Estimating the Impact of BGP Prefix Hijacking

Pavlos Sermpezis [1], Vasileios Kotronis [2],
Konstantinos Arakadakis [2,3], Athena Vakali [1]

[1] DataLab, Informatics Dept., Aristotle University of Thessaloniki, Greece
[2] Institute of Computer Science, FORTH, Greece
[3] Computer Science Dept., University of Crete, Greece

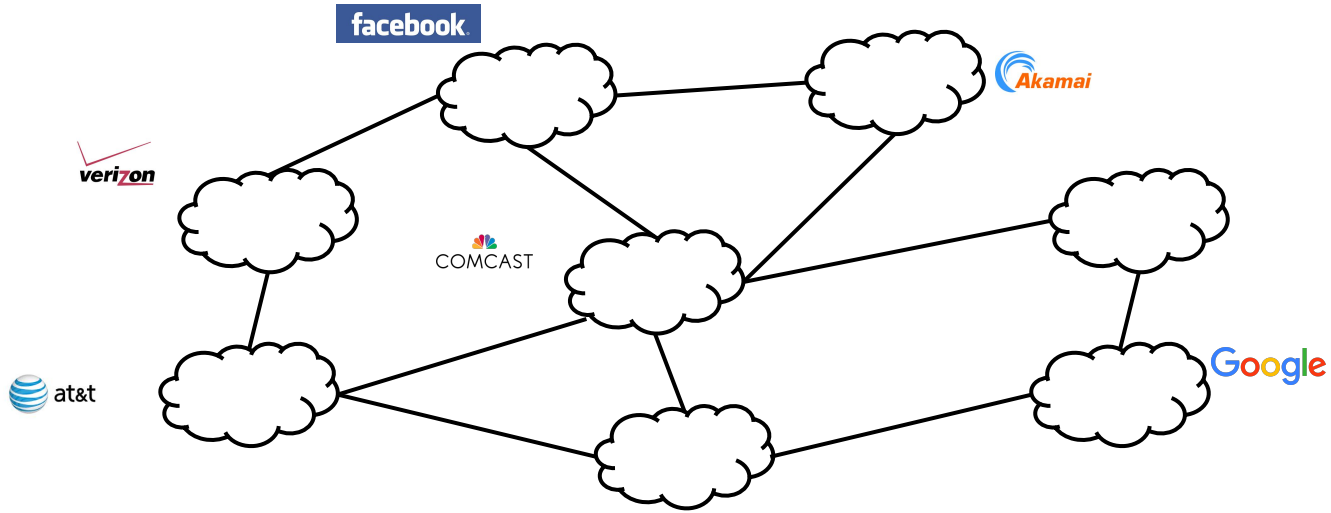
DATA&WEB SCIENCE LABORATORY

ARISTOTLE UNIVERSITY OF THESSALONIKI

FORTH INSTITUTE OF COMPUTER SCIENCE

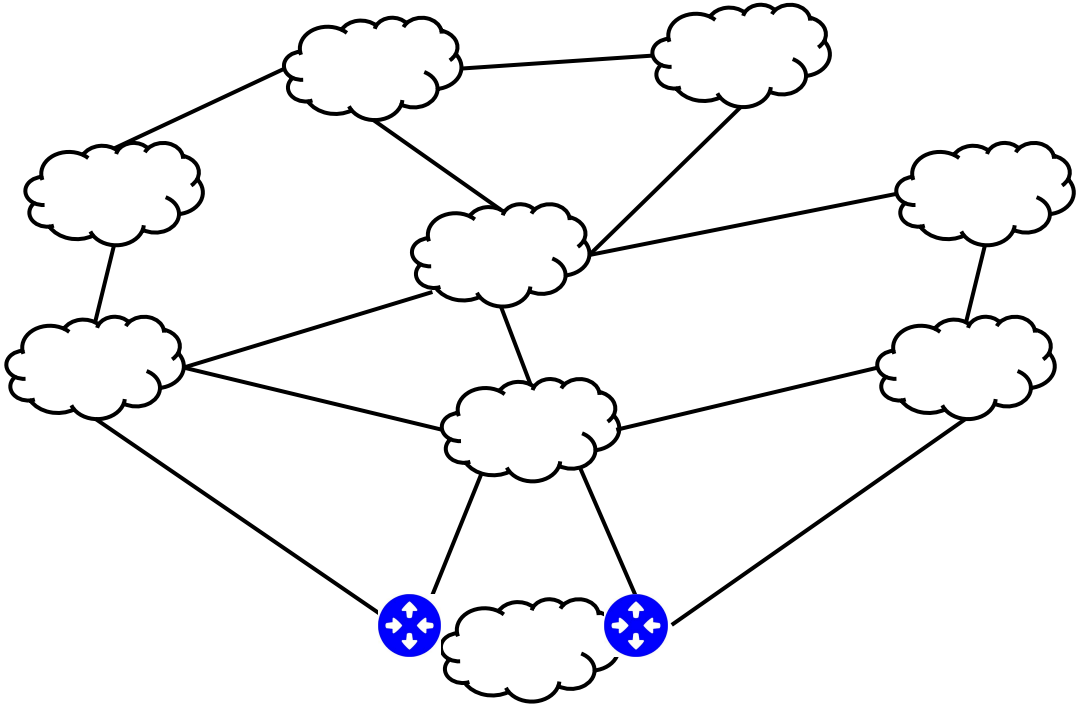
UNIVERSITY OF CRETE

The Internet



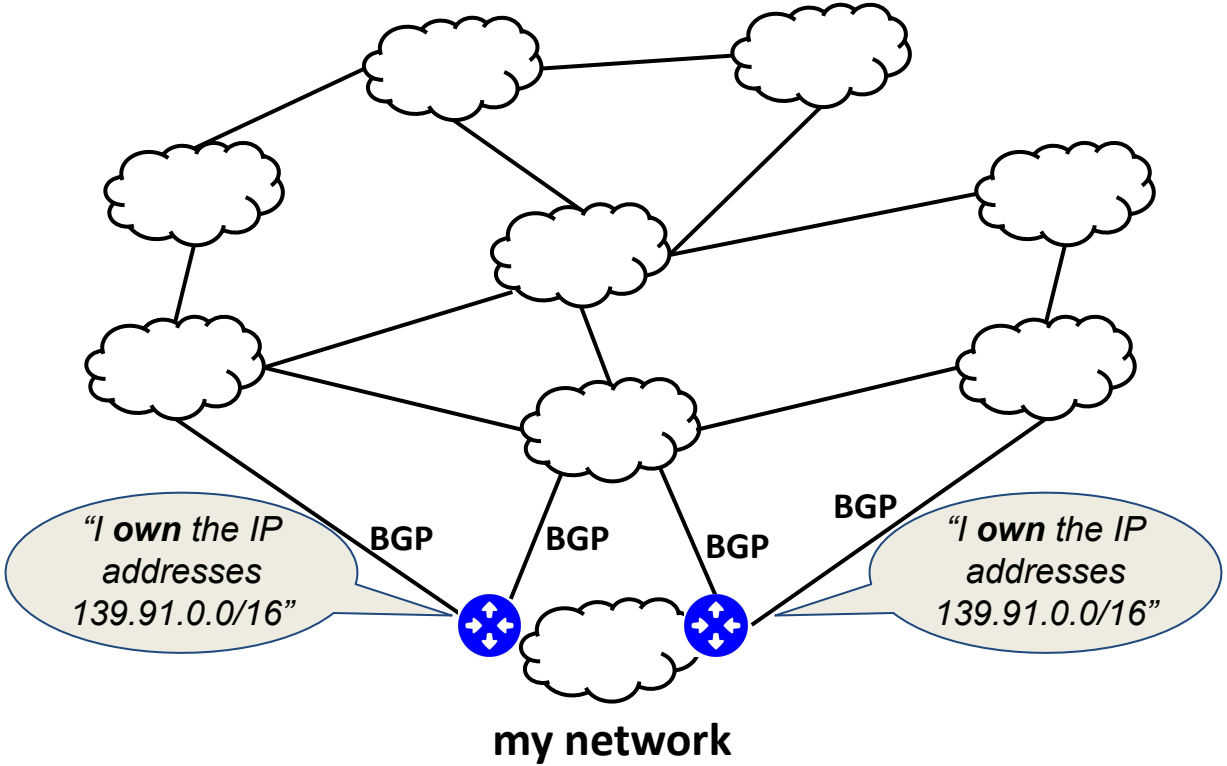
- The Internet is a network of networks or "**Autonomous Systems (AS)**"
- today ***~70k*** ASes

