

Polyèdres et reconfiguration dynamique d'autocommutateurs répartis

R. SIRDEY (Nortel, Heudiasyc)

Réunion du Gotha, 12 mai 2006

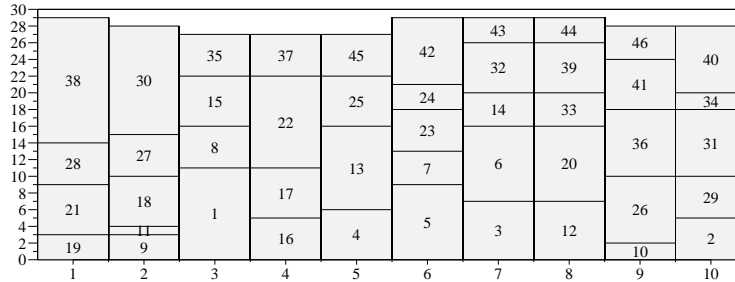
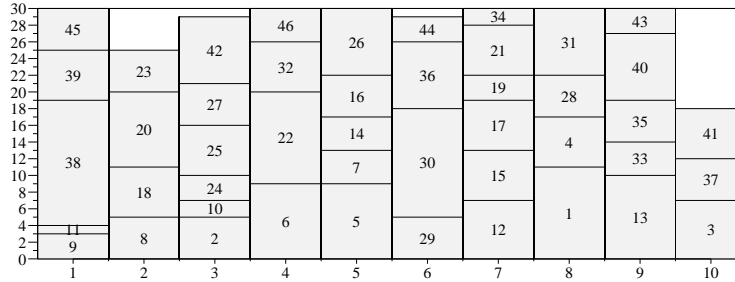
Problématique	3
Le polytope des sous-tournois sans circuit	13
Le polytope des programmes de déplacement	23
Recherche arborescente polyédrique	28
Perspectives : reroutage dans les réseaux MPLS	34

Problématique [15]

- Soit un autocommutateur composé de processeurs sur lesquels tournent des processus.
- Un état du système consiste en une distribution d'un sous-ensemble des processus sur les processeurs qui respecte des contraintes de capacité (erlangs, ports, ...).
- On cherche à faire passer le système d'un état *initial* à un état *terminal*, en déplaçant des processus et ce sans violation des contraintes de capacité¹.

¹D'autres variantes sont envisageables.

Exemple



- Un processus peut être :
 - Déplacé d'un processeur vers un autre sans impact sur le service rendu² (*migration*).
 - Arrêté, puis ultérieurement redémarré sur un autre processeur (*interruption*).
- L'interruption permet de résoudre les blocages (les interruptions sont réalisées au début), l'objectif étant d'y avoir le moins recours possible.

²Durant une migration, le processus consomme des ressources sur les processeurs source et cible du déplacement.

- Coffman et al. [6, 7] (1984) : transferts de fichiers.
- Carlier [4, 5] (1984) : paiements de dettes.
- Hall et al. [10], Anderson et al. [2] (2001) : reconfiguration de baies de stockage.
- Aggarwal et al. [1] (2003) : rééquilibrage de charge.
- Józsa et Makai [11] (2003) : reroutage dans les réseaux MPLS³.

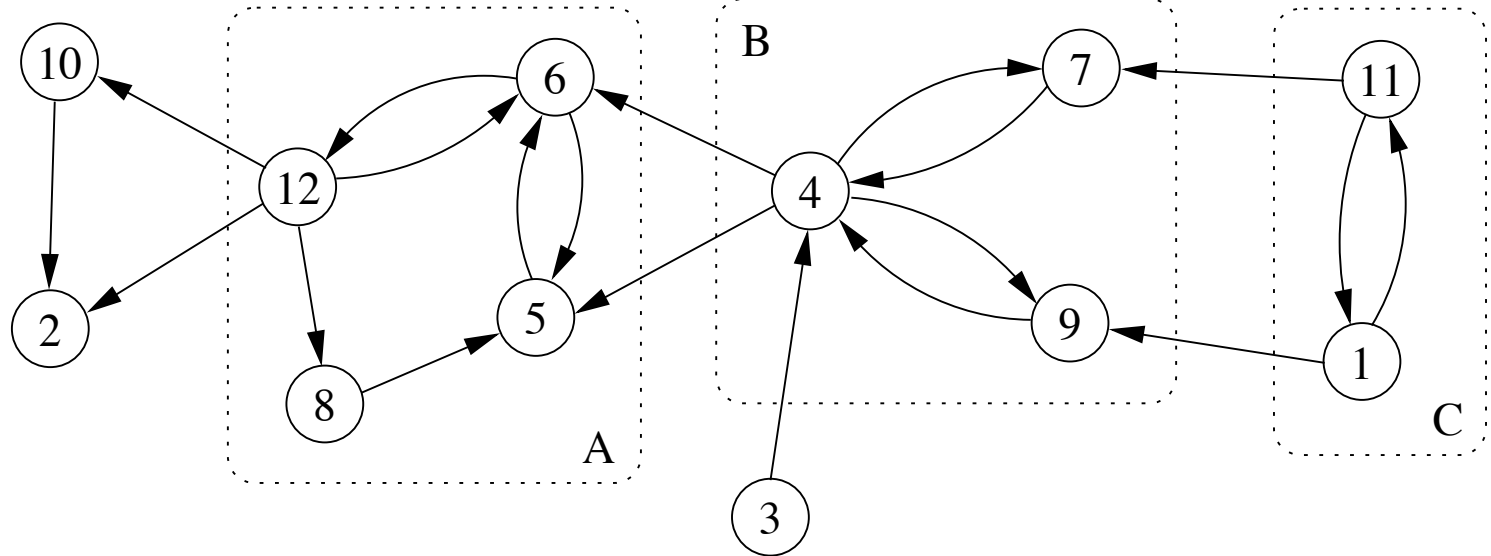
³L'auteur remercie O. Klopfenstein qui a porté ce problème à sa connaissance.

