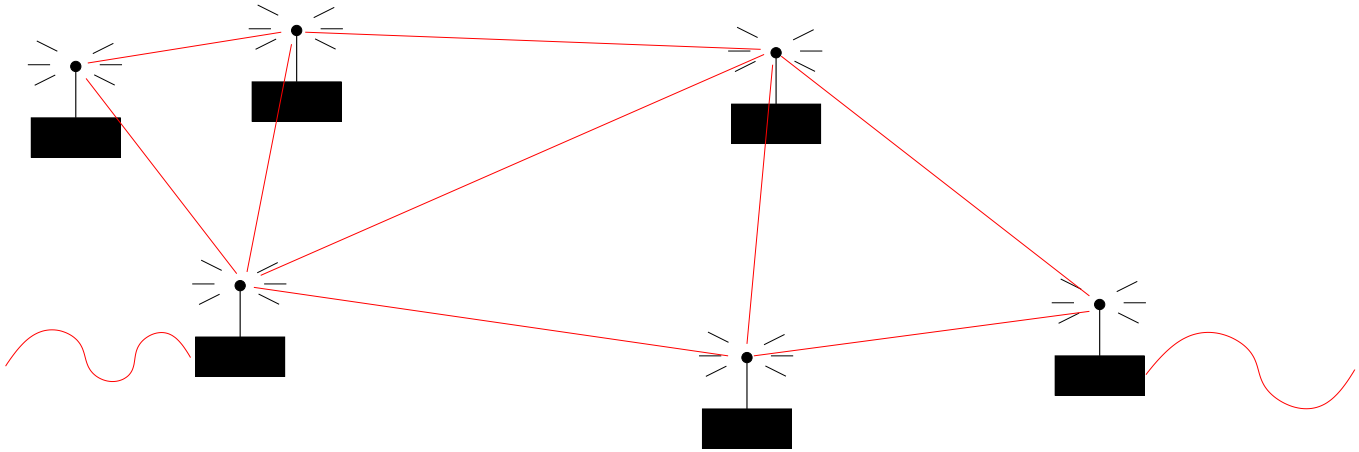
Babel
Un protocole de routage pour réseaux hybrides

Juliusz Chroboczek
Laboratoire PPS
Université Paris-Diderot (Paris 7)

16 janvier 2012

Réseaux maillés sans fil

Un réseau maillé est un réseau sans structure *a priori*, où tout chemin est utilisable : pas de distinction hôte/routeur (ou STA/AP).



Intérêt depuis les années 2000 : routeurs *wifi* peu chers.

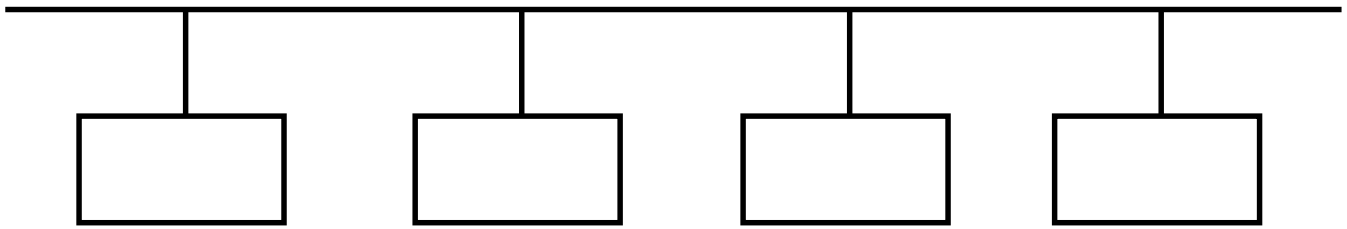
Réseaux hybrides

En routage IP (couche 3), les réseaux maillés sans-fil purs n'existent pas :

- passerelles Internet (ou vers d'autres préfixes) ;
- préfixe injectés à des points multiples ;
- liaisions filaires (JBOLs).