The Internet



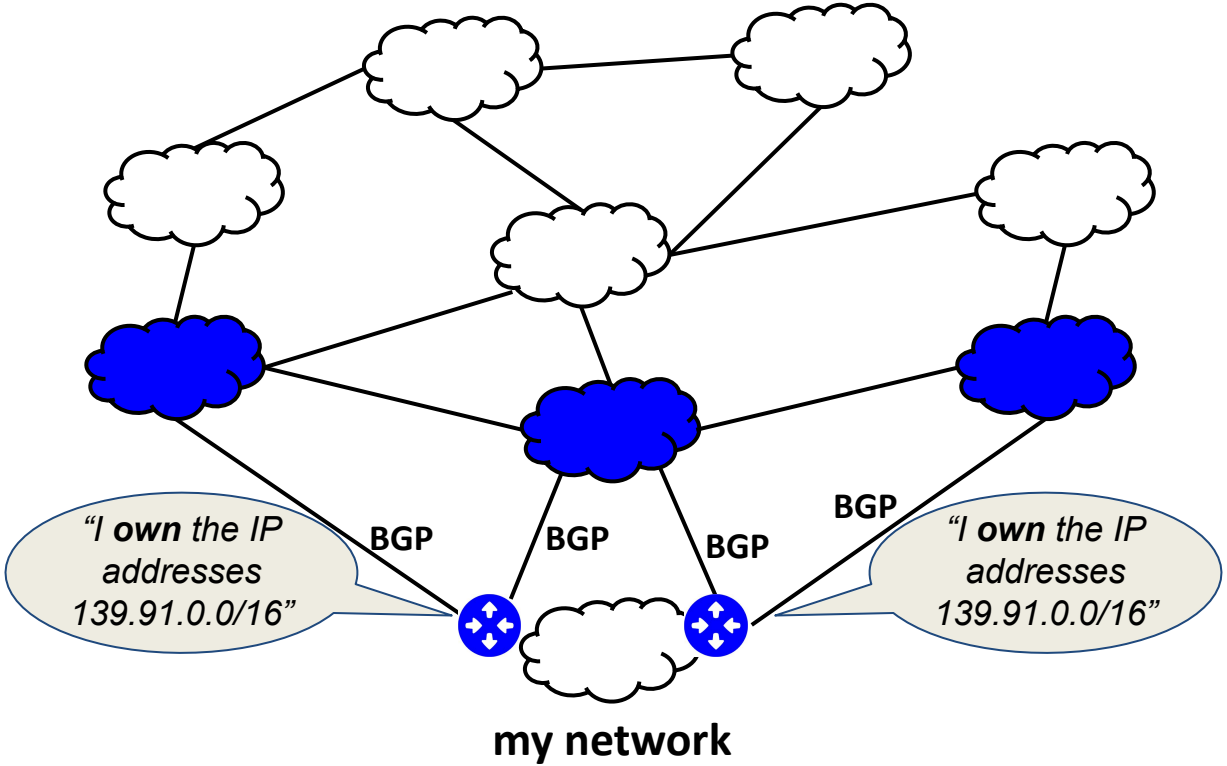
my network

Internet routing



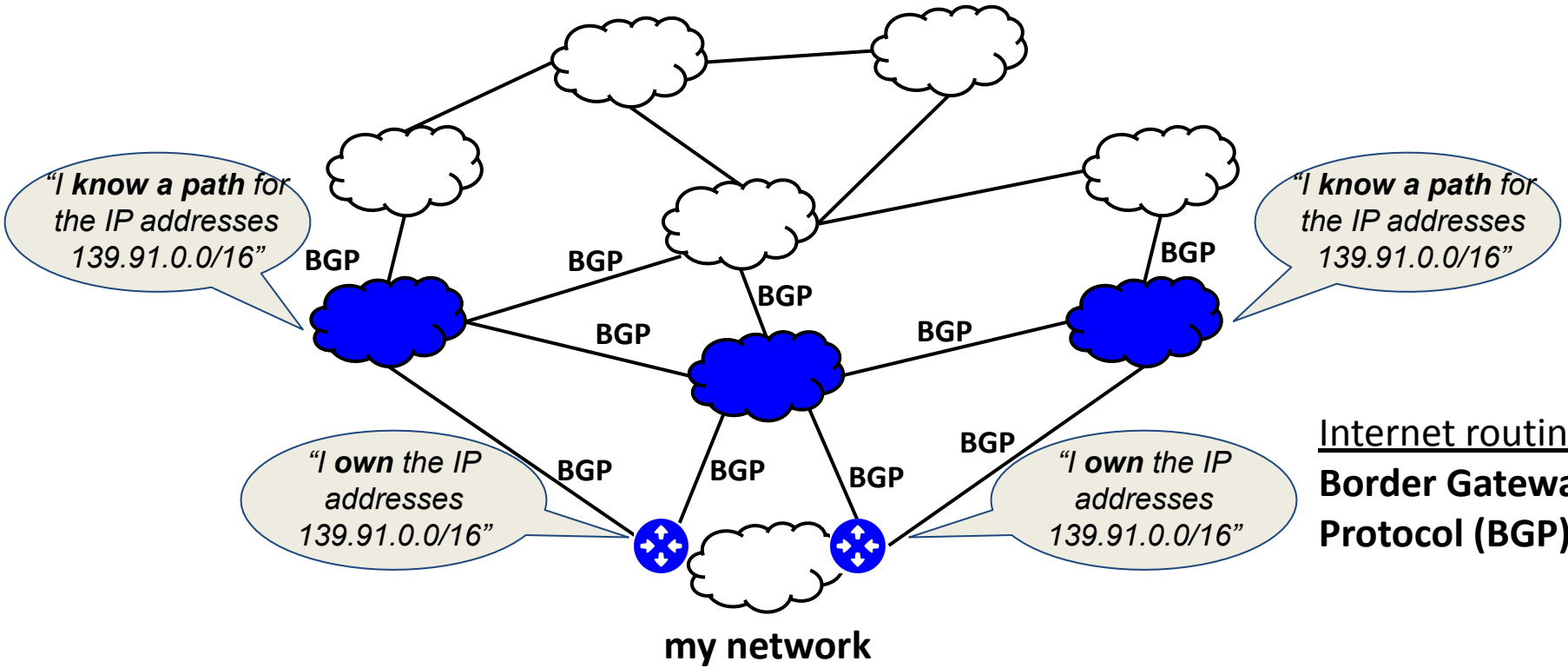
Internet routing:
Border Gateway Protocol (BGP)

