www.german-lab.de

Integrated Communication Systems

G-Lab Experimental Facility

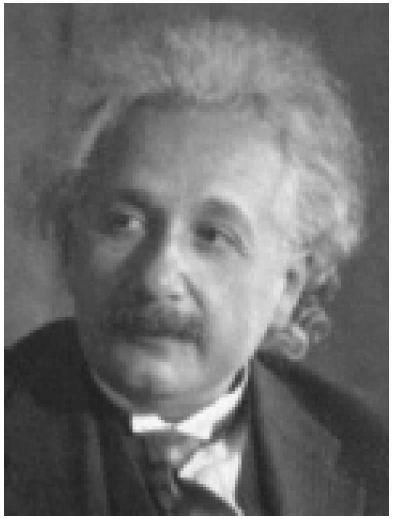
5th DFN-Forum Kommunikationstechnologien
– Verteilte Systeme im Wissenschaftsbereich –
21. – 22. Mai Regensburg

Paul Mueller / Dennis Schwerdel
Integrated Communication Systems Lab
Dept. of Computer Science
University of Kaiserslautern
Paul Ehrlich Bld. 34, D-67663 Kaiserslautern, Germany
Tel.+49 631 205 2263, Fax. +49 631 205 3056
www.ICSY.de

SPONSORED BY THE
Federal Ministry of Education and Research

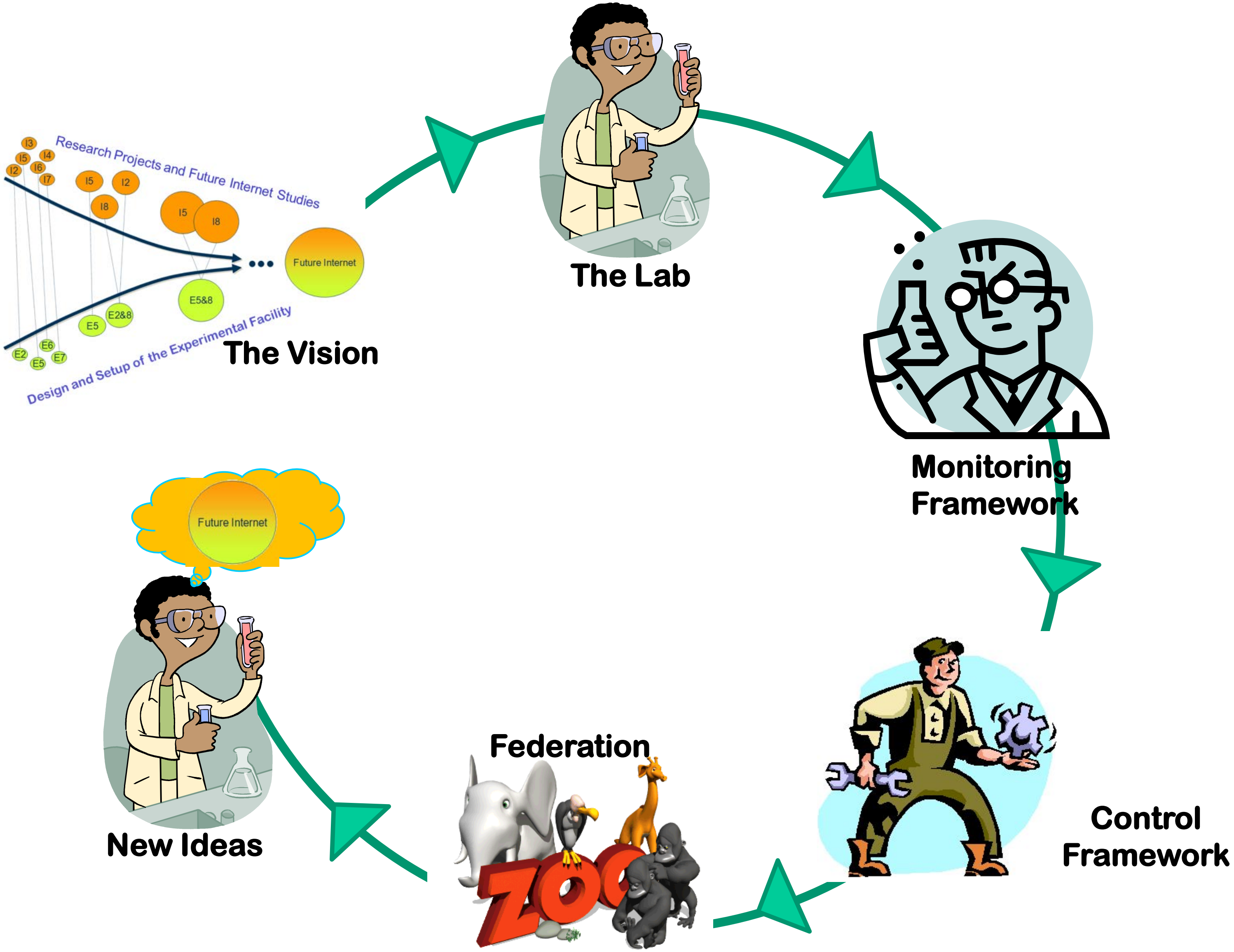
The challenge …

„Probleme kann man niemals mit derselben Denkweise lösen, durch die sie entstanden sind.“



“We can't solve problems by using the same kind of thinking we used when we created them.”

Content



Research Projects and Future Internet Studies
I3
I4
I5
I6
I2
I7
I5
I2
I8
I5
I8
Future Internet
E5&8
E2&8
E5
E6
E2
E5
E7
Design and Setup of the Experimental Facility
Future Internet
ZOO