- Problème NP -difficile au sens fort : Sirdey, Carlier, Kerivin et Nace [15].
- Cas particuliers polynomiaux⁴ : *ibid.*
- Recherche arborescente « combinatoire » : *ibid.*
- Résolution approchée :
 - Recuit simulé : Sirdey, Carlier et Nace [16].
 - Grasp : Sirdey, Carlier et Nace [17].
- Méthodes polyédriques : Sirdey et Kerivin [19, 13, 18].

⁴En particulier le cas homogène à une ressource.

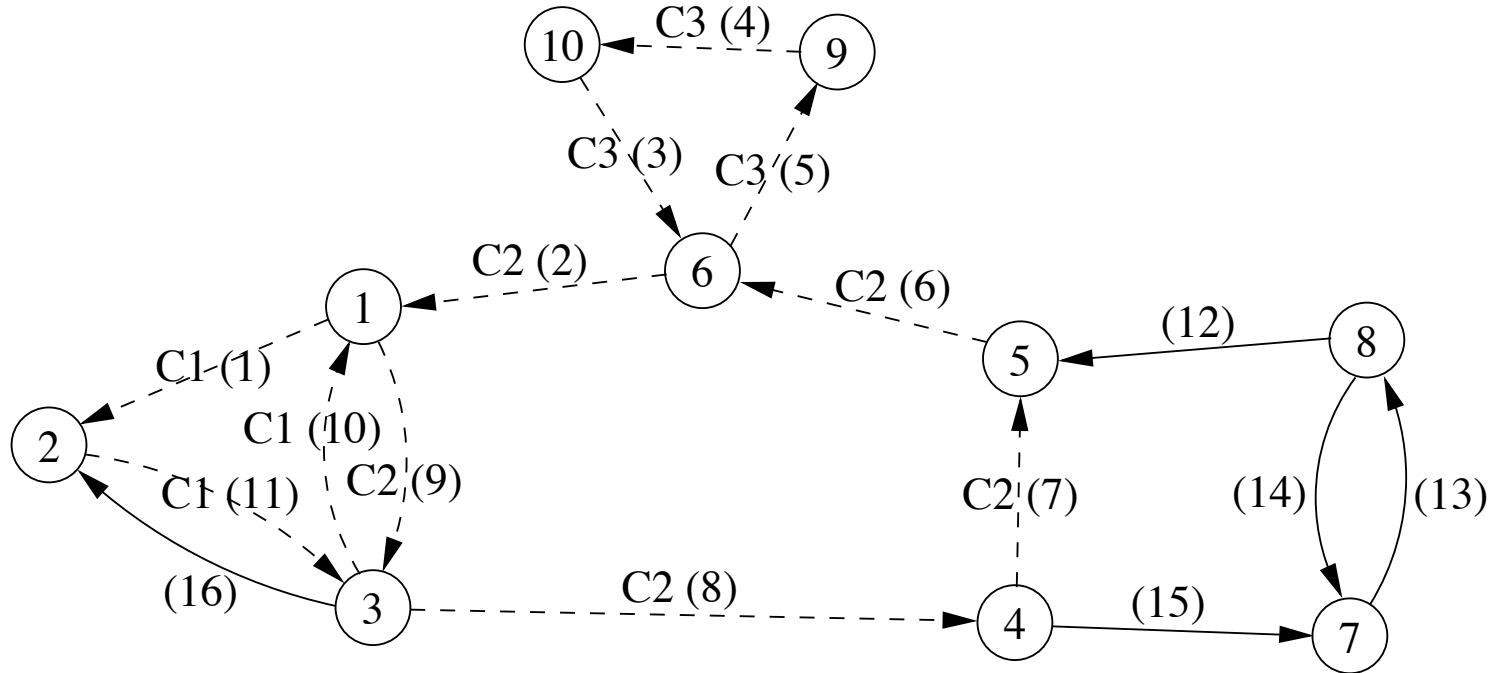
Connexité forte et décomposition

9/46



Le cas homogène (polynomial)

10/46



- U : ensemble des processeurs.
- M : ensemble des déplacements.
- w_m , s_m et t_m : respectivement, la charge déplacée et les processeurs source et cible du déplacement m .
- c_m : le coût de réaliser m par interruption.
- $S(u) = \{m \in M : s_m = u\}$, $u \in U$.
- $T(u) = \{m \in M : t_m = u\}$, $u \in U$.
- K_u : la capacité résiduelle initiale du processeur u .
- (I, σ) : un programme de déplacement de processus.

