

Extending PrimoGENI for Symbiotic Distributed Network Emulation

Jason Liu
Florida International University

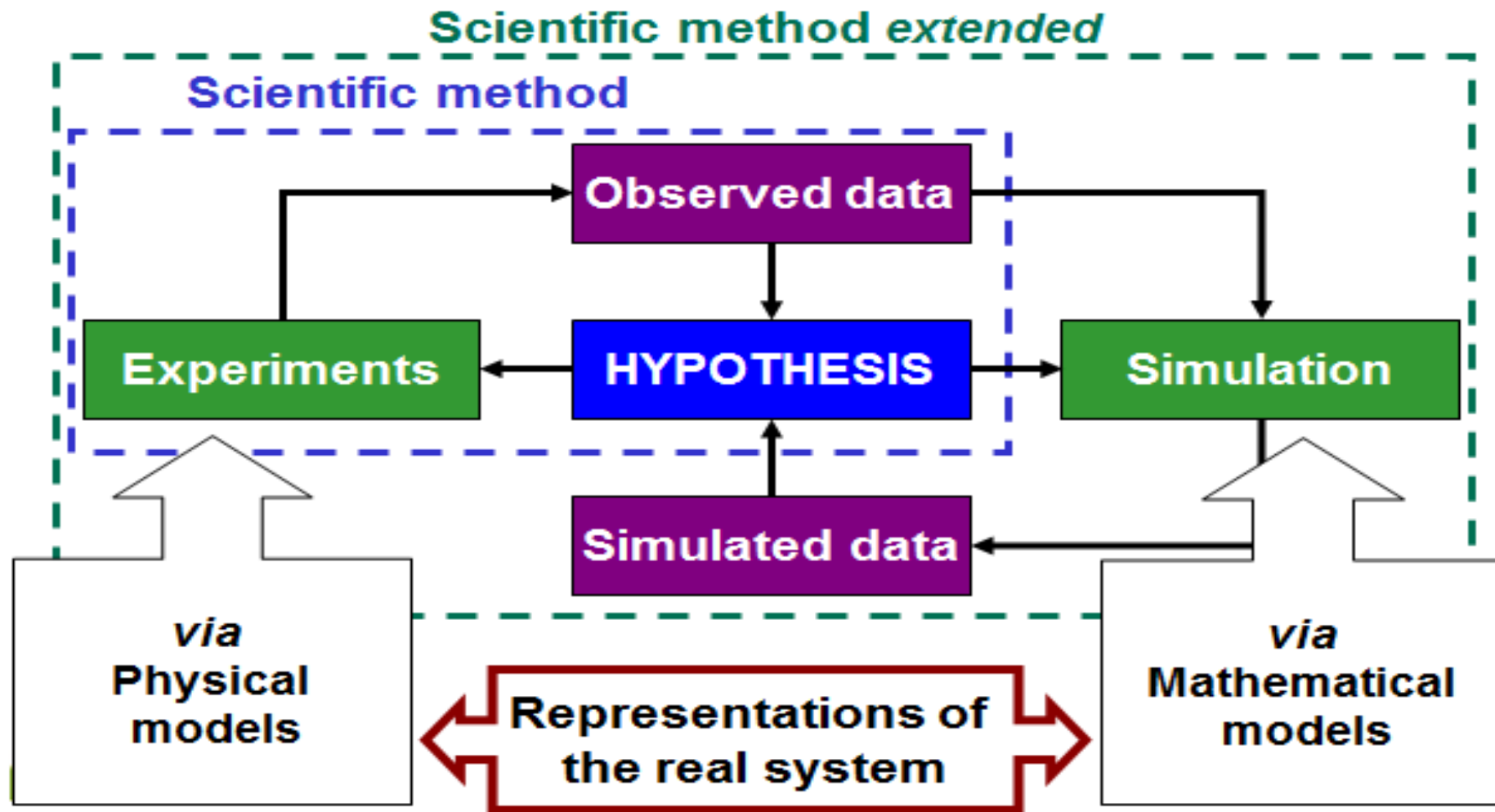


GENI Regional Workshop, Miami, Florida, March 13, 2017

What's *Prí* *eni*

- Add network simulation to GENI experiments
- Allow hybrid network experiments, including *simulated*, *emulated*, and *physical* components, on GENI

Modeling and Simulation



James Webb Telescope, price tag: \$8.7B

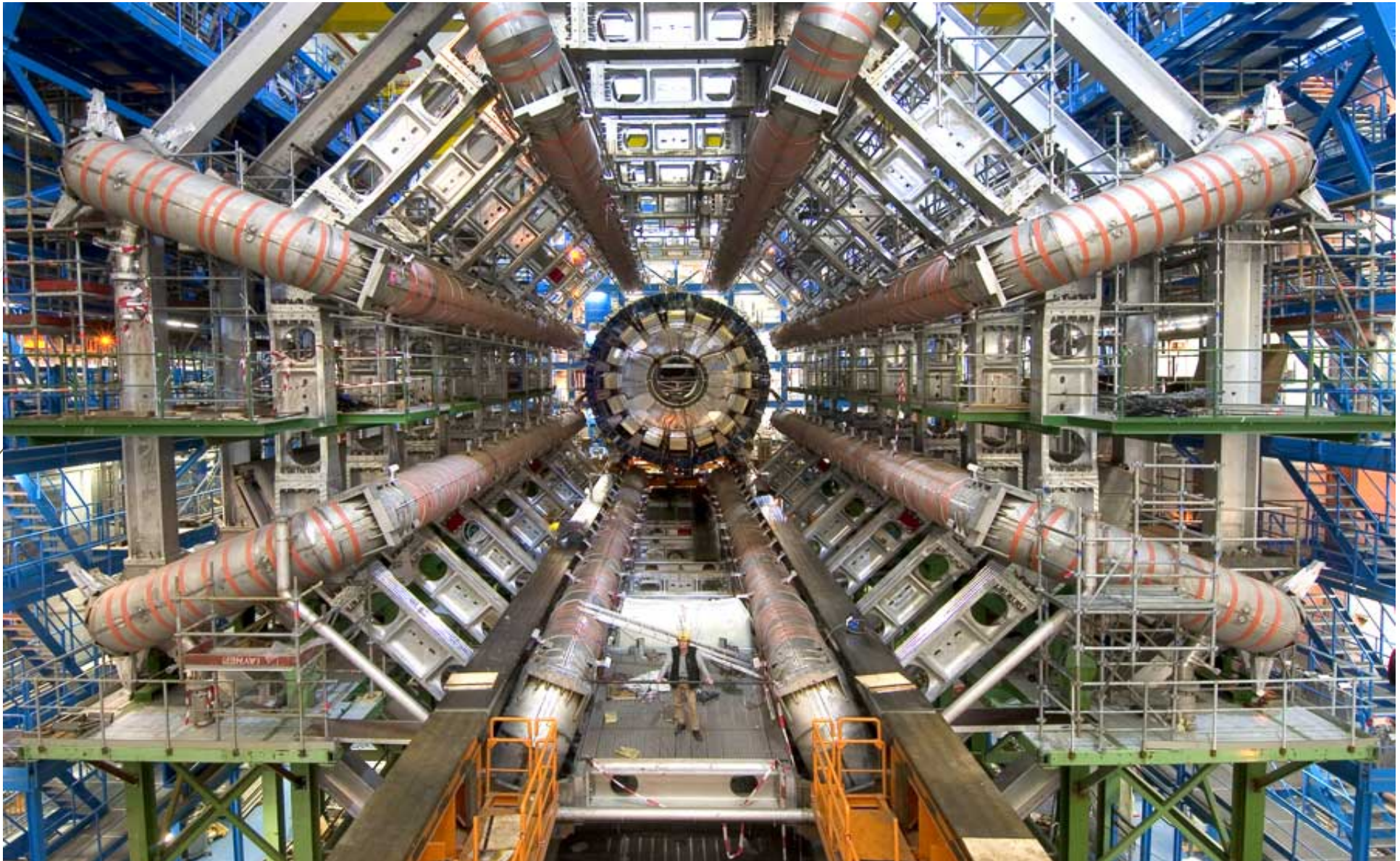
4



http://www.sciencemag.org/sites/default/files/styles/article_main_large/public/images/gallery-webb-telescope_0.jpg?itok=PlI4je6v

