IODA-NP: Detecting outages affecting the Internet's edge

Ramakrishna Padmanabhan, Alistair King, Philipp Winter,
Marina Fomenkov, Alberto Dainotti

UC San Diego

Measuring Internet outages is important

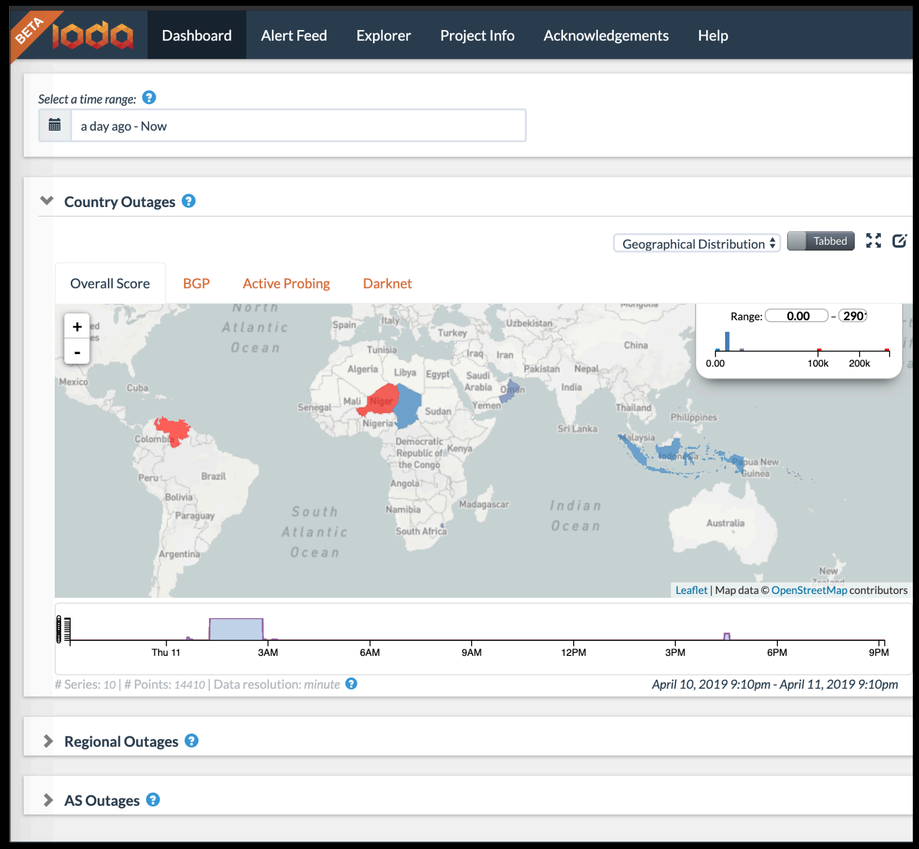
- Governments can monitor critical infrastructure to obtain situational awareness
- Users can compare reliability across providers
- ISPs can identify and diagnose problems

