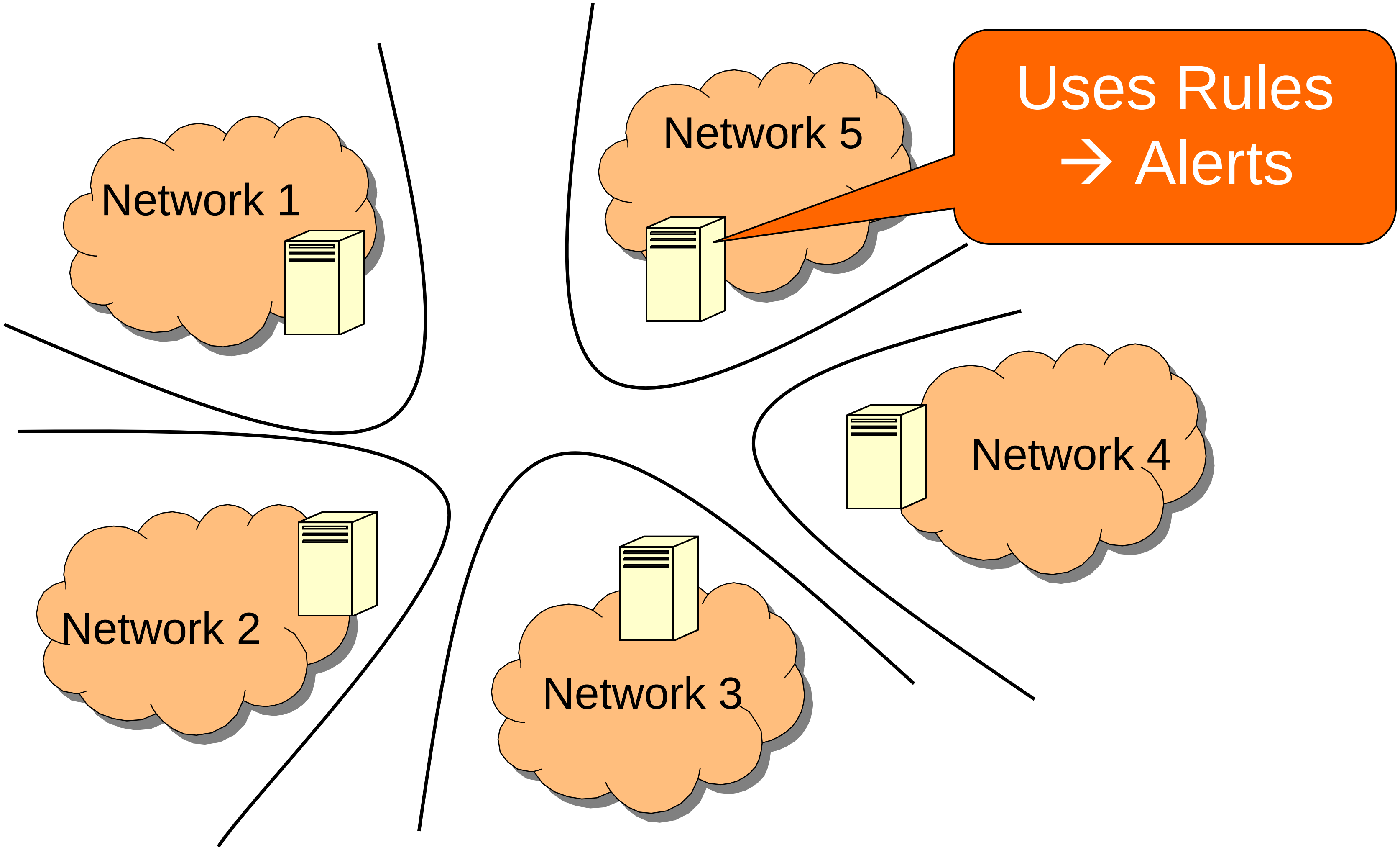
Collaborating Against Common Enemies

Sachin Katti

Balachander Krishnamurthy and Dina Katabi
AT&T Labs-Research & MIT CSAIL

Current Intrusion Detection



How about collaborating?

Potential reasons for collaboration:

- Provides global picture of attack
- Detecting low rate distributed attackers
- Detecting stepping stones

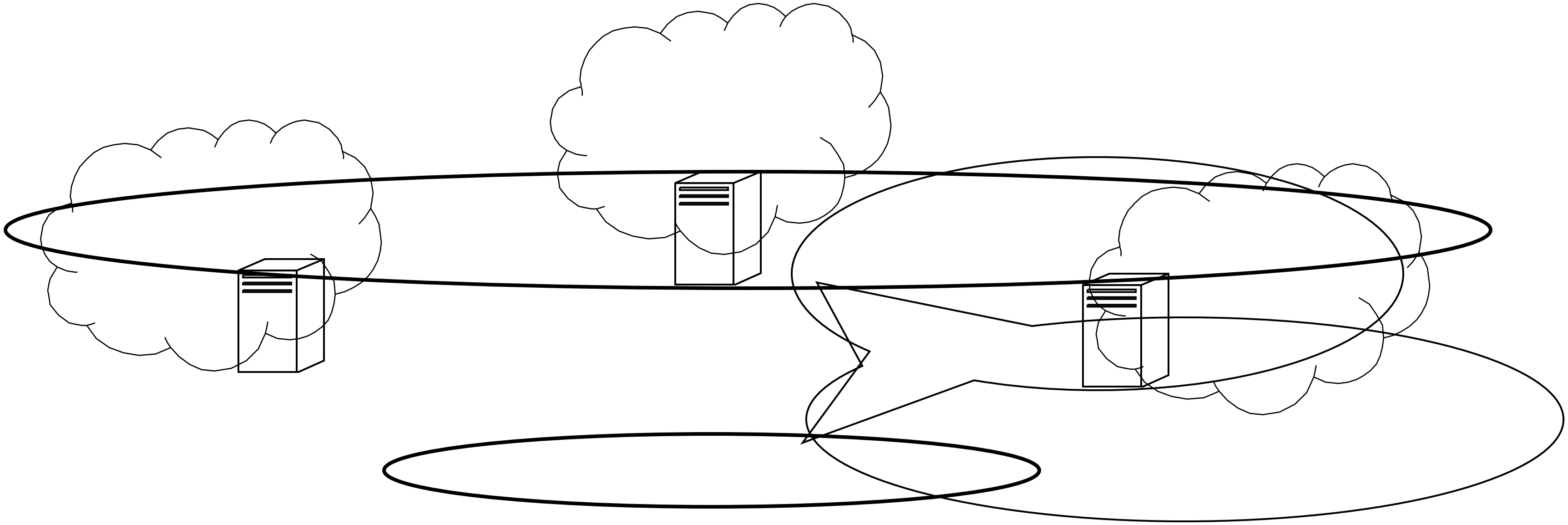
But benefit depends on networks/IDSs seeing Correlated Attacks?