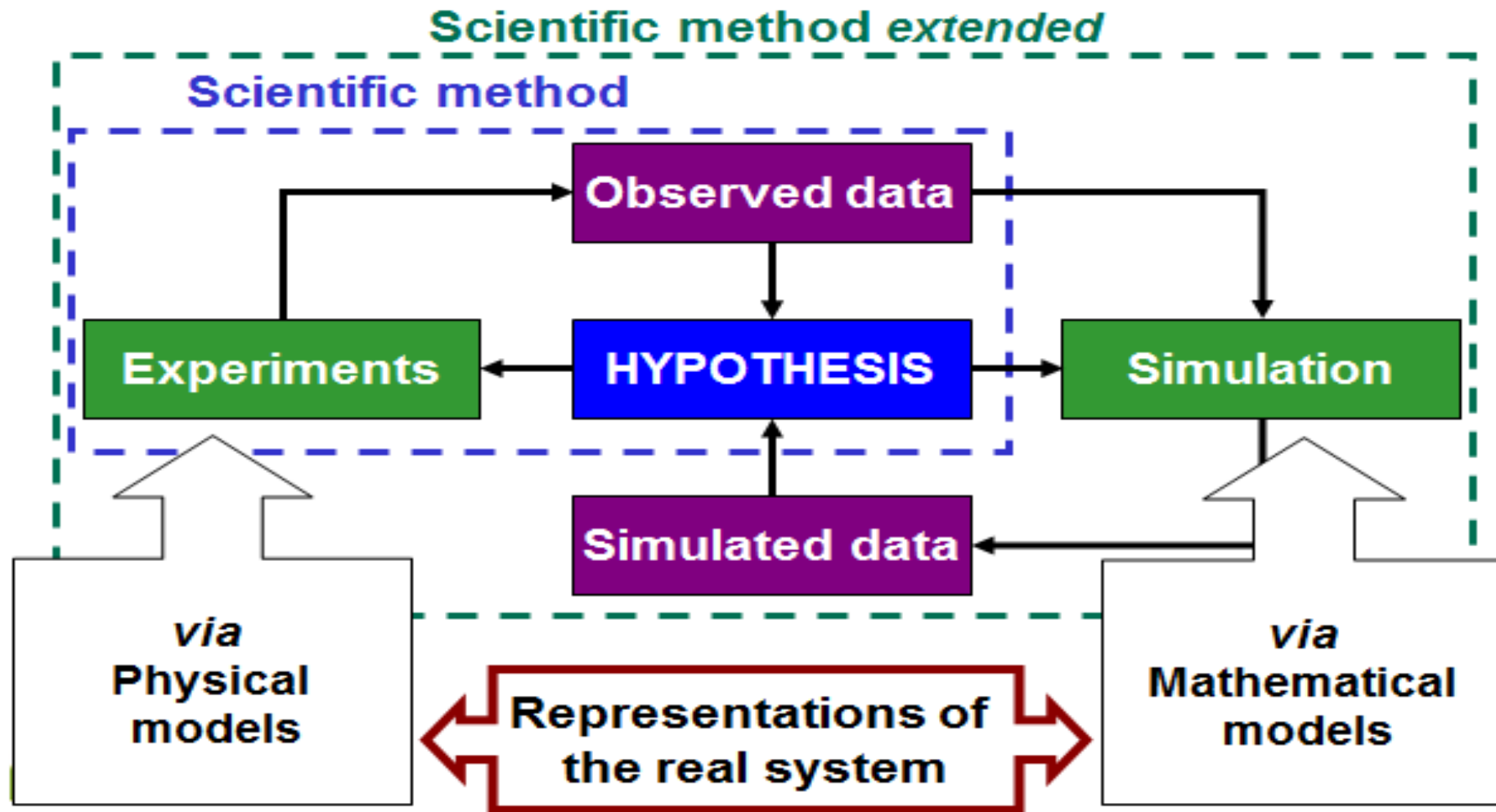
Large Hadron Collider, price tag: €7.5B

5



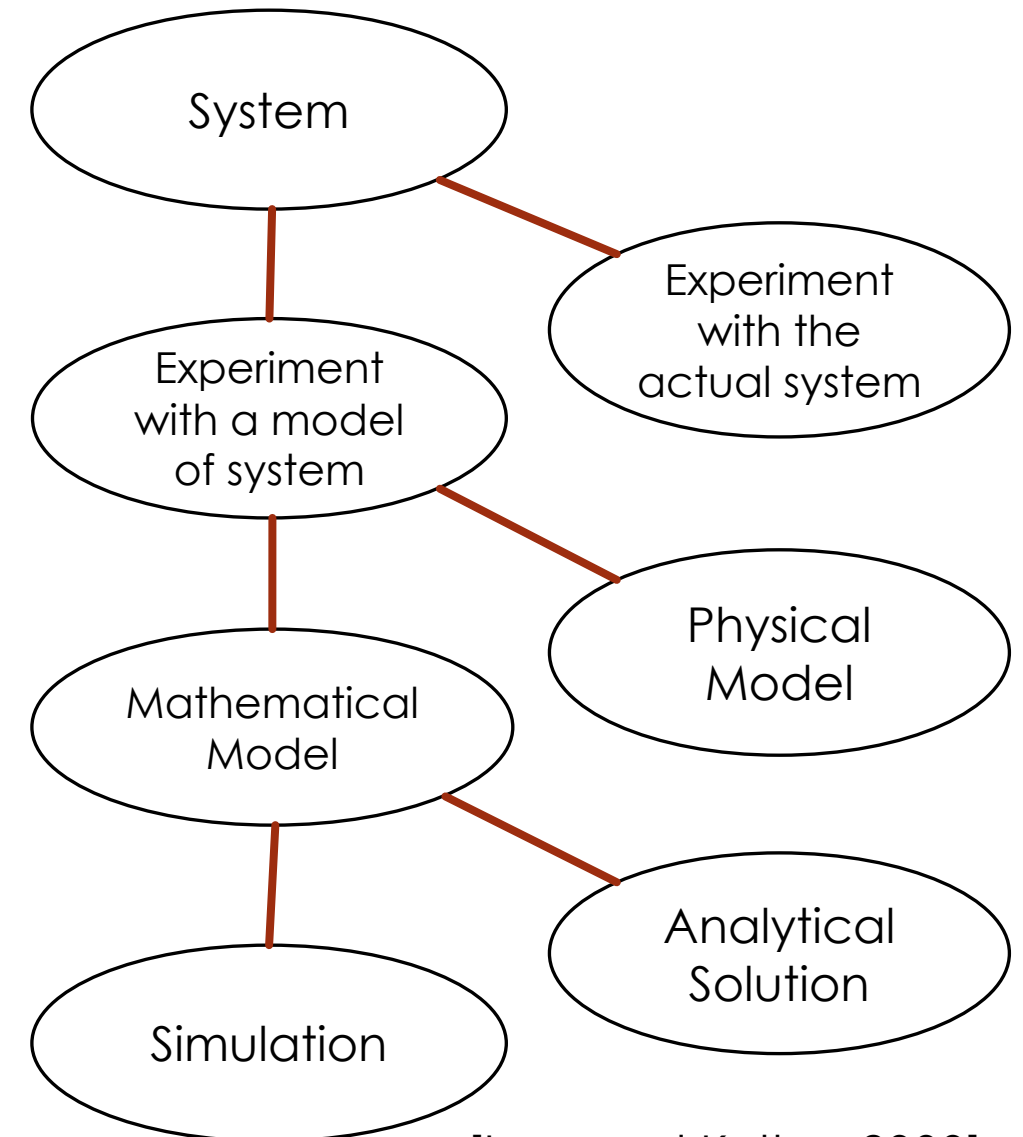
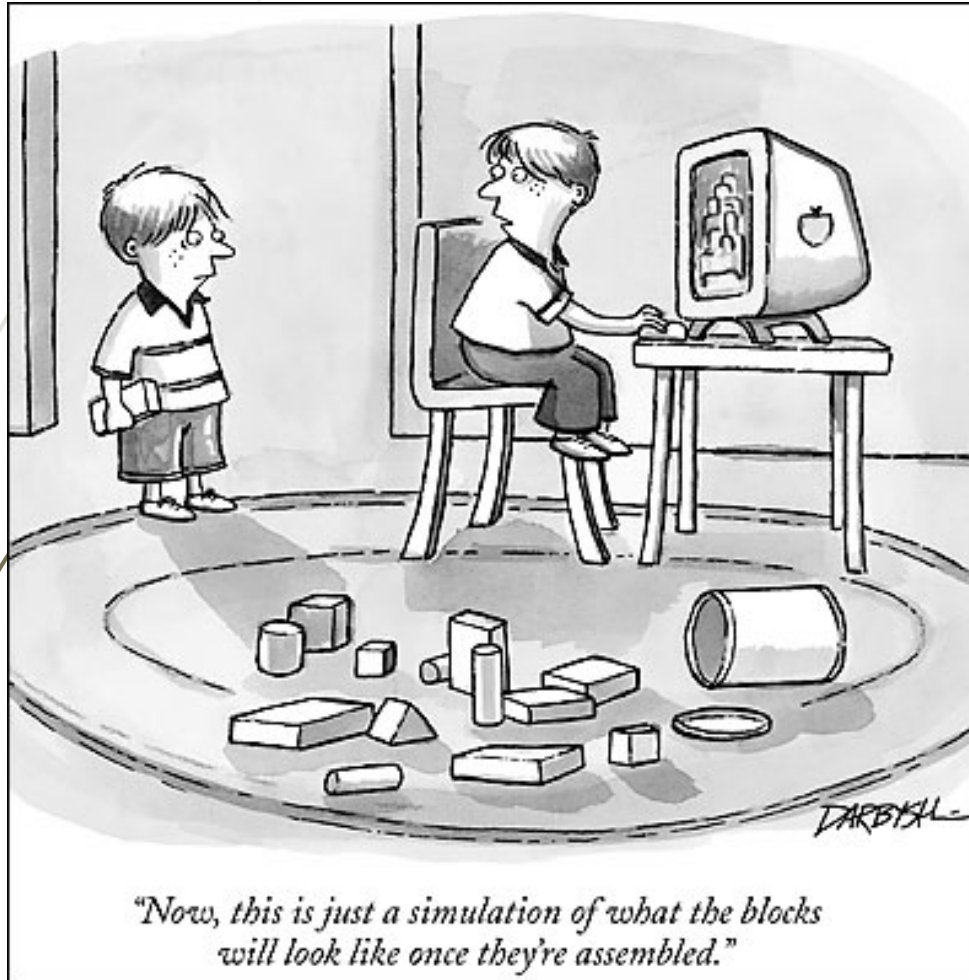
<http://apod.nasa.gov/apod/ap080225.html>

Modeling and Simulation Extends Scientific Method



What is Modeling and Simulation (M&S)?

7

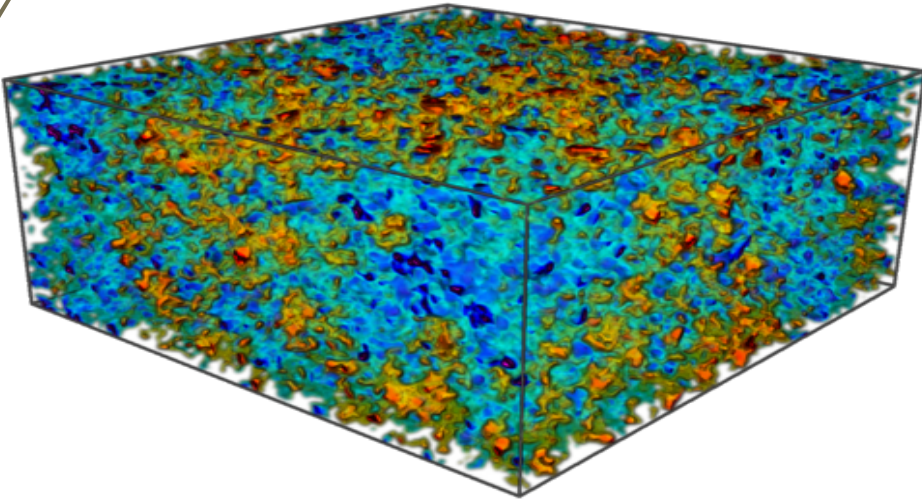


[Law and Kelton 2000]

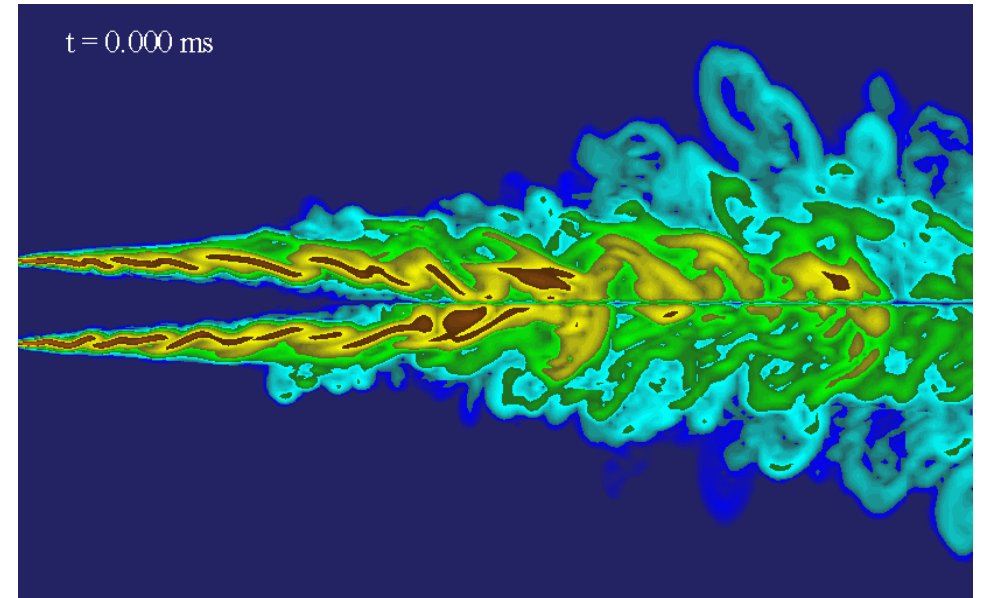
M&S Applications



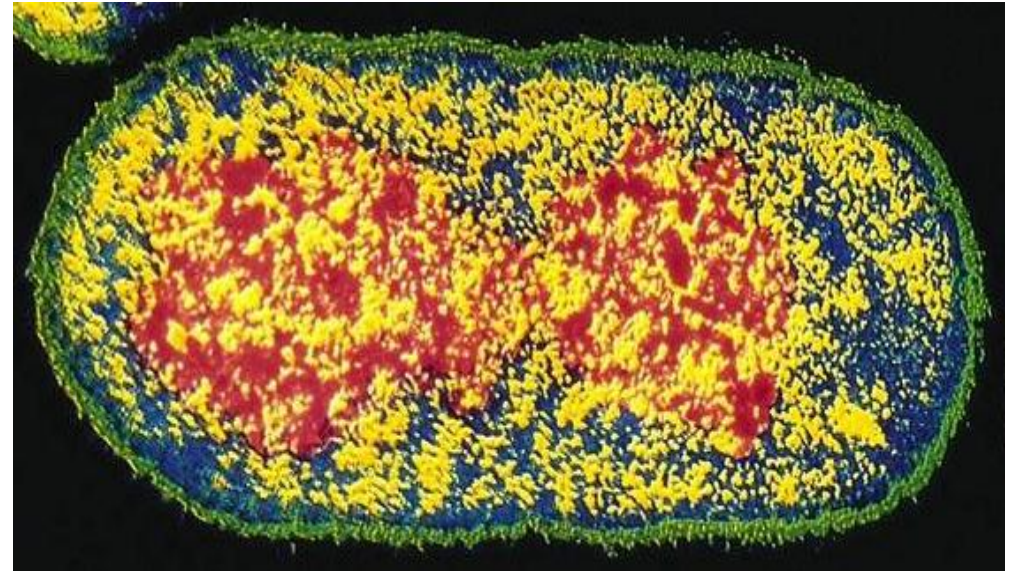
<http://www.gettyimages.com/>



<http://computation.llnl.gov/sites/default/files/public/dawnProgram-LLNL.png.png>

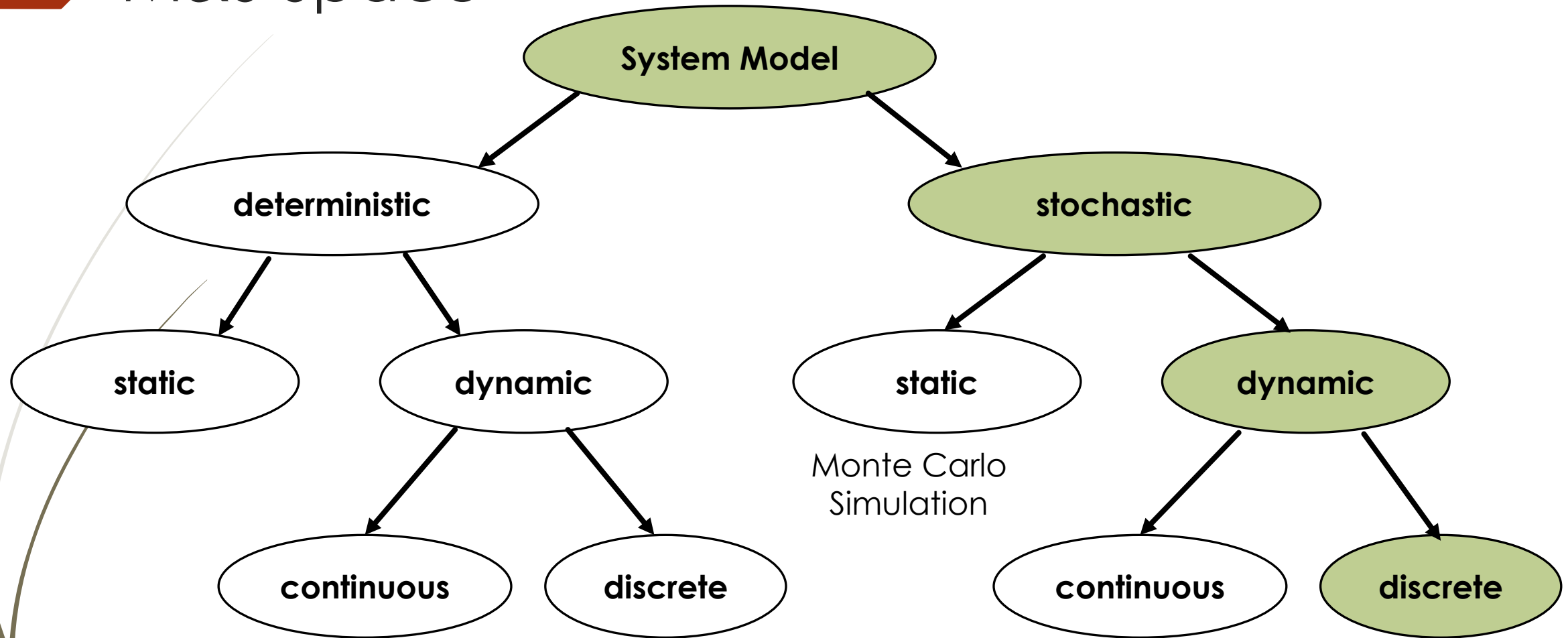


[http://www.yang.gatech.edu/projects/Liquid%20Rocket%20Injector%20Dynamics...%20\(Nan\)/index.htm](http://www.yang.gatech.edu/projects/Liquid%20Rocket%20Injector%20Dynamics...%20(Nan)/index.htm)



http://www.sandia.gov/~sjplimp/images/ecoli_expt.jpg

M&S Space

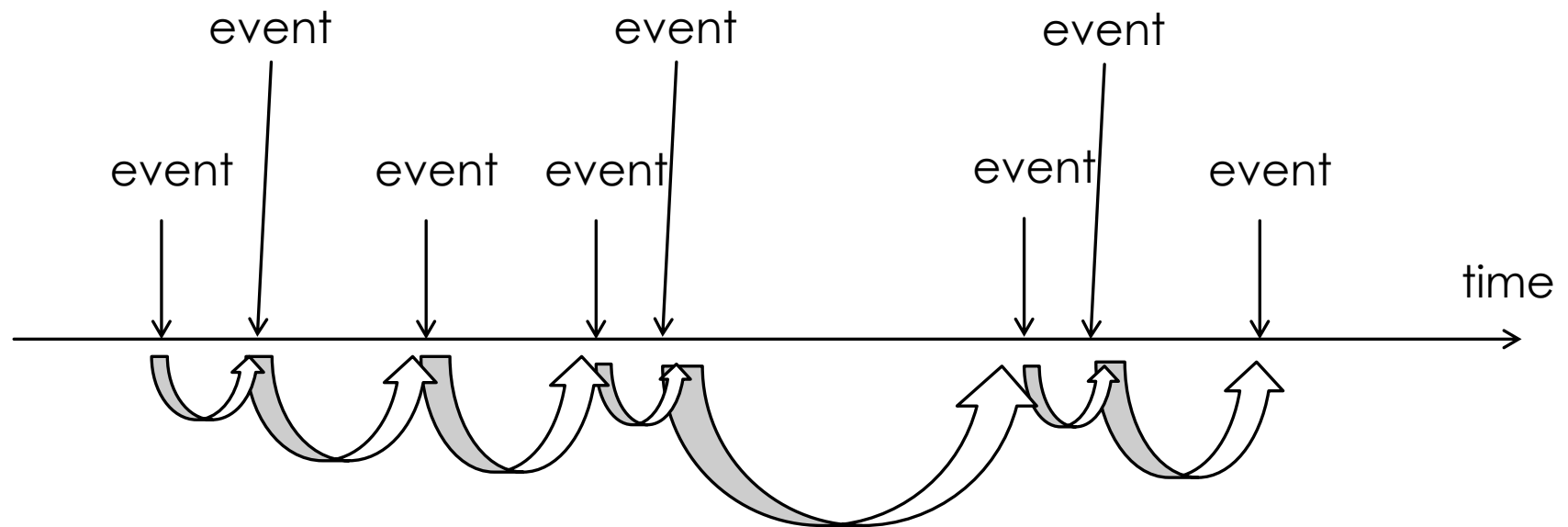


[Leemis and Park 2006]

