



Reflections on Active Network Measurements in Academia

Tobias Fiebig

Max Planck Institute for Informatics

Email: contact@as59645.net

Phone: [REDACTED]

Network Measurement



- Network Measurements:
The thing we do
- Important tool for academics
(getting papers) and practitioners
(getting something useful to
improve protocols)
- Come in active or passive
- Especially active ones are difficult
to do well



© Constanze Dietrich



Email 1991

- Mail: RFC821, RFC822
- DNS: RFC1032, RFC1033, RFC1034, RFC1035
- If you like X.400 just a handful more

Email 2022

- Mail: ~500 RFCs
- DNS: ~300 RFCs
- HTTP (MTA-STS): ~Too Many
- TLS/Cert: Yes
- IPv4/IPv6: Welcome to the MTU world!

Related Publication:

Holzbauer, F., Ullrich, J., Lindorfer, M., & Fiebig, T. (2022). *Not that Simple: Email Delivery in the 21st Century*. In USENIX ATC '22, USENIX ATC. Carlsbad, CA, USA: USENIX Association.

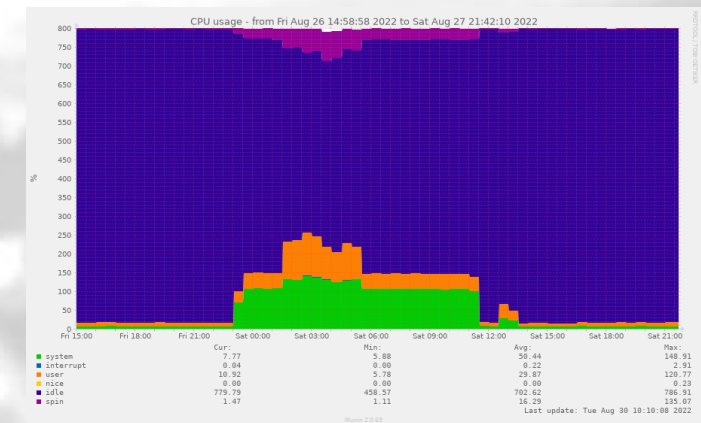
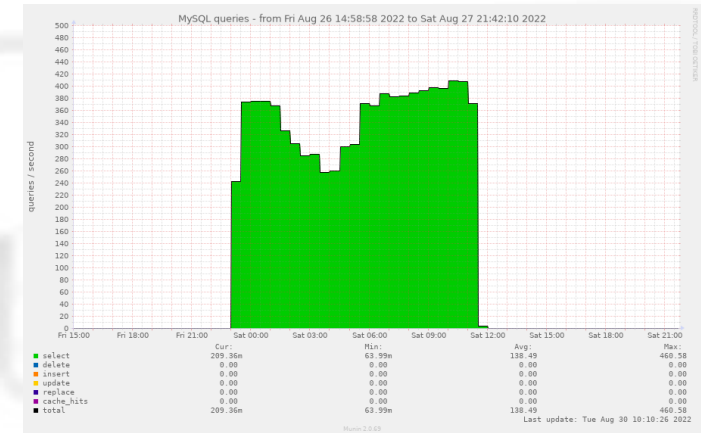
Writing Reliable Measurement Software



Writing Reliable Measurement Software



Current Status: **CRITICAL** (for 0d 0h 0m 7s)
Status Information: connect to address mail.aperture-labs.org and port 25: Connection refused
SMTP CRITICAL - 0.003 sec. response time



Writing Reliable Measurement Software

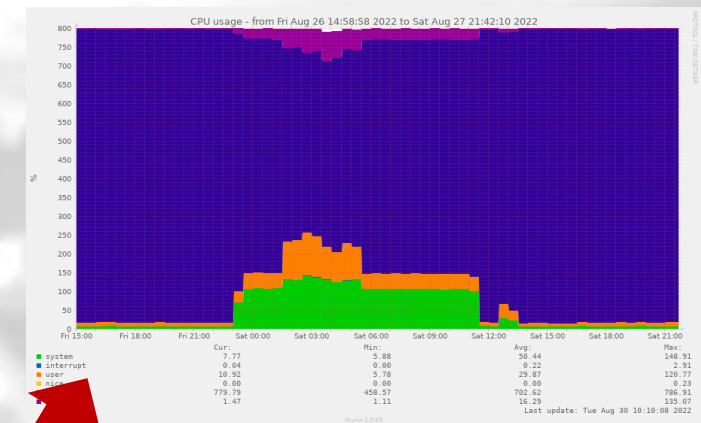
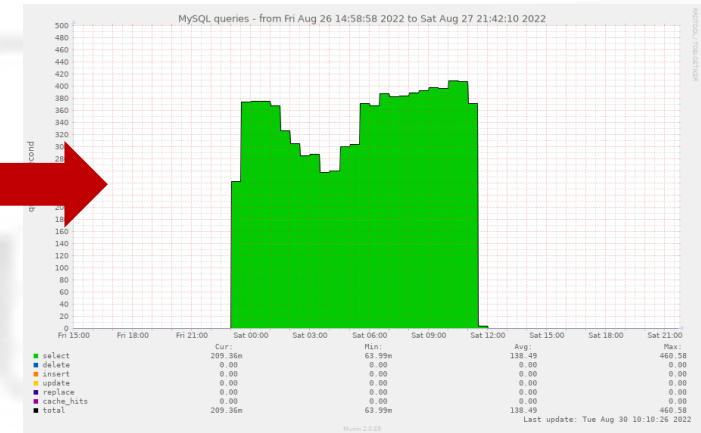


No SMTP



Current Status: **CRITICAL** (for 0d 0h 0m 7s)
Status Information: connect to address mail.aperture-labs.org and port 25: Connection refused
SMTP CRITICAL - 0.003 sec. response time

MySQL @ 400 qps



OpenSMTPd @ 100% CPU

Writing Reliable Measurement Software



- The Internet is full of corner cases: Account for all!
- Be aware of all (unwritten) rules of your protocol of choice.
- Implement a *reliable* measurement tool, ideally reusing as much existing (tested) software as possible
- Be, in general, a good and experienced programmer able to write software able to interact with all systems on the Internet (not breaking them even if you just do standard-compliant things when interacting with them)
- Version Control, tests... proper development!