Les protocoles pour réseaux maillés purs ont donc un intérêt surtout théorique.

Réseaux filaires

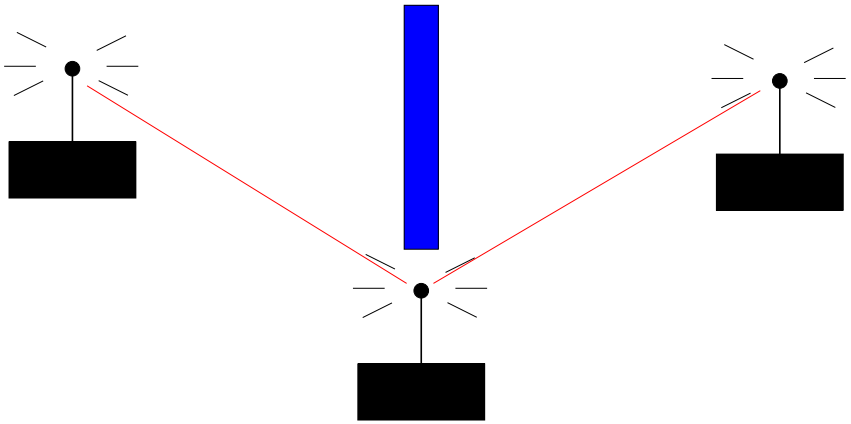


Les protocoles de routage habituels sont conçus pour des *liens filaires à diffusion* :

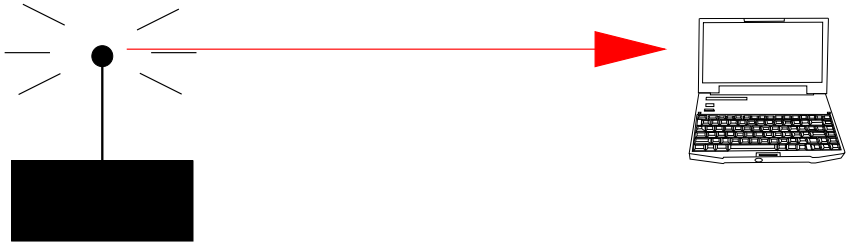
- communication symétrique
- communication transitive
- communication bimodale

La radio n’est pas un Ethernet

- la communication n’est pas transitive



- la communication n’est pas symétrique



- la communication n’est pas bimodale (liens marginaux).

Réseaux maillés communeautaires

Réseaux maillés communeautaires :

- urbains ;
- ruraux.

Réseaux urbains :

- nombre raisonnable de nœuds (100 à 1000) ;
- liens sans fil (wifi) et liens filaires (cuivre et fibre) ;
- liens marginaux (> 50% de pertes) productifs ;
- chemins longs (7–8 sauts) productifs.

Ce sont des réseaux longs et fins.

Parenthèse : problèmes d'OLSR

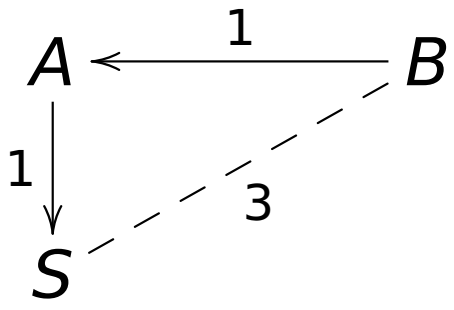
L'expérience montre que OLSR marche mal dans les réseaux hybrides épars :

- OLSR fait du routage « shortest hop »

