Princeton P4 Campus:

Building and Running Novel Network Applications on Campus

https://p4campus.cs.princeton.edu

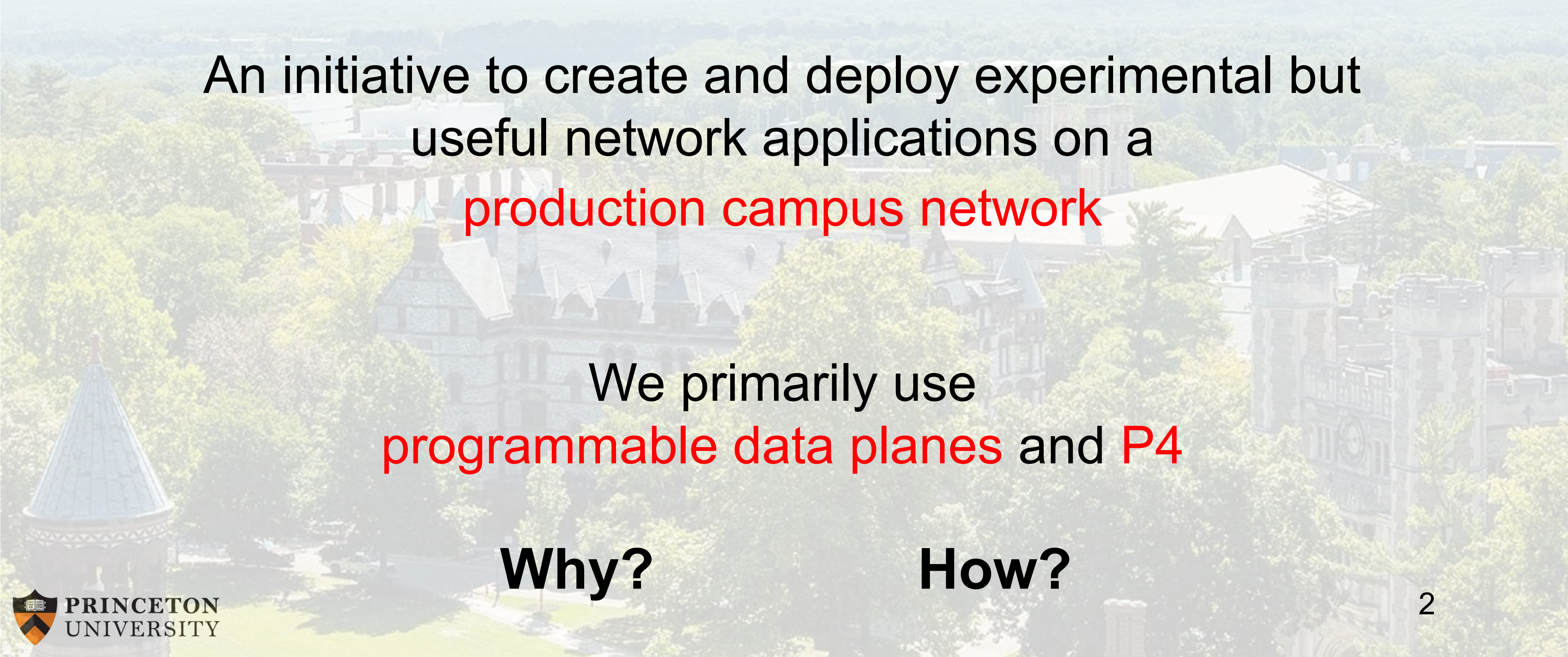
Hyojoon (Joon) Kim

Princeton University

Programmable Switches Workshop
February 23, 2022

PRINCETON
UNIVERSITY

P4 Campus



An initiative to create and deploy experimental but useful network applications on a
production campus network

We primarily use
programmable data planes and P4

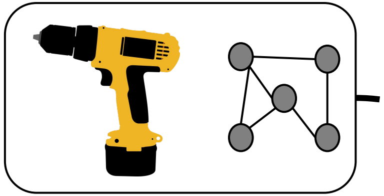
Why? **How?**

The “Gap”