Running Reliable Network Measurements



```
root@msrmnt.example.com:/opt/yolo-colo# docker compose up -d
root@msrmnt.example.com:/opt/yolo-colo# docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
33ecc4cc2bfb	yolo-colo/scan	"/docker-entrypoint...."	21 hours ago	Up 21 hours	25/tcp	scan_1
8b86f724e188	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	rdns_1
15d771655355	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	adns_1

Running Reliable Network Measurements



```
root@msrmnt.example.com:/opt/yolo-colo# docker compose up -d
root@msrmnt.example.com:/opt/yolo-colo# docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
33ecc4cc2bfb	yolo-colo/scan	"/docker-entrypoint...."	21 hours ago	Up 21 hours	25/tcp	scan_1
8b86f724e188	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	rdns_1
15d771655355	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	adns_1

Running Reliable Network Measurements



```
root@msrmnt.example.com:/opt/yolo-colo# docker compose up -d
root@msrmnt.example.com:/opt/yolo-colo# docker ps
```

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
33ecc4cc2bfb	yolo-colo/scan	"/docker-entrypoint...."	21 hours ago	Up 21 hours	25/tcp	scan_1
8b86f724e188	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	rdns_1
15d771655355	yolo-colo/rdns	"/bin/sh -c /entrypo..."	21 hours ago	Up 21 hours	53/udp	adns_1

- No TCP for DNS
- Maybe rdns_1 stops resolving
- Maybe DNS is not delegated to adns_1
- Maybe rdns_1 just stubs vs. q1/q8/q9 and we do not know if queries we see at adns_1 just come from... us...



© Constanze Dietrich

Running Reliable Network Measurements



- Be an experienced SysOp
- Know about all the things involved (and available tools!)
- Monitor your stack
 - Historic for bottlenecks (you may just measure your IOPS)
 - Real-time for reliability
- Have an end-to-end understanding
- Make sure the setup is self-contained

Running Ethical Measurements



- Consider all possible unintended harms
 - “Yeah, we know, the Internet is made from duct tape and bubble gum, and this would be an issue; So we just don’t talk about it!”
- Get ethics approval
- Do probe attribution
 - rDNS, RIR Whois, running webserver etcetcetc.
- Handle 24/7 opt-out and abuse
- Have a maintained block-list

The PhD we Need



- Thoroughly understand the protocol stack they are measuring, including operational lore and lived experience since the inception of these protocols
- Be versed in the domain of available implementations to identify components they can use to construct their measurement setup
- Be experienced programmers and versed in software development in general to follow development best practices and produce tested and reliable code
- Be experienced system administrators—or have such institutional support—to setup the measurement system, including all basic services the system depends on, including historic and real-time monitoring of all components

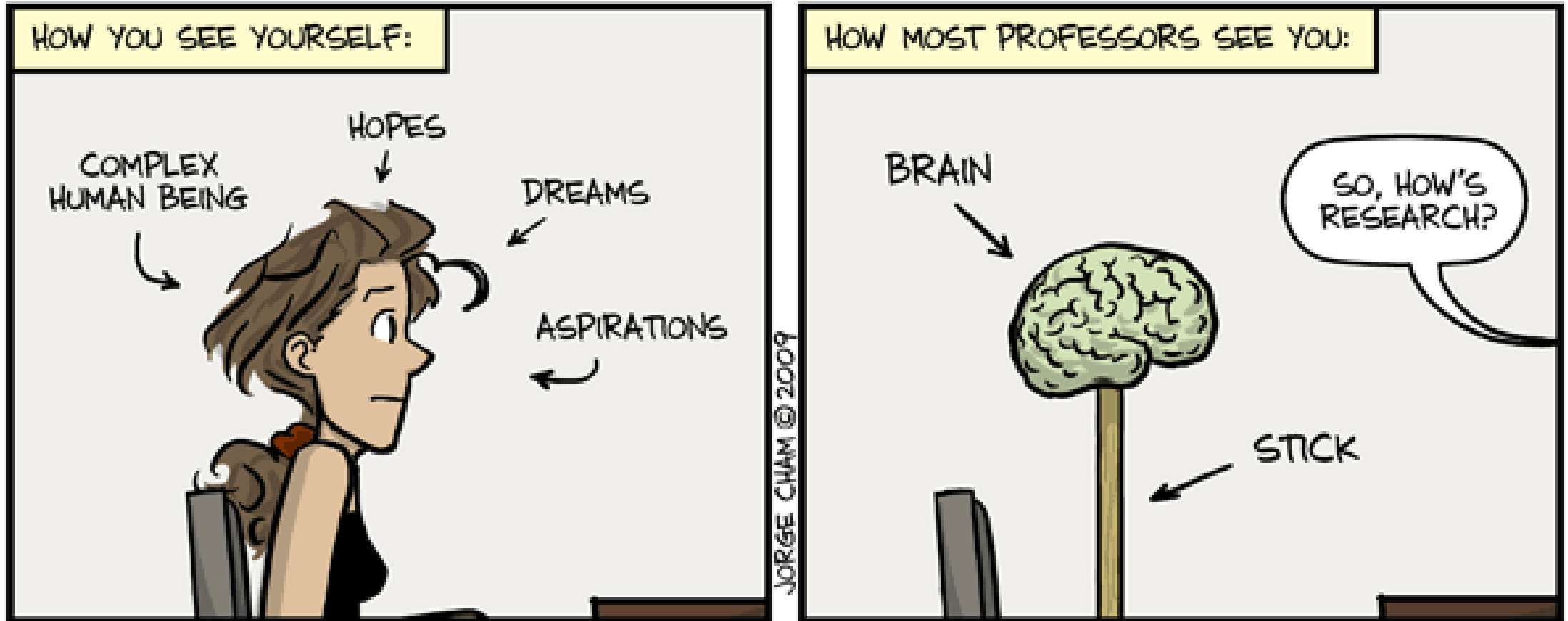


- Thoroughly understand the protocol stack they are measuring, including operational lore and lived experience since the inception of these protocols

