

Thèse pour obtenir le grade de

Docteur en sciences de l'Université Pierre et Marie Curie

spécialité Informatique

Une approche pour l'estimation fiable des propriétés de la topologie d'Internet

Elie Rotenberg

Rapporteurs

Bertrand Jouve

Directeur de recherches, CNRS

Jean-Jacques Pansiot

Professeur émérite, Strasbourg

Examineurs

Clémence Magnien

Chargée de recherches, CNRS

Pascal Mérindol

Maître de conférences, Strasbourg

Philippe Owezarski

Directeur de recherches, CNRS

Directeur

Matthieu Latapy

Directeur de recherches, CNRS

Co-directeur

Christophe Crespelle

Maître de conférences, UCBL

Organisation de l'exposé

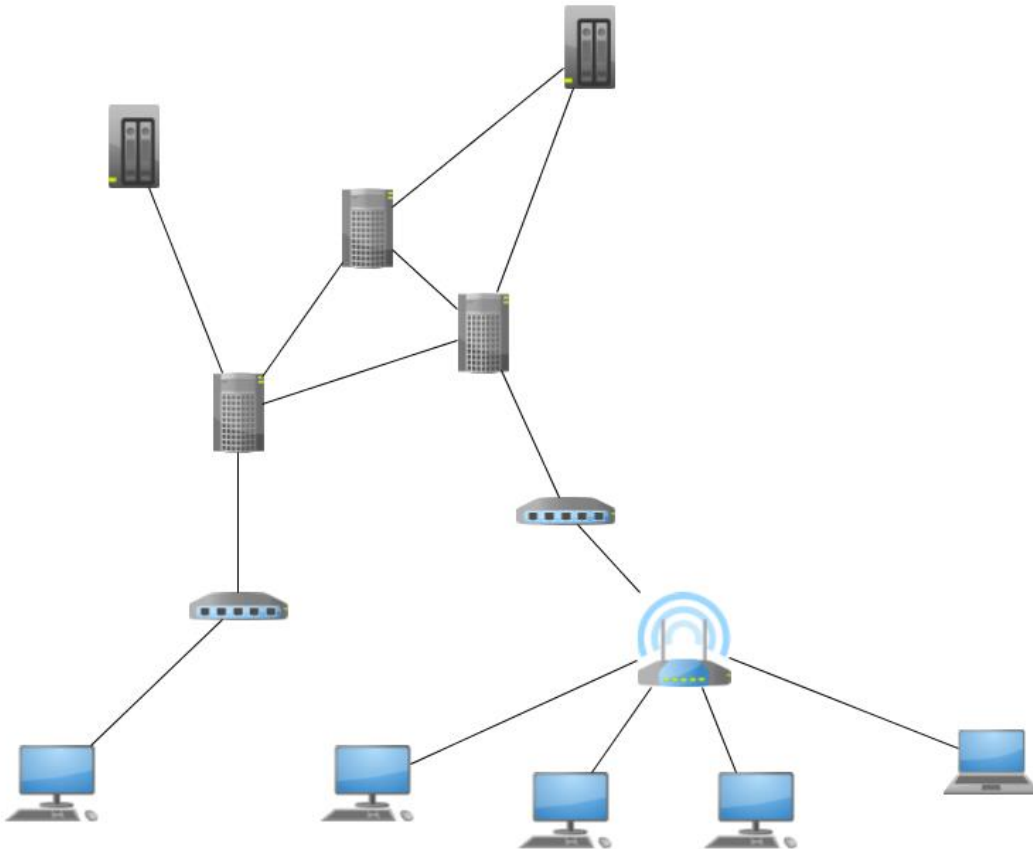
1. Une nouvelle approche de la métrologie d'Internet
2. Contributions
 1. Topologie logique
 2. Topologie physique
 1. Distribution de degré des routeurs du cœur
 2. Tables de transmission
3. Conclusions et perspectives

Organisation de l'exposé

1. Une nouvelle approche de la métrologie d'Internet
2. Contributions
 1. Topologie logique
 2. Topologie physique
 1. Distribution de degré des routeurs du cœur
 2. Tables de transmission
3. Conclusions et perspectives

Une nouvelle approche de la métrologie d'Internet

Topologie d'Internet



- Modélisation
- Routage
- Résilience aux pannes
- Résistance aux attaques
- ...

Approches historiques

- 1) Cartes basées sur les déclarations des autorités administratives
ARPANET Map, BNN Technologies, 1977
- 2) Modèles « *bottom-up* » basés sur une connaissance *a priori* des éléments
A better model for generating test networks, Doar, Nexion, 1996
- 3) Interprétation de cartes issues de mesures avec des outils de diagnostic
DIMES: Let the Internet Measure Itself, Shavir, Shir, 2005
- 4) Notre approche : mesure orientée propriété

Notre approche

- **Description rigoureuse** de nos objets et de nos outils
- Mesure **directe** d'observables topologiques locaux
- **Echantillonnage** rigoureux du réseau
- Méthode d'inférence d'une propriété globale
- Pas de cartes intermédiaires !



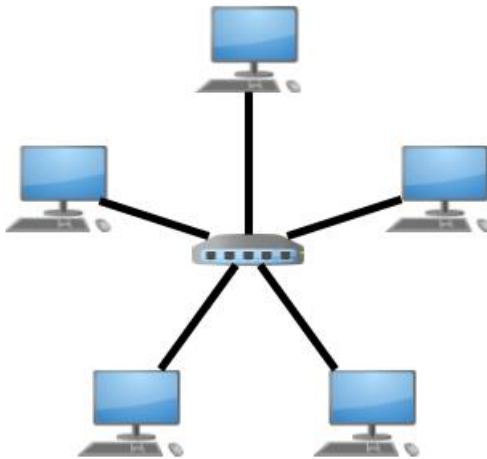
Estimation fiable d'une propriété topologique du réseau

Organisation de l'exposé

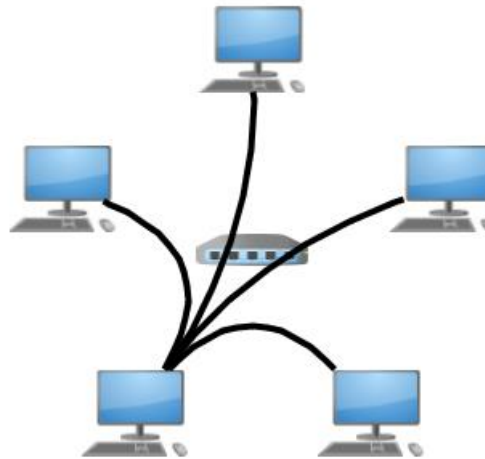
1. Une nouvelle approche de la métrologie d'Internet
2. Contributions
 1. Topologie logique
 2. Topologie physique
 1. Distribution de degré des routeurs du cœur
 2. Tables de transmission
3. Conclusions et perspectives

Contributions

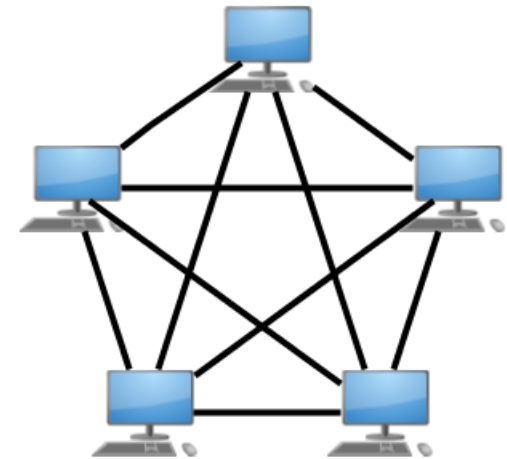
Topologie physique et topologie logique



Topologie physique



Liens logiques induits



Topologie logique

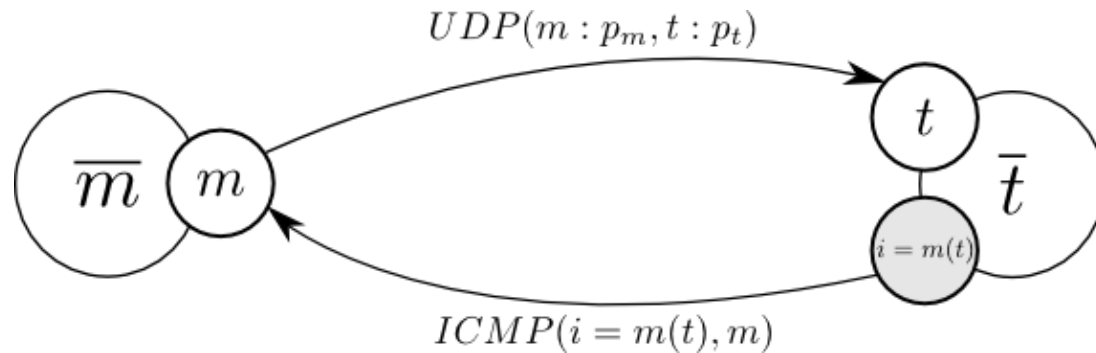
- Machines physiques
- Câbles (ou équivalents)