Bad idea

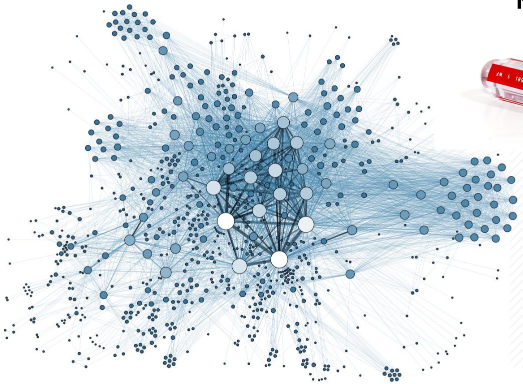
New idea

Small, unrealistic experiments



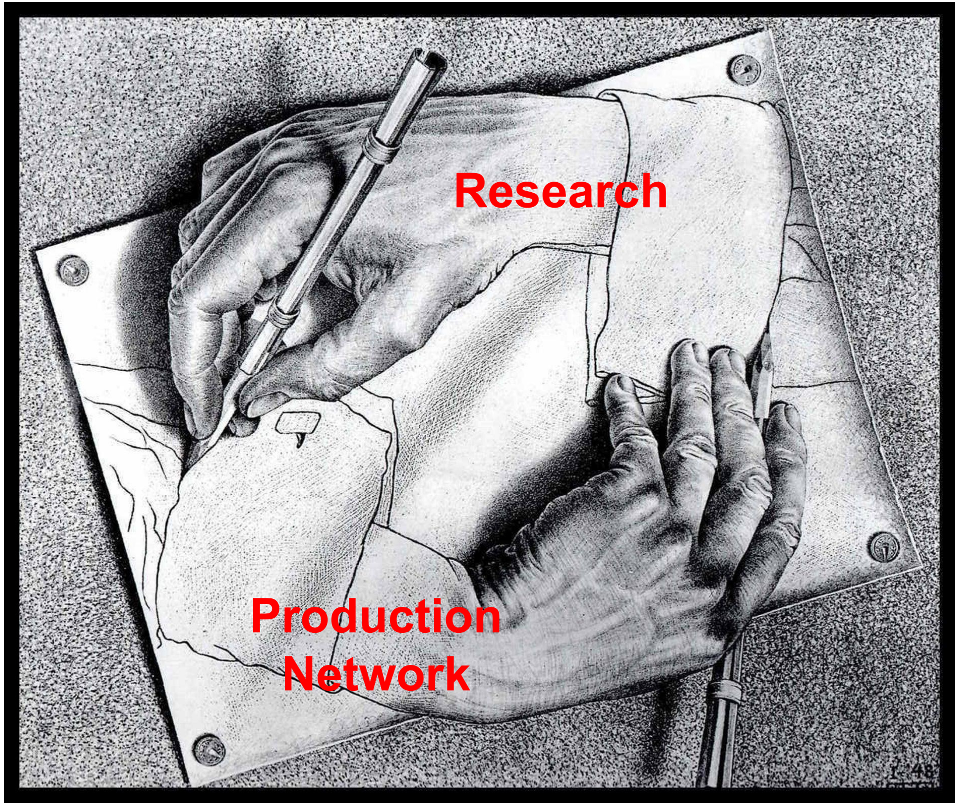
Research

Outdated tools & practices

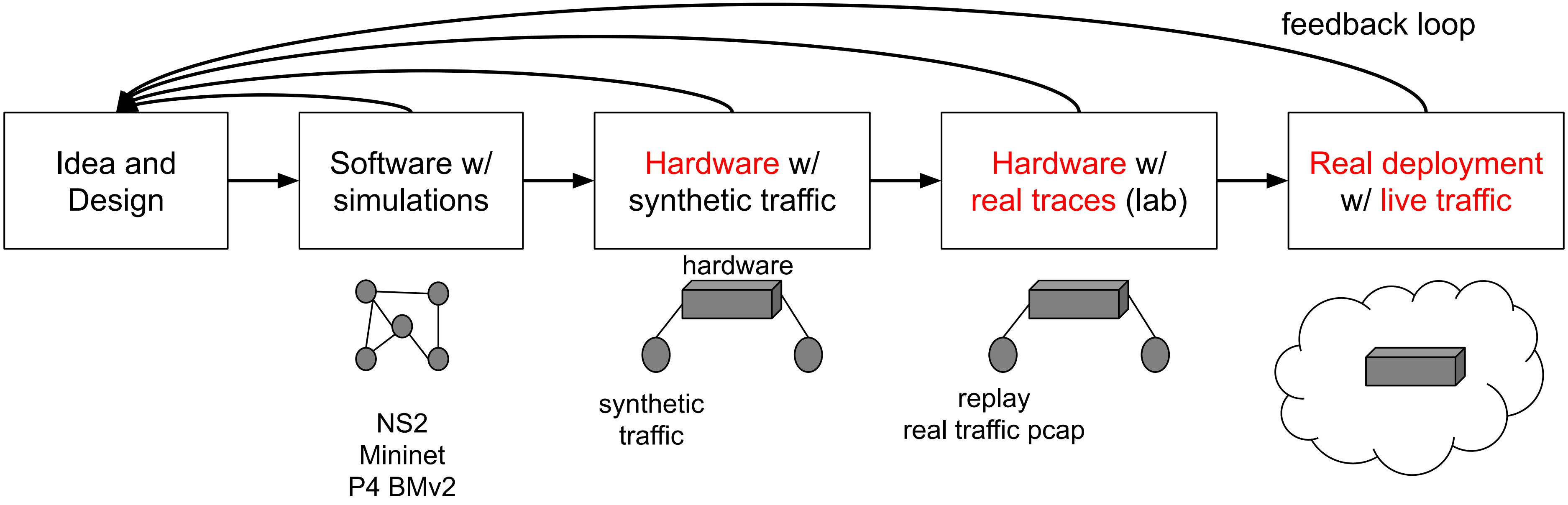


Production Network

Positive Feedback Loop Missing



Network Research Pipeline



Idea and Design

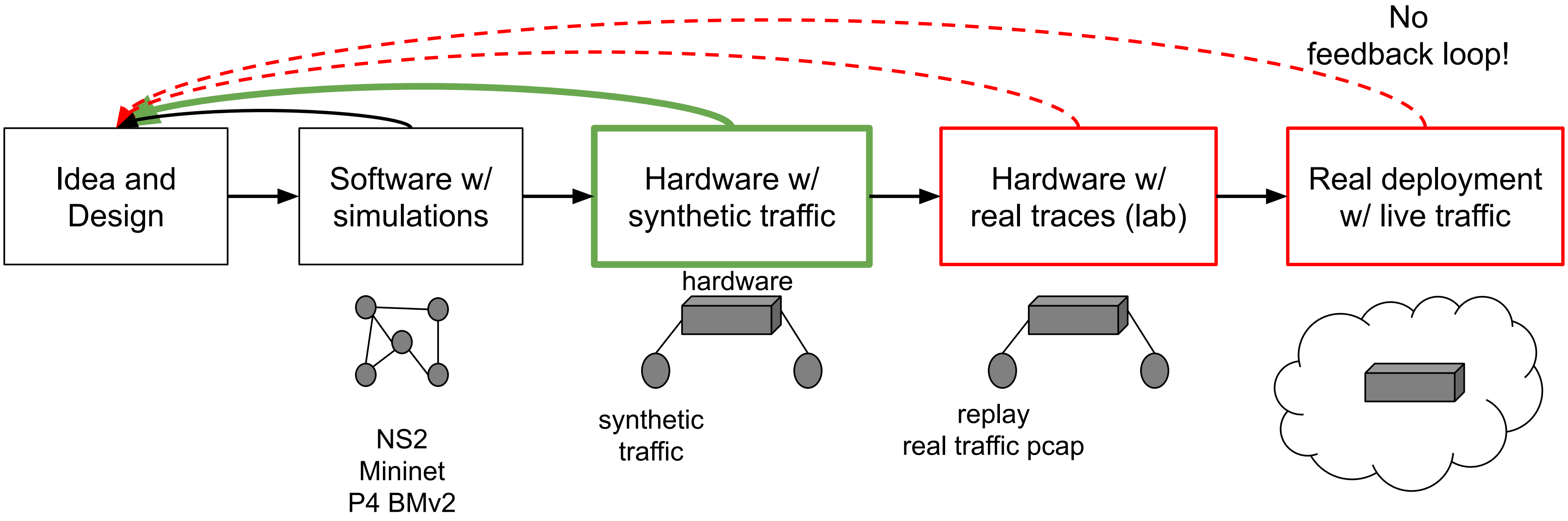
Software w/ simulations

Hardware w/ synthetic traffic

Hardware w/ real traces (lab)

Real deployment w/ live traffic

Network Research Pipeline



No feedback loop!

Idea and Design

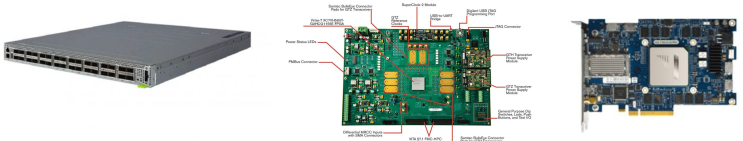
Software w/ simulations

Hardware w/ synthetic traffic

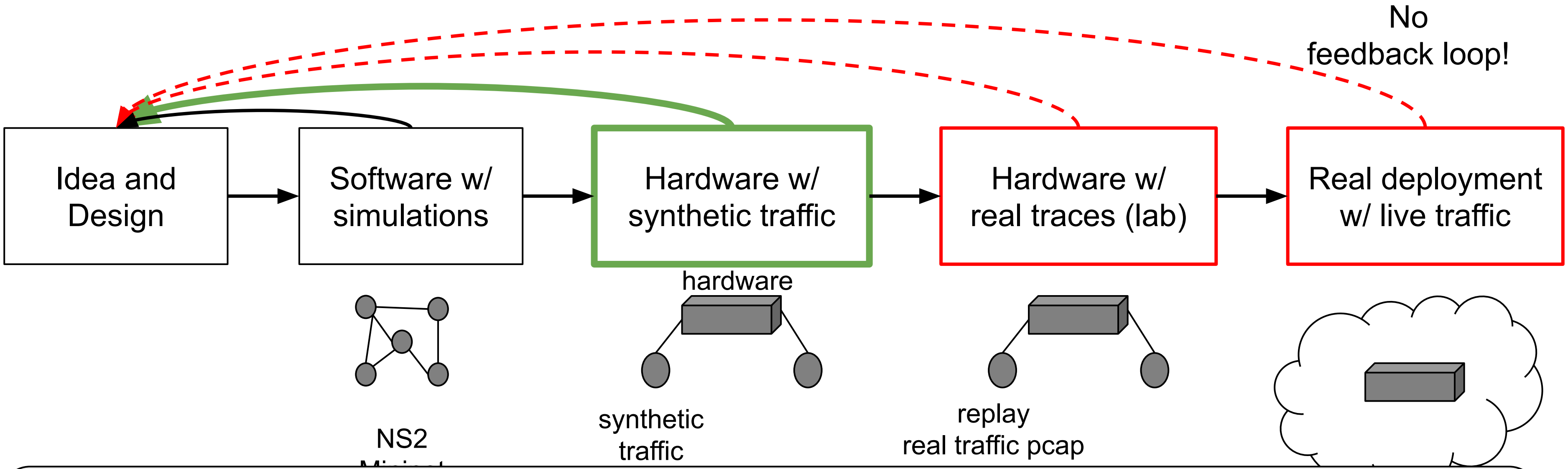
Hardware w/ real traces (lab)

Real deployment w/ live traffic

Programmable hardware



Network Research Pipeline



**Without the last two stages,
new ideas barely see light outside of a lab**

Real Traces & Deployment is Hard

- Disruptive
- User privacy (PII in production traffic)

