Optimal deployment of security virtual functions in

Software-Defined Networks (SDN)

Sonia Vanier*, Céline Gicquel**,

° KahinaLazri, Alexandros Papadimitriou

* Université de Paris I Panthéon-Sorbonne, ** Université Paris Saclay

° Orange Labs Products & Services, Security Department

CTW 2020

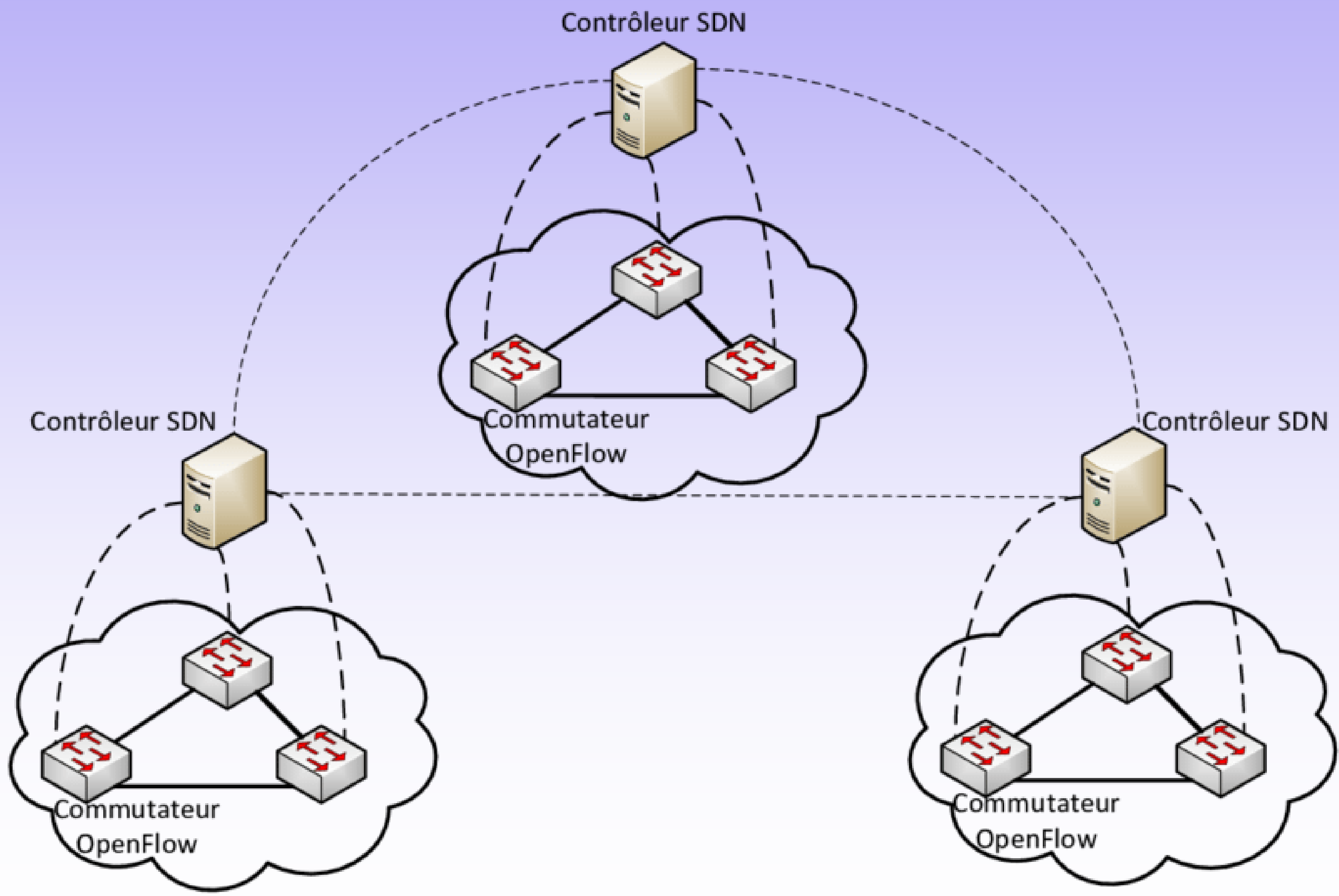
UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE

Summary

- **Problem description:**
  - Software Defined Networking (SDN)
  - Virtual Network Functions for security services
  - What is a Distributed Denial of Service attack?
- **Mathematical Formulation**
- **Solution Approach**
- **Numerical Results**
- **Conclusion**

Software Defined Networking (SDN)

- **SDN** attempts to **centralize network management** in one network component by disassociating the forwarding process of network packets from the routing process.