IODA: Internet Outage Detection and Analysis

Country

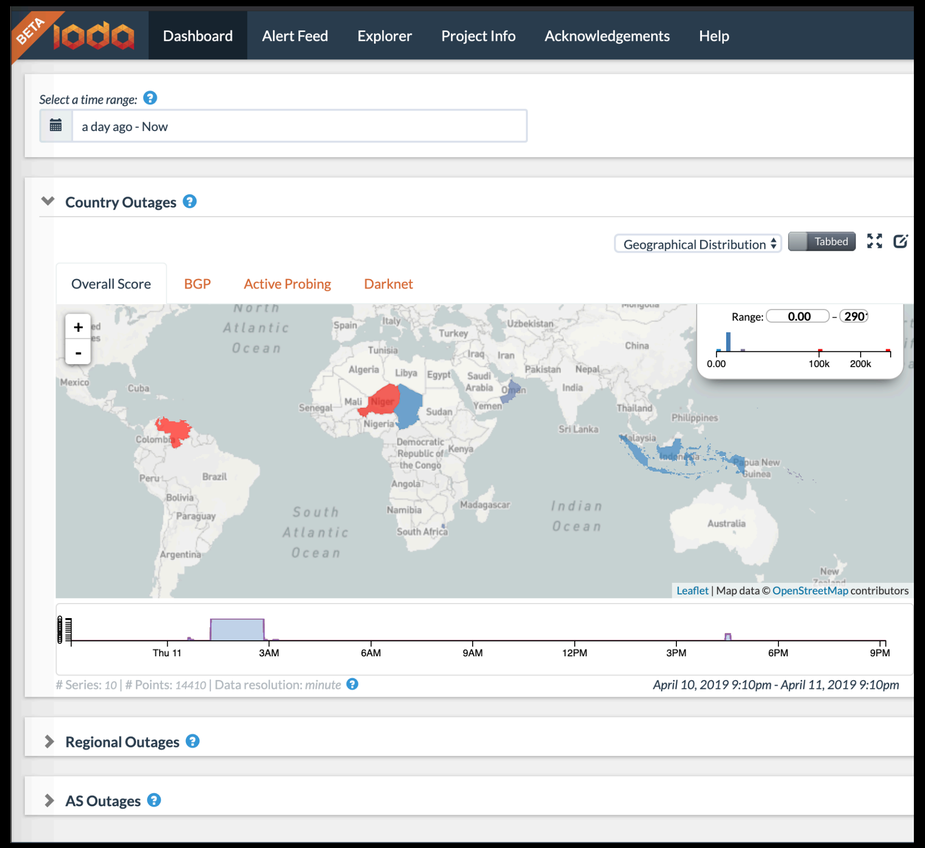
Region

AS



ioda.caida.org

IODA detects outages using three complementary data sources

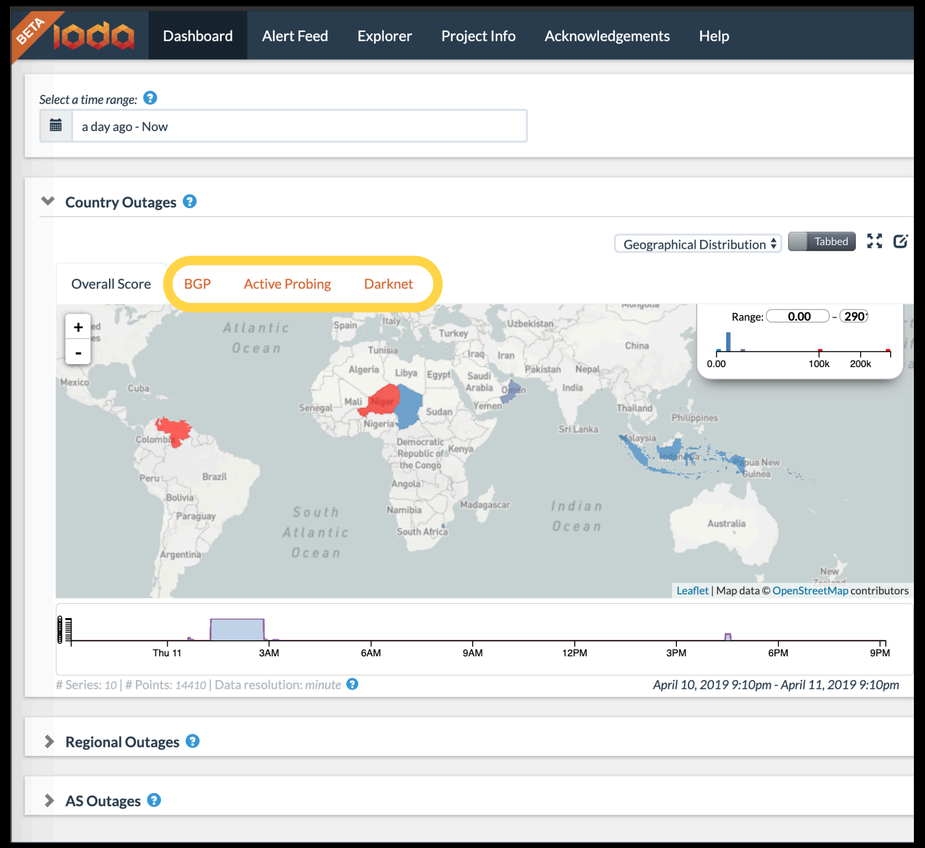


BGP: Detect when prefixes belonging to an aggregate lose control-plane connectivity

