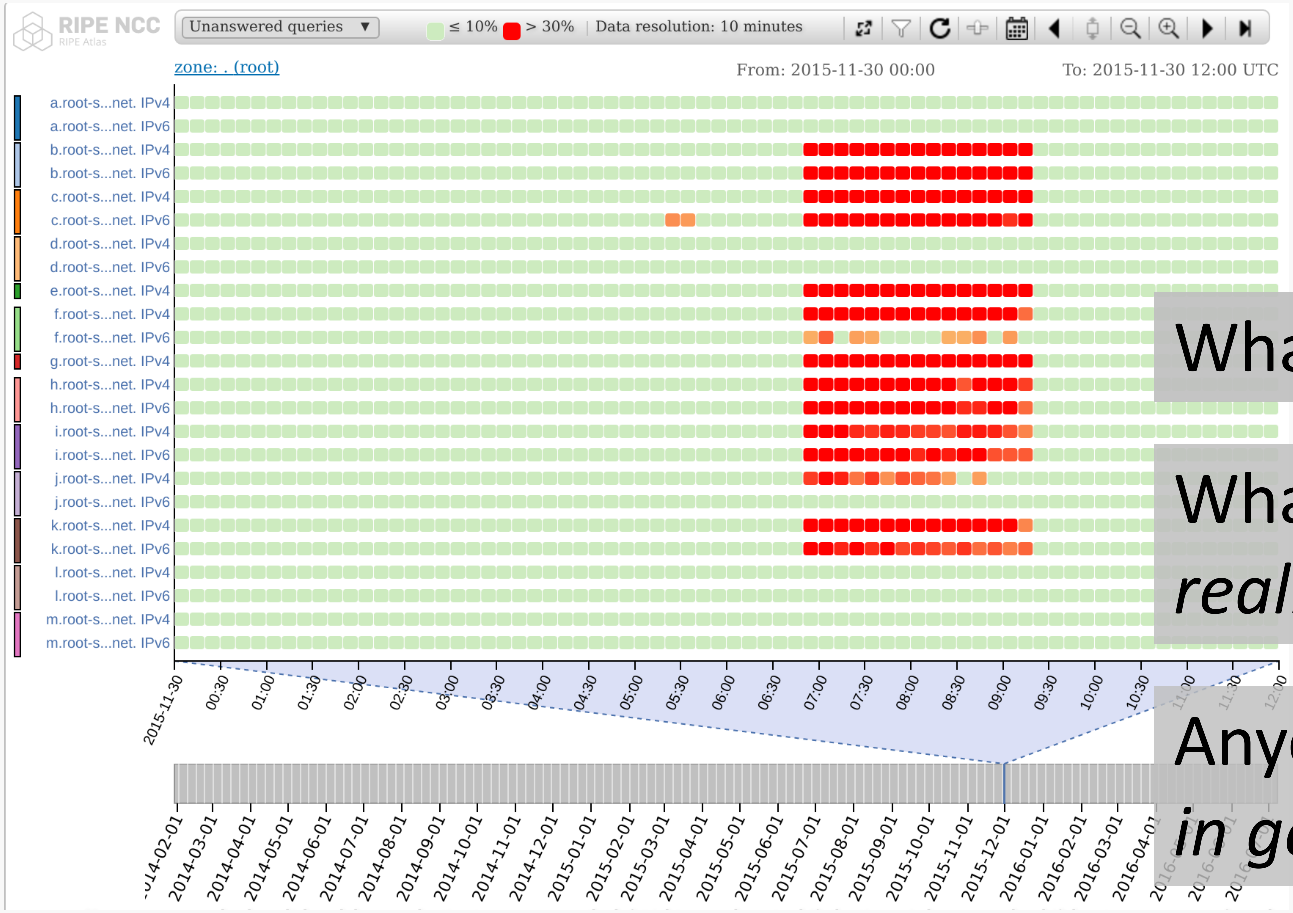
Anycast vs. DDoS: Evaluating Nov. 30

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

At NMRG IETF 97 Seoul 2016-11-15

A Bad Day at the Root…



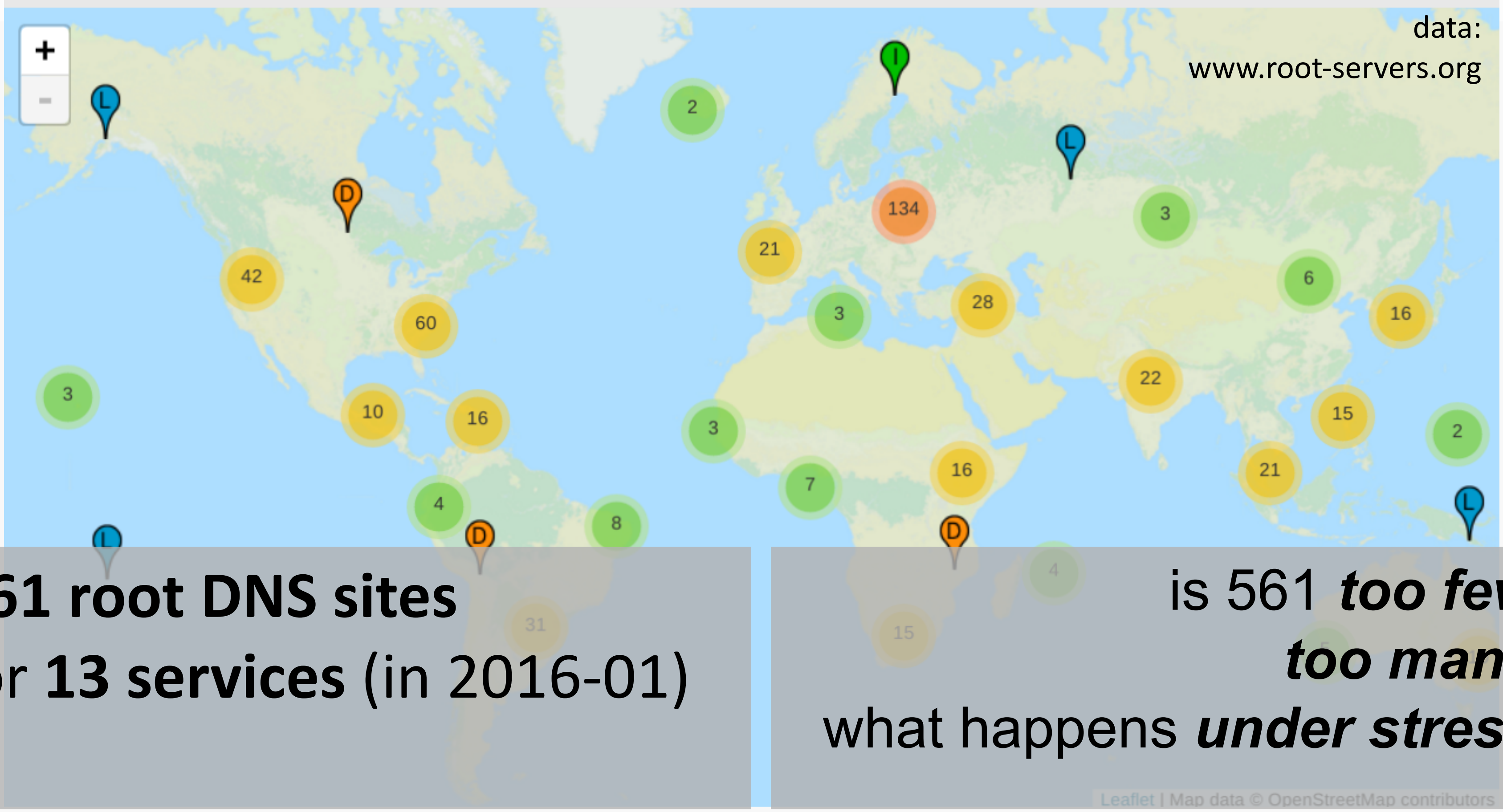
data: RIPE DNSmon
red: >30% loss
(some sites ~99% loss!)

