Anycast vs. DDoS: Evaluating Nov. 2015 Root DNS event

Giovane C. M. Moura[1], Ricardo de O. Schmidt[2], John Heidemann[3], Wouter B. de Vries[2], Moritz Müller[1], Lan Wei[3], Cristian Hesselman[1]

[1]SIDN Labs [2]University of Twente [3]USC/ISI

2016-07-18

IEFT 96 – MAPRG/IRTF

IP Anycast

* It's simple: “ making a particular Service Address available in multiple, discrete, autonomous locations” (RFC4786, 7094)

* Improves performance and resilience (1 IP → Many services, 1 down, others operate)

* Widely use in DNS (and also CDNs)

DDoS

* Getting bigger (400Gbps +)

* Getting cheaper (booters, few dollars)

* Happening more often

* Core idea: bring down services

* Question: *How anycast behaves during a DDoS attack?*

* Case study: Root DNS events Nov 2015

USC University of Southern California
UNIVERSITY OF TWENTE.
SIDN LABS
ant.isi.edu

The Root DNS system

- List the records that points to all TLDs (.com, .nl, .net...)

letter	operator	sites (global, local)	architecture
A	Verisign	5 (5, 0)	anycast
B	USC/ISI	1 (1, 0)	single site
C	Cogent	8 (8, 0)	anycast
D	U. Maryland	87 (18, 69)	anycast
E	NASA	12 (1, 11)	anycast
F	ISC	59 (5, 54)	anycast
G	U.S. DoD	6 (6, 0)	anycast
H	ARL	2 (2, 0)	primary/backup
I	Netnod	49 (49, 0)	anycast
J	Verisign	98 (66, 32)	anycast
K	RIPE	33 (15, 18)	anycast
L	ICANN	144 (144, 0)	anycast
M	WIDE	7 (6, 1)	anycast

Table: The 13 Root Letters, each operating a separate DNS service, and their number of sites and architecture as of 2015-11-18.

A Bad Day at the Root…

data: RIPE DNSmon
red: >30% loss
(some sites ~99% loss!)

What happened?

What does "red" *really* mean?

Anycast vs. DDoS *in general*?

Summary of the Events

Two events

- 2015-11-30t06:50 for 2h40m
- 2015-12-01t05:10 for 1h

affected 10 of 13 letters

about 5M q/s or 3.5Gb/s per affected letter

- aggregate: 155Gb/s

real DNS queries, common query names, from spoofed source IPs

implications:

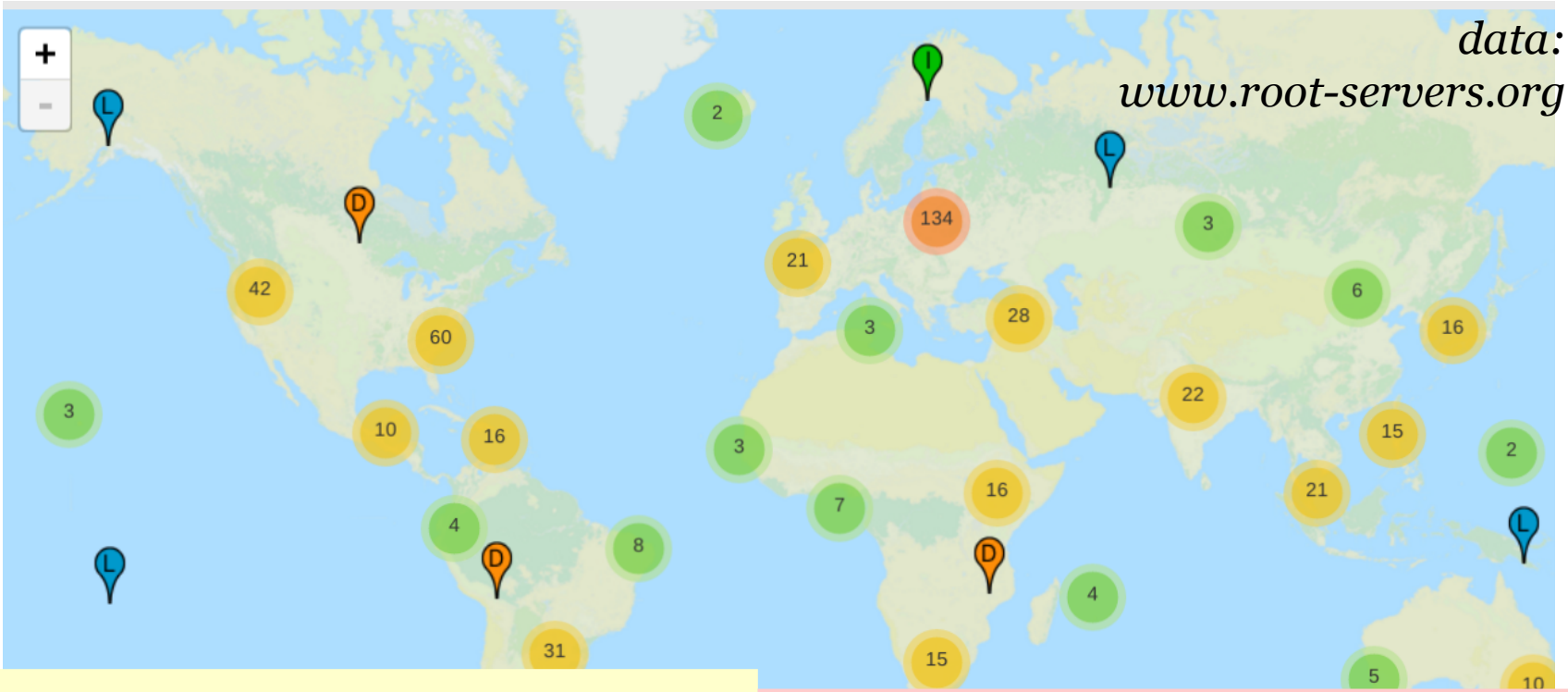
some letters had high loss

overall, though DNS worked fine

- **clients retried other letters (as designed)**

data:
A-Root had full view (Verisign presentation);
RSSAC-002 reports

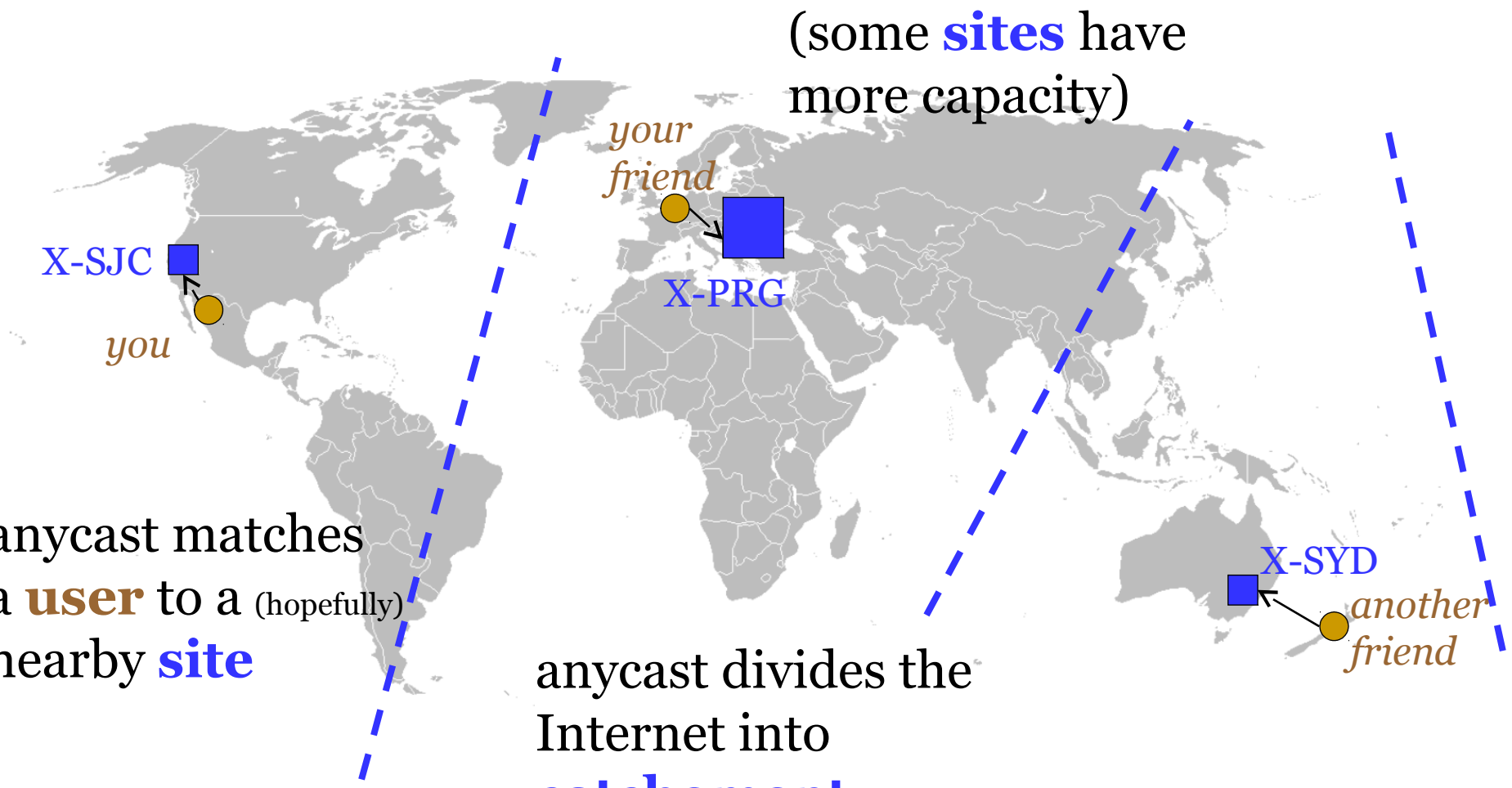
How Well Does Anycast Defend?



561 root DNS locations
for **13 services** (in 2016-01)