Talk Is About Correlated Attacks

Define Correlated Attacks: as attacks from the same sources IP on different IDSs/networks

Talk Is About Correlated Attacks

Define Correlated Attacks: as attacks from the same source IP on different IDSs/networks



This Talk

Logs from 1700 IDSs show:

- 40% of alerts are correlated
- Correlated attacks within 10min
- An IDS sees correlated attacks with 8 IDSs (out of 1700), and the group does not change

→ Collaboration is useful

→ Realtime

→ Collaborate with a few IDSs

Collaboration with correlated IDSs increases detection by 75% and as good as collaborating with all.

Dataset

Full packet headers, unanonymized src/dest addresses

Anonymized dest IP; no packet headers or alert type

Method

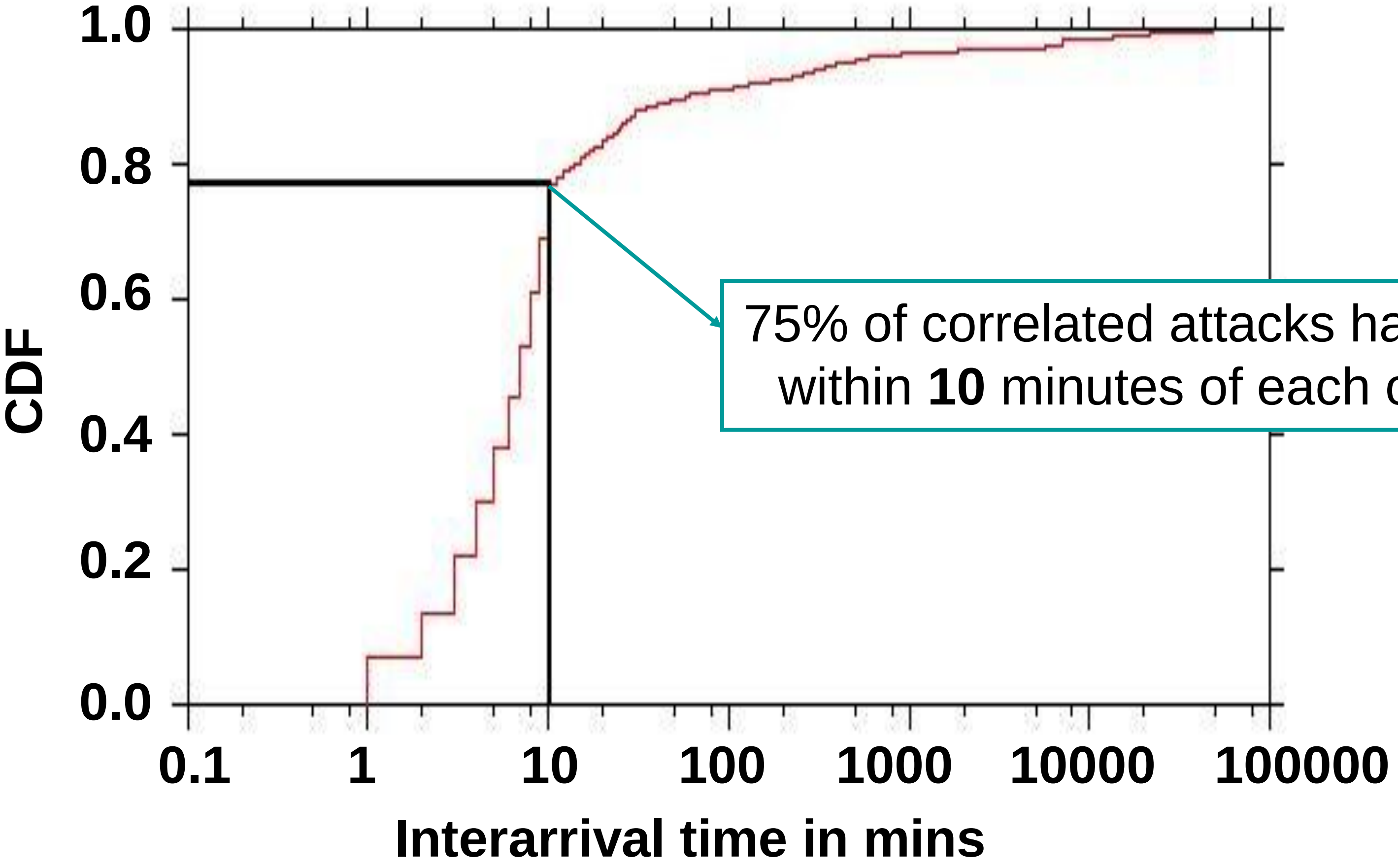
- Correlation is based on sharing the same source IP
  - Adding info about attack type and dest port did not matter
- Correlated IDSs – IDSs for which more than 10% of their attacks are correlated

Do IDSs see Correlated Attacks?