What happened?

What does “red” *really* mean?

Anycast vs. DDoS *in general*?

How *Well* Does Anycast Defend?



data:
www.root-servers.org

Leaflet | Map data © OpenStreetMap contributors

561 root DNS sites

for **13 services** (in 2016-01)

is 561 ***too few?***

too many?

what happens ***under stress?***

Contributions

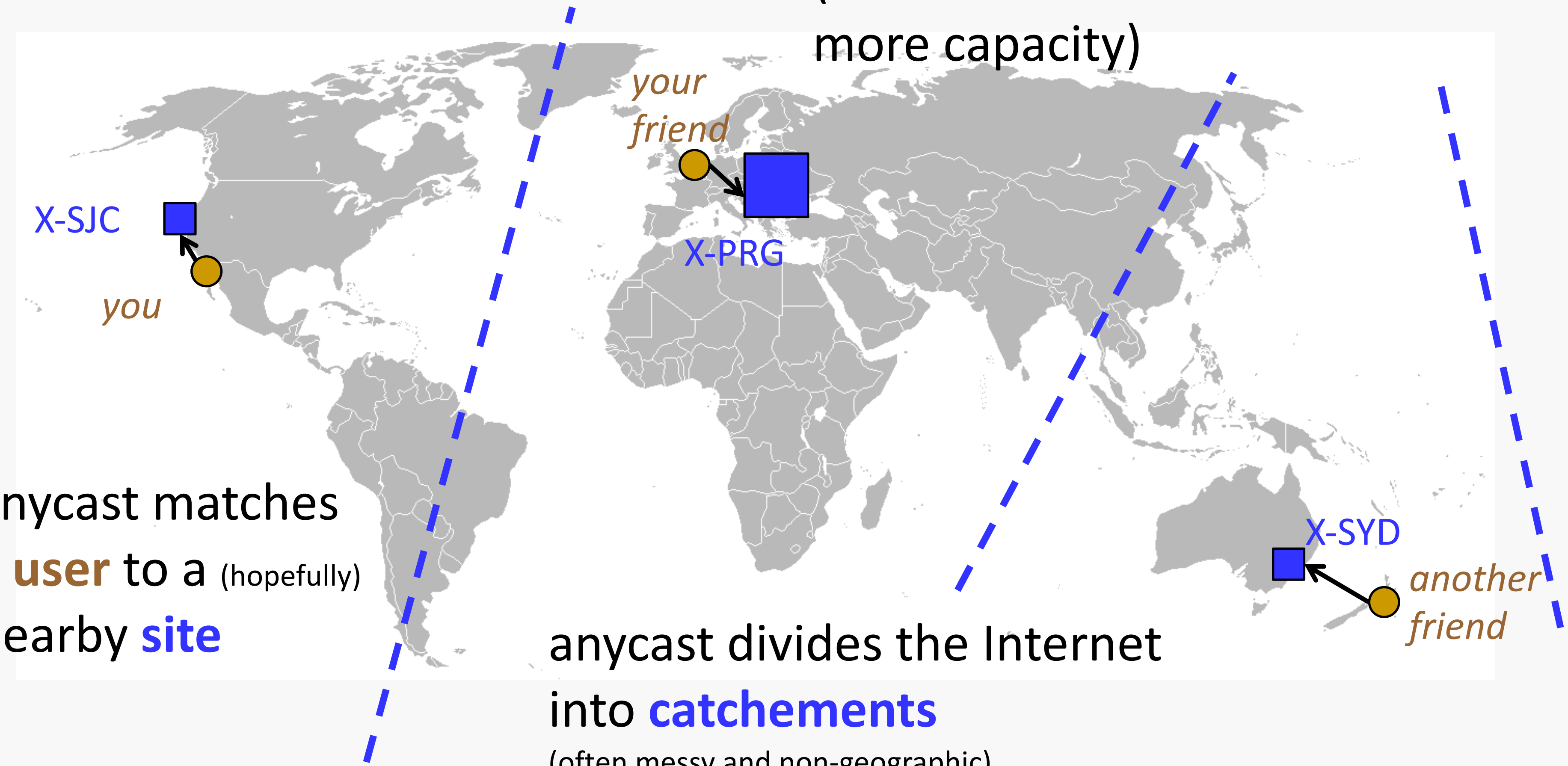
- public evaluation of anycast under stress
- public articulation of design options
- evaluation of collateral damage

prior work for *all*, but in *private*

- goals:
  - public discussion → greater transparency
  - expectation setting
  - possible future defenses

Anycast in Good Times

(some **sites** have more capacity)