Personally Identifiable Information

- Lack of collaboration and communication

Alternatives: Dedicated Testbed

PLANETLAB

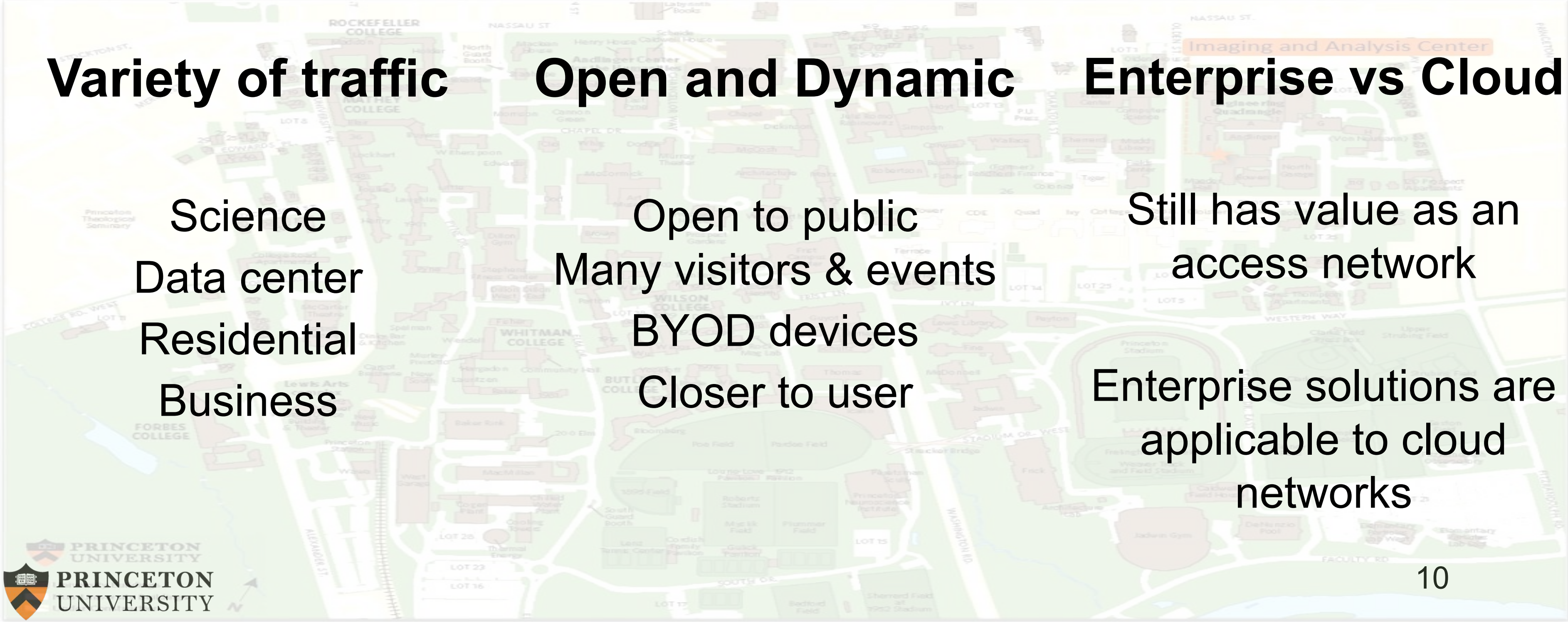
geni
Exploring Networks
of the Future

FABRIC

Real network, WAN-scale

Limited access to production traffic

Campus Network as Lab



Variety of traffic

Science

Data center

Residential

Business

Open and Dynamic

Open to public

Many visitors & events

BYOD devices

Closer to user

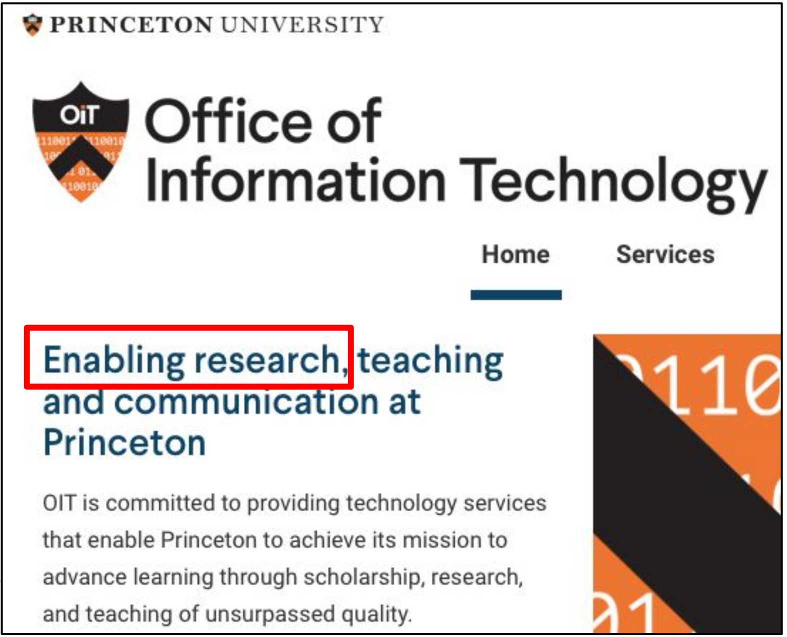
Enterprise vs Cloud

Still has value as an access network

Enterprise solutions are applicable to cloud networks

Campus Network as Lab

Research-friendly



PRINCETON UNIVERSITY

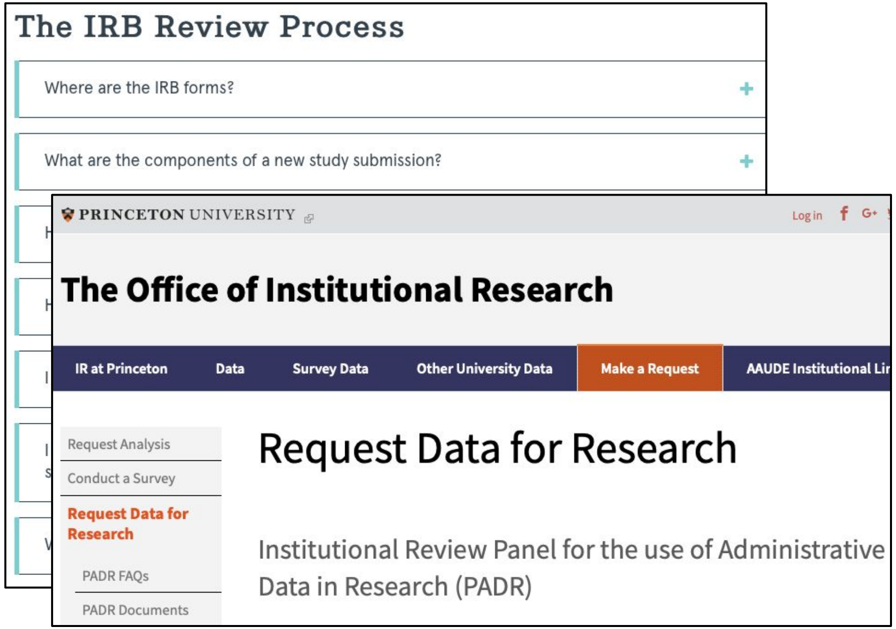
Office of Information Technology

Home Services

Enabling research, teaching and communication at Princeton

OIT is committed to providing technology services that enable Princeton to achieve its mission to advance learning through scholarship, research, and teaching of unsurpassed quality.

Existing mechanisms



The IRB Review Process

Where are the IRB forms? +

What are the components of a new study submission? +

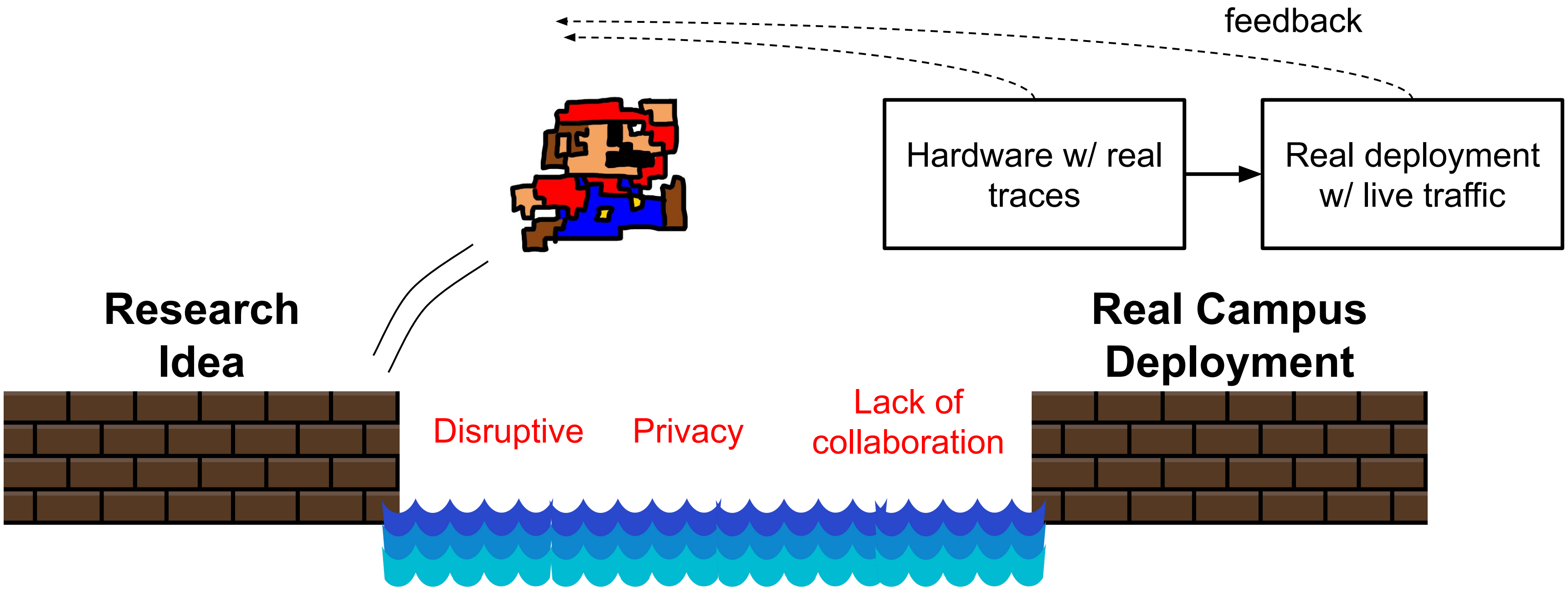
PRINCETON UNIVERSITY

The Office of Institutional Research

Request Data for Research

Institutional Review Panel for the use of Administrative Data in Research (PADR)

