

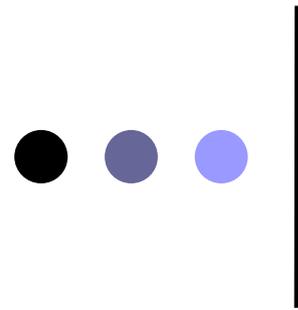
Estimation du Débit dans une Grille de Calcul

● ● ●
Carlos BARRIOS-HERNÁNDEZ, Yves
DENNEULIN et Michel RIVEILL

Laboratoire de Informatique de
Grenoble
INRIA
Equipe Mescal
Montbonnot-St Martin

Laboratoire de Informatique,
Signaux et Systèmes de Sophia
Antipolis
Equipe Rainbow
Sophia-Antipolis

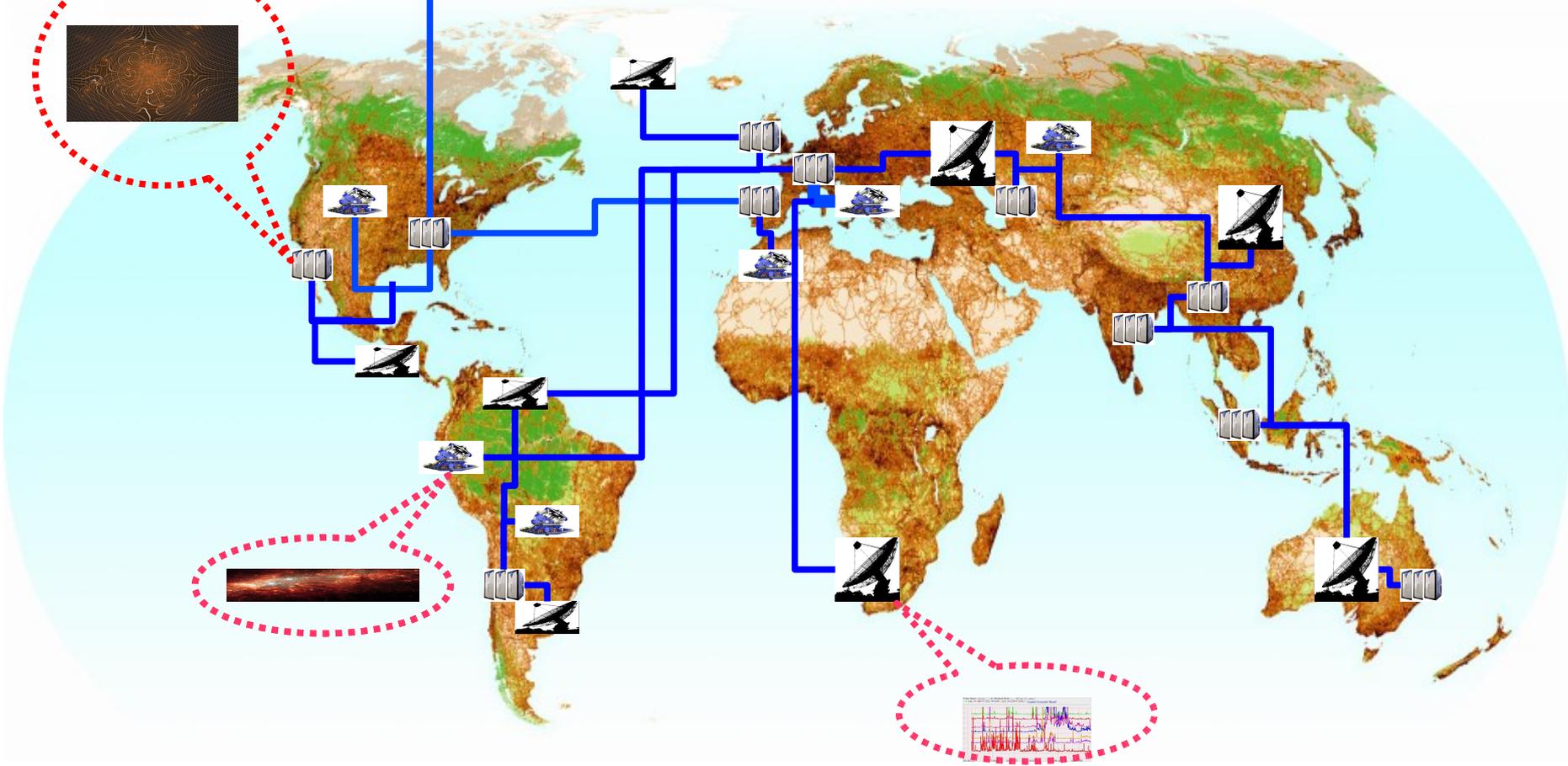
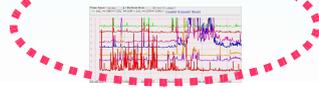
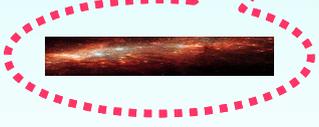
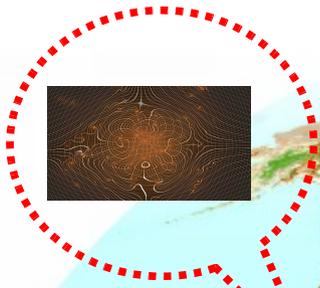




Plan

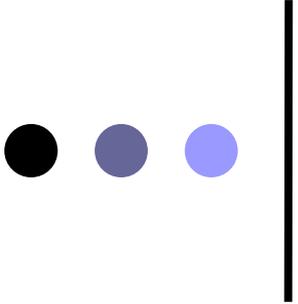
- Introduction
- Expérimentation
 - Description
- Estimation du Débit
- Analyse du Débit
 - Évolution du débit du lien
 - Évolution du débit du switch
 - Analyse
- Discussion
- Perspectives