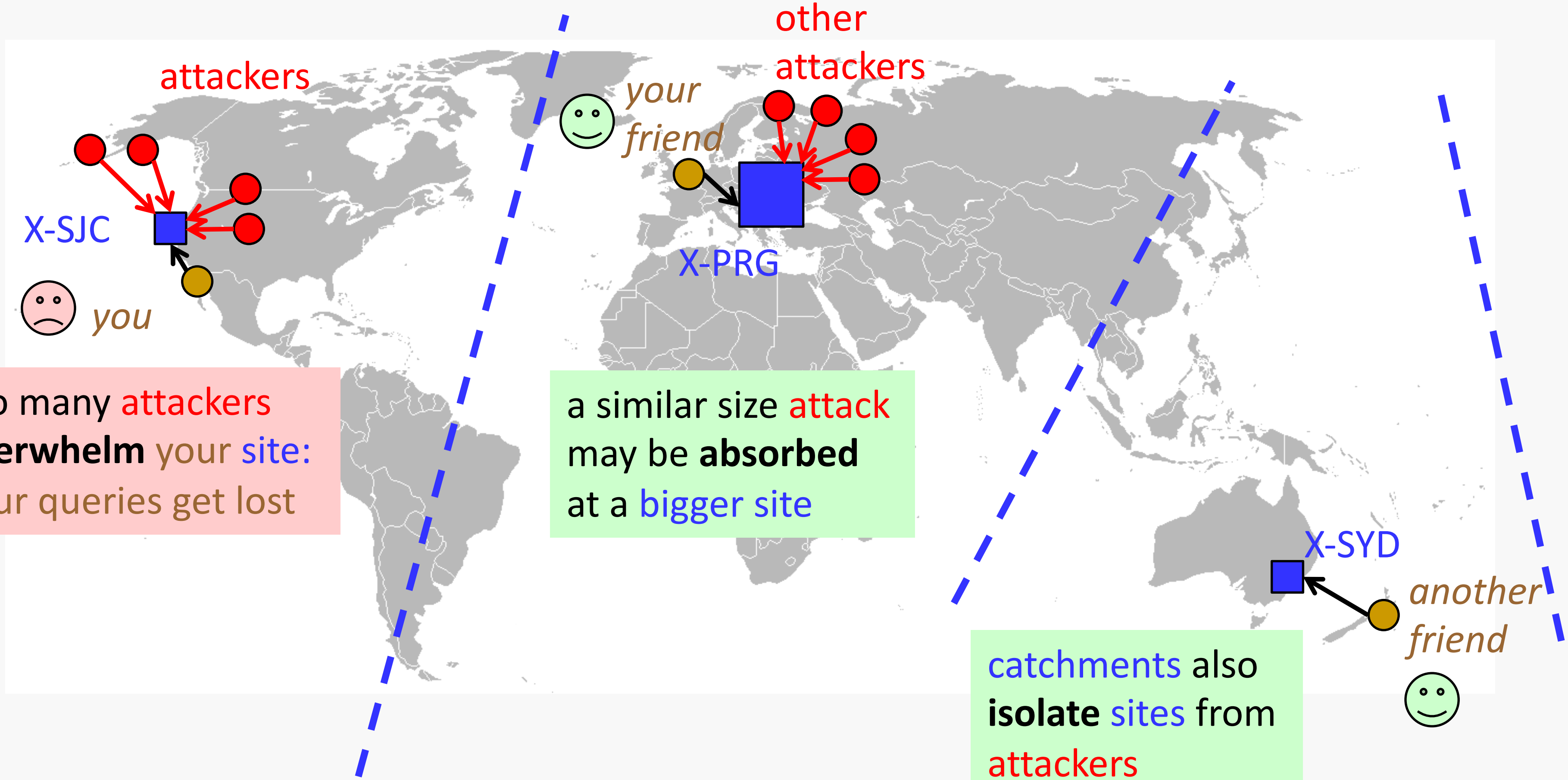


anycast matches a **user** to a (hopefully) nearby **site**

anycast divides the Internet into **catchements**

(often messy and non-geographic)

Anycast Under Stress



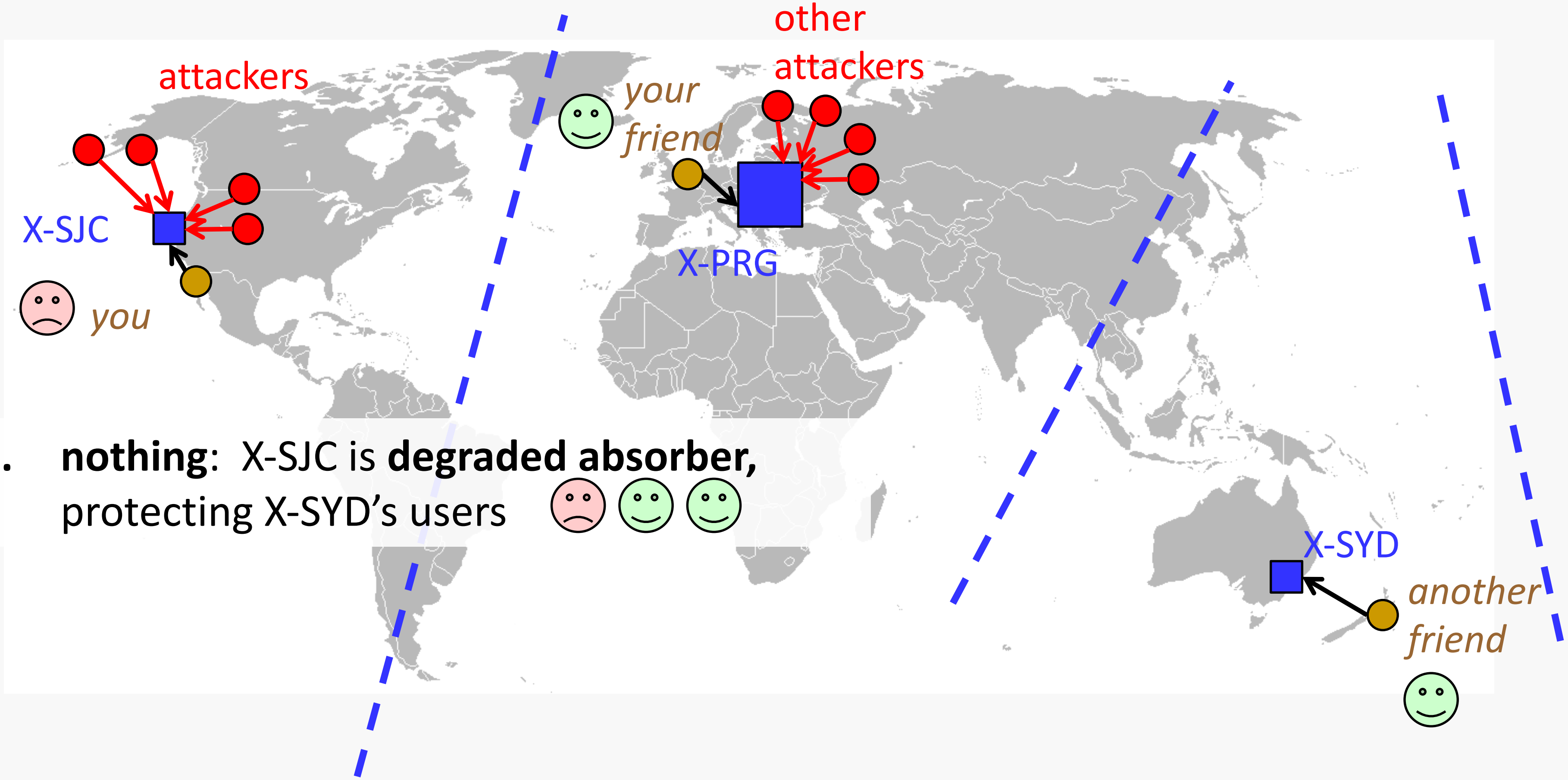
too many attackers **overwhelm** your site: your queries get lost