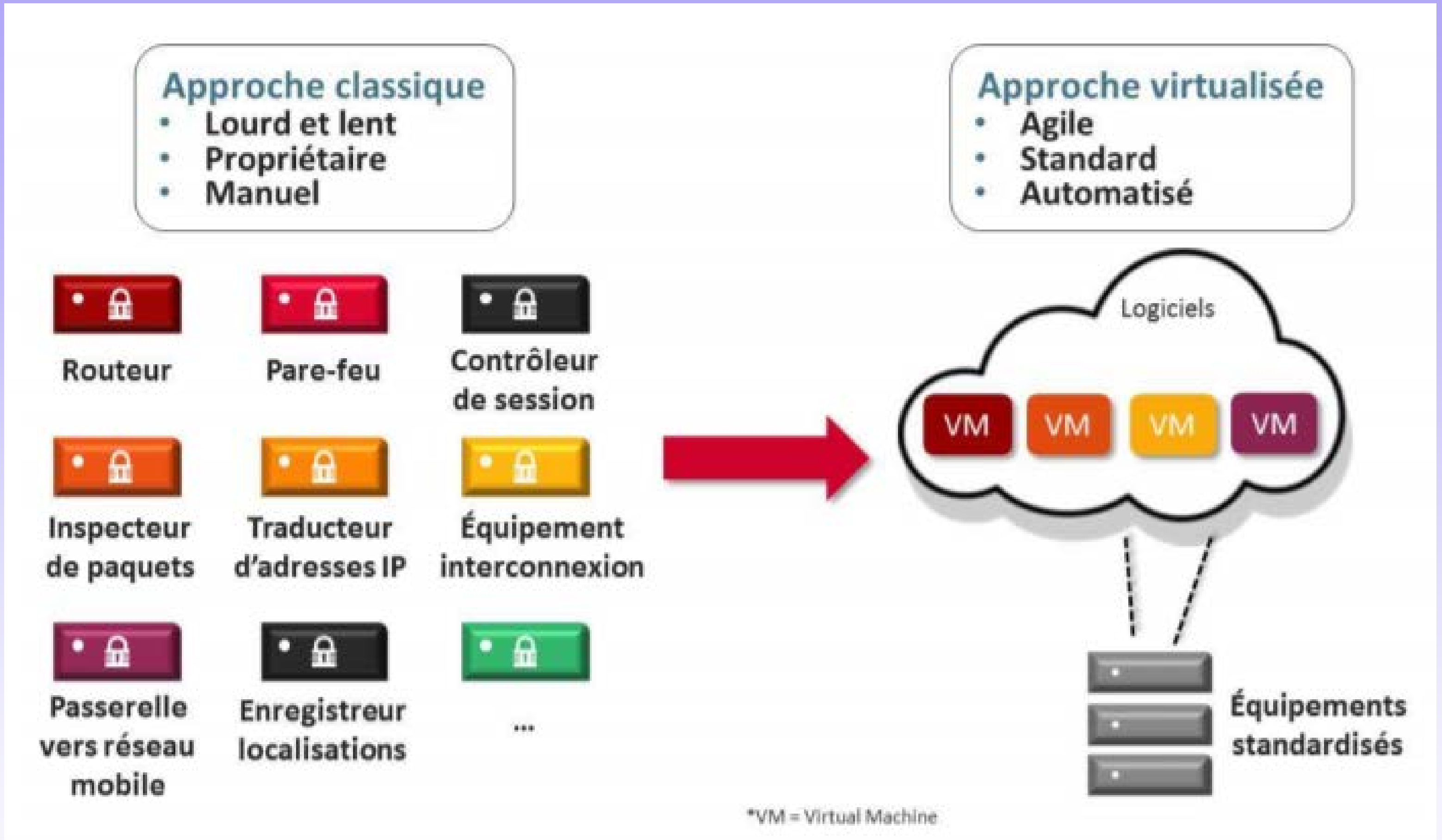
Software Defined Networking (SDN)

- SDN architecture is an approach to cloud computing that facilitates network management and enables programmatically efficient network configuration in order to **improve network performance and monitoring.**

- SDN is an architecture that provides support for virtual machine mobility independent of the physical network.

- **Network Function Virtualization (NFV)**: may consist of one or more virtual machines running different software and processes, on top of standard high-volume servers, switches and storage devices, or even cloud computing infrastructure, instead of having custom hardware appliances for each network function.

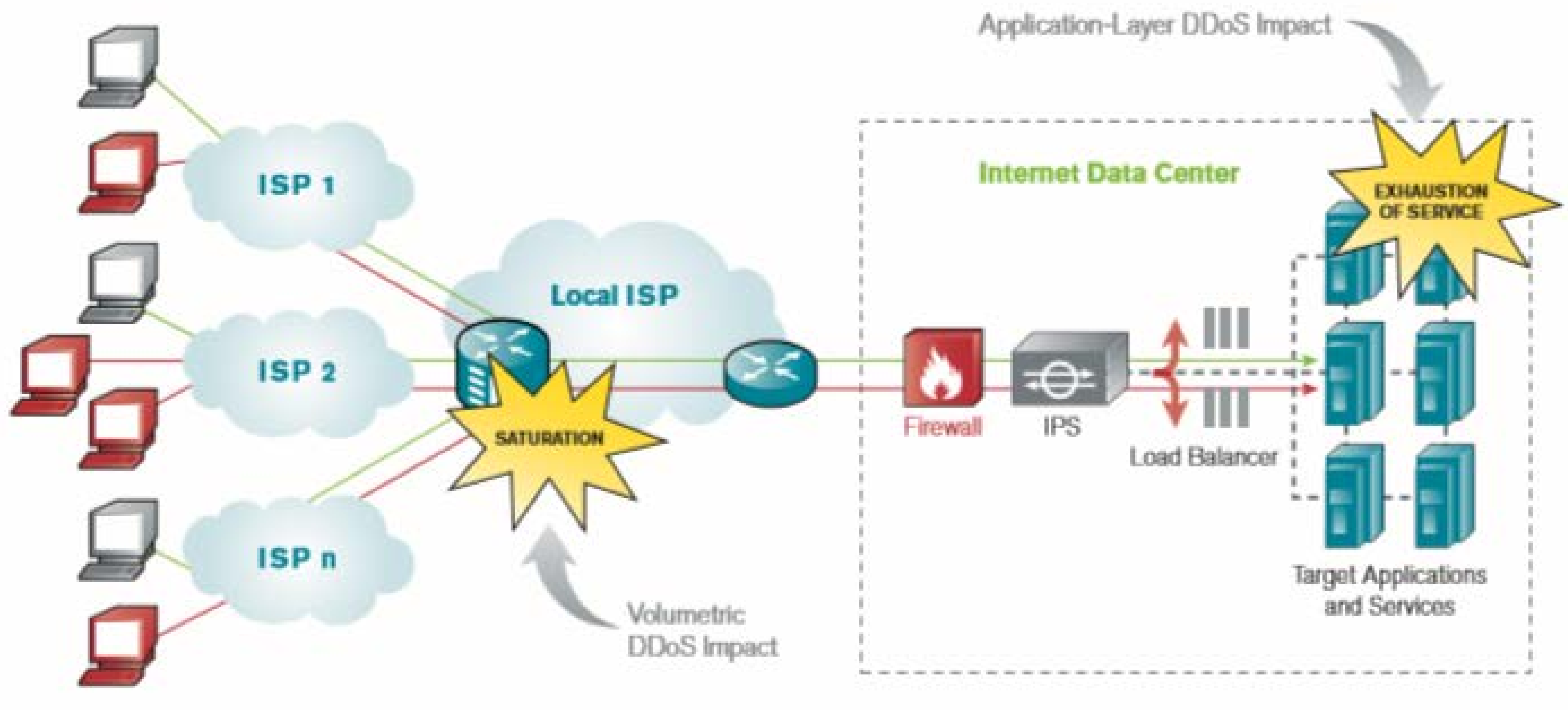
Network Functions Virtualization (NFV)