Introduction



● ● ● | Introduction (2/3)

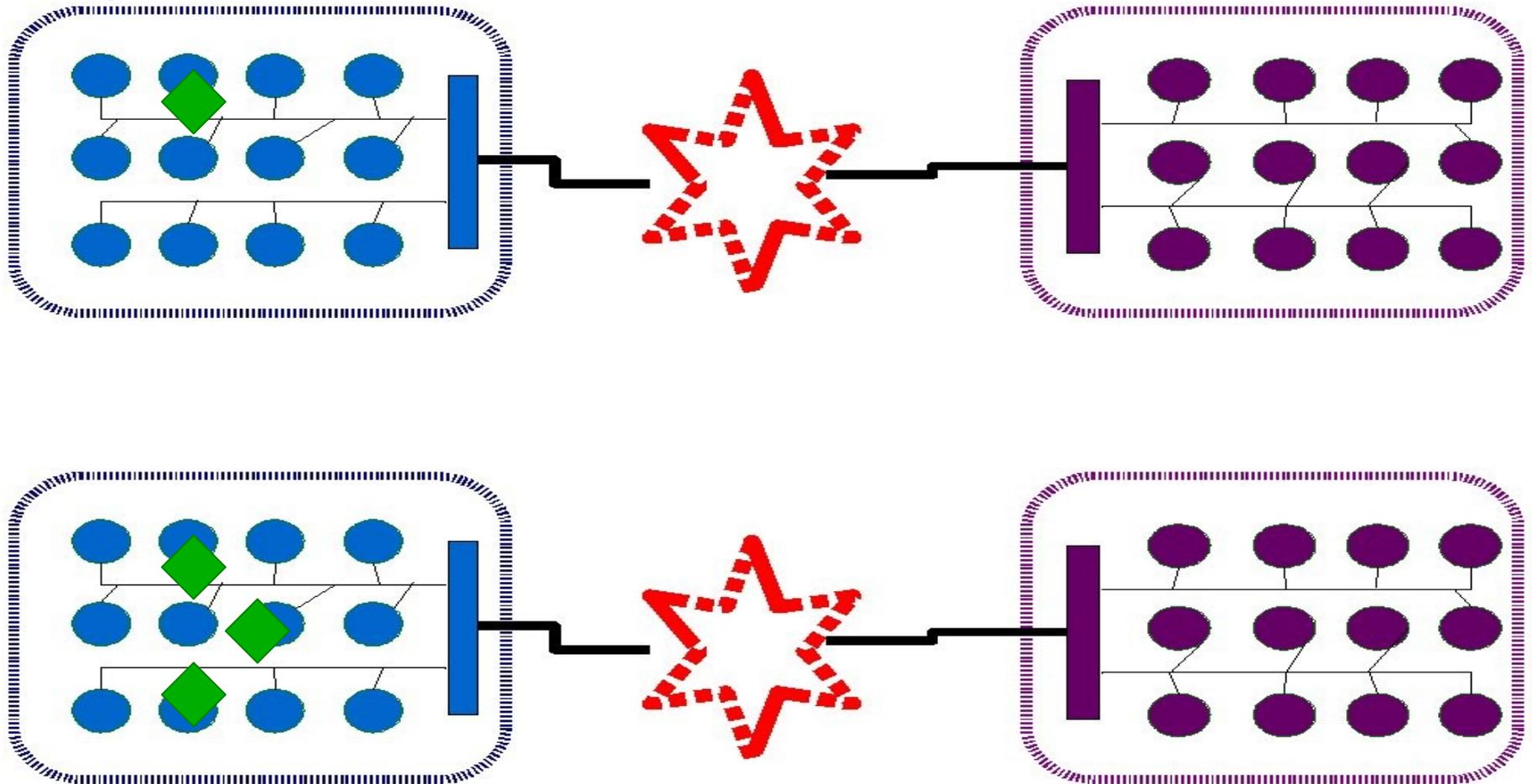
- Transfert des données sur la Grille de Calcul
 - Massive
 - Intensive
- Transfert Haut Débit sur Grilles de Calcul
 - Plateformes Distribuées
 - Différents Architectures (Hétérogénéité)
 - Changements de réseaux
 - Multiple Utilisation
 - Différents Utilisateurs
 - Différente Utilisation