is 561 ***too few? too many?***
what happens ***under stress?***

Anycast in Good Times



anycast matches a **user** to a (hopefully) nearby **site**

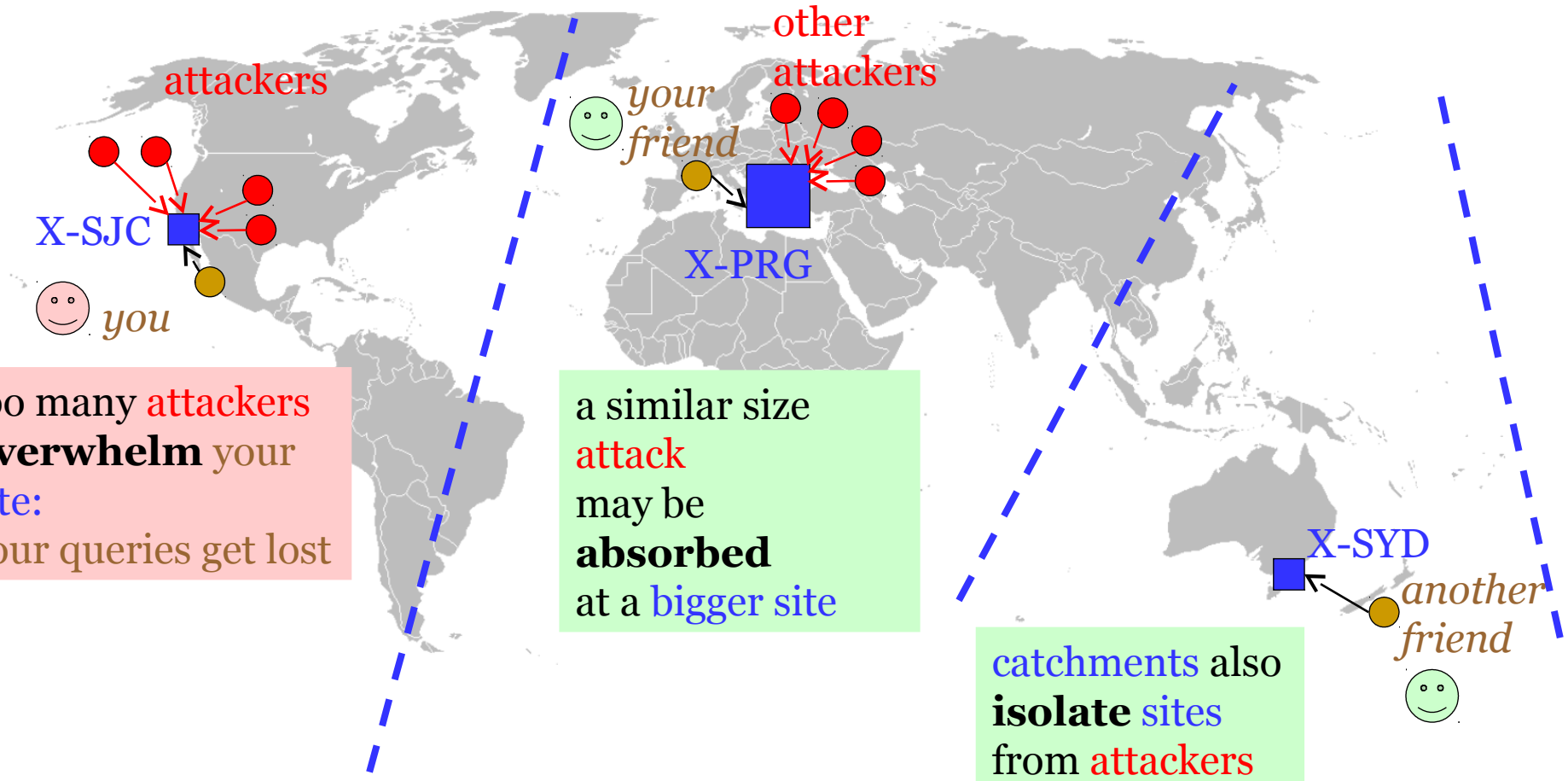
anycast divides the Internet into **catchement** (often messy and non-geographic)

USC University of Southern California

UNIVERSITY OF TWENTE.