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

catchments also **isolate** sites from attackers

Anycast Reactions to Stress

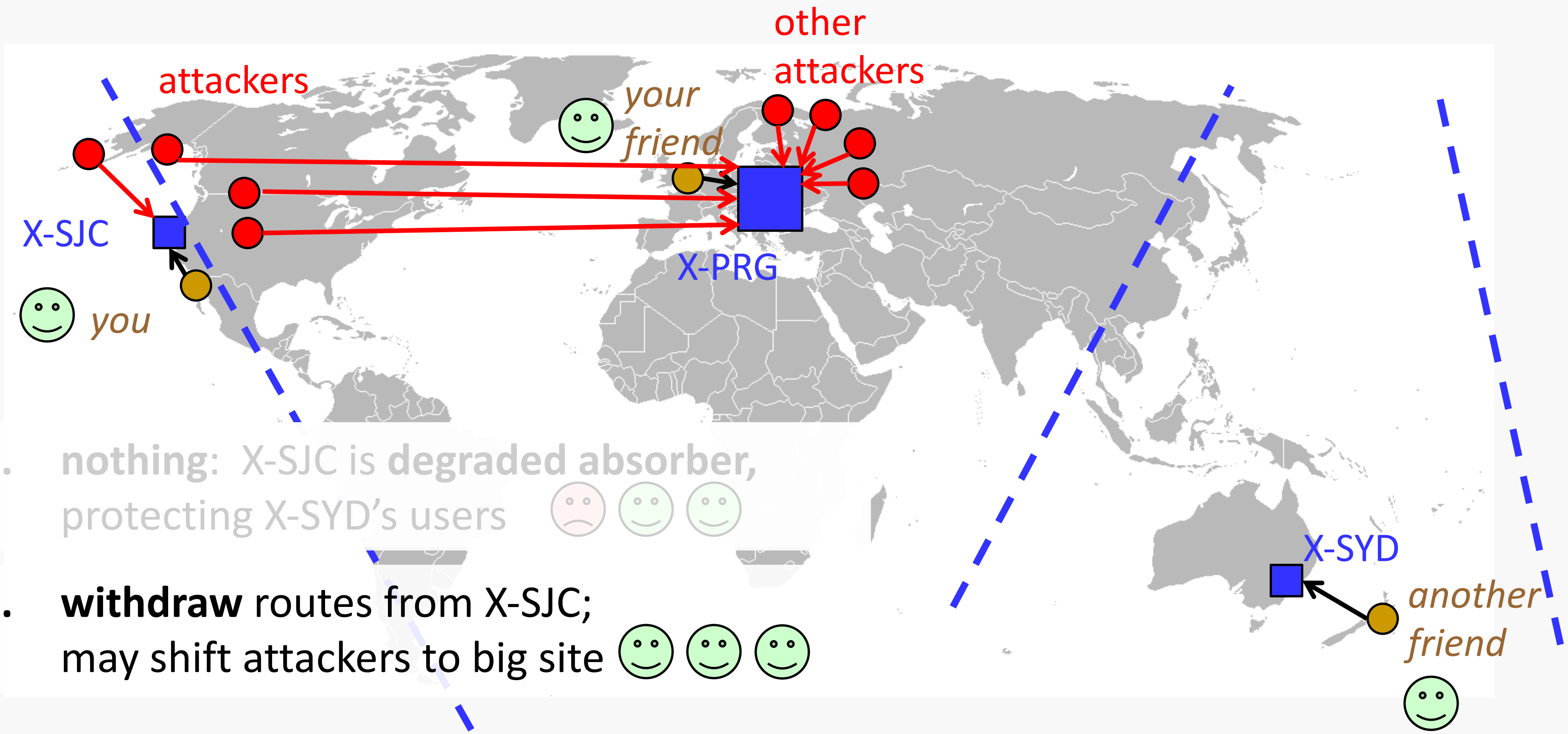
(do nothing?)



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

Anycast Reactions to Stress

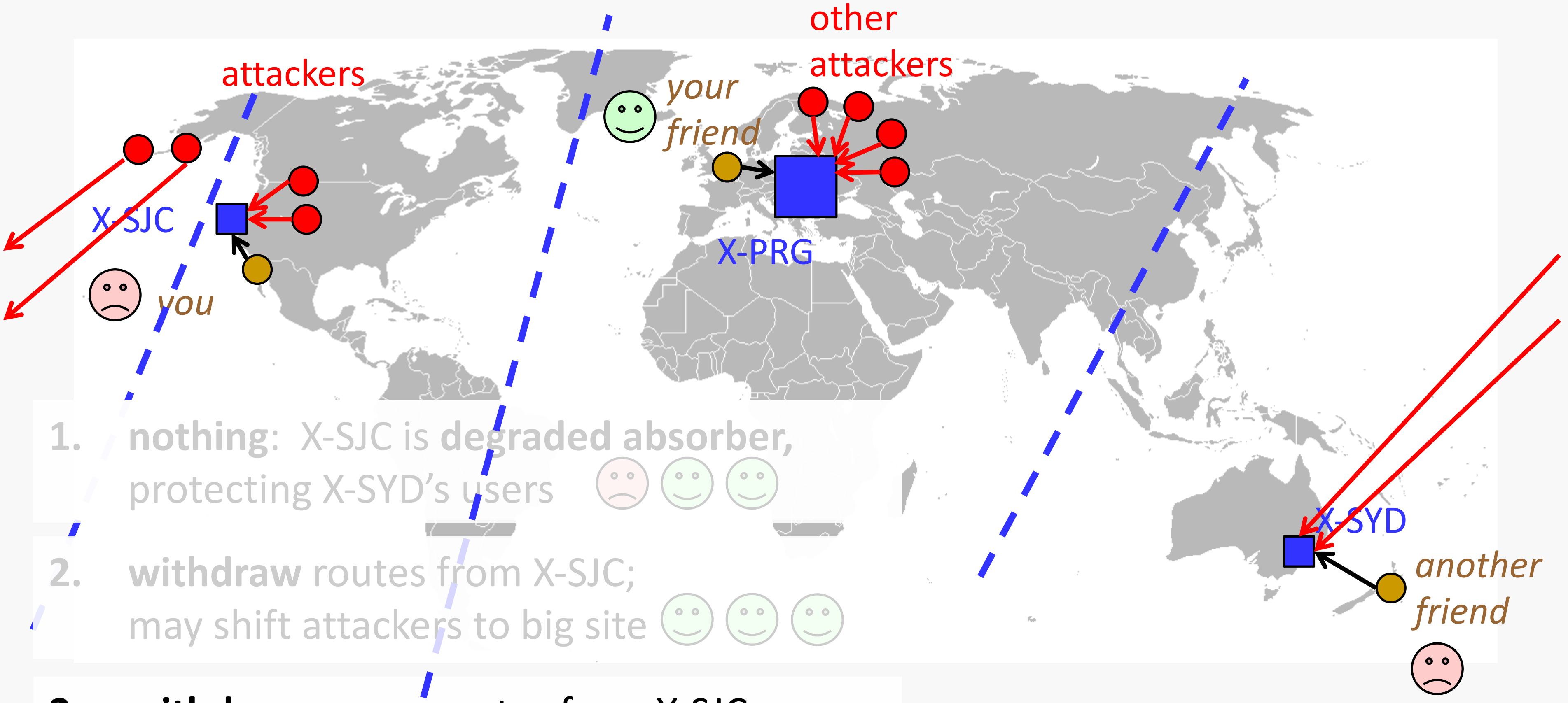
(withdraw some routes?)