Discrete-Event
Simulation (DES)

Discrete-Event Simulation

- Use computers to mimic operations and processes over **discrete** instances of time (events)



Components of a DES

- **The state of a system:** collection of variables necessary to describe the system at any time, relative to the objectives of the study, e.g., the number of jobs in queue.
- **An event:** an instantaneous occurrence that may change the system state, e.g. a new job arrival or service time completion
- **Future event list:** a priority queue (e.g., binary heap, or splay tree) to store the events
- **Simulation clock:** current simulation time (has nothing to do with the wall-clock time)

```
While(!eventlist.isEmpty()) {  
    Event e = eventlist.getMinEvent();  
    clock = e.time;  
    process(e) # which may generate more future events in eventlist  
}
```

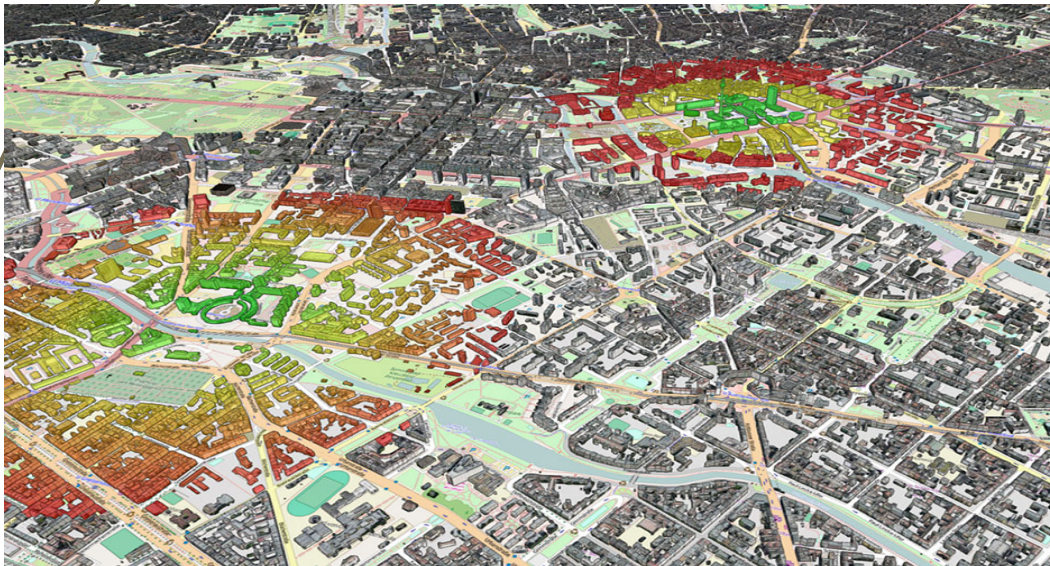
Parallel Simulation

- Divide simulation model into sub-models that can be distributed among different processors/cores
 - Each sub-model is called a **Logical Process (LP)**
- The global event-list is replaced by many event-lists, one for each LP
 - **No global simulation clock!!!**
- LPs communicate via explicit messages
 - No shared state variables

DES Applications



<http://corporate.comcast.com/news-information/news-feed/internet-of-things>



<https://www.3dcontentlogistics.com>



<https://digitalunite.com/guides/using-internet-0/connecting-internet/how-connect-internet>



<http://news.filehippo.com/wp-content/uploads/2013/06/dell-data-centre-1024x640.jpg>

Network Experiments

➤ Physical Testbeds



ESnet

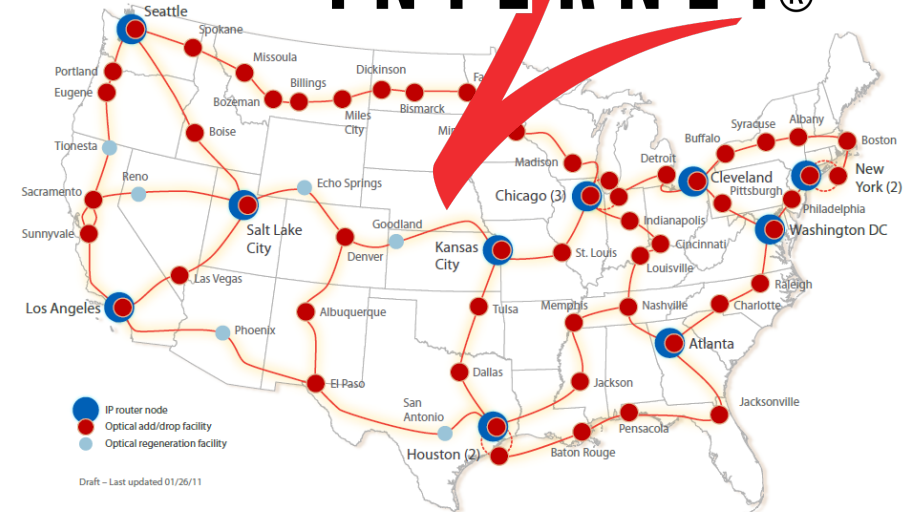
ENERGY SCIENCES NETWORK

ESnet—Connecting DOE Labs to the World of Science

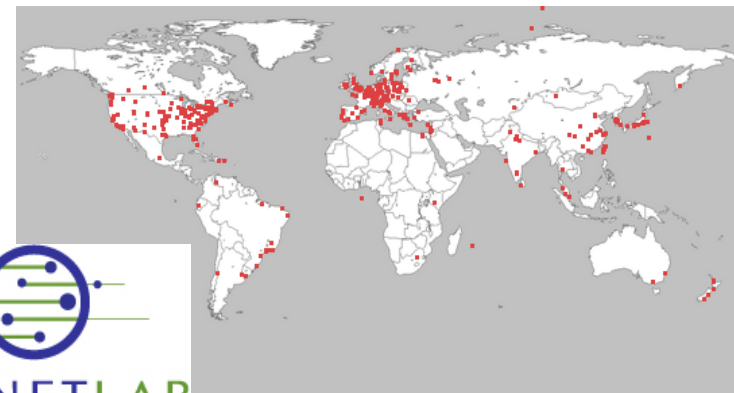


<http://www.es.net>

INTERNET®



<http://www.internet2.edu>



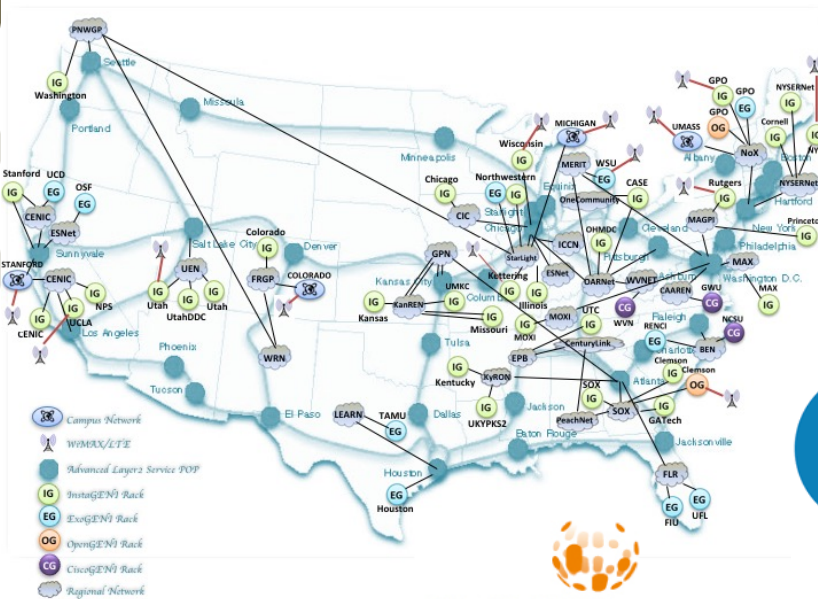
<http://www.planet-lab.org>

Network Experiments

- Physical Testbeds
 - Real
 - Small
 - Lack diversity

Network Experiments

- Physical Testbeds
- Emulation Testbeds



geni
Exploring Networks
of the Future

CloudLab
Chameleon



WINLAB



Network Experiments

- Physical Testbeds
- Emulation Testbeds
 - Real apps or real transmission
 - Some flexibility (from virtualization)
 - Limited scale
 - Limited scenarios

Network Experiments

- Physical Testbeds
- Emulation Testbeds
- Simulation Testbeds



GTNeTS
ROSSNet

SSFNet

And more ...

Network Experiments

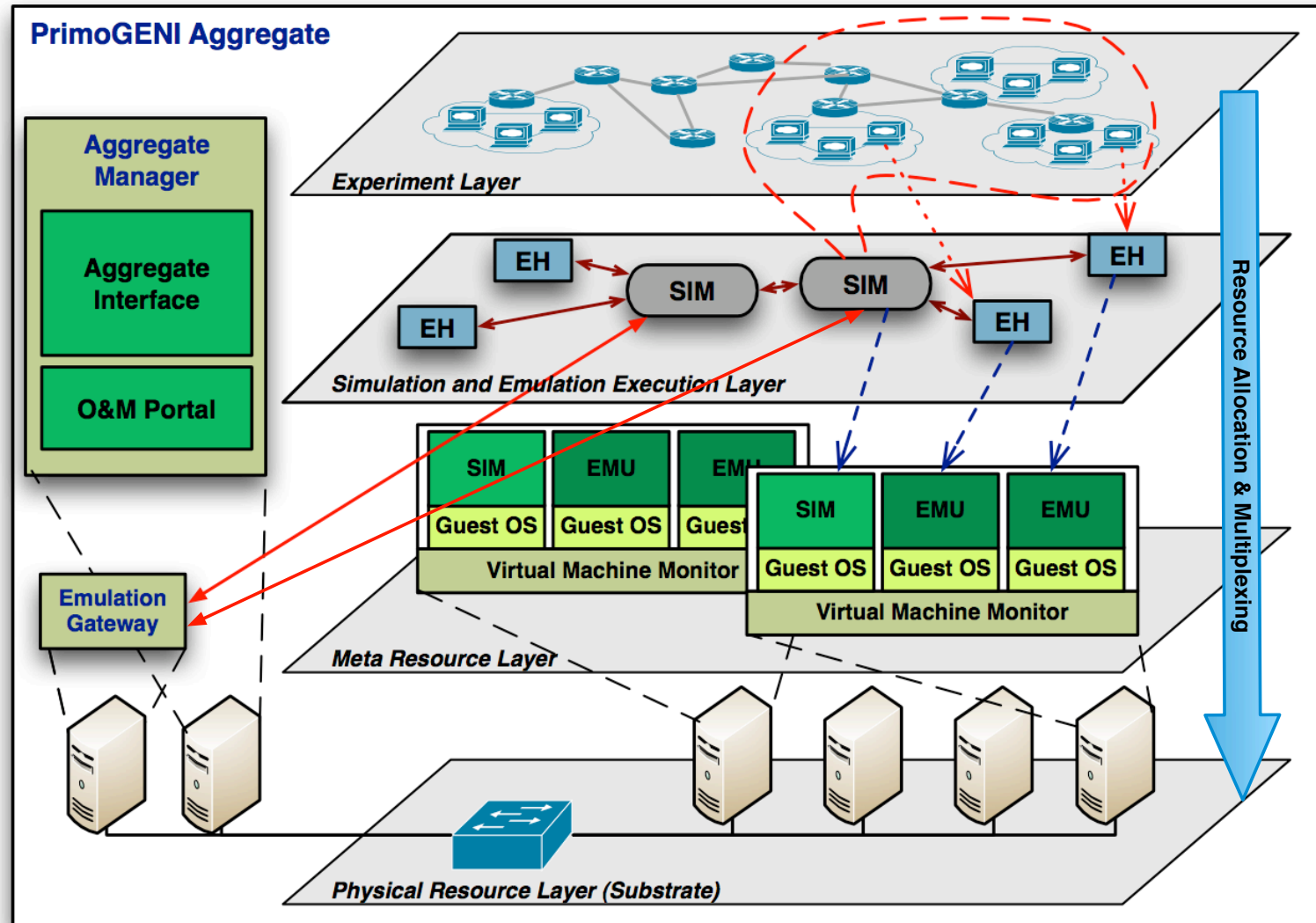
- Physical Testbeds
- Emulation Testbeds
- Simulation Testbeds
 - Flexible
 - Scalable
 - Diverse
 - Lack credibility
 - Significant model development effort