YES, Many

- 20% of attacking IPs are common attackers
- 40% of the attacks are correlated
- On average, 1500 correlated attackers/ day/IDS

Interarrival of Correlated Attacks



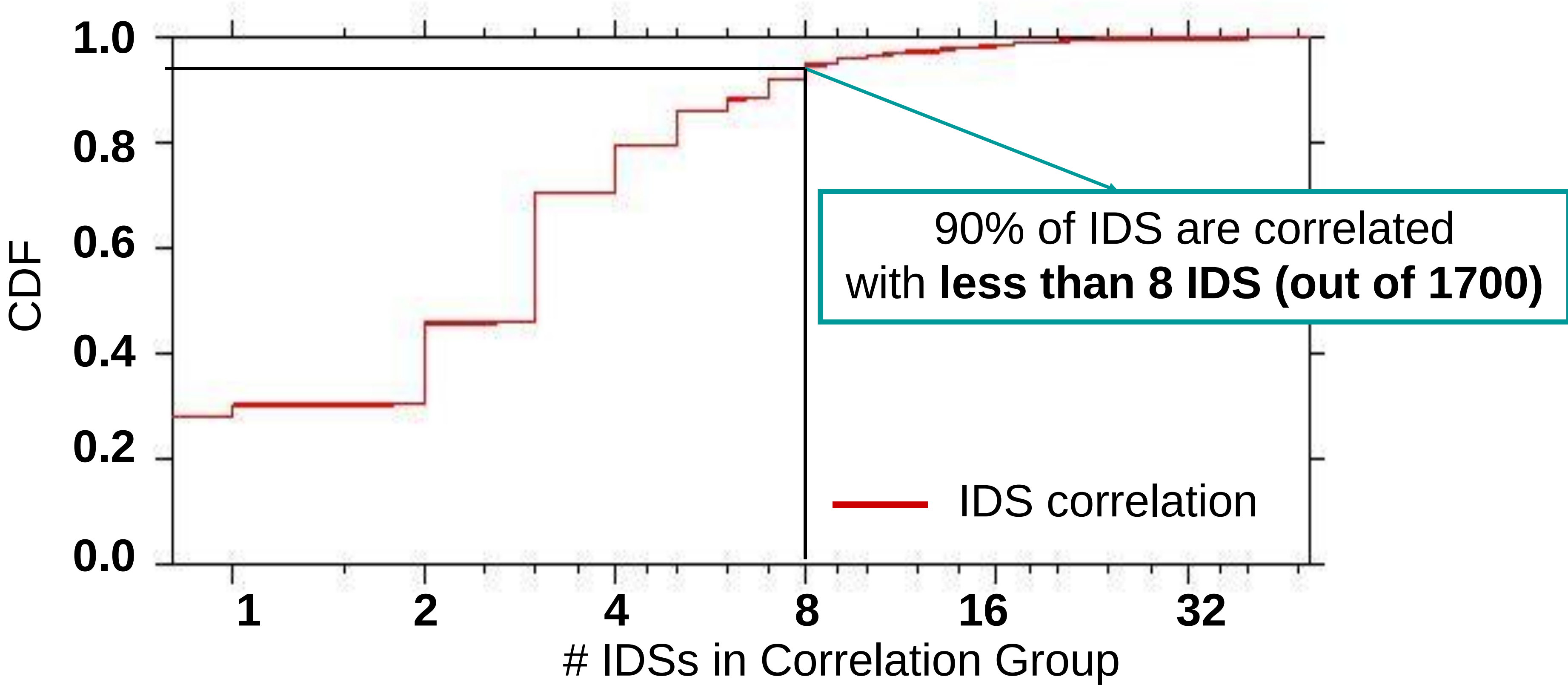
75% of correlated attacks happen within **10** minutes of each other

Correlated attacks within a few minutes

→ Need realtime collaboration!

Size of Correlation Groups

For each IDS compute the # of IDSs with which it is correlated

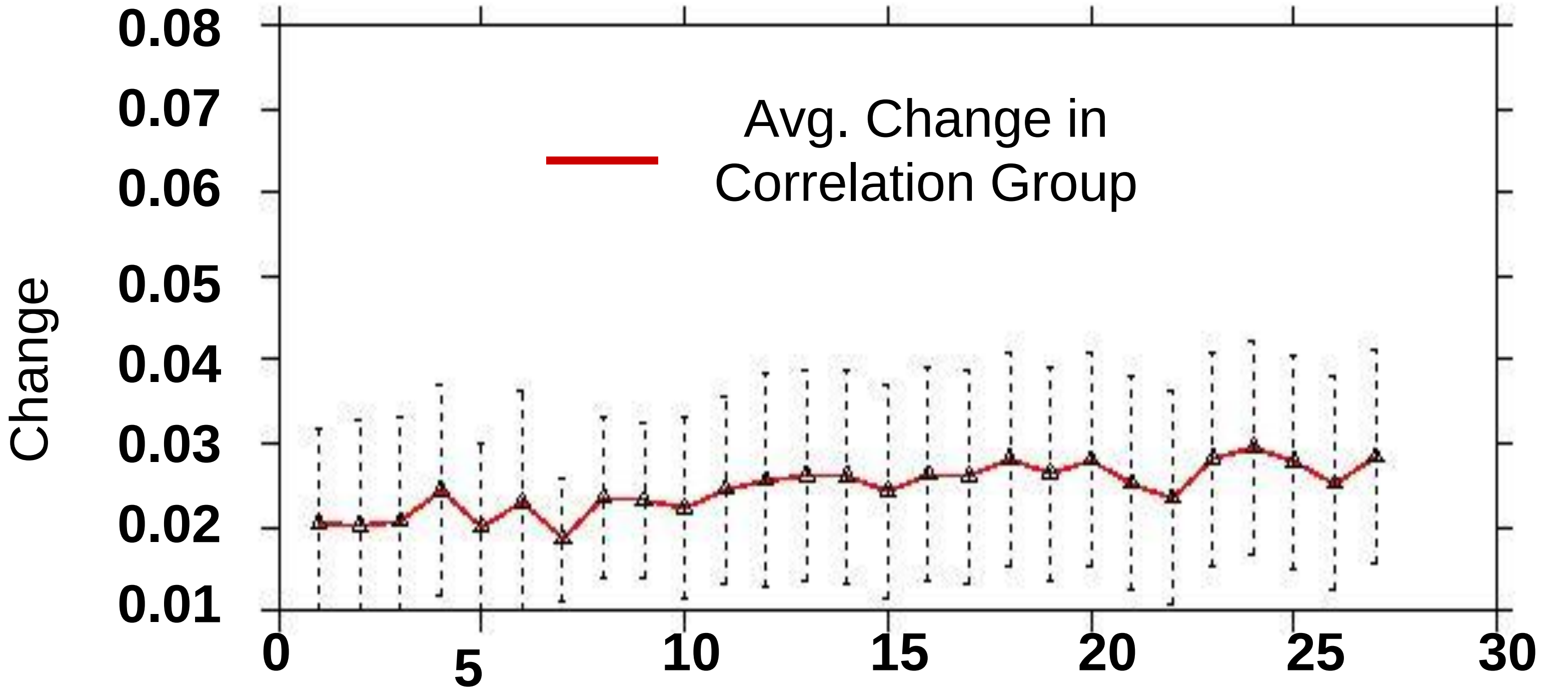


90% of IDS are correlated with **less than 8 IDS (out of 1700)**

IDS correlate within small groups!
→ Scalable collaboration

Do Correlation Groups Change?

