

Calcul parallèle Topologie des architectures parallèles

Florian Barbarin
Abdelkader Beldjilali
Nicolas Holvoet

30 décembre 2016



Table des matières

Introduction	3
1 Algorithmes de Dijkstra et A*	4
2 Les grilles	6
3 Les arbres et grilles d'arbres	9
4 Les hypercubes	10
4.1 Comparaison des différentes architectures	11
4.2 Template de communication	11
4.3 Tri bitonique	11
4.3.1 <i>Compare-split</i>	12
4.3.2 <i>Parallel merge</i>	12
4.4 Quelques autres exemples d'algorithmes	12
Références	13

Introduction

bla

1 Algorithmes de Dijkstra et A*

Algorithmes de Dijkstra : Edgser Wybe Dijkstra (EWD), Physicien Néerlandais reconverti à l'informatique en 1955, a proposé en 1959 un algorithme de recherche de chemin minimum dans un graphe dont la complexité est en $O(n)$.

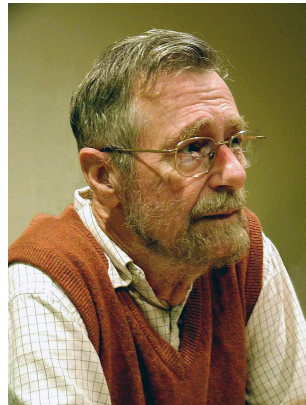


FIGURE 1 – Edgser Wybe Dijkstra (1930-2002)

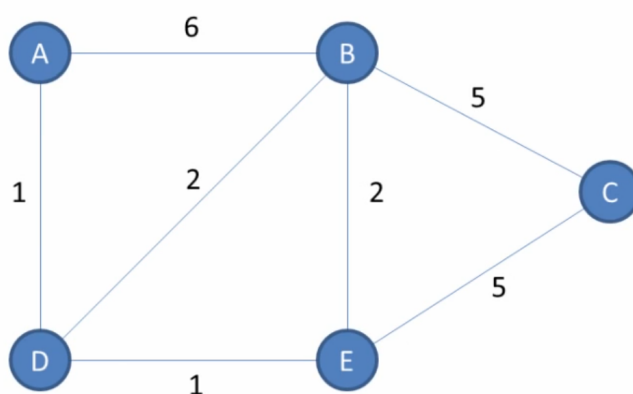
On doit à Dijkstra, qui avait la réputation d'avoir mauvais caractère et qui était notoirement allergique au "GOTO", quelques citations¹ telles que :

« Il est pratiquement impossible d'enseigner la bonne programmation aux étudiants qui ont eu une exposition antérieure au BASIC : comme programmeurs potentiels, ils sont mentalement mutilés, au-delà de tout espoir de régénération. »

« Le plus court chemin d'un graphe n'est jamais celui que l'on croit, il peut surgir de nulle part, et la plupart du temps, il n'existe pas. »

« La programmation par objets est une idée exceptionnellement mauvaise qui ne pouvait naître qu'en Californie. »

L'algorithme donne le plus court chemin de la source à *tous les sommets* d'un graphe connexe pondéré (orienté ou non) dont le poids lié aux arêtes est positif ou nul.



Vertex	Shortest distance from A	Previous vertex
A	0	
B	3	D
C	7	E
D	1	A
E	2	D

Visited = [A, D, E, B, C] Unvisited = []

FIGURE 2 – Exemple de calcul des plus courts chemins à partir du noeud A.

Sans entrer dans les détails d'implémentation que l'on trouve nombreuses sur la toile et dans la littérature, l'algorithme de Dijkstra est un algorithme glouton qui utilise l'hypothèse qu'une décision prise sur la base d'un critère d'optimalité locale conduira à un optimum global. Ainsi, à chaque itération, l'algorithme choisit, parmi les noeuds non traités, le noeud du réseau dont la distance au noeud de départ est la plus faible. Ainsi, dans l'image 2, on voit que le plus court chemin pour aller de A à C est égal à 7. On détermine ensuite de proche en proche le chemin le plus court grâce à la mémorisation des noeuds précédents. Ici, C à pour "previous vertex" E, E -> D et D -> A. Le plus court chemin pour aller de A à C est donc ADEC. On utilise la même technique pour déterminer tous les chemins

1. source : https://fr.wikipedia.org/wiki/Edsger_Dijkstra

optimaux d'origine A. Mais on peut n'avoir besoin que du calcul du chemin entre 2 sommets. De plus, la myopie de l'algorithme limite forcément ses performances. D'où l'idée d'utiliser une heuristique visant à guider les choix locaux. Ce qui nous amène à l'algorithme A* présenté brièvement dans la suite.