The Vision

The Lab

Monitoring Framework

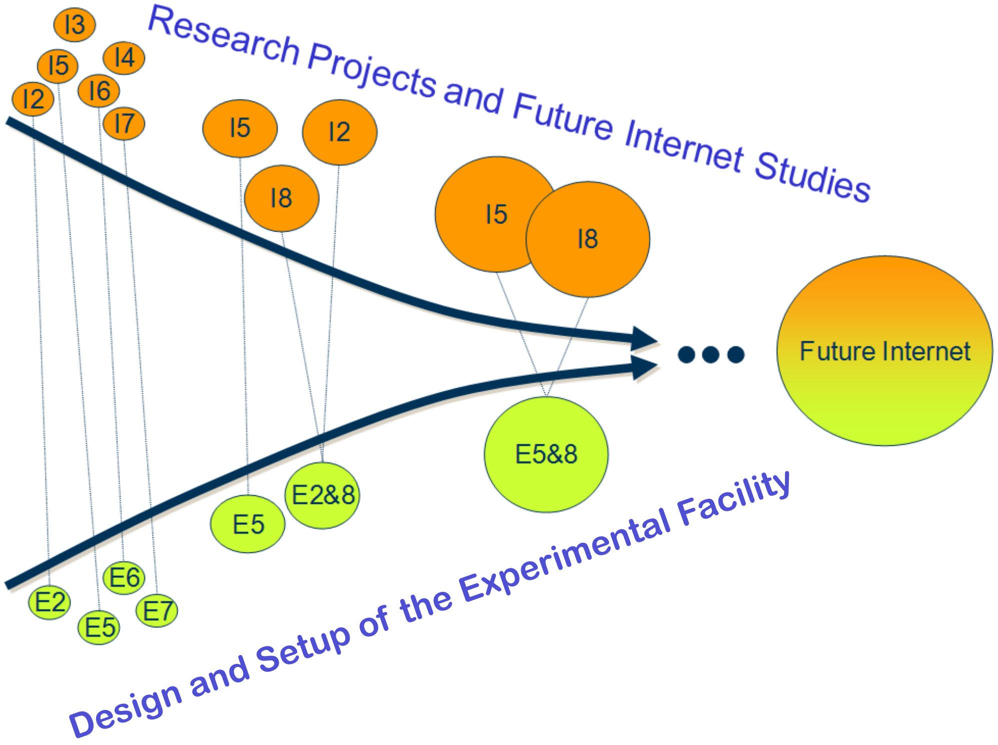
Control Framework

Federation

New Ideas

G-Lab: Vision of the Future Internet

- Closing the loop between ***research*** and ***real-world experiments***
- Provide an **experimental facility** for studies on architectures, mechanisms, protocols and applications towards Future Internet
- Investigate interdependency of theoretical studies and **prototype development**


Research Projects and Future Internet Studies
I3
I4
I5
I6
I2
I7
I5
I2
I8
I5
I8
Future Internet
E5&8
E2&8
E5
E6
E2
E5
E7
Design and Setup of the Experimental Facility

G-Lab Environment

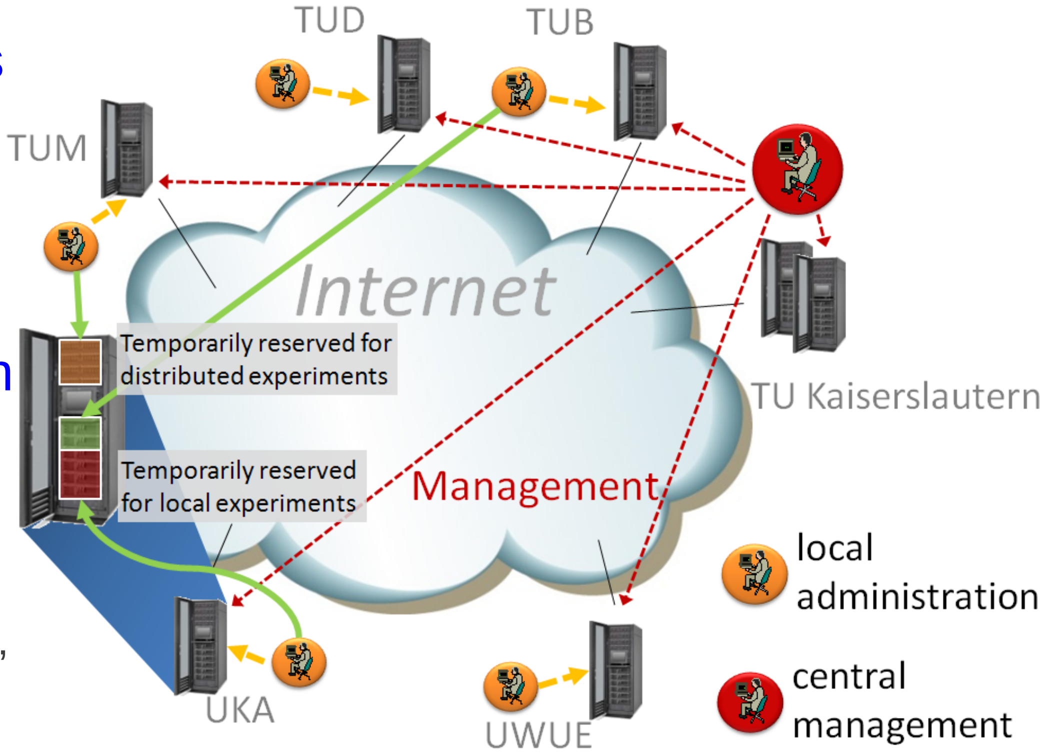
- **Testbed:**
  - Real not simulated
  - Specific purpose
  - Focused goal
  - Known success criteria
  - Limited scale

Not sufficient for clean slate design

- **Experimental facility:**
  - Purpose:
    - **explore yet unknown architectures**
    - **expose researchers to real thing**
    - **breakable infrastructure**
  - Larger scale (global?)
  - Success criteria: unknown
  - **Closed User Group**

The Lab

- Full control over the resources
  - Reservation of single resource should be possible
  - Elimination of side effects
  - Testing scalability
- Exclusive resource reservation
  - Testing QoS / QoE
  - Decentralized Resources can be independently used
  - Tests on the lower layers of the network without affecting the “life” network
- Extended functionality
  - New technologies (Wireless, Sensor,…)
  - Interfaces to other testbeds (GENI, PlanetLab Japan, WinLab, …)
  - OpenFlow setup


TUD
TUB
TUM
Internet
Temporarily reserved for distributed experiments
Temporarily reserved for local experiments
Management
TU Kaiserslautern
UKA
UWUE
local administration
central management

TUB	TU Berlin
TUD	TU Darmstadt
TUKL	TU Kaiserslautern
TUM	TU München
UKA	University Karlsruhe KIT
UWUE	University Wurzburg