- Entités logiques
- Communication en un saut IP

Distribution de degrés au niveau physique

UDP Ping

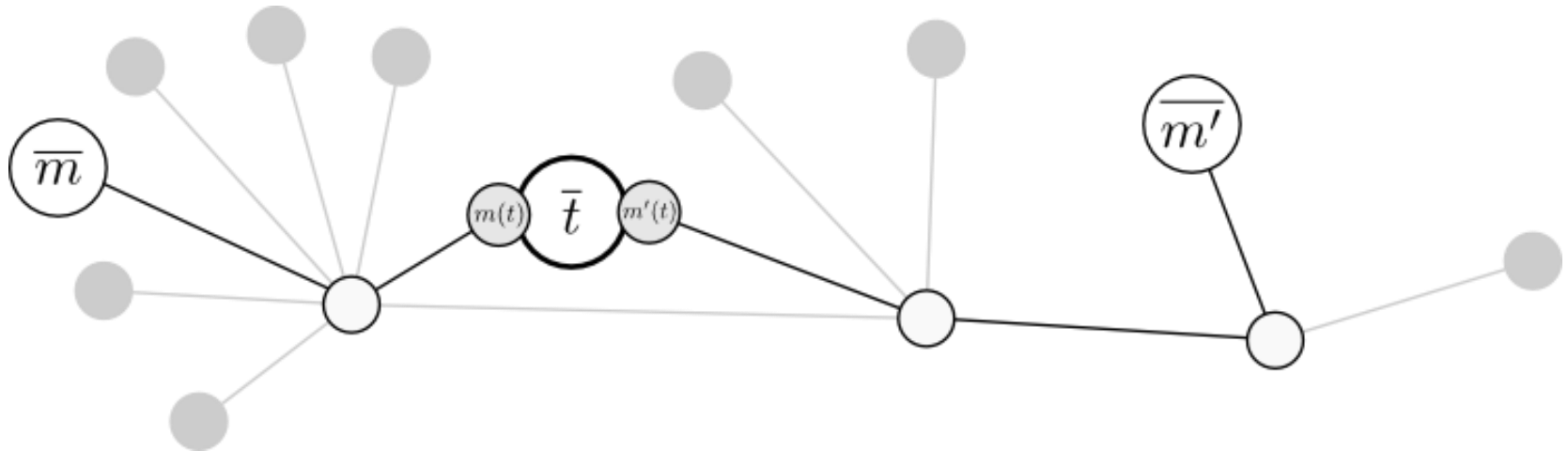


UDP Ping(m, t)

m envoie un message X malformé à t
 t reçoit X (expéditeur: m)
 t génère un message d'erreur Y
 t choisit son interface i pour envoyer Y à m
 m reçoit Y (expéditeur: i)

L'interface de réponse i appartient à t .

UDP Ping



UDP Ping(m, t)

m envoie un message X malformé à t

t reçoit X (expéditeur: m)

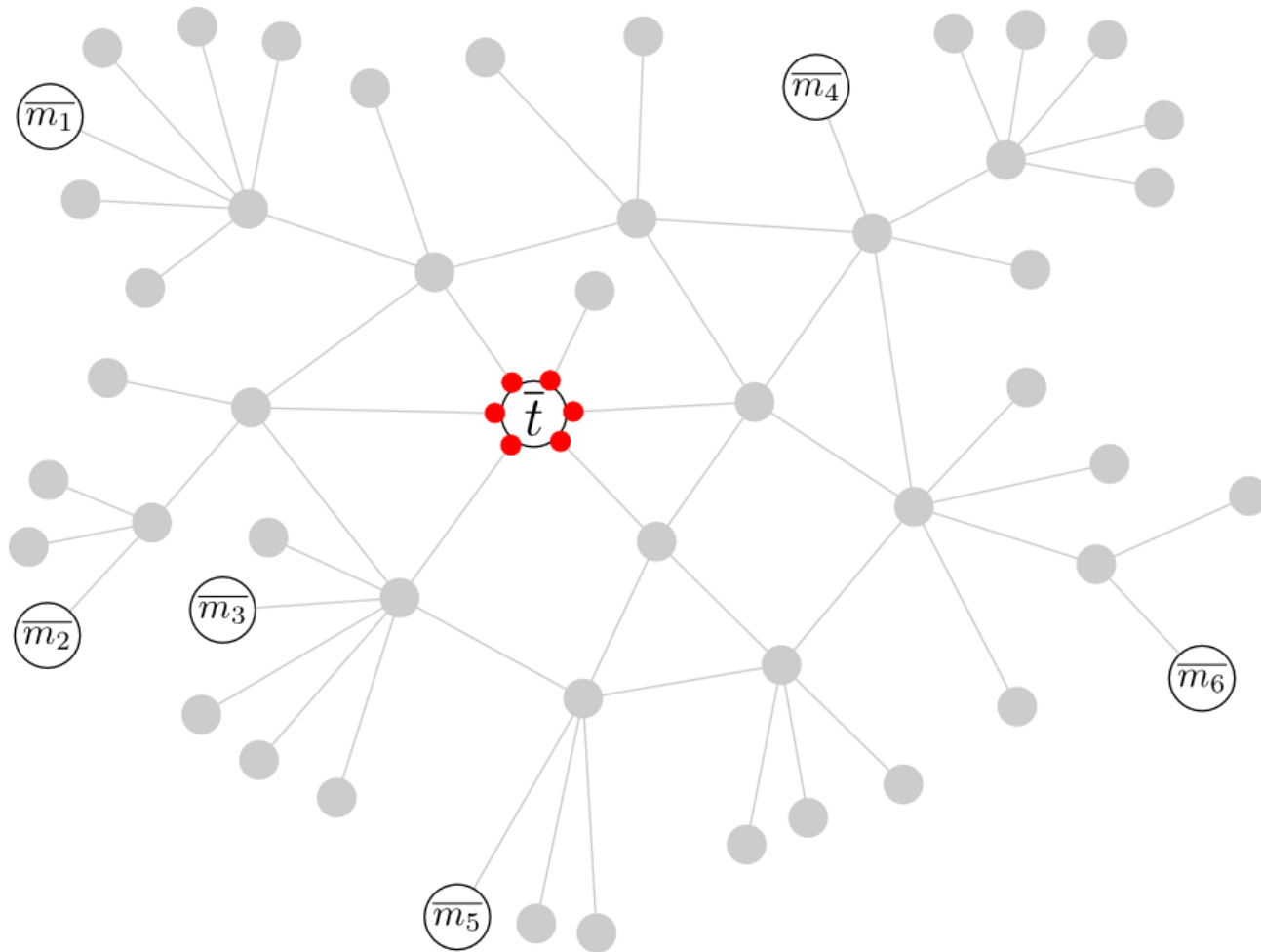
t génère un message d'erreur Y

t choisit son interface i pour envoyer Y à m

m reçoit Y (expéditeur: i)