If an IDS is correlated with 4 other IDS and the group changes by one, the percentage change is 25%



Correlation is persistent!

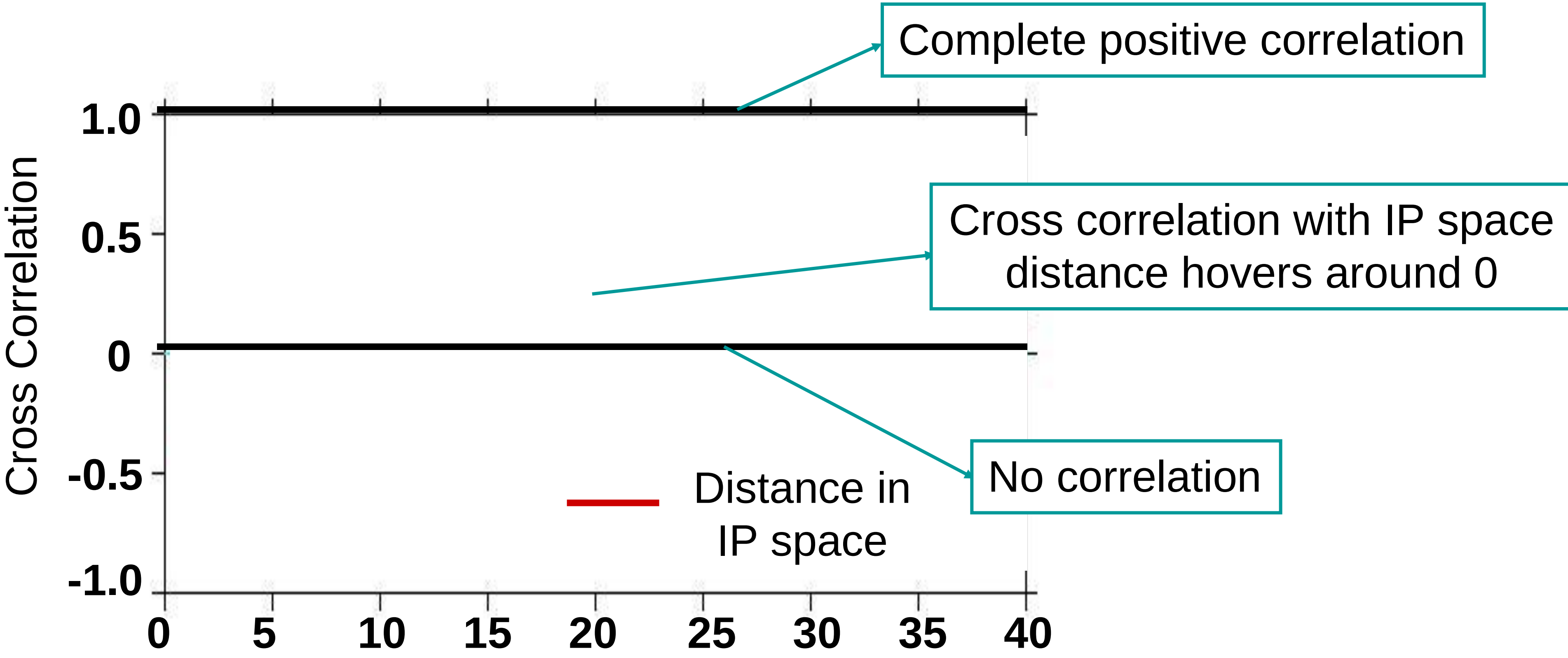
→ Establish trust out of band

Why IDS correlate?

- Is it proximity in IP space?

Is Proximity in IP Space the Reason?