Internet routing



Internet routing:
Border Gateway Protocol (BGP)

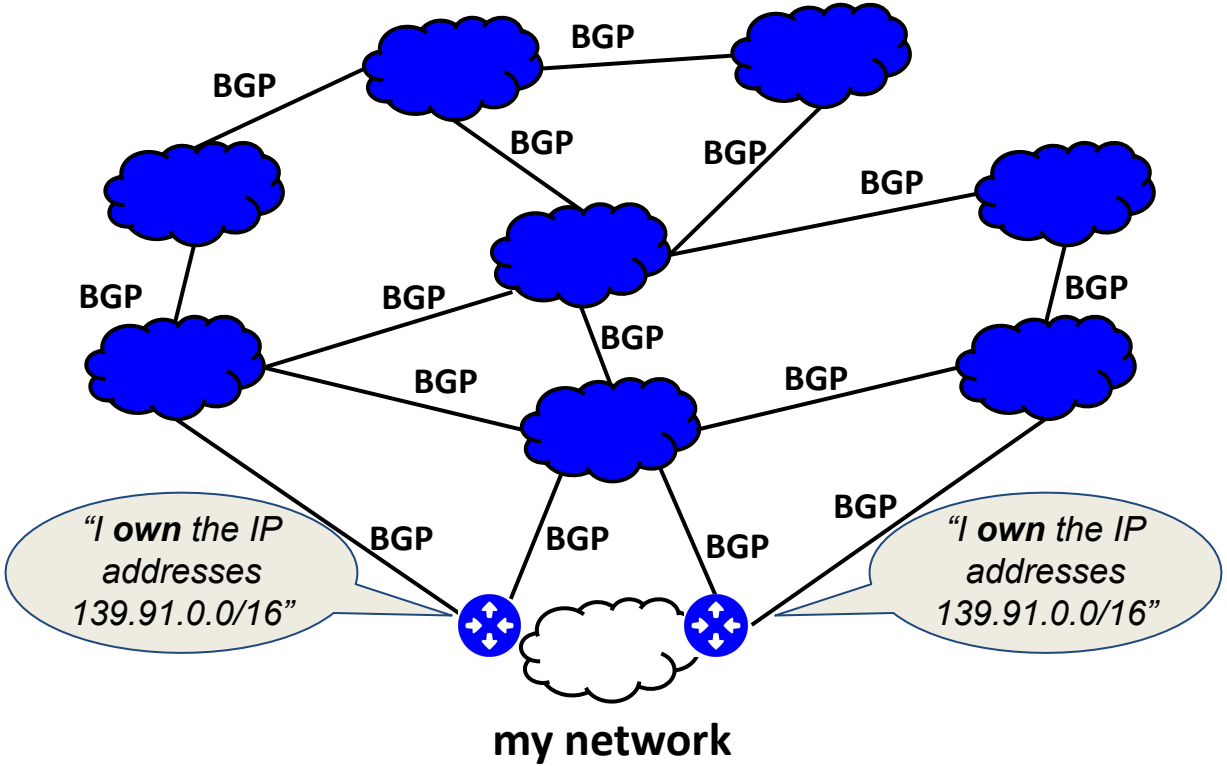
Internet routing



Internet routing:

Border Gateway Protocol (BGP)

