PlanetLab: A Blueprint for Introducing Disruptive Technology into the Internet

Larry Peterson
Princeton University / Intel Research

PLANETLAB

PlanetLab vs Grid

- The Grid is the next generation supercomputer

- PlanetLab is the next generation Internet

Innovator's Dilemma

- The Internet is an enormous success story
  - commercially
  - impact on our daily lives
  - global reach

- Success has an unexpected cost: *ossification*
  - difficult to deploy disruptive technologies
    - correct vulnerabilities
    - introduce new capabilities

PLANETLAB

Today's Internet

Best-Effort Packet Delivery Service

<u>Limitations</u>

- the Internet is "opaque" making it difficult to adapt to current network conditions

- applications cannot be widely distributed (typically split into two pieces: client and server)
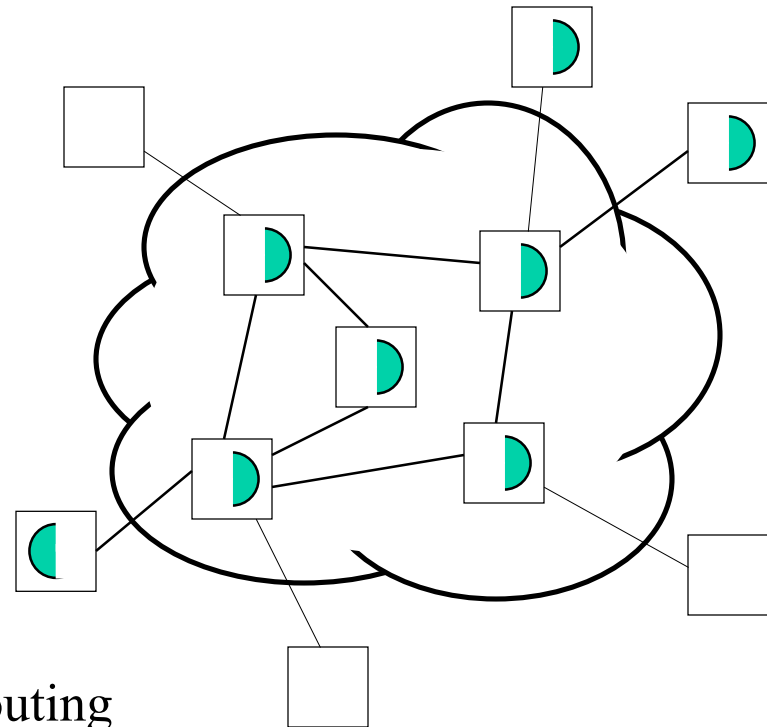
PLANETLAB
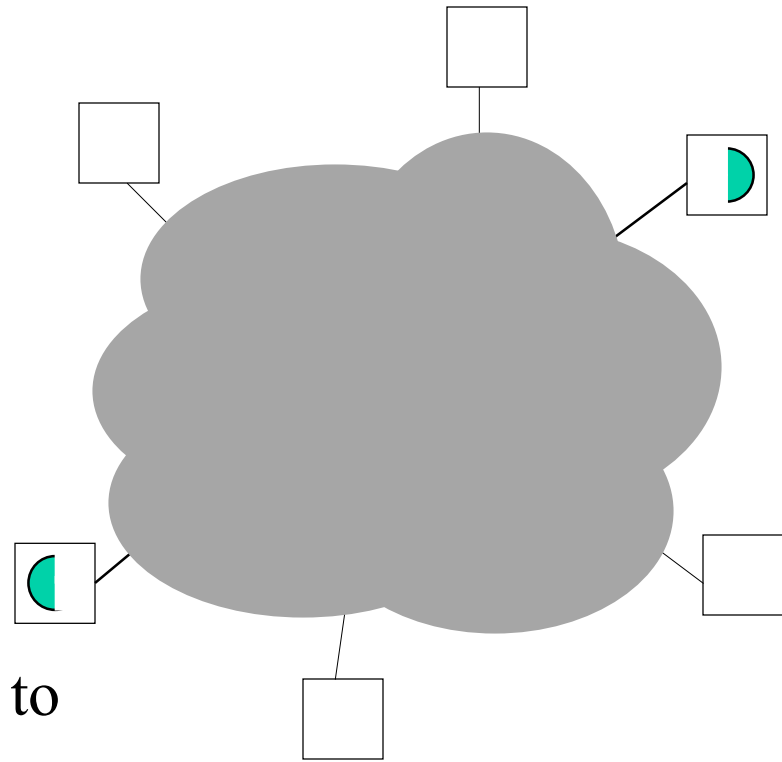
Tomorrow's Internet
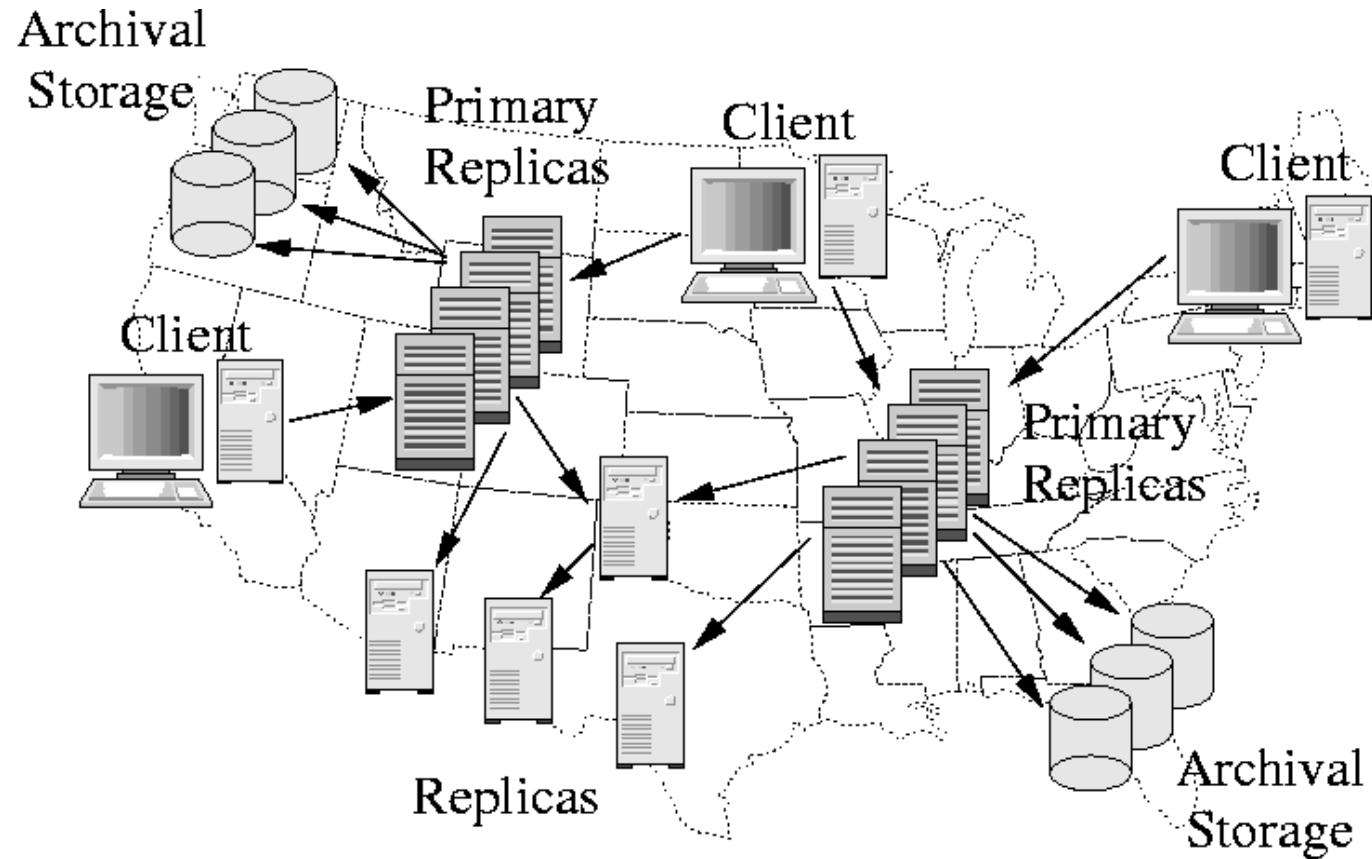
Collection of Planetary-Scale Services