Prímení

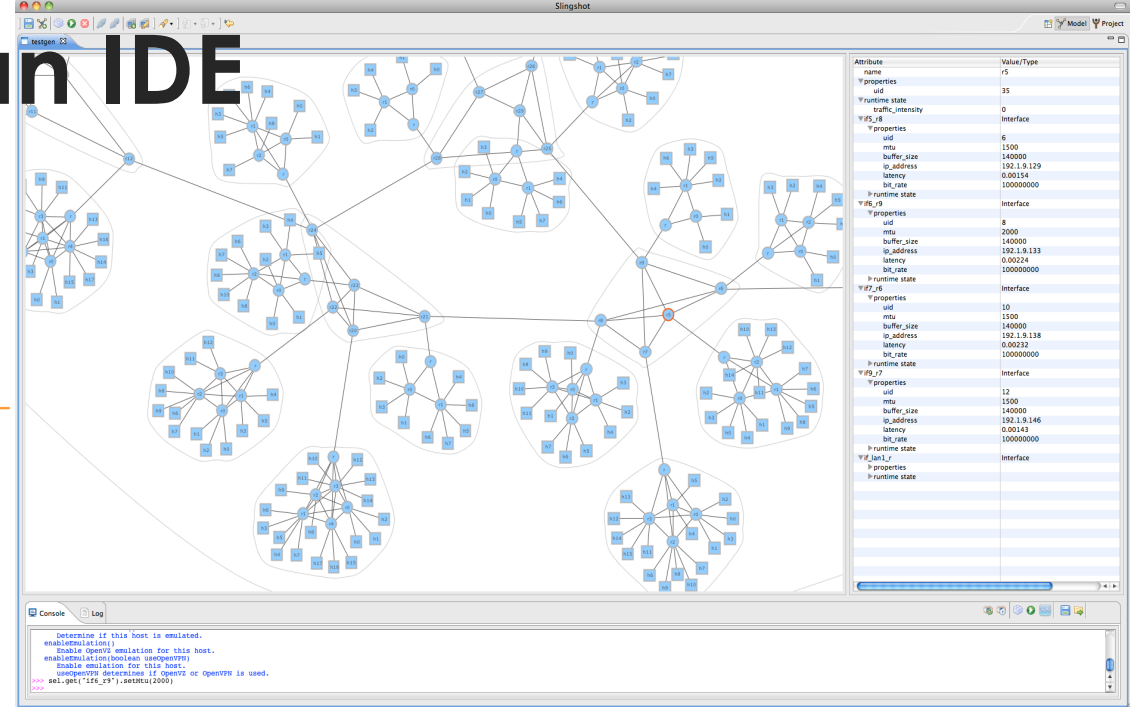
- Enables hybrid network experiments
- Include **simulated**, **emulated**, and **physical** components
- **Simulation**
 - Experimentation at scale
 - Modeling abstractions
 - Flexibility
- **Emulation**
 - Real applications
 - Resource multiplexing
- **Physical Networks**
 - Real traffic



PrimoGENI is a GENI Aggregate

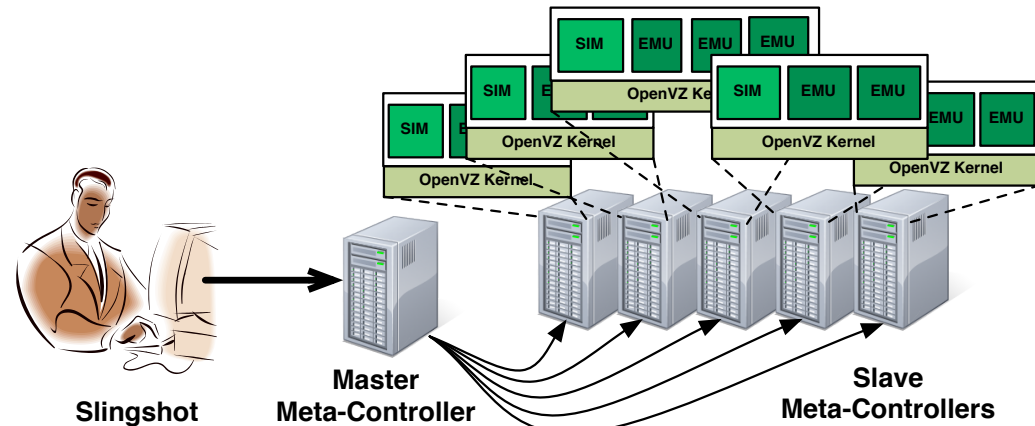


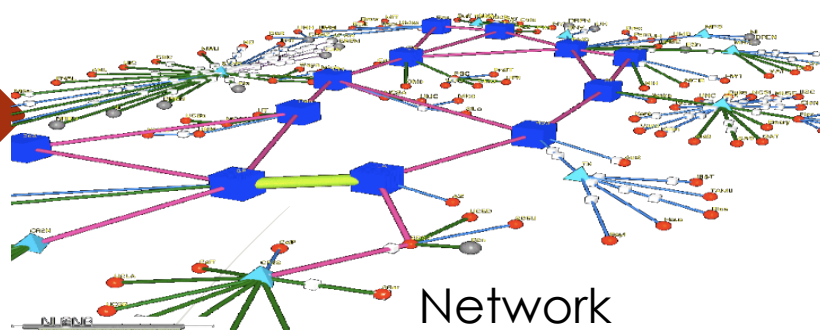
PrimoGENI has an IDE



➔ Manage the *life cycle* of network experiments

- Model Configuration
- Resource Specification
- Deployment, Execution
- Online Control & Monitoring
- Visualization
- Data Collection





Network
Visualization



Interactive
Console

Java
Model

Python
Model

XML
Model

In-Memory
Model
(JAVA)

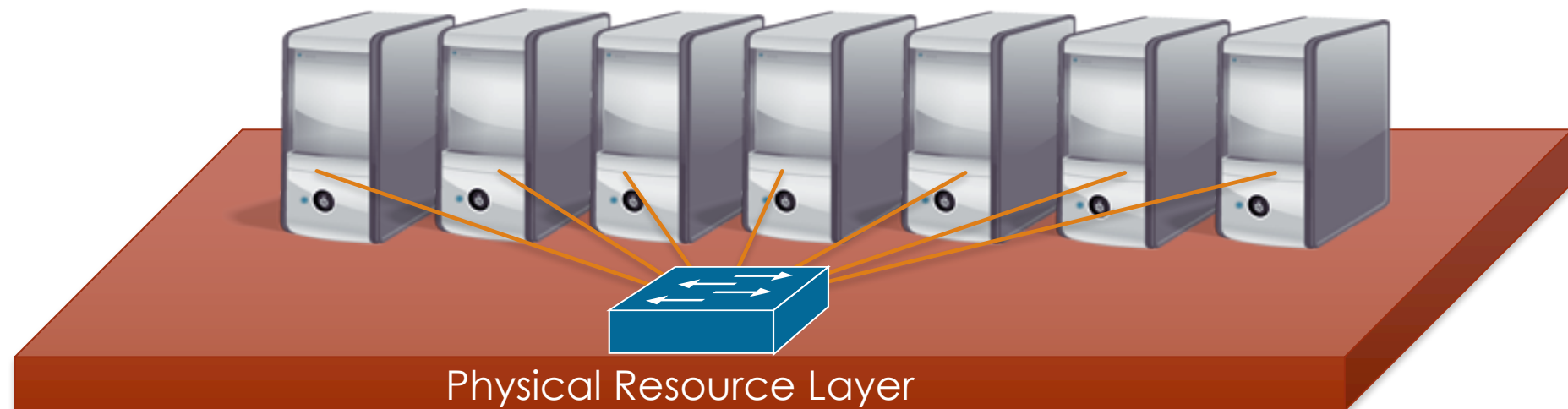
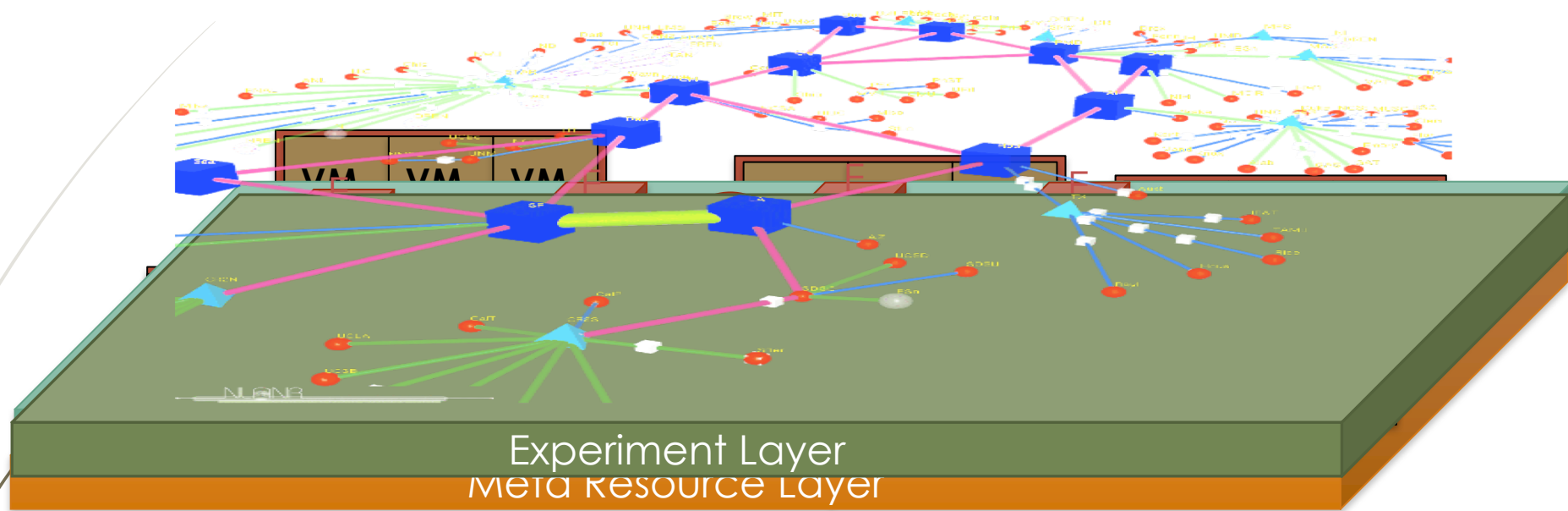
Model
Compilation

Model
Partitioning

Execution
Model

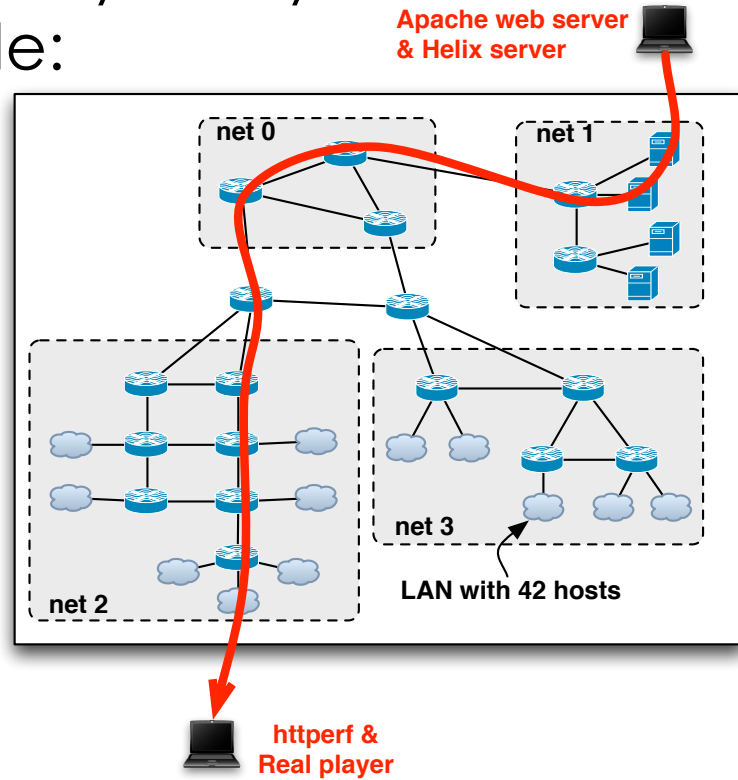
DBMS





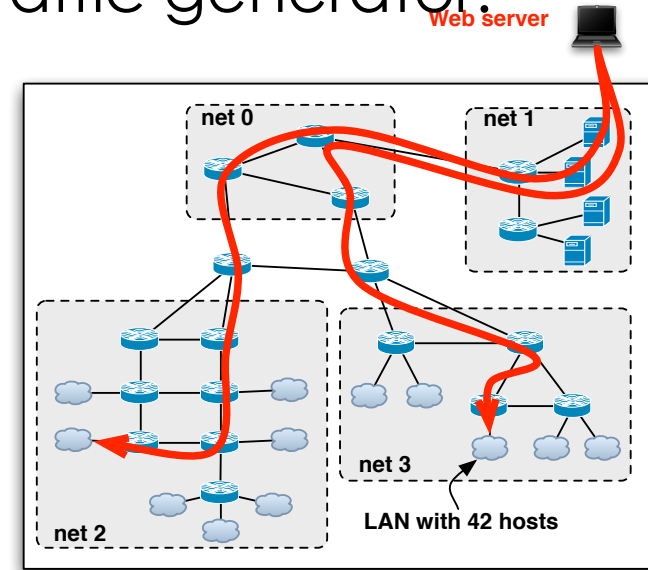
What's the Use?