- Compute cross correlation between proximity in IP space and correlated IDS



Complete positive correlation

Cross correlation with IP space distance hovers around 0

No correlation

Attack Correlation is independent of proximity in IP space

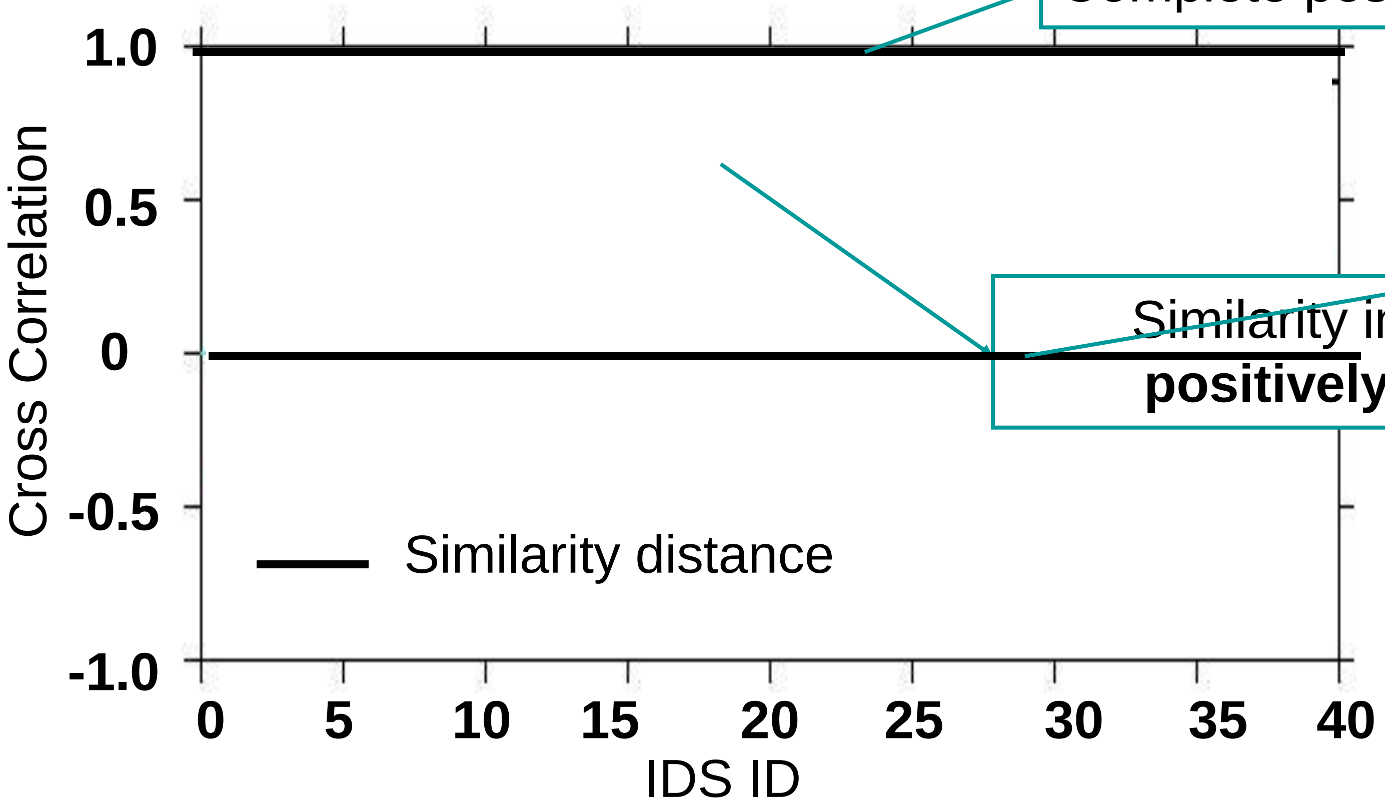
Why IDS correlate?

- ~~Is it proximity in IP space?~~
- Is it because attackers target sites with similar software and services (e.g., Santy worm) ?

More than 60% of attacks in a correlation group target particular service (e.g. SMTP groups, IBM Tivoli, IIS servers)

Is Similarity in Software the Reason?

- Compute cross correlation between similarity in software & attack correlation



Complete positive correlation

Similarity in software is **positively correlated**

No correlation

Decreasing similarity ⇨ **Decreasing correlation**

So, what does it mean for Collaborative Intrusion Detection?

Issues for IDS collaboration across networks

- Is it useful?
- How often should IDS exchange information?
- How to make it scale?
- How does an IDS trust its collaborators to protect the privacy of its information and not lie?

Exploiting Correlation for collaboration

- 40% of alerts are correlated
- Correlated attacks within 10min
- An IDS sees correlated attacks with small correlation groups (8 out of 1700 IDS)
- The correlation group does not change

→ Collaboration is useful

→ Realtime

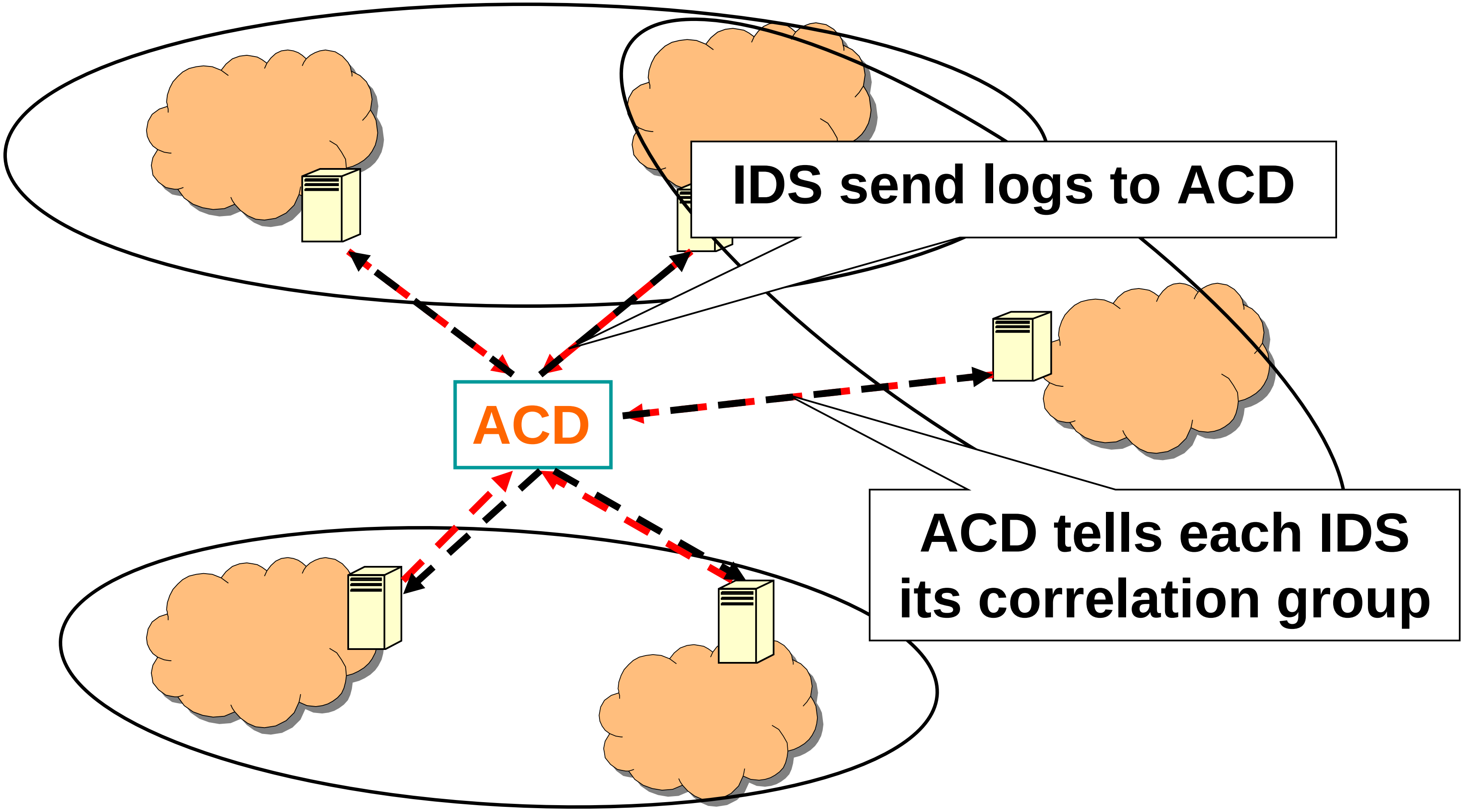
→ Scale by collaborating with IDS in same correlation group