Opportunities

- multiple vantage points
  - anomaly detection, robust routing
- proximity to data sources/sinks
  - content distribution, data fusion
- multiple, independent domains
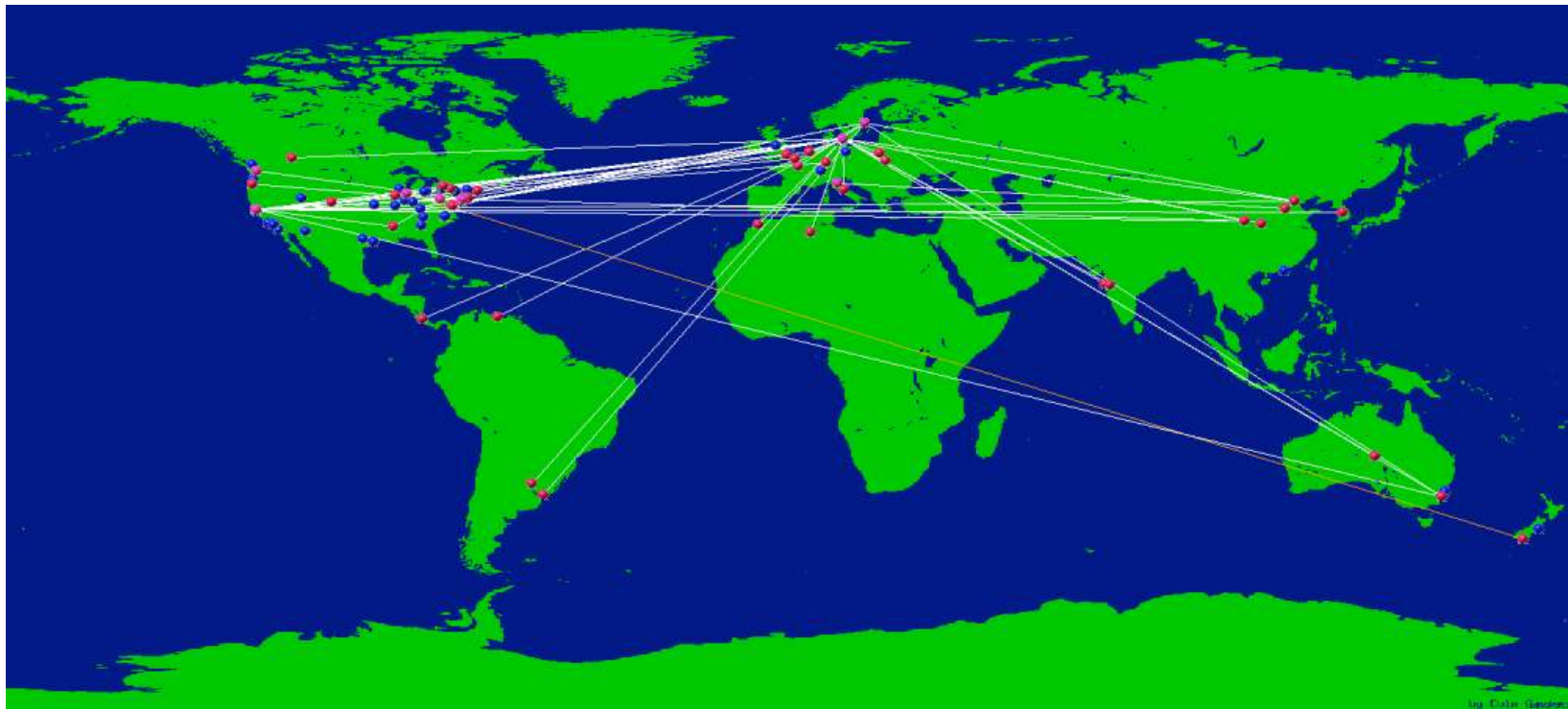 - survivable storage