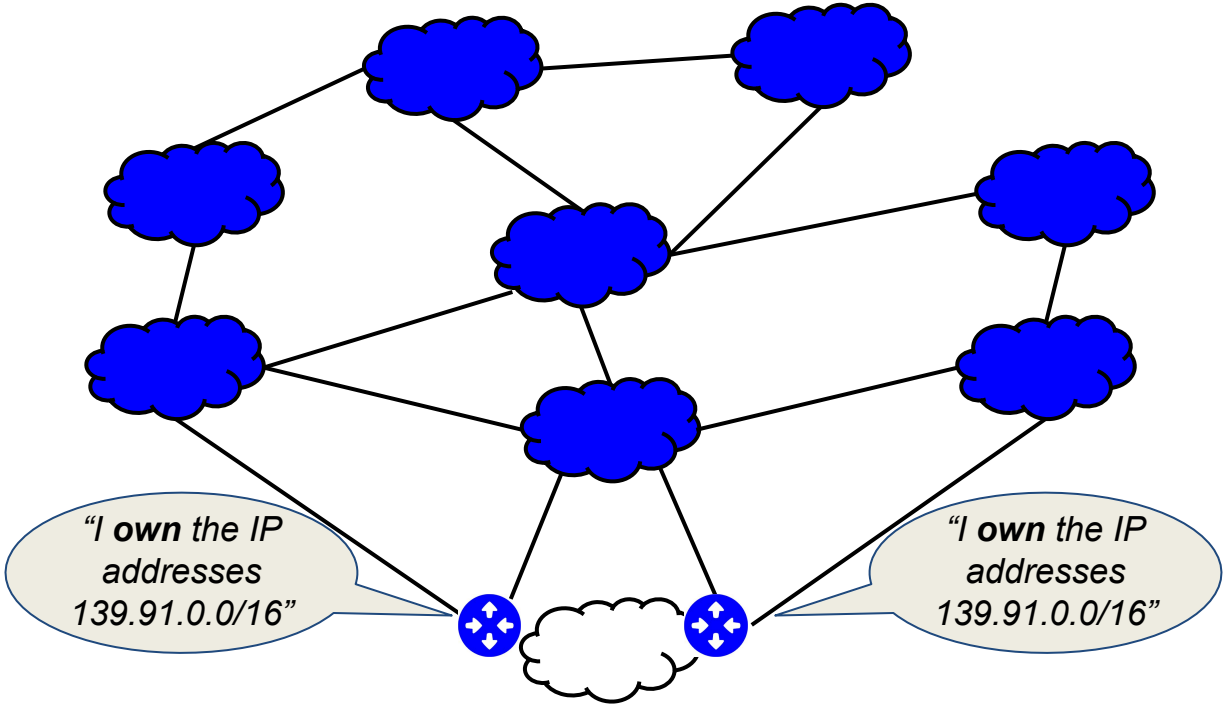
Internet routing



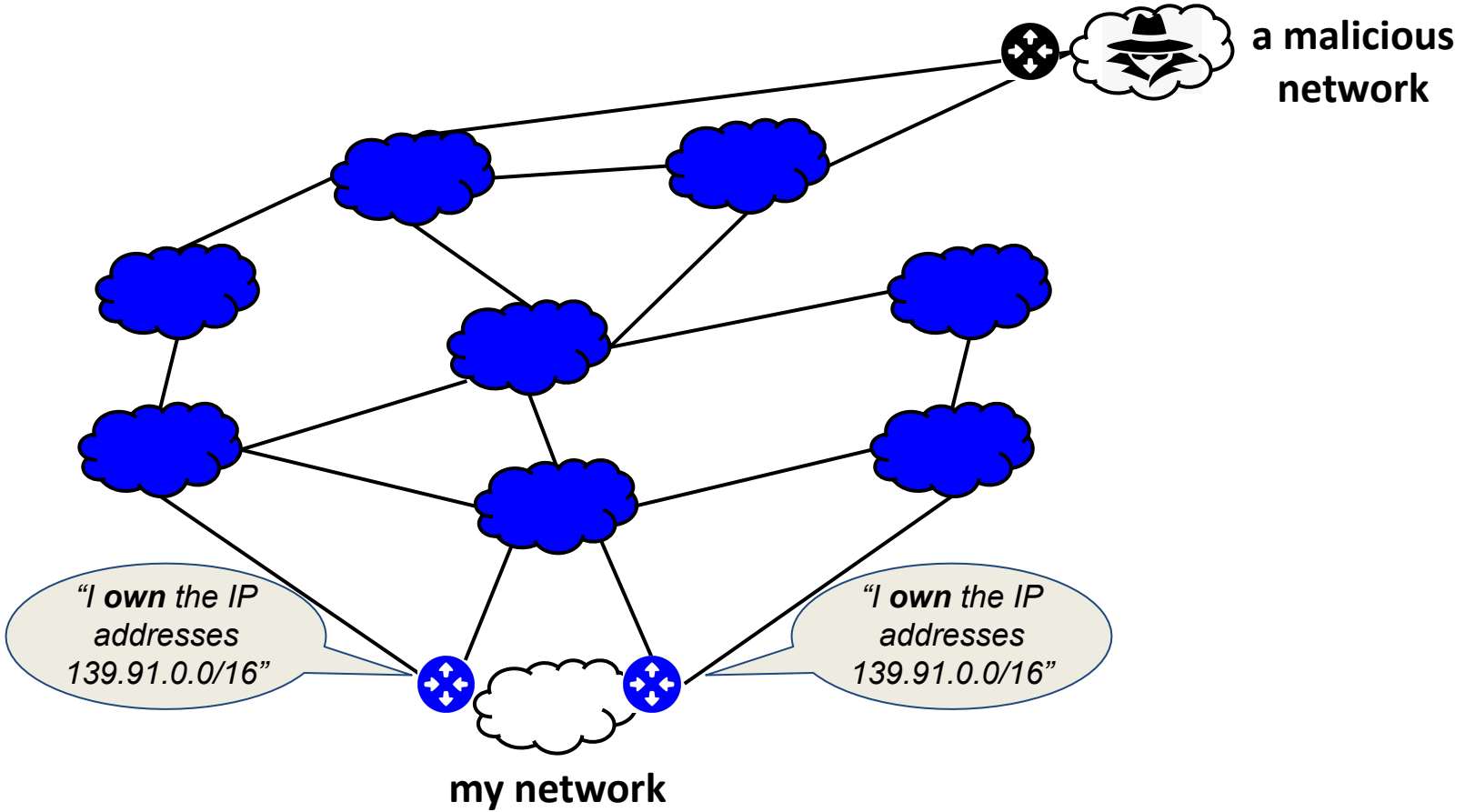
Internet routing:

Border Gateway Protocol (BGP)