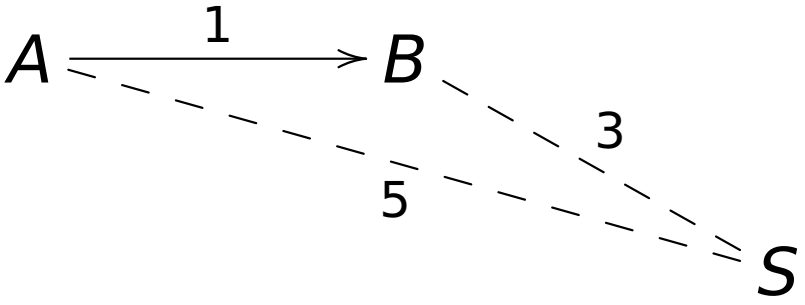
  ⟶ pas de préférence pour les liens filaires ;
  ⟶ selon les réglages,
    - soit le réseau perd sa connexité,
    - soit OLSR choisit des liens marginaux.
- les bases de données de lien se désynchronisent

  ⟶ boucles de routage transitoires.

Problèmes d'OLSR (2)



A route directe vers *S*
B passe par *A*



A bascule vers *B* que *B* bascule

Cette situation persiste jusqu'à ce que la nouvelle topologie soit inondée jusqu'à *B*.

Dans Babel, *A* délaie le changement de route jusqu'à ce que *B* ait basculé.

Parenthèse : problèmes d'OLSR (3)

OLSR dans le réseau communautaire *Funkfeuer* : OLSR modifié pour prendre en compte la qualité de lien : OLSR-ETX.

Pour éviter la désynchronisation :

- MPR redundancy = 9,
- Hello interval = 5 s,
- TC interval = 2 s.

(Non, ce n'est pas une typo.)

Pour éviter l'inondation trop fréquente, algorithme fish-eye.

Babel : résumé

Babel est un protocole à vecteur de distance (comme RIP), mais avec un certain nombre d'améliorations :

- détection explicite des voisins avec calcul de qualité de lien ;
- flexible : métriques arbitraires ;
- mises à jour partielles et irrégulières *(triggered updates)* ;
- mécanisme de prévention de boucles : pas de comptage à l'infini.

À la différence des protocoles pour réseaux maillés purs, Babel supporte le routage par préfixe et les passerelles multiples.

Babel : détection des voisins

Un nœud Babel A diffuse plus ou moins périodiquement

$$\text{Hello}(\text{seqno}, \text{interval})$$

un paquet Hello équipé d'un numéro de séquence et de l'intervalle avant le prochain hello.
Le taux de perte de paquets est déterminé de façon robuste même si les intervalles sont différents ou variables.
Pour chaque voisin B, un nœud A envoie

$$\text{IHU}(B, \text{taux})$$

un paquet I Heard You contenant le taux de succès du transfert de paquets Hello émis par B.

Babel : qualité de lien

Pour chaque voisin B, un nœud Babel A connaît

- le taux de succès α de paquets de A vers B (IHU),
- le taux de succès β de paquets de B vers A (Hello).

$$A \underset{\beta}{\overset{\alpha}{\rightleftarrows}} B$$

L'implémentation actuelle utilise la métrique ETX pour les liens sans-fil :

$$c_{AB} = \frac{1}{\alpha\beta}$$

Sur les liens filaires, Babel utilise la métrique 2-3 :

$$\begin{aligned} c_{AB} &= 1 && \text{si 2 des derniers 3 hellos ont été reçus} \\ c_{AB} &= \infty && \text{sinon.} \end{aligned}$$

Rappel : Bellman-Ford distribué

Initialement,

$$\mathrm{d}(S) = 0 \qquad \mathrm{d}(X) = \infty$$

Assez souvent, Y diffuse $\mathrm{d}(Y)$ à ses voisins.
Lorsque X reçoit $\mathrm{d}(Y)$,