**This is not a PhD student;
This is a whole IT department.**

- Be versed in the domain of available implementations to identify components they can use to construct their measurement setup
- Be experienced programmers and versed in software development in general to follow development best practices and produce tested code
- Be experienced system administrators—or have such institutional support—to setup the measurement system, including all basic services the system depends on, including historic and real-time monitoring of all components

The PhD we Need





WWW.PHDCOMICS.COM

The Reality of a PhD



- 4-8 Years
- ~4 'Top-Tier' papers
- New research advancing the field
- Embedded in related work (Meaning: You have to read it!)
- Joining after a bachelor's degree (US) or master's degree (most-other-ish)

	MARRIAGE	vs. The Ph.D.
		
	<u>Marriage</u>	<u>Ph.D.</u>
Typical Length:	7.5 years	7 years
Begins with:	A proposal	A thesis proposal
Culminates in a ceremony where you walk down an aisle dressed in a gown:	✓	✓
Usually entered into by:	Foolish young people in love	Foolish young people without a job
50% end in:	Bitter divorce	Bitter remorse
Involves exchange of:	Vows	Know-how
Until death do you part?	If you're lucky	If you're lazy



JORGE CHAM © 2010

WWW.PHDCOMICS.COM

The Reality of a PhD



- 4-8 Years
- ~4 'Top-Tier' papers
- New research advancing the field
- Embedded in related work (Meaning: You have to read it!)
- Joining after a bachelor's degree (US) or master's degree (most-other-ish)
- **First paper should be under submission after ~1 year!**

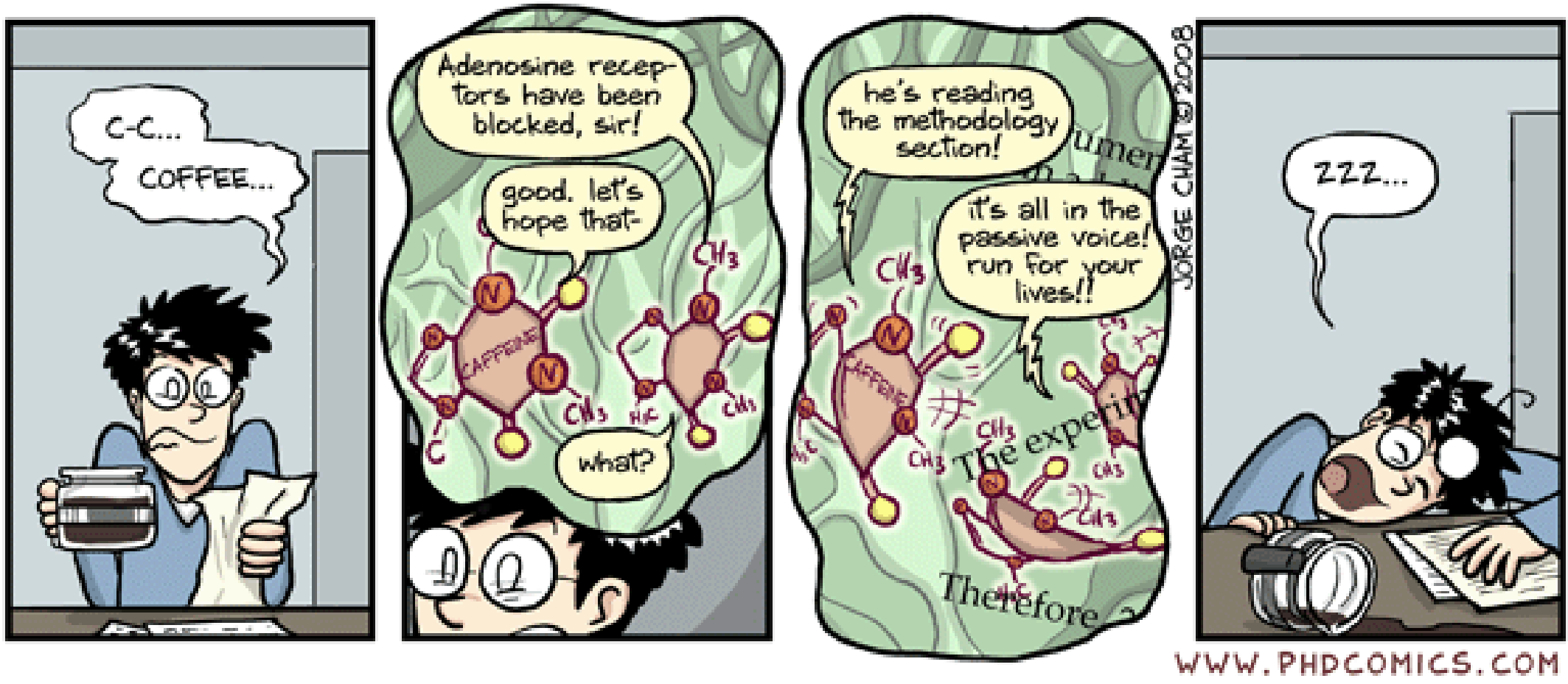
	MARRIAGE	vs. The Ph.D.
		