PLANETLAB

Berkeley: OceanStore

RAID distributed over the whole Internet

PLANETLAB
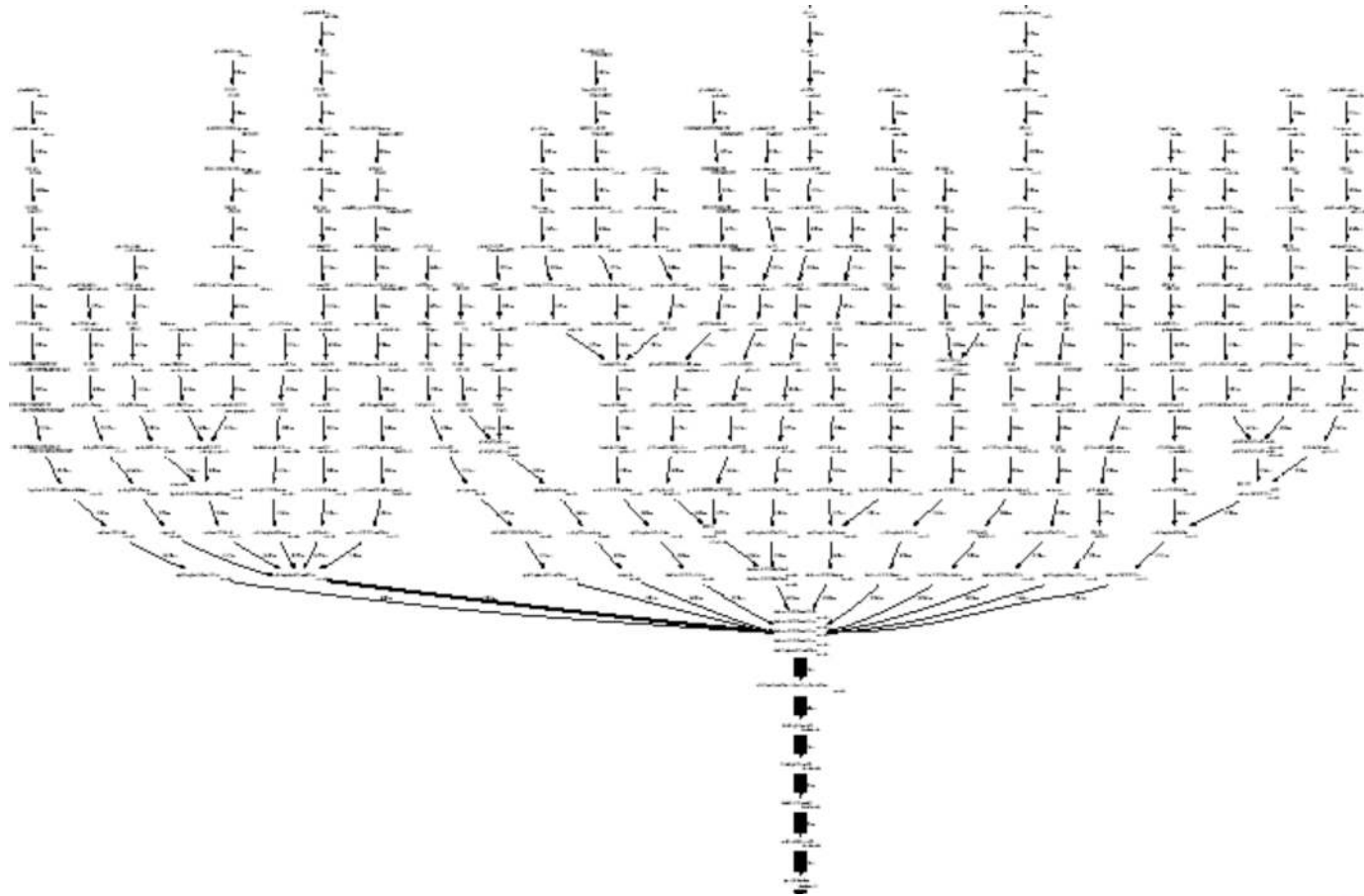
Intel: Netbait

Detect and track Internet worms globally

PLANETLAB

Washington: ScriptRoute

Internet Measurement Tool

PLANETLAB

Princeton: CoDeeN

Open Content Distribution Network

PLANETLAB
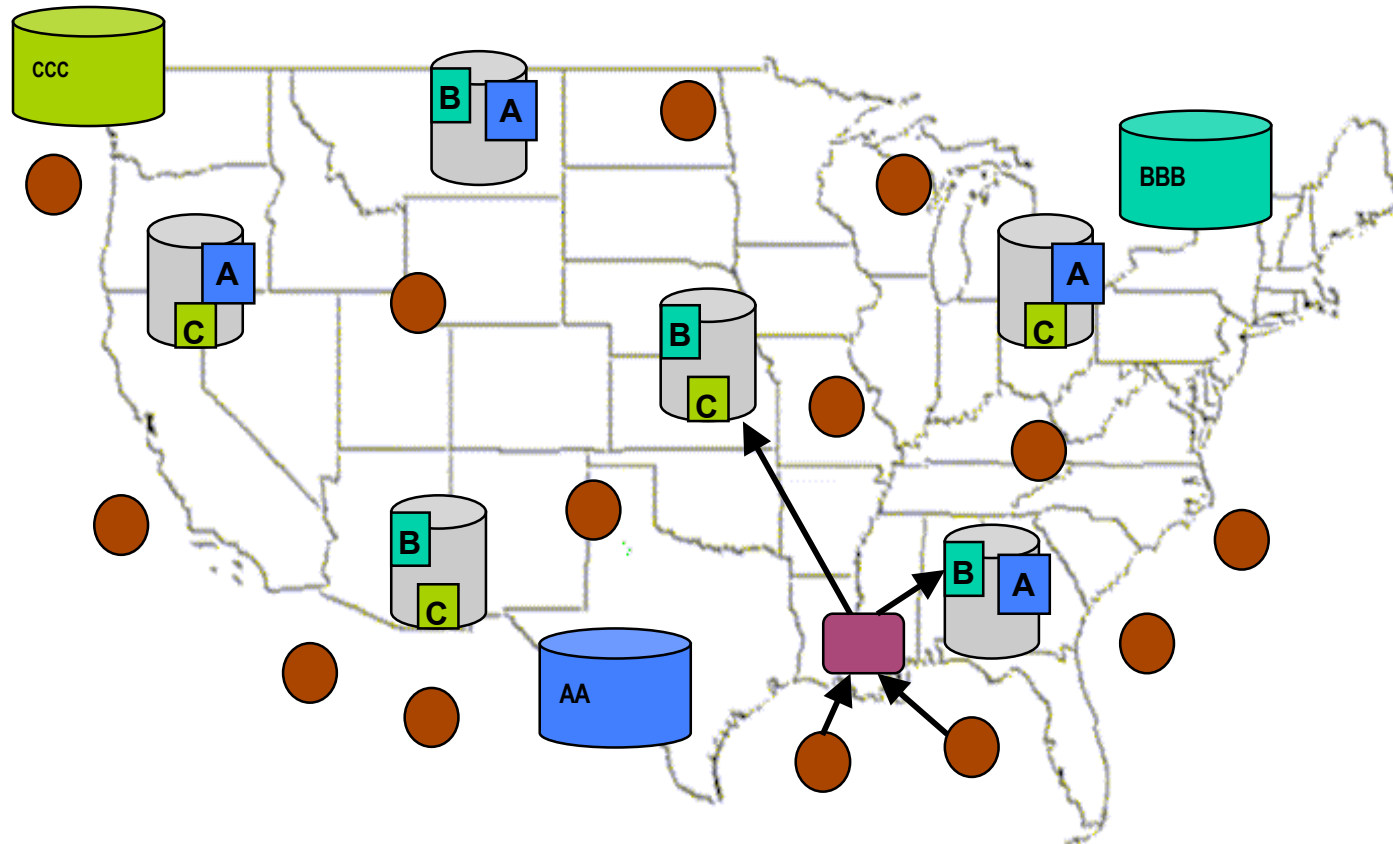
Evolving the Internet

- Add a new layer to the network architecture
  - overlay networks
    - purpose-built virtual networks that use the existing Internet for transmission