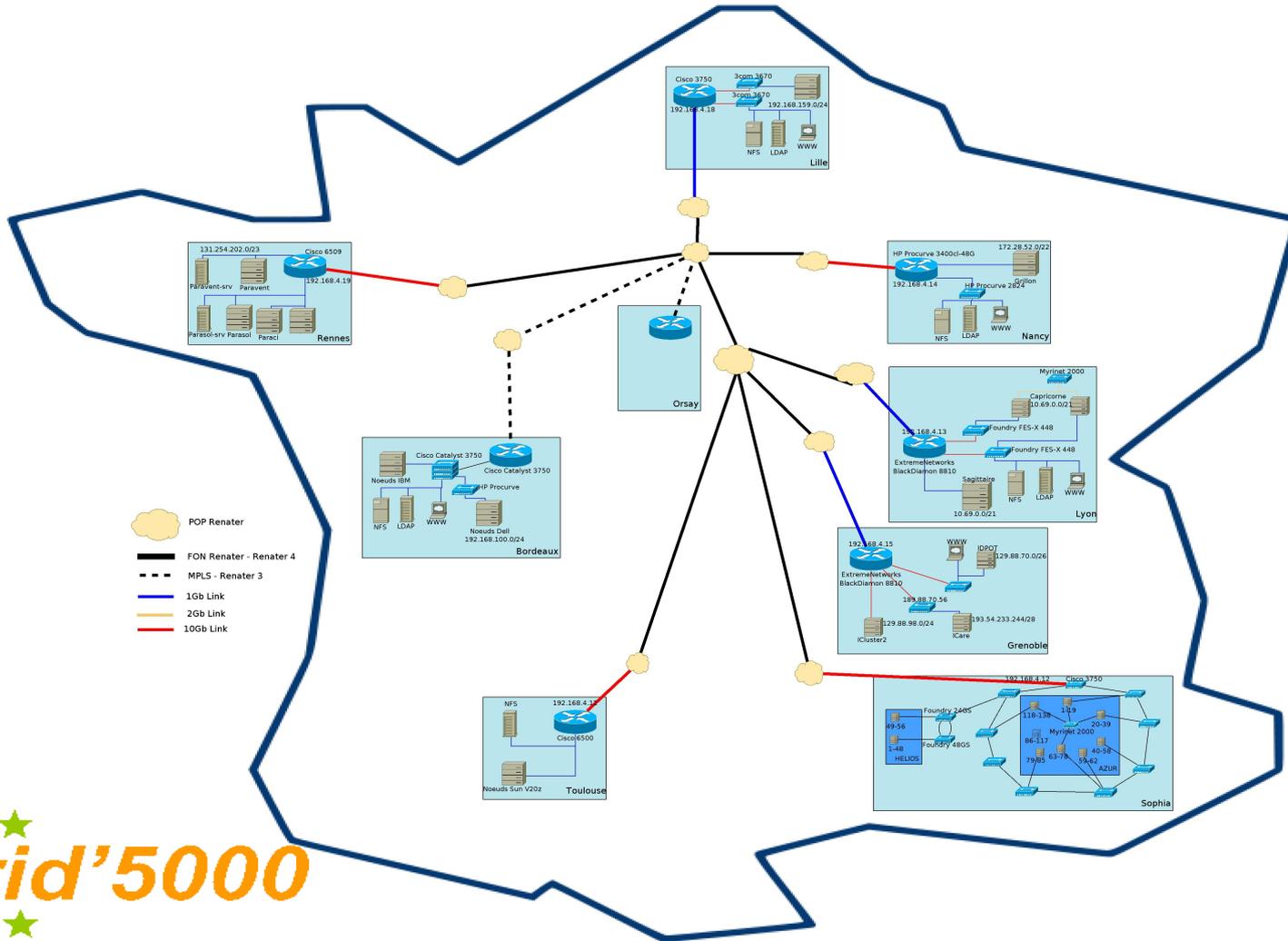
Introduction (3/3)

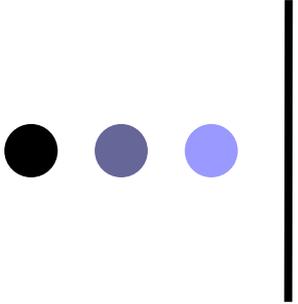
- Transfert des Données comme un Processus Critique
 - Caractérisation de transfert en fonction des éléments architecturaux
 - Observation depuis tests -dit réelles
 - Simulation
 - « Monitoring »
 - Senseurs

Expérimentation (1/3)



Experimentation (2/3)