- si $\mathrm{nh}(X) = Y$,

$$\mathrm{d}(X) := c_{XY} + \mathrm{d}(Y)$$

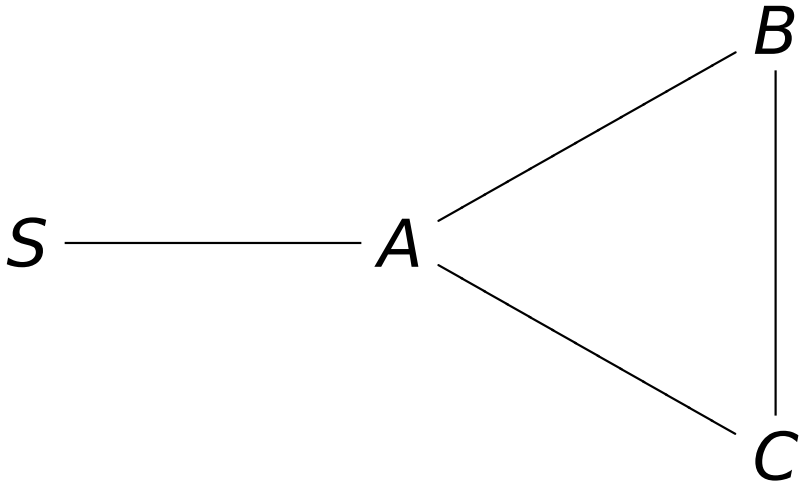
- si $c_{XY} + \mathrm{d}(Y) < \mathrm{d}(X)$

$$\mathrm{d}(X) := c_{XY} + \mathrm{d}(Y) \qquad \mathrm{nh}(X) := Y$$

Timeout : si $nh(X) = Y$, et Y cesse de diffuser,

$$\mathrm{d}(X) := \infty \qquad \mathrm{nh}(X) := \perp$$

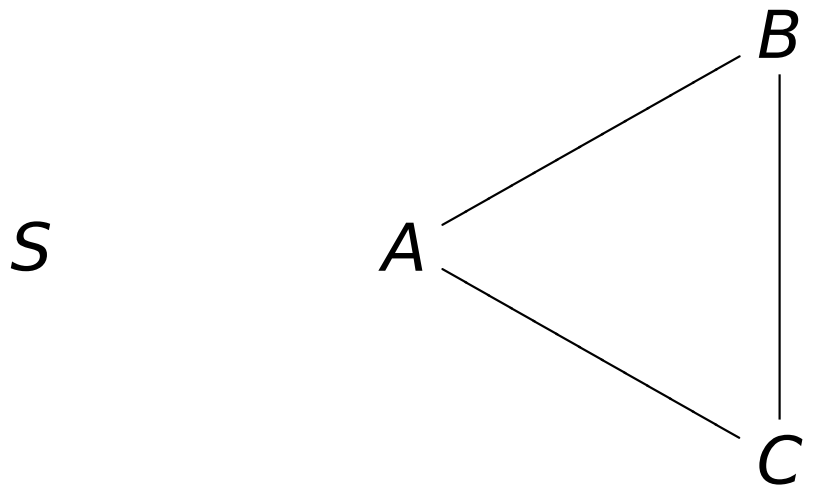
BF distribué : convergence



S	0	0	0	0
A	∞	1, nh $= S$	1, nh $= S$	1, nh $= S$
B	∞	∞	2, nh $= A$	2, nh $= A$
C	∞	∞	2, nh $= A$	2, nh $= A$

Converge en $O(\Delta)$.

BF distribué : comptage à l'infini



A	1, nh $= S$	3, nh $= B$	3, nh $= B$	3, nh $= B$
B	2, nh $= A$	2, nh $= A$	3, nh $= C$	3, nh $= C$
C	2, nh $= A$	2, nh $= A$	2, nh $= A$	4, nh $= A$

Converge en $O(\infty)$. (Dans RIP, $\infty = 16$.)

Avant convergence, on a une boucle de routage.