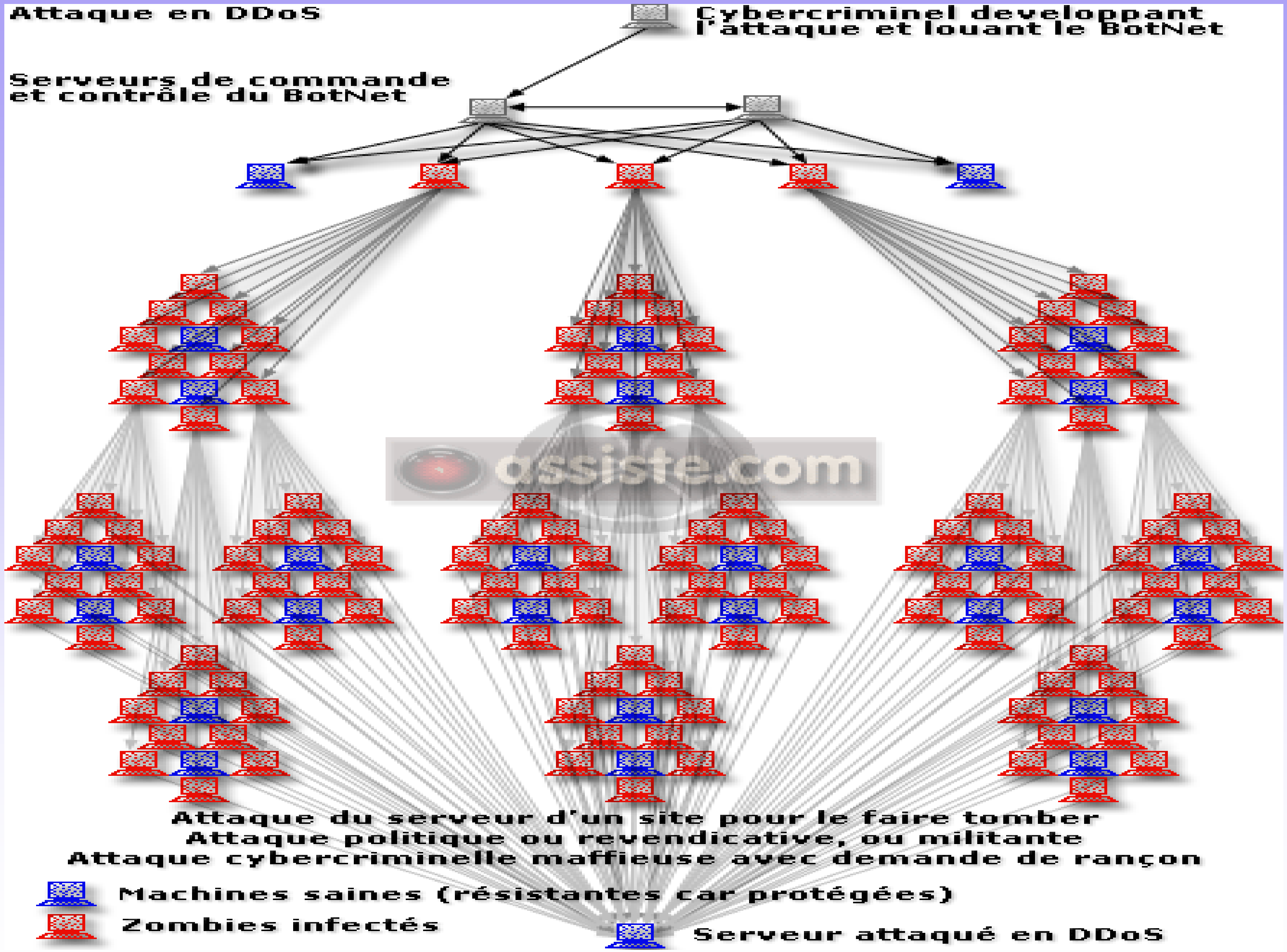
Virtualisation de fonctions réseaux

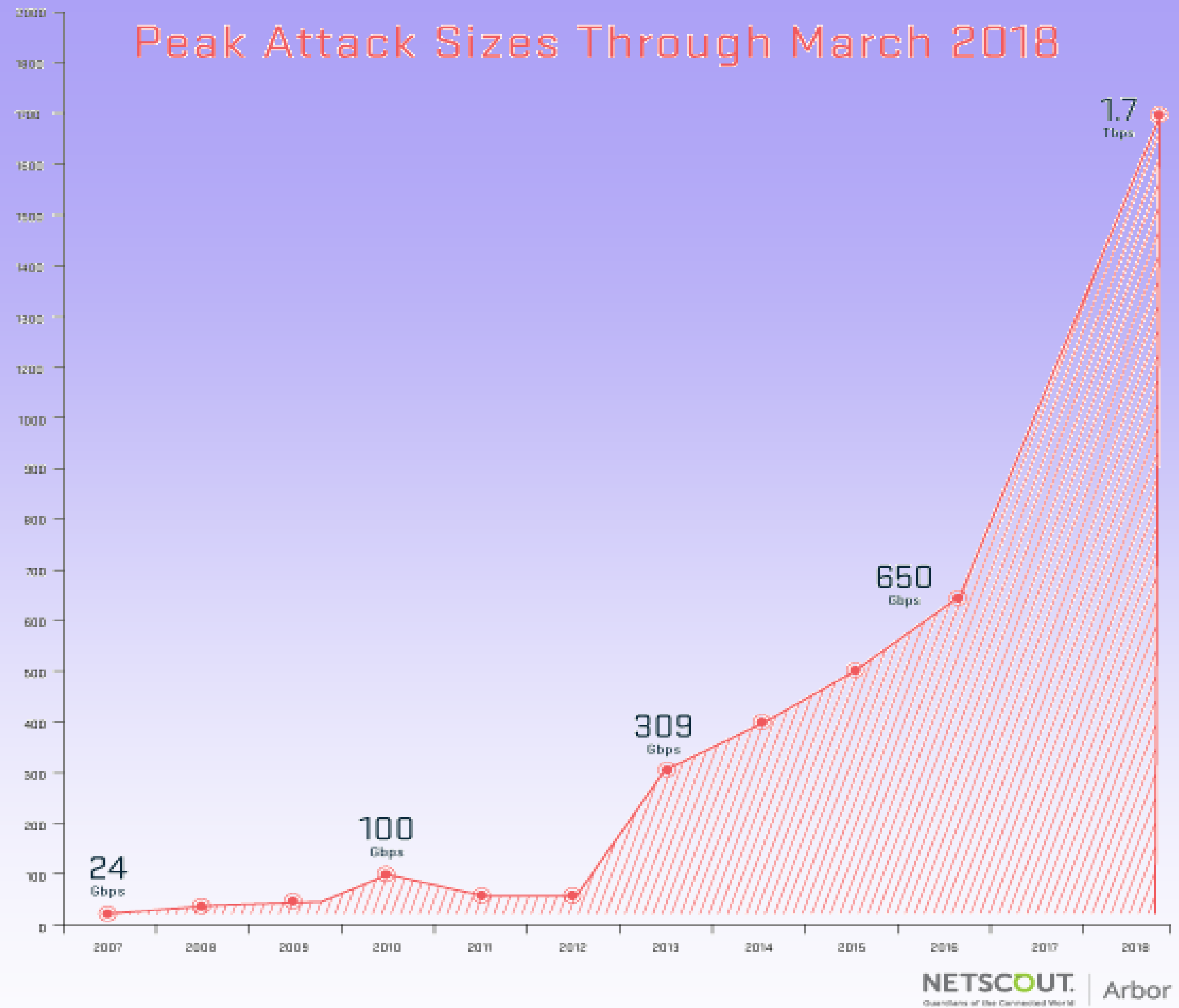
What is a DDoS Attack?

What is a DDoS attack?



Volumetric and application-layer DDoS attacks


Attaque en DDoS
Cybercriminel developpant l'attaque et louant le BotNet
Serveurs de commande et contrôle du BotNet
assiste.com
Attaque du serveur d'un site pour le faire tomber
Attaque politique ou revendicative, ou militante
Attaque cybercriminelle maffieuse avec demande de rançon
Machines saines (résistantes car protégées)
Zombies infectés
Serveur attaqué en DDoS



The record saved by Arbor Networks

NFV placement for Defense against DDoS

- **Problem Formulation:**
  - **Given the topology of the network, with transmission capacities on links and computing resources at nodes**
  - **A set of source-target DDoS attacks**
  - **We are interested in the problem of optimal placing filtering NFVs on network nodes**
  - **NFVs filter the malicious traffic by differentiating legitimate packets from illegitimate packets.**
  - **Placement of NFVs must progressively and completely filter attacking traffic while minimizing the total cost of deployed NFVs.**

Mathematical models

- **Routing of DDoS attacks is not known**
  - With the advent of 5G/6G networks and ISPs (Internet Service Providers), operators are preparing to lease parts of their physical networks to service providers. Service providers will apply their own routing algorithms to route traffic on their leased network.
- **First part of the project:**
  - We developed single level MILP model allowing high level of security but can induce high investments costs, resolved with constraints generation approach
  - *S Haddad-Vanier, C Gicquel, L Boukhatem, K Lazri, and P Chaignon. Virtual network functions placement for defense against distributed denial of service attack. In Proceedings of the 8th ICORES, pages 42-150, 2019.*