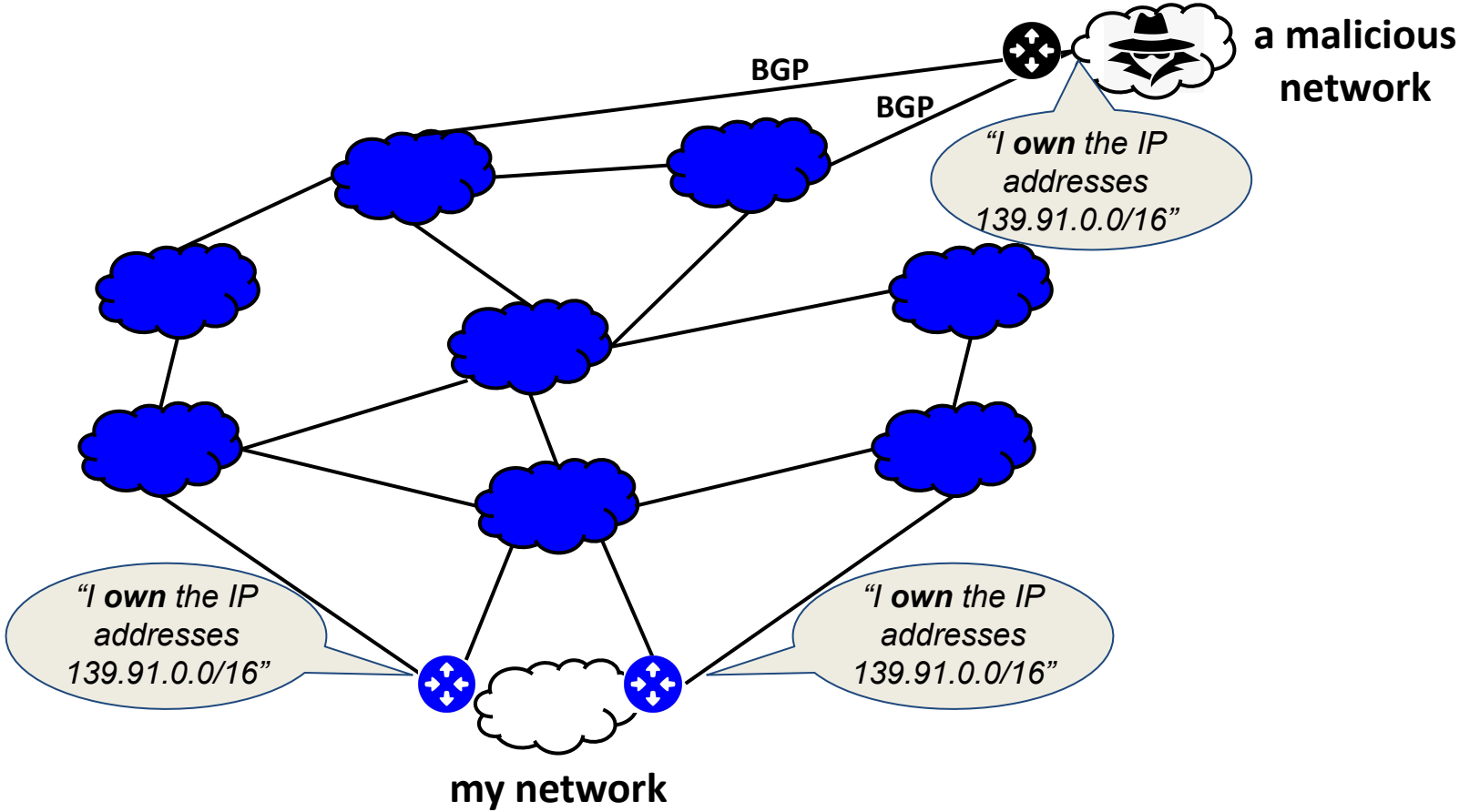
BGP prefix hijacking



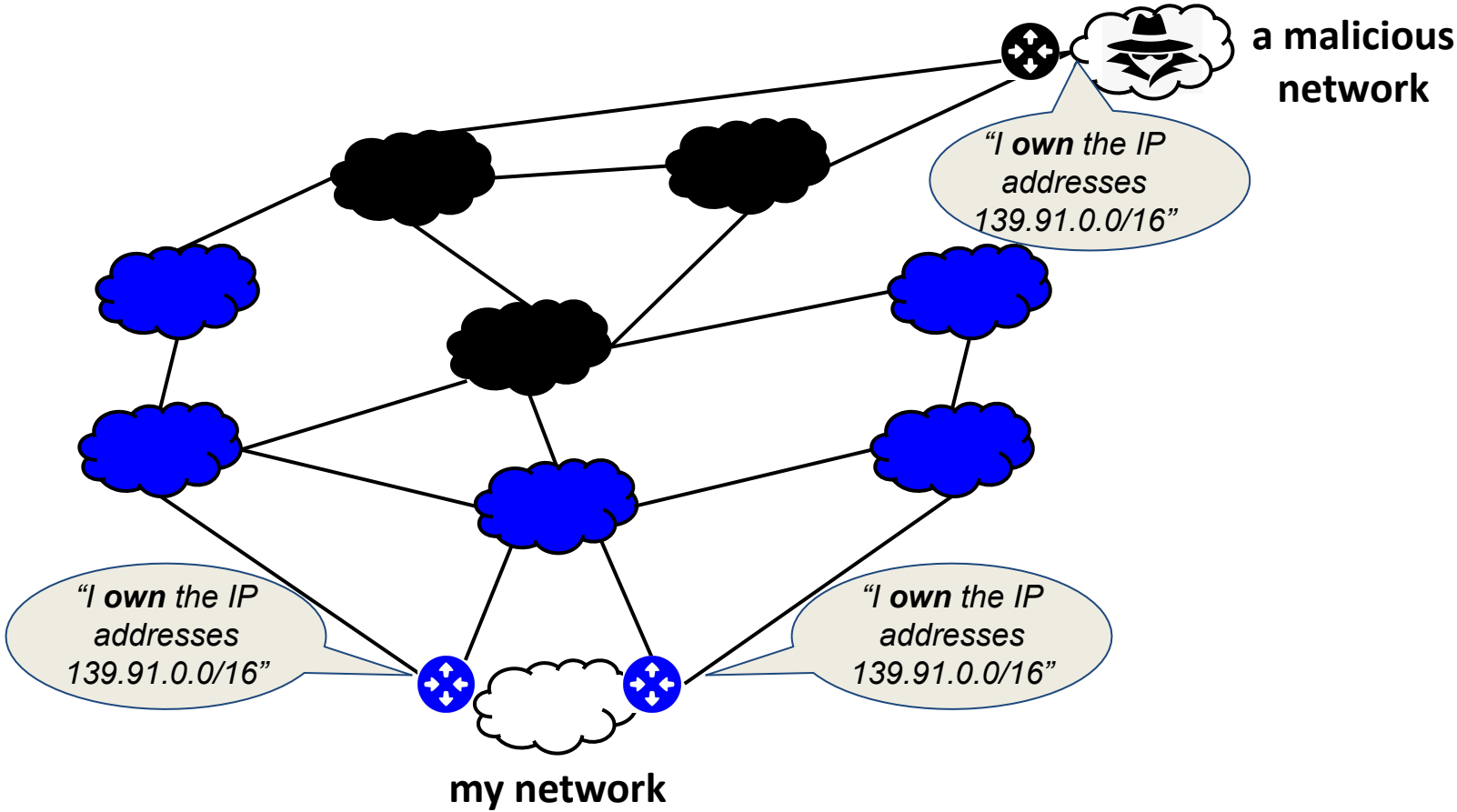
BGP prefix hijacking



BGP prefix hijacking



BGP prefix hijacking



BGP prefix hijacking

- **It's an important problem today! a few facts:**
  - ~2500 (reported) prefix hijacking events in 2020 [1]
  - examples of affected networks: Google, Amazon, Akamai, Visa, etc.

- **service outages & traffic interception**
  - can last for hours [2]
  - can cost $100k (or more) *per minute*!

- **no effective (proactive) defence**
  - RPKI: limited adoption & efficiency [2,3]
  - defences based upon detection & countermeasures [2]

[1] APNIC, *"BGP, RPKI, and MANRS: 2020 in review"*, Feb 2021, https://blog.apnic.net/2021/02/05/bgp-rpki-and-manrs-2020-in-review/
[2] P. Sermpezis, et. al., *"A survey among Network Operators on BGP Prefix Hijacking"*, in ACM SIGCOMM CCR, Jan 2018.
[3] NIST RPKI Monitor, https://rpki-monitor.antd.nist.gov/

How do we defend?

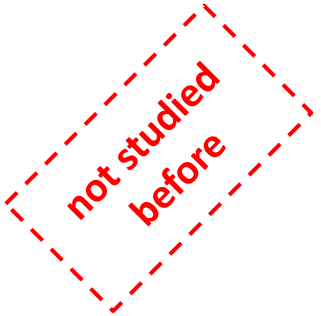
- **How do we defend against hijacks? → (mostly) reactively:**
  - *Step 1:* detect the hijack
  - *Step 2:* proceed to mitigation action
    - filtering, deaggregation, outsourcing (e.g., blackholing, anycast from large ISPs)
- **Detection**
  - ✓ a lot of research the last ~10 years
  - ✓ public monitoring infrastructure (RIPE RIS, RouteViews, etc.)
  - ✓ state-of-the-art: near real-time detection (in a few seconds) [1]
- **Mitigation**
  - ❓ different actions → different costs… *which one to choose?*
  - ❓ ok, I took an action… *was it effective? is the problem solved?*

→ we need to know the impact of the hijack (before/after its mitigation) !!!

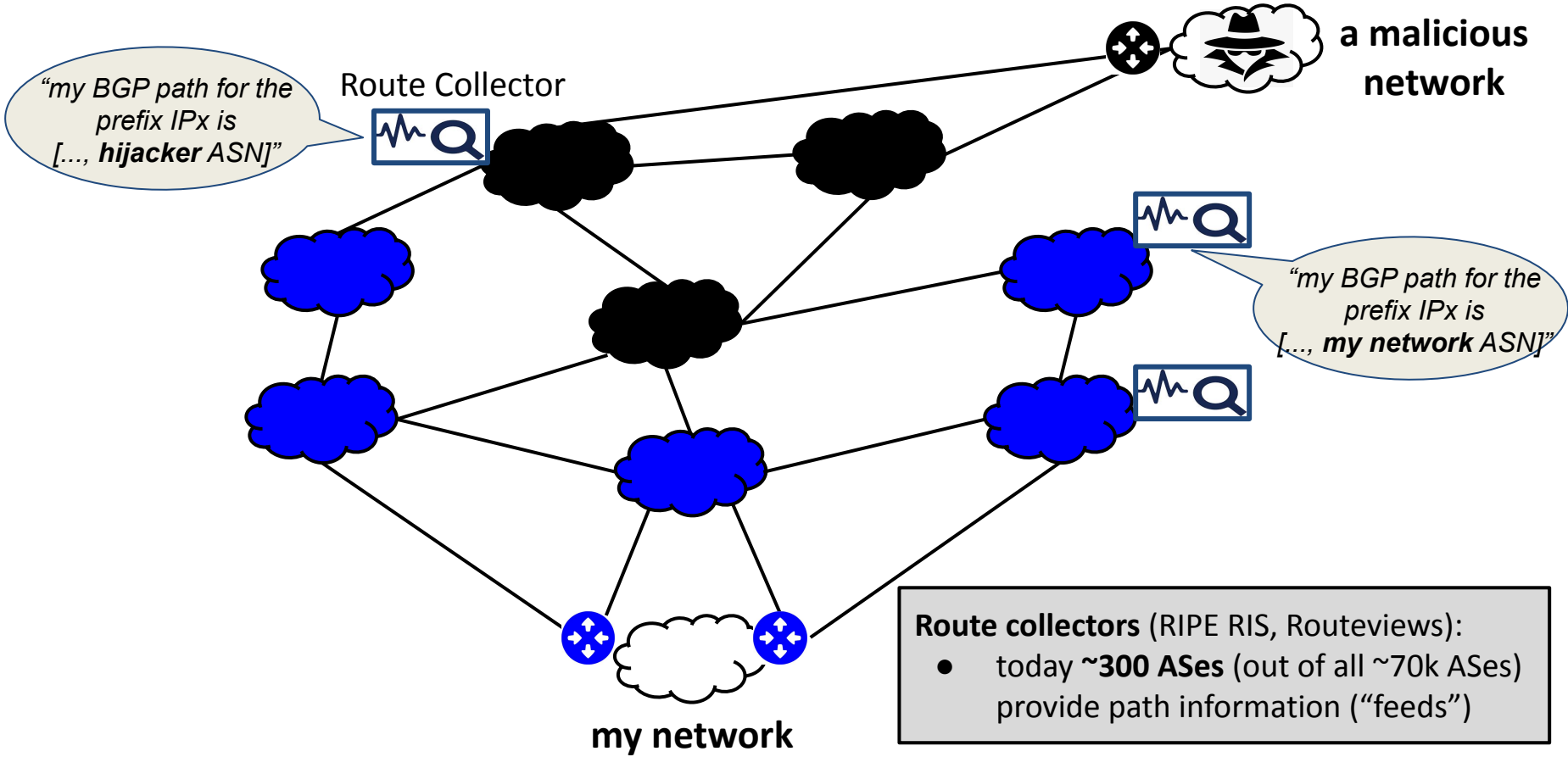
[1] ARTEMIS, open-source software, https://bgpartemis.org/

In this paper...