SIDN LABS

Anycast Under Stress

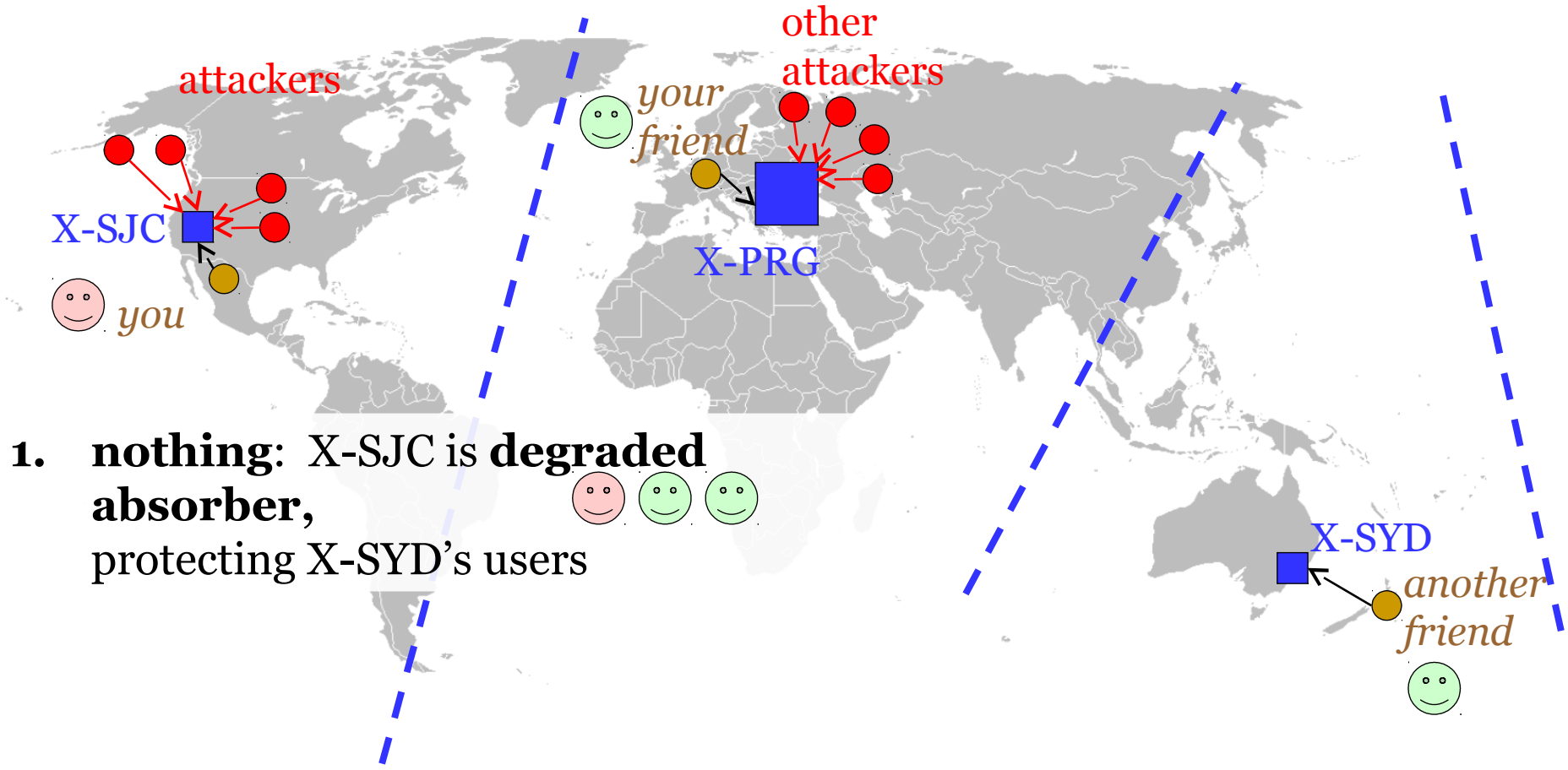


too many attackers **overwhelm** your site:
your queries get lost

a similar size attack may be **absorbed** at a bigger site

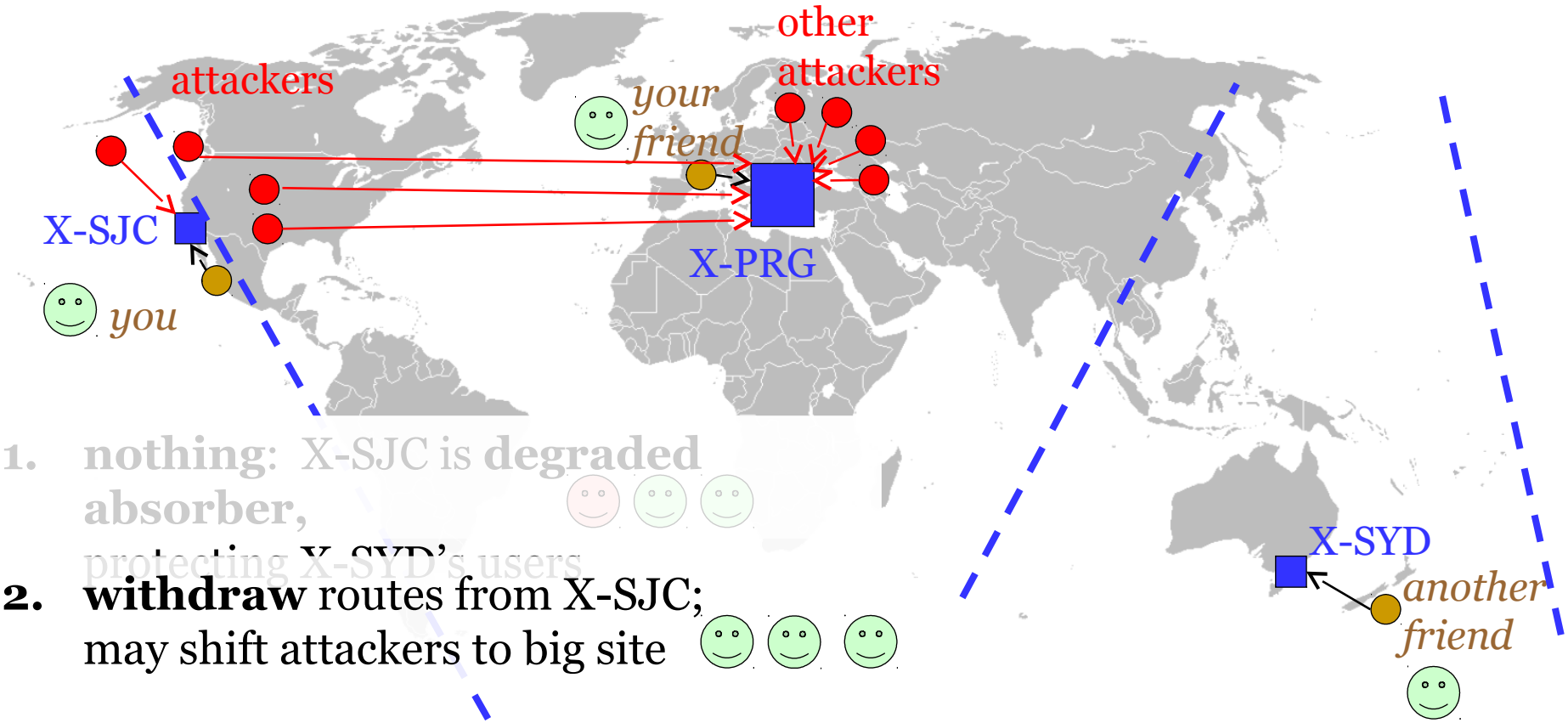
catchments also **isolate** sites from attackers

Anycast Under Stress (do nothing)



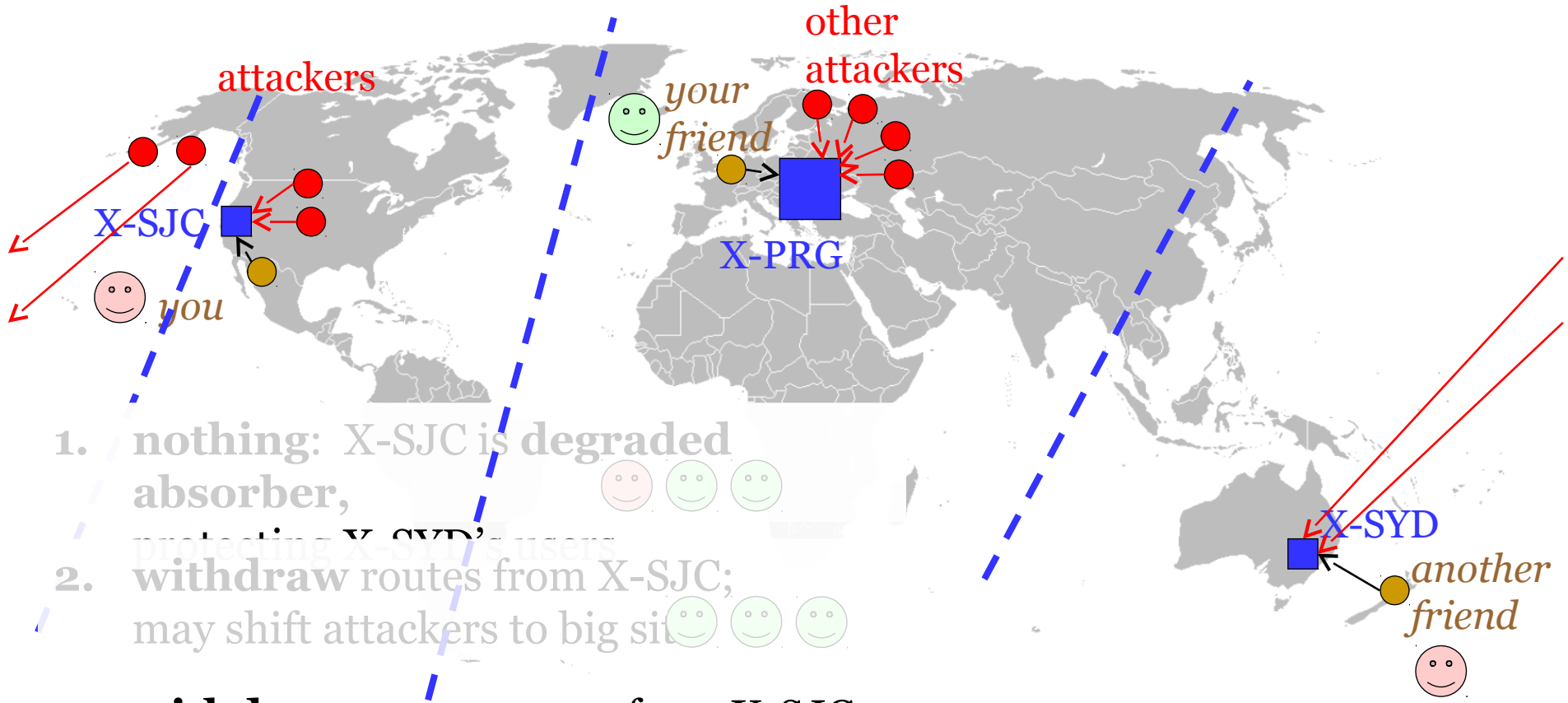
1. **nothing**: X-SJC is **degraded absorber,**
   protecting X-SYD's users