- **Second part of the project:**
  - We propose a bilevel formulation which offers a compromise between the achieved security and the costs of NFVs. This approach reduces costs while ensuring a satisfactory level of security.

Mathematical models

- Since the routing of DDoS attacks is unknown, each NFV placement solution is evaluated against the "worst routing" of the malicious flow.

- **Bilevel model**: security in "the worst case“

  - **First level**: Leader problem
    - Decide the NFV placement

  - **Second level**: Follower problem
    - According to the NFVs installed at the first level,
    - determine a routing of the malicious flow allowing as much traffic as possible to reach its target.

Problem Formulation

- **Given**
  - G = (V, E) which models the topology of the network
  - b_e : transmission capacity of each link **e**
  - Set **A** of DDoS attacks, $a \in A$ corresponds to an illegitimate traffic of value ψ^a Mbps from source s^a to target t^a
  - Let $\wp^a$ be the set of all potential paths between s^a and t^a
  - A set of ***N*** type of NFV to install
    - ϕ^n the filtering capacity of each NFV $n \in N$
    - γ^{rn} resource consumption related to NFV deployment
    - K^n cost of NFN $n \in N$
  - cap^r_v: resources available at each node (CPU, Memory etc..)

- **Decide the optimal placement of NFVs such as**
  - All attacking traffic is filtered
  - The total cost is minimal

Mathematical Formulation

- **Decision variables:**
  - X^n_v = number of NFVs of type **n** placed at node **v**
  - φ^a_v = filtering capacity dedicated for attack **a** at node **v**

  - f^a_p = illegitimate flow of attack **a** routed on path **p**
  - z^a_v = 1 if there is a positive flow of attack **a** filtered at node v, 0 otherwise.