Active Probing: Detect lack of ping responses from /24 blocks in an aggregate

Darknet: Detect when traffic from an aggregate of addresses ceases

IODA detects outages using three complementary data sources

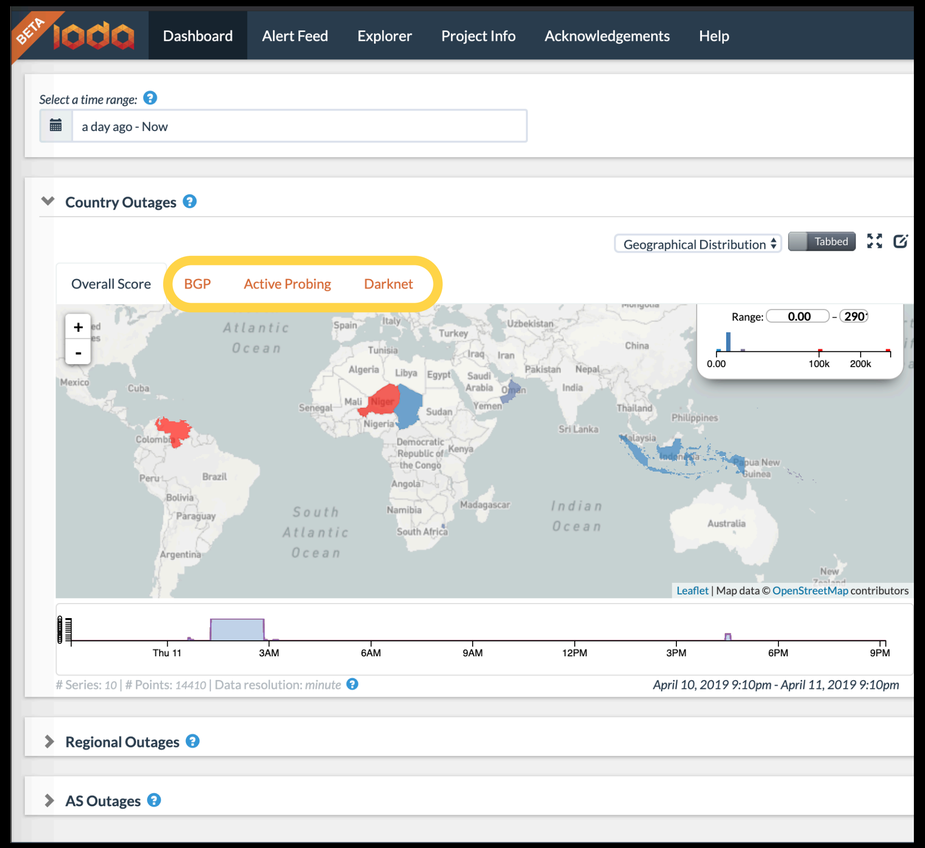


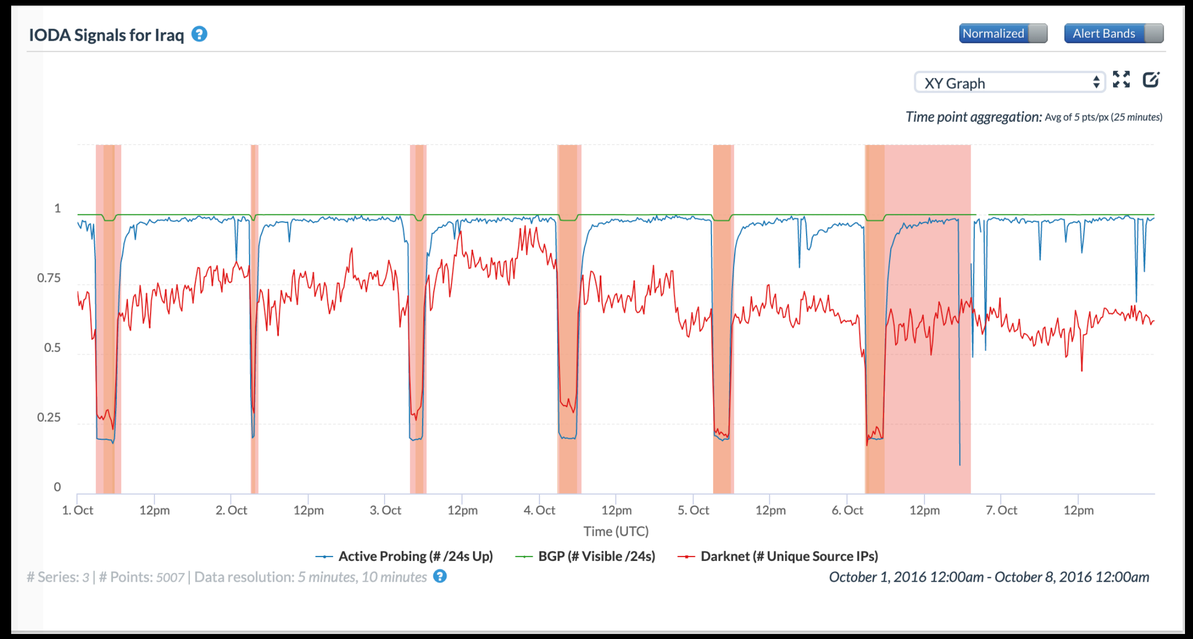
BGP: Detect when prefixes belonging to an aggregate lose control-plane connectivity