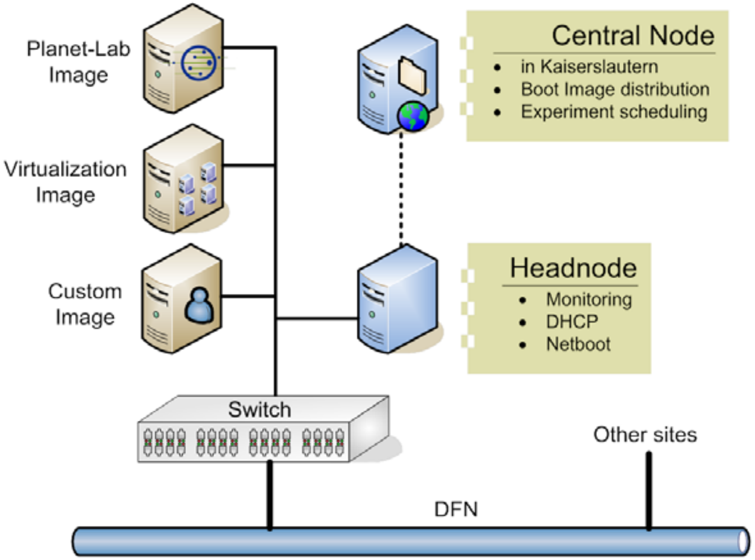
Hardware Equipment

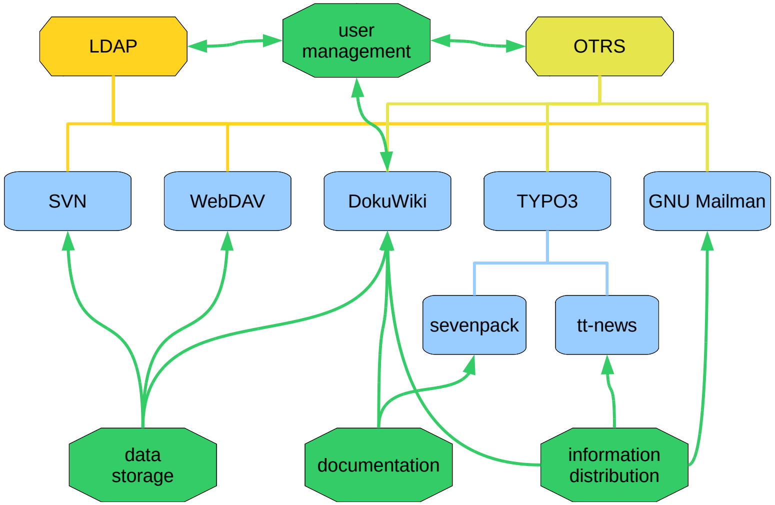
- Normal Node
  - 2x Intel L5420 Quad Core 2,5 GHz
  - 16 GB Ram, 4x 146 GB disk
  - 4x Gbit-LAN
  - ILOM Management Interface (separate LAN)
- Network Node
  - 4x extra Gbit-Lan
- Headnode
  - 2x Intel E5450 Quad Core 3,0 GHz
  - 12x 146 GB disk
- Switch Fabric CISCO 45xx
- Site requirements
  - 1 public IP address per Node
    - IPv4 and/or IPv6 addresses.
    - Virtualized nodes need additional addresses
  - Direct Internet access
    - No firewall or NAT
    - Nodes must be able to use public services (NTP, public software repositories)
  - Dedicated Links
    - dark fiber, λ wavelength, MPLS

- 174 Nodes in total (1392 cores total)

Site	Head	Network	Normal	
Kaiserslautern	1	2	47+9	Phase I
Würzburg	1	2	22	
Karlsruhe	1	2	22	
München	1	2	22	
Darmstadt	1	2	22	
Berlin	1	2	12	
Passau	1	2	2	Phase II
Hannover	1		1	
Hamburg	1		1	
Lübeck	1			
Stuttgart	1			
Total	185			

G-Lab Structure

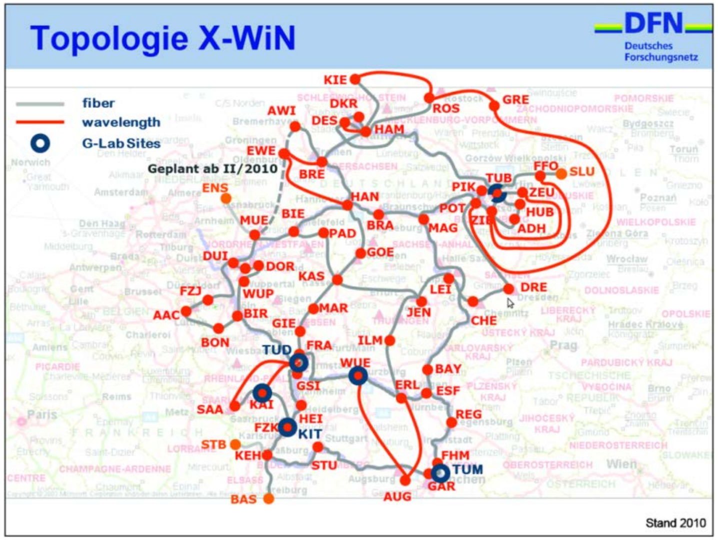




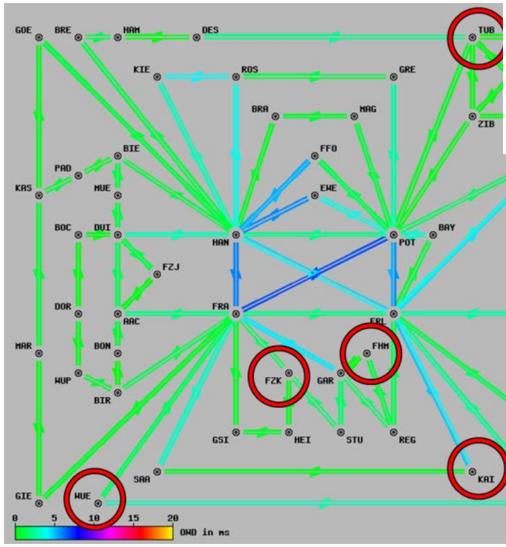
- Central Node
  - Resource management
    - Experiment scheduling
    - Resource provisioning
  - Boot Image management
    - Distributes Images
    - Assigns Images to nodes
- Each site has a Headnode
  - Manages local nodes
    - DHCP
    - Netboot
    - Monitoring
    - ILOM access
  - Executes orders from Central node
    - Local overrides possible
- G-Lab Central Services
  - Overall user management
  - Not an open platform
  - Trouble ticket system (OTRS)
  - WiKi, data storage, …
  - Based on TYPO3 (CMS)

G-Lab Network Topology

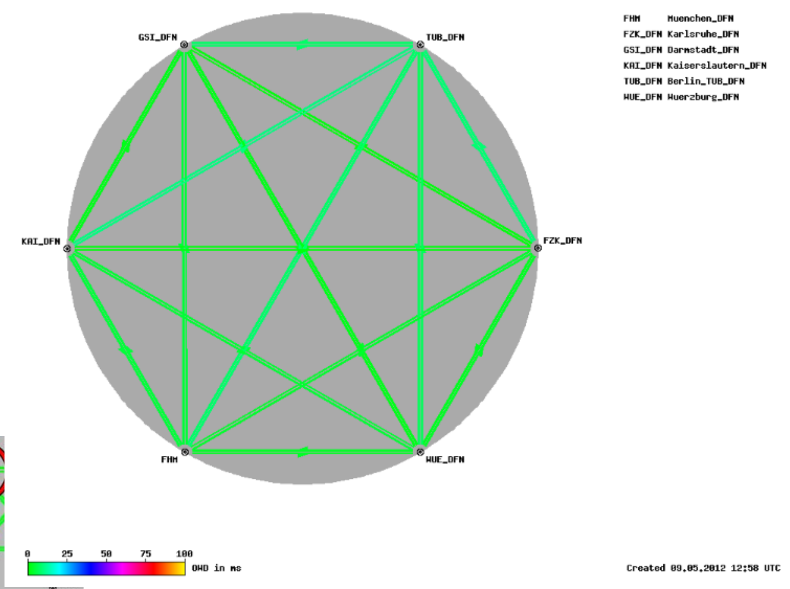
G-Lab · MyOps · PNP4Nagios · CoMon · Nagios · NagVis · NAGIOS ADMINISTRATOR