Jumping The Research Chasm



feedback

Hardware w/ real traces

Real deployment w/ live traffic

Research Idea

Real Campus Deployment

Disruptive

Privacy

Lack of collaboration

1. Sharing Our Experience

Less Disruptive

Preserve Privacy

More Collaboration with OIT

2. Successful Deployments

live traffic anonymization

*ONTAS, *Flow*

enablers

monitor RTT

OS fingerprint

Precision, ConQuest, P4-RTT, P4Of, P4-DNS

passive traffic analytics

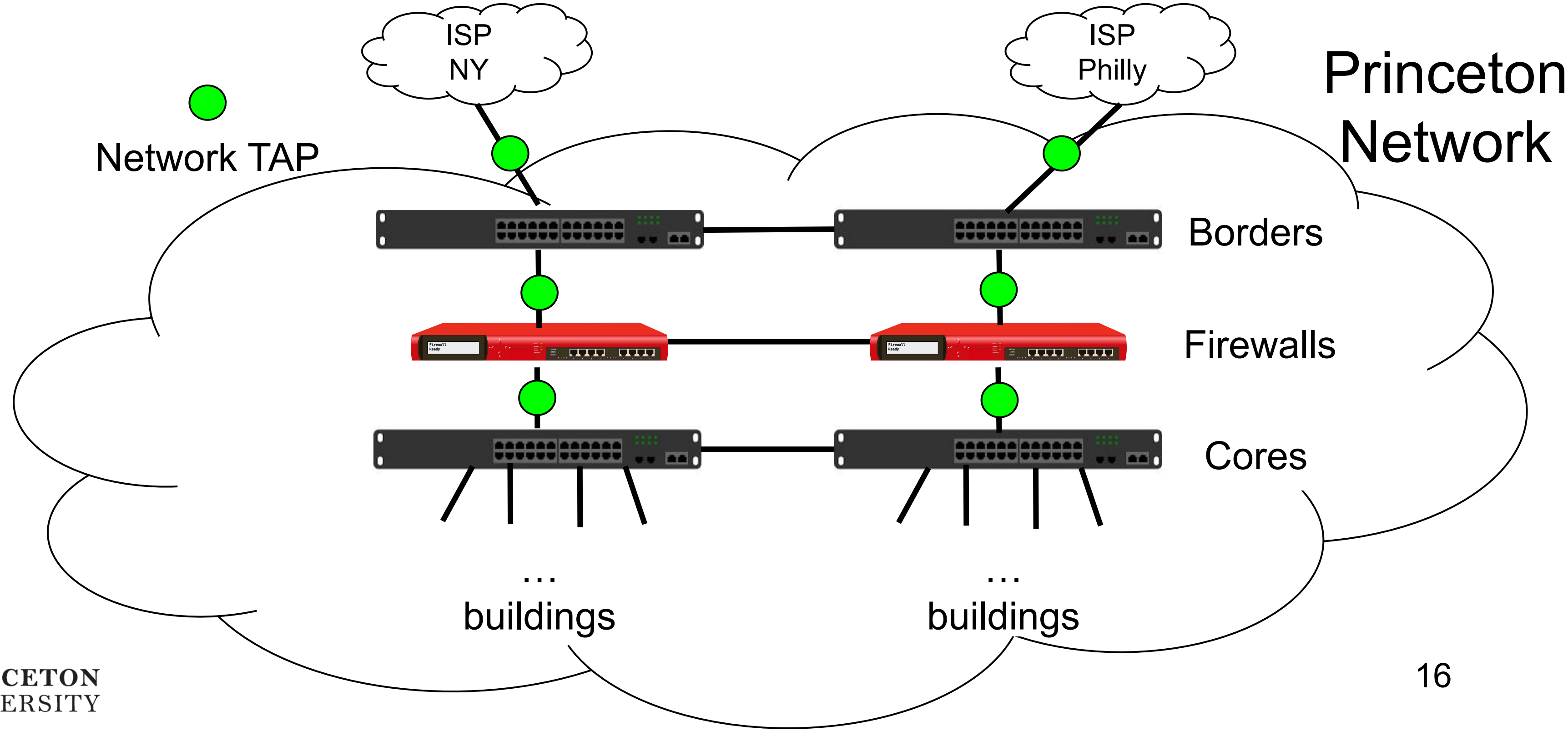
PINOT

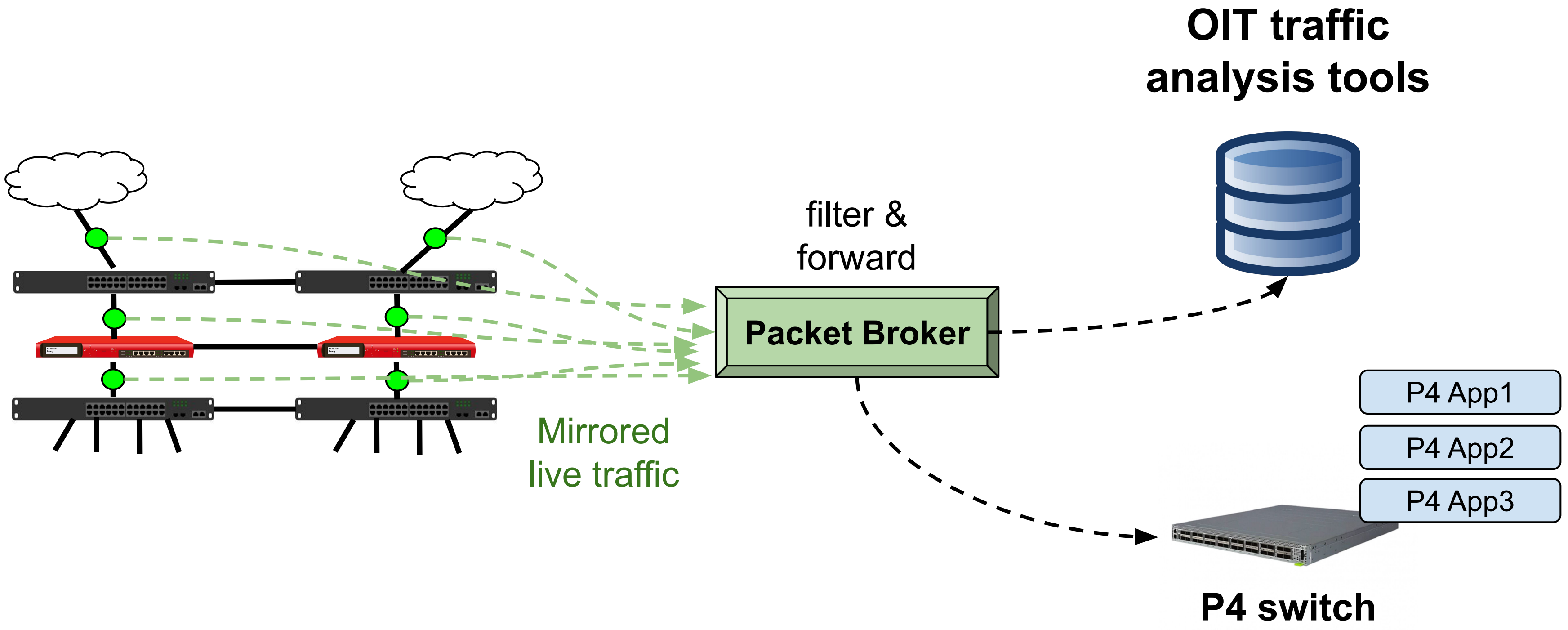
active traffic experiment

Becoming Less Disruptive

Less Disruptive	**Preserve Privacy**	**More Collaboration with OIT**
● Work w/ mirrored traffic ● Passive monitoring as a “gateway drug”		

Traffic Mirroring




OIT traffic analysis tools
filter & forward
Packet Broker
Mirrored live traffic
P4 App1
P4 App2
P4 App3
P4 switch

Working with Mirrored Traffic

- Low risk
  - Little or no disruption
- High return
  - Real-time traffic analysis is a nice “gateway drug”

Some Tips on Mirrored Traffic

Test Access Point

Switch Port ANalyzer

- TAP is better than port mirroring (SPAN)
  - SPAN burns a port and resources on a switch

- Modern *packet broker systems* can
  - Apply filtering policies
  - Remove/mask payload
  - Remove duplicate packets

Preserving User Privacy

Less Disruptive

- Work w/ mirrored traffic
- Passive monitoring as a "gateway drug"

Preserve Privacy

- IRB
- Data anonymization

More Collaboration with OIT

PRINCETON UNIVERSITY

Navigating Campus Traffic Data Access

Institutional Review Board (IRB)

Rights, Privacy, Welfare
of Human Subjects

Institutional Review Panel for the use of Administrative Data in Research (PADR)

Feasibility, Value,
Risk, and Compliance

Prepping IRB Applications

State that you will remove/anonymize PII

- MAC and <Your Institute> IP addresses will be anonymized
- Payload will be removed
- If not, operator will run scripts/programs and provide aggregated results

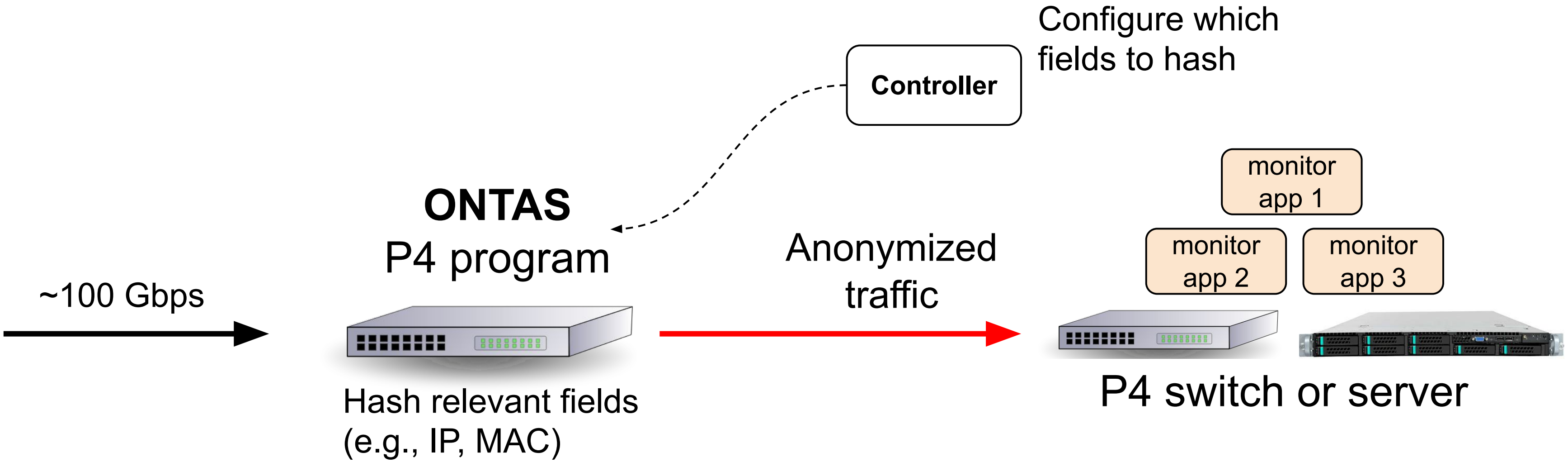
Show you will take good care of the data

- Data will be stored and processed at machines managed by IT staff

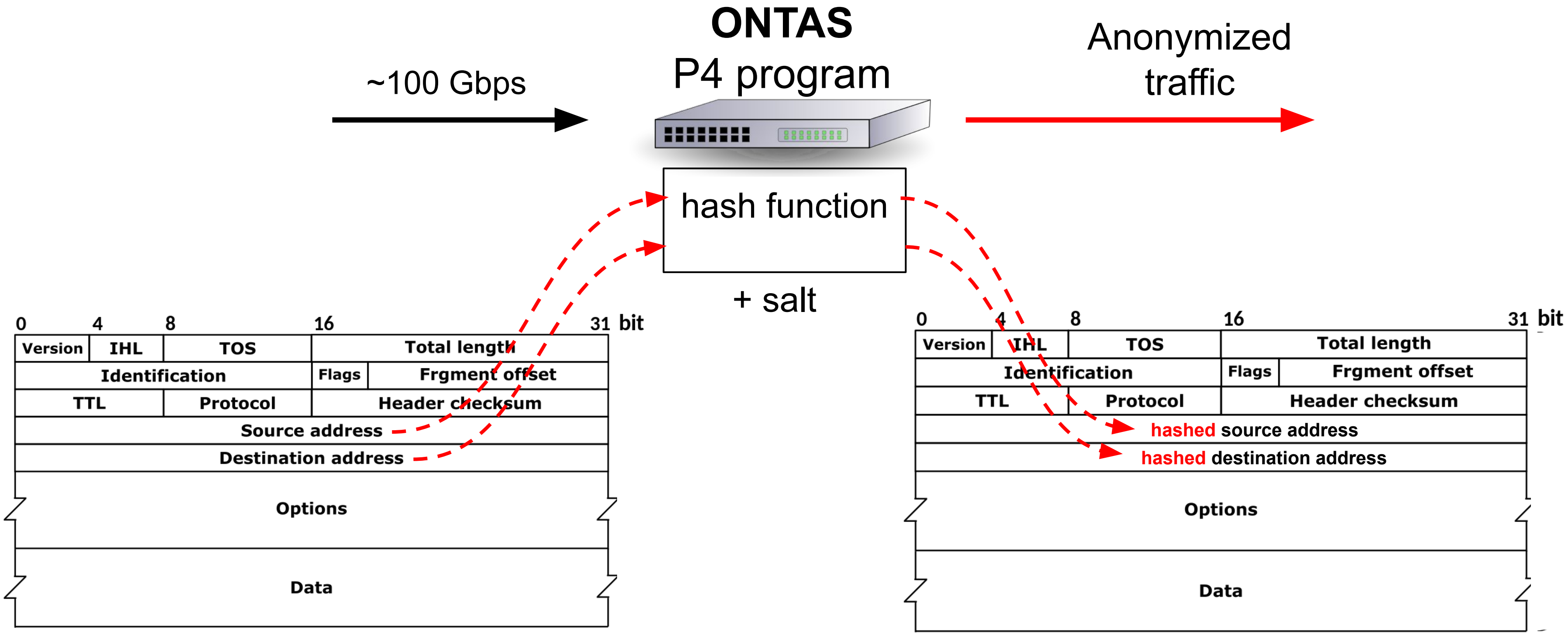
Offline Data Anonymization

- CAIDA’s best practices and offline tools
    - https://www.caida.org/projects/predict/anonymization/