- **Constraints:**
  - computing resources available at each node
  - NFV filtering capacities at each node
  - Filtering constraints of all malicious traffic

- **Goal:**
  - Minimize the total cost of deploying NFVs

Bilevel Programming Formulation

Leader problem

- **Decide NVF placement:**
  compute X^{n}_{v} and φ^{a}_{v} values
- **Minimize NVF total Cost**

Follower problem

- **Given φ^{a}_{v} decided by the Leader**
- **Compute the worst routing f^{a}_{p} of attacks**
- **Maximize total unfiltered malicious flow**

Bilevel Programming Formulation

- **Objectif: $Z = \min \sum_v \sum_n K^n X^n_v$**

Bilevel Programming Formulation

- Objectif: $Z = \min \sum_v \sum_n K^n_v X^n_v$
- Under the constraints:
  - Consumption constraints of each resource r at each node v

$$\sum_n \gamma^{rn} X^n_v \leq cap^r_v \;\; \forall\, v \in V \text{ and } \forall\, r \in R$$

Bilevel Programming Formulation

- **Objectif: Min $z = \sum_v \sum_n K^n_v \, X^n_v$**
- **Under the constraints:**
  - **Consumption constraints of each resource r at each node v**

$$\sum_n \gamma^{rn} \, X^n_v \leq cap^r_v \qquad \forall\, v \in V \text{ and } \forall\, r \in R$$

  - **NFV filtering capacity constraints at each node v**

$$\sum_a \varphi^a_v \leq \sum_n \phi^n \, X^n_v \qquad \forall\, v \in V$$

$$\varphi^a_t = 0 \qquad \forall a \in A$$

Bilevel Programming Formulation

- Objectif: Min $z = \sum_v \sum_n K^n_v X^n_v$
- **Under constraints:**
  - Consumption constraints of each resource r at each node v

    $$\sum_n \gamma^{rn} X^n_v \leq cap^r_v \qquad \forall v \in V \text{ and } \forall r \in R$$

  - NFV filtering capacity constraints

    $$\sum_a \varphi^a_v \leq \sum_n \phi^n X^n_v \qquad \forall v \in V$$

    $$\varphi^a_t = 0 \qquad \forall a \in A$$

  - Filtering constraints of total damage inflicted to the attack targets

    $$D = 0$$

  - X^n_v integer, $\varphi^a_v \geq 0$, $(X, \varphi) \in Z^{NV}_+ \times R^{AV}_+$

Bilevel Programming Formulation

- D the optimal value of follower problem:

$$\text{Max} \sum_a \left(\sum_p f^a_p - \sum_v z^a_v \varphi^a_v\right)$$

Subject to

Link capacity constraints $\forall\, e \in E$

$$\sum_a \sum_p f^a_p \leq b_e$$

Routing constraints for each attack $a \in A$

$$\sum_p f^a_p \leq \psi^a$$

$$z^a_v \geq \left[\sum_{v \in p} f^a_p \,/\, \psi^a\right] \forall\, a \in A\ \forall\, v \in V$$

- $z^a_v \in \{0,1\},\ f^a_p \geq 0\ \forall\ v \in V,\ a \in A,\ p \in \wp^a$

$$Z^*_{BL} = \min \sum_{v \in V} \sum_{n \in N} K^n x^n_v$$

$$\sum_{n \in N} \gamma^{rn} x^n_v \leq Cap^r_v$$

$$\sum_{a \in A} \varphi^a_v \leq \sum_{n \in N} \phi^n x^n_v$$

$$\varphi^a_{t^a} = 0$$

$$D = 0$$

$$(x, \varphi) \in \mathbb{Z}^{NV}_+ \times \mathbb{R}^{AV}_+$$

Leader

Follower

$$D = \begin{cases} \max \sum_{a \in A} \Big(\sum_{p \in P^a} f^a_p - \sum_{v \in V} z^a_v \varphi^a_v \Big) & \\ \sum_{a \in A} \sum_{e \in p, p \in P^a} f^a_p \leq b_e & \forall e \in E \\ \sum_{p \in P^a} f^a_p \leq \psi^a & \forall a \in A \\ z^a_v \geq \dfrac{\sum_{v \in p \in P^a} f^a_p}{\psi^a} & \forall v \in V \\ z^a_v \in \{0, 1\} & \forall v \in V \end{cases}$$

Decomposition approach

Leader

current filtering capacity allocation $\overline{\varphi}^a$

Follower

$$D = \begin{cases} Z^*_{AP}(\overline{\varphi}^a) = \max \sum_{p \in \mathcal{P}^a} f_p^a - \sum_{v \in V} \overline{\varphi}_v^a z_v^a \\ \sum_{a \in A} \sum_{e \in p, p \in P^a} f_p^a \leq b_e & \forall e \in E \\ \sum_{p \in P^a} f_p^a \leq \psi^a & \forall a \in A \\ z_v^a \geq \frac{\sum_{v \in p \in P^a} f_p^a}{\psi^a} & \forall v \in V \\ z_v^a \in \{0, 1\} & \forall v \in V \end{cases}$$

Decomposition approach

Leader

current filtering capacity allocation $\overline{\varphi}^a$

Follower

$$D = \begin{cases} Z^*_{AP}(\overline{\varphi}^a) = \max \sum_{p \in \mathcal{P}^a} f^a_p - \sum_{v \in V} \overline{\varphi}^a_v z^a_v & \\ \sum_{a \in A} \sum_{e \in p, p \in P^a} f^a_p \leq b_e & \forall e \in E \\ \sum_{p \in P^a} f^a_p \leq \psi^a & \forall a \in A \\ z^a_v \geq \frac{\sum_{v \in p \in P^a} f^a_p}{\psi^a} & \forall v \in V \\ z^a_v \in \{0, 1\} & \forall v \in V \end{cases}$$

$\overline{P}^a_i \subset P^a$

paths p such that $\overline{f}^a_p > 0$

Leader

current filtering capacity allocation $\overline{\varphi}^a$

Follower

$\sum_{v \in V(\overline{P}_i^a)} \varphi_v^a \geq \sum_{p \in \overline{P}_i^a} \overline{f}_p^a$