Physical Topology



IP Topology



OWD

Flexibility

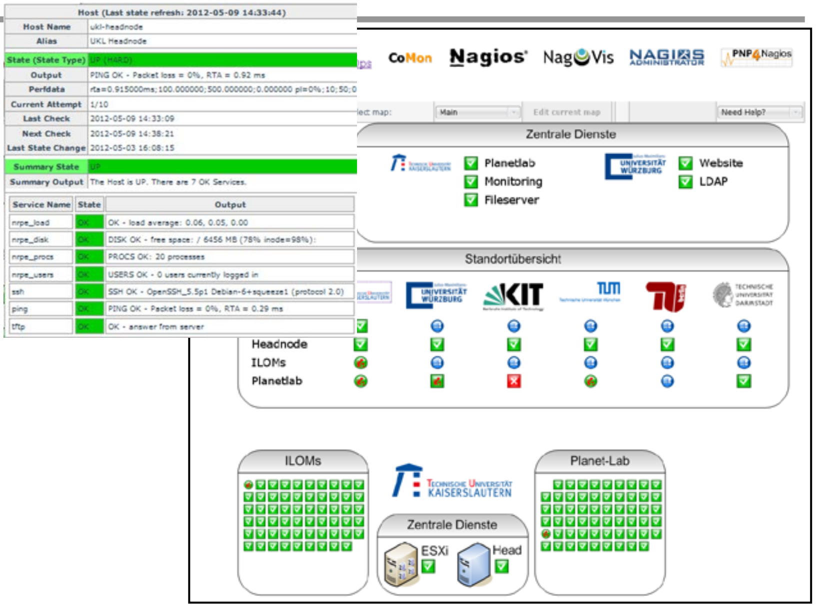
- Experimental Facility is part of research experiments
  - Facility can be modified to fit the experiments needs
  - Researchers can run experiments that might break the facility
    - Experimental facility instead of a testbed
- Research is not limited by
  - Current software setup
  - Current hardware setup
  - Restrictive policies

Mobility
Energy Efficiency
Sensornetworks
…

- Experimental Facility is evolving
  - Cooperative approach
    - „When you need it, build it“
    - Core team helps
  - Cooperation with other facilities (e.g. Planet-Lab, GENI, …)
  - Sustainability (as a non profit organization) / Federation

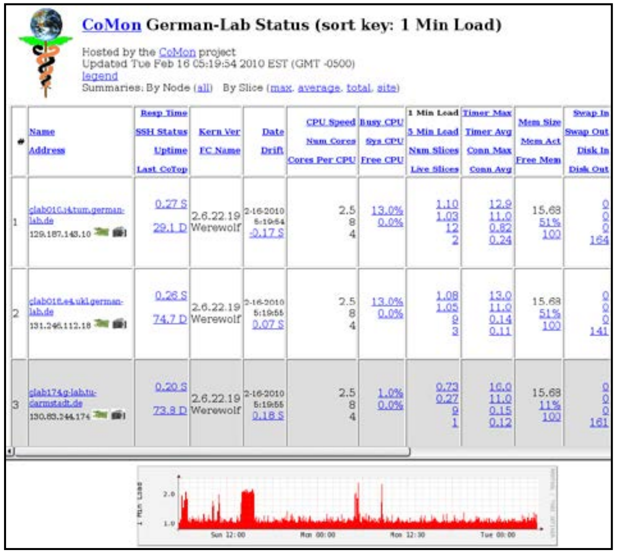
G-Lab Monitoring Framework

- **Nagios**
  - Central monitoring in Kaiserslautern
  - Obtain information from other sites via NRPE proxy on the head-node
    - nrpe (Nagios Remote Plugin Executor)
  - Checks
    - Availability of Nodes
    - Status of special services
    - Hardware status (via ILOM)
    - http://nagios.german-lab.de
- **CoMon**
  - Planet-Lab specific monitoring
    - In cooperation with Planet-Lab, Princeton
  - Monitors nodes from within
    - CPU, Memory, IO
  - Slice centric view
    - Monitors experiments



Host (Last state refresh: 2012-05-09 14:33:44)	
Host Name	ukl-headnode
Alias	UKL Headnode
State (State Type)	UP (HARD)
Output	PING OK - Packet loss = 0%, RTA = 0.92 ms
Perfdata	rta=0.915000ms;100.000000;500.000000;0.000000 pl=0%;10;50;0
Current Attempt	1/10
Last Check	2012-05-09 14:33:09
Next Check	2012-05-09 14:38:21
Last State Change	2012-05-03 16:08:15
Summary State	UP
Summary Output	The Host is UP. There are 7 OK Services.

Service Name	State	Output
nrpe_load	OK	OK - load average: 0.06, 0.05, 0.00
nrpe_disk	OK	DISK OK - free space: / 6456 MB (78% inode=98%):
nrpe_procs	OK	PROCS OK: 20 processes
nrpe_users	OK	USERS OK - 0 users currently logged in
ssh	OK	SSH OK - OpenSSH_5.5p1 Debian-6+squeeze1 (protocol 2.0)
ping	OK	PING OK - Packet loss = 0%, RTA = 0.29 ms
tftp	OK	OK - answer from server



CoMon German-Lab Status (sort key: 1 Min Load)

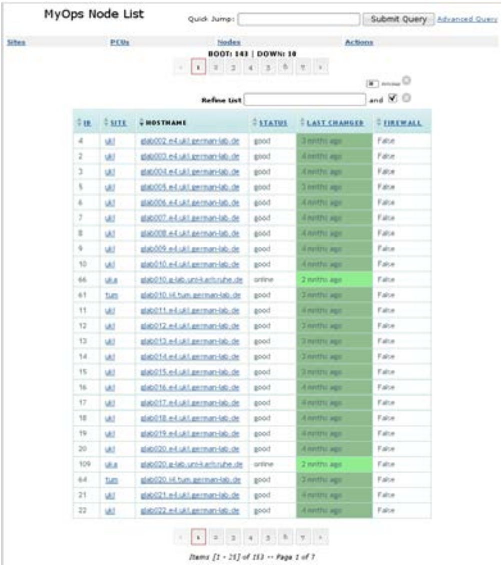
Hosted by the CoMon project
Updated Tue Feb 16 05:19:54 2010 EST (GMT -0500)
legend
Summaries: By Node (all) By Slice (max, average, total, site)