	<u>Marriage</u>	<u>Ph.D.</u>
Typical Length:	7.5 years	7 years
Begins with:	A proposal	A thesis proposal
Culminates in a ceremony where you walk down an aisle dressed in a gown:	✓	✓
Usually entered into by:	Foolish young people in love	Foolish young people without a job
50% end in:	Bitter divorce	Bitter remorse
Involves exchange of:	Vows	Know-how
Until death do you part?	If you're lucky	If you're lazy

JORGE CHAM © 2010

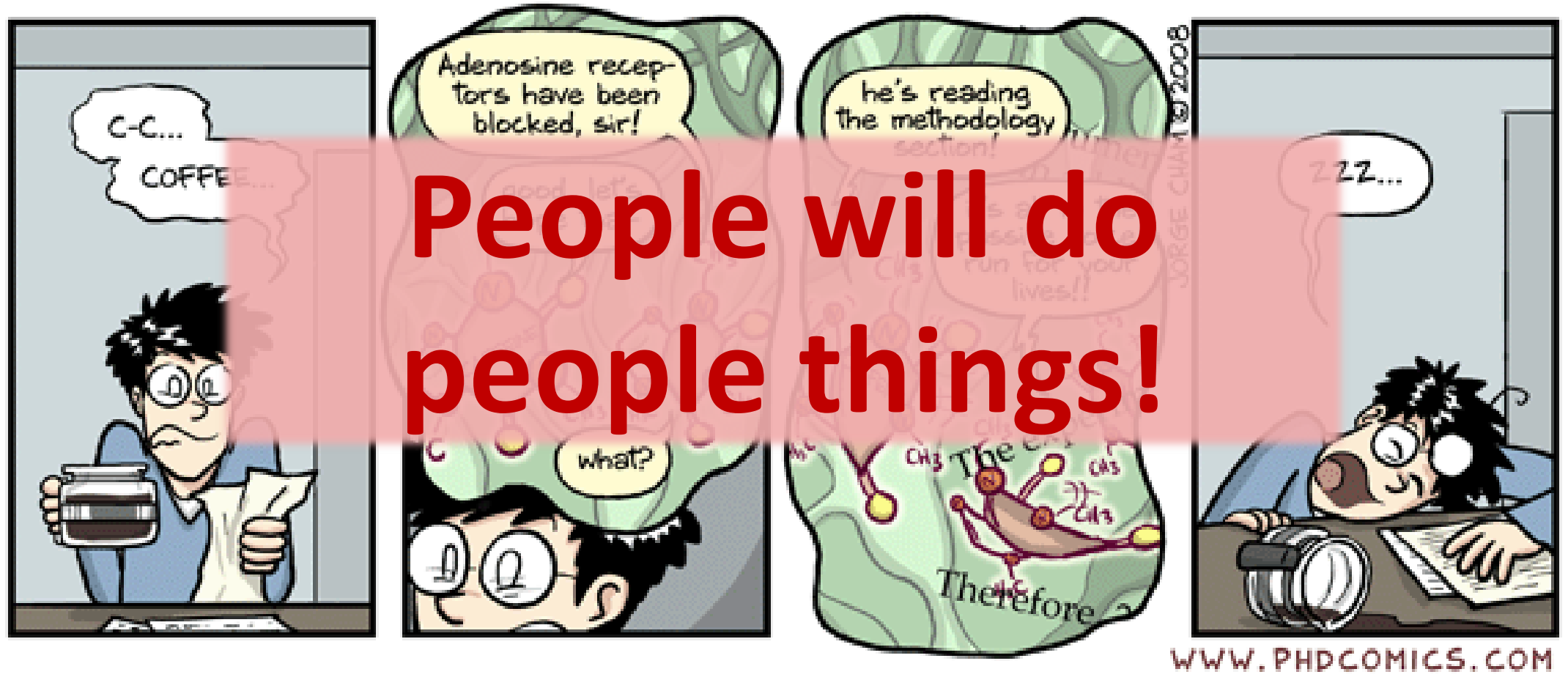