→ Check trust out-of band

Correlation Based Collaboration (CBC)

- **Attack Correlation Detector (ACD)** for finding correlation groups (e.g., DShield)
- Since groups persist for months → ACD computation scale
- It is up to each network to decide whether to collaborate or not

Correlation Based Collaboration (CBC)



IDS send logs to ACD

ACD tells each IDS its correlation group

Evaluation of CBC Blacklisting

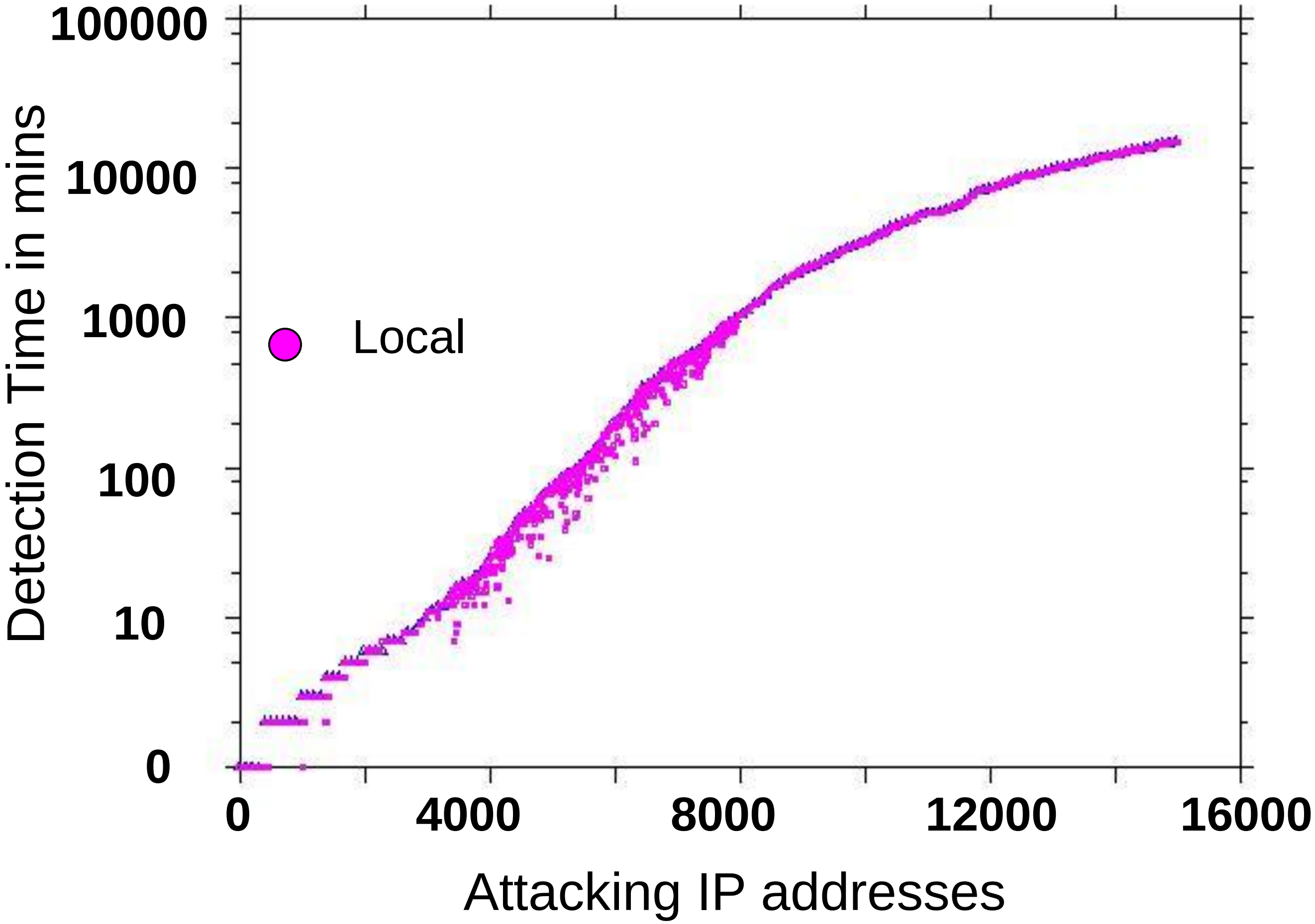
- Flag an attacking IP address if # alerts cross a threshold
- Compare with
  - Local detection
  - Collaborating with all IDSs
  - Random Collaboration - Collaborating with the same sized random subset as the correlation group

Evaluation Method

- IDS queries its collaborators when # alerts from an IP exceeds `Querying Threshold`
- IDS blacklists IP if aggregate # alerts exceeds `Blacklisting Threshold`
- Thresholds picked to minimize false positives (for ISP dataset)