    - the Internet was once deployed as an overlay on top of the telephony network

overlay

Internet

- Challenge
  - how to innovate & deploy at scale

PLANETLAB

The Story so Far

- The Internet is a tremendous success, but…
  - the architecture has fundamental limits
  - its very success makes it hard to change
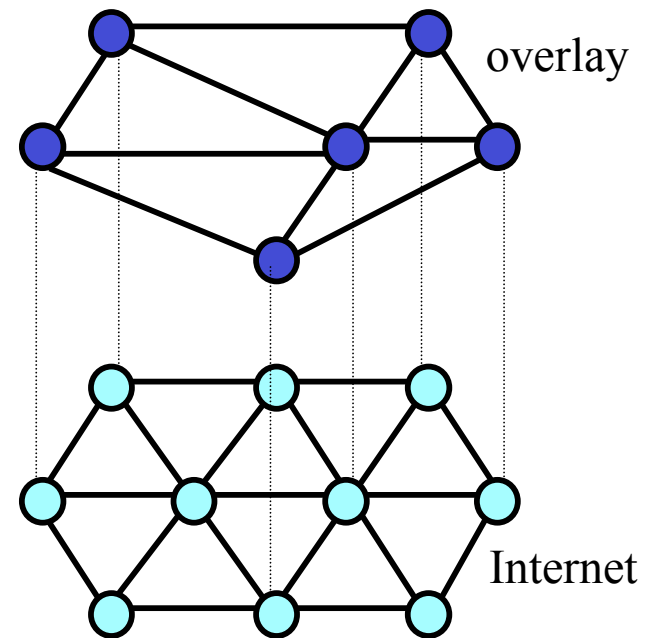
- The research community is teeming with innovative planetary-scale services
  - exploit multiple points-of-presence throughout the net