#	Name Address	Resp Time SSH Status Uptime Last CoTop	Kern Ver FC Name	Date Drift	CPU Speed Num Cores Cores Per CPU	Busy CPU Sys CPU Free CPU	1 Min Load 5 Min Load Num Slices Live Slices	Timer Max Timer Avg Conn Max Conn Avg	Mem Size Mem Act Free Mem	Swap In Swap Out Disk In Disk Out
1	glab010.i4.tum.german-lab.de 129.187.143.10	0.27 S 29.1 D	2.6.22.19 Werewolf	2-16-2010 5:19:54 -0.17 S	2.5 8 4	13.0% 0.0%	1.10 1.03 12 2	12.9 11.0 0.82 0.24	15.68 51% 100	0 0 0 164
2	glab016.e4.ukl.german-lab.de 131.246.112.18	0.26 S 74.7 D	2.6.22.19 Werewolf	2-16-2010 5:19:55 0.07 S	2.5 8 4	13.0% 0.0%	1.08 1.05 9 3	13.0 11.0 0.14 0.11	15.68 51% 100	0 0 0 141
3	glab174.g-lab.tu-darmstadt.de 130.83.244.174	0.20 S 73.8 D	2.6.22.19 Werewolf	2-16-2010 5:19:55 0.18 S	2.5 8 4	1.0% 0.0%	0.73 0.27 9 1	16.0 11.0 0.15 0.12	15.68 11% 100	0 0 0 161

G-Lab Monitoring Framework

- **MyOps**
  - Planet-Lab specific tool
    - In cooperation with Planet-Lab, Princeton
  - Detects common Planet-Lab problems
  - Reacts to problems

- **In/Out Network traffic**
  - Based on DFN connectivity
  - Important to control the lab at runtime to avoid interference with operational systems
  - Traffic patterns can be stored and related to the experiments
    - Quality assurance of the experiments
  - Further developments
    - MPLS or wavelength links

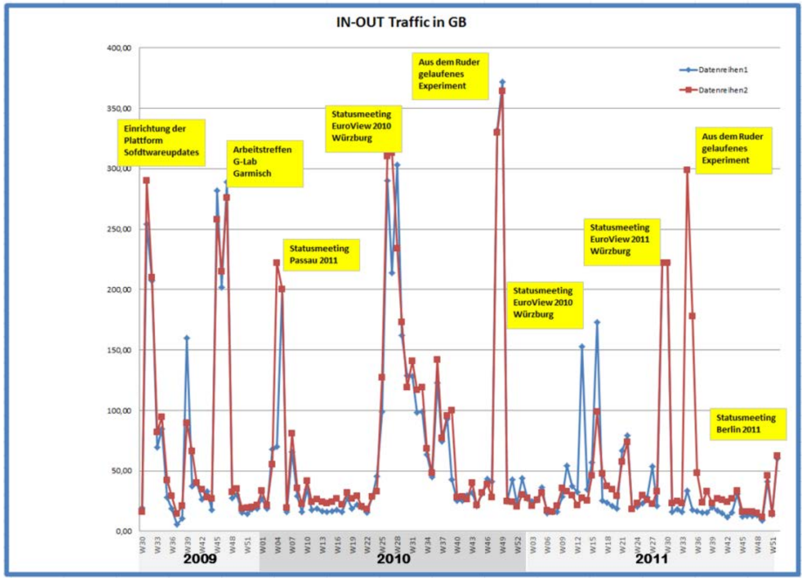


MyOps Node List

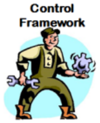
BOOT: 143 | DOWN: 10

ID	SITE	HOSTNAME	STATUS	LAST CHANGED	FIREWALL
4	ukl	glab002.e4.ukl.german-lab.de	good	3 mnths ago	False
2	ukl	glab003.e4.ukl.german-lab.de	good	4 mnths ago	False
3	ukl	glab004.e4.ukl.german-lab.de	good	4 mnths ago	False
5	ukl	glab005.e4.ukl.german-lab.de	good	3 mnths ago	False
6	ukl	glab006.e4.ukl.german-lab.de	good	4 mnths ago	False
7	ukl	glab007.e4.ukl.german-lab.de	good	4 mnths ago	False
8	ukl	glab008.e4.ukl.german-lab.de	good	4 mnths ago	False
9	ukl	glab009.e4.ukl.german-lab.de	good	4 mnths ago	False
10	ukl	glab010.e4.ukl.german-lab.de	good	4 mnths ago	False
66	uka	glab010.g-lab.uni-karlsruhe.de	online	2 mnths ago	False
61	tum	glab010.i4.tum.german-lab.de	good	3 mnths ago	False
11	ukl	glab011.e4.ukl.german-lab.de	good	4 mnths ago	False
12	ukl	glab012.e4.ukl.german-lab.de	good	3 mnths ago	False
13	ukl	glab013.e4.ukl.german-lab.de	good	3 mnths ago	False
14	ukl	glab014.e4.ukl.german-lab.de	good	3 mnths ago	False
15	ukl	glab015.e4.ukl.german-lab.de	good	3 mnths ago	False
16	ukl	glab016.e4.ukl.german-lab.de	good	4 mnths ago	False
17	ukl	glab017.e4.ukl.german-lab.de	good	4 mnths ago	False
18	ukl	glab018.e4.ukl.german-lab.de	good	4 mnths ago	False
19	ukl	glab019.e4.ukl.german-lab.de	good	4 mnths ago	False
20	ukl	glab020.e4.ukl.german-lab.de	good	4 mnths ago	False
109	uka	glab020.g-lab.uni-karlsruhe.de	online	2 mnths ago	False
64	tum	glab020.i4.tum.german-lab.de	good	3 mnths ago	False
21	ukl	glab021.e4.ukl.german-lab.de	good	4 mnths ago	False
22	ukl	glab022.e4.ukl.german-lab.de	good	4 mnths ago	False

Items [1 - 21] of 153 -- Page 1 of 7



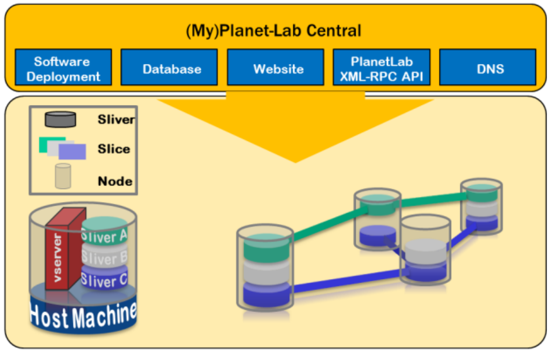
IN-OUT Traffic in GB



Control Framework

Planet-Lab

- Easy management of testbed-„silce“
- Lightweight virtualization
- Flat network
- Rich tool support (monitoring, experiment control)



ToMaTo

- Topology-oriented
- Multiple virtualization options
- Virtualized and emulated networks

ToMaTo

Seattle

- For algorithm testing
- Executes code in custom python dialect
- Federated with GENI Seattle