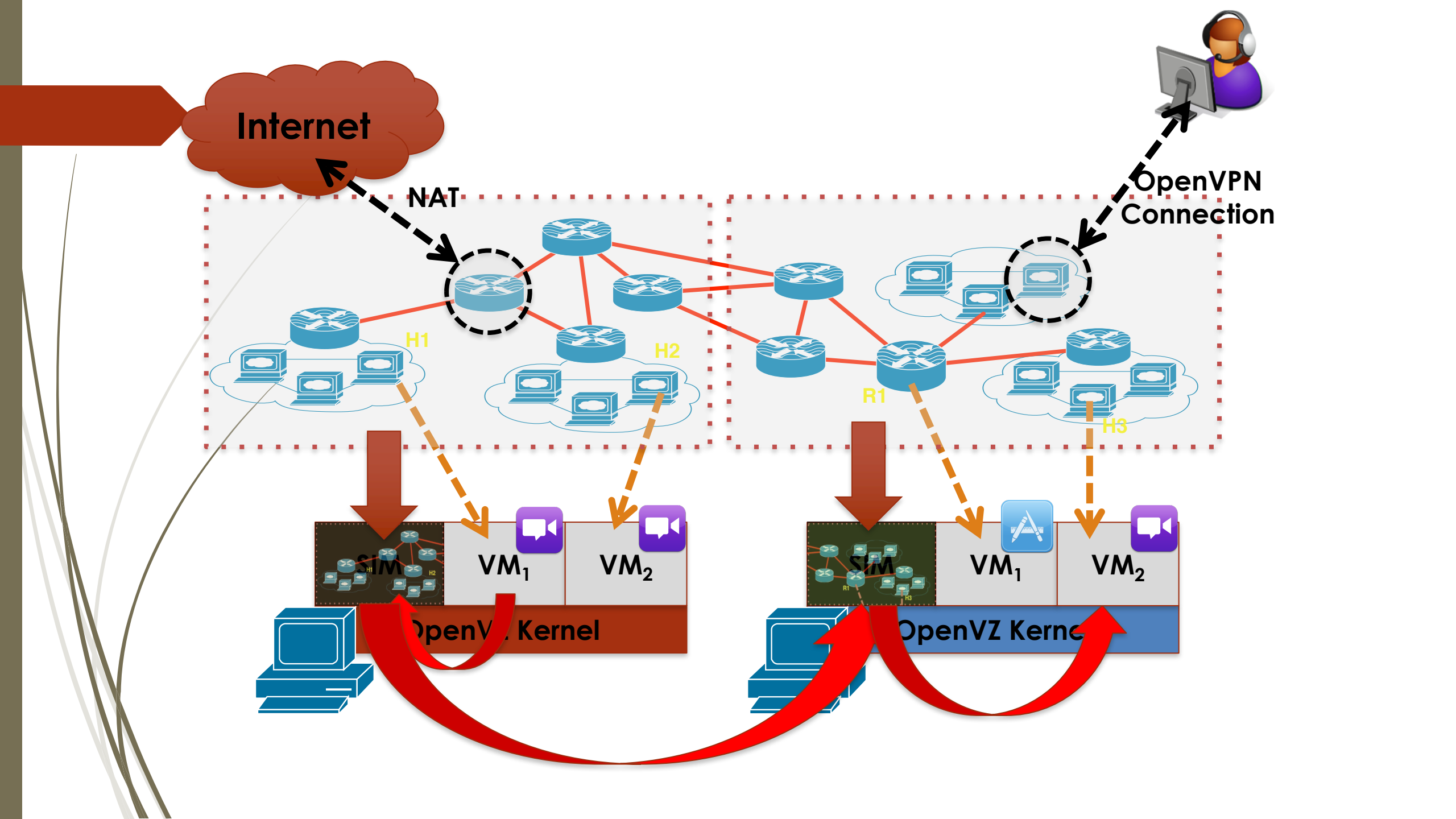
A fancy delay node:



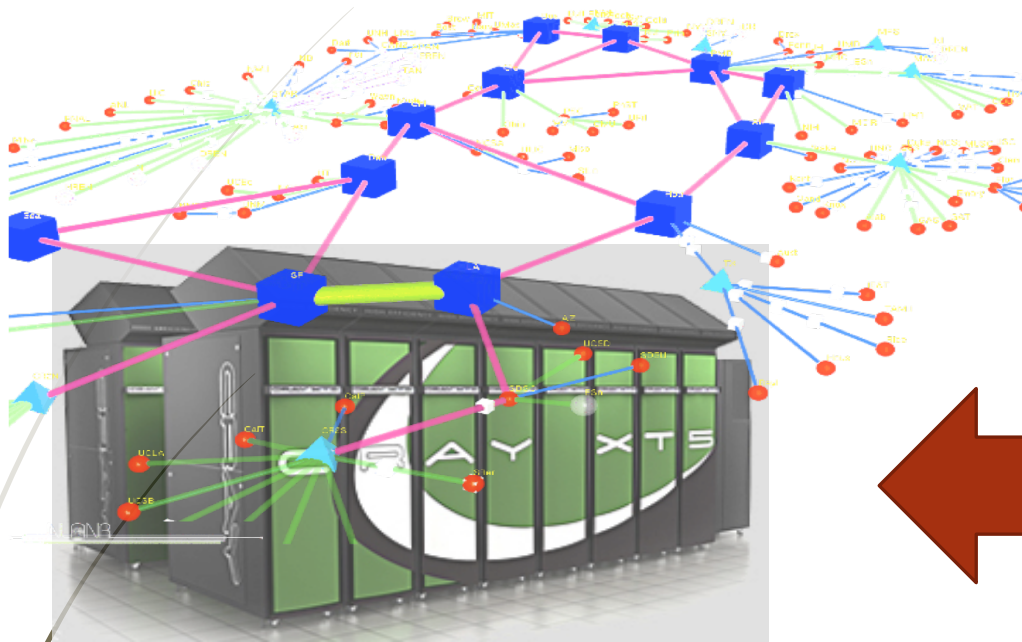
A virtual distributed environment:

A traffic generator:





Problem: Scale can still be limited due to intense interactions with the simulated component.



Network Simulation

- Full-scale network model
- Detailed topology and protocols

**Capture large-scale network behavior
and global traffic conditions**



Real Time Traffic



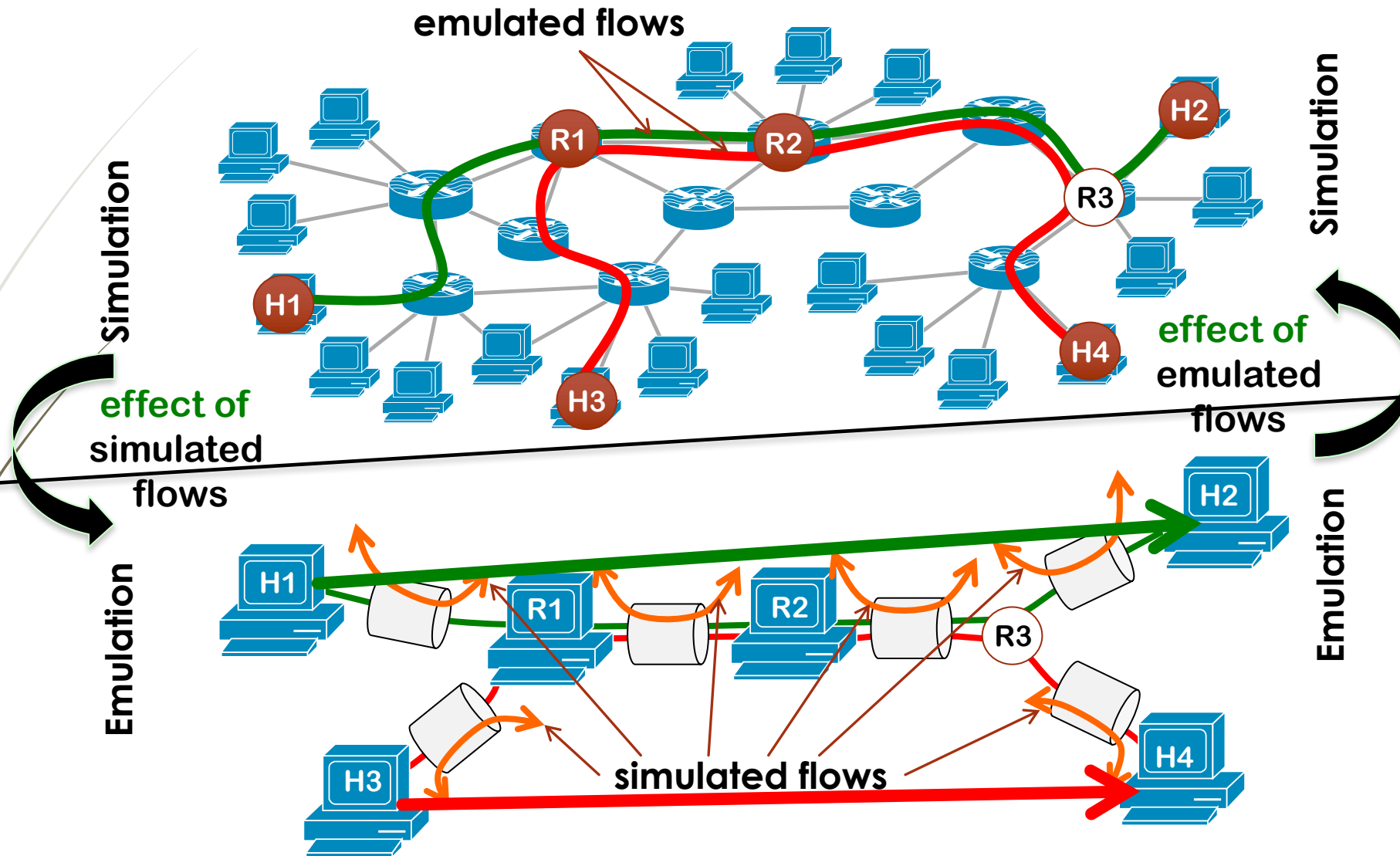
Network Emulation

- Real execution environment (operating system, network stack, software tool chain)
- Unmodified applications

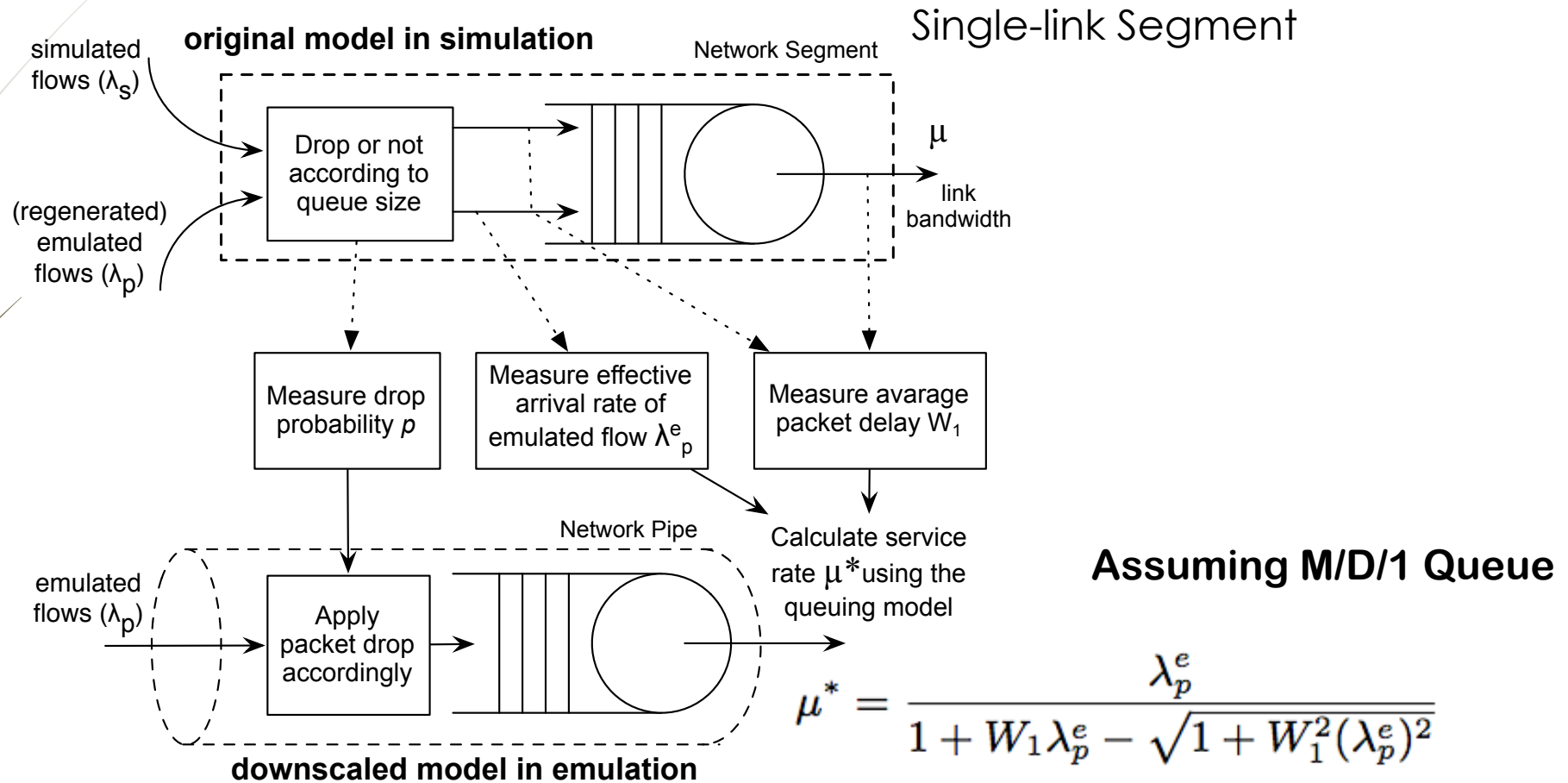
**Represent real application
traffic behavior**

Our Symbiotic Approach in a Nutshell

29



Steady-State Queuing Model



Does It Work?

- Well, steady-state does **not** work!
- Closed-form solution for transient effect is rather elusive, even for Poisson arrivals
- We invent a “control knob” to dynamically adjust μ^* from measurements

Adjust for Transient Effect

excess queue length avg pkt delay in emulation avg pkt delay in simulation

$$\Delta L(t) = \mu^*(t)(W_2(t) - W_1(t))$$

excess service rate

$$\Delta \mu^*(t) = \frac{\Delta L(t)}{\Delta T} = \frac{\mu^*(t)(W_2(t) - W_1(t))}{\Delta T}$$

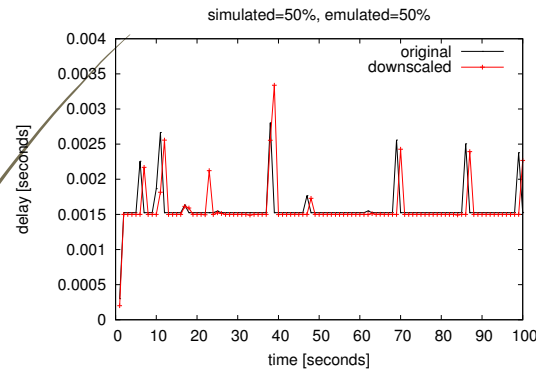
$$\begin{aligned} \hat{\mu}(t) &= \mu^*(t) + \Delta \mu^*(t) \\ &= \frac{\lambda_p^e(\Delta T + W_2(t) - W_1(t))}{\Delta T(1 + W_1(t)\lambda_p^e(t) - \sqrt{1 + W_1^2(t)\lambda_p^e(t)^2})} \end{aligned}$$

- The adjustment effectually forces the emulation system to “track” the simulated network conditions at each update interval.
- The result can be extended to network segments with multiple links.