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

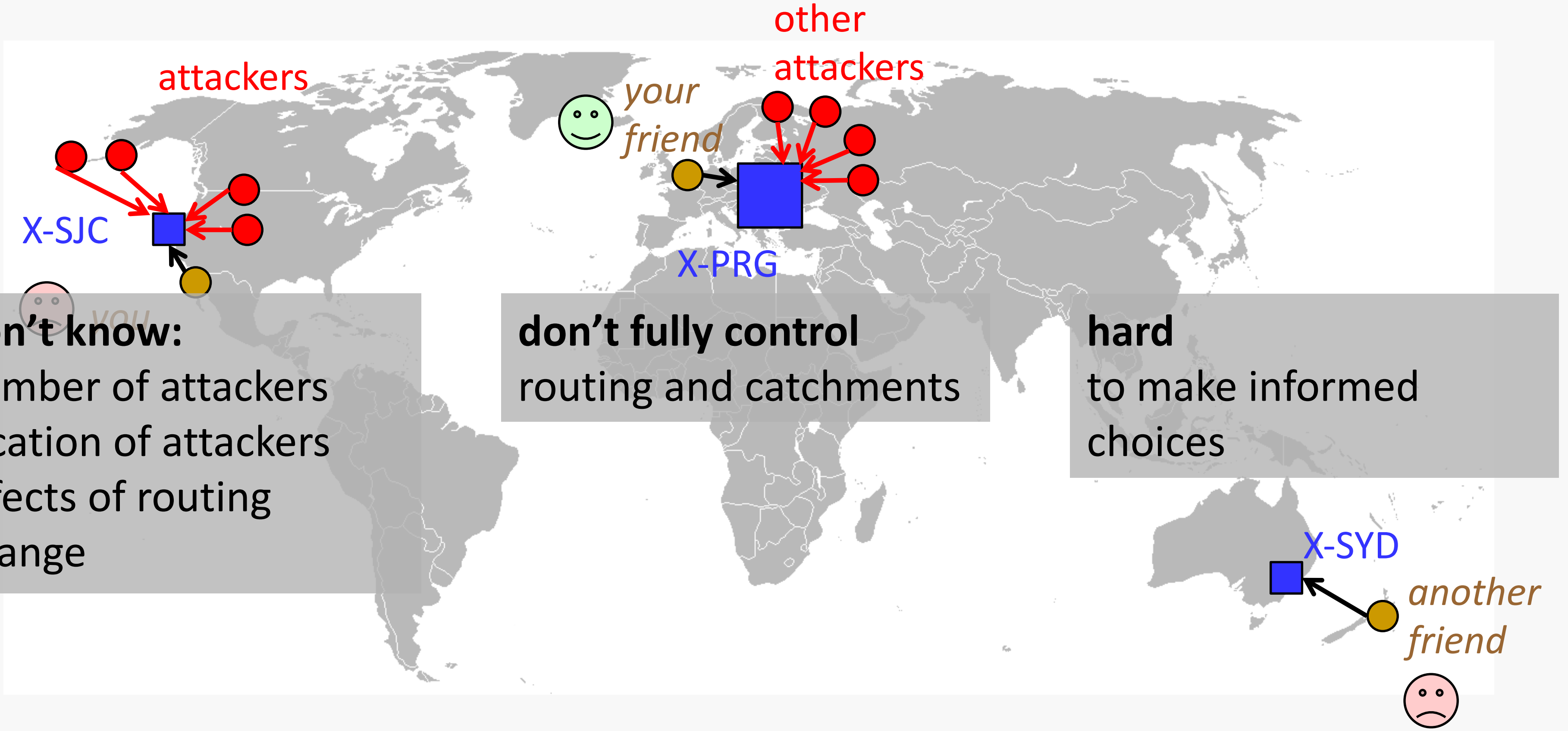
Anycast Reactions to Stress

(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 site ☺ ☺ ☺
3. **withdraw** wrong routes from X-SJC; may shift attackers to other site ☹ ☹ ☺

Best Reaction to Stress?
You Don't Know



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

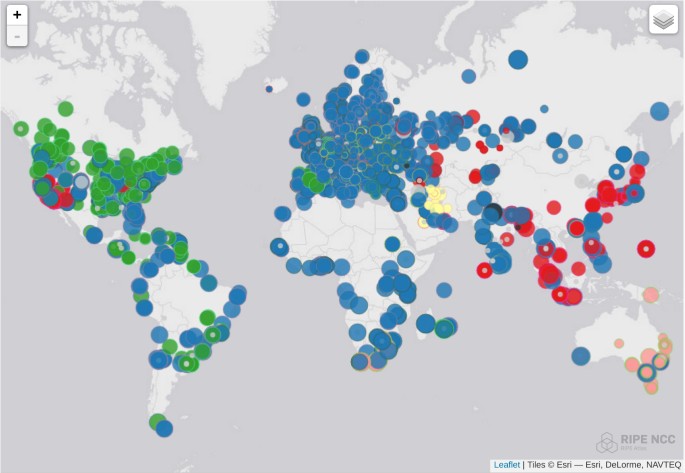
don't fully control
routing and catchments

hard
to make informed choices

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
    - CHAOS query identifies *server* and implies *site*
    - targets *letters*, not Root DNS (cannot switch letter)
  - global, but heavily biased to Europe
  - we map *server->site*
    - map will be public dataset
- RSSAC-002 reports
  - self-reports from letters
  - not guaranteed when under stress
- BGPmon routing
  - control plane

6996 RIPE Atlas VPs on 2015-11-30 (looking at K-Root)



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: 34Gb/s (unreflected)
- 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)
  - but want to do better

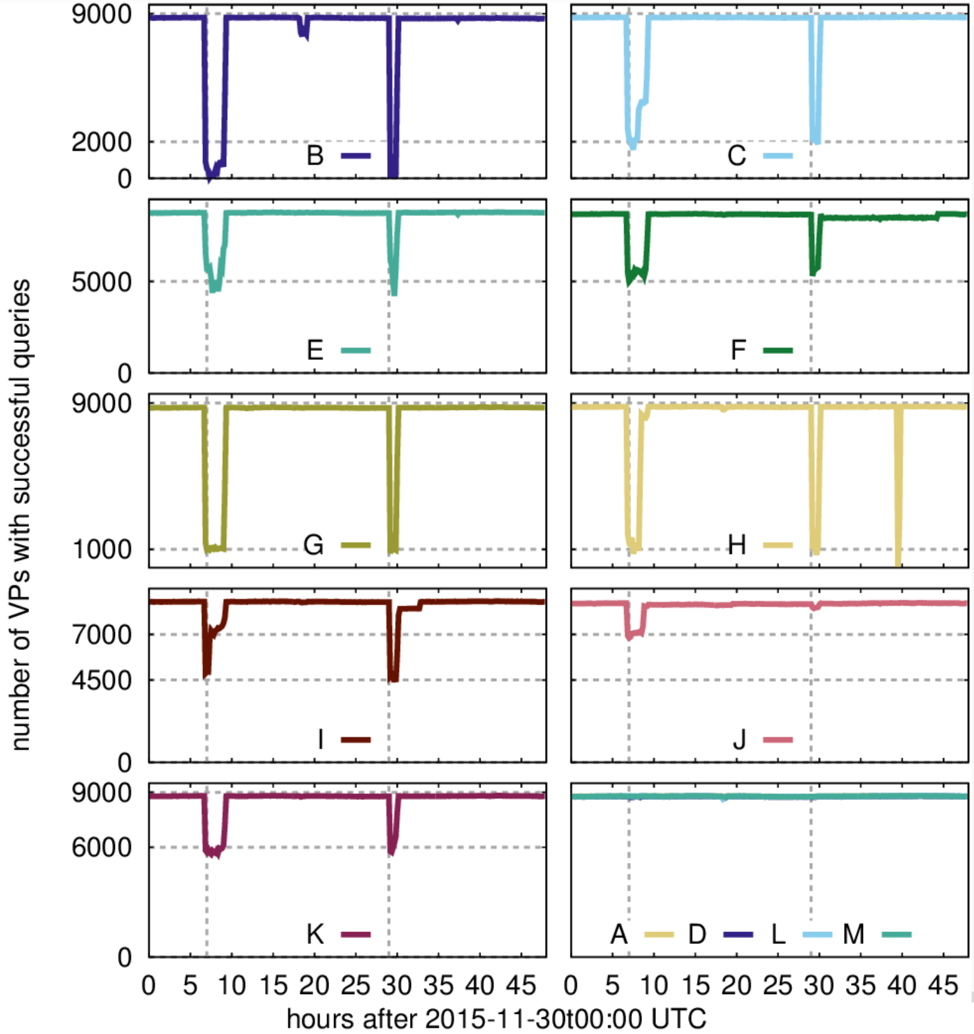
data:
A-Root had full view
(Verisign presentation);
RSSAC-002 reports

