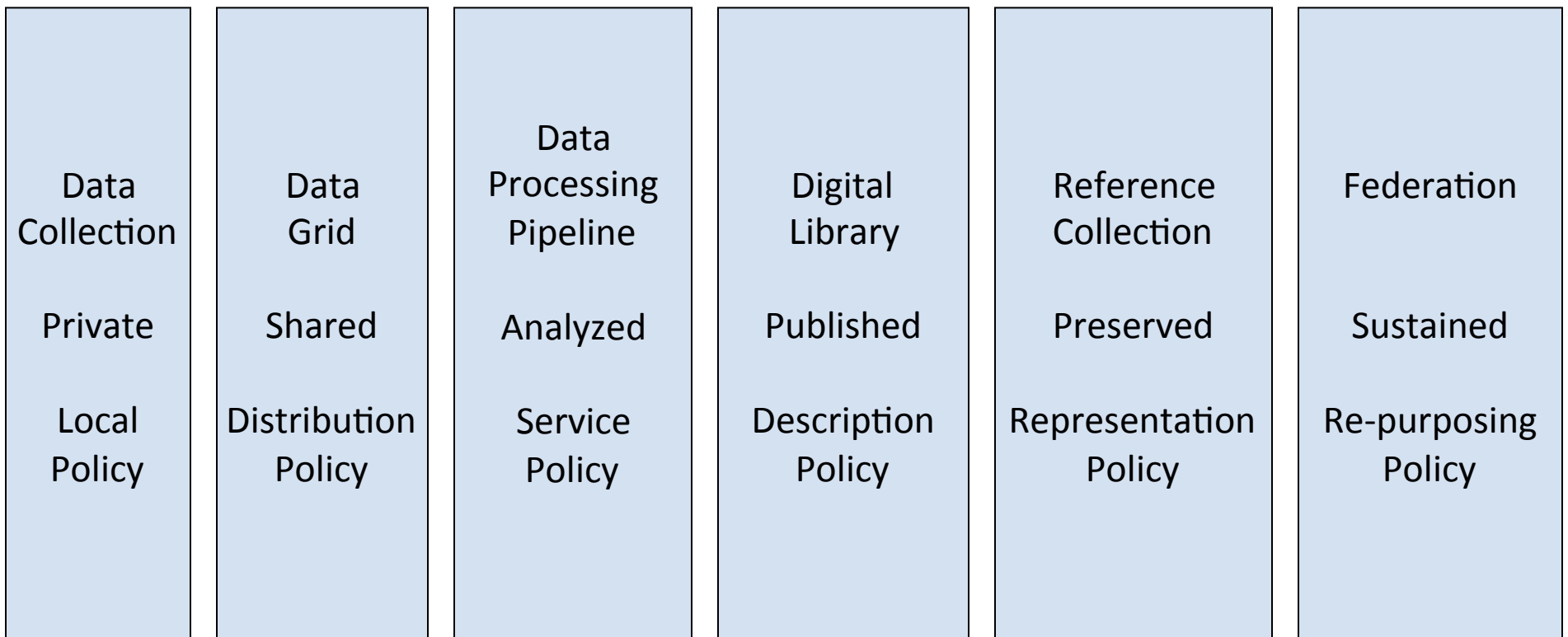
Applications

- International projects
  - Cyber Square Kilometer Array (radio astronomy), Cinegrid (movies)
- National data grids
  - Australia, New Zealand, Portugal, UK, France
- Federal agency archives
  - NASA Center for Climate Simulation, National Optical Astronomy Observatories, Ocean Observatories Intiative
- Institutional repositories
  - French National Library, Carolina digital repository, Broad Institute genomics data, Sanger Institute

iRODS and GIMI

- Measurement data is automatically stored in user's iRODS home directory.

- Retrieve file from your iRODS home directory into user workspace:

  ```
  $ iget <file_name>
  ```

- Store file in you iRODS home directory:

  ```
  $ iput <file_name>
  ```

iRODS and GIMI

- iRODS web interface: https://www.irods.org/web/index.php

- Host/IP: emmy8.casa.umass.edu
  Port: 1247
  Username: as given on printout
  Password: as given on printout

IREEL

- Internet Remote Emulation Experiment Laboratory

- GIMI portal

- Will allow you to run experiments in ExoGENI slice

- Manage experiment and data through IREEL

- Make experiments truly repeatable

IREEL and GIMI

- ExoGENI slice allocated through Flukes (as performed earlier in the tutorial).

- Execute experiment from IREEL (ED is actually identical to the one we used earlier).

- Visualize data within IREEL.

- GIMI's IREEL:
  http://emmy9.casa.umass.edu:8080/

Creating an Experiment
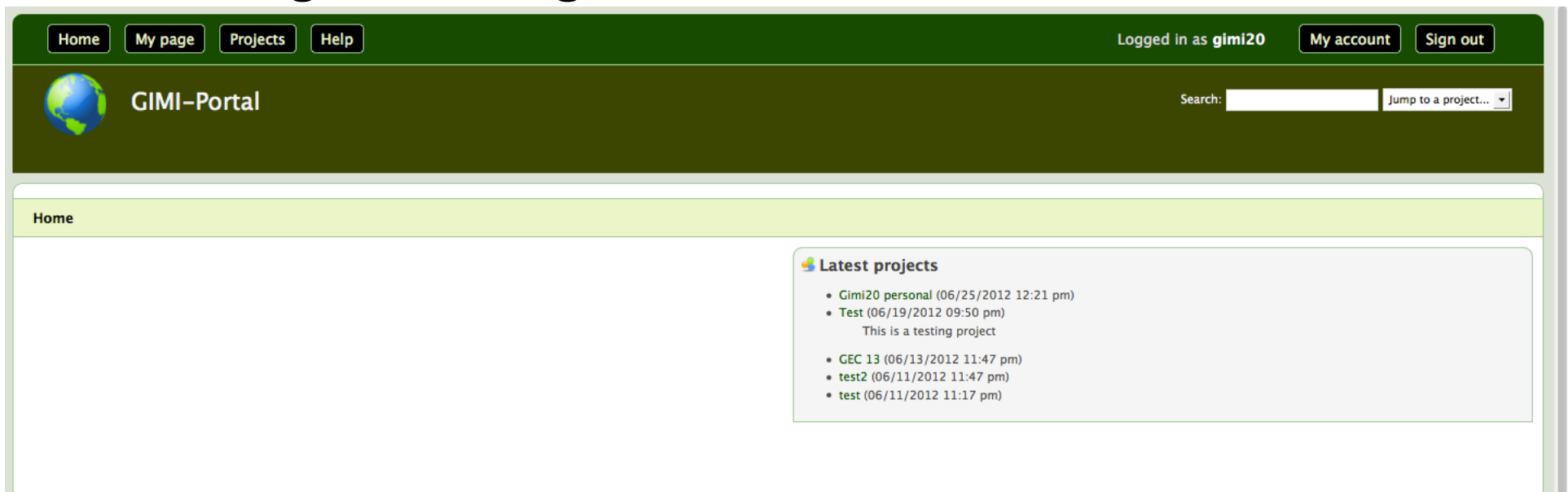
Running Experiment with IREEL (1)

- Reserve resources via Flukes

- Open IREEL in Browser
    - http://emmy9.casa.umass.edu:8080
    - Register or login

Running Experiment with IREEL (1)

- Upload experiment script using *Experiment Scripts* button

- Create experiment using *New Experiment* button

- Click *Create button* to run experiment

- Wait to see result display from IREEL!

Presenting Results

Mailing List

Your place to get help:

- geni-gimi-user Google group

Thank You
Questions?