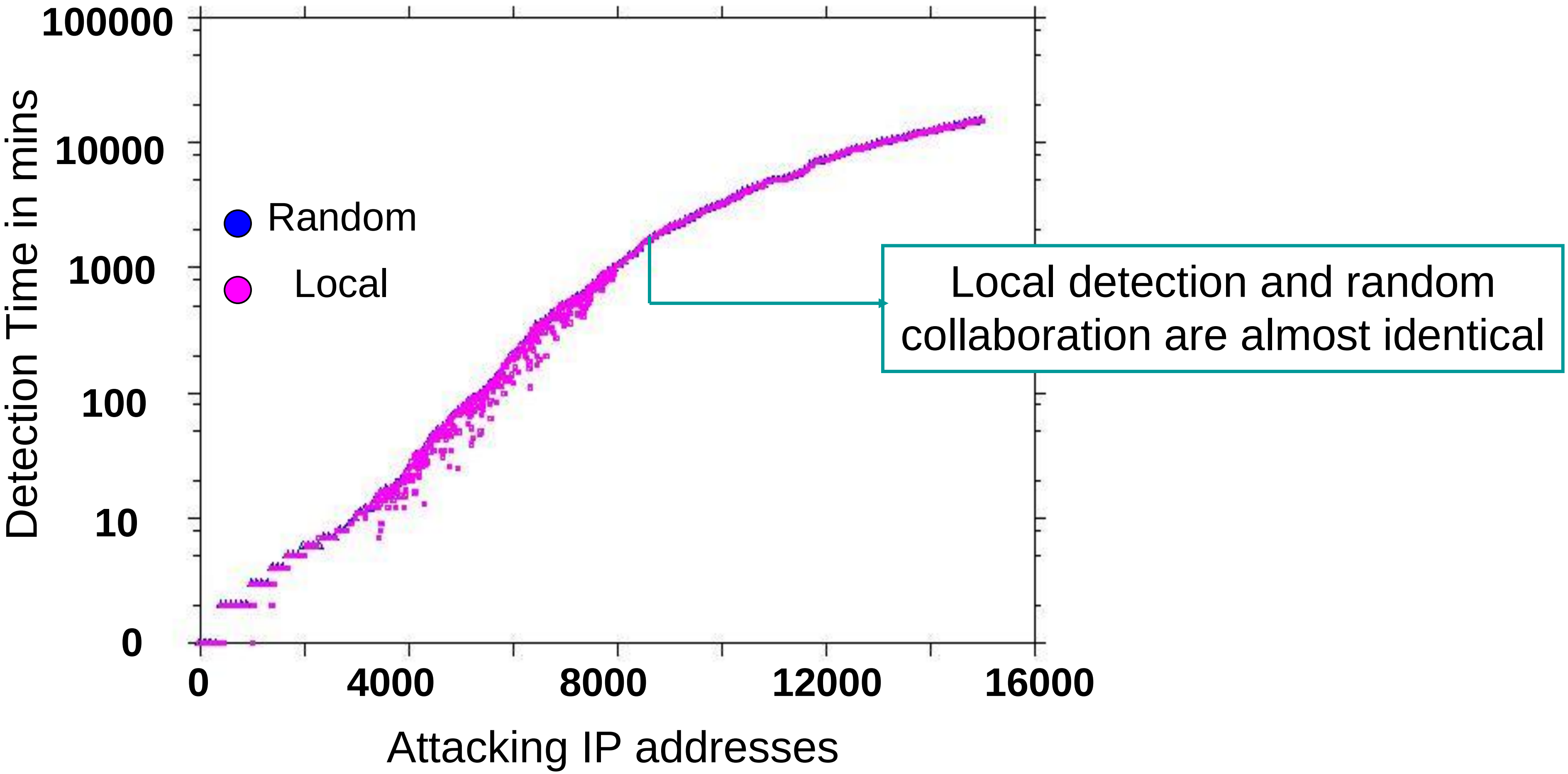
Speed!

- Compute time taken to blacklist a source in each scheme



Speed!

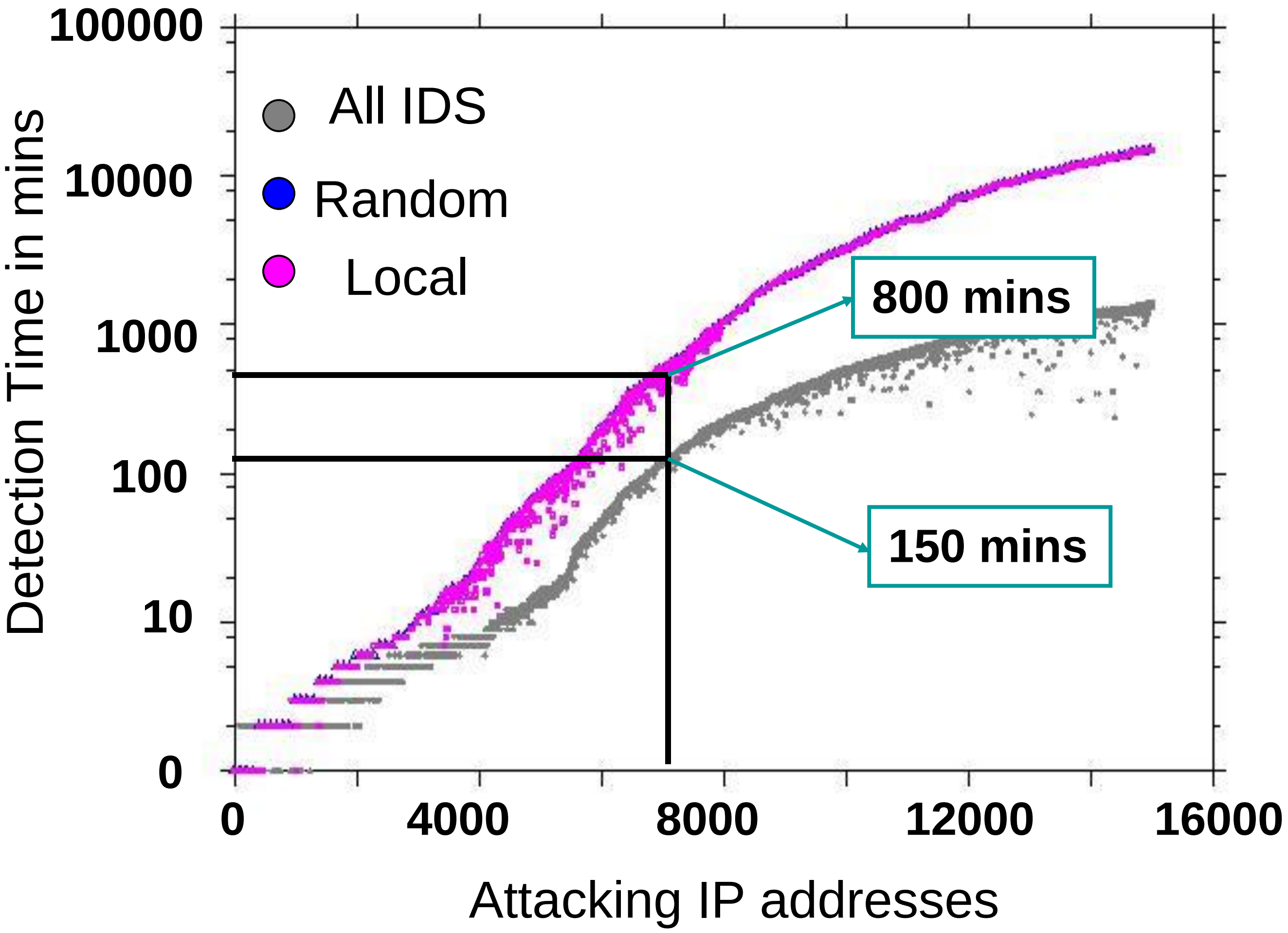
- Compute time taken to blacklist a source in each scheme