- **Goal → Estimate the impact of an ongoing hijack through measurements**
  - *sub-goal 1:* understand/characterise the estimation problem
    - types of measurements & public infrastructure
    - challenges & limitations
    - accuracy
  - *sub-goal 2:* design efficient estimation methodologies
    - with public infrastructure
    - without public infrastructure



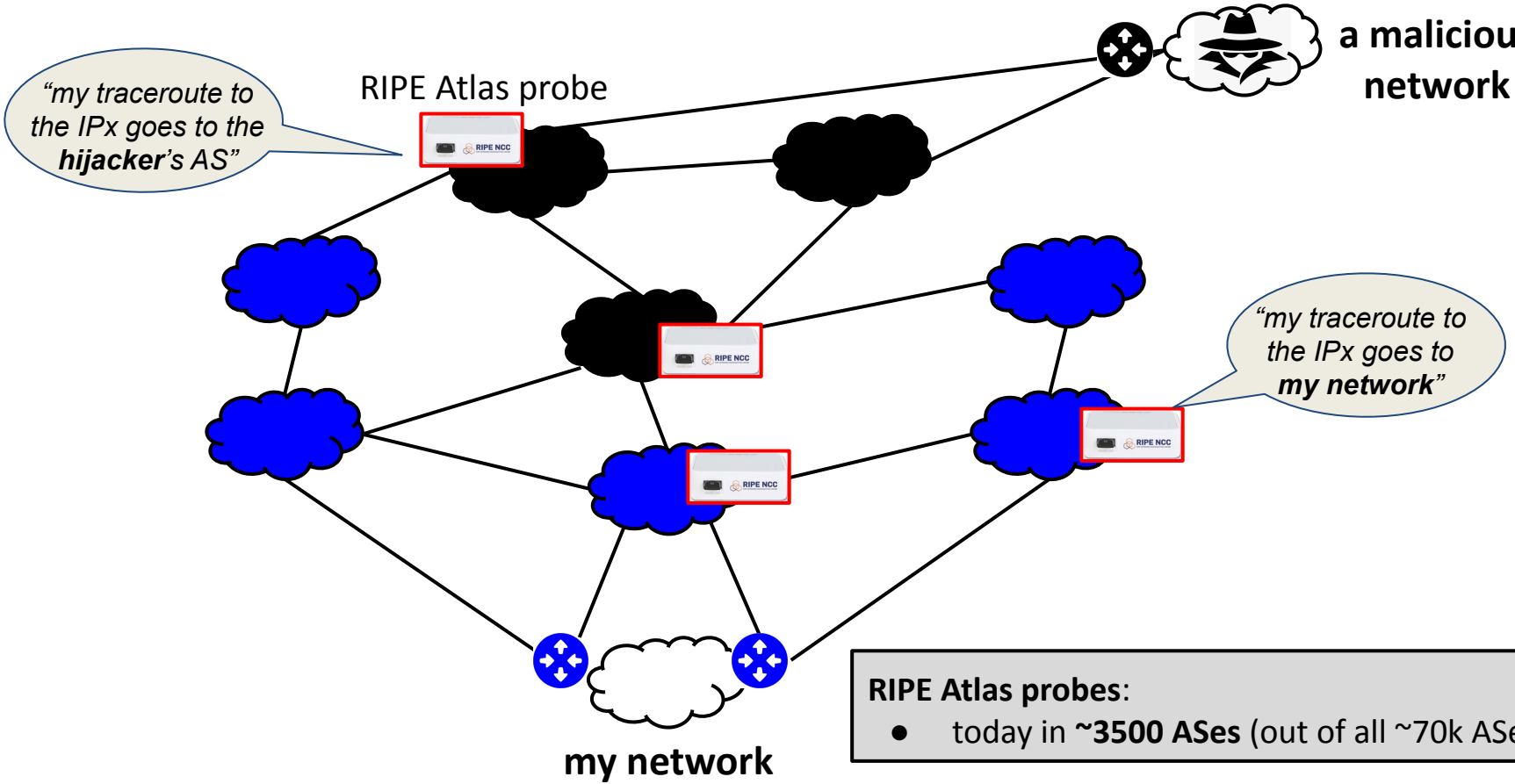
Measurements: BGP paths



Route collectors (RIPE RIS, Routeviews):

- today **~300 ASes** (out of all ~70k ASes) provide path information ("feeds")