Active Probing: Detect lack of ping responses from /24 blocks in an aggregate

Darknet: Detect when traffic from an aggregate of addresses ceases

IODA detects outages using three complementary data sources





Internet outages in Iraq in Apr 2016

IODA-NP: DHS-funded project for the Next Phase

- Define the scope of the outages IODA can detect
- Detect outages at finer geographic granularity (such as county)
- Evaluate accuracy of detected outages
- Detect outages in near real-time

Prerequisite: Characterize Outages

- IP address dimension Do outages typically affect addresses from the same /24 block?
- Geographic dimension How are the addresses affected by an outage related by geography?
- Time dimension How long do outages last?

Existing systems allow only partial characterization

- Detecting Internet outages requires broad measurements
- Existing systems deal with this challenge by taking a top-down approach
  - They have some expectations about how outages will occur
  - They design systems to capture these outages

Existing systems allow only partial characterization

- Trinocular looks for outages that span an entire /24 block
- Thunderping detects outages occurring during times of predicted severe weather
- IMC ’18 work using CDN logs focuses upon detecting outages that last a full calendar hour
- IODA detects outages using the network telescope when many addresses in an aggregate stop contacting it

Measurements with active probes have evolved since the early 2010s

- Trinocular and Thunderping probe conservatively
- Recent work with active probing suggests we can be less conservative

Characterize outages using active probes but with minimal assumptions

- Some addresses should respond to active probes
- Outages will last at least X minutes

Towards a better understanding of outages

1. Measure broadly:

- Probe all addresses
- Probe regularly

2. Handle noise:

- Addresses can “fail” due to user action
- Use statistical tests to discard noise

3. Characterize outages along:

- IP dimension
- Geographic dimension
- Time dimension

4. Correlate with related data sources:

- Weather data
- Power outage data

Towards a better understanding of outages

1. Measure broadly:
- Probe all addresses
- Probe regularly

2. Handle noise:
- Addresses can “fail” due to user action
- Use statistical tests to discard noise

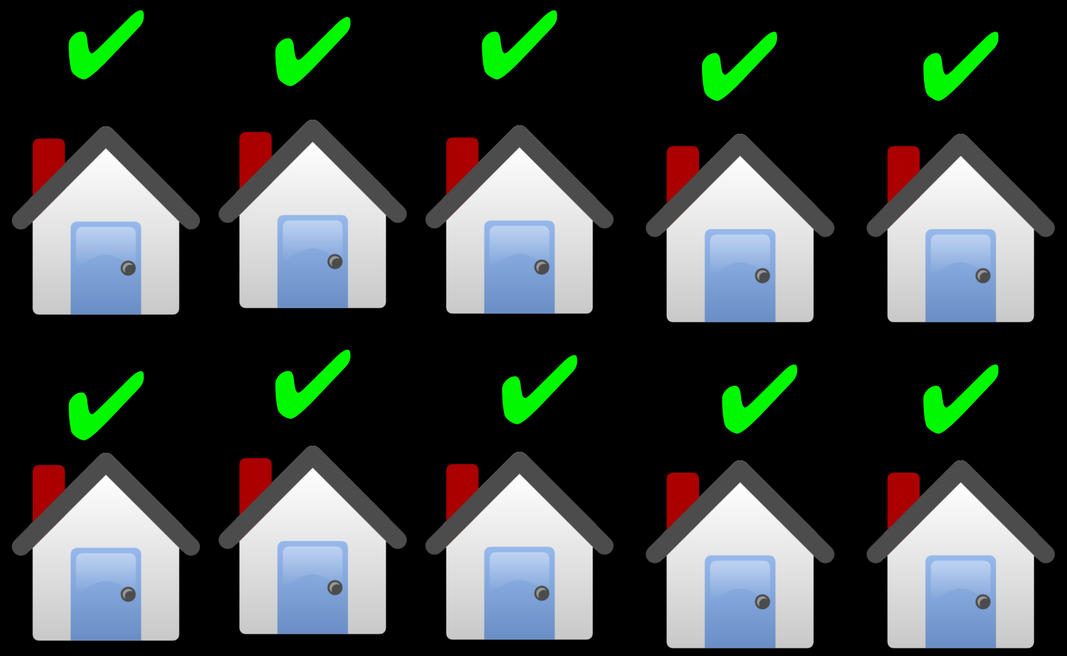
3. Characterize outages along:
- IP dimension
- Geographic dimension
- Time dimension