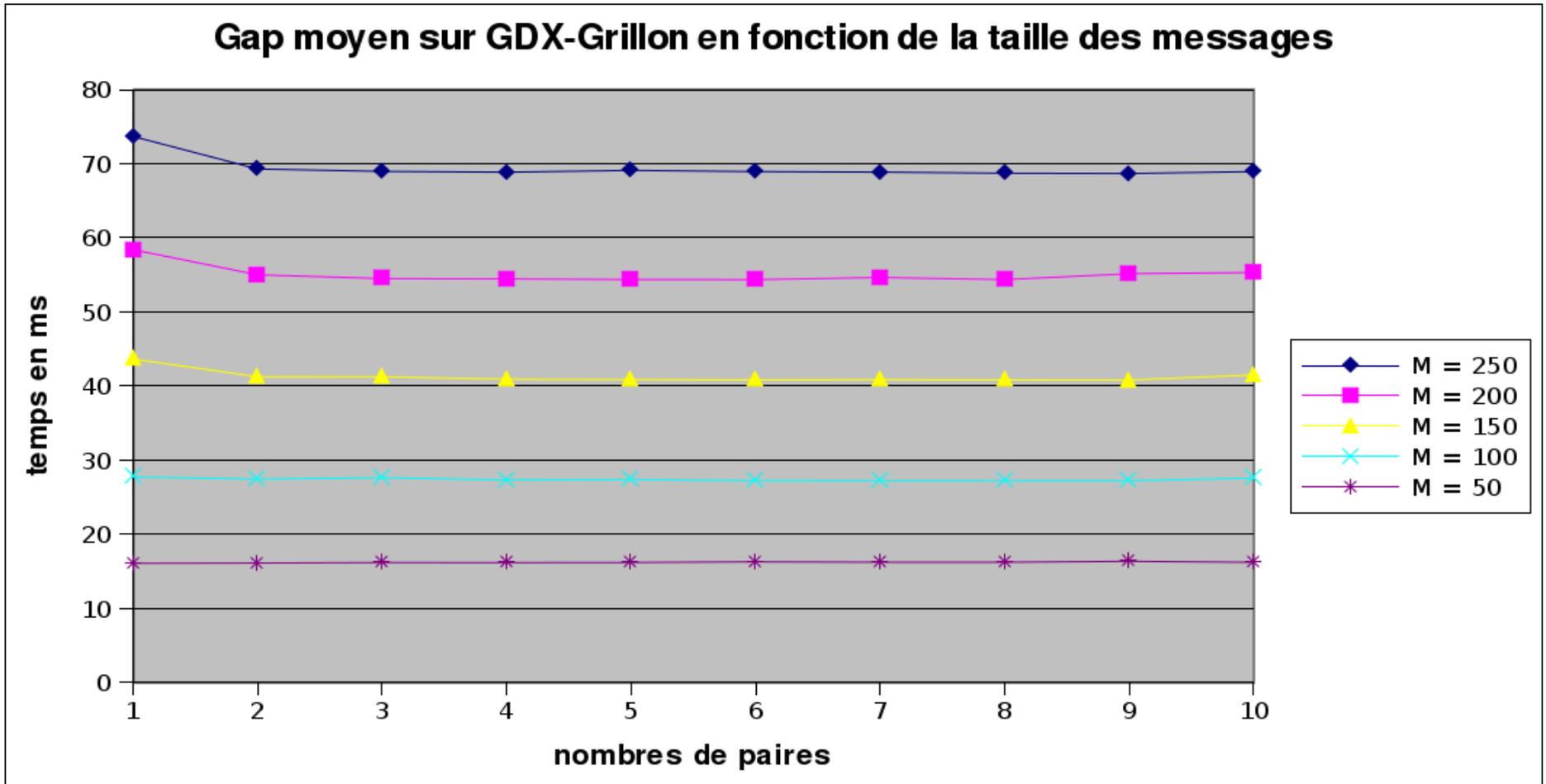




Expérimentation (3/3)

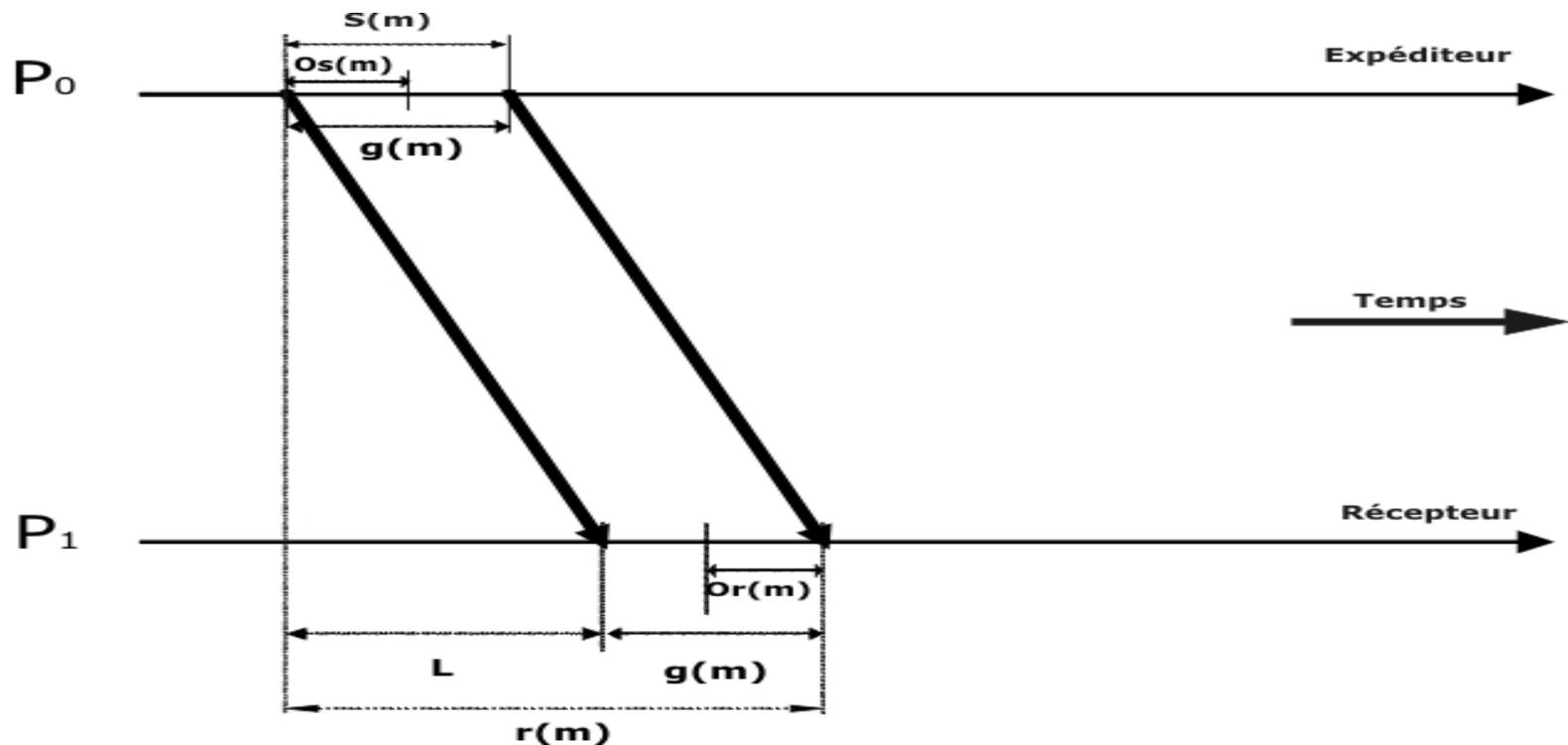
- Envoi et Réception des messages
 - Utilisation des outils pour générer trafic des données (MPIe Benchmark Tool et Netpipe)
 - Transfert Synchronisée des messages de taille fixe sur paires des noeuds (10Mb -> 250Mb par lien)
 - Incrément des noeuds (2 -> 20 ; 2 -> 50; 2 -> 100)
 - TCP
- Mesures
 - Espace de temps entre messages (gap)
 - Latence
 - Temps d'envoie et temps de réception (surcoût)