Algorithme A* : A* est l'algorithme qu'on utilise intuitivement pour se déplacer d'un endroit à un autre dans une ville connaissant la direction à prendre. Grâce à cette information supplémentaire, A* va privilégier le noeud qui minimise la somme de la distance déjà parcourue et de la distance estimée restant à parcourir qui sera ici la distance à vol d'oiseau.

Cet algorithme a été proposé pour la première fois par Peter E. Hart, Nils John Nilsson et Bertram Raphael en 1968. Il s'agit d'une extension de l'algorithme de Dijkstra et s'applique comme ce dernier à des graphes munis d'une distance positive. En outre, A* ne teste pas tous les chemins ; il ne fournit donc qu'un des meilleurs chemins. A* est cependant très performant dans le cas où le graphe comporte une faible densité de noeuds mais pas ou peu efficace dans le cas de parcours de labyrinthes par exemple.

Il faut donc choisir, en fonction du problème, une heuristique h , c'est à dire ici une fonction qui estime la distance restante entre chaque noeud et l'arrivée, estimation par défaut qui ne doit jamais surestimer cette distance. La précision du résultat final dépendra de la précision de l'estimation. On peut citer les fonctions estimatrices suivantes : la distance euclidienne (norme 2) ou à "vol d'oiseau" déjà citée et la distance de Manhattan (norme 1) qui, sur une grille, correspond à la somme des cases verticales et horizontales qui sépare la position courante de l'arrivée. Les deux distances sont utilisables dans A* car elles sous-estiment la distance à l'arrivée.

A chaque étapes, l'algorithme calcule g , la distance déjà parcourue pour chaque noeud n non visité puis la fonction $f(n) = g(n) + h(n)$. Le noeud sélectionné est celui pour lequel $f(n)$ est minimale.

Utilisation Dijkstra est utilisé dans le routage dynamique OSPF, on peut donc envisager optimiser un réseau de noeuds distribués.

Comment les utiliser pour transférer de façon optimale une donnée d'un noeud à un autre.?????

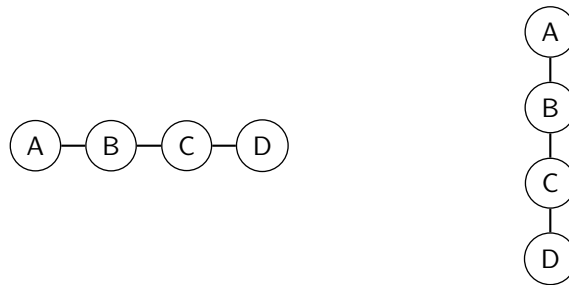
2 Les grilles

Définition (Grille). Une grille de dimension d possédant N nœuds suivant chaque coordonnée est le produit cartésien de d chaînes ($d > 1$) de N sommets. On note cette grille $M(N)^d$ que l'on dira de côté N .

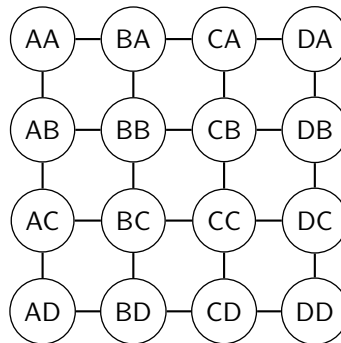
Remarque. Si l'on considère le produit cartésien de deux graphes, le graphe résultant est tel que :

- l'ensemble de ses nœuds est le produit cartésien des nœuds des deux premiers graphes ;
- deux de ses nœuds sont voisins s'ils sont composés de nœuds qui étaient voisins dans l'un des deux premiers graphes.