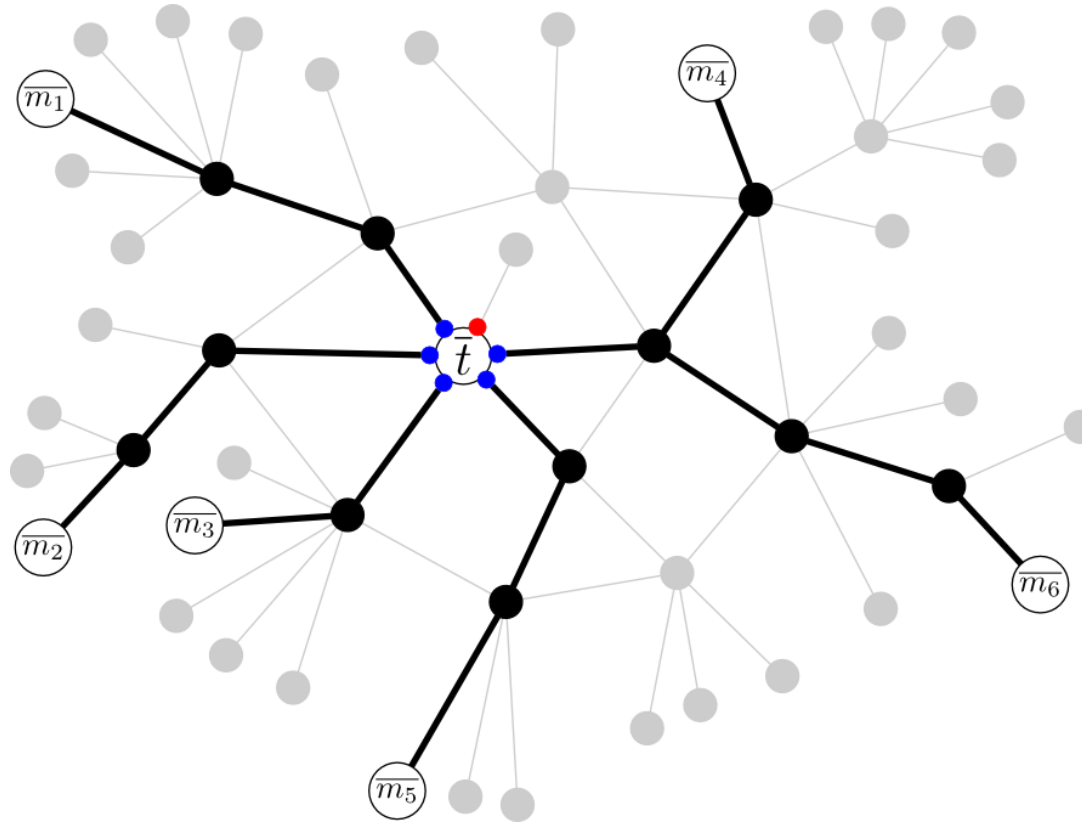
Des moniteurs distincts peuvent permettre
d'observer des interfaces distinctes.

UDP Ping distribué



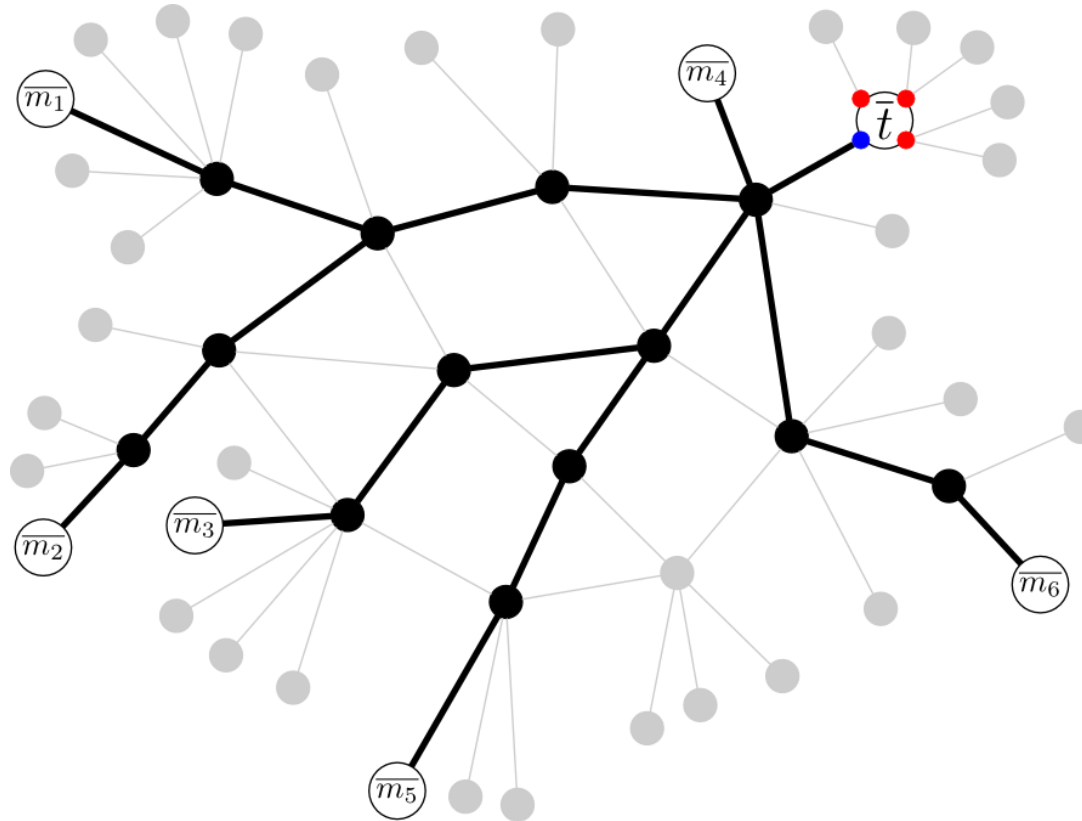
Peut-on observer toutes les interfaces d'une cible ?

Cas d'une cible dans le cœur



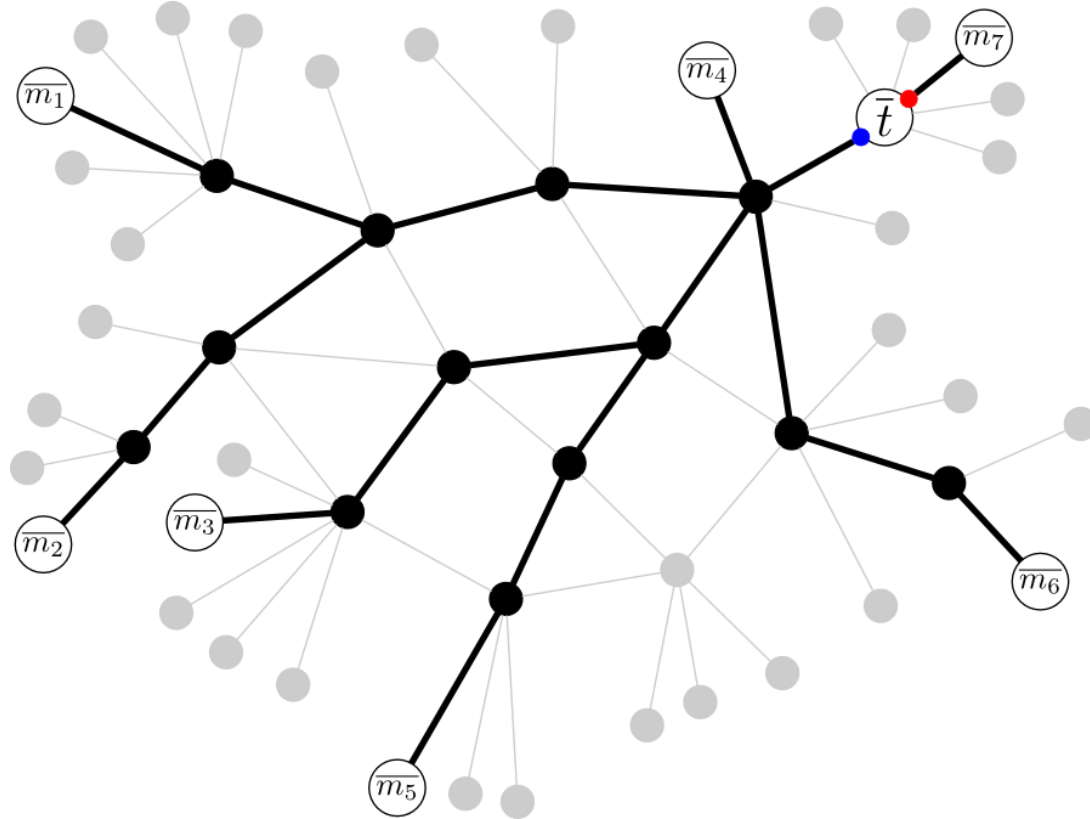
Les interfaces tournées vers le cœur sont toutes observées.
Les interfaces tournées vers le bord ne sont pas observées.