Anonymize Live Traffic



Hyojoon Kim et. al., "**ONTAS: Flexible and Scalable Online Network Traffic Anonymization System.**"
2019 SIGCOMM Workshop on Network Meets AI & ML



~100 Gbps

ONTAS
P4 program

Anonymized traffic

hash function

+ salt

0	4	8	16	31 bit
Version	IHL	TOS	Total length	
Identification			Flags	Frgment offset
TTL		Protocol	Header checksum	
Source address				
Destination address				
Options				
Data				

0	4	8	16	31 bit
Version	IHL	TOS	Total length	
Identification			Flags	Frgment offset
TTL		Protocol	Header checksum	
hashed source address				
hashed destination address				
Options				
Data				

- **Online (not offline)**
- **Line-rate**
- **Customizable** (e.g., select IPs, preserve prefix, etc)

PRINCETON UNIVERSITY

Collaboration With The IT Group

Less Disruptive	Preserve Privacy	More Collaboration with OIT
● Work w/ mirrored traffic ● Passive monitoring as a “gateway drug”	● IRB ● Data anonymization	● Tackle problems that matter ● Joint positions

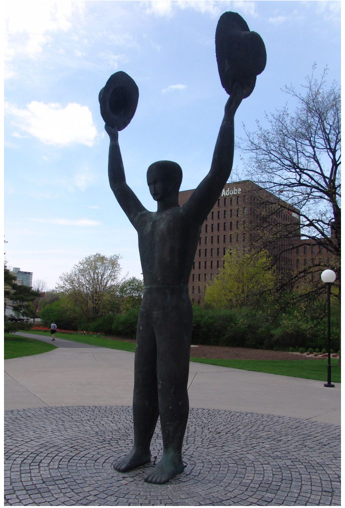
Collaboration with OIT

Find problems that matter

- Operator is tired of anonymizing traffic for researchers. Harder for live traffic.
  - ONTAS: traffic anonymization
- Occasional packet drops at switch, but don't know why.
  - ConQuest: Queue monitoring
- Having latency problems. Where is the bottleneck?
  - P4-RTT: Measure RTT at different vantage points

Joint position (CS & OIT)

Research projects



Access to campus network

PRINCETON UNIVERSITY

Successful Deployments

Real-time
OS fingerprinting

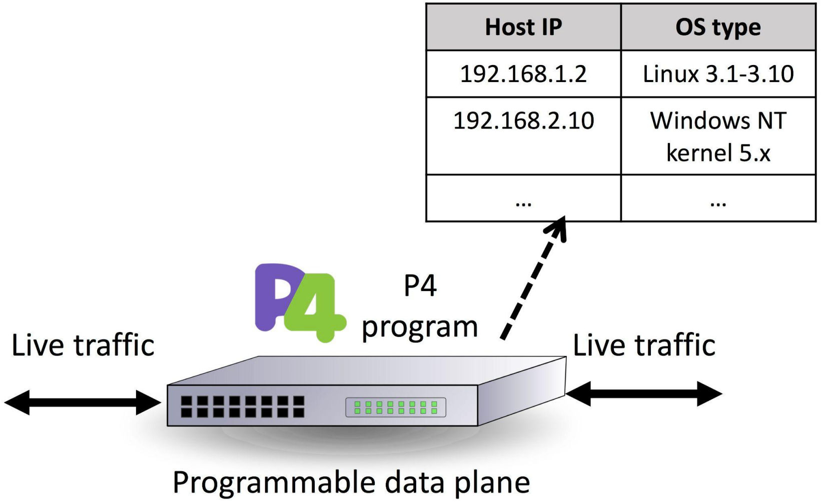
Continuous
RTT monitoring

P40f: OS Fingerprinting with P4

- Fingerprint OS type in the data plane
- Higher abstraction than IP addresses

Write policies based on OS type

- “Block all traffic from Windows XPs”
- “Rate-limit traffic to/from Echo Dot”
- “Monitor OS distribution in real time”
- …



Host IP	OS type
192.168.1.2	Linux 3.1-3.10
192.168.2.10	Windows NT kernel 5.x
…	…

The *p0f* Tool

- Each OS uses a unique combination of IP/TCP header values

- *p0f* signature

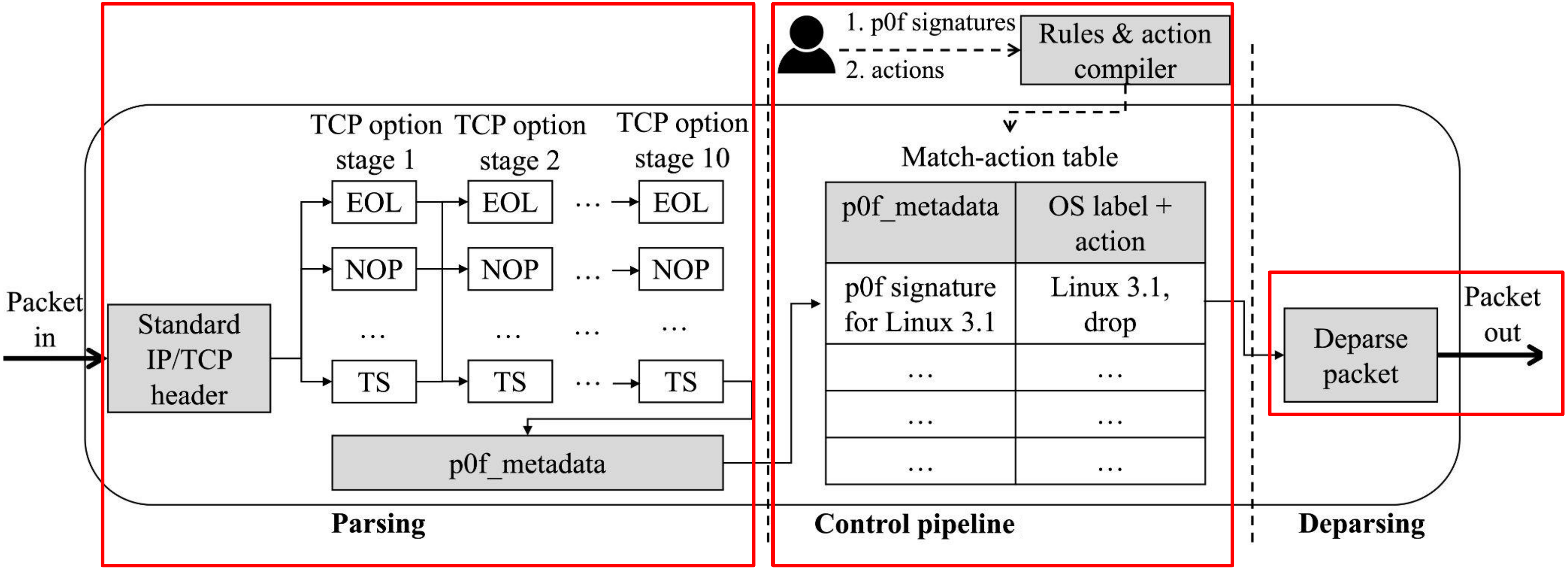
  IPv4/IPv6 : TTL : IP option length : MSS : (window size,scale) :
  TCP option layout : quirks : pclass