WWW.PHDCOMICS.COM

Source:
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=1296>

The Reality of a PhD

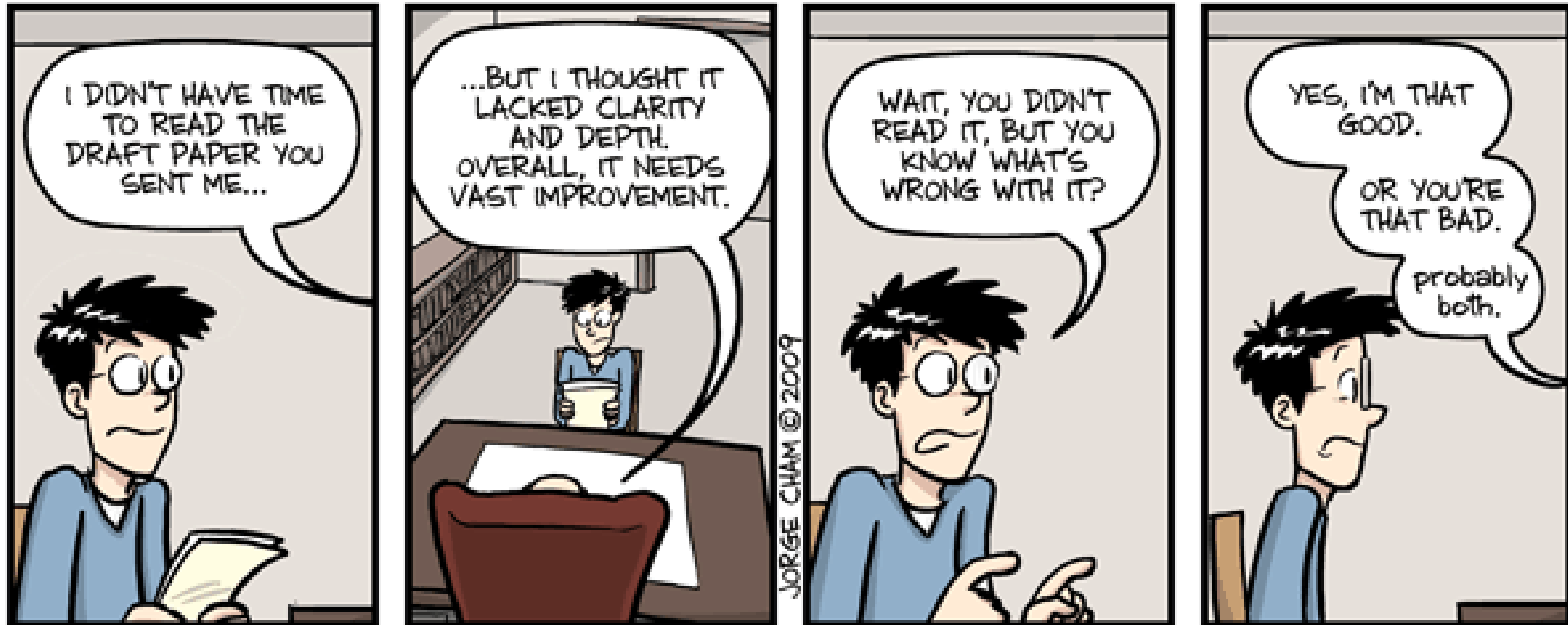


The Reality of a PhD



Source:
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=983>

Faculty can help!



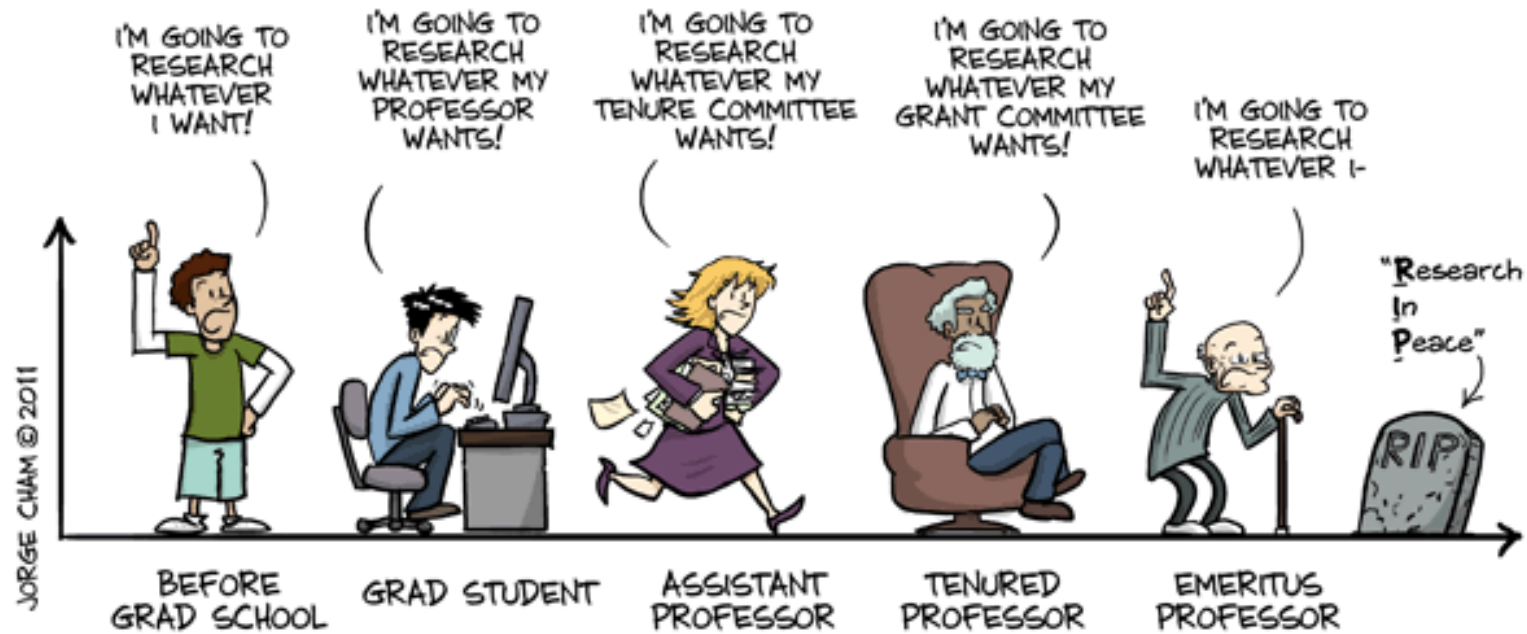
WWW.PHDCOMICS.COM

Source:
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=1143>

Faculty can help!



THE EVOLUTION OF INTELLECTUAL FREEDOM



WWW.PHDCOMICS.COM

Source:

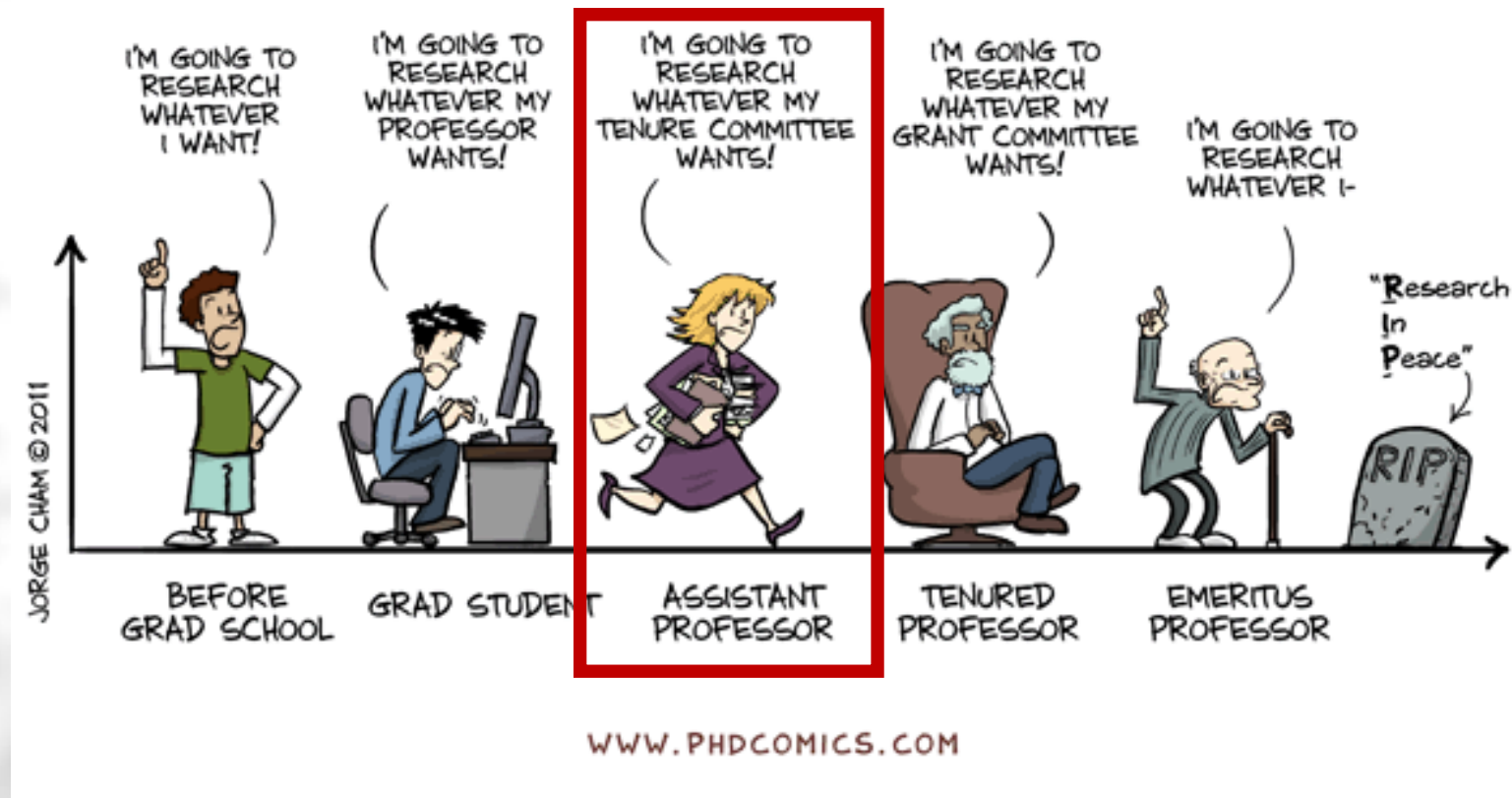
"Piled Higher and Deeper" by Jorge Cham

<https://phdcomics.com/comics/archive.php?comid=1436>

Faculty can help!



THE EVOLUTION OF INTELLECTUAL FREEDOM

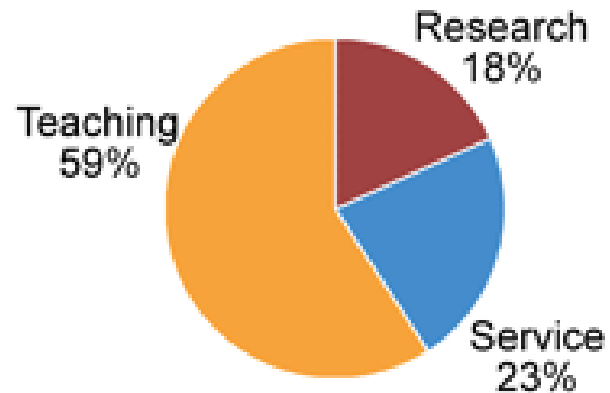


Faculty can help!



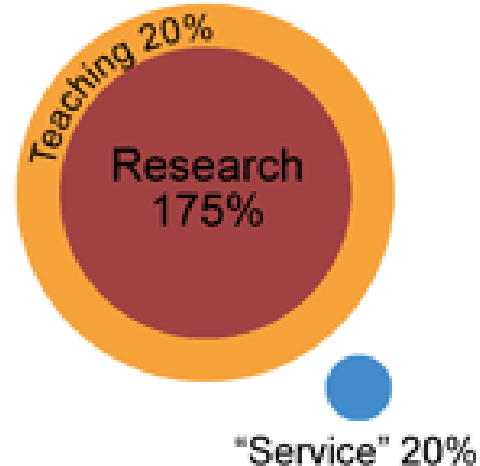
HOW PROFESSORS SPEND THEIR TIME

How they actually spend their time:



Source: Higher Education Research Institute Survey (1999)

How departments expect them to spend their time:



How Professors would *like* to spend their time:

Don't tell me what to do

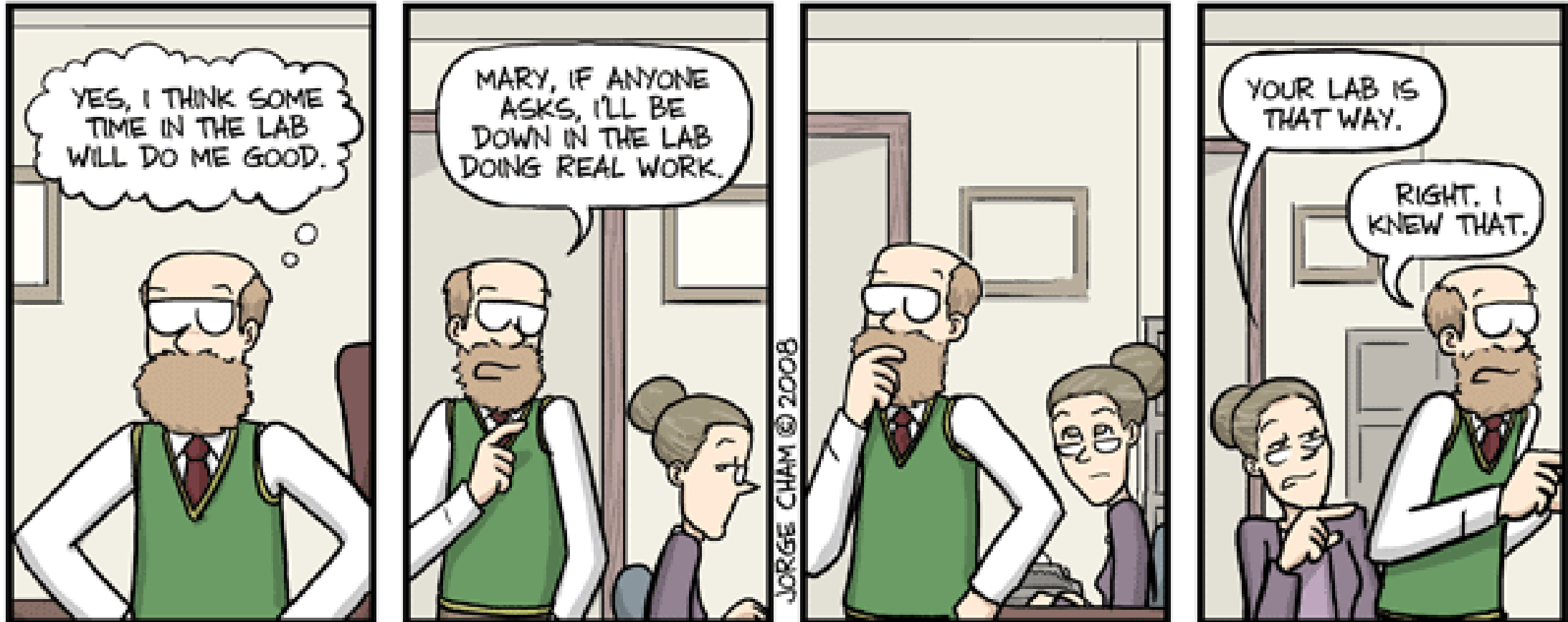
JORGE CHAM © 2008

WWW.PHDCOMICS.COM

Source:

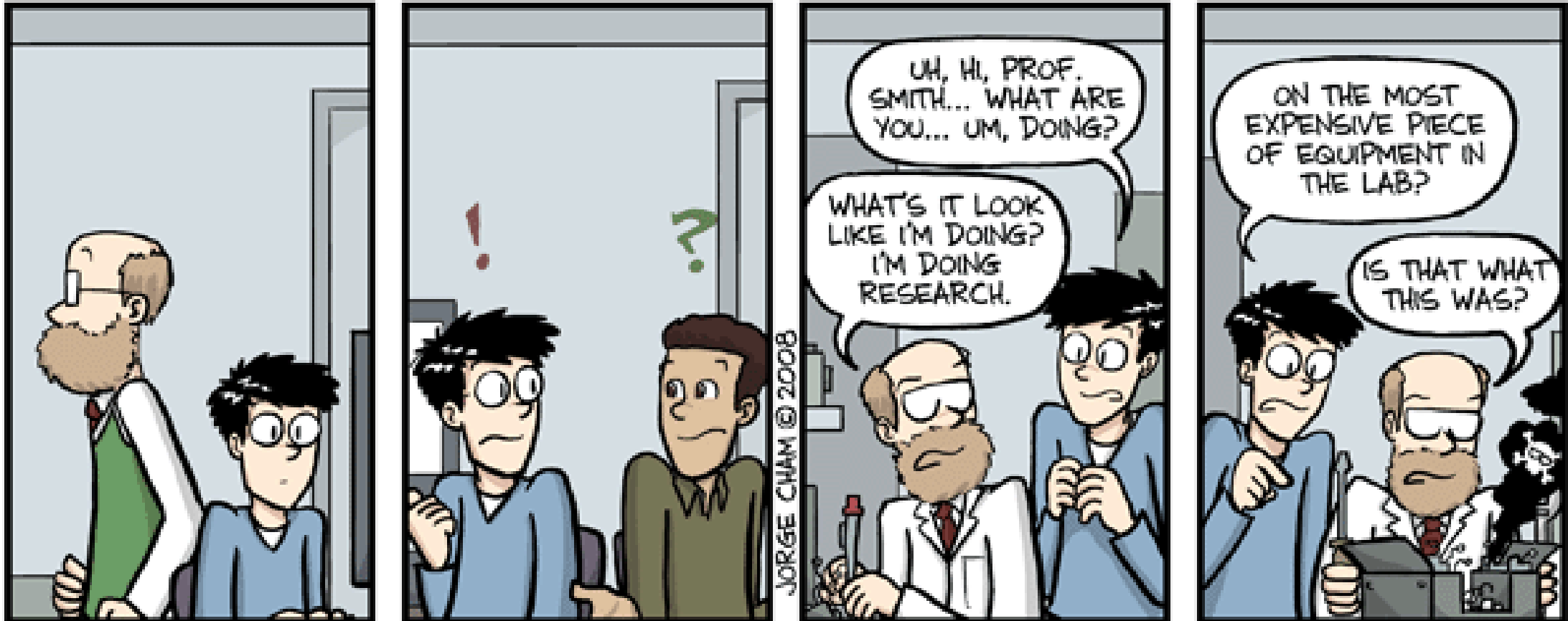
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=1060>

Faculty can help!