4. Correlate with related data sources:
- Weather data
- Power outage data

Simultaneous outages could occur due to a common cause

An individual outage is hard to interpret

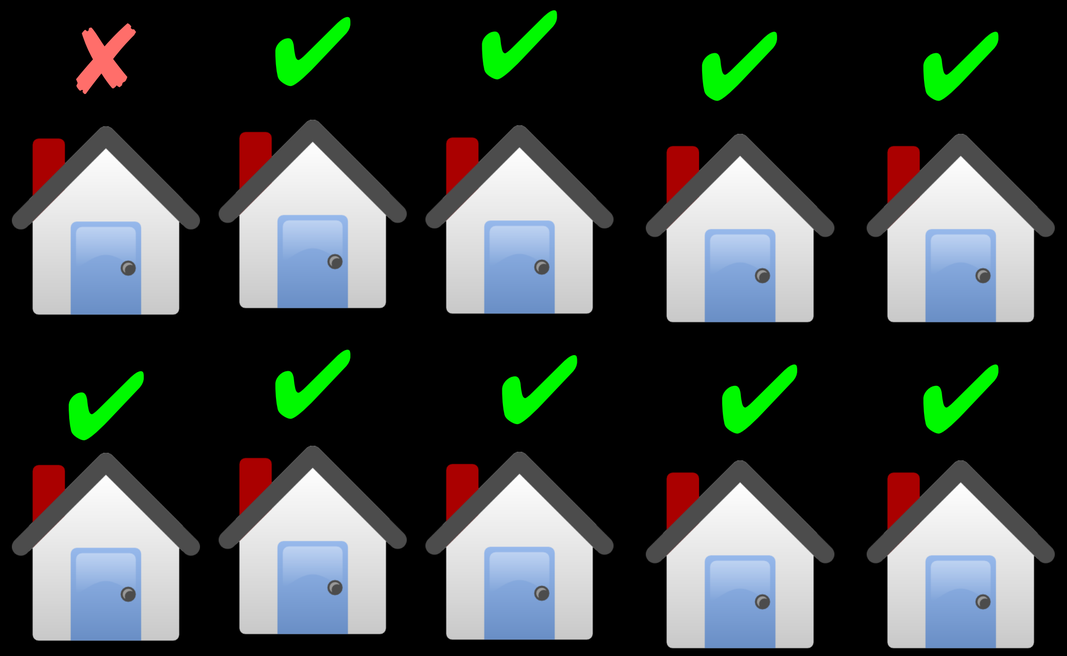
Common underlying cause would result in simultaneous outages