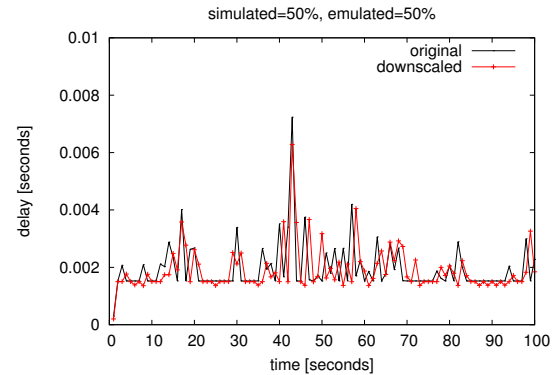
Validation Results

- Low, medium, heavy traffic conditions
- Various proportions of emulated vs. simulated traffic
- Mixed arrivals: exponential, constant, real traces

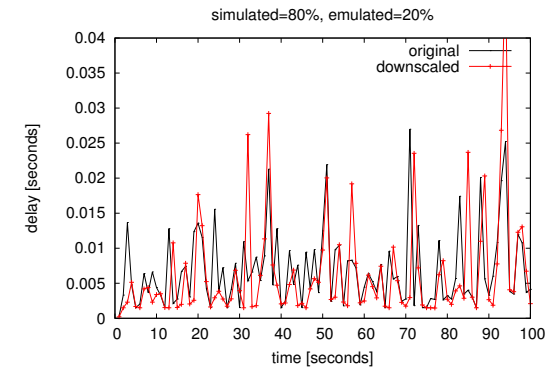
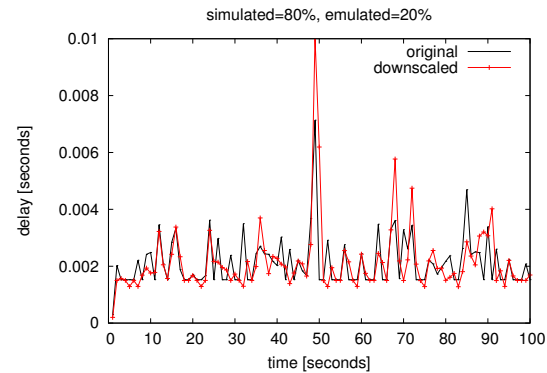
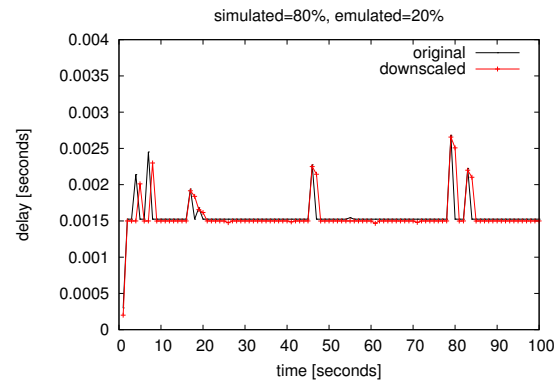
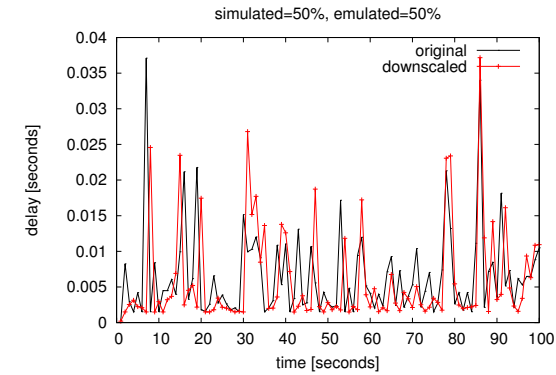
10% utilization



50% utilization

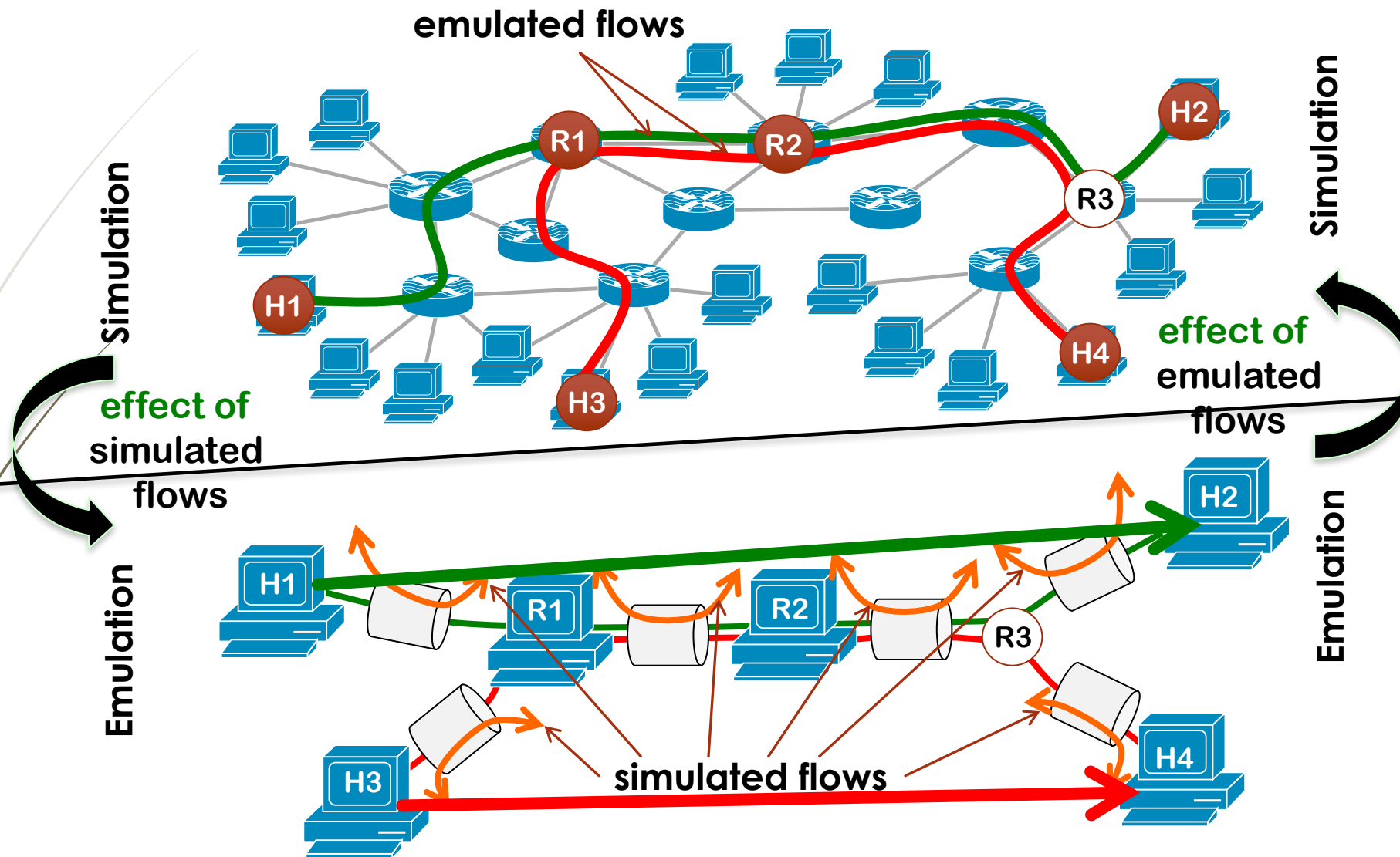


90% utilization



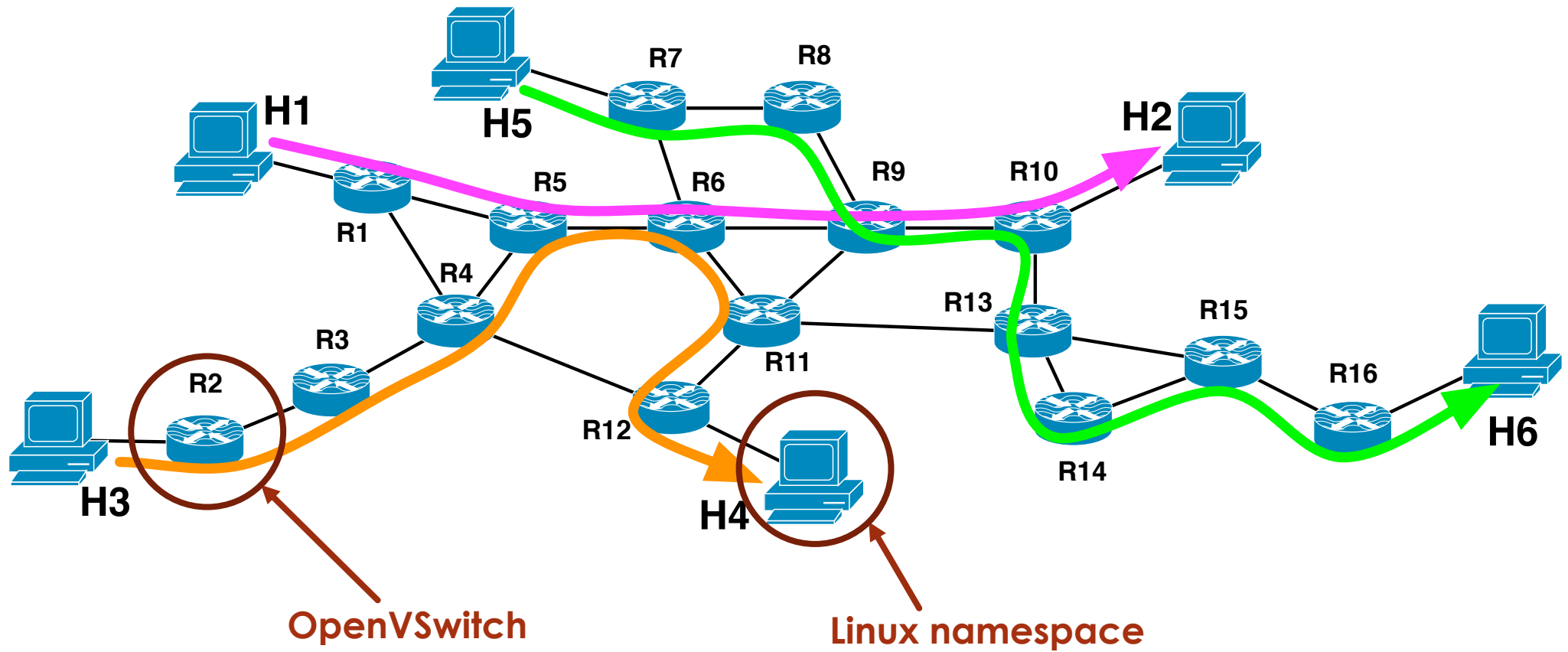
Our Symbiotic Approach in a Nutshell

34

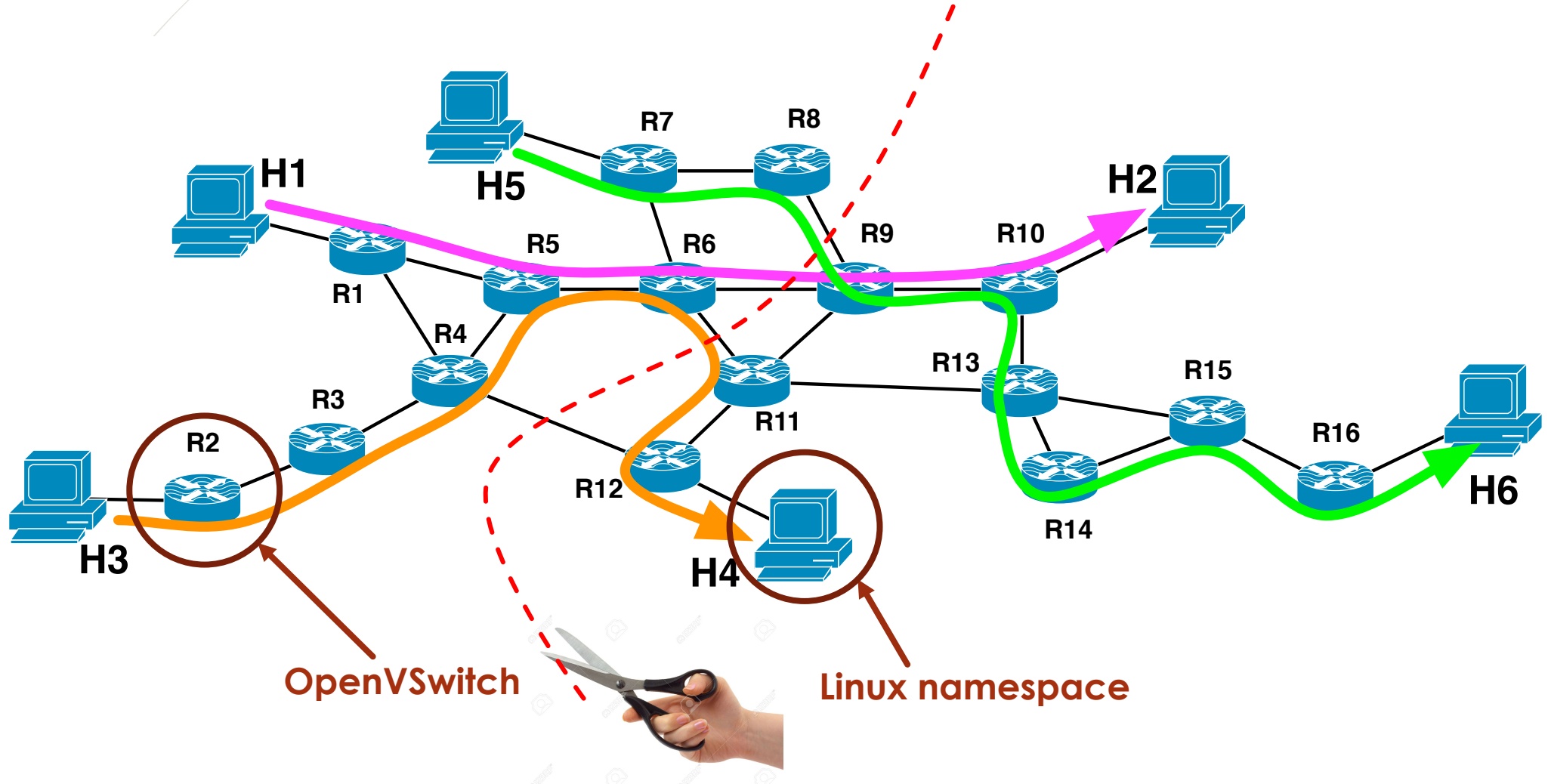


Symbiosis works, but can we extend the method further?