Add the filtering constraint

$\overline{P}_i^a \subset P^a$

paths p such that $\overline{f}_p^a > 0$

$$D = \begin{cases} Z^*_{AP}(\overline{\varphi}^a) = \max \displaystyle\sum_{p \in \mathcal{P}^a} f_p^a - \sum_{v \in V} \overline{\varphi}_v^a z_v^a & \\ \displaystyle\sum_{a \in A} \sum_{e \in p, p \in P^a} f_p^a \leq b_e & \forall e \in E \\ \displaystyle\sum_{p \in P^a} f_p^a \leq \psi^a & \forall a \in A \\ z_v^a \geq \dfrac{\sum_{v \in p \in P^a} f_p^a}{\psi^a} & \forall v \in V \\ z_v^a \in \{0,1\} & \forall v \in V \end{cases}$$

$$Z^*_{BL} = \min \sum_{v \in V} \sum_{n \in N} K^n x^n_v$$

$$\sum_{n \in N} \gamma^{rn} x^n_v \leq Cap^r_v$$

$$\sum_{a \in A} \varphi^a_v \leq \sum_{n \in N} \phi^n x^n_v$$

$$\varphi^a_{t^a} = 0$$

$$\sum_{v \in V(\overline{P}^a_i)} \varphi^a_v \geq \sum_{p \in \overline{P}^a_i} \overline{f}^a_p$$

Leader

current filtering capacity allocation $\overline{\varphi}^a$

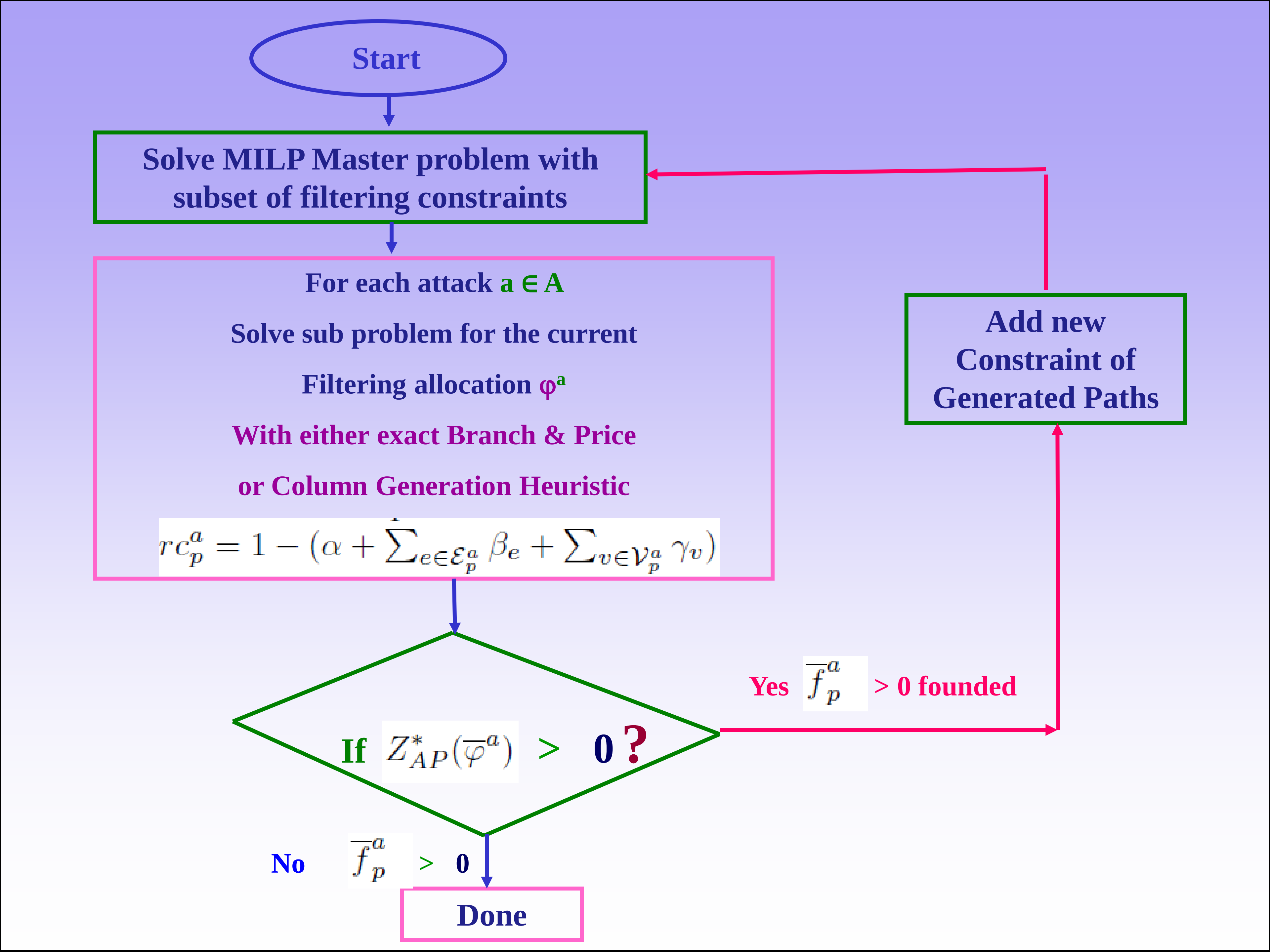
Follower

Add the filtering constraint

$$\overline{P}^a_i \subset P^a$$

paths p such that $\overline{f}^a_p > 0$

$$D = \begin{cases} Z^*_{AP}(\overline{\varphi}^a) = \max \sum_{p \in \mathcal{P}^a} f^a_p - \sum_{v \in V} \overline{\varphi}^a_v z^a_v & \\ \sum_{a \in A} \sum_{e \in p, p \in P^a} f^a_p \leq b_e & \forall e \in E \\ \sum_{p \in P^a} f^a_p \leq \psi^a & \forall a \in A \\ z^a_v \geq \dfrac{\sum_{v \in p \in P^a} f^a_p}{\psi^a} & \forall v \in V \\ z^a_v \in \{0, 1\} & \forall v \in V \end{cases}$$