Cas d'une cible dans le bord



Les interfaces tournées vers le cœur sont toutes observées.
Les interfaces tournées vers le bord ne sont pas observées.

Cas problématique



Les interfaces tournées vers le cœur sont toutes observées.
L'une des interfaces tournées vers le bord est observée.

Domaine de pertinence d'UDP Ping distribué

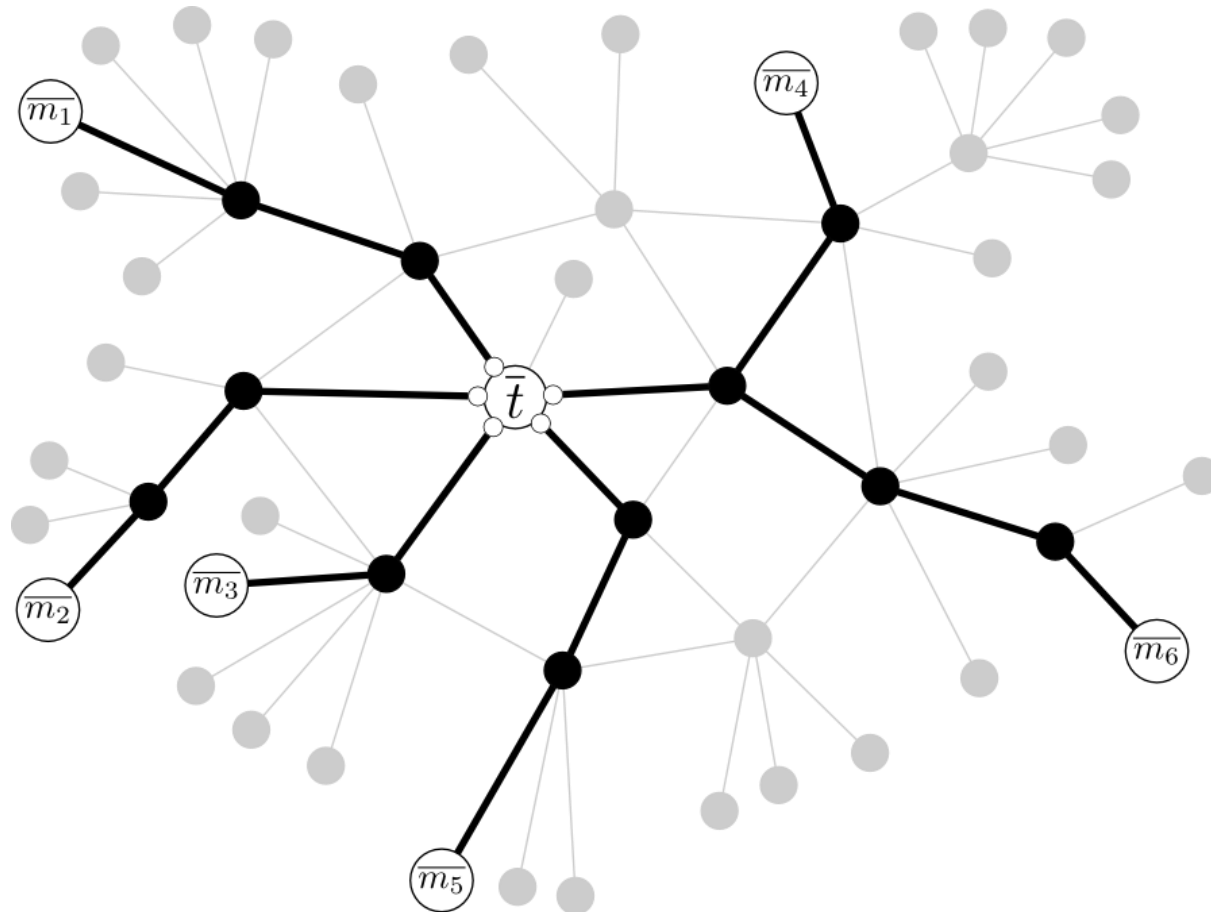
Hypothèses

1. Notre ensemble de moniteurs est suffisamment grand et bien réparti
2. La cible est dans le cœur et répond à UDP Ping
3. On ignore les interfaces de réponse « problématiques »

Résultat

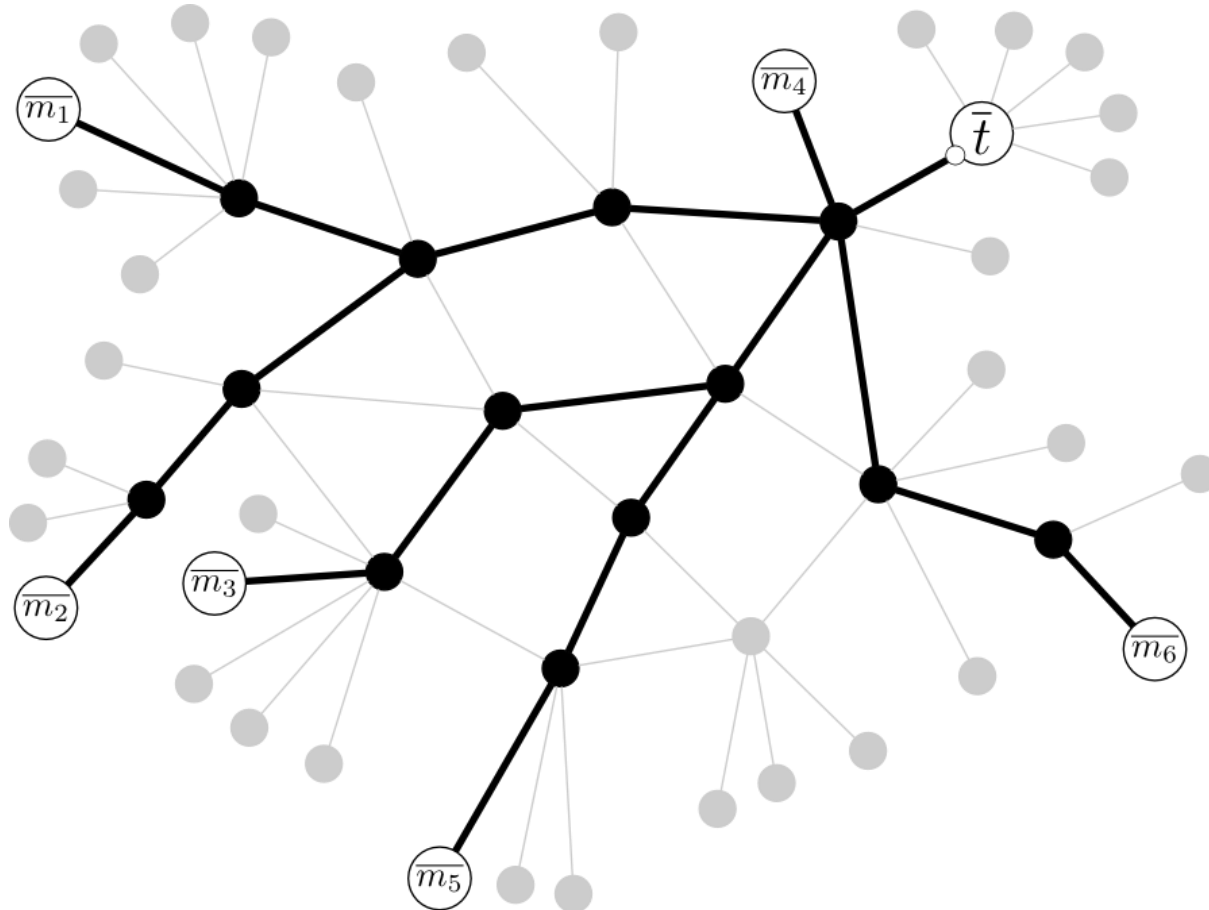
- La liste des interfaces **tournées vers le cœur** de la cible.