Programme linéaire en nombres entiers 12/46

– Variables : $\delta_{mm'} = 1$ ($m \neq m'$) ssi m précède m' ,
 $\delta_{mm} = 1$ ssi m est interrompu.

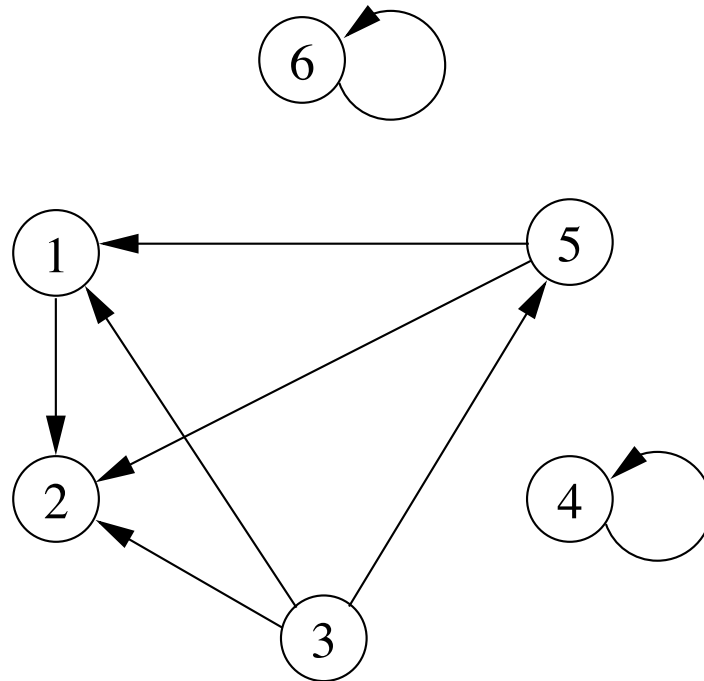
– Énoncé :

$$\left\{ \begin{array}{l} \text{Minimiser } \sum_{m \in M} c_m \delta_{mm} \\ \text{s. l. c.} \\ \delta_{mm'} + \delta_{m'm} + \delta_{mm} + \delta_{m'm'} \geq 1, \forall \{m, m'\} \subseteq M \\ \delta_{mm'} + \delta_{m'm} + \delta_{mm} \leq 1, m, m' \in M, m \neq m' \\ \delta_{mm'} + \delta_{m'm''} - \delta_{mm''} \leq 1, m, m', m'' \in M, m \neq m' \neq m'' \neq m \\ (1 - \delta_{mm})w_m \leq K_{t_m} + \sum_{m' \in S(t_m)} w_{m'}(\delta_{m'm'} + \delta_{m'm}) - \sum_{m' \in C(t_m) \setminus \{m\}} w_{m'} \delta_{m'm}, \forall m \in M \\ \delta_{mm'} \in \{0, 1\}, m, m' \in M \end{array} \right.$$

Le polytope des sous-tournois sans circuit [19]

- Enveloppe convexe, notée P_{PLO}^n , des vecteurs caractéristiques des ordonnancements d'un sous-ensemble des nœuds du graphe orienté complet à n nœuds, les nœuds exclus étant munis d'une boucle.
- Enveloppe entière de

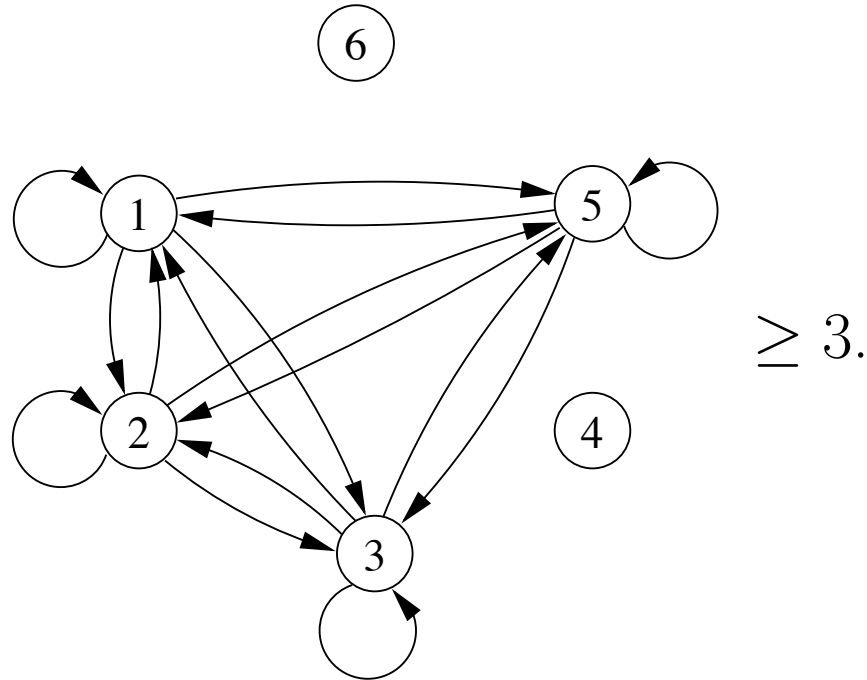
$$\begin{cases} \delta_{ij} + \delta_{ji} + \delta_{ii} + \delta_{jj} \geq 1 & i < j \text{ (minicliques)} \\ \delta_{ij} + \delta_{ji} + \delta_{ii} \leq 1 & i \neq j \text{ (minimonocycles)} \\ \delta_{ij} + \delta_{jk} - \delta_{ik} \leq 1 & i \neq j \neq k \neq i \text{ (trans.)} \\ 0 \leq \delta_{ij} \leq 1 \end{cases}$$