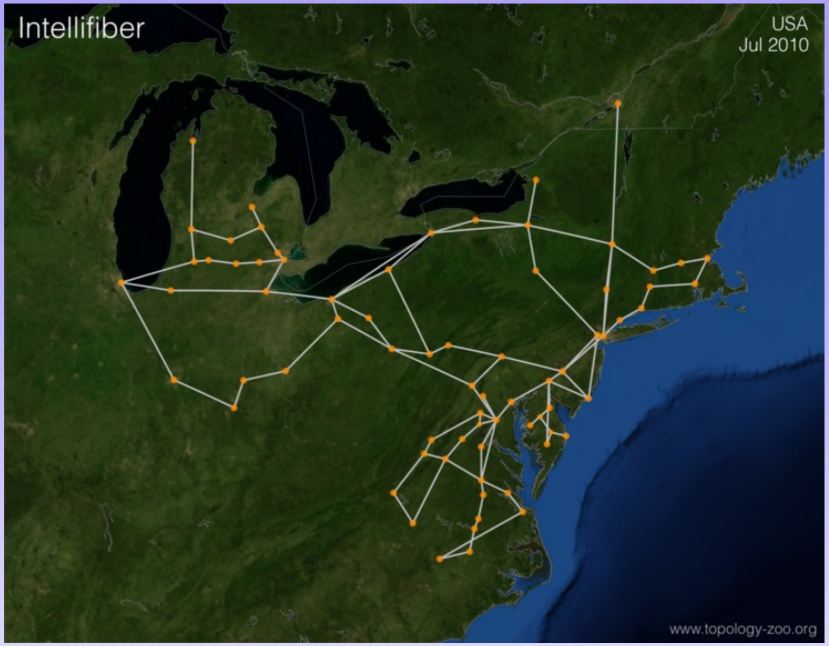


Start
Solve MILP Master problem with subset of filtering constraints
For each attack a ∈ A
Solve sub problem for the current
Filtering allocation φ^a
With either exact Branch & Price
or Column Generation Heuristic
$rc_p^a = 1 - (\alpha + \sum_{e \in \mathcal{E}_p^a} \beta_e + \sum_{v \in \mathcal{V}_p^a} \gamma_v)$
Add new Constraint of Generated Paths
If $Z_{AP}^*(\overline{\varphi}^a) > 0$?
Yes $\overline{f}_p^a > 0$ founded
No $\overline{f}_p^a > 0$
Done

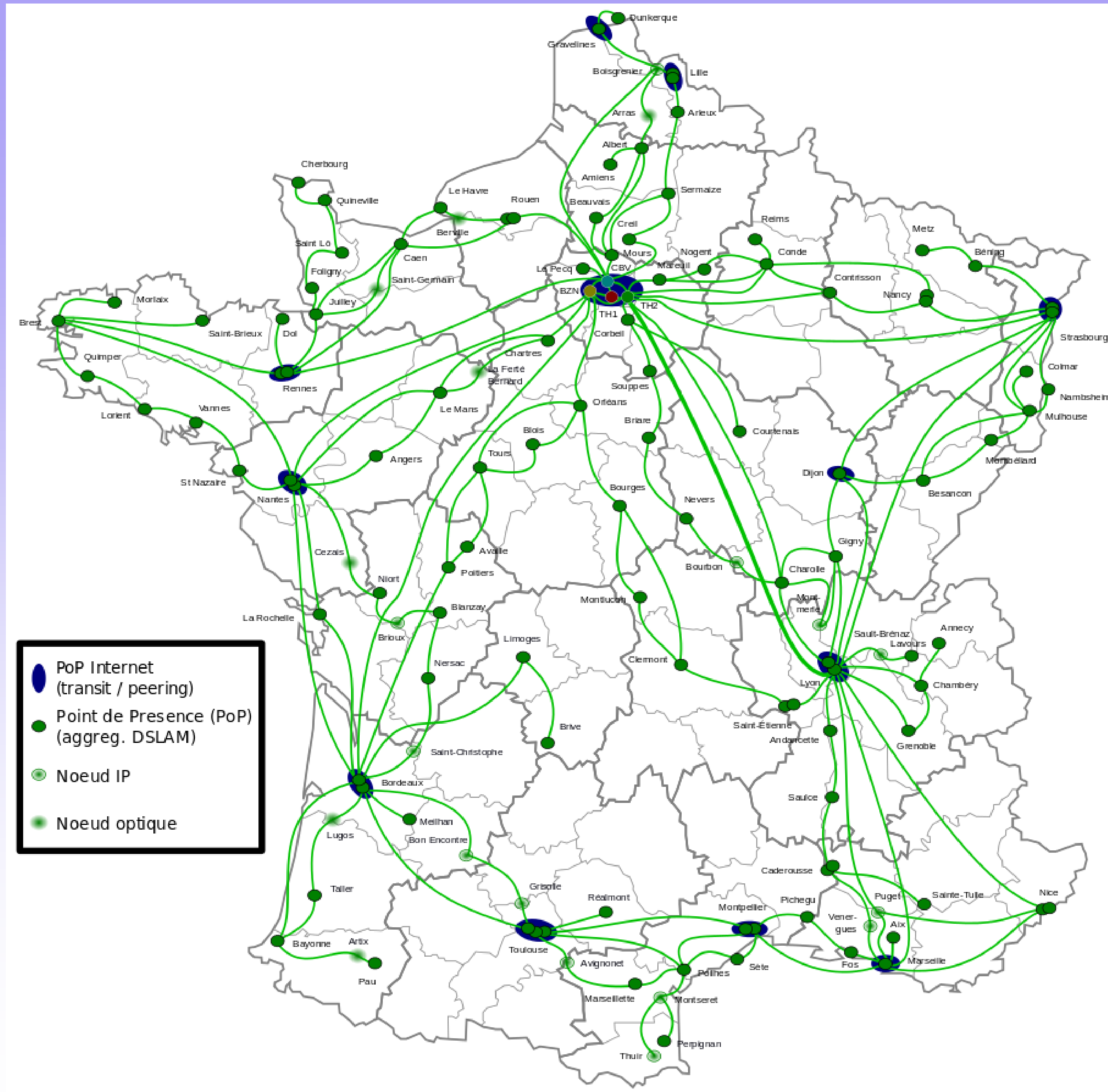
Numerical Results

- **Instances:**
  - https://github.com/Orange-OpenSource/synthetic-tm-generator
  - https://fr.wikipedia.org/wiki/Free_(entreprise)#/media/File:Proxad_FR.svg
  - Internet Topology Zoo, http://www.topology-zoo.org/dataset.html
  - On-Demand EC2 prices, https://aws.amazon.com/fr/ec2/pricing/on-demand/
  - All tests were run on a an Intel Core i5 (1.9GHz) with 16 GB of RAM, running under Windows 10 using Cplex 12.8.9.