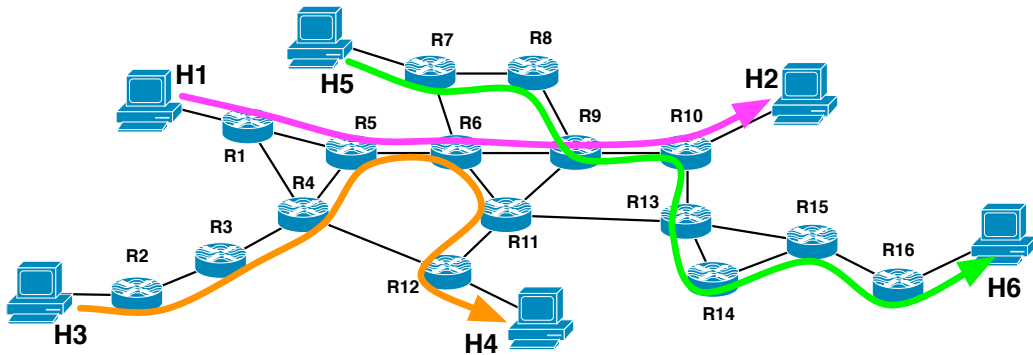
Mininet is an emulation testbed for OpenFlow



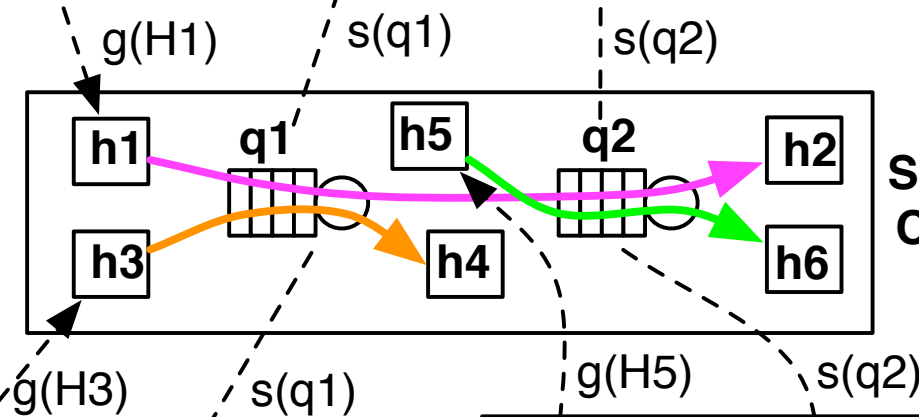
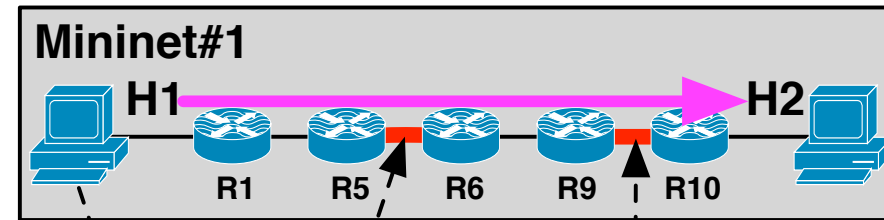
Distributed Mininet



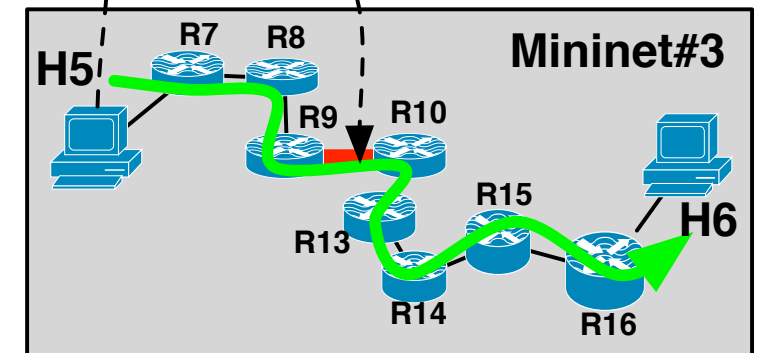
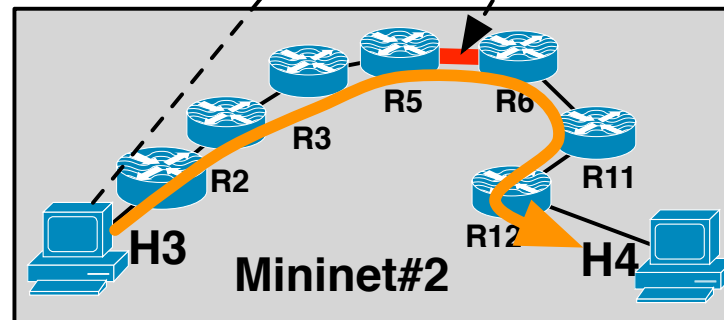
Distributed Mininet with Symbiosis



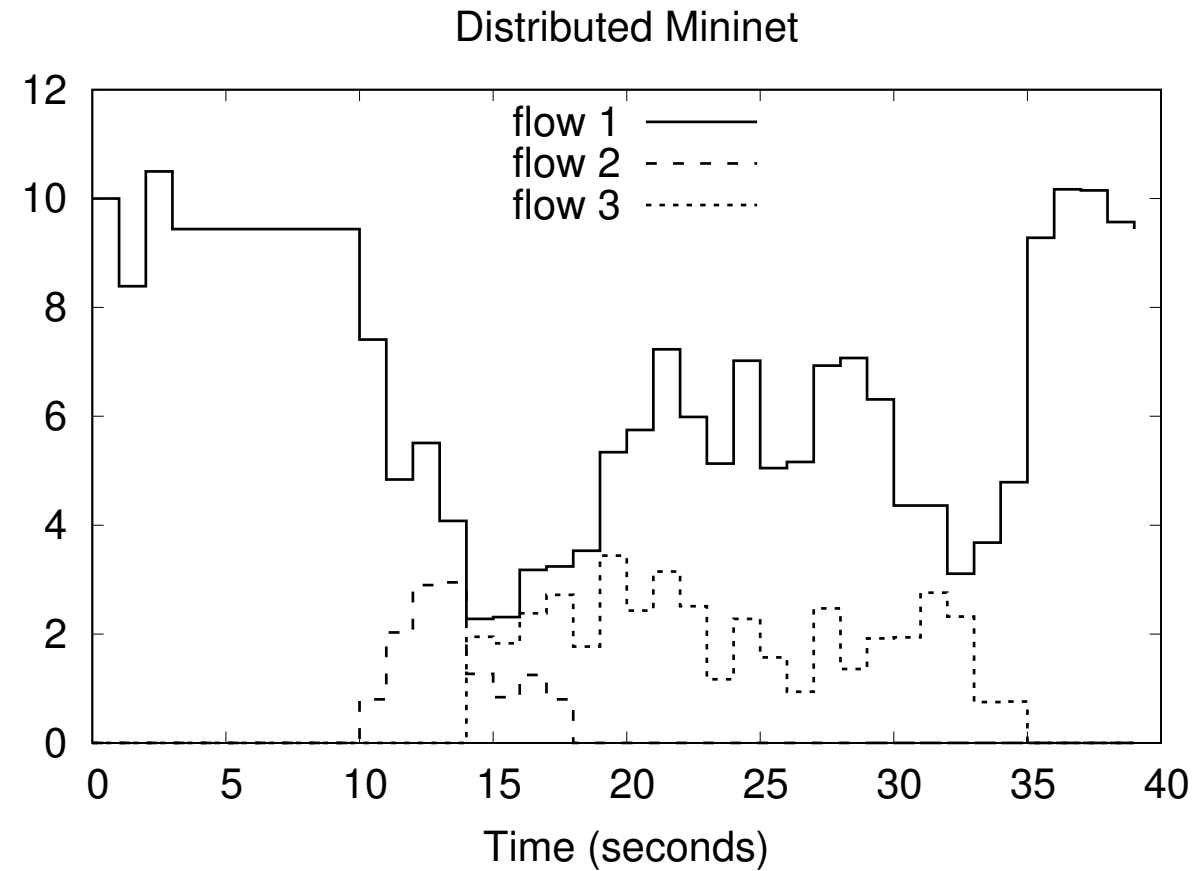
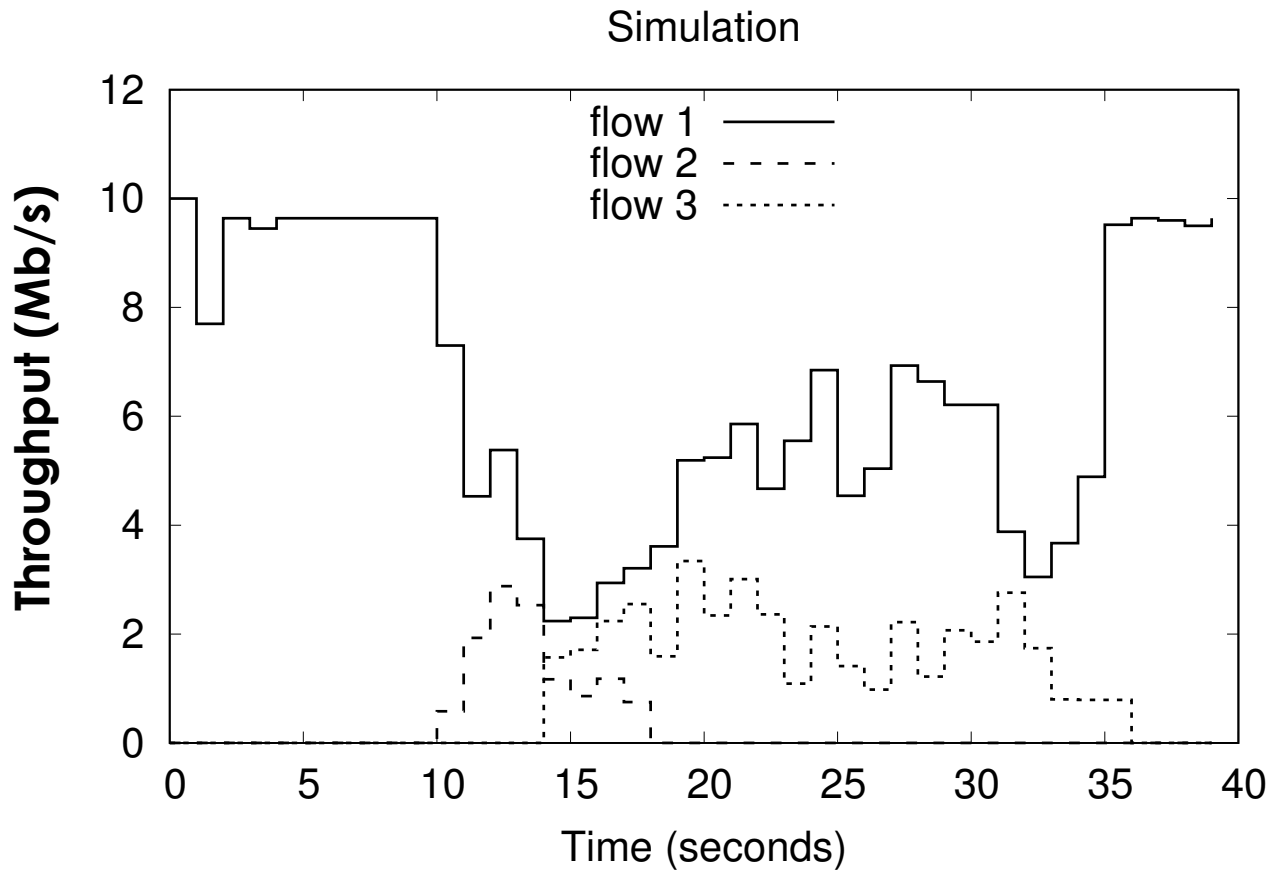
$$\mu_i = \frac{\lambda_i(\Delta T + \hat{w}_i - w_i)}{\Delta T \left(1 + w_i \lambda_i - \sqrt{1 + w_i^2 \lambda_i^2}\right)}$$