- Example: Linux v3.11 or higher

  * : 64 : 0 : * : (20,10) : (mss,sok,ts,nop,ws) : (df,id+) : 0

The p0f tool cannot run against
live traffic with high data rate

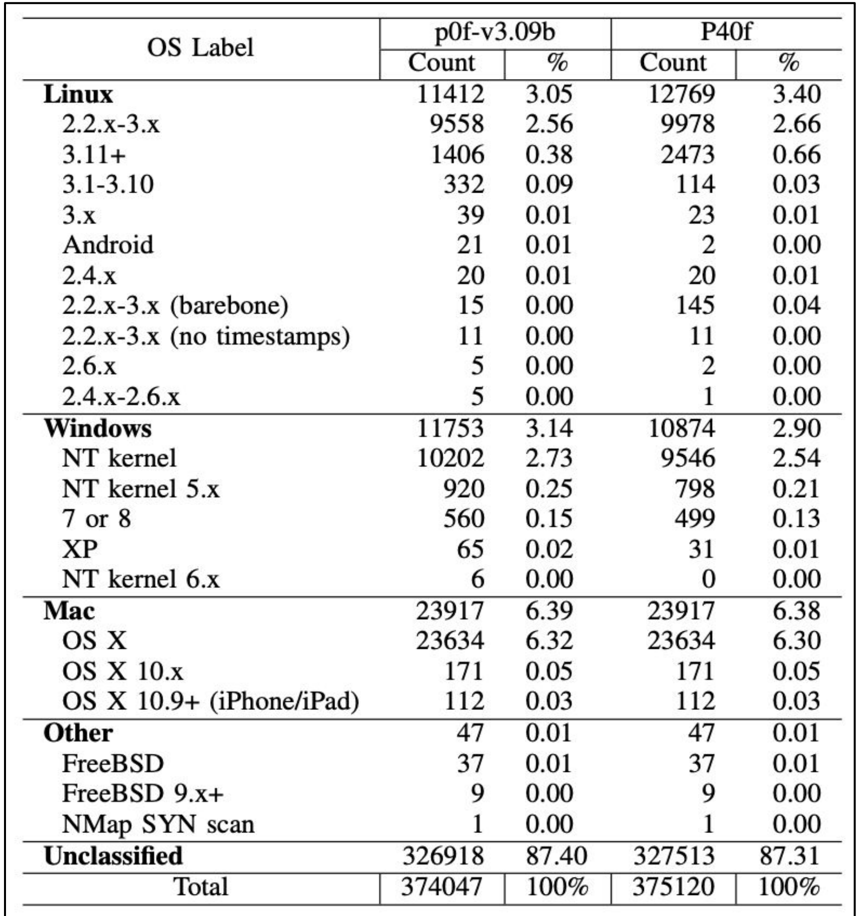
P40f: Let’s Do This in the Switch



PRINCETON UNIVERSITY

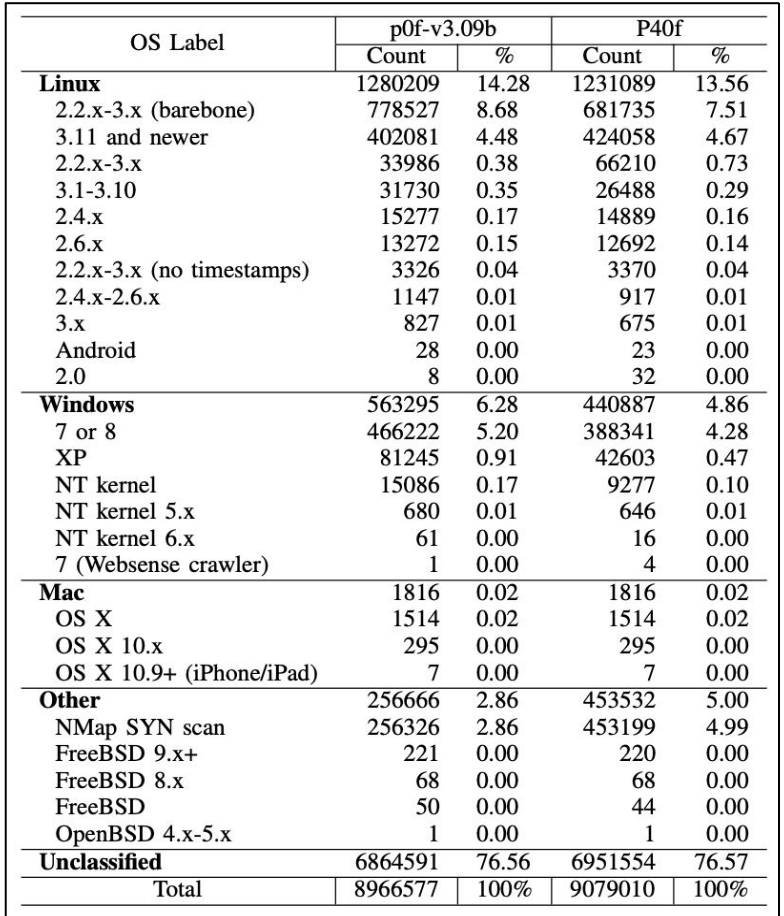
Against 3-hour campus trace

Internal hosts



OS Label	p0f-v3.09b		P40f	
	Count	%	Count	%
Linux	11412	3.05	12769	3.40
2.2.x-3.x	9558	2.56	9978	2.66
3.11+	1406	0.38	2473	0.66
3.1-3.10	332	0.09	114	0.03
3.x	39	0.01	23	0.01
Android	21	0.01	2	0.00
2.4.x	20	0.01	20	0.01
2.2.x-3.x (barebone)	15	0.00	145	0.04
2.2.x-3.x (no timestamps)	11	0.00	11	0.00
2.6.x	5	0.00	2	0.00
2.4.x-2.6.x	5	0.00	1	0.00
Windows	11753	3.14	10874	2.90
NT kernel	10202	2.73	9546	2.54
NT kernel 5.x	920	0.25	798	0.21
7 or 8	560	0.15	499	0.13
XP	65	0.02	31	0.01
NT kernel 6.x	6	0.00	0	0.00
Mac	23917	6.39	23917	6.38
OS X	23634	6.32	23634	6.30
OS X 10.x	171	0.05	171	0.05
OS X 10.9+ (iPhone/iPad)	112	0.03	112	0.03
Other	47	0.01	47	0.01
FreeBSD	37	0.01	37	0.01
FreeBSD 9.x+	9	0.00	9	0.00
NMap SYN scan	1	0.00	1	0.00
Unclassified	326918	87.40	327513	87.31
Total	374047	100%	375120	100%

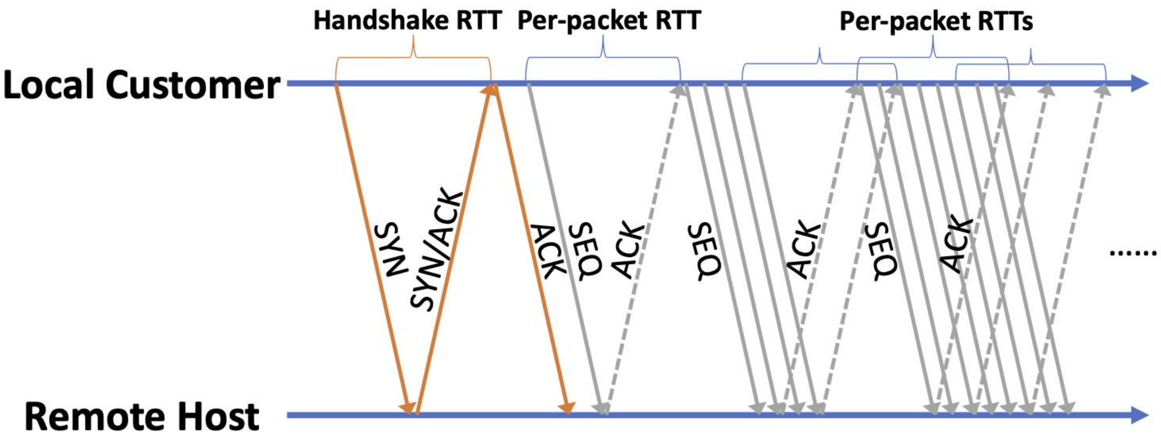
External hosts



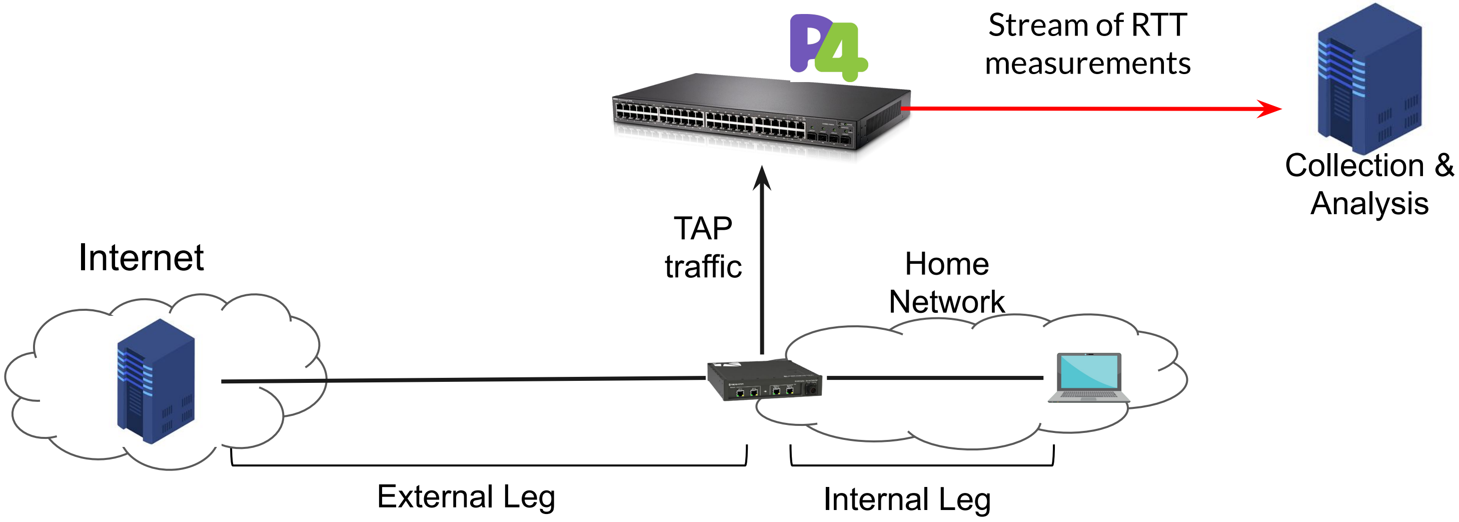
OS Label	p0f-v3.09b		P40f	
	Count	%	Count	%
Linux	1280209	14.28	1231089	13.56
2.2.x-3.x (barebone)	778527	8.68	681735	7.51
3.11 and newer	402081	4.48	424058	4.67
2.2.x-3.x	33986	0.38	66210	0.73
3.1-3.10	31730	0.35	26488	0.29
2.4.x	15277	0.17	14889	0.16
2.6.x	13272	0.15	12692	0.14
2.2.x-3.x (no timestamps)	3326	0.04	3370	0.04
2.4.x-2.6.x	1147	0.01	917	0.01
3.x	827	0.01	675	0.01
Android	28	0.00	23	0.00
2.0	8	0.00	32	0.00
Windows	563295	6.28	440887	4.86
7 or 8	466222	5.20	388341	4.28
XP	81245	0.91	42603	0.47
NT kernel	15086	0.17	9277	0.10
NT kernel 5.x	680	0.01	646	0.01
NT kernel 6.x	61	0.00	16	0.00
7 (Websense crawler)	1	0.00	4	0.00
Mac	1816	0.02	1816	0.02
OS X	1514	0.02	1514	0.02
OS X 10.x	295	0.00	295	0.00
OS X 10.9+ (iPhone/iPad)	7	0.00	7	0.00
Other	256666	2.86	453532	5.00
NMap SYN scan	256326	2.86	453199	4.99
FreeBSD 9.x+	221	0.00	220	0.00
FreeBSD 8.x	68	0.00	68	0.00
FreeBSD	50	0.00	44	0.00
OpenBSD 4.x-5.x	1	0.00	1	0.00
Unclassified	6864591	76.56	6951554	76.57
Total	8966577	100%	9079010	100%