- Overlays offer an attractive way to introduce disruptive technology into the Internet, but…
  - there is a high barrier-to-entry

PLANETLAB
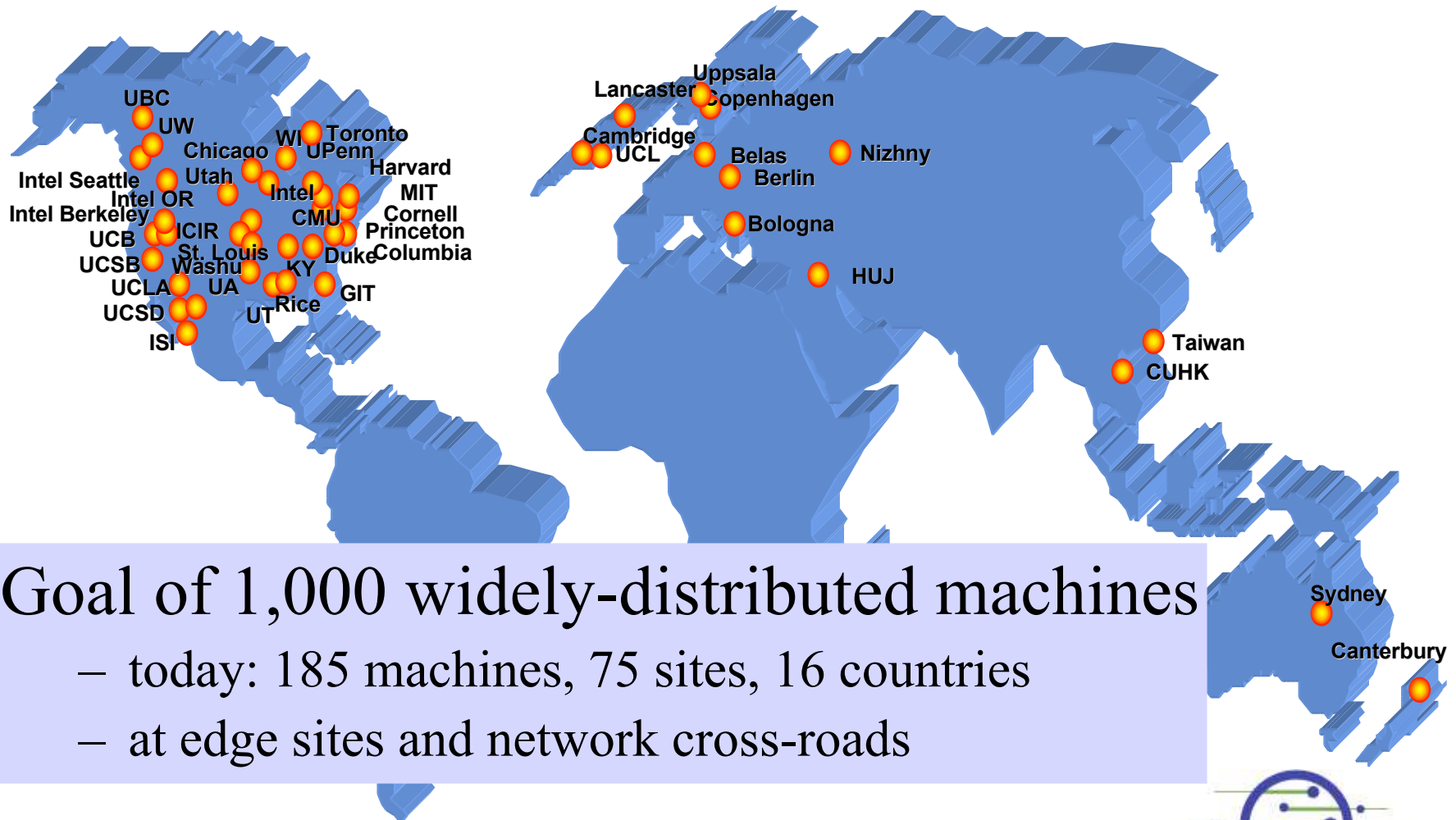
PlanetLab is…

Goal of 1,000 widely-distributed machines
- today: 185 machines, 75 sites, 16 countries
- at edge sites and network cross-roads

PlanetLab is…

A common software package

- Main components
  - Linux kernel w/ extensions to support isolation
  - bootstrapping and software distribution mechanisms
  - collection of *unbundled management* services

- Collectively support distributed virtualization
  - run many overlay networks simultaneously
  - each service (overlay) runs in a *slice* of PlanetLab's global resources