Exemple. Soient deux chaînes composées des nœuds appartenant à l'ensemble $E = \{A, B, C, D\}$. Le produit cartésien de ces deux chaînes ($d = 2$) de taille $N = \text{Card}(E) = 4$:



A pour résultat la grille $M(4)^2$:

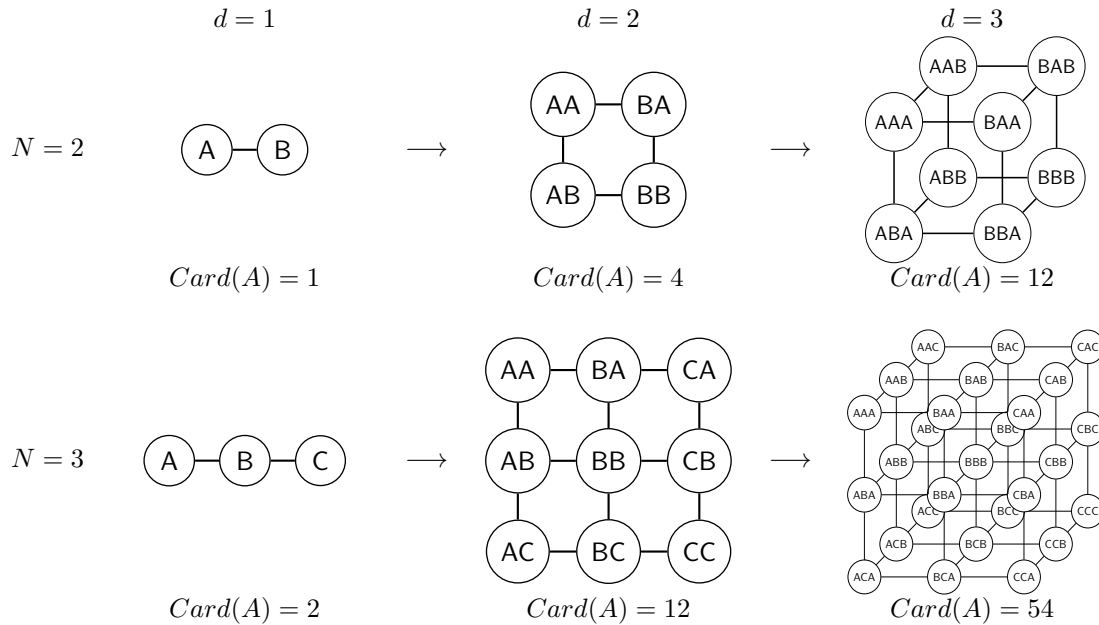


Nombre total de nœuds : Il résulte de la définition ci-dessus que le nombre total de nœuds est égal au cardinal du produit cartésien de l'ensemble des nœuds de départ :

$$\text{Card}(S) = \text{Card}(\underbrace{E \times E \times \dots \times E}_{p \text{ fois}}) = \prod_1^p \text{Card}(E) = N^p$$

où S est l'ensemble des noeuds de la grille

Nombre total d'arêtes : Soit A l'ensemble des arêtes de la grille. Voici pour $N = 2$ et $N = 4$ le passage de la dimension 1 aux dimensions supérieures (2 et 3).



On remarque à présent que, pour N fixé, le passage d'une dimension d à une dimension $d + 1$ se fait en deux étapes :

- on "copie" N fois la grille de dimension d ;
- on relie, par une arête, les nœuds de la grille 1 avec les nœuds correspondants de la grille 2 puis les nœuds de la grille 2 avec les nœuds correspondants de la grille 3, et ainsi de suite jusqu'à la grille N .

Cette méthode de construction nous donne une relation de récurrence pour le nombre d'arêtes :

$$Card(A)_{d+1} = \underbrace{N \times Card(A)_d}_{\text{On copie } N \text{ fois la grille de dimension } d} + \underbrace{(N-1) \times Card(S)_d}_{\text{On relie les arêtes de chaque grille}}$$

On ne peut aisément déterminer une expression générale de la suite ci-dessus. Or, les exemples précédents nous permettent de déduire une expression du nombre d'arêtes en fonction de d et de N :

$$Card(A)_d = d \times (N-1) \times N^{d-1}$$

Nous pouvons démontrer par récurrence cette expression.