How About the Letters?

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

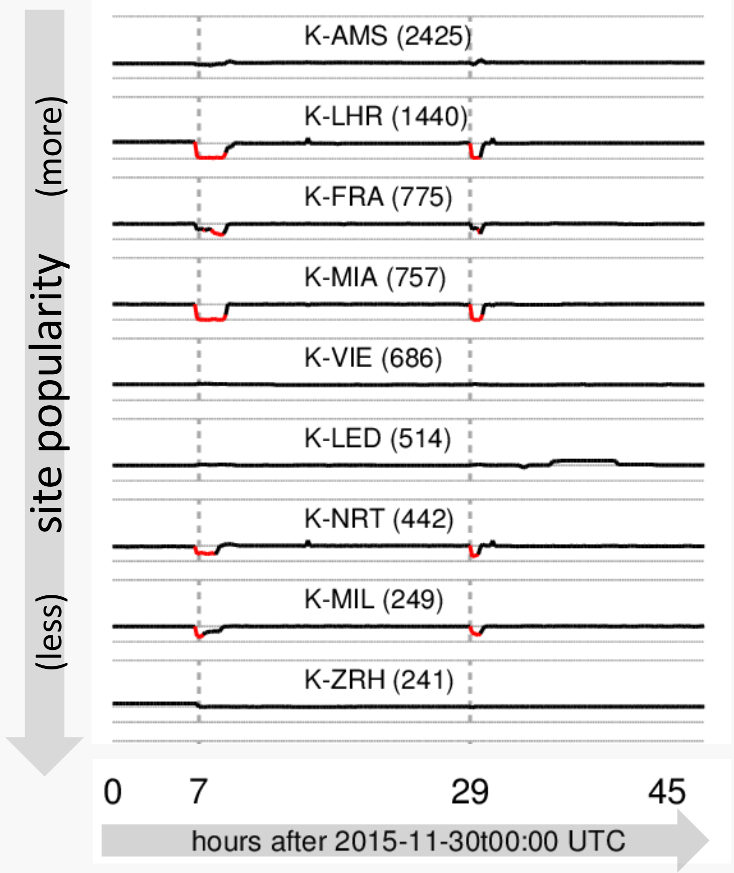
most suffered:
a bit (E, F, I, J, K)
or a lot (B, C, G, H)

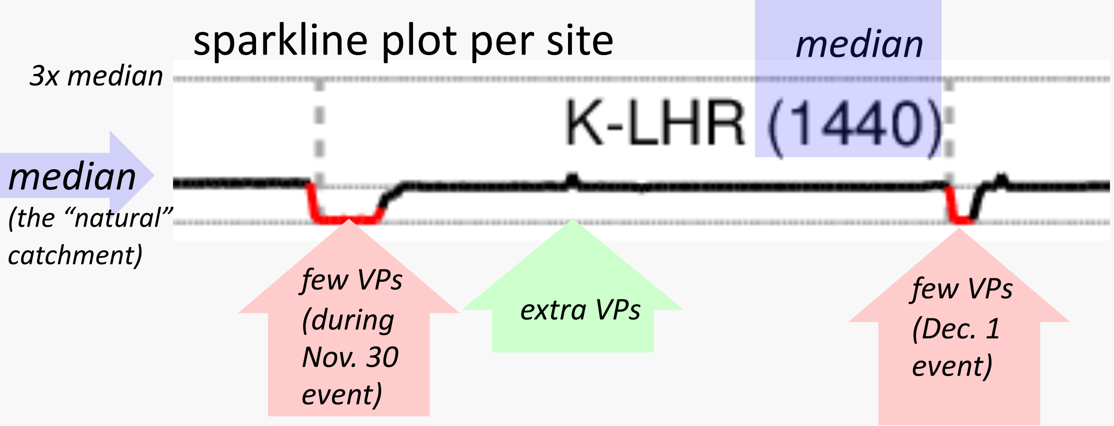
but does “x%”
measure what
users actually see?



[Moura16a, figure 3; data: RIPE Atlas]