P4-RTT

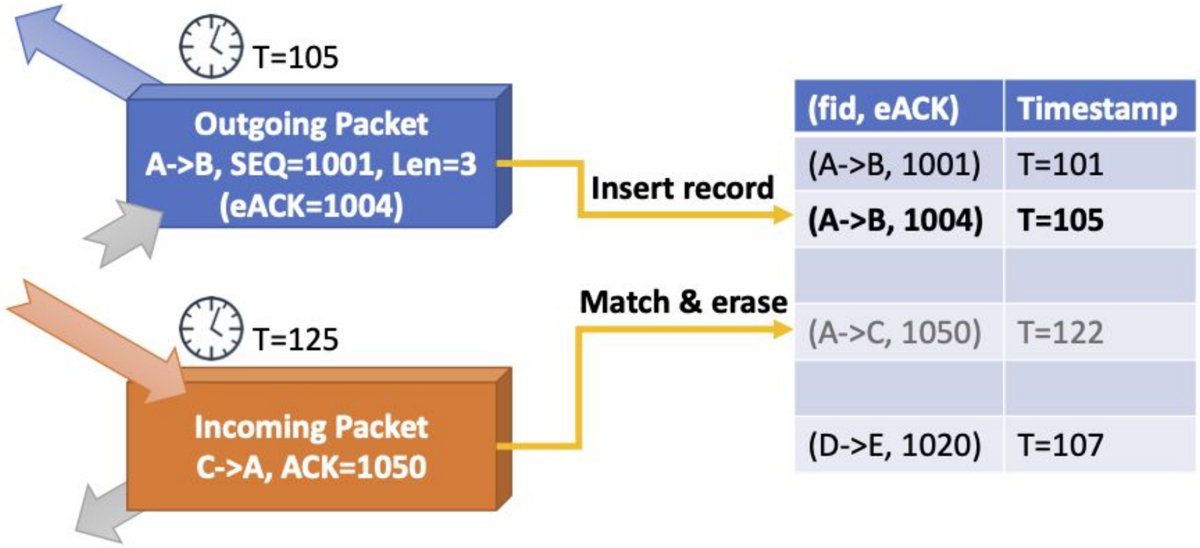
Continuous round trip time monitoring *beyond* the TCP handshake



How P4-RTT Operates



Register as Hash Table



(fid, eACK)	Timestamp
(A->B, 1001)	T=101
(A->B, 1004)	**T=105**
(A->C, 1050)	T=122
(D->E, 1020)	T=107

Memory limit

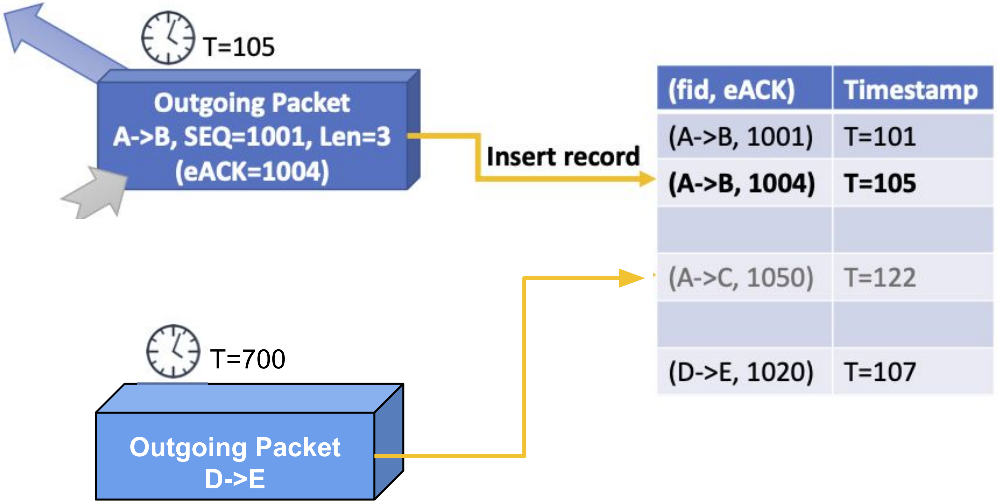
Hash collisions

TCP quirks

***JOIN* of *outgoing* and *incoming* packet streams in the data plane**

Overcoming The Memory Limit

Many TCP packets don’t receive a corresponding ACK

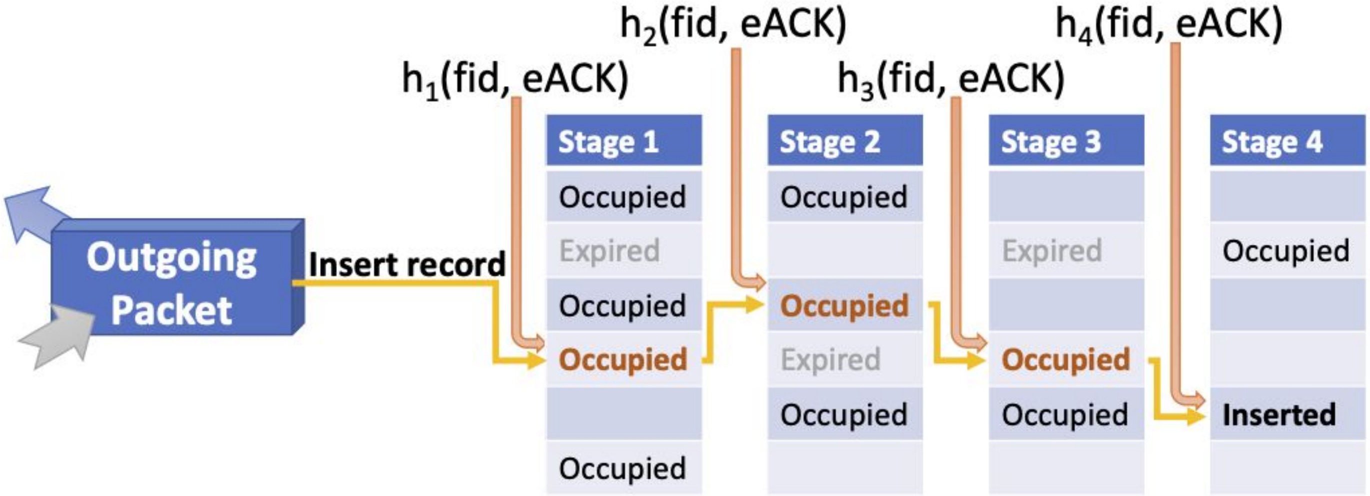


(fid, eACK)	Timestamp
(A->B, 1001)	T=101
(A->B, 1004)	**T=105**
(A->C, 1050)	T=122
(D->E, 1020)	T=107

Lazily expire entries with a *threshold*

Threshold: 99th percentile RTT (500ms)

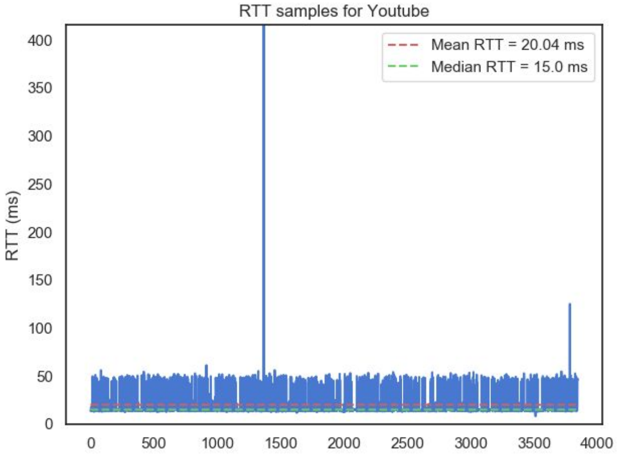
Multi-stage Hash Table (Registers)