Mesures



Estimation de Débit (1/2)

- Modele de Base: LogP ... et compagnie [*]
 - $N(m)p = (L(m), O(m), g(m), P)$



[*] LogGP (Alexandrov et al, pLogP (Kielman et al)

● ● ● | Estimation de Débit (2/2)

- Depuis l'utilisation de modèle de base:

$$\mathbf{T}(m) = \mathbf{O}_s(m) + \mathbf{O}_r(m) + (k-1)*g(m)$$

- D'après les mesures:

$$\hat{g} = (k-1)*g$$
$$\hat{G} = \sum \hat{g} / m$$

- Le débit (depuis le modèle de base):

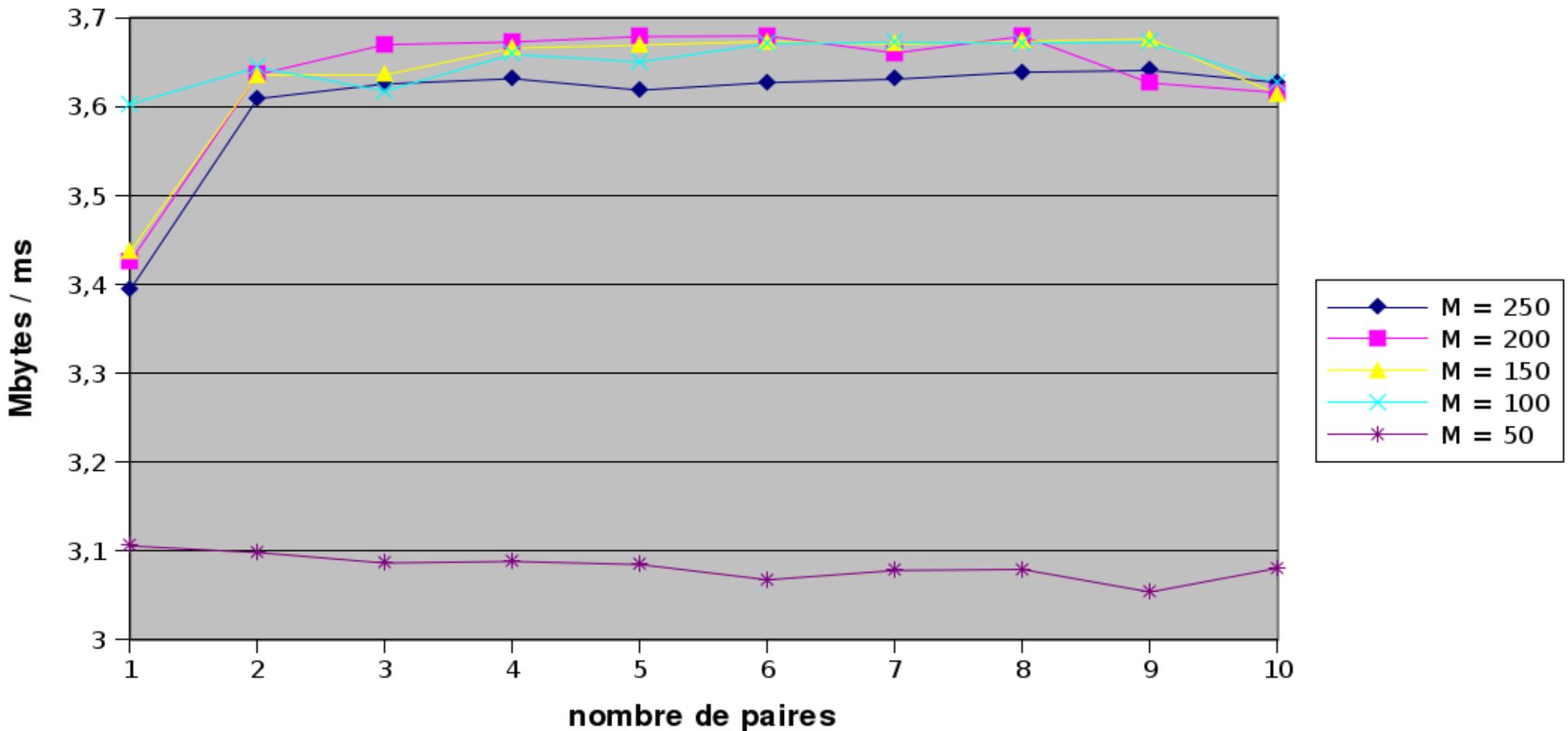
$$\mathbf{B} = m / (\mathbf{O}_s(m) + \mathbf{O}_r(m) + (k-1)*g(m))$$