PLANETLAB

Slices
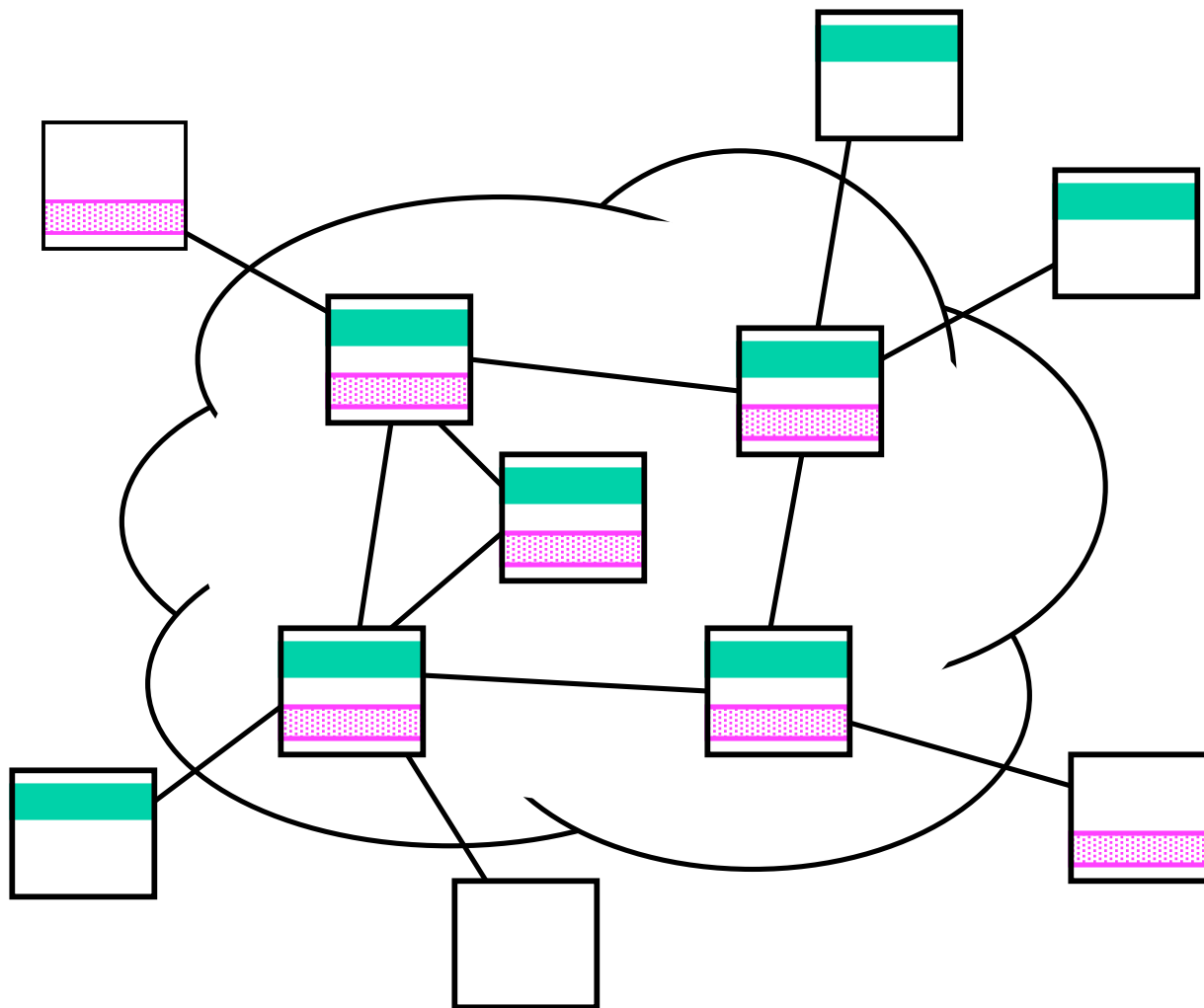
PLANETLAB

PlanetLab is…

A test-bed for experimenting with network services

- 120+ active research projects
- Advantages
  - experiment at scale
  - experiment under real-world conditions
  - potential for real workloads and users