Anycast Under Stress (withdraw some routes)



1. **nothing**: X-SJC is **degraded absorber**, protecting X-SYD's users
2. **withdraw** routes from X-SJC; may shift attackers to big site

Anycast Under Stress (withdraw other routes)



1. **nothing**: X-SJC is **degraded absorber**, protecting X-SYD's users
2. **withdraw** routes from X-SJC; may shift attackers to big sit
3. **withdraw** wrong routes from X-SJC; may shift attackers to other sit

Best reaction to stress: you don' t know

don’t know:
number of attackers
location of attackers
affects of routing change

don’t fully control
routing and catchments

hard to make informed choices

1. **nothing**: X-SJC is **degrade** **absorber,**
2. **withdraw** routes from X-SJ may shift attackers to big si
3. **withdraw** wrong routes from X-SJC; may shift attackers to other sit

-SYD
another friend

What Actually Happens?

studying Nov. 30

we see **withdrawals** and **degraded absorbers**

some clients loose service

- results vary by anycast deployment

Data About Nov. 30

RIPE Atlas

- 9000 vantage points (RIPE Atlas probes)
- try every letter every 4 minutes
  - except A-root, at this time, was every 30 minutes
- data-plane queries
- global, but heavily biased to Europe

RSSAC-002 reports

- self-reports from letters
- not guaranteed when under stress

BGPmon routing

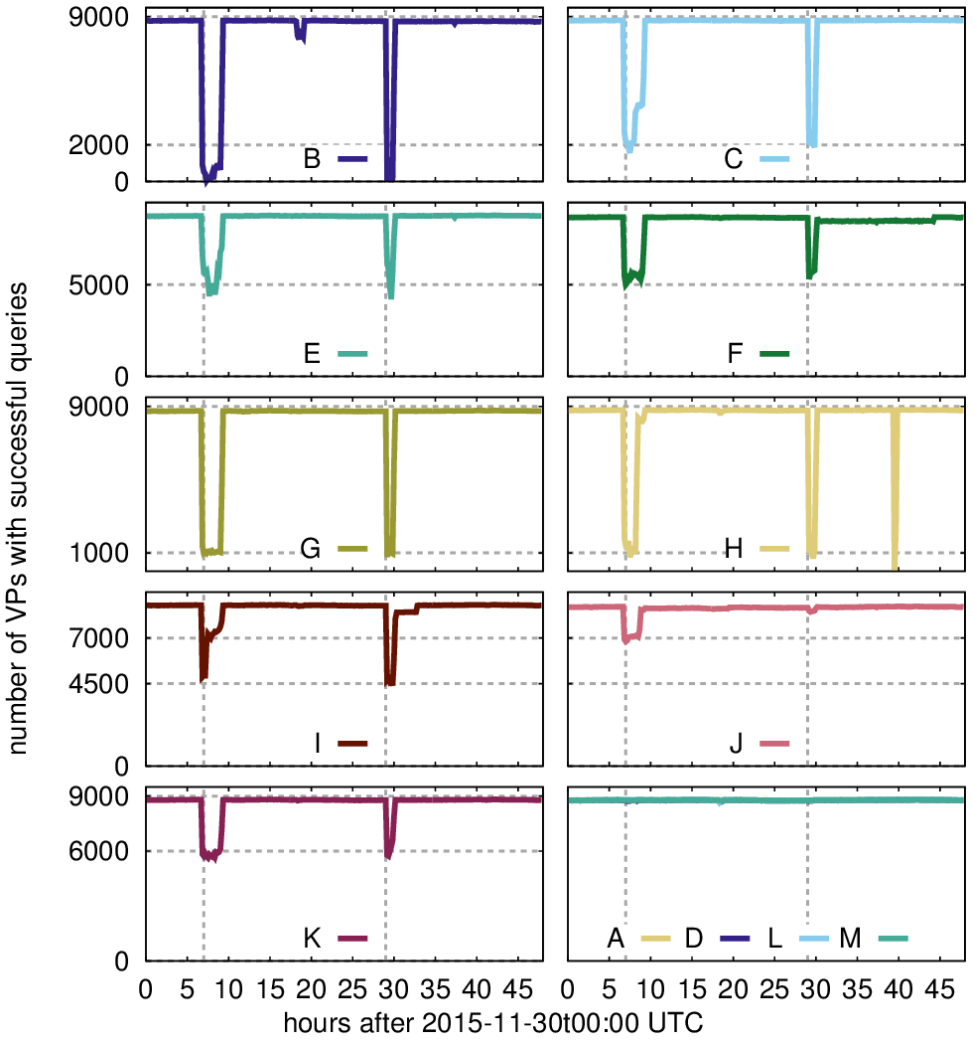
- control plane

How About the Letters?