- *Mais comme il y a une latence très petite, on fait $O_s(m), O_r(m) = 0$*

$$\mathbf{B} = m/\hat{g}(m)$$

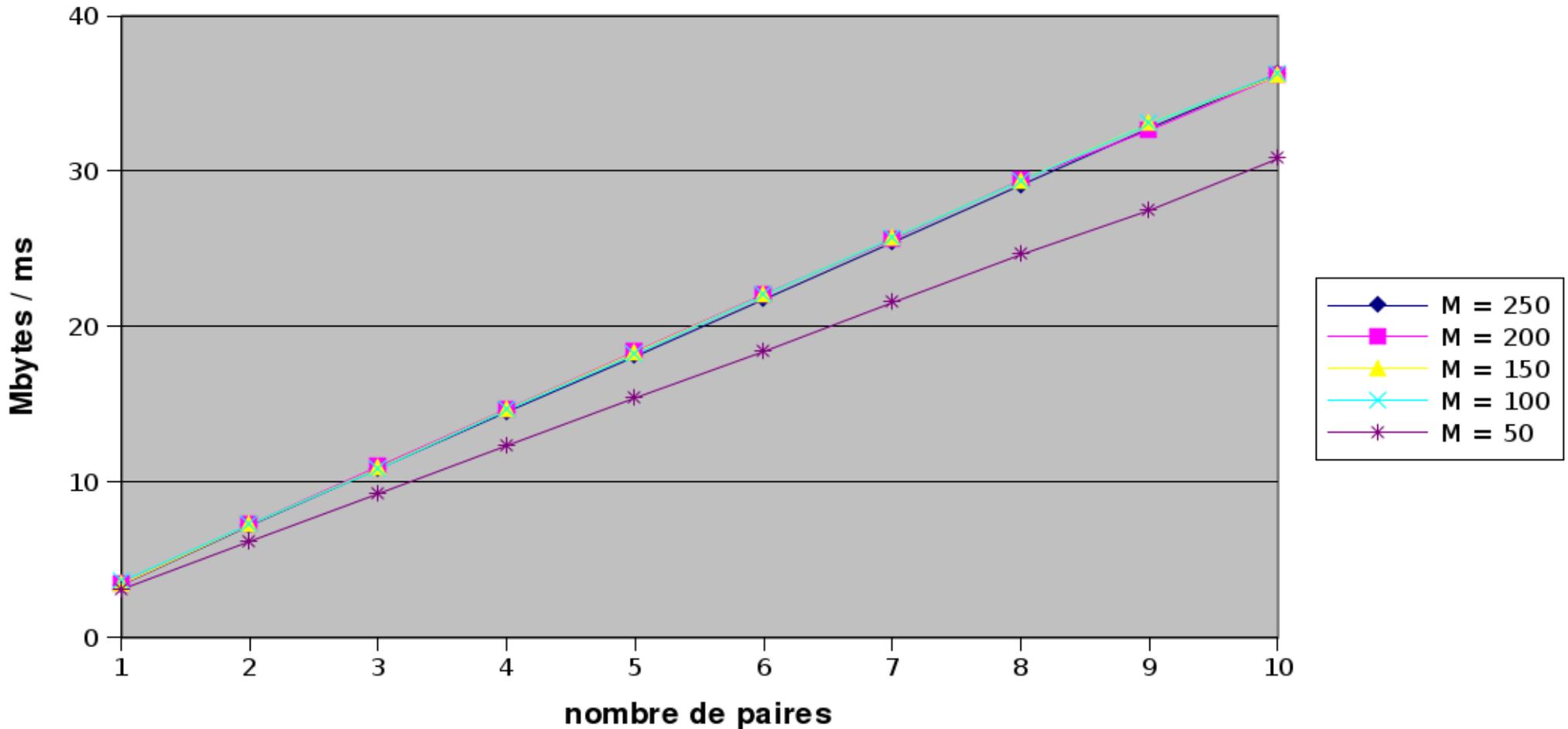
Évolution du Débit de lien

Débit d'un lien sur GDX-Grillon avec Gap moyen en fonction de la taille des messages