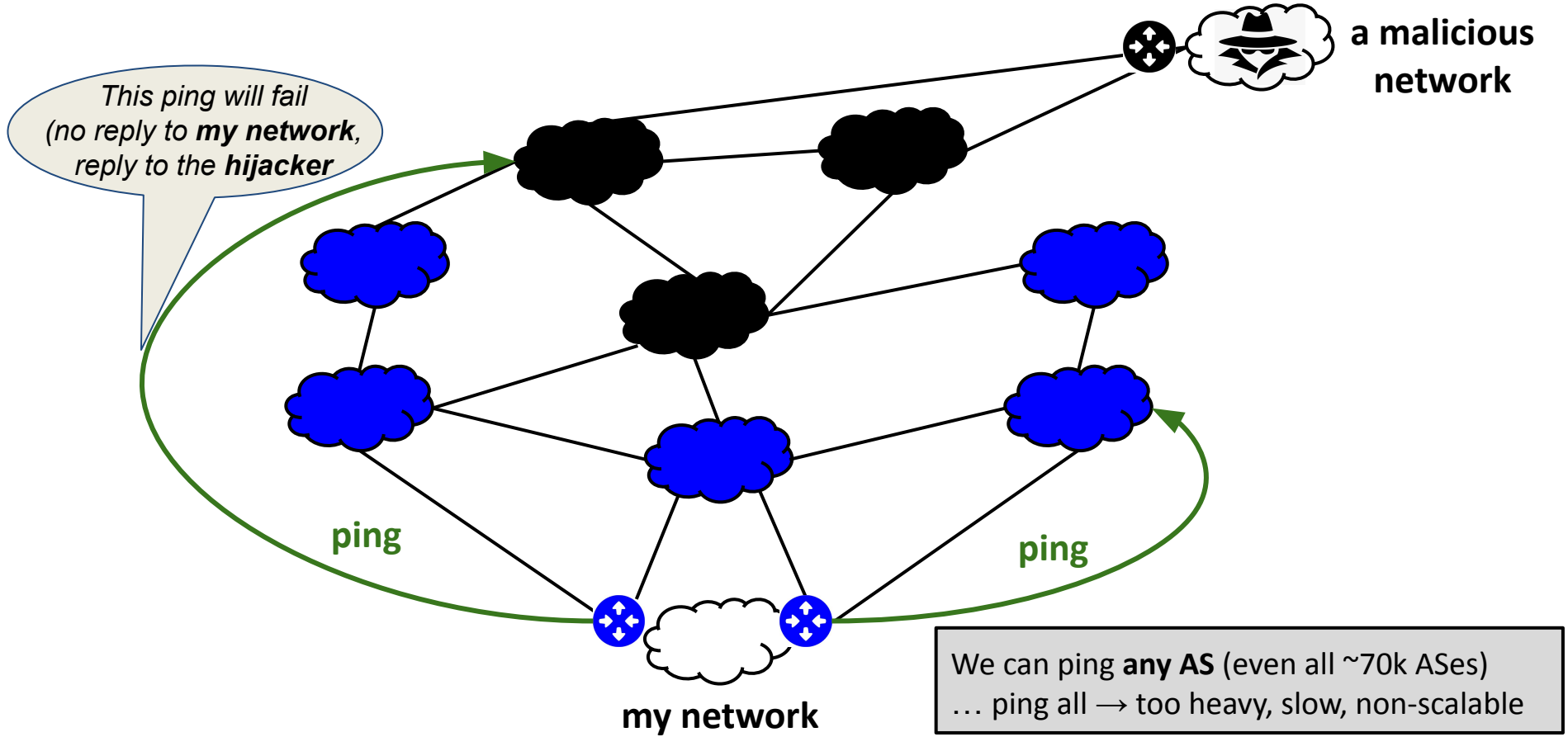
Measurements: traceroutes



RIPE Atlas probes:

- today in **~3500 ASes** (out of all ~70k ASes)

Measurements: pings



We can ping **any AS** (even all ~70k ASes)
… ping all → too heavy, slow, non-scalable

Hijack impact estimation with measurements

- Hijack impact == number of infected ASes
  - ***"infected AS"*** == an AS that routes its traffic to the hijacker AS

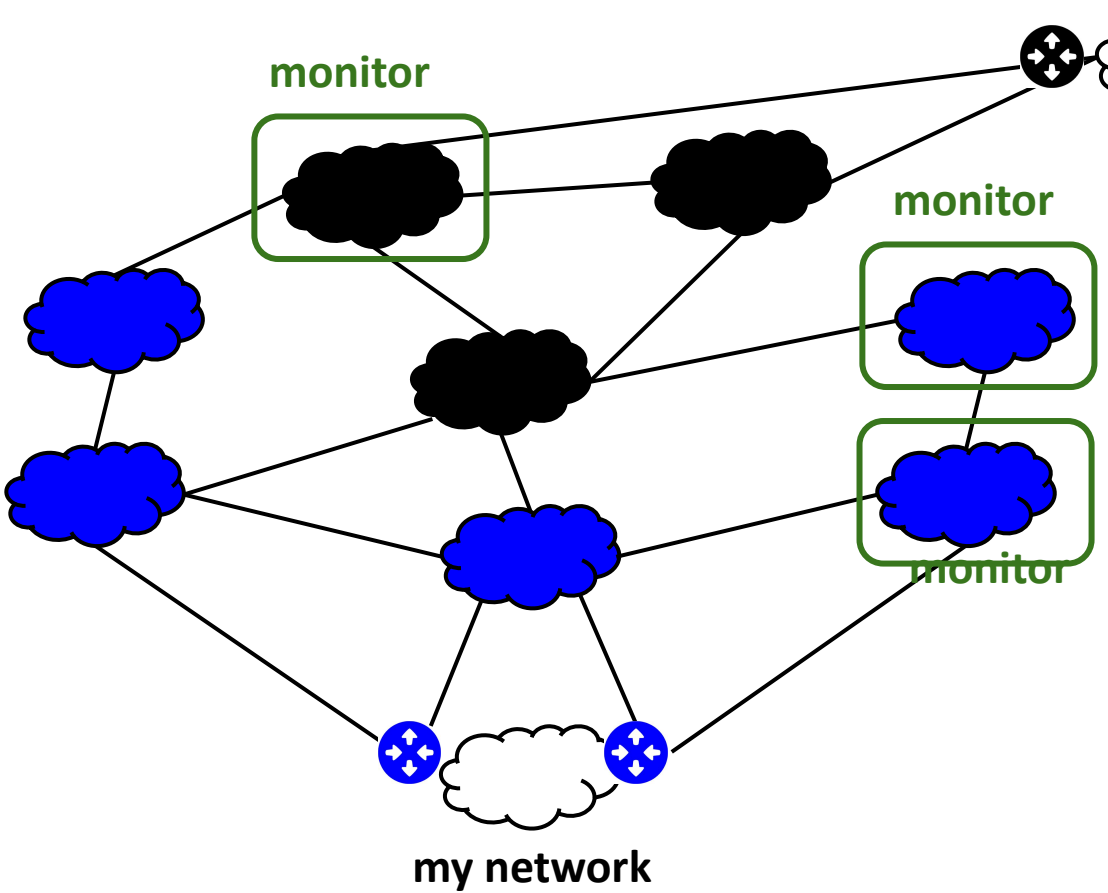
$$\textbf{actual impact} = \frac{\#\ \text{infected } \textbf{ASes}}{\#\ \text{total } \textbf{ASes}}$$

- Measurements for hijack impact estimation
  - measure some ASes
  - measured AS == ***"monitor"***
  - any measurement type: BGP path (route collector), traceroute (RIPE Atlas probe), ping

- Estimate hijack impact
  - from the number/percentage of "infected monitors"

$$\textbf{estimated impact} = \frac{\#\ \text{infected } \textbf{monitors}}{\#\ \text{total } \textbf{monitors}}$$