some did great:
D, L, M: not attacked
A: no visible loss

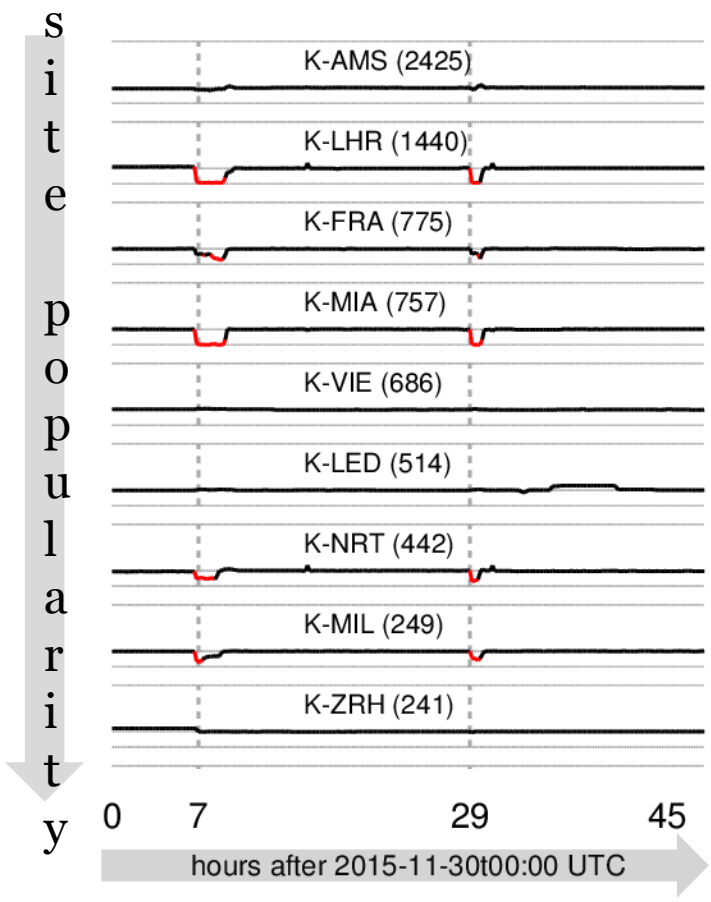
most suffered:
a bit (E, F, I, J, K)
or a lot (B, C, G, H)

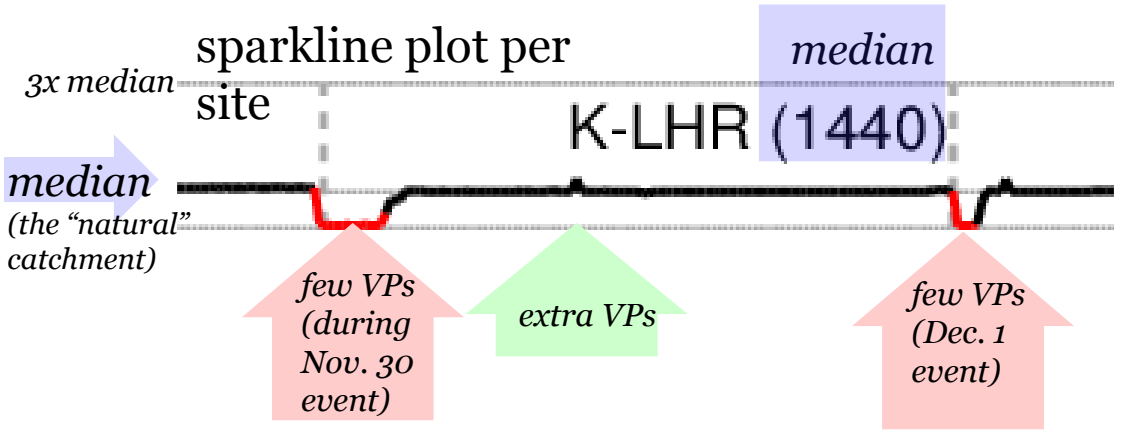
but does “x%”
measure what
users actually see?



[Moura16a, figure 3; data: RIPE Atlas]