« Good news travel fast, bad news travel forever. »

BF : condition de faisabilité

BF est robuste, on peut refuser certaines mises à jour si elles risquent de causer une boucle.
Lorsque X reçoit $(\mathrm{d}(Y), f)$,

- si $\mathrm{nh}(X) = Y$ et $\mathrm{feasible}(Y, \mathrm{d}(Y), f)$

$$\mathrm{d}(X) := c_{XY} + \mathrm{d}(Y)$$

- si $c_{XY} + \mathrm{d}(Y) < \mathrm{d}(X)$ et $\mathrm{feasible}(Y, \mathrm{d}(Y), f)$

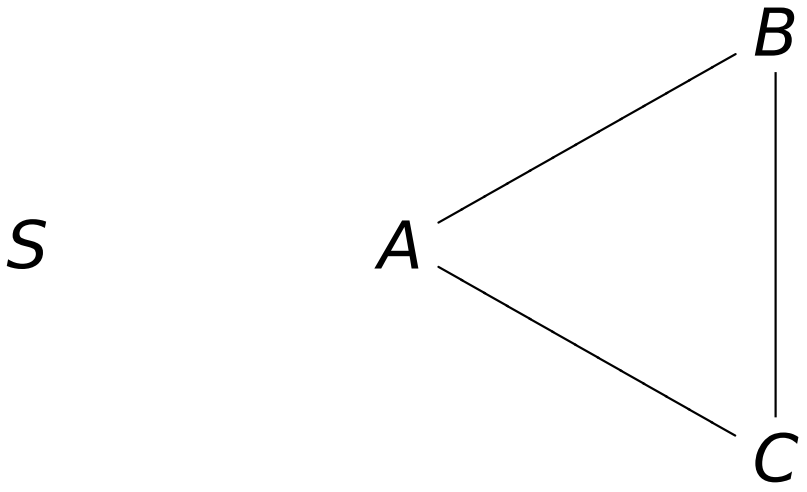
$$\begin{aligned} \mathrm{d}(X) &:= c_{XY} + \mathrm{d}(Y) \\ \mathrm{nh}(X) &:= Y \end{aligned}$$

où feasible est une fonction garantissant l'absence de boucles.

Conditions de faisabilité

BGP, ***Path Vector*** :
f est le chemin complet ;
feasible(f) $=$ self $\notin f$.
DSDV, **AODV** :
$feasible(d) \equiv c + d \leq \mathrm{d}(\mathrm{self})$
Invariants : $\mathrm{d}(X) \searrow$ et si $A \leftarrow B$ alors $\mathrm{d}(A) < \mathrm{d}(B)$.
EIGRP/DUAL, **Babel** :
On maintient $\mathrm{rd}(X) = \min_{t \leq \mathrm{now}} \mathrm{d}(X, t)$.
$feasible(d) \equiv d < \mathrm{rd}(\mathrm{self})$
Invariants : $\mathrm{rd}(X) \searrow$ et si $A \leftarrow B$ alors $\mathrm{rd}(A) < \mathrm{rd}(B)$.

Faisabilité : exemple

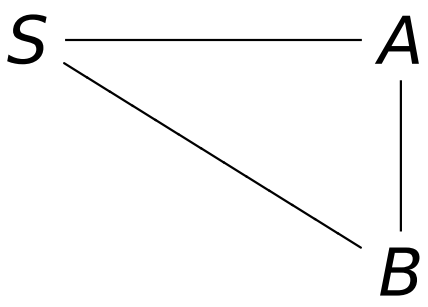


A	1, rd $= 1$	∞, rd $= 1$	∞, rd $= 1$	∞, rd $= 1$
B	2, rd $= 2$	2, rd $= 2$	∞, rd $= 2$	∞, rd $= 2$
C	2, rd $= 2$	2, rd $= 2$	∞, rd $= 2$	∞, rd $= 2$

Convergence en $O(\Delta)$.

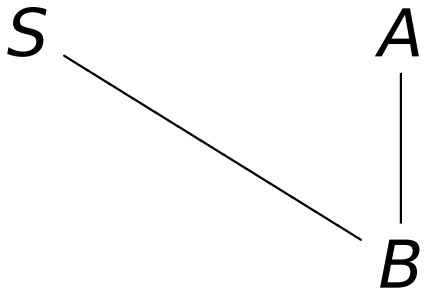
Faisabilité : famine

Les conditions de faisabilité (1) et (2) causent une famine.