- Pleine dimension.
- Les inégalités $\delta_{ij} \geq 0$ ($i \neq j$), les minicliques et les minimonocycles définissent des facettes.
- Les inégalités $\delta_{ij} \geq 0$ ($i = j$), $\delta_{ij} \leq 1$ et les contraintes de transitivité n'en définissent pas, on peut néanmoins remplacer ces dernières par les *contraintes de transitivité étendues* (qui, elles, définissent des facettes) :

$$\delta_{ij} + \delta_{jk} - \delta_{ik} + \delta_{jj} \leq 1 \quad i \neq j \neq k \neq i$$

- Toute facette de P_{PLO}^n définit une facette de P_{PLO}^{n+1} .



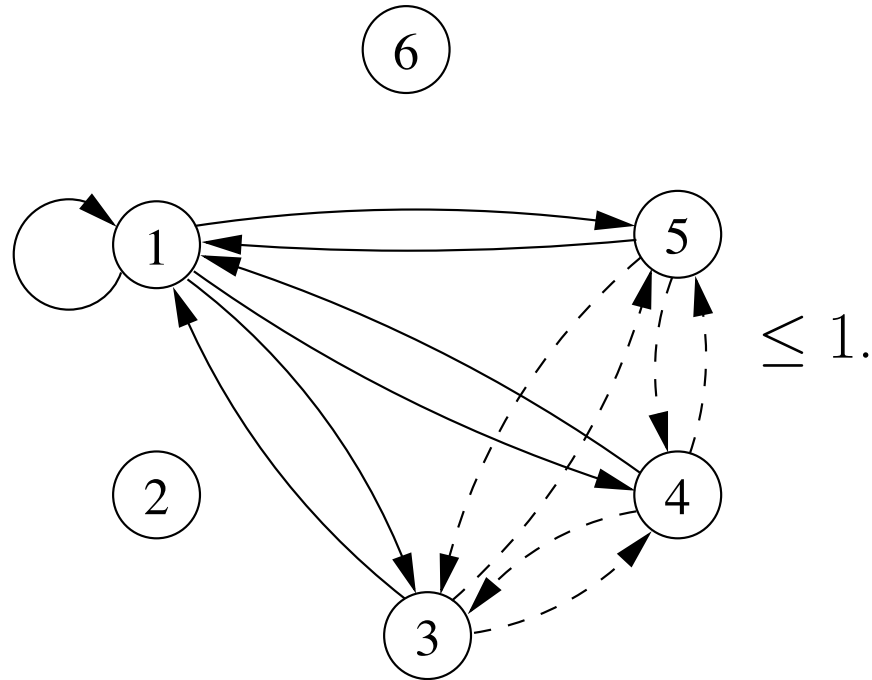
Théorème 1 Soit $I \subset \{1, \dots, n\}$ avec $|I| = k \geq 2$, l'inégalité

$$\sum_{i \in I} \sum_{j \in I} \delta_{ij} \geq |I| - 1$$

définit une facette de P_{PLO}^n .

- Généralisent les minicliques.
- Polynomialement séparables pour $|I|$ fixé.
- Problème de séparation NP -difficile au sens fort⁵.

⁵Par restriction au problème de la coupe de plus fort poids .