Caractérisation des cibles dans le cœur



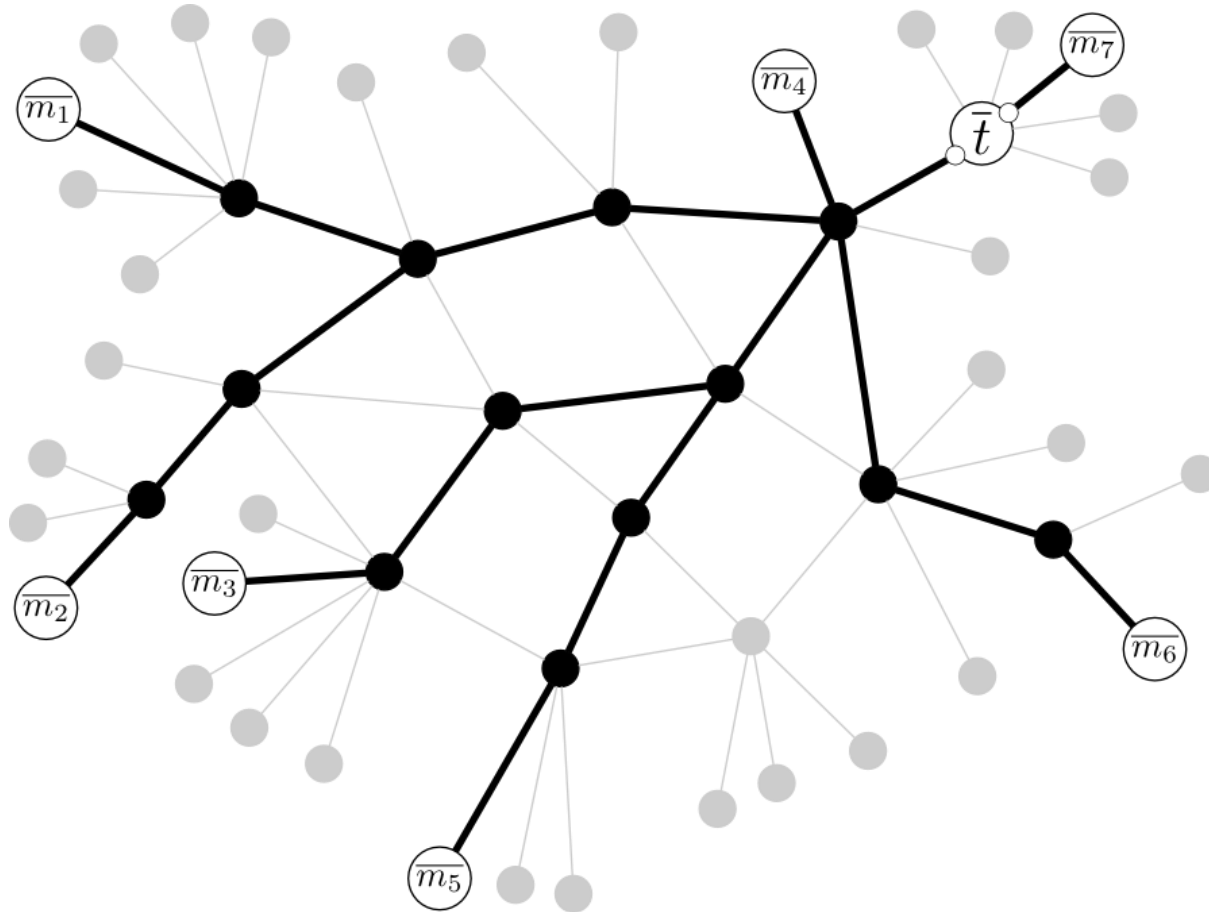
La cible t est dans le cœur : on observe au moins deux interfaces tournées vers le cœur.

Caractérisation des cibles dans le cœur



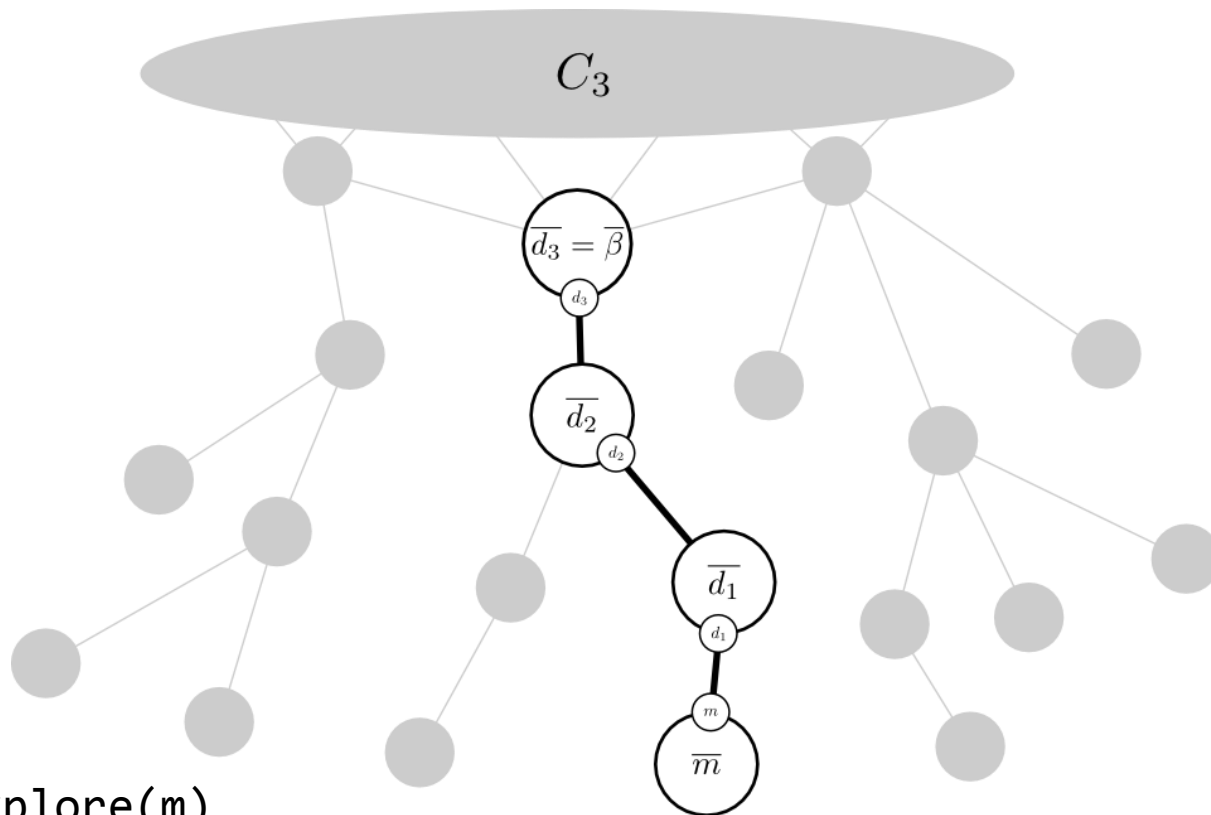
La cible t est dans le bord : on observe une unique interface tournée vers le cœur.

Caractérisation des cibles dans le cœur



Cas problématique : un moniteur est situé « derrière » une cible et deux interfaces sont observées.

Caractérisation des cibles problématiques



UDP Explore(m)