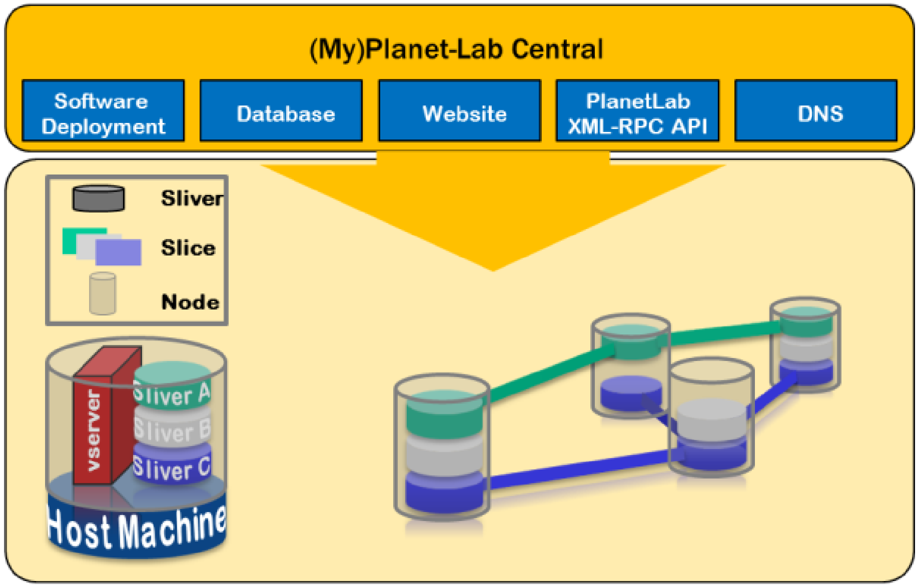
Seattle
Open peer-to-peer computing

Custom Boot-Images

- Software comes as boot image
- Either booted directly on hardware or in virtualization

Planet-Lab Structure

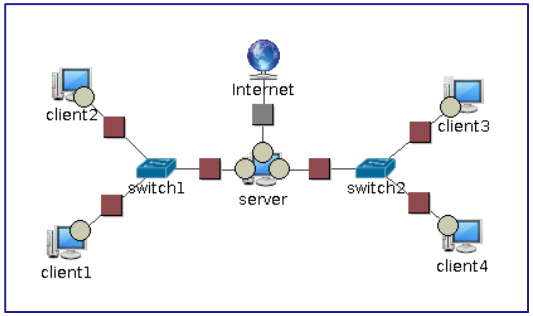
- Planet-Lab
  - Testbed and software by Princeton
  - Only Software is used
  - Extended in Cooperation with Princeton
- Uses Virtualization
  - Provides virtual node access called „Sliver“
  - Slivers across several nodes form a „Slice“
- Central configuration
  - Planet-Lab Central (PLC) in Kaiserslautern
  - User management
  - Sliver management



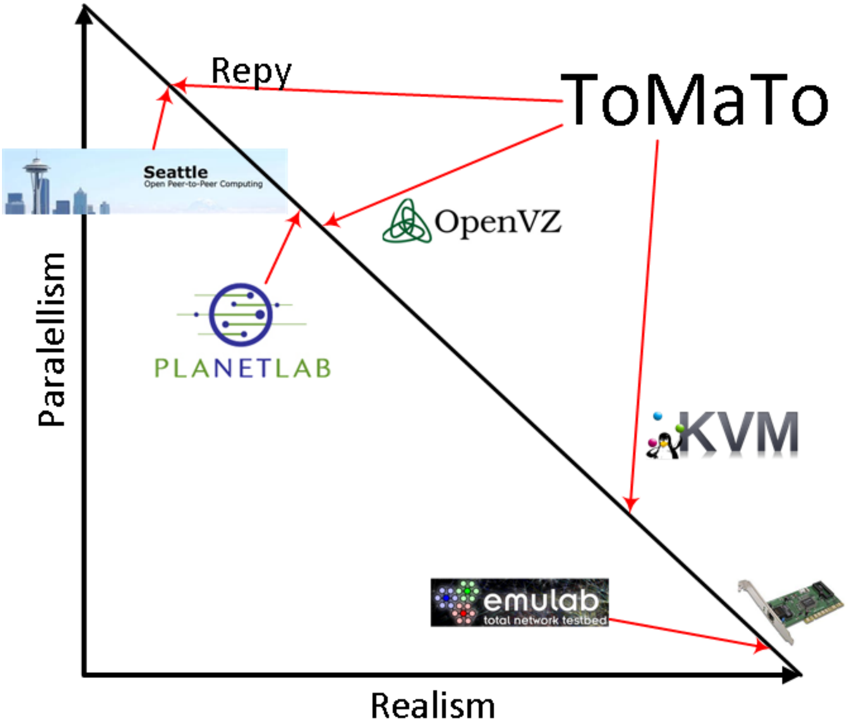
ToMaTo - A network experimentation tool[1]

[1]: http://dswd.github.com/ToMaTo/

- ToMaTo „Topology Management Tool“



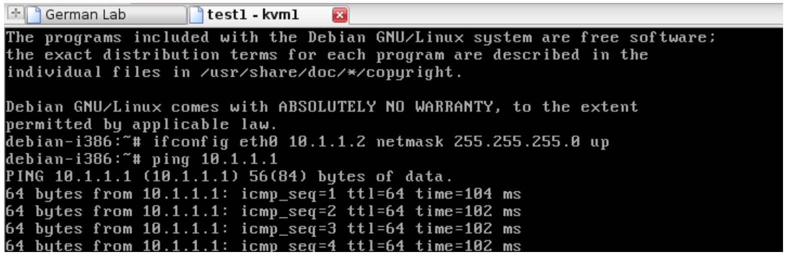
- Topology contains
  - **Devices**: produce and consume data; can run software
    - **Three kinds of devices**
      - KVM devices
      - OpenVZ devices
      - Programmable devices
  - **Connectors** forward and manipulate data and connect devices
    - **Two kinds of connectors**
      - VPN networks (based on Tinc)
      - External networks