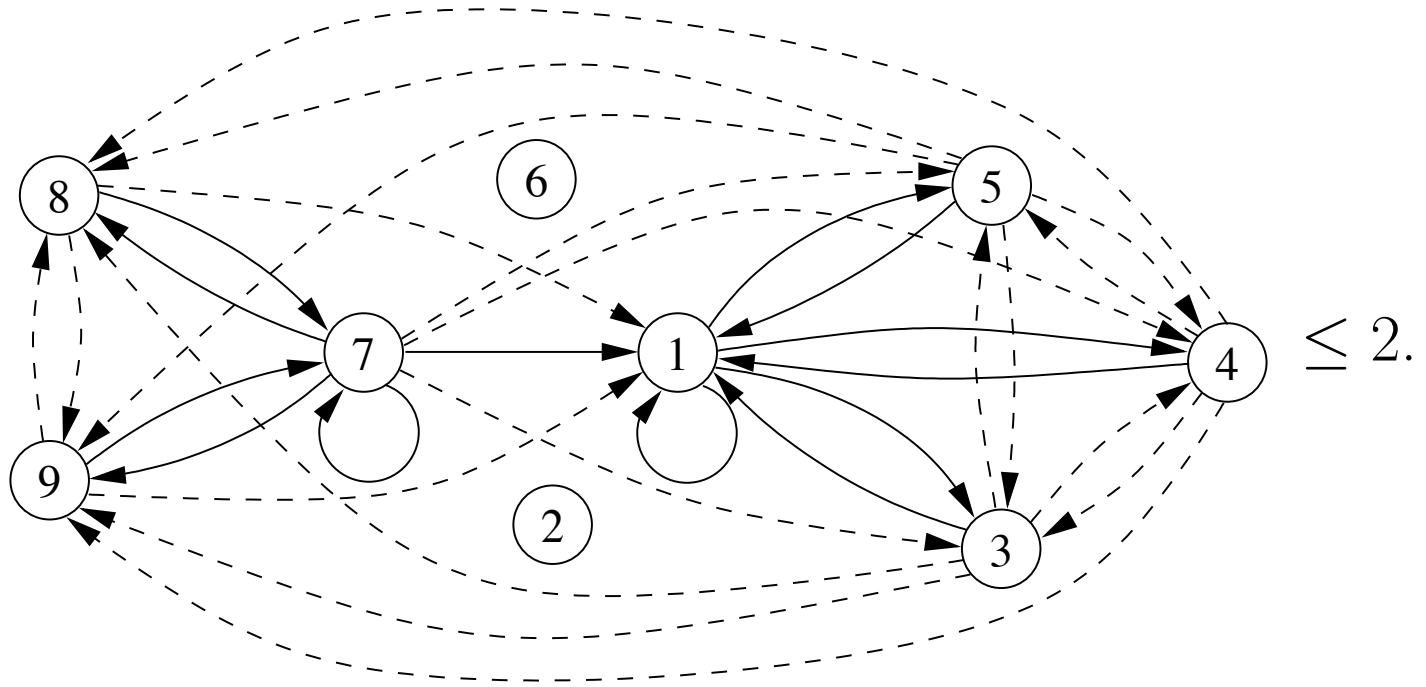


Théorème 2 Soit $I \subset \{1, \dots, n\}$, $|I| = k$, et $i_0 \in \{1, \dots, n\} \setminus I$, l'inégalité

$$\delta_{i_0 i_0} + \sum_{i \in I} (\delta_{i i_0} + \delta_{i_0 i}) - \sum_{i \in I} \sum_{j \in I \setminus \{i\}} \delta_{ij} \leq 1$$

définit une facette P_{PLO}^n .

- Généralisent les minimonocycles.
- Polynomialement séparables pour $|I|$ fixé.
- Problème de séparation NP -difficile au sens fort.



Théorème 3 Soit $\{i_0, j_0\} \subset \{1, \dots, n\}$, $\emptyset \subset I \subset \{1, \dots, n\} \setminus \{i_0, j_0\}$ et $\emptyset \subset J \subset \{1, \dots, n\} \setminus \{i_0, j_0\}$ avec $|I| = k$, $|J| = l$ et $I \cap J = \emptyset$, l'inégalité

$$\begin{aligned} & \delta_{i_0 i_0} + \delta_{j_0 j_0} + \delta_{i_0 j_0} + \sum_{i \in I} (\delta_{i i_0} + \delta_{i_0 i} - \delta_{i j_0}) + \sum_{j \in J} (\delta_{j j_0} + \delta_{j_0 j} - \delta_{i_0 j}) \\ & - \sum_{i \in I} \sum_{i' \in I \setminus \{i\}} \delta_{i i'} - \sum_{j \in J} \sum_{j' \in J \setminus \{j\}} \delta_{j j'} - \sum_{i \in I} \sum_{j \in J} \delta_{j i} \leq 2 \end{aligned}$$

définit une facette de P_{PLO}^n .

Le polytope des programmes de déplacement [13]

- Pleine dimension⁶ (hypothèse pour la suite).
- Les inégalités $\delta_{mm'} \geq 0$ ($m \neq m'$), les contraintes de transitivité étendues, les k -cliques, les k -monocycles et les k - l -bicyclettes⁷ définissent des facettes de P_{PMP} .
- Les inégalités $\delta_{mm} \geq 0$, $\delta_{mm'} \leq 1$, les contraintes de transitivité et les contraintes de capacité n'en définissent pas.

⁶Si tous les programmes qui interrompent tous les déplacements sauf deux sont admissibles.

⁷Sous des hypothèses raisonnablement peu restrictives.

Théorème 4 Soit $m_0 \in M$, $\emptyset \subset A \subseteq T(s_{m_0})$ et $B \subseteq S(s_{m_0}) \setminus \{m_0\}$ tels que

$$\sum_{m \in A} w_m > K_{s_{m_0}} + \sum_{m \in \bar{B}} w_m$$

avec $\bar{B} = S(s_{m_0}) \setminus (B \cup \{m_0\})$, alors l'inégalité

$$\sum_{m \in A} \delta_{mm_0} + \sum_{m \in B} \delta_{m_0m} \leq (|A| + |B| - 1)(1 - \delta_{m_0m_0})$$

est valide pour P_{PMP} .