Pour chaque k à partir de 1,

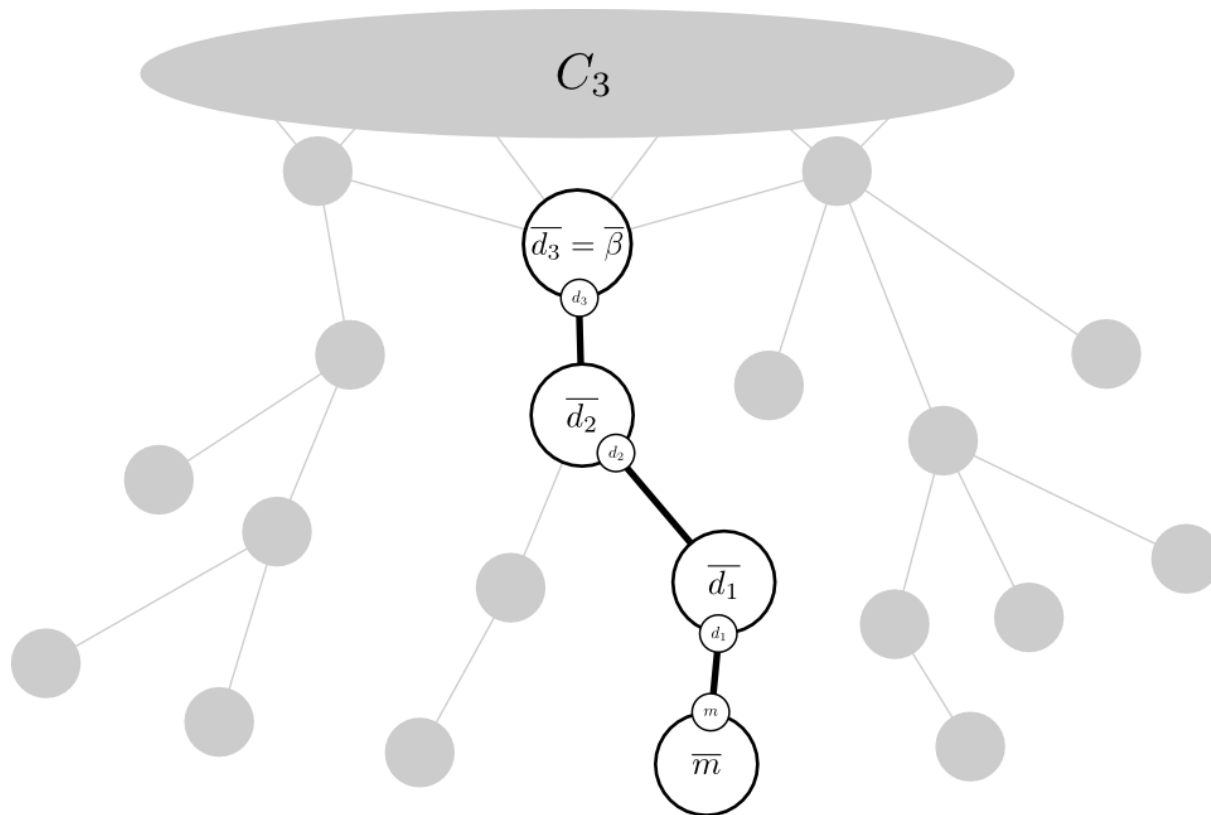
envoyer N sondes de TTL = k vers des cibles aléatoires

soit L_k la liste des interfaces ayant répondu

tant que $|L_k| = 0$ ou $|L_k| = 1$

Renvoyer la réunion des L_k .

Caractérisation des cibles problématiques



On obtient la liste des interfaces dans le bord observables par UDP Ping depuis m .

Mesure du degré dans le cœur d'une cible

Hypothèses

1. Notre ensemble de moniteurs est suffisamment grand et bien réparti
2. La cible est dans le cœur et répond à UDP Ping
3. On ignore les interfaces de réponse « problématiques »

Résultat

- La liste des interfaces **tournées vers le cœur** de la cible.
- Le nombre de ces interfaces est le **degré dans le cœur** de la cible.

Echantillonnage d'adresses de routeurs du cœur

- 1) Effectuer un tirage aléatoire uniforme d'entiers de 32 bits
- 2) Supprimer les entiers ne correspondant pas à des adresses valides (RFC)
- 3) Exécuter UDP Explore depuis chaque moniteur
- 4) Exécuter UDP Ping depuis chaque moniteur vers chacune des adresses valides
- 5) Supprimer les interfaces observées par UDP Explore des résultats
- 6) Supprimer les cibles ayant moins de 2 interfaces dans le cœur

Correction du biais de sélection

Echantillonner des **adresses** \neq échantillonner des **routeurs**

- Sélection uniforme sur les adresses
- Probabilité de tirage d'un routeur = proportionnelle à son nombre d'adresses

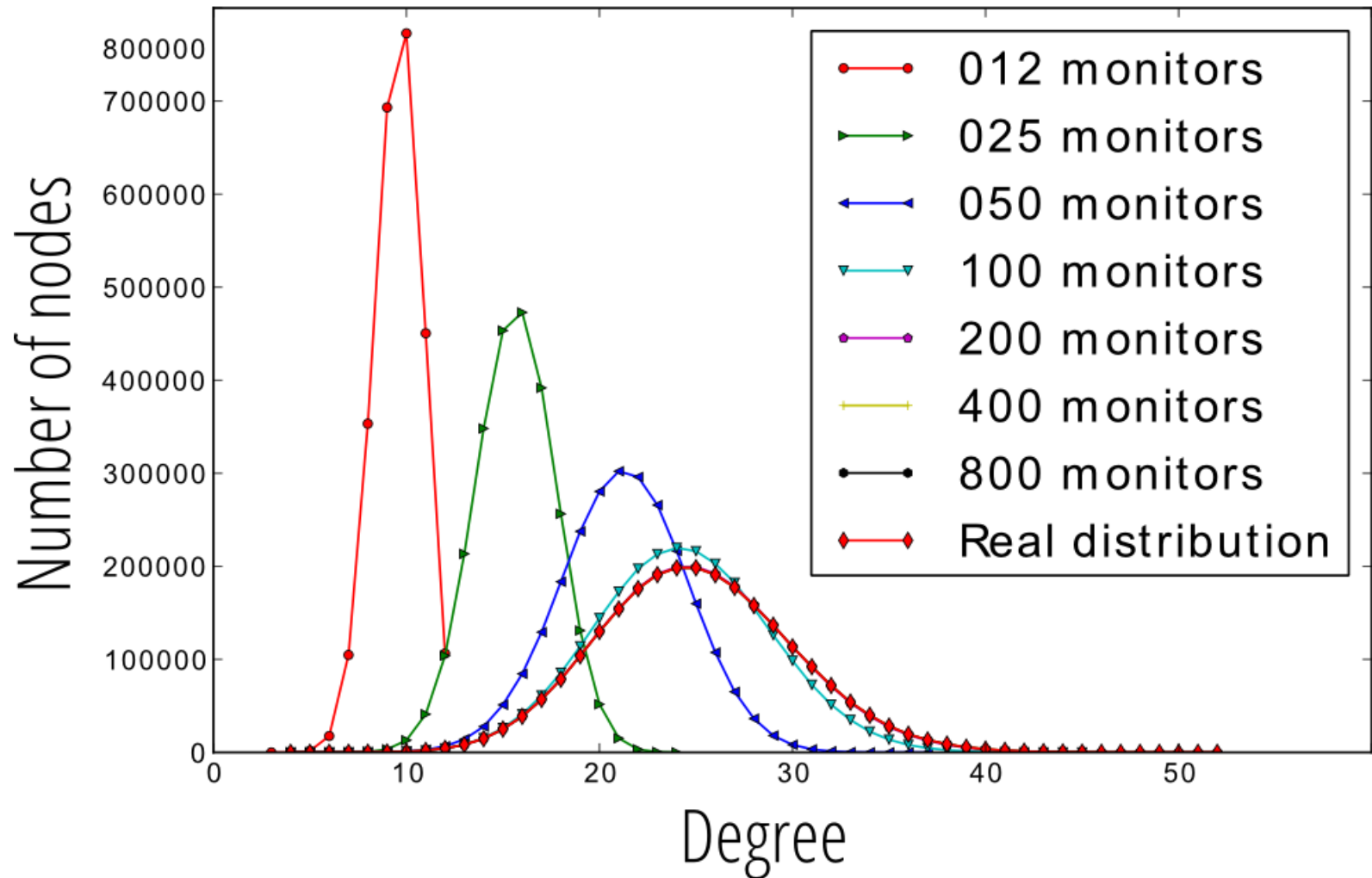
➤ Transformation de correction du biais :

$$p_k = \frac{p'_k}{k} \cdot \frac{1}{\sum_i \frac{p'_i}{i}}$$

Validation de la méthode : protocole

- 1) Générer des graphes aléatoires selon des modèles usuels
- 2) Tirer un grand nombre de nœuds, les *moniteurs*
- 3) Considérer que *tous les nœuds* sont des cibles
- 4) Simuler *UDP Ping* par des plus courts chemins
- 5) Calculer le degré de chaque cible
- 6) Comparer les distributions de degrés mesurée et réelle