http://dswd.github.com/ToMaTo/presentations.html

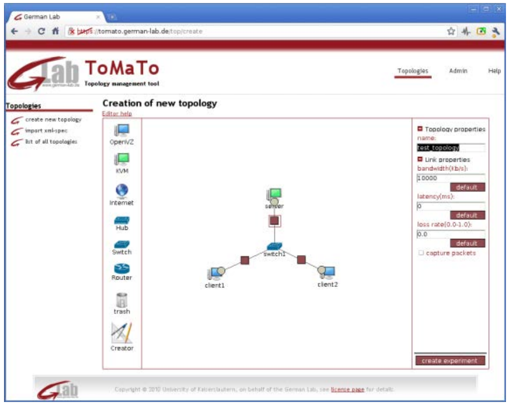
ToMaTo – Features and editor

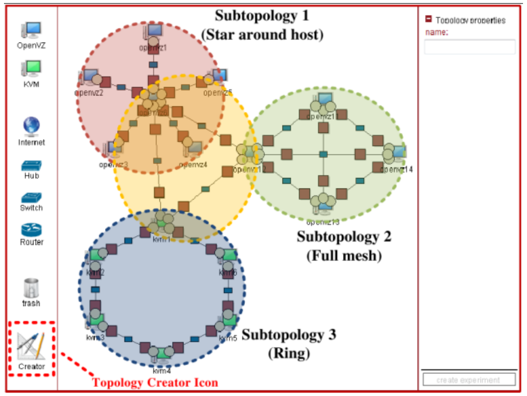
- Administrator/Developer features
  - Intelligent load-balancing
  - Open xml-rpc interface
  - Administrator tools
  - LDAP integration
- User features
  - Automatic network interface configuration
  - Changes to running topologies
  - Console access
  - Image up/download
  - Pcap capturing (packet capturing)
- ToMaTo graphical editor
  - Automatically creates topologies
  - Ring-, Star- and Full mesh topologies
  - Connects topologies
- Configures network interfaces
  - IP addresses
  - Netmasks
- DEMO Video gives a short introduction



```
The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
debian-i386:~# ifconfig eth0 10.1.1.2 netmask 255.255.255.0 up
debian-i386:~# ping 10.1.1.1
PING 10.1.1.1 (10.1.1.1) 56(84) bytes of data.
64 bytes from 10.1.1.1: icmp_seq=1 ttl=64 time=104 ms
64 bytes from 10.1.1.1: icmp_seq=2 ttl=64 time=102 ms
64 bytes from 10.1.1.1: icmp_seq=3 ttl=64 time=102 ms
64 bytes from 10.1.1.1: icmp_seq=4 ttl=64 time=102 ms
```





ToMaTo Application Area

Access layer experiments

- Consider lower layers and hardware
  - Example: Mobile handover
- Requirements
  - Hardware access
  - Custom operating systems (Realtime)
  - Heterogeneous access technologies (3G, Wifi, etc.)
- Needs specialized **testbeds** depending on hardware NO ToMaTo support
  - DES Testbed, Wisebed

Network layer experiments

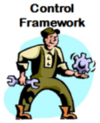
- Focus on TCP/IP suite
  - Example: IPv6 extensions, TCP substitutes
- Requirements
  - Deep OS access (modified kernels, etc.)
  - Small but complex topologies, link emulation
- ToMaTo **offers**
  - Full kernel access via KVM
  - Complex topologies
  - Link emulation
  - Packet capturing (for analysis)
  - Easy setup of topologies

Algorithm/Protocol experiments

- Work on top of network layer
  - Example: P2P-Networks
- Requirements
  - Huge but simple topologies
  - Link emulation
  - No hardware or OS access
- ToMaTo **offers**
  - Lightweight virtualization with OpenVZ
  - Link emulation
  - Federation with other testbeds via Internet

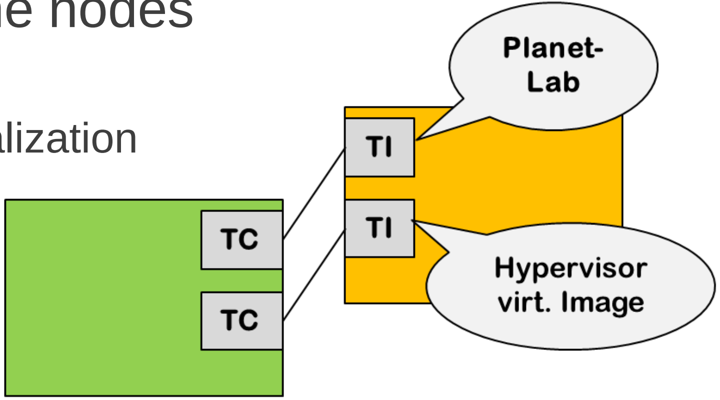
Legacy software experiments

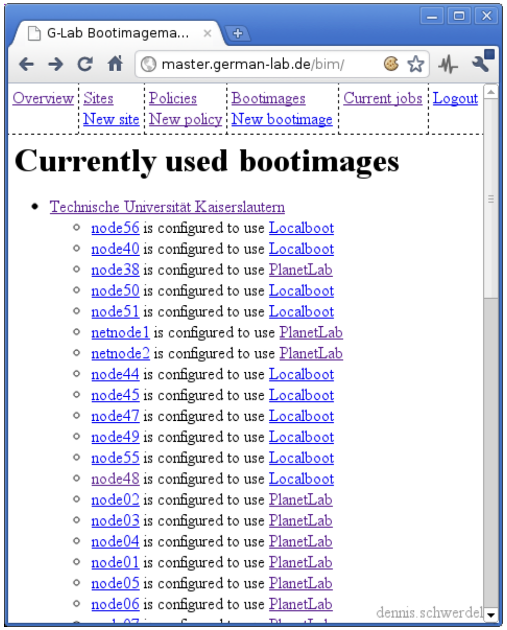
- Considers legacy software
  - „Legacy software“ refers to any widespread software with undocumented or unpublished behavior
  - Example: Skype and Windows
- Requirements
  - Special environments, custom operating systems
  - Small but complex topologies
  - Link emulation and external packet capturing
- ToMaTo **offers**
  - Custom operating systems with KVM (Windows
  - Access to external service via Internet connector
  - Packet capturing independent of guest OS



Boot Images

- Researchers can run any software on the nodes
  - Software comes as boot image
  - Either booted directly on hardware or in virtualization
- Three types of boot image
  1. Planet-Lab
     - Access for everybody
     - Restricted hardware access
  2. Hypervisor virtualization image
     - Access for everybody
     - Unrestricted access to virtual hardware
     - Topology management via ToMaTo
  3. Custom boot image
     - Access can be restricted to specific research group
     - Unrestricted access to real hardware
- Access regulated by policy
  - Favors generic images with open access over specific images with restricted access
  - Policy does not over-regulate





Currently used bootimages

- Technische Universität Kaiserslautern
  - node56 is configured to use Localboot
  - node40 is configured to use Localboot
  - node38 is configured to use PlanetLab
  - node50 is configured to use Localboot
  - node51 is configured to use Localboot
  - netnode1 is configured to use PlanetLab
  - netnode2 is configured to use PlanetLab
  - node44 is configured to use Localboot
  - node45 is configured to use Localboot
  - node47 is configured to use Localboot
  - node49 is configured to use Localboot
  - node55 is configured to use Localboot
  - node48 is configured to use Localboot
  - node02 is configured to use PlanetLab
  - node03 is configured to use PlanetLab
  - node04 is configured to use PlanetLab
  - node01 is configured to use PlanetLab
  - node05 is configured to use PlanetLab
  - node06 is configured to use PlanetLab