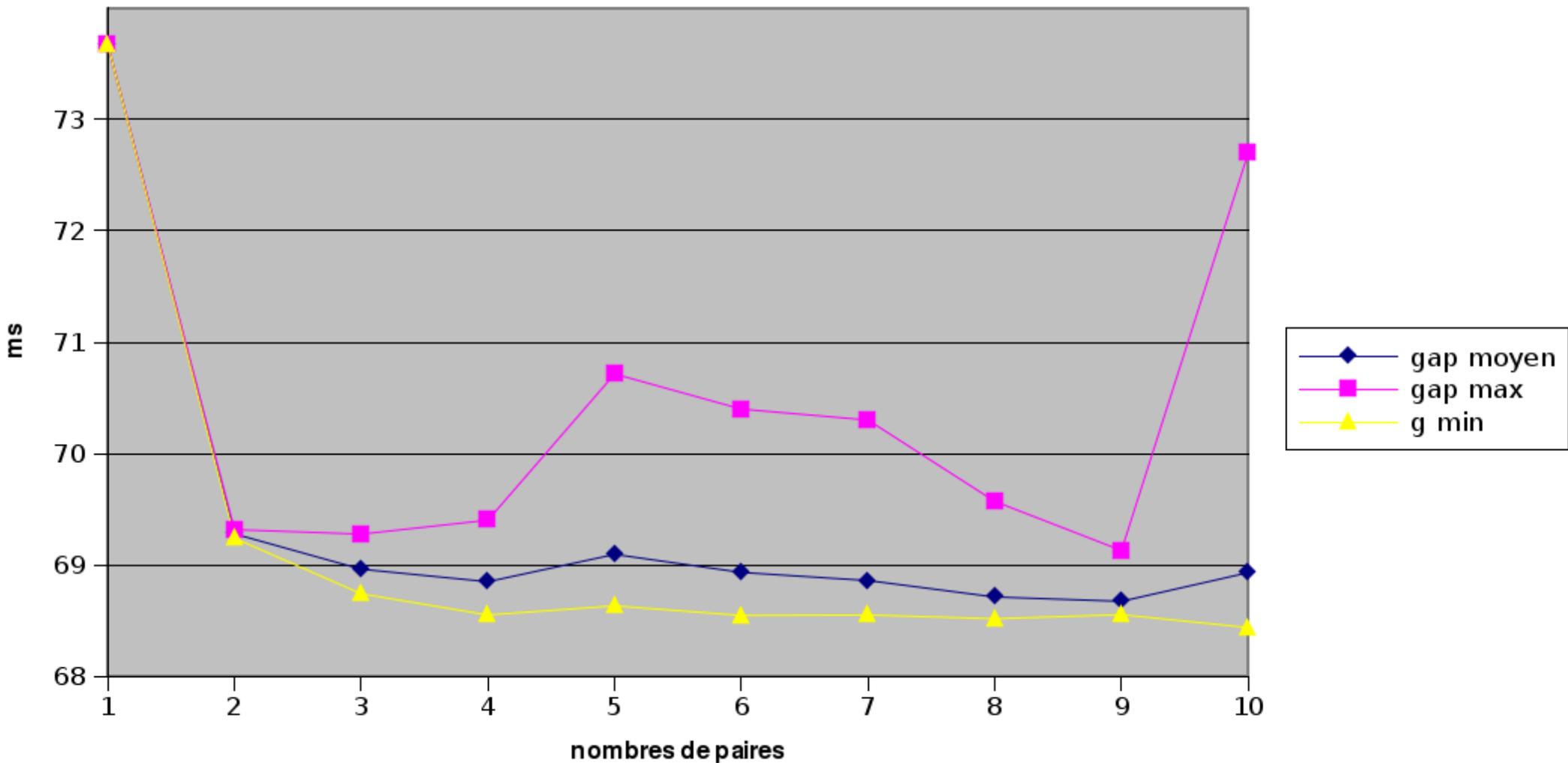
Évolution de Débit du Switch

Débit du switch sur GDX-Grillon avec Gap moyen en fonction de la taille des messages



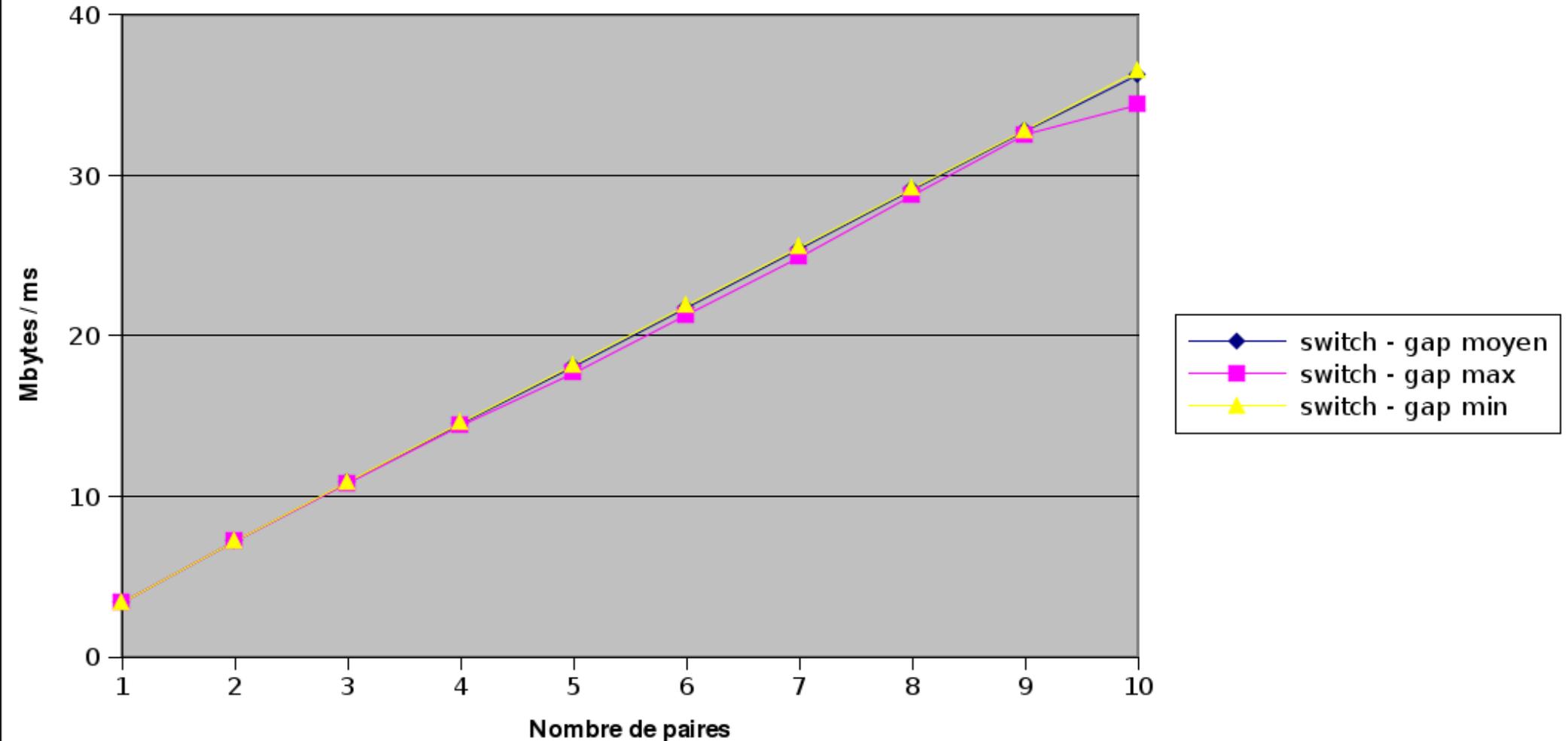
Analyse du Débit

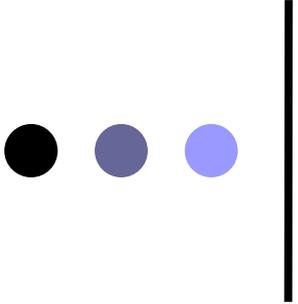
Relation de la valeur du Gap (moyen, min, max) pour m = 250Mb