Reachability at K's Sites

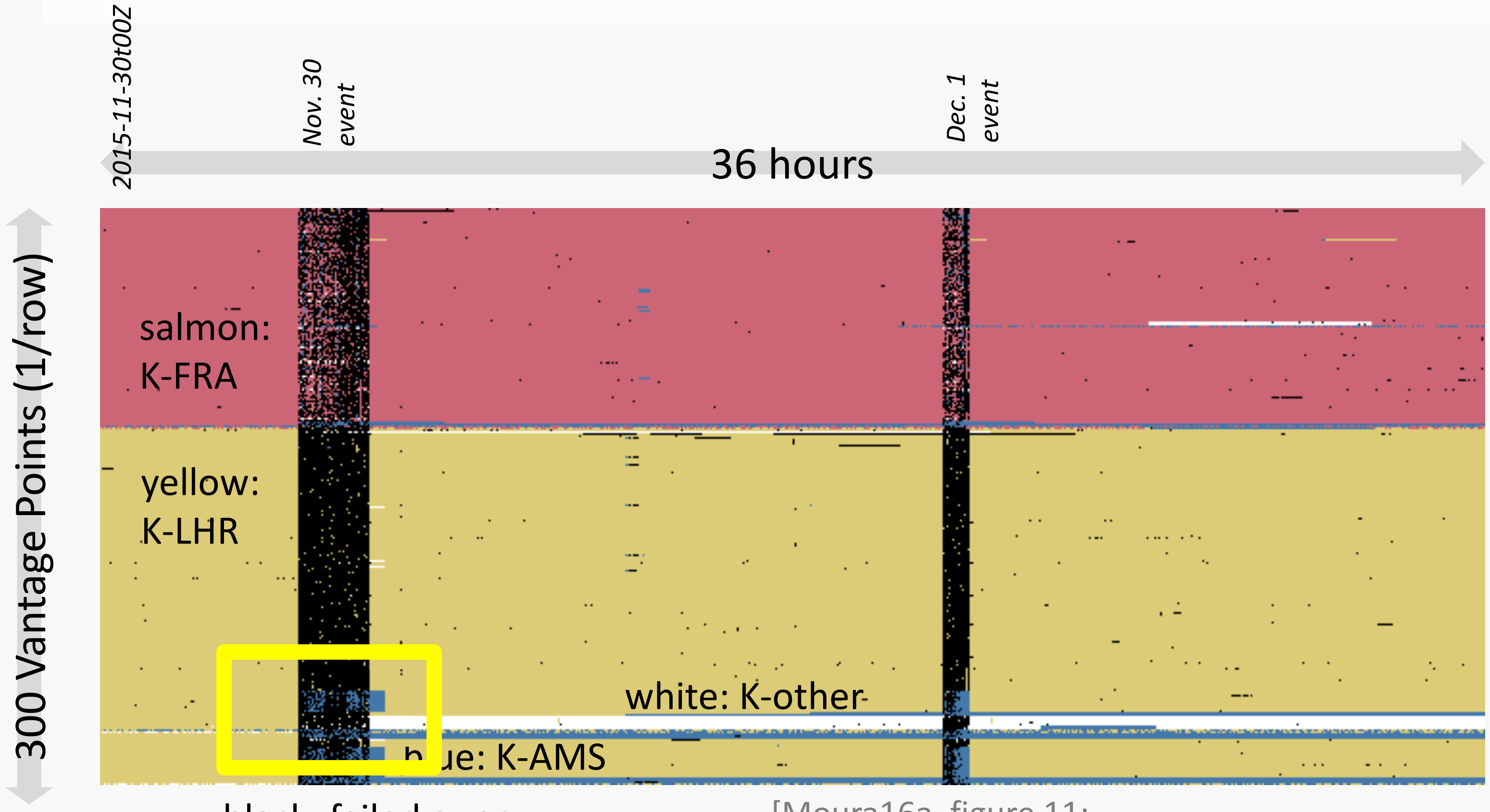




sites see fewer VPs, but why?

- query loss?
- route change?

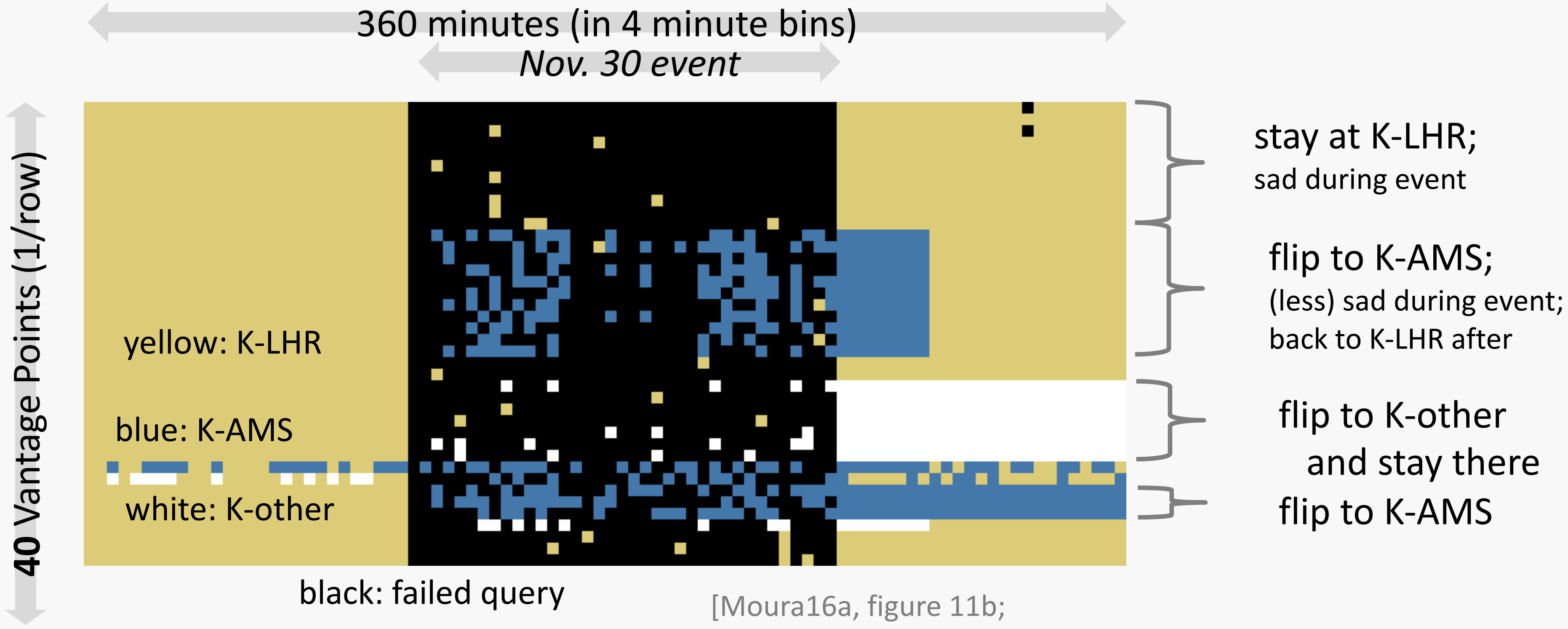
Site *Flips* from Routing Changes



black: failed query

[Moura16a, figure 11; data: RIPE Atlas]

Site *Flips* from Routing Changes



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

Flips: Implications

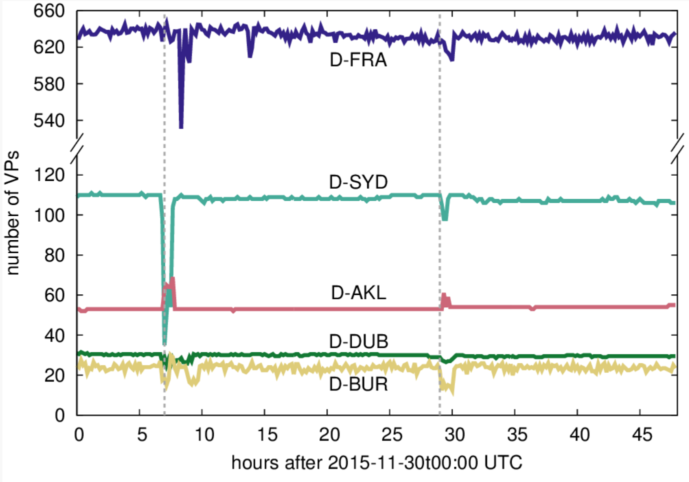
- some ISPs are “sticky” and won’t flip
  - will suffer if their site is overloaded
- some ISPs will flip
  - but new site may not be much better
- result depends on many factors
  - actions taken by root operator
  - routing choices by operator *and peer*
    - and perhaps *peer’s peers*, depending on congestion location
  - implementation choices
    - DNS, routing

During An Event: Active Routing Changes or Not?