Simultaneous outages could occur due to a common cause

An individual outage is hard to interpret

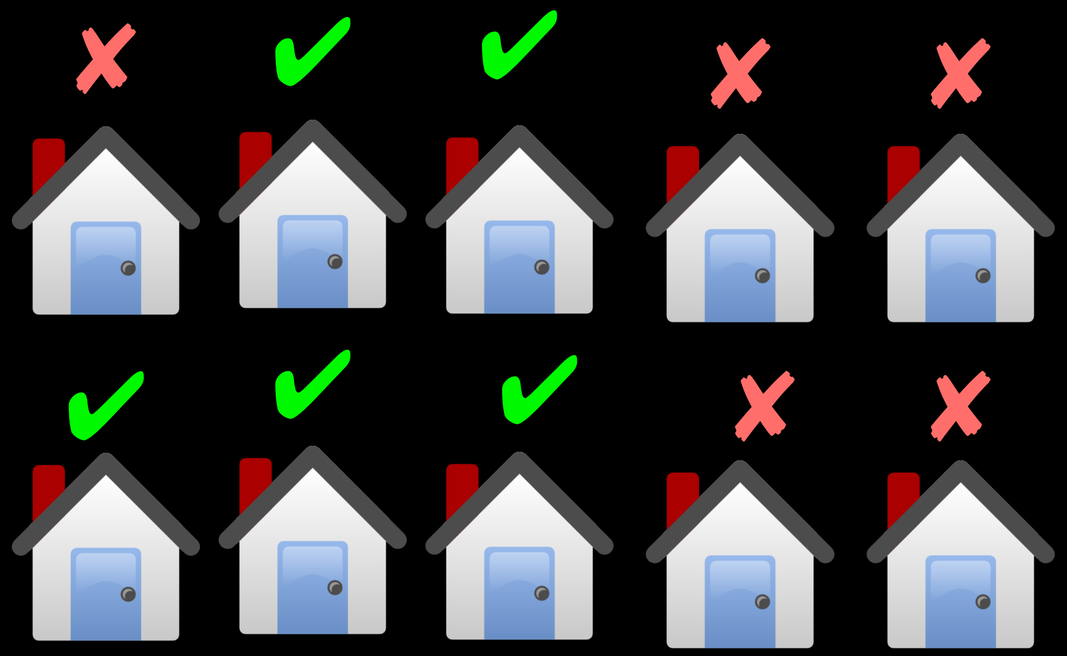
Common underlying cause would result in simultaneous outages