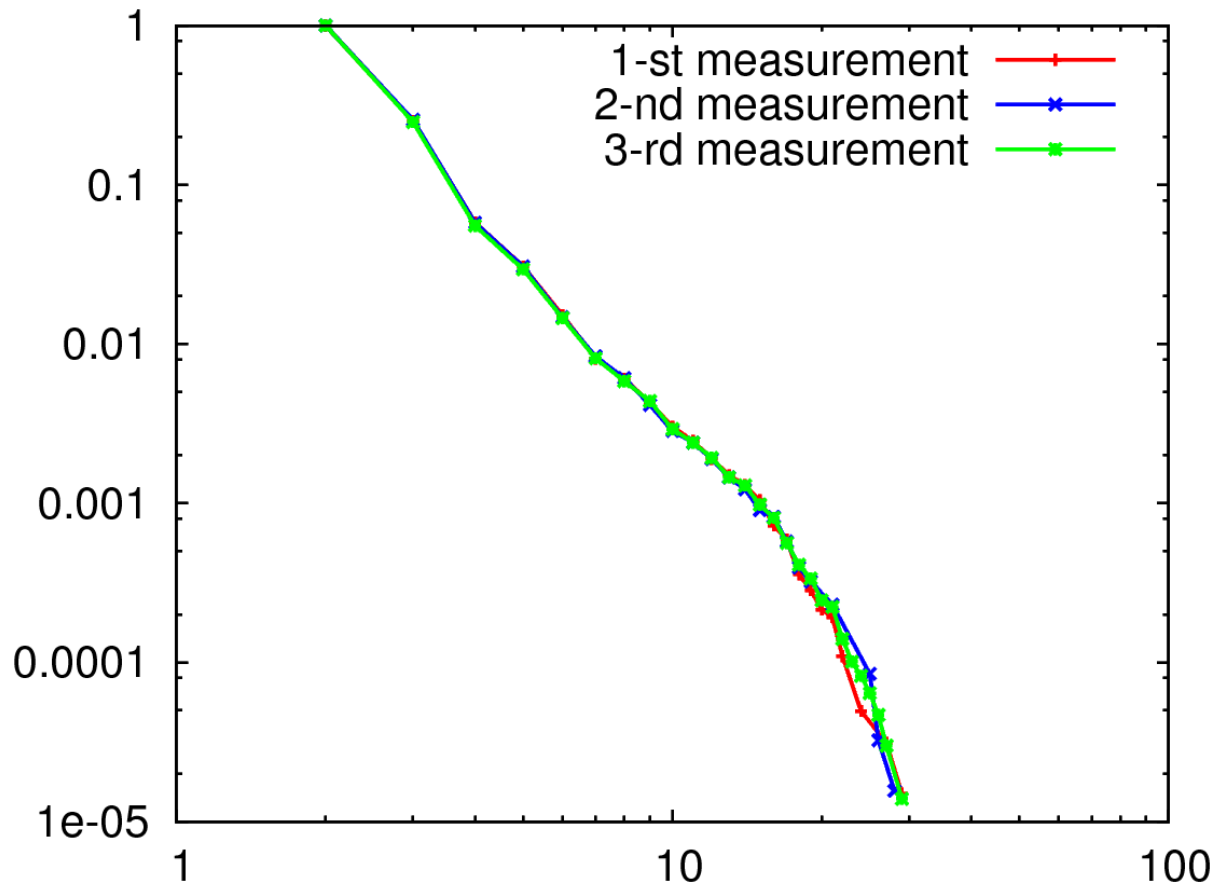
Validation de la méthode : résultats



Mesure réelle sur PlanetLab

- 700 moniteurs initialement
- $3 \cdot 10^6$ cibles initiales échantillonnées en 10 heures
- Mesure répétée 3 fois, chacune durant 4 heures
- 5600 cibles dans le cœur après filtrage

Mesure réelle sur PlanetLab



x : nombre d'interfaces (degré physique dans le cœur)

y : fraction des routeurs du cœur

Validation des résultats

- Réinjection de la distribution mesurée dans les simulations (*bootstrapping*)
- Evaluation de la qualité de l'ensemble des moniteurs
 - Classes de colocalisation
 - Convergence des résultats

Limites de l'approche

- Relativement peu de cibles répondent à UDP Ping
- Précision limitée pour la fraction des nœuds de fort degré
- Dépendance à la répartition d'un ensemble de moniteurs
- Pertinence limitée aux routeurs du cœur

Conclusions et perspectives

Organisation de l'exposé

1. Une nouvelle approche de la métrologie d'Internet
2. Contributions
 1. Topologie logique
 2. Topologie physique
 1. Distribution de degré des routeurs du cœur
 2. Tables de transmission
3. Conclusions et perspectives

Contributions

- Définition rigoureuse des objets et des outils de la topologie d'Internet
- Une nouvelle approche : la mesure orientée propriété
- Outils de mesure : UDP Ping, UDP Ping distribué, UDP Explore
- Estimation de la distribution de degré de la topologie physique du cœur
- Travaux préliminaires sur la topologie logique
- Travaux qualitatifs sur les tables de transmission des routeurs du cœur