PLANETLAB

PlanetLab is…

A deployment platform

- Continuously-running services
    - CoDeeN content distribution network (Princeton)
    - Sophia distributed query processing engine (Princeton)
    - PIER distributed query processing engine (Berkeley)
    - ScriptRoute network measurement tool (Washington)
    - NetBait worm detection service (Intel)
    - Chord scalable object location service (MIT, Berkeley)
    - OceanStore storage system (Berkeley)

design → deploy → measure →

PLANETLAB

PlanetLab is…

A microcosm of the next Internet

- Fold services back into PlanetLab
  - evolve core technologies to support overlays and slices
- Examples
  - Sophia used to monitor health of PlanetLab nodes
  - Chord provides scalable object location
- Long-term goals
  - develop open protocols and standards
    - allow federation of public & private "PlanetLabs" to co-exist
  - discover common sub-services

PLANETLAB

PlanetLab is…

A research community

- Started as a grass-roots effort
  - 35 researchers gathered in March of 2002
  - Intel provided seed funding
  - self-organized into five working groups
- Next Phase: Academic/Industrial Consortium
  - hosted by Princeton (w/ Berkeley and Washington)
  - build-out and operate the infrastructure
  - lower the barrier to entry for research and teaching

PLANETLAB

Software Architecture

- Support distributed virtualization
  - *slice*: a network of virtual machines
  - multiple services run concurrently (some long-lived)
  - deploy version *i* of PlanetLab on version *i+1*
- Per-Node Components
  - create and isolate virtual machines
- Global Components (Services)
  - create slice across a set of nodes
  - monitor node health
  - routing underlay

PLANETLAB

Per-Node Components

- Node Manager
  - responds to requests to create a virtual machine
  - defines spec for VM
    - resources consumed
    - network name space consumed
  - performs admission control
- Vserver: virtualizes at system call interface
  - each vserver runs in its own security context
    - private UID/GID name space
    - limited superuser capabilities (e.g., no CAP_NET_RAW)
  - uses `chroot` for file system isolation
  - scales to hundreds of vservers per node

PLANETLAB

Per-Node (cont)

- plkmod: kernel module that enforces VM isolation
  - processor and link scheduling
  - virtualizes the network
    - safe raw sockets
    - port-space isolation
    - address space sandboxing
- Sensors: uniform interface to node status info
  - HTTP-based
  - core set + user-defined
- Admin Slice: local admin control
  - set bw limits
  - run tcpdump

PLANETLAB

Creating Slices

- Two-stage process
  - discover available resources
    - use monitoring service
  - create virtual machine on each selected node
    - contact broker for rights to resources (receive *tickets*)
    - contact node manager to redeem tickets
    - node manager implements admission control

- Status
  - prototypes of mechanisms
  - simple policies in the near-term
  - create a market for resources in the long-term

PLANETLAB

Monitoring Services

- Serve several purposes
  - discover/select resources for a slice
  - monitor node/network health
  - measure/monitor Internet activity
- Exploit sensors
  - local state + local view of the network
- Multiple services being built
  - Sophia: distributed Prolog engine
  - PIER: distributed SQL query processor
  - IrisNet: XML-based queries
  - service-specific mechanisms (e.g., CoDeeN)

PLANETLAB

Routing Underlay

Overlay Services

___

Library of Routing Services
- k-disjoint paths
- k-nearest neighbors

___

Topology Probing Layer
- peering graph
- path from x to y
- latency from x to y

___

Raw Topology Information
- local BGP feed

PLANETLAB

Current Institutions

Academia Sinica, Taiwan	Princeton University	University of Kansas
Boston University	Purdue University	University of Kentucky
Caltech	Rensselaer Polytechnic Inst.	University of Maryland
Carnegie Mellon University	Rice University	University of Massachusetts
Chinese Univ of Hong Kong	Rutgers University	University of Michigan
Columbia University	Stanford University	University of North Carolina
Cornell University	Technische Universitat Berlin	University of Pennsylvania
Datalogisk Institut Copenhagen	The Hebrew Univ of Jerusalem	University of Rochester
Duke University	University College London	USC / ISI
Georgia Tech	University of Arizona	University of Technology Sydney
Harvard University	University of Basel	University of Tennessee
HP Labs	University of Bologna	University of Texas
Intel Research	University of British Columbia	University of Toronto
Johns Hopkins	UC Berkeley	University of Utah
Lancaster University	UCLA	University of Virginia
Lawrence Berkeley Laboratory	UC San Diego	University of Washington
MIT	UC Santa Barbara	University of Wisconsin
Michigan State University	University of Cambridge	Uppsala University, Sweden
National Tsing Hua Univ.	University of Canterbury	Washington University in St Louis
New York University	University of Chicago	Wayne State University
Northwestern University	University of Illinois	

PLANETLAB

More Information

www.planet-lab.org

PLANETLAB