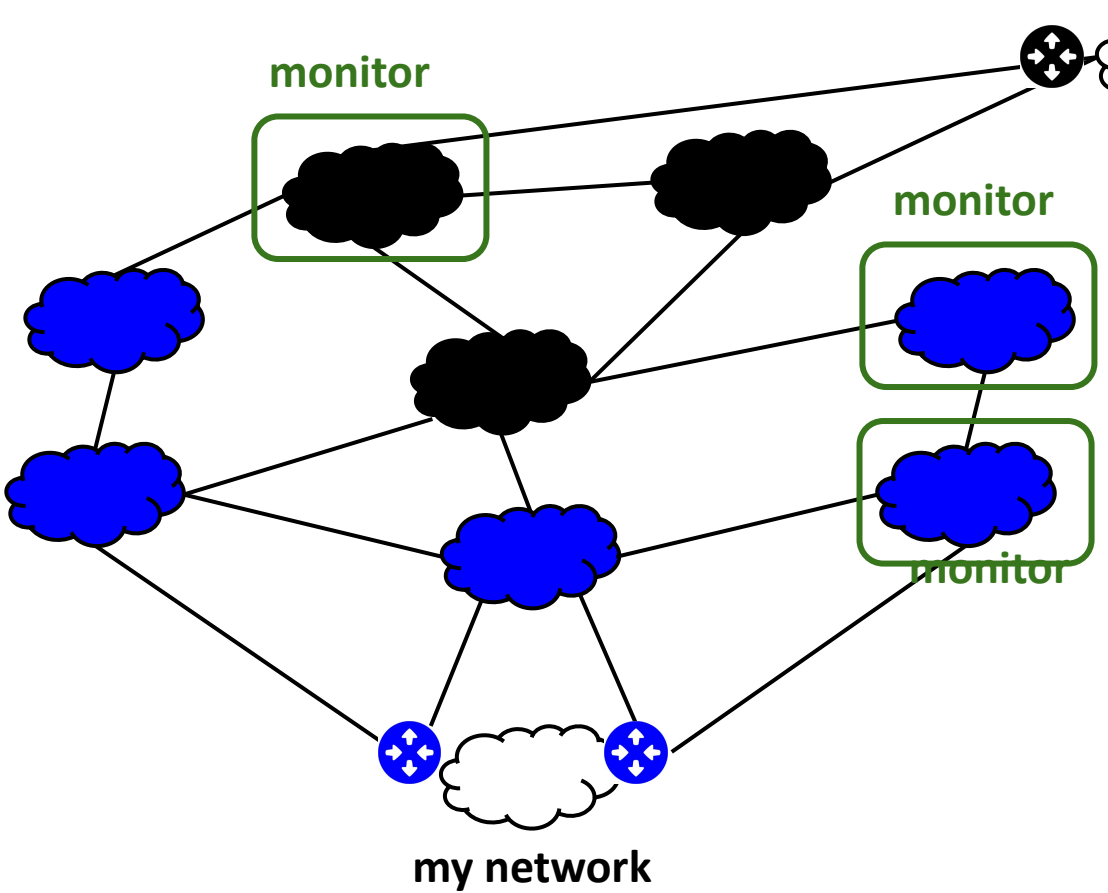
Impact estimation: an example



$$\text{actual impact} = \frac{3}{5 + 3} = 0.375$$

$$\text{estimated impact} = \frac{1}{2 + 1} = 0.333$$

Impact estimation: an example



$$\text{actual impact} = \frac{3}{5+3} = 0.375$$

$$\text{estimated impact} = \frac{1}{2+1} = 0.333$$

actual vs. estimated value
→ estimation error

Goal 1: study the accuracy of estimation

- **Sampling in theory…**
  - The estimation error (RMSE) decreases with the number of samples/monitors (M)

Theorem 1. *Under a randomly selected set of monitors $\mathcal{M}$, the bias and root mean square error of NIE are given by*

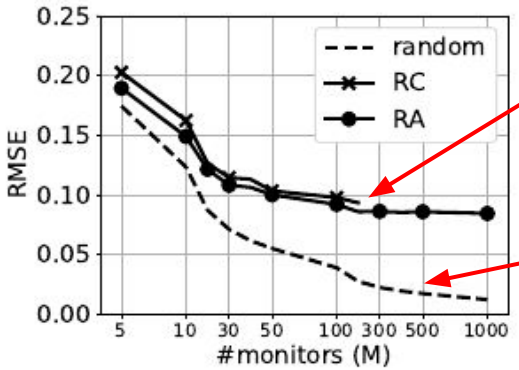
$$Bias_{NIE} = 0 \qquad RMSE_{NIE} = \frac{1}{\sqrt{M}} \cdot c_I$$

where $c_I = \int_0^1 \sqrt{I \cdot (1-I)} \cdot f(I) \cdot dI$, *is a constant that depends on the impact distribution* $f(I)$.

Goal 1: study the accuracy of estimation

- **Sampling in theory…**
  - The estimation error (RMSE) decreases with the number of samples/monitors (M)

- **Sampling in practice…**
  - with public infrastructure (Route Collectors, RIPE Atlas probes)



public infrastructure

theory (random sampling)

Key findings:

- The error of public infrastructure does not decrease less than 9%
  → due to location bias; public infrastructure is not deployed uniformly around the world

- 50 samples from public monitors are enough
  → insights for lightweight measurements

Goal 1: study the accuracy of estimation

- **Sampling in theory…**
  - The estimation error (RMSE) decreases with the number of samples/monitors (M)

- **Sampling in practice…**
  - with public infrastructure (Route Collectors, RIPE Atlas probes)

  - what about ping measurements?

Goal 1: study the accuracy of estimation

- **Sampling in theory…**
  - The estimation error (RMSE) decreases with the number of samples/monitors (M)

- **Sampling in practice…**
  - with public infrastructure (Route Collectors, RIPE Atlas probes)

  - what about ping measurements?
    - ✓ we can have random sampling!
    - ✗ but... high measurement failures (*> 90%* non pingable IP addresses)

Theorem 2.
RMSE vs. failure probability ***p.***

Key findings:
- ping measurements end-up being less accurate than public infrastructure (for ***p*** > 20%)
- we would need at least ***p*** < 10%

Goal 2: design accurate estimators

- Approach 1: based on **ping measurements**…
  - *Goal*: we need to decrease the failure probability ***p***
- Approach 2: based on **public-infrastructure**…
  - *Goal*: we need to remove the measurement bias

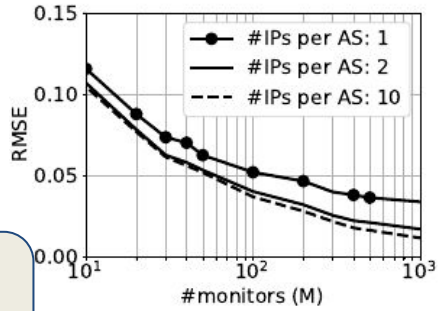
Ping-based impact estimator

- Approach 1: based on **ping measurements**…
  - *Goal*: we need to <u>decrease the failure probability ***p***</u>

<u>Ping-based impact estimator</u>

1. Find "pingable" IP addresses for every AS [*ANT Lab's IP hitlist*]
2. Ping multiple (N_{IP}) IP addresses per AS
3. If at least one ping reply from an AS → the AS is not affected by the hijack

N_{IP} *(nb of pinged IPs per AS)*	1	2	3	…	10
p *(failure probability per AS)*	12.8%	4.2%	2.1%	…	0%
RMSE *(estimation error; M=100)*	7.9%	4.7%	4.1%	…	3.9%



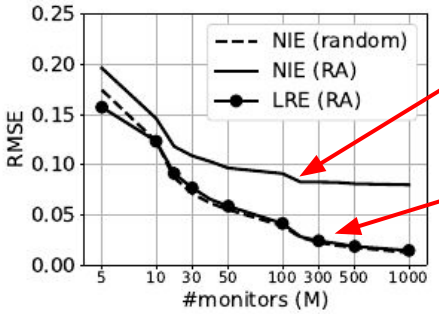
Key findings:
- N_{IP} **>= 2** for low error
- no need for N_{IP}**>3**

Public infrastructure based estimator

- Approach 2: based on **public-infrastructure**…
  - *Goal*: we need to <u>remove the measurement bias</u>

<u>Linear-regression estimator (LRE)</u>

1. Collect past measurements of (public infrastructure) monitors
2. Fit a least-square estimator → give to each monitor ***i*** a weight w_i
3. Collect measurements m_i for the ongoing hijack
4. Estimate the impact as: $\Sigma_i m_i * w_i$



public infrastructure

public infrastructure with LRE

Key findings:

- LRE eliminates the bias in public infrastructure measurements & achieves close-to-theory efficiency
- Only a few past measurements are needed for fitting the LRE (e.g., it worked quite well even with 20 past events in our experiments)

Summarizing...

Estimating the impact of BGP prefix hijacking

- Important for network operations (e.g., mitigation actions)
- Not studied before
- We studied fundamental (limits, trade-offs, etc.) and practical aspects (use of public infrastructure, measurement failures, etc.)
  - theory (insights) & simulations (generality) & experiments (realism/verification)

Future research directions

- ML-based estimators (but… lack of labelled datasets)
- Generality of results - beyond BGP prefix hijacking
  - de-bias public infrastructure measurements
  - identify key locations for expanding public infrastructure