BICS
Europe
Jan 2011
www.topology-zoo.org


Intellifiber
USA
Jul 2010
www.topology-zoo.org


Cogent
USA, Europe
Aug 2010
www.topology-zoo.org



Dunkerque
Gravelines
Boisgrenier
Lille
Arras
Arleux
Albert
Amiens
Sermaize
Beauvais
Creil
Cherbourg
Quineville
Le Havre
Rouen
Berville
Saint Lô
Caen
Foligny
Saint-Germain
Julley
Morlaix
Brest
Saint-Brieux
Dol
Quimper
Rennes
Lorient
Vannes
St Nazaire
Nantes
Reims
Metz
Béning
Mours
Nogent
Conde
CBV
Mareuil
La Pecq
BZN
TH1
TH2
Corbeil
Contrisson
Nancy
Strasbourg
Chartres
La Ferté Bernard
Le Mans
Souppes
Orléans
Colmar
Nambsheim
Mulhouse
Briare
Courtenais
Blois
Tours
Angers
Montbéliard
Dijon
Besancon
Bourges
Nevers
Gigny
Cezais
Availle
Poitiers
Bourbon
Charolle
Niort
Montluçon
Mont-merle
La Rochelle
Blanzay
Brioux
Limoges
Sault-Brénaz
Annecy
Lavours
Clermont
Lyon
Chambery
Nersac
Saint-Etienne
Andancette
Brive
Grenoble
Saint-Christophe
Bordeaux
Saulce
Meilhan
Lugos
Bon Encontre
Caderousse
Taller
Grisolle
Réalmont
Sainte-Tulle
Nice
Puget
Montpellier
Pichegu
Vener-gues
Aix
Bayonne
Artix
Toulouse
Pau
Avignonet
Sète
Fos
Marseille
Polines
Marseillette
Montseret
Perpignan
Thuir
PoP Internet (transit / peering)
Point de Presence (PoP) (aggreg. DSLAM)
Noeud IP
Noeud optique

Numerical Results

We use 4 internet network topologies

- **BICS**: $|V|=32$, $|E|=48$
- **IntelliFiber**: $|V|=73$, $|E|=96$
- **Colt Telecom**: $|V|=153$, $|E|= 179$
- **Cogentco**: $|V|=196$, $|E|=245$

And one Free network topology **Free**: $|V|=120$, $|E|= 167$

- Number of source-target pairs was set to $A \in \{5,10,15,20\}$
  - We considered different sources and a single target randomly selected.

- For each considered network topology and value of A, we randomly generated 5 instances, leading to a total of 100 instances.

Numerical Results

		EXACT				HEUR					
Topology	A	$Cost$	$\#IT$	$\#FC$	$Time$(s)	$Cost$	$\#IT$	$\#FC$	$Time$(s)	$\#Inf$	$Max\%UF$
BICS	5	1144	12	39	7	1144	12	35	1	0	0.00%
	10	4004	11	41	9	4004	12	43	3	2	0.76%
	15	5590	11	56	12	5590	11	58	58	1	1.81%
	20	7150	15	77	23	7150	16	81	4	0	0.00%
IntelliFiber	5	1508	10	25	7	1508	10	26	2	0	0.00%
	10	2392	13	55	19	2392	13	53	4	0	0.00%
	15	3874	16	83	34	3874	17	85	9	3	0.43%
	20	4290	14	103	39	4290	15	108	11	0	0.00%
Free	5	1482	7	16	5	1482	7	16	5	0	0.00%
	10	3042	8	32	13	3042	9	34	6	1	0.13%
	15	3822	9	49	21	3822	8	47	9	0	0.00%
	20	4680	8	67	26	4680	11	73	16	1	0.11%
Colt Telecom	5	1768	11	31	21	1768	10	27	9	0	0.00%
	10	2522	14	65	35	2522	12	51	14	3	0.29%
	15	3042	19	117	81	3042	18	115	35	2	0.44%
	20	5564	22	153	124	5564	25	159	77	3	1.00%
Cogentco	5	1040	26	73	65	1040	21	66	26	0	0.00%
	10	2002	53	231	586	2002	40	194	3534	3	1.71%
	15	3978	46	268	494	3978	43	249	696	3	3.02%
	20	4680	53	419	765	4680	50	379	2319	4	1.12%

Numerical Results

			LCG		DEC			
Topology	V	E	$Cost$	$Time$(s)	$Cost$	$\#IT$	$\#FC$	$Time$(s)
Goodnet	17	31	1473	0.3	858	6	22	1.2
BICS	33	48	1690	0.6	754	12	51	3.2
IntelliFiber	73	96	1737	2.4	1040	10	44	5.2
Free	120	167	1560	4.5	1014	6	22	6.6
Cogentco	197	245	1950	12.5	767	24	99	44.8

Conslusion

- **In summary we talked about:**
  - **Security mechanisms against DDoS attacks that use the flexibility and efficiency of network virtualization SDN and NFV.**
  - **The proposed bilevel formulation of the problem offers a compromise between the achieved security and the costs of NFVs. This approach reduces significantly costs while ensuring a high level of security.**
  - **The decomposition algorithm efficiently solved the generated instances**

Future work :

– Develop a polyhedral study to improve the decomposition approach, generate new inequalities valid for both the master problem and the sub problem.

– Improve problem formulation in order to reduce costs, deepen research on bilevel and robust optimization approaches for security issues.

– Extend this work to the deployment of different virtual functions for new services in future heterogeneous networks.