Perspectives

- Approfondissement d'*UDP Ping*

Validation, dynamique, mesure longue, autres ensembles de moniteurs

- Echantillonnage orienté propriété

Application à d'autres réseaux, marche aléatoire orientée propriété

- Nouveaux objets d'intérêt

Réseau de routage pondéré, topologie égo-centrée, routes longues

Thèse pour obtenir le grade de

Docteur en sciences de l'Université Pierre et Marie Curie

spécialité Informatique

Une approche pour l'estimation fiable des propriétés de la topologie d'Internet

Elie Rotenberg

Rapporteurs

Bertrand Jouve

Directeur de recherches, CNRS

Jean-Jacques Pansiot

Professeur émérite, Strasbourg

Examineurs

Clémence Magnien

Chargée de recherches, CNRS

Pascal Mérindol

Maître de conférences, Strasbourg

Philippe Owezarski

Directeur de recherches, CNRS

Directeur

Matthieu Latapy

Directeur de recherches, CNRS

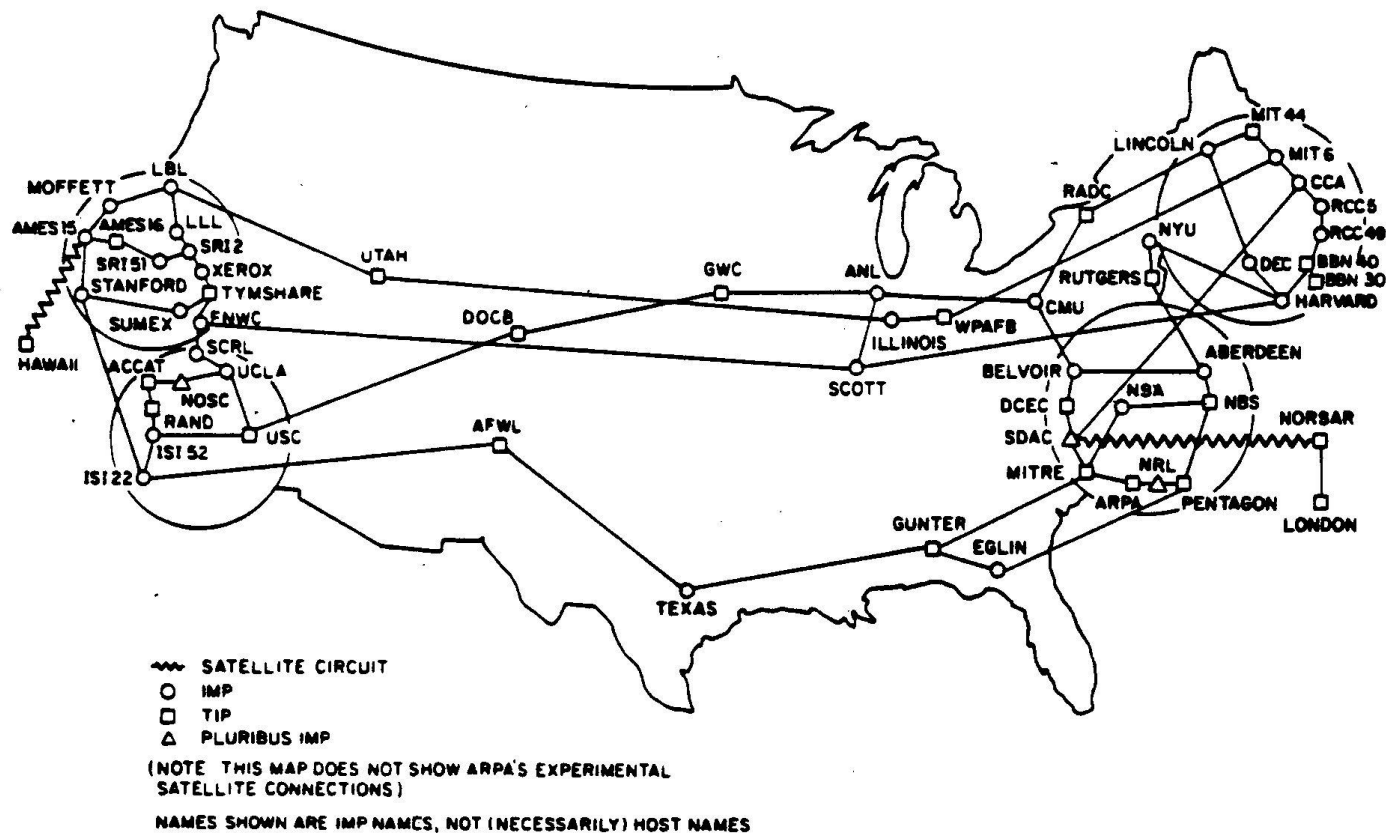
Co-directeur

Christophe Crespelle

Maître de conférences, UCBL

Approches historiques

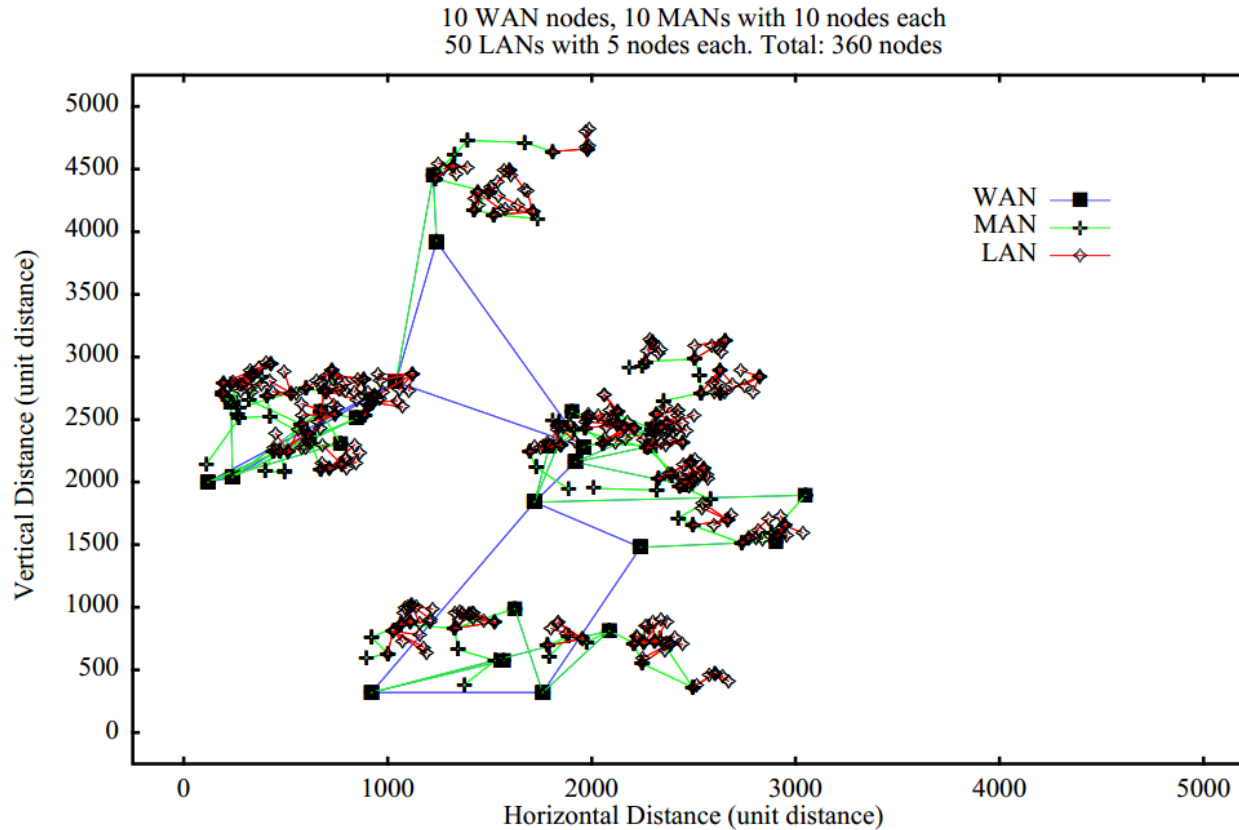
1) Cartes basées sur les déclarations des autorités administratives



Carte du réseau ARPANET, *BNN Technologies*, 1977

Approches historiques

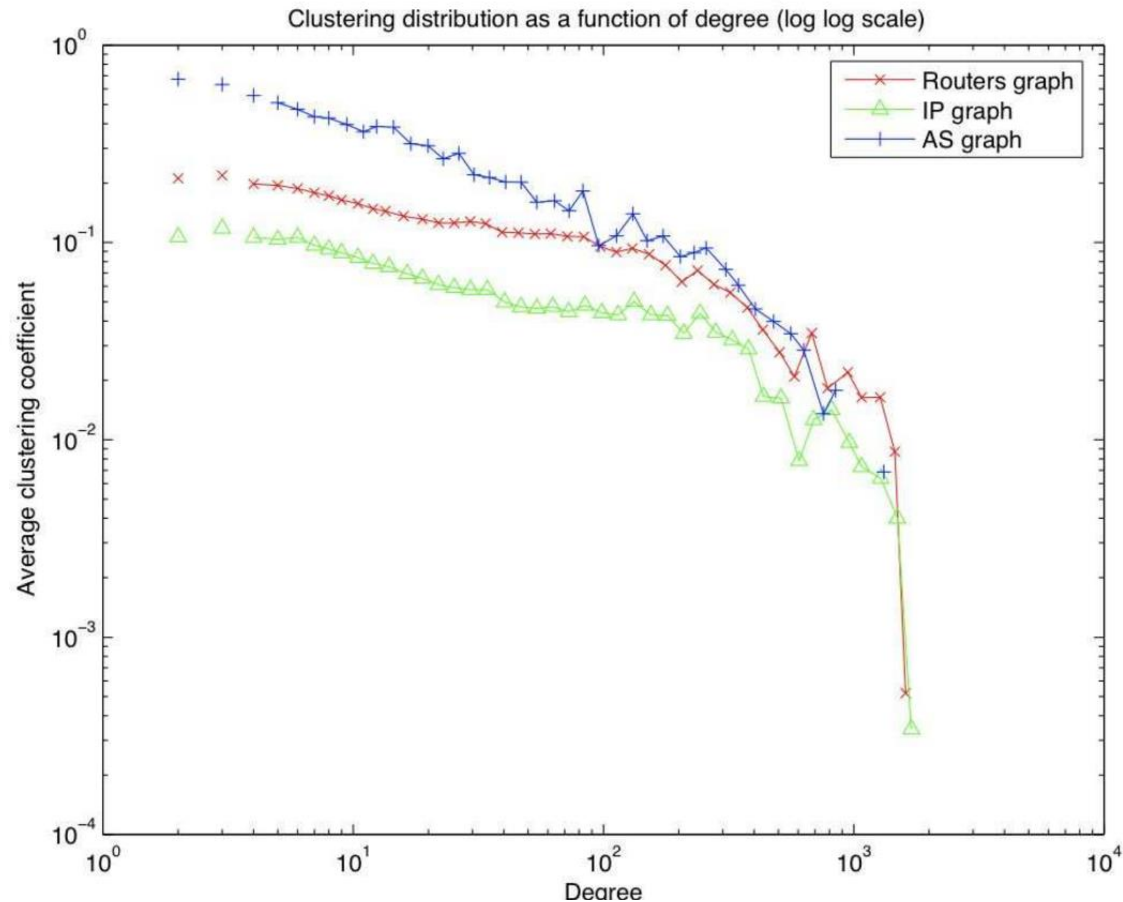
2) Graphes générés à partir d'une connaissance *a priori* des éléments du réseau



Réseau généré par simulation, Doar *et al.*, 1996

Approches historiques

3) Cartes déduites de mesures avec des outils de diagnostic (*traceroute*, *tracetree*...)



Propriété extraite d'une carte *traceroute*, DIMES, Shavitt *et al.*, 1999