Reachability at K-sites

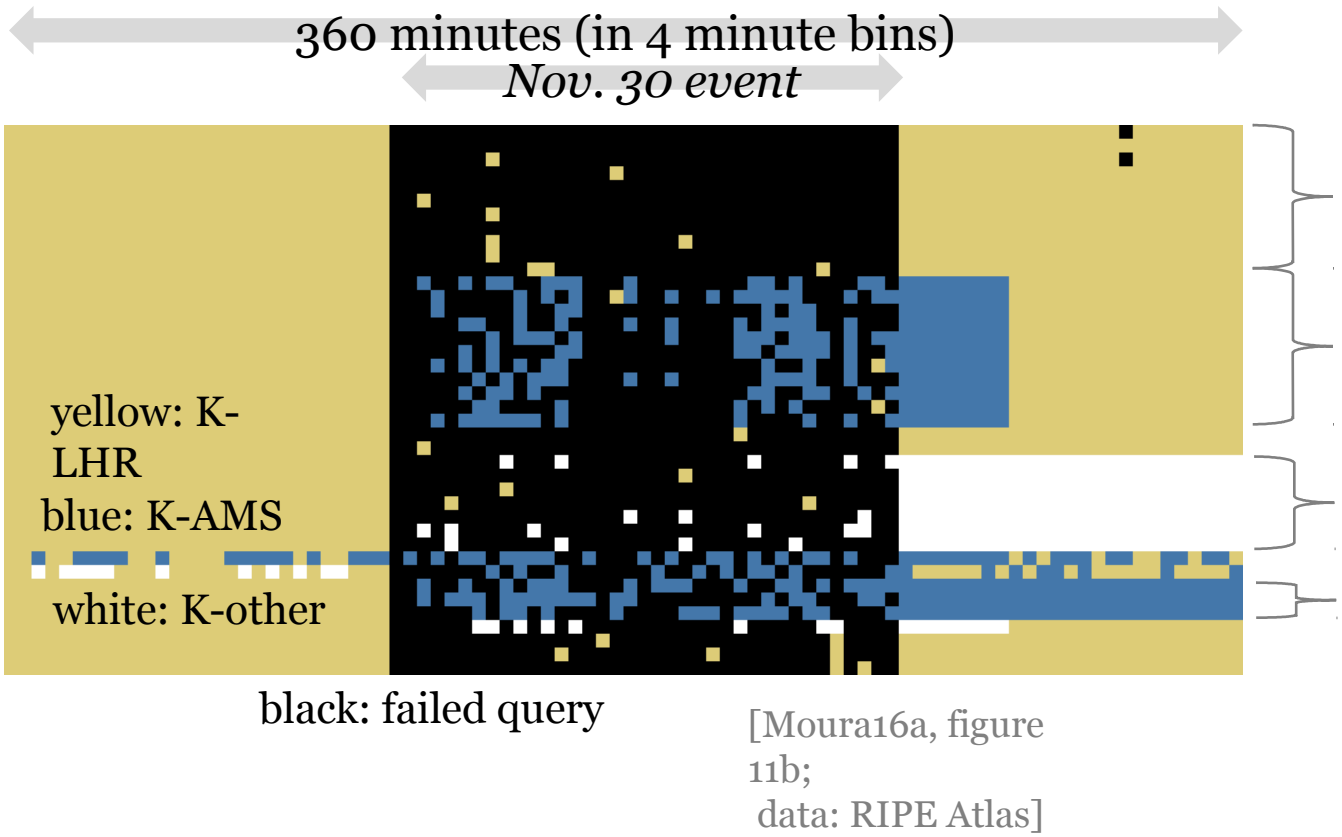




sites see fewer VPs, but why?

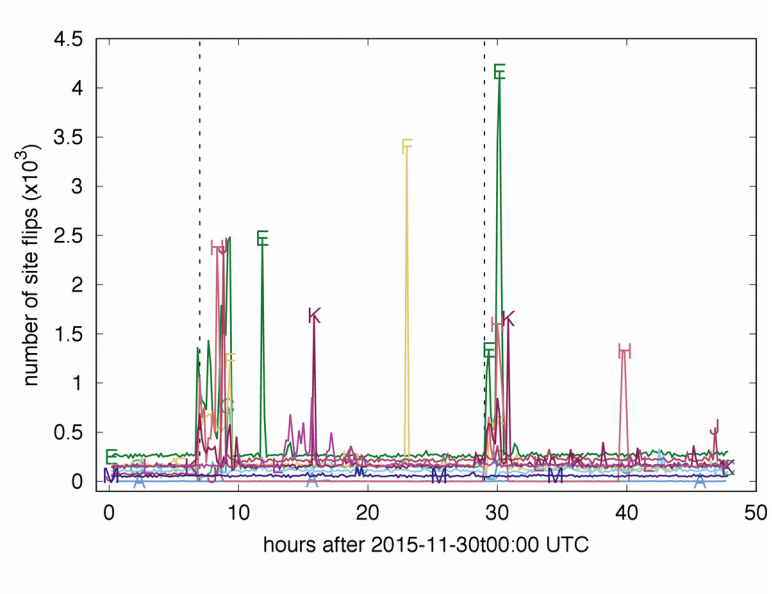
- query loss? site absorbs attack,
  but sad customers
- route change? who? why? where?

Site *Flips* from Routing Changes

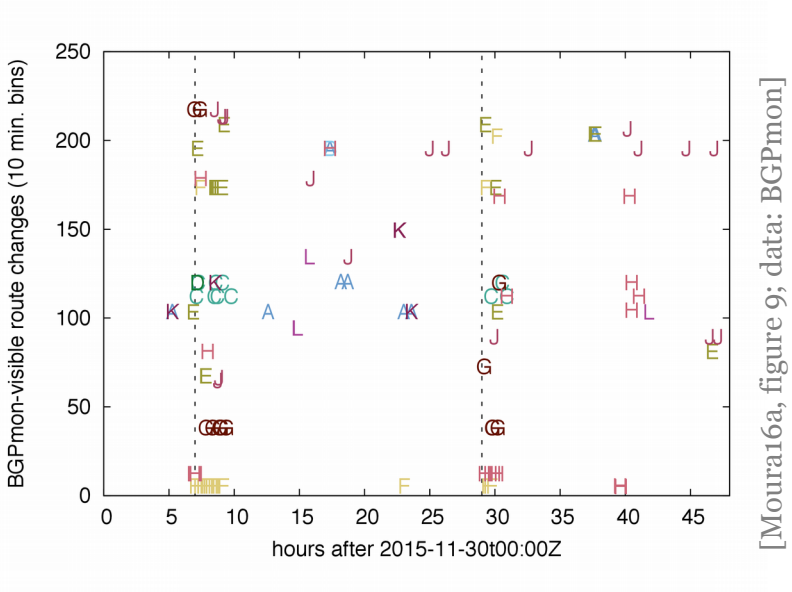


[Moura16a, figure 11b;
data: RIPE Atlas]

Confirming flips in BGP



[Moura16a, figure 8; data: RIPE Atlas]



[Moura16a, figure 9; data: BGPmon]