Simultaneous outages could occur due to a common cause

An individual outage is hard to interpret

Common underlying cause would result in simultaneous outages

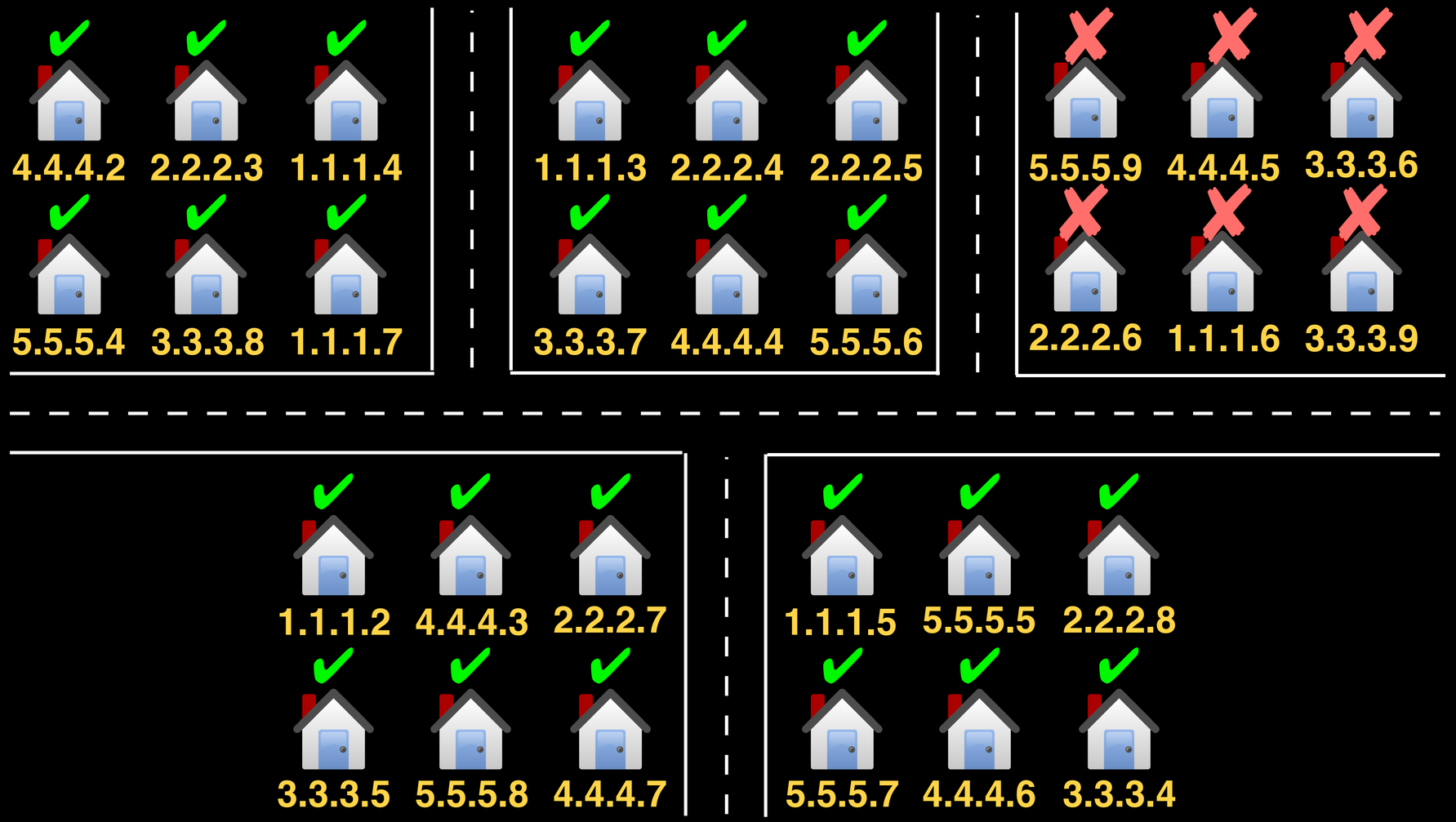


But simultaneous outages can also occur by random chance

But simultaneous outages can also occur by random chance

We thus identify simultaneous outages that are statistically unlikely

We thus identify simultaneous outages that are statistically unlikely



Binomial distribution gives the probability that D independent outages occur

$$\Pr[D \text{ independent outages}] = \binom{N}{D} \cdot P_d^D (1 - P_d)^{N-D}$$

- ***N***: # addresses in a bin of time that can potentially experience an outage
- P_d: Probability of independent outages of an address in a bin of time

Apply the Binomial test to identify statistically unlikely events

$$\Pr[D \text{ independent outages}] = \binom{N}{D} \cdot P_d^D (1 - P_d)^{N-D}$$

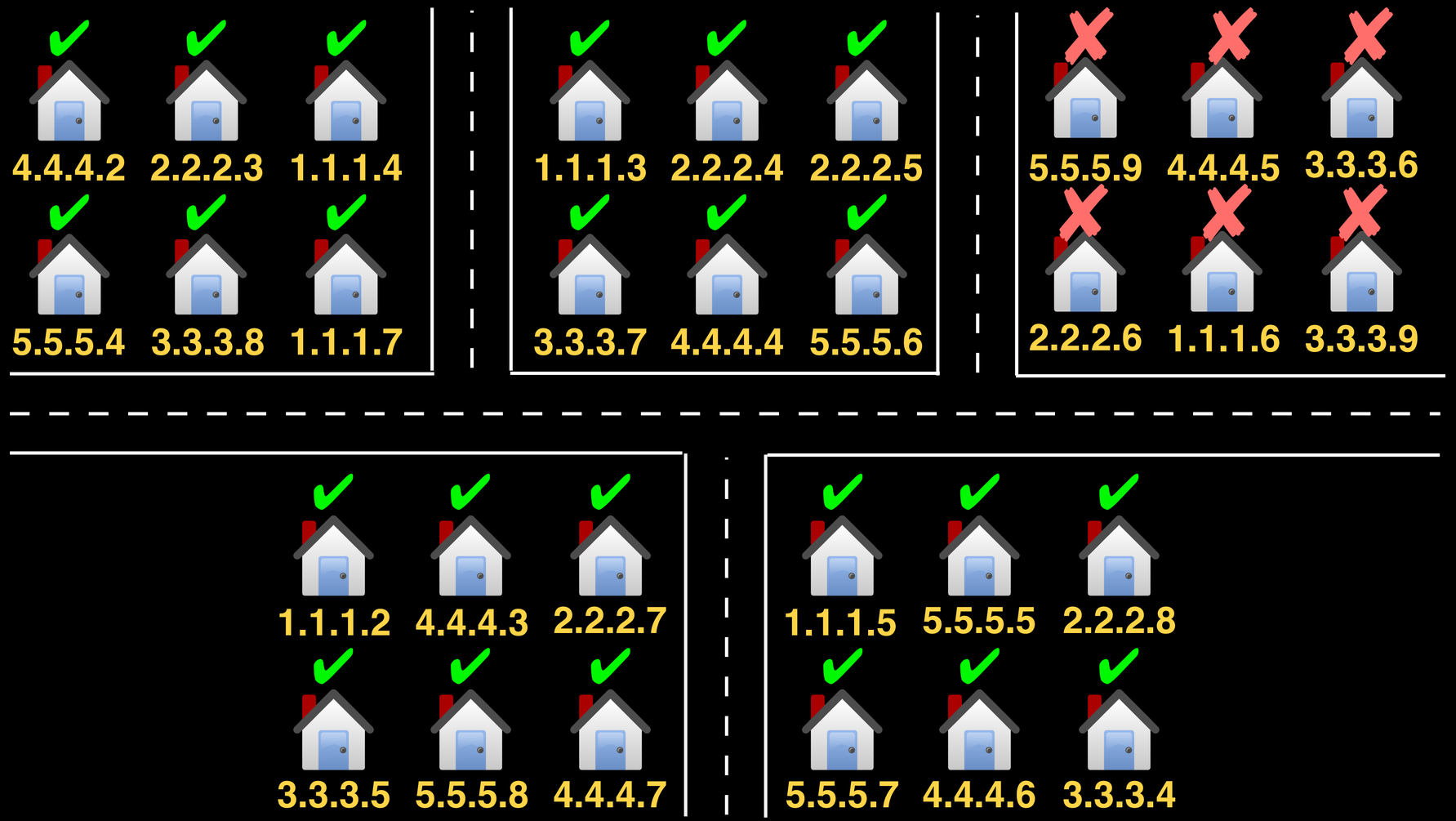
- We find D_{min} such that D_{min} or more independent outages occur with very small probability
- Proof by contradiction to find dependent events:
  - If D_{min} or more outages occur, the outages are highly likely to be dependent