Local detection and random collaboration are almost identical

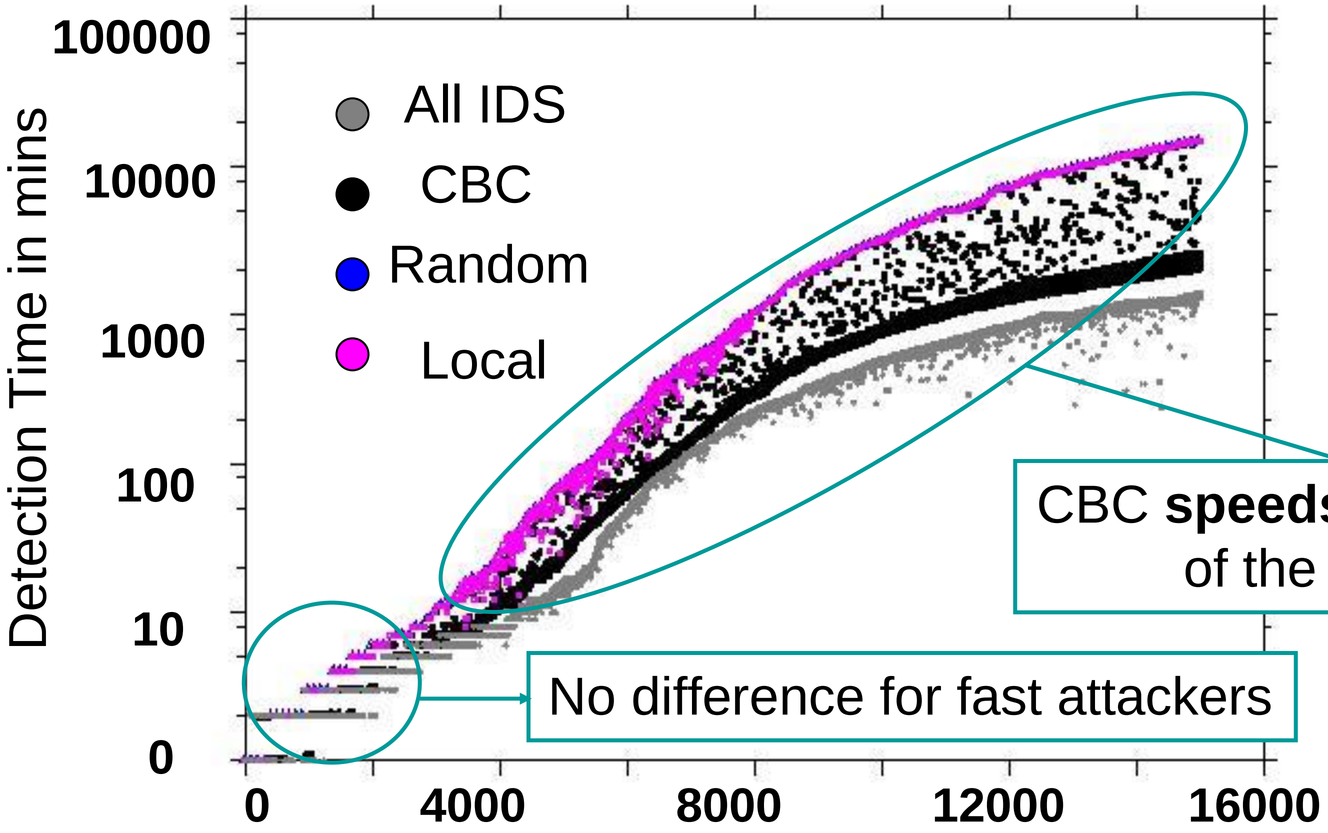
Speed!

- Compute time taken to blacklist a source in each scheme



Speed!

- Compute time taken to blacklist a source in each scheme



CBC **speeds** up detection for **75%** of the studied sources

No difference for fast attackers

CBC performs almost as well as collaborating with all IDS

Significant Reduction in Alert Volume

	CBC	Local Detection	Random	All IDSs
Alert Reduction	73.44%	35.48%	37.77%	80.56%

CBC halves the volume of the alert logs a network administrator has to examine!

Low Overhead

	CBC	Local Detection	Random	All IDSs
Alert Reduction	73.44%	35.48%	37.77%	80.56%
Overhead (query/min)	1.3	-	1.3	454.9

CBC requires orders of magnitude less querying overhead for the same benefits!

Conclusions

- 40% of alerts are correlated
- Correlated attacks within 10min
- An IDS sees correlated attacks with small correlation groups (8 out of 1700 IDS)
- The correlation group does not change

→ Collaboration is useful

→ Realtime

→ Scale by collaborating with IDS in same correlation group

→ Check trust out-of band

CBC exploits the above; is as good as collaborating with all but with 0.3% of the overhead.