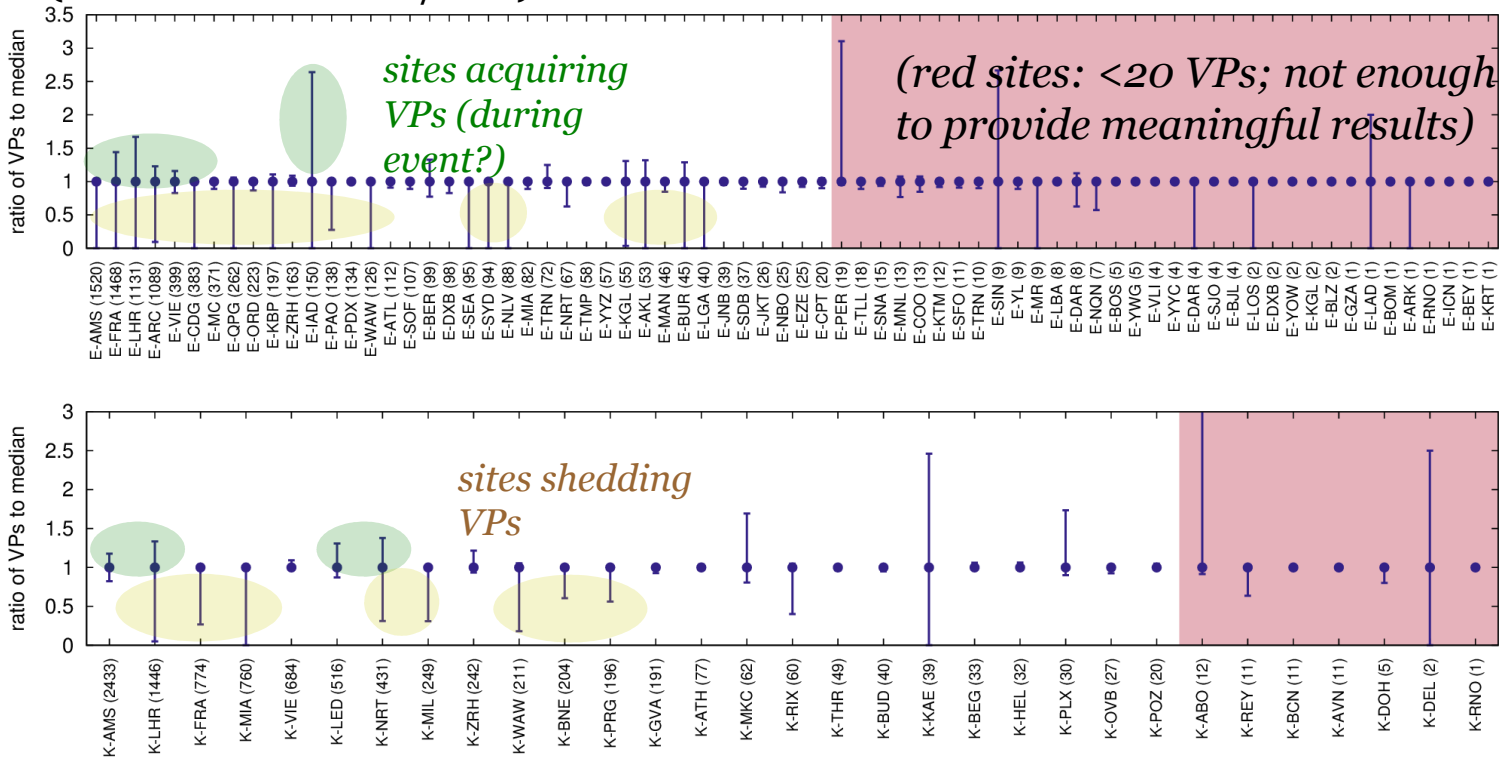
flips common during events for most letters

flips seen in BGP

Flips Across Letters: E and K

to evaluate flips over two days:
compare *minium* and
maximum cachement
(measured in VPs/site)

normalize to median
(natural catchment)



[Moura16a, figure 5; data: RIPE Atlas]

Flips Implications

some ISPs are “sticky” and won’t flip

- will suffer if their site is overloaded

some ISPs will flip

- but new site may not be much better

result depends on many factors

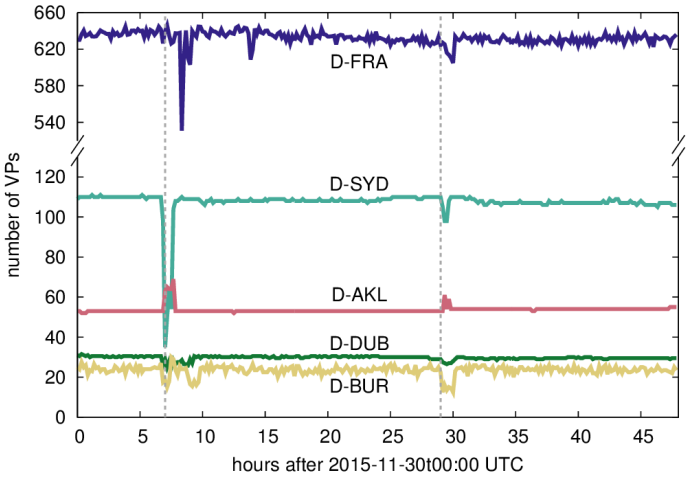
- *actions* taken by root operator
- routing choices by operator *and peer*
  - and perhaps *peer’s peers*, depending on congestion location implementation choices

- DNS, routing

Aside: Collateral Damage

can an event hurt non-targets?

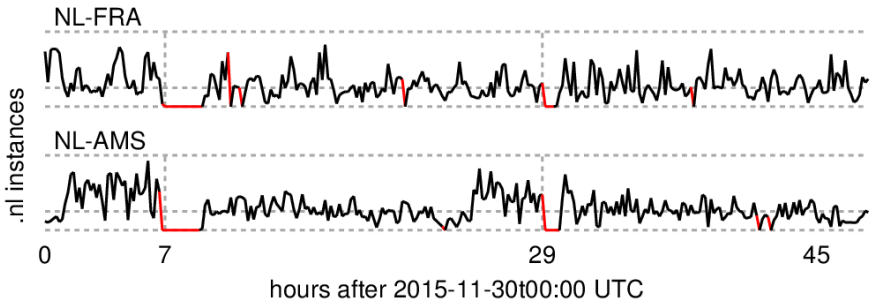
yes! ...a risk of shared datacenters



[Moura16a, figure 14; data: RIPE atlas]

D-FRA and D-SYD: less traffic

(even though D was not directly attacked)



[Moura16a, figure 15; data: SIDN]

.NL-FRA and .NL-AMS: *no* traffic

In other attacks, B-Root's ISP saw loss to other customers

Conclusions

anycast under stress is complicated

- some users will see persistent loss
- “x% loss” is not complete picture

reactions depend on design and implementation choices

- many not under operator control

more info:

paper: http://www.isi.edu/~johnh/PAPERS/Moura16a/

data: https://ant.isi.edu/anycast/