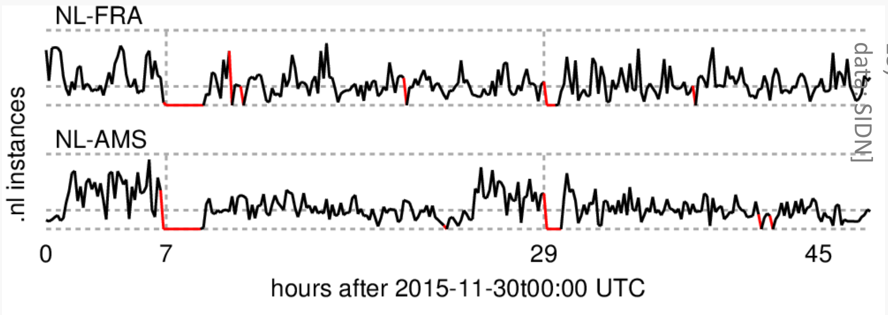
- no active routing changes
  - should expect partial loss in future attacks
    - inevitable: non-uniform attacker and defender capacity
  - overloaded catchments will suffer during attack
  - need to pre-deploy excess capacity
  - *operators understand and are doing these; but what about user expectations?*
- active routing changes
  - important when aggregate attack and defense capacity is similar
    - if one exceeds the other, no need to bother
  - requires *much* better measurement and route control
    - seems like a research problem; AFAIK no tools today
  - important to reduce client losses at smaller sites
  - *seems necessary to get to 0% loss*

Aside: Collateral Damage

- can an event hurt non-targets?
- *yes!* …a risk of shared datacenters



[Moura16a, figure 14; data: RIPE atlas]



[Moura16a, figure 15; data: SIDN]

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

D-FRA and D-SYD: less traffic

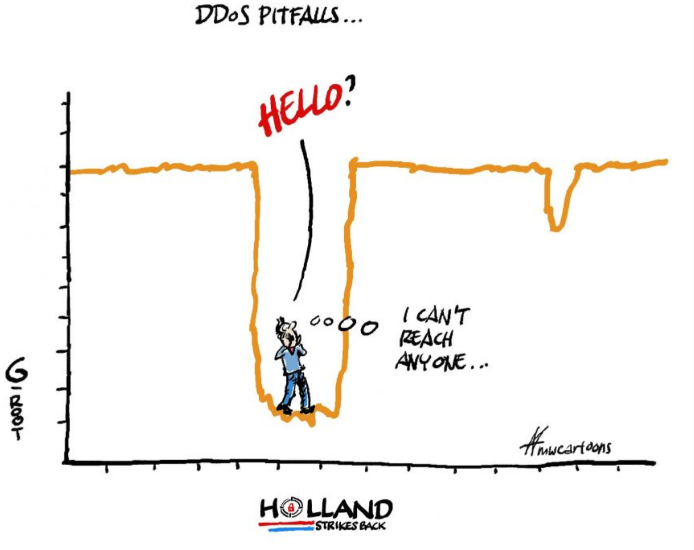
(even though D was not directly attacked)

Recommendations

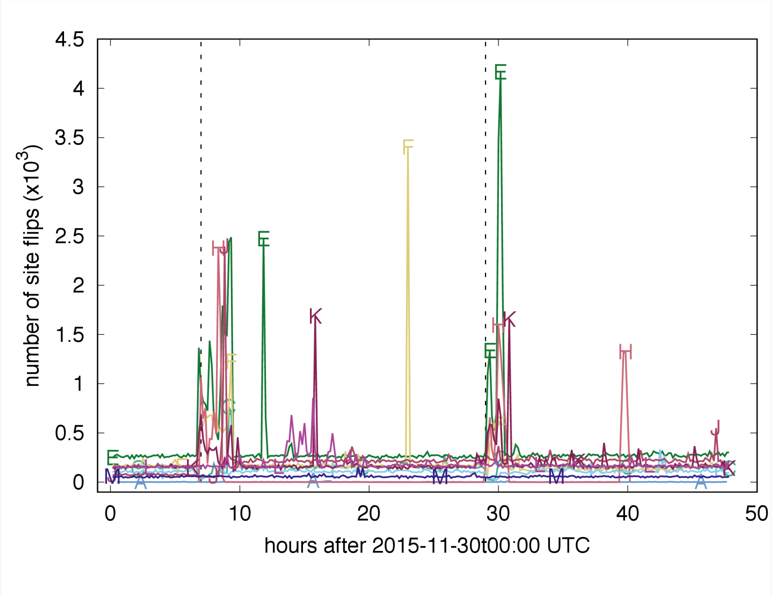
- current approach reasonable
  - build out capacity in advance
  - no active re-routing during attack
  - should expect some loss during each attack
- need true diversity to avoid collateral damage
- longer-term
  - need research to improve measurement and control
  - active control can improve loss during some attacks
- how many sites needed?
  - there is a *lot* of capacity already
  - many small sites seem to increase partial outages

More Info

- paper: http://www.isi.edu/~johnh/PAPERS/Moura16b
- data: https://ant.isi.edu/datasets/anycast/

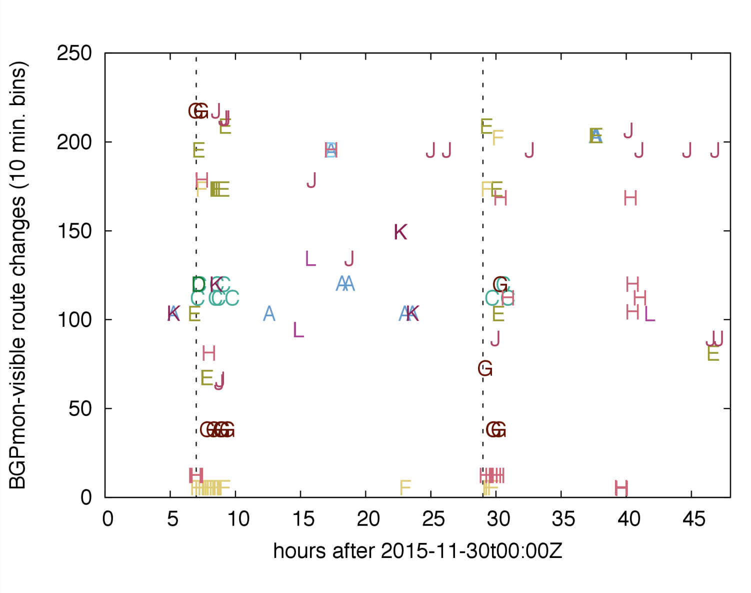


Confirming Flips in BGP



[Moura16a, figure 8; data: RIPE Atlas]

flips common during events for most letters



[Moura16a, figure 9; data: BGPmon]

flips seen in BGP