Proof of concept on the Thunderping dataset [PAM ’19]

- Applied the binomial test to identify statistically unlikely outages of multiple addresses
- Studied their properties
  - The majority of dependent outages recover within an hour
  - They often do not affect entire /24 address blocks

Geographic neighbors aren’t necessarily neighbors in the address space

Geographic neighbors aren’t necessarily neighbors in the address space



Towards a better understanding of outages

1. Measure broadly:
- Probe all addresses
- Probe regularly

2. Handle noise:
- Addresses can “fail” due to user action
- Use statistical tests to discard noise

3. Characterize outages along:
- IP dimension
- Geographic dimension
- Time dimension

4. Correlate with related data sources:
- Weather data
- Power outage data

Towards a better understanding of outages

1. Measure broadly:
- Probe all addresses
- Probe regularly

2. Handle noise:
- Addresses can “fail” due to user action
- Use statistical tests to discard noise

3. Characterize outages along:
- IP dimension
- Geographic dimension
- Time dimension

4. Correlate with related data sources:
- Weather data
- Power outage data

Measure broadly: Zeusping

- We expect to ping ~150M ping-responsive addresses in the U.S.
  - Each address will be pinged from 3 vantage points, once every 10 minutes
  - Each address will receive 432 pings a day
  - Total pings that will be sent in a day: 65 Billion
- We are investigating which infrastructure to run these measurements from
  - Ideally, we would have tens of vantage points and probing volume is spread across them

Towards a better understanding of outages

1. Measure broadly:

- Probe all addresses
- Probe regularly

2. Handle noise:

- Addresses can “fail” due to user action
- Use statistical tests to discard noise

3. Characterize outages along:

- IP dimension
- Geographic dimension
- Time dimension

4. Correlate with related data sources:

- Weather data
- Power outage data

Backup Slides

Comparison with related work

Prior Work	Failure Scale	Min Failure Duration	Scale
IODA	Detects “macroscopic” (at the moment)	10 minutes	Internet-wide
Trinocular	Detects when most addresses in a /24 are disrupted	11 minutes	Internet-wide
Richter et. al.	Detects when majority of active addresses in a /24 are disrupted	60 minutes	Internet-wide
Disco	Detects bursts of RIPE Atlas probe disconnects	O(seconds)	10,000 probes
Thunderping	Detects when a few individual addresses are disrupted	11 minutes	50,000 U.S. addresses