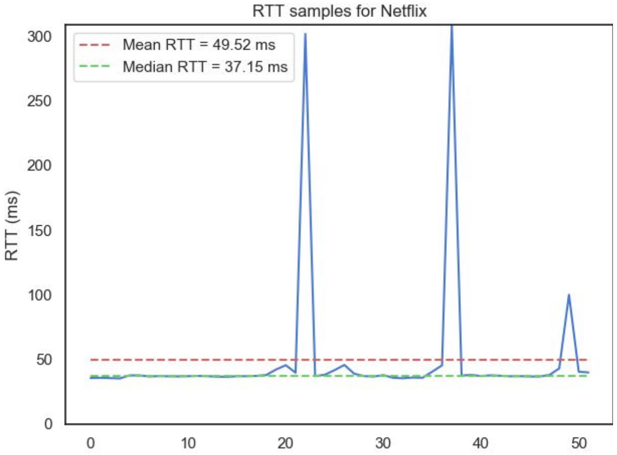
Overcomes the memory limit per register

Overcomes RTT sample loss due to hash collision

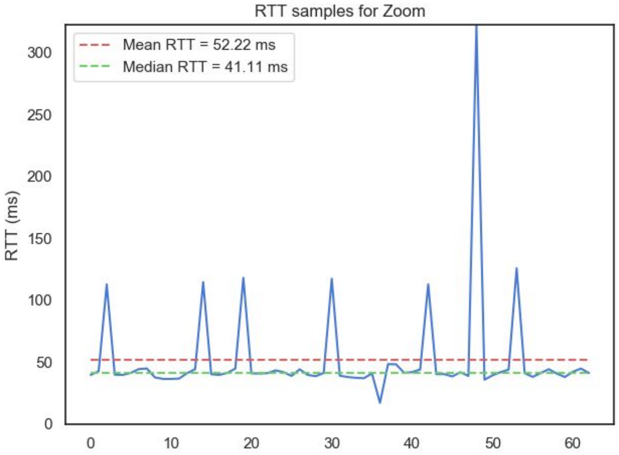
Per Application RTT



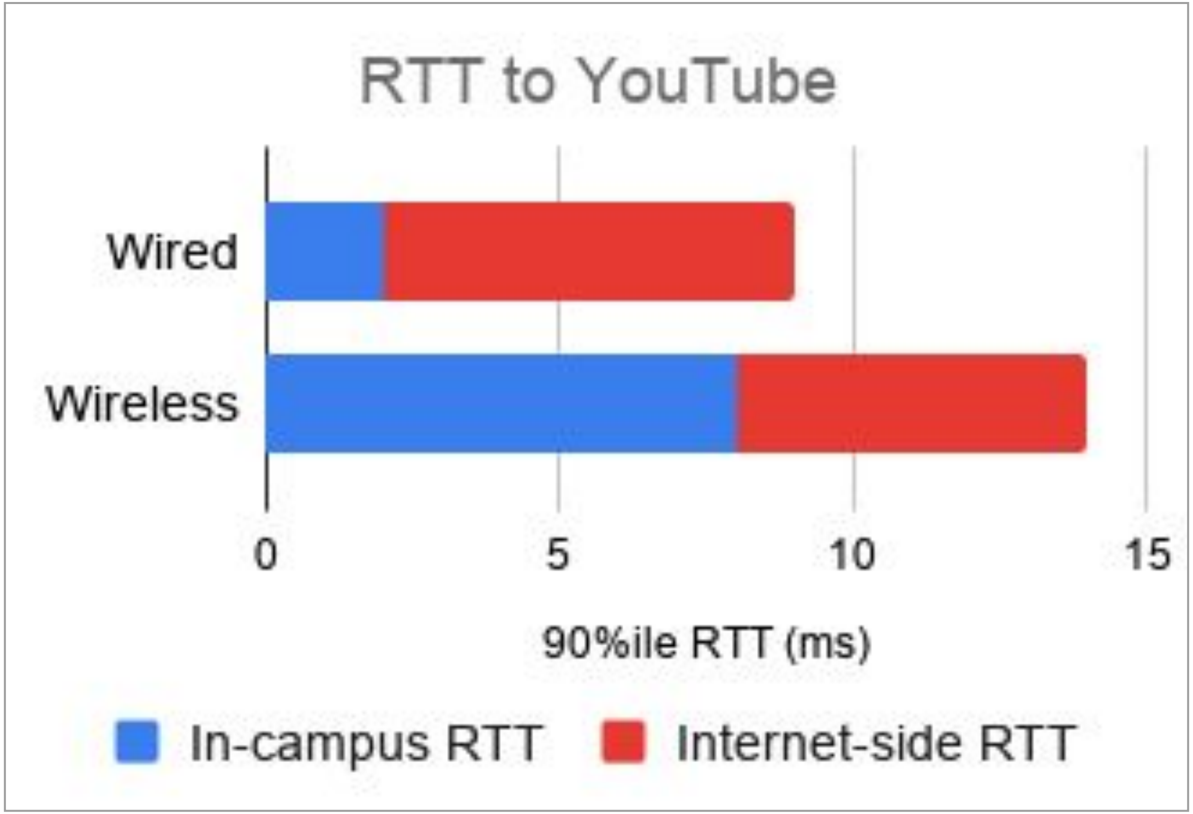
Youtube



Netflix



Zoom



Impact of wired vs. wireless infrastructure on 90%ile RTT to YouTube

Summary

Jumping The Research Chasm

- **Less disruptive**
  - Passive traffic monitoring
- **Preserve user privacy**
  - IRB prep (and more)
  - Anonymization tools
- **More collaboration**
  - Joint position
  - problems that matter

Real Deployment Successes

- **ONTAS**
  - anonymized data collection
- **P40f**
  - Real-time OS fingerprinting
- **P4 RTT**
  - Continuous RTT monitoring

Experience-Driven Research

Experience-Driven Research on Programmable Networks

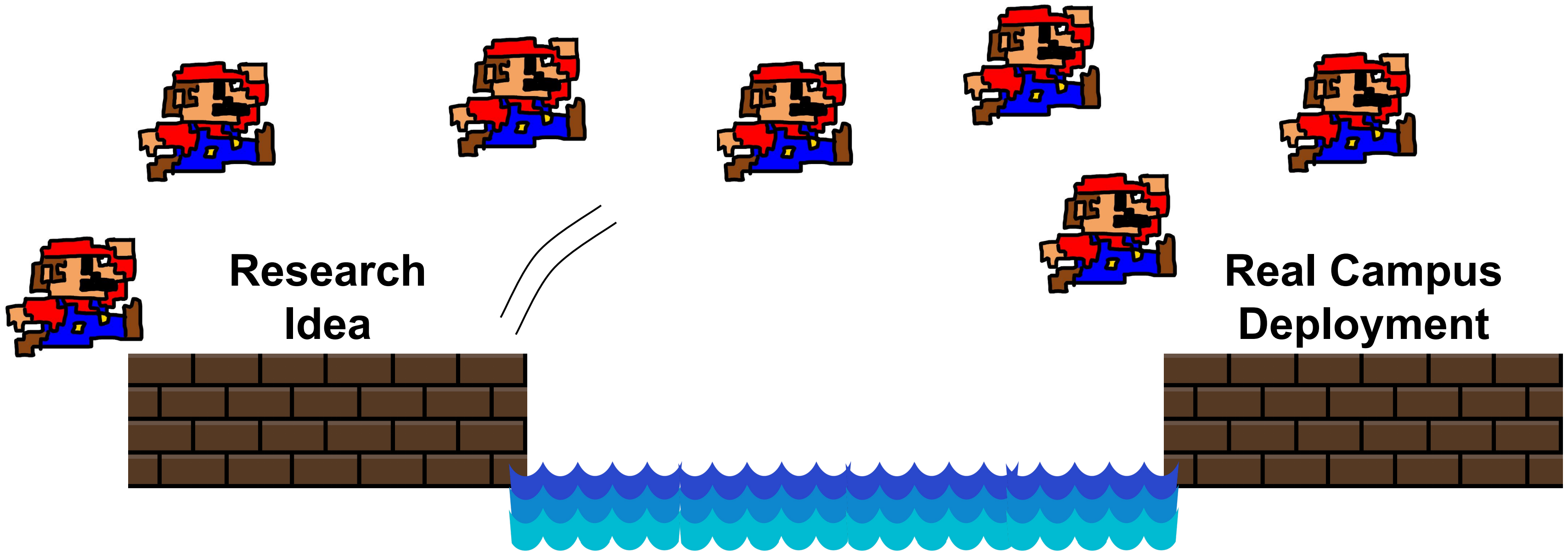
Hyojoon Kim, Xiaoqi Chen, Jack Brassil, and Jennifer Rexford

ACM SIGCOMM Computer Communications Review. January 2021.

Github repo for our P4 projects:

https://github.com/Princeton-Cabernet/p4-projects

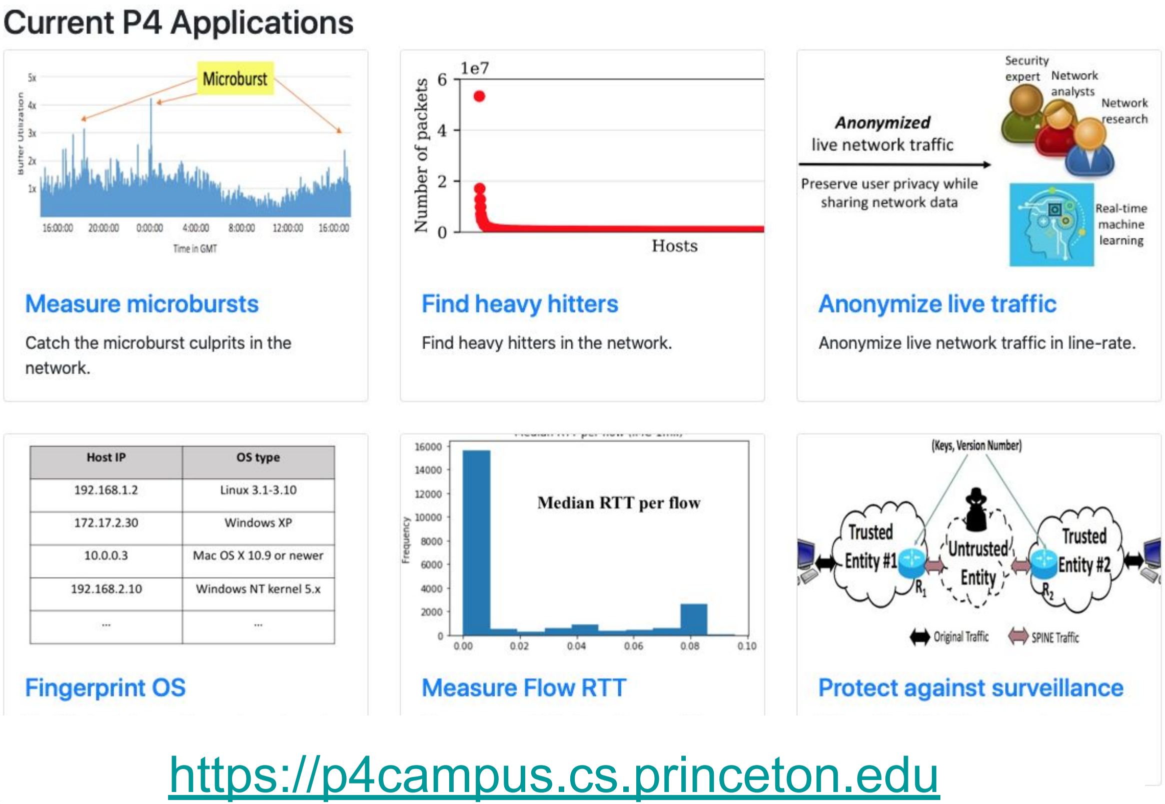
Please Join Our Effort!



Research Idea

Real Campus Deployment

More Campus Applications



Current P4 Applications

Measure microbursts

Catch the microburst culprits in the network.

Find heavy hitters

Find heavy hitters in the network.

Anonymize live traffic

Anonymize live network traffic in line-rate.

Host IP	OS type
192.168.1.2	Linux 3.1-3.10
172.17.2.30	Windows XP
10.0.0.3	Mac OS X 10.9 or newer
192.168.2.10	Windows NT kernel 5.x
...	...

Fingerprint OS

Measure Flow RTT

Protect against surveillance

...

More from you!

https://p4campus.cs.princeton.edu

Thank You!

P4 Campus Website:

https://p4campus.cs.princeton.edu

Reach me at:

hyojoonk@cs.princeton.edu