WWW.PHDCOMICS.COM

Faculty can help!



WWW.PHDCOMICS.COM

Source:

"Piled Higher and Deeper" by Jorge Cham

<https://phdcomics.com/comics/archive.php?comid=1054>

Operational Support



- IT departments dislike measurement setups in their network
- Some researchers scan from public clouds (hard to block)
- Ethics committees / IRBs do not understand lore & technology
- Middleboxes everywhere
- Infrastructure often does not survive students' departure



Source:
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=768>

A measurement.network



Build a community governed scan infrastructure available to researchers, where measurement tools are reviewed by people with actual experience before they are let loose on the Internet. Provide systems to take the overhead (operation, monitoring, self-contain, abuse, blocklist maintenance, attribution) away from researchers, while making infrastructure accessible even if a local IT department likes its Fortigate, predictable (blockable), and (more) reliable; Also enable others to easily get access to reproduction (artifact evaluation!) and gathered data.

FAQ: Why are you doing this alone?



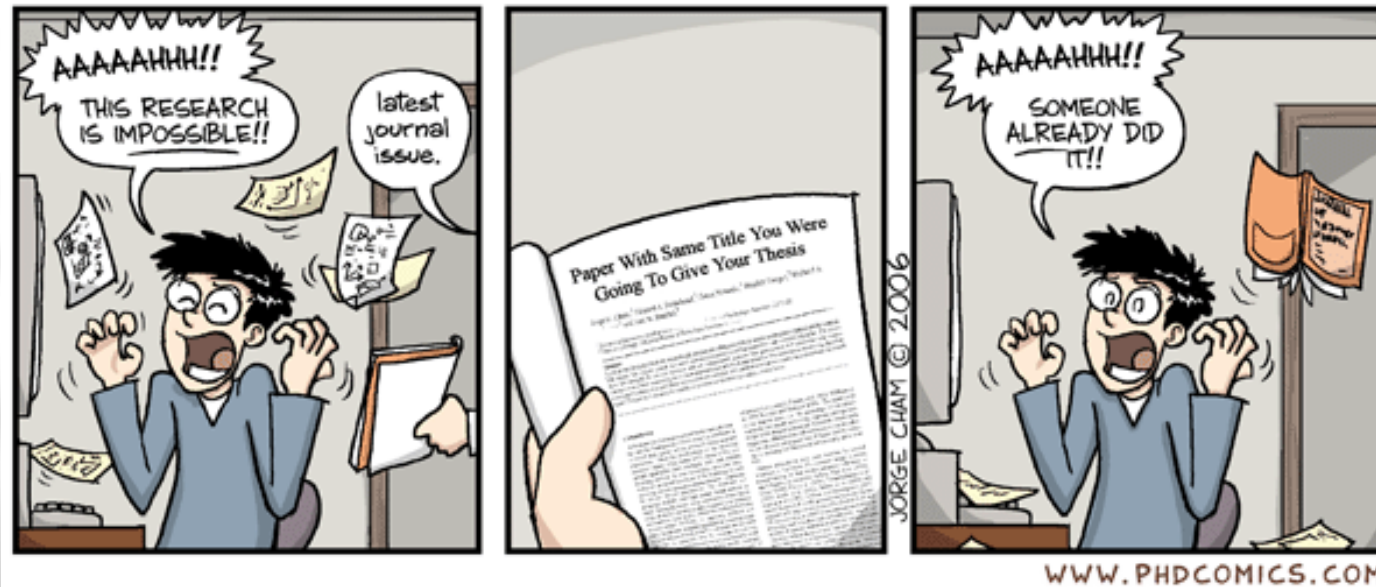
“Alone for now; Usually, it is much easier to convince people of joining an effort if something is already *there*; Also, to prevent bike shedding.

Let me know if you want to join:
tfiebig@measurement.network”

FAQ: Shouldn't <US R1 Group> run this?



“The infrastructure should not be ‘owned’ by an entity that publishes for a living itself. This is also why I want to hand this off when the foundation is there! Researchers are paranoid and do not want to play with others’ toys.”



Source:
"Piled Higher and Deeper" by Jorge Cham
<https://phdcomics.com/comics/archive.php?comid=789>

FAQ: Who could be running this?



“I guess the RIPE NCC/RIPE Measurement WG, or the MAPRG in the IRTF would be great places to start looking.”

FAQ: Isn't this easier to block?



“That is the point of it. People should be able to block a prefix and be sure there is not a core-component of the University’s LMS showing up on the same IP a week later.”

FAQ: What if this doesn't work?



“I burned my own time and money, and no one else is hurt.”

FAQ: How will this be paid for?



“For now, this is supported by the ‘Tobias-Fiebig-Personal-Bank-Account-Foundation for Doing Things I Consider Useful’-grant (TFPBAffDTICU).

Additionally, two operators (indirectly) sponsor upstream

(AS50629  LWL.COM, AS58299  OPENFACTORY)

and I got a /22 IPv4 (LEGACY) from MPI-INF for the project.

With things being set up, I will try to motivate more entities to pitch in; But, as said before, getting things to work first. Overall costs should be manageable (~\$6k/Y).”

FAQ: Where can we find it?



Resources

ASN: AS211286

IPv4: 141.39.220.0/22

IPv6: 2a0d:8d04::/32

PoPs:

- Düsseldorf (MyLoc DUS01)
- Berlin (Speedbone/AK)
- Planned: MPI-INF, AMS

Needs: Colo, hardware, L2

Web

<https://measurement.network/>

Existing services (to be merged):

<https://www.email-security-scans.org/>

<https://bttf-whois.as59645.net/>

<https://v6only-resolver.measurement.network/>