Théorème 5 Soit $m_0 \in M$, $A \subseteq T(t_{m_0}) \setminus \{m_0\}$ et $\emptyset \subset B \subseteq S(t_{m_0})$ tels que

$$w_{m_0} + \sum_{m \in A} w_m > K_{t_{m_0}} + \sum_{m \in \bar{B}} w_m$$

avec $\bar{B} = S(t_{m_0}) \setminus B$, alors l'inégalité

$$\sum_{m \in A} \delta_{mm_0} + \sum_{m \in B} \delta_{m_0m} \leq (|A| + |B| - 1)(1 - \delta_{m_0m_0})$$

est valide pour P_{PMP} .

- Dans les deux cas, les processus associés aux déplacements de $A \cup B \cup \{m_0\}$ ne peuvent pas tous cohabiter sur s_{m_0} (respectivement t_{m_0}).
- Sous certaines hypothèses (raisonnablement peu restrictives) les contraintes de s - et de t -recouvrement définissent des facettes de P_{PMP} .
- Dans les deux cas, le problème de séparation se réduit à un problème de sac à dos lorsque m_0 est fixé. Le problème de séparation est donc pseudopolynomial.

Recherche arborescente polyédrique [18]

- Obtention d'une « bonne » solution initiale⁸.
- Résolution de la relaxation⁹ comprenant :
 - Les $|M|$ contraintes de capacité.
 - Les $O(|M|^2)$ minicliques.
 - Les $O(|M|^2)$ minimonocycles.
 - Les $O(|M|^3)$ contraintes de transitivité étendues.
 - Les contraintes de s - et de t -recouvrement.
- Recherche arborescente.

⁸À l'aide d'un recuit simulé conçu pour fournir des solutions 5%-acceptables [16].

⁹Théoriquement réalisable en temps pseudopolynomial.

- Une solution ω d'un problème de minimisation est (α, β) -acceptable si $P(c(\omega) \leq e_1 + \beta(e_P - e_1)) \geq \alpha$.
- C'est le cas des solutions issues de la loi stationnaire de l'algorithme de Metropolis à la température

$$T_f(z) = \frac{\beta(e_P - z)}{\log N - \log(1 - \alpha)}, \quad e_1 \leq z \leq e_P.$$

- Il en découle un algorithme pseudopolynomial capable *en pratique* (vs en théorie) de fournir des solutions $(95\%, 5\%)$ -acceptables ($0.9408 \leq \hat{\alpha} \leq 0.9676$).

- Solveur utilisé : COIN-OR LP¹⁰.
- Séparation des contraintes¹¹ :
 - Les minicliques, les minimonocycles et les contraintes de transitivité étendues par énumération.
 - Les contraintes de s - et t -recouvrement par résolution de $2|M|$ sacs à dos (algorithme de Bellman [12]).
- À chaque itération, toutes les contraintes inactives sont éliminées¹¹ [9].

¹⁰www.coin-or.org.

¹¹À l'exception des $|M|$ contraintes de capacité.

$ U $	$ M $	z_0	$d(z_0)$	GLB	$d(\text{GLB})$	# itér.	# cont.	z^*	# nœuds	Temps ^a
25	45	84	0.00%	78	0.33%	144	3667	84	5	1 m 53.88 s
25	38	173	3.40%	110	0.00%	53	2394	110	3	36.35 s
25	50	161	2.26%	110	0.24%	225	4352	115	13	5 m 30.18 s
25	37	218	3.51%	155	0.00%	48	2529	155	1	49.78 s
25	36	206	2.96%	153	0.00%	54	2263	153	4	31.28 s
25	41	68	0.00%	68	0.00%	≥ 99	-	68	1	1 m 10.85 s
25	40	118	2.01%	77	0.22%	73	2657	81	15	52.79 s
25	55	74	1.23%	49	0.00%	325	4652	49	5	5 m 29.06 s
25	47	158	1.27%	135	0.00%	151	3700	135	6	2 m 10.71 s
25	42	128	3.30%	69	0.00%	43	2671	69	2	39.09 s
	43.1		1.99%		0.08%					

^aRésultats obtenus à l'aide d'une station Sun Ultra 10 équipée d'un microprocesseur Sparc cadencé à 440 MHz, de 512 Mo de mémoire vive et du système d'exploitation Solaris 5.8. Le temps total d'exécution est limité à 4 heures.