$\mathrm{d}(A) = 1, \mathrm{rd}(A) = 1$

$\mathrm{d}(B) = 1, \mathrm{rd}(B) = 1$



$\mathrm{rd}(A) = 1$

$\mathrm{d}(B) = 1$

La seule route disponible n'est pas faisable.

Résoudre la famine

Idée : lorsqu'aucune route n'est faisable, on reboote tout le réseau.

DUAL/EIGRP effectue une synchronisation globale (des routes vers S).

DSDV, AODV et Babel utilisent des routes séquencées.

Résoudra la famine : routes séquencées

Les annonces de routes sont équipées d'un numéro de séquence :

$$(s, \mathrm{d}(B))$$

où $s \in \mathbf{N}$ est incrémenté par la source :

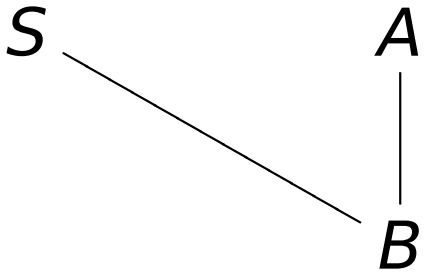
$$\begin{aligned} \mathrm{d}(S) &= (s, 0) \qquad (s \nearrow) \\ c + (s, m) &= (s, c + m) \end{aligned}$$

On définit

$$(s, m) \leq (s', m') \quad \text{lorsque} \quad \begin{array}{l} s > s' \text{ ou} \\ s = s' \text{ et } m \leq m' \end{array}$$

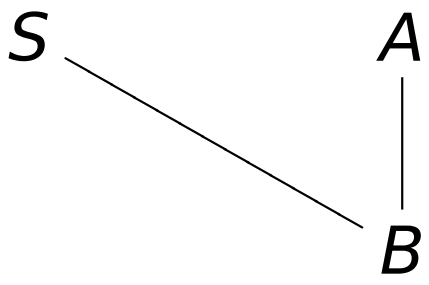
$$\mathrm{feasible}(s, m) \equiv (s, m) < \mathrm{rd}\,.$$

Routes séquencées : exemple



S	$(1,0)$	$(2,0)$	$(2,0)$
A	$\infty, \mathrm{rd} = (1,1)$	$\infty, \mathrm{rd} = (1,1)$	$(2,2), \mathrm{rd} = (2,2)$
B	$(1,1), \mathrm{rd} = (1,1)$	$(2,1), \mathrm{rd} = (2,1)$	$(2,1), \mathrm{rd} = (2,1)$

Famine temporaire

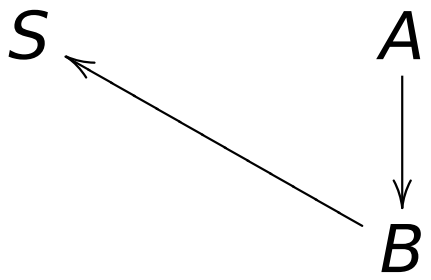


$$\begin{aligned} \mathrm{d}(S) &= (1,0) \\ \mathrm{d}(B) &= (1,1) \\ \mathrm{d}(A) &= \infty \qquad \mathrm{rd}(A) = (1,1) \end{aligned}$$

A doit attendre que S génère un nouveau seqno, et que celui-ci se propage.
Dans Babel, cette famine temporaire est signalisée explicitement par A ($\neq$ DSDV).

Résolution de la famine temporaire

Lorsqu'un nœud Babel souffre d'une famine temporaire
(routes disponibles mais non-faisables)
il envoie une requête explicite pour un nouveau seqno.



À la différence d'AODV, cette requête n'est pas multidiffusée, ce qui évite de faire une recherche par cercles concentriques — un simple *hop count* suffit.