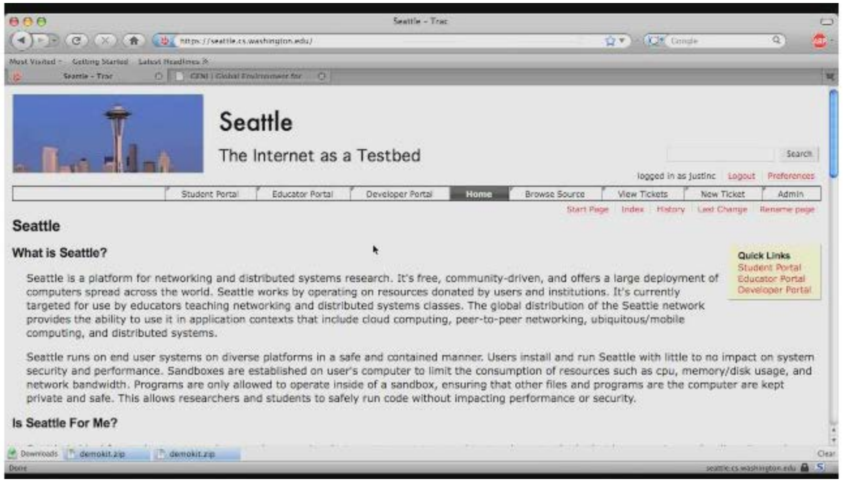
dennis.schwerdel

Seattle

Open peer-to-peer computing

- Testbed for python code
- Very lightweight, no virtualization, just sandbox
- Very comfortable experiment control
- Fully federated with Seattle GENI (over 1000 nodes)
- Wide variety of network types accessible
  - Sensors
  - Cell phones
  - Mobile nodes
- Coming soon in G-Lab, early tests running
- Algorithm testing
  - https://seattle.cs.washington.edu
  - Developed by Justin Cappos (University of Washington)

- Demo Video
  - https://seattle.cs.washington.edu/wiki/UnderstandingSeattle/DemoVideo



Seattle

The Internet as a Testbed

Seattle

What is Seattle?

Seattle is a platform for networking and distributed systems research. It's free, community-driven, and offers a large deployment of computers spread across the world. Seattle works by operating on resources donated by users and institutions. It's currently targeted for use by educators teaching networking and distributed systems classes. The global distribution of the Seattle network provides the ability to use it in application contexts that include cloud computing, peer-to-peer networking, ubiquitous/mobile computing, and distributed systems.

Seattle runs on end user systems on diverse platforms in a safe and contained manner. Users install and run Seattle with little to no impact on system security and performance. Sandboxes are established on user's computer to limit the consumption of resources such as cpu, memory/disk usage, and network bandwidth. Programs are only allowed to operate inside of a sandbox, ensuring that other files and programs are the computer are kept private and safe. This allows researchers and students to safely run code without impacting performance or security.

Is Seattle For Me?

- This five-minute demo video should help get you acquainted with the Seattle project.

Why Federation

Zoo

- **Controlled environment**
  - Host systems
  - Network
  - Users

- **Controlled environment for**
  - development, deployment and testing of new algorithms
  - Breakable infrastructure
- **Repeatable experiments**
  - When dealing with new algorithms for routing, security, mobility, …
  - Improve scientific quality

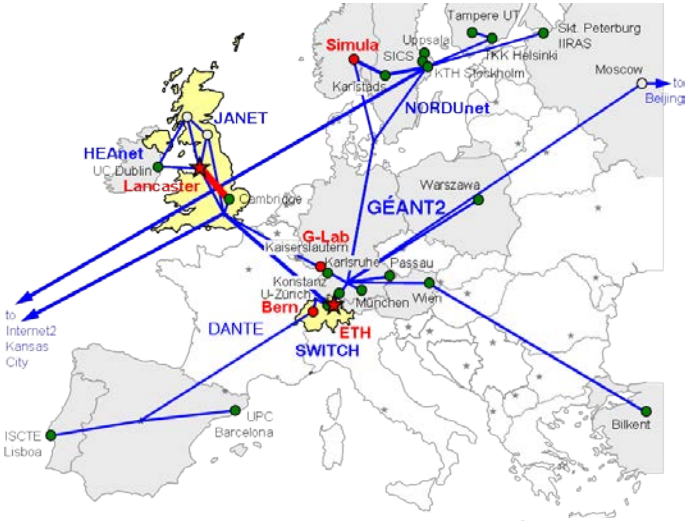
Wilderness

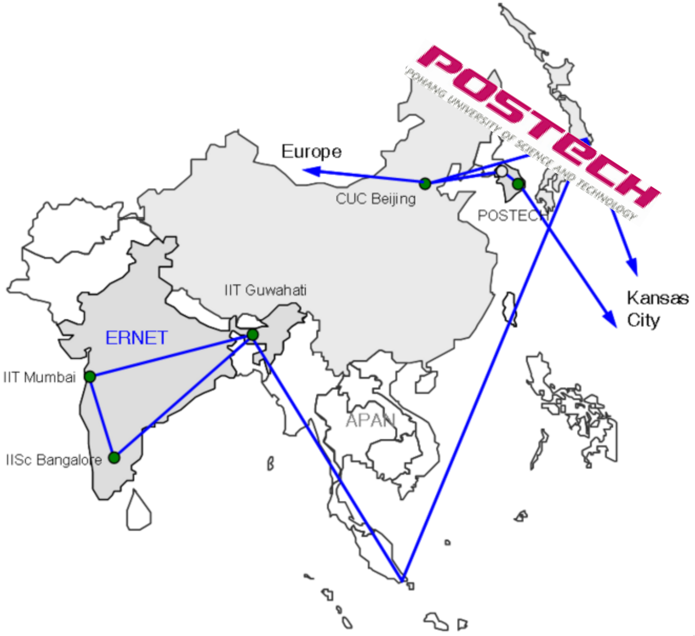
- **Scalability**
  - Hoe does new algorithms behave in the wilderness?

"Wow, Jimmy, that's pretty good!"

Federations

- GENI Federation
  - GENI connection by 1Gbit/s link from Starlink/Geant/DFN for GEC10
- GpENI *„Great Plains Environment for Network Innovation”*
  - US-based network testbed
  - Kaiserslautern is fan-out location for central European sites
  - Connection to G-Lab possible
- GpENI Asian flows use L2TPv3 and IP tunnels over **Internet2** to **APAN** (Asia-Pacific Advanced Network), which interconnects Asian regional and national research networks.
  - In Korea, **POSTECH** (Pohang University of Science and Technology) is connected to GpENI (**J. Won-Ki Hong)**





Conclusion

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

Prof. Dr. Paul Mueller

Integrated Communication Systems ICSY

University of Kaiserslautern
Department of Computer Science
P.O. Box 3049
D-67653 Kaiserslautern

Phone:	+49 (0)631 205-2263
Fax:	+49 (0)631 205-30 56
Email:	pmueller@informatik.uni-kl.de
Internet:	http://www.icsy.de