Résultats empiriques (suite)

33/46

$ U $	$ M $	# inst.	# rés.	$\bar{d}(z_0)$	$\bar{d}(\text{GLB})$	$\bar{d}(z_f)$
15	23.2	10	10	1.92%	0.07%	0.00%
20	34.1	10	10	1.20%	0.33%	0.00%
25	43.1	10	10	1.99%	0.08%	0.00%
30	53.1	10	6	$\leq 2.01\%$	$\leq 1.05\%$	$\leq 0.74\%$
35	61.8	10	7	$\leq 2.38\%$	$\leq 0.98\%$	$\leq 0.86\%$
40	67.5	10	6	$\leq 2.23\%$	$\leq 0.48\%$	$\leq 0.43\%$
45	76.5	10	6	$\leq 2.50\%$	$\leq 1.15\%$	$\leq 1.02\%$
50	88.5	10	3	$\leq 2.69\%$	$\leq 1.80\%$	$\leq 1.72\%$
55	93.6	10	4	$\leq 3.22\%$	$\leq 1.34\%$	$\leq 1.27\%$
60	108.4	10	2	$\leq 2.80\%$	$\leq 2.52\%$	$\leq 2.52\%$
75	117.7	10	0	$\leq 3.69\%$	$\leq 3.69\%$	$\leq 3.69\%$
70	122.3	10	1	$\leq 3.32\%$	$\leq 3.02\%$	$\leq 3.00\%$
75	131.3	10	0	$\leq 4.08\%$	$\leq 4.08\%$	$\leq 4.08\%$
80	137.4	10	0	$\leq 3.30\%$	$\leq 3.30\%$	$\leq 3.30\%$

Perspectives : reroutage dans les réseaux MPLS [14]

- MPLS¹² est une nouvelle technologie de commutation de paquets pour les grands réseaux publics.
- Cette technologie est basée sur la notion de tunnel ou LSP¹³.
- Les LSP sont associés à des services et consomment la bande passante des liens du réseau sous-jacent.

¹²*Multi-Protocol Label Switching.*

¹³*Label Switched Path.*

- Au cours de la vie d'un réseau, des LSP sont régulièrement ajoutés, modifiés ou supprimés.
- Après un certain temps d'opération, les ressources du réseau peuvent être utilisées de manière sous-optimale.
- D'où un besoin de reconfiguration, généralement réalisée en deux temps [11] :
 - Une meilleure configuration est calculée.
 - Le réseau est reconfiguré en reroutant des LSP, sans violations des contraintes de capacité sur les liens.

- Rerouter un LSP consiste à
 - Établir le nouveau chemin.
 - Basculer le trafic sur le nouveau chemin.
 - Supprimer l'ancien chemin.
- Les blocages sont résolus en supprimant des LSP pendant la reconfiguration, l'objectif étant d'avoir le moins recours possible à ce mécanisme.
- Reconfigurer un autocommutateur revient à reconfigurer un réseau MPLS avec 2 nœuds et $|U|$ liens.

- V : l'ensemble des nœuds du réseau.
- L : l'ensemble des liens du réseau.
- P : l'ensemble des LSP établis sur le réseau.
- C_l : la bande passante du lien l .
- w_p : la bande passante requise par le LSP p .
- S_p et T_p : respectivement, les liens utilisés par le LSP p dans les configurations initiale et finale.
- S_l et T_l : respectivement, les LSP qui utilisent le lien l dans les configuration initiale et finale.

- $K_l = C_l - \sum_{p \in S_l} w_p$.
- $S_p^\Delta = S_p \setminus T_p$, $T_p^\Delta = T_p \setminus S_p$.
- $S_l^\Delta = S_l \setminus T_l$, $T_l^\Delta = T_l \setminus S_l$.
- $\Delta = \{p \in P : S_p \neq T_p\}$.
- $\delta_{pp'} = 1$ ($p, p' \in \Delta$, $p \neq p'$) si et seulement si p est rerouté avant p' et ni p ni p' ne sont interrompus.
- $\delta_{pp} = 1$ ($p \in \Delta$) si et seulement si p est interrompu.

– Énoncé :

$$\left\{ \begin{array}{ll}
 \text{Minimiser } \sum_{p \in \Delta} \gamma_p \delta_{pp} & \\
 \text{s. l. c.} & \\
 \delta_{pp'} + \delta_{p'p} + \delta_{pp} + \delta_{p'p'} \geq 1 & \forall \{p, p'\} \subseteq \Delta, \\
 \delta_{pp'} + \delta_{p'p} + \delta_{pp} \leq 1 & p \neq p' \in \Delta, \\
 \delta_{pp'} + \delta_{p'p''} - \delta_{pp''} \leq 1 & p \neq p' \neq p'' \neq p \in \Delta, \\
 (1 - \delta_{pp})w_p \leq K_l + \sum_{p' \in S_l^\Delta} w_{p'}(\delta_{p'p'} + \delta_{p'p}) - \sum_{p' \in T_l^\Delta \setminus \{p\}} w_{p'} \delta_{p'p} & \forall p \in \Delta, \forall l \in T_p^\Delta, \\
 \delta_{pp'} \in \{0, 1\} & p, p' \in \Delta.
 \end{array} \right.$$

– Le polytope associé s'appelle le *polytope des programmes de reroutage* et se note P_{RSP} .