Analyse du Débit

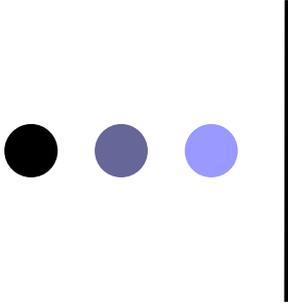
Relation de la valeur du gap sur le débit du switch pour M = 250





Discussion

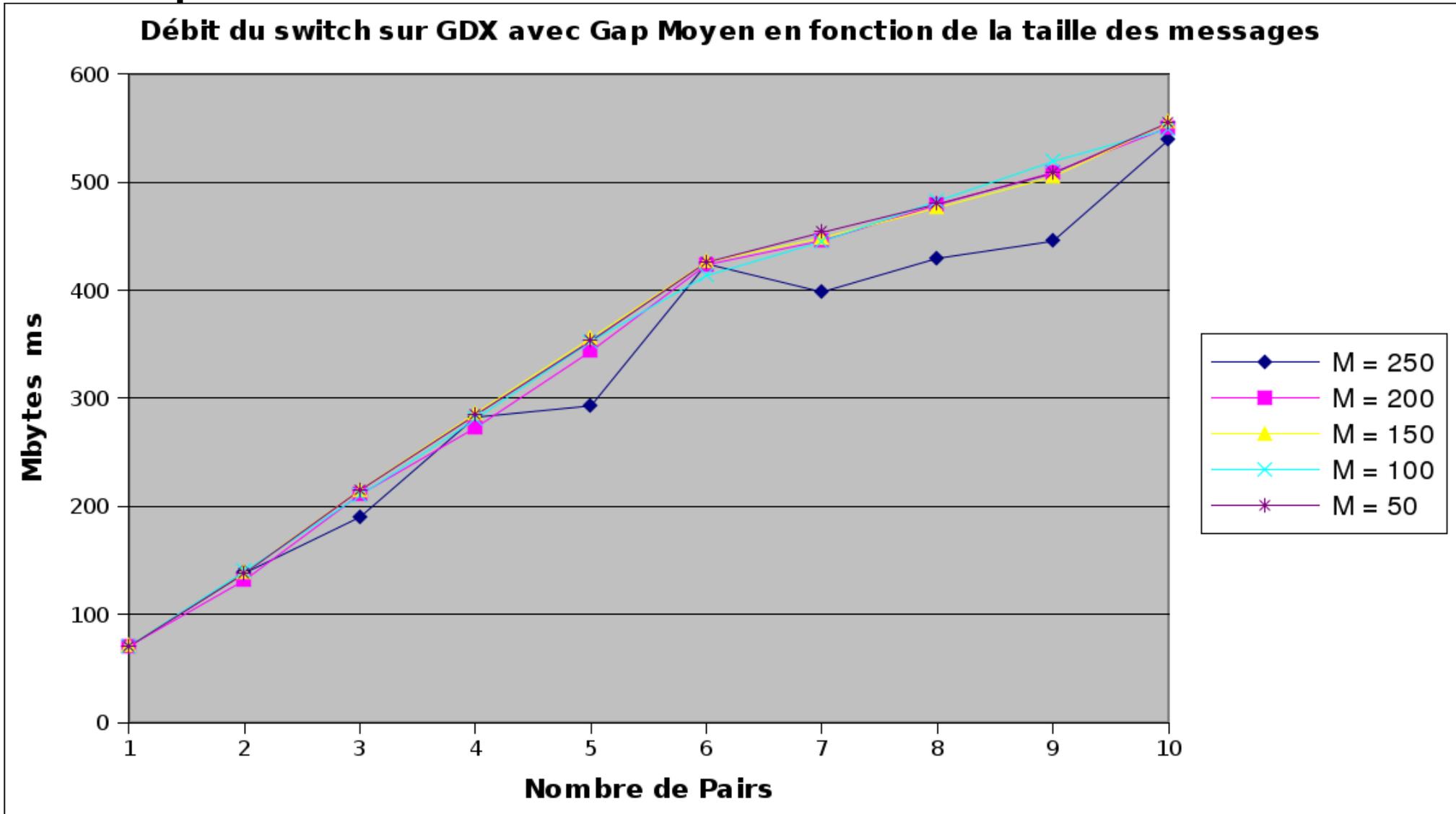
- Dans le cadre de « transfert haut débit »:
 - Analyse de Performance à partir de mesures:
 - Utilisation du « gap » comme mesure de référence.
 - Estimation du débit.
 - Ré-définition du concept de « haut débit »
 - Transfert intensif et Transfert massive ... qui représente une variation significative dans le coût de transfert.
 - Transfert qu'implique une important variation que conduisent à la saturation des ressources (dans ce cas, les liens entre noeuds)



Perspectives et travaux associés

- Re-définition et use du « surcoût » (Os, Or) pour les architectures hétérogènes .
- Utilisation de latence dans la cadre de « meilleur cas » et « pire cas » pour prédire la performance.
- Evaluation du « modèle analytique » vs « mesures » vs « simulation ».
- Développement des outils de « monitoring » de transfert pour des applications distribuées vers la grille de calcul.

Annexe: Débit des grappes



● ● ● | Débit des Grappes

Débit du switch sur Grillon avec Gap moyen en fonction de la taille des messages