Simulation
Controller



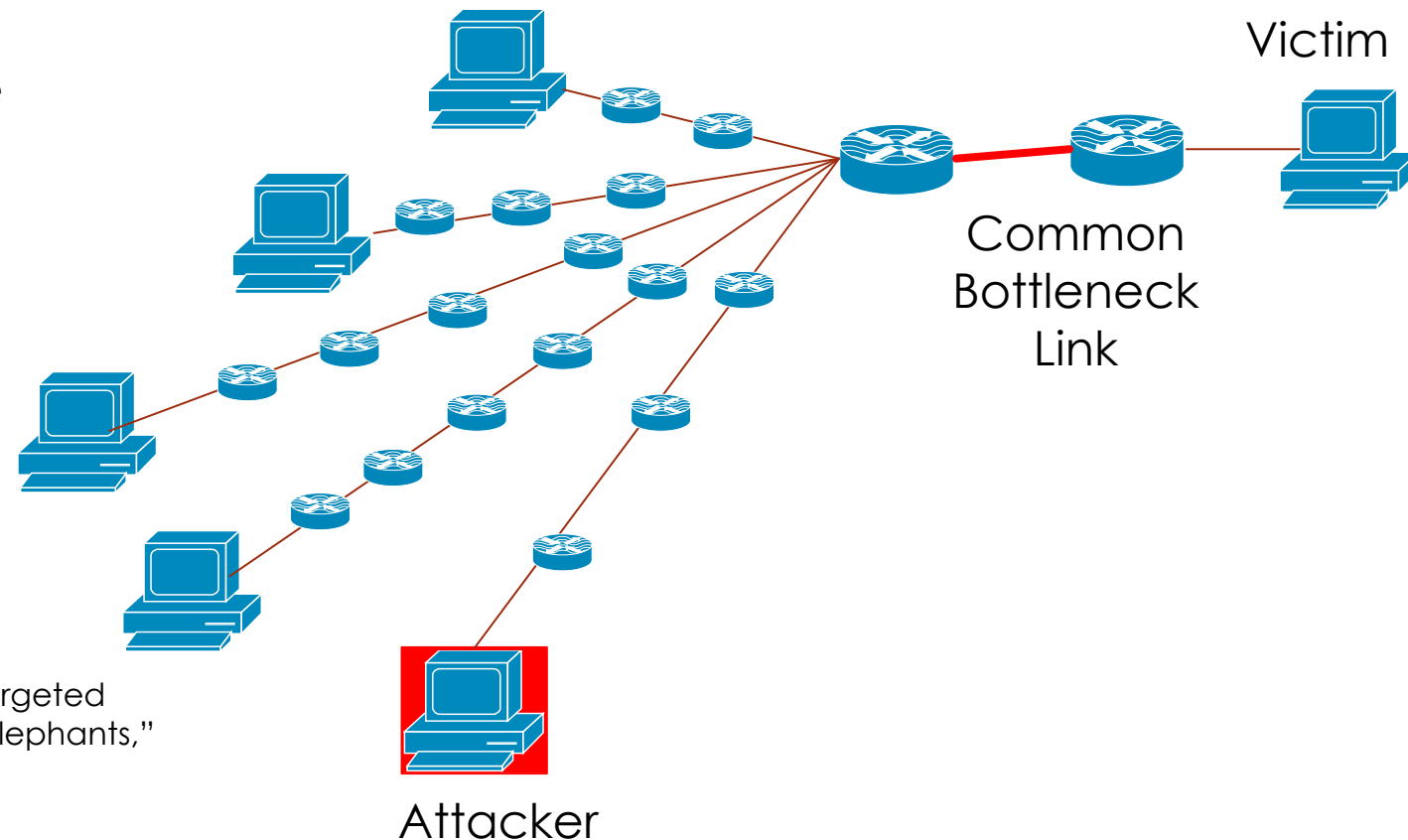
Validation Experiment



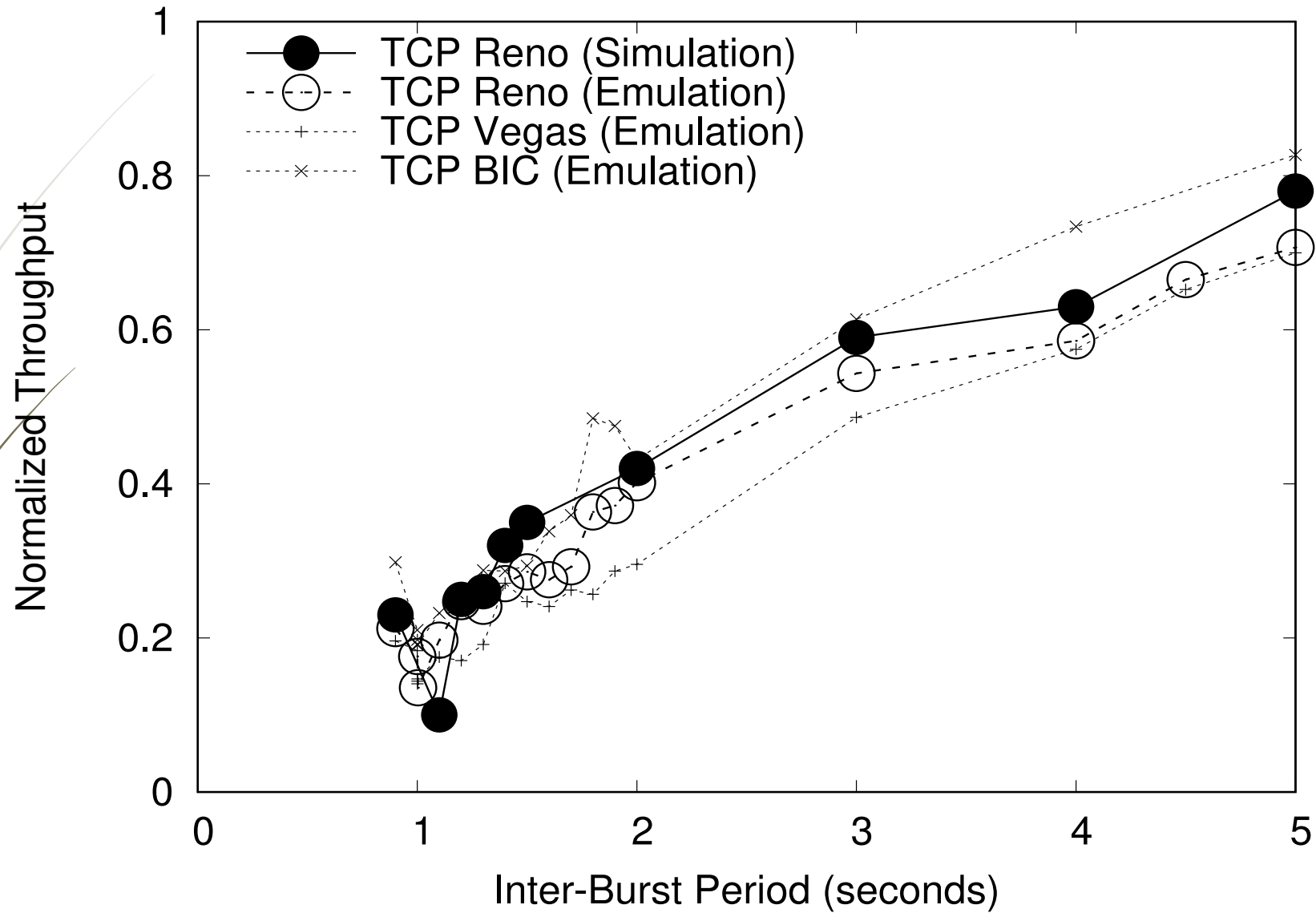
A Case Study: Shrew Attack

- Attacker sends bursts of data at a regular interval to an over-committed bottleneck link^[1]
- When burst intervals synchronize with RTO of TCP connections sharing the link, they can trigger TCP timeouts and consequently strangle throughput
- Difficult to detect since avg. traffic rate of the attack is low

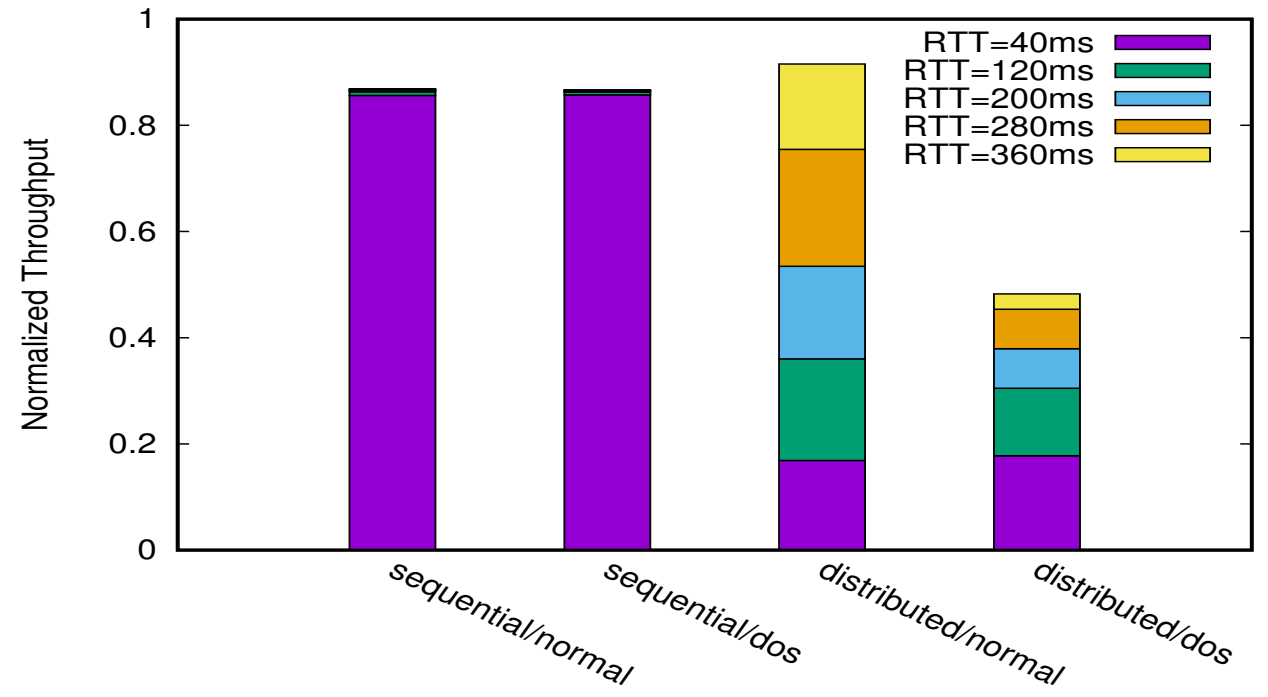
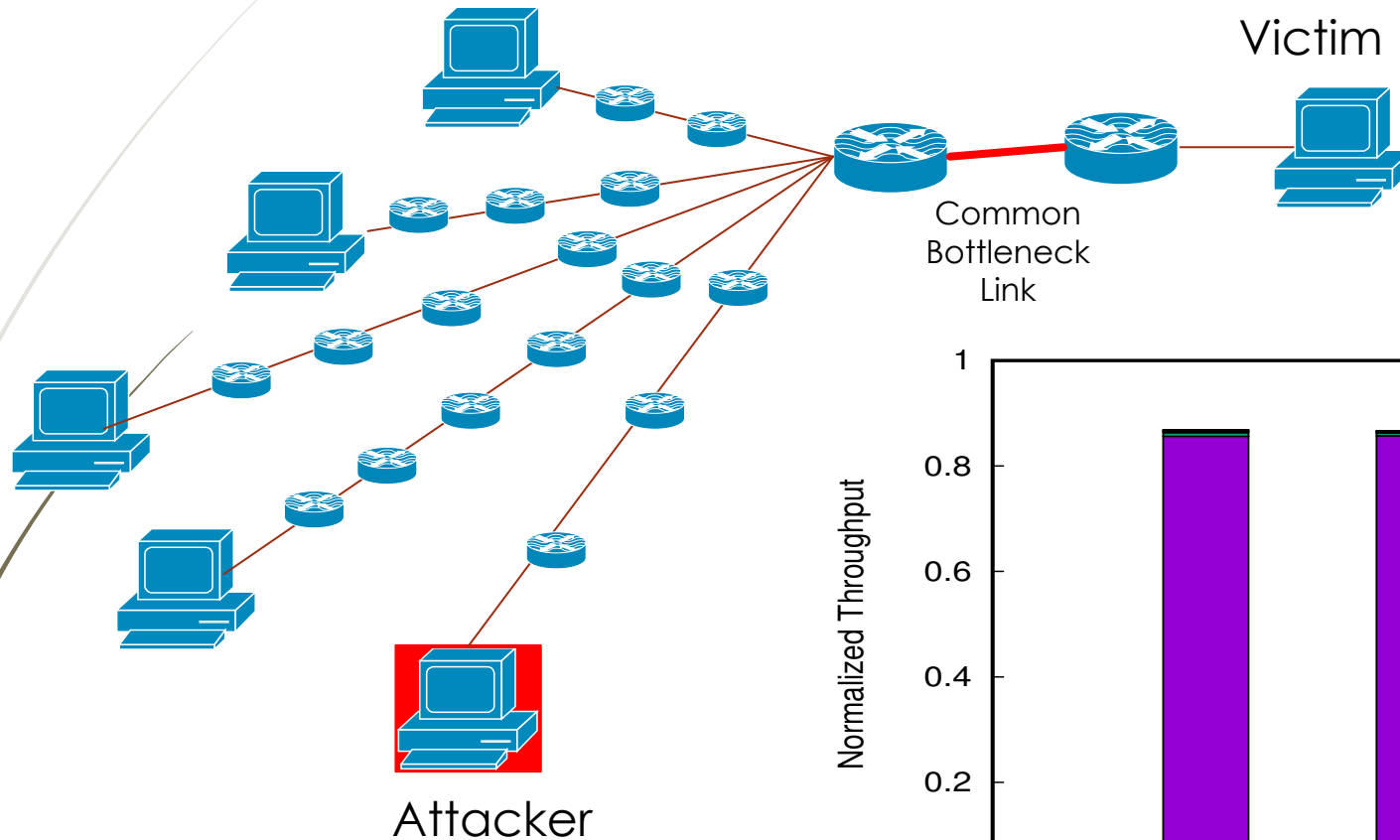
[1] A. Kuzmanovic and E. W. Knightly, "Low-rate TCP-targeted denial of service attacks: The shrew vs. the mice and elephants," SIGCOMM 2003, pp. 75–86.



One "Good" Flow over 10 Mb/s Bottleneck Link



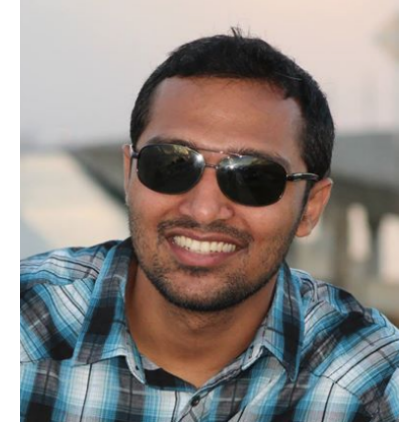
Experiment Results



Conclusions

- Physical or emulation experiment is limited by scale
 - Simulation can scale better
- Network simulation is unreal
 - Combine simulation and emulation
- Real-time simulation is limited in capacity
 - Interaction in-between is bottleneck
- Symbiosis creates an ideal separation between the two
- Symbiosis method has its limitations
 - Cross system traffic is not allowed
 - Need to combine with real-time method
- This is a GENI talk:
 - The potential of GENI is still not fully tapped
 - Experiment design and tools are still being made
 - Look for convergence -> KC's talk next!

THANK YOU!



- **Leveraging symbiotic relationship between simulation and emulation for scalable network experimentation**, M. Erazo and J. Liu. SIGSIM-PADS, 2013.
- **Symbiotic network simulation and emulation**, M. Erazo, R Rong, J Liu. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 26 (1), 25 pages, 2015
- **Toward scalable emulation of future internet applications with simulation symbiosis**. J. Liu, C. Marcondes, M. Ahmed, R. Rong. *IEEE/ACM 19th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*, pages 68-77, 2015.
- **Distributed Mininet with symbiosis**. R. Rong and J. Liu. *ICC'17*. To appear.