- Les résultats concernant la dimension ou les facettes simples obtenus pour P_{PMP} se généralisent aisément à P_{RSP} , les k -cliques et les k -monocycles en définissent des facettes (*quid* des k - l -bicyclettes?).
- Les contraintes de s - et t -recouvrement se généralisent naturellement à P_{RSP} , à quelles conditions en définissent-elles des facettes?
- *Quid* de la pertinence pratique d'une relaxation linéaire utilisant ces inégalités?

Service d'architecture BSC (PC 12A7)

Nortel GSM Access R&D

Parc des activités de Magny-les-Hameaux

78928 Châteaufort Cedex 09, France

Tél. : +33 (0)1 69 55 41 18 (ESN : 574-4118)

Fax : +33 (0)1 69 55 31 25 (ESN : 574-3125)

Courrier électronique : *renauds@nortel.com*

- [1] G. AGGARWAL, R. MOTWANI et A. ZHU. « The load rebalancing problem ». In *Proceedings of the Fifteenth Annual ACM Symposium on Parallel Algorithms and Architectures*, pages 258–265, 2003.
- [2] E. ANDERSON, J. HALL, J. HARTLINE, M. HOBBS, A. R. KARLIN, J. SAIA, R. SWAMINATHAN et J. WILKES. « An experimental study of data migration algorithms ». In *Proceedings of the 5th International Workshop on Algorithm Engineering*, Lecture Notes in Computer Science, page 145. Springer, 2001.
- [3] D. O. AWDUCHE. « MPLS and traffic engineering in IP networks ». *IEEE Communication Magazine*, 37(12) :42–47, 1999.
- [4] J. CARLIER. « Le problème de l’ordonnancement des paiements de dettes ». *R.A.I.R.O. Recherche Opérationnelle/Operations Research*, 18(1), février 1984.
- [5] J. CARLIER. *Problèmes d’ordonnancement à contraintes de ressources : algorithmes et complexité*, tome 40 de *Méthodologie & Architecture des Systèmes Informatiques*. Université P. et M. Curie et CNRS, 1984.

- [6] E. G. COFFMAN, M. R. GAREY, D. S. JOHNSON et A. S. LAPAUGH. « Scheduling file transfers in distributed networks ». In *Proceedings of the 2nd Annual ACM Symposium on Principles of Distributed Computing*, pages 254–266, 1983.
- [7] E. G. COFFMAN, M. R. GAREY, D. S. JOHNSON et A. S. LAPAUGH. « Scheduling file transfers ». *SIAM Journal on Computing*, 14(3), 1985.
- [8] F. GLOVER, B. ALIDAEI, C. REGO et G. KOCHENBERGER. « One-pass heuristics for large-scale unconstrained binary quadratic problems ». *European Journal of Operational Research*, 137 :272–287, 2002.
- [9] M. GRÖTSCHEL, M. JÜNGER et G. REINELT. « A cutting plane algorithm for the linear ordering problem ». *Operations Research*, 32(6) :1195–1220, 1984.
- [10] J. HALL, J. HARTLINE, A. R. KARLIN, J. SAIA et J. WILKES. « On algorithms for efficient data migration ». In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 620–629, 2001.
- [11] B. G. JÓZSA et M. MAKAI. « On the solution of reroute sequence planning problem in MPLS networks ». *Computer Networks*, 42 :199–210, 2003.
- [12] H. KELLERER, U. PFERSCHY et D. PISINGER. *Knapsack problems*. Springer, 2004.

- [13] H. KERIVIN et R. SIRDEY. « Polyhedral combinatorics of a resource-constrained ordering problem part II : on the process move program polytope ». Rapport technique PE/BSC/INF/017913 V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France (submitted to Mathematical Programming), 2006.
- [14] R. SIRDEY. « A polyhedral approach to reroute sequence planning in MPLS networks ». Rapport technique PE/BSC/INF/xxxx V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France, 2006.
- [15] R. SIRDEY, J. CARLIER, H. KERIVIN et D. NACE. « On a resource-constrained scheduling problem with application to distributed systems reconfiguration ». Rapport technique PE/BSC/INF/015593 V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France (submitted to the European Journal of Operational Research), 2005.
- [16] R. SIRDEY, J. CARLIER et D. NACE. « Approximate resolution of a resource-constrained scheduling problem ». Rapport technique PE/BSC/INF/016550 V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France (submitted to the Journal of Heuristics), 2005.
- [17] R. SIRDEY, J. CARLIER et D. NACE. « A fast heuristic for a resource-constrained scheduling problem ». Rapport technique PE/BSC/INF/017254 V01 EN, Service d'ar-

chitecture BSC, Nortel GSM Access R&D, France, 2005.

- [18] R. SIRDEY et H. KERIVIN. « A branch-and-cut algorithm for a resource-constrained scheduling problem ». Rapport technique PE/BSC/INF/017633 V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France (submitted to RAIRO—Operations Research), 2006.
- [19] R. SIRDEY et H. KERIVIN. « Polyhedral combinatorics of a resource-constrained ordering problem part I : on the partial linear ordering polytope ». Rapport technique PE/BSC/INF/017912 V01 EN, Service d'architecture BSC, Nortel GSM Access R&D, France (submitted to Mathematical Programming), 2006.