Passerelles multiples

En général, on veut qu'il soit possible d'avoir plusieurs nœuds Babel qui annoncent le même préfixe
sans besoin de synchroniser les numéros de séquence.
Babel distingue entre source et destination.
Une annonce Babel contient un triplet

$$(s, d, id)$$

où *id* identifie de façon unique le routeur introduisant la route. Les distances de référence sont indexées par source et destination.

Passerelles multiples : boucles

En présence de passerelles multiples,
Babel ne garantit plus l'absence de boucles.

$$S_1 \text{———} A \text{———} B \text{———} S_2$$

$$\mathrm{d}(A) = (17, 1) \qquad \mathrm{d}(B) = (43, 1)$$
$$\mathrm{rd}(A, S_1) = (17, 1) \qquad \mathrm{rd}(B, S_2) = (43, 1)$$

On garantit par contre qu'une boucle disparaît en $O(n)$, où n est la taille de la boucle.

Routes non-disjointes

Une boucle de routage peut aussi provenir de la collision de deux routes vers des préfixes distincts.

$$0.0.0.0/0 \text{ ——— } A \text{ ——— } B \text{ ——— } C$$

Le lien entre B et C disparaît :

$$0.0.0.0/0 \text{ ——— } A \text{ ——— } B \qquad C$$

Si B bascule sur A, il y a une boucle de routage temporaire. Ce problème ne peut avoir lieu qu'après une rétraction.
Babel respecte un hold time après une rétraction de route. C'est la partie la moins satisfaisante de Babel.

Variantes non-standard de Babel

Babel est un protocole robuste : la correction de Babel (convergence, absence de boucles) ne dépend pas de la métrique ou de la politique de sélection de routes.

Conditions de correction :

$$m \preceq c \cdot m \qquad \text{(monotonicité stricte)}$$

$$\text{si } m \preceq m' \text{ alors } c \cdot m \preceq c \cdot m' \qquad \text{(isotonicité)}$$

Parenthèse : interférences internes

Lorsqu'une route utilise plusieurs fois la même fréquence, on a des interférences entre les sauts :

$$A \text{——} B \text{——} C$$

Le débit décroit de façon exponentielle avec le nombre de sauts :

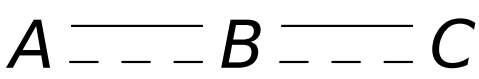
$A \text{——} B$	100 %
$A \text{——} B \text{——} C$	50 %
$A \text{——} B \text{——} C \text{——} D$	25 %
...	
7 hops	< 10%

(Au-delà de 7 sauts environ, le débit devient constant.)

Sensible à la fréquence radio

Babel-Z

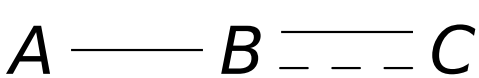
Solution : utiliser plusieurs fréquences.



La variante dite *Babel-Z* utilise une métrique sensible à la fréquence radio :

$$c \oplus m = c + m \quad \text{si } r \text{ et } l \text{ interfèrent}$$
$$c \oplus m = \frac{1}{2}c + m \quad \text{sinon}$$

Cette métrique n'est pas isotone :



Sensible au délai

Une nouvelle variante (pas encore nommée) implémente une métrique sensible au délai. Utile en présence de tunnels.

En général, une telle métrique mène à des oscillations non bornées. Avec les bon paramètres oscillations bornées.

Conclusions

Babel est

- décrit en détail (RFC 6126) ;
- implémenté (*babeld*, *babel-quagga*) ;
- partiellement démontré ;
- Babel-Z est implémenté et disponible.

Travail en cours :

- finir de rédiger la preuve de correction ;
- intégrer *babel-quagga* à *quagga* (avec Matthieu Boutier) ;
- finir les mesures expérimentales sur *Babel-Z* (avec Gabriel Kerneis) ;
- finir la variante sensible au délai.