Preuve de la relation.

Initialisation : Pour $d = 1$, on a $Card(A)_1 = 1 \times (N-1) \times N^{1-1} = N-1$ ce qui correspond bien au nombre d'arêtes dans une chaîne.

Hypothèse de récurrence : On fait l'hypothèse qu'il existe un rang d tel que $Card(A)_d = d \times (N-1) \times N^{d-1}$. Montrons que cette relation est vraie au rang $d + 1$.

Hérédité : Nous avons :

$$\begin{aligned}
 Card(A)_{d+1} &= N \times Card(A)_d + (N-1) \times N^d \\
 &= N \times d \times (N-1) \times N^{d-1} + N^d \times (N-1) \\
 &= N^d \times d \times (N-1) + N^d \times (N-1) \\
 &= (d+1) \times (N-1) \times N^d
 \end{aligned}$$

Nous retrouvons bien l'hypothèse de récurrence au rang $d + 1$. On en déduit que $\forall d \in \mathbb{N}^*$, on a $Card(A)_d = d \times (N-1) \times N^{d-1}$. □

D'après l'ensemble des éléments qui précèdent, le nombre d'arêtes d'une grille est telle que :

$$\forall d \in \mathbb{N}^*, Card(A)_d = d \times (N-1) \times N^{d-1}$$

Diamètre de la grille : La construction de la grille telle qu'envisagée jusque là nous permet de faire l'observation suivante : un nœud de la grille diffère exactement d'une seule coordonnée de ses voisins. Pour reprendre l'exemple déjà vu ci-dessus avec $N = 2$ et $d = 2$, le nœud AA est voisin des nœuds AB et BA . Dans un cas plus général où les nœuds de la grille seraient formés de coordonnées entières, le nœud de coordonnées $(x_0, x_1, \dots, x_{d-1})$ aurait pour voisin dans la dimension i le nœud de coordonnées $(y_0, y_1, \dots, y_i, \dots, y_{d-1})$ avec $y_i = x_i \pm 1$.

Or par définition, le diamètre d'un graphe est la plus grande distance entre deux sommets. Dans le cas général d'une grille composée de coordonnées entières, les deux nœuds les plus éloignés sont le nœud $(0, 0, \dots, 0)$ et le nœud (N, N, \dots, N) . Le plus long chemin entre ces deux nœuds est donc de passer par l'ensemble des coordonnées possibles, ce qui correspond à q fois $N - 1$ possibilités.

Au final, le diamètre D d'une grille est :

$$D = q \times (N - 1)$$

Bissection : La bissection, ou plus précisément la largeur de la bissection est le nombre minimum d'arête qu'il faut enlever à la grille pour la diviser en deux moitiés avec un nombre de nœuds identique.

On remarque déjà que ce problème dépend de la parité du nombre de nœuds de la grille : dans le cas où celui-ci est pair, il est possible de diviser la grille en deux avec un nombre de nœuds identiques ; dans le cas où celui-ci est impair, les deux grilles résultantes auront un nombre de nœuds égal à plus ou moins un nœud près.

Cas où N est pair : La séparation en deux grilles avec un nombre de nœuds identique est triviale. La question est cependant de connaître le nombre exacte d'arêtes à enlever pour obtenir cette séparation.

Or on a vu précédemment que passer de la dimension $d - 1$ à la dimension d se faisait en deux étapes : recopie N fois de la grille de dimension $d - 1$ et création des nouvelles arêtes entre chaque nœuds correspondants des grille, c'est à dire N^{d-1} .

Ainsi, lorsque l'on sépare une grille en deux, on la sépare à une jonction entre deux grille de dimension $d - 1$ ce qui nous conduit à enlever exactement N^{d-1} .

On peut donc dire que la largeur de la bissection dans le cas où N est de N^{d-1} .

Cas où N est impair : On peut montrer que dans le cas où N est impair, la largeur de la bissection est :

$$\sum_{i=0}^{d-1} N^i = \frac{N^d - 1}{N - 1}$$

3 Les arbres et grilles d'arbres

Diamètre : Un arbre est constitué d'un ensemble de niveaux, le niveau 0 étant occupé par la racine de l'arbre. Les niveaux se répartissent donc de 0 à h où h est par définition la hauteur de l'arbre. Chaque niveau i est occupé par k^i noeuds dans un arbre k -naire. Un arbre binaire complet aura donc $n = 2^h$ feuilles qui représente aussi le nombre de processeurs. La plus grande distance entre deux noeuds, ie. le diamètre de l'arbre, est donc le nombre d'arrêtes du chemin passant par la racine et joignant deux feuilles donc $D = 2h = 2 * \log_2 n$.

Bissection : La bissection est le nombre minimal d'arrêtes à supprimer pour qu'un graphe initialement connexe se sépare en deux composantes connexes possédant le même nombre de noeuds à une unité près. Pour un arbre binaire, ce nombre minimal est obtenu en supprimant une des deux arrêtes liées à la racine. On trouve donc dans ce cas une bissection égale à 1.

4 Les hypercubes

Un hypercube est une grille de dimension d ne possédant que deux sommets selon chaque coordonnée. Ainsi, il possède 2^d nœuds de degré d et $d2^{d-1}$ arêtes. On construit un hypercube de dimension d récursivement à partir de deux hypercubes de dimension $d-1$ en connectant les sommets similaires, un hypercube de dimension 0 correspondant à un nœud de calcul unique. On obtient ainsi ces quatre premières grilles :

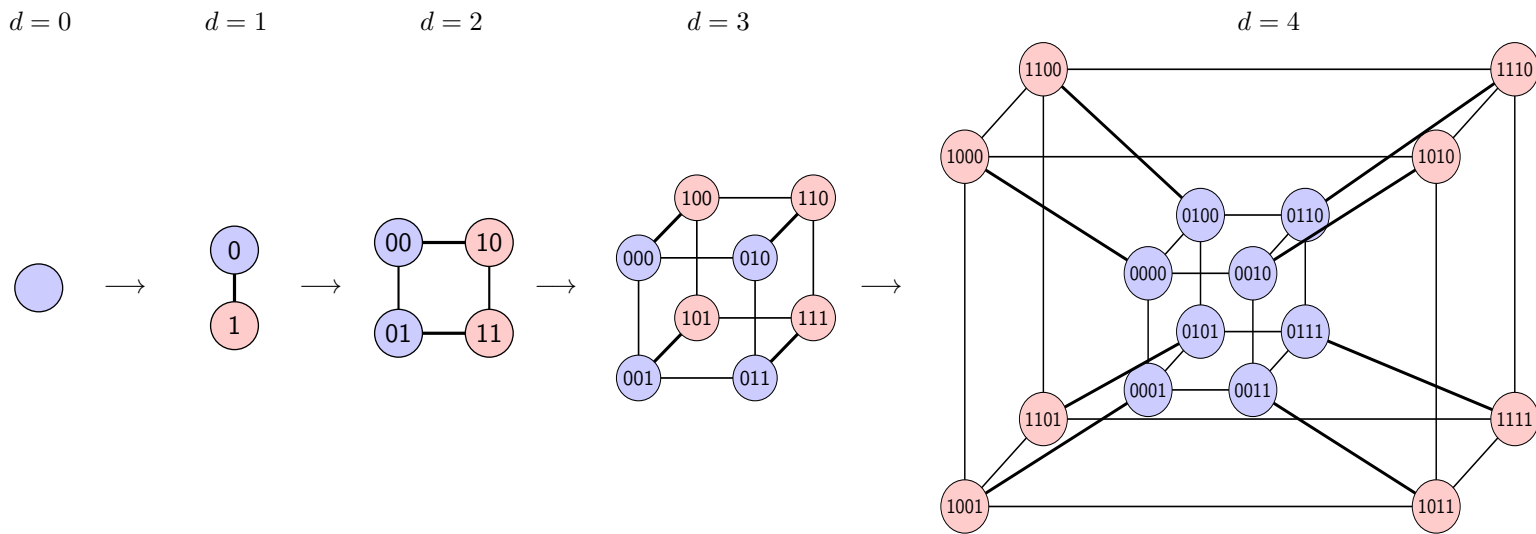


FIGURE 3 – Construction des hypercubes de dimension 0 à 4

On étiquette généralement les nœuds par une séquence de d bits. Pour étiqueter un hypercube de dimension d , on part d'un hypercube de dimension $d-1$ dont les sommets auront été préfixés d'un 0 (en bleu sur la figure) puis on ajoute un hypercube de dimension $d-1$ dont les sommets auront été préfixés d'un 1 (en rouge). On a alors deux propriétés intéressantes :

- Chaque bit correspond à une dimension de l'hypercube. Ainsi, dans notre exemple $d=3$ et par construction, le premier bit (en partant de la droite) correspond à la coordonnée classique z (apparue dans l'hypercube $d=1$), le deuxième correspond à la coordonnée y (issue de l'hypercube $d=2$) et enfin le troisième à x . On peut donc voir les étiquettes sous une forme « xyz ».
- Un nœud est voisin d'un autre nœud si et seulement si leurs étiquettes diffèrent d'un seul bit. Ainsi, la distance entre deux nœuds correspond à la distance de Hamming de leurs étiquettes.

Ces propriétés permettent d'implémenter un algorithme de routage très simple dans lequel un message transite via les nœuds dont les étiquettes se rapprochent à chaque étape d'un bit de l'étiquette destination. Enfin, cela permet, à l'aide de codes de Gray bien définis, de plonger une grille classique ou un arbre dans un hypercube. .

Diamètre Comme nous l'avons vu pour les grilles, le diamètre d'un hypercube de dimension d correspond à la distance entre le nœud $(\underbrace{0, 0, \dots, 0}_{d \text{ fois}})$ et le nœud $(\underbrace{1, 1, \dots, 1}_{d \text{ fois}})$ qui, compte tenu de la deuxième propriété énoncée ci-dessus, est égale à d .

Bissection D'après les résultats obtenus sur les grilles, la bissection d'un hypercube de dimension d est de 2^{d-1} (on partitionne l'hypercube en deux hypercubes de dimension $d-1$ en retirant les 2^{d-1} arêtes qui ont permis de les lier).

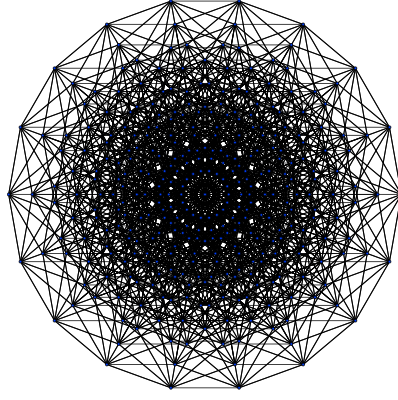


FIGURE 4 – Une projection de l'enneract (9-cube)

4.1 Comparaison des différentes architectures

À partir des résultats précédents, on peut manipuler les équations pour obtenir une estimation du diamètre, du degré des sommets, du nombre d'arêtes et de la bisection des différentes architectures étudiées jusqu'ici, directement en fonction du nombre de nœuds de calcul.

Architecture	Diamètre	Degré	Nombre d'arêtes	Bisection
Graphe complet	1	$n - 1$	$\frac{n(n-1)}{2}$	$\approx \frac{n^2}{4}$
Grille de dimension d	$d(n^{1/d} - 1)$	$2d$	$dn(1 - n^{-1/d})$	$\approx n^{1-1/d}$
Grille de dimension 2	$2(\sqrt{n} - 1)$	4	$2(n - \sqrt{n})$	$\approx \sqrt{n}$
Grille d'arbres bidimensionnelle	•	•	•	•
Hypercube	$\log_2 n$	$\log_2 n$	$\frac{1}{2}n \log_2 n$	$\frac{n}{2}$

FIGURE 5 – Comparaison des architectures pour n nœuds

On constate que l'hypercube offre les meilleurs diamètre et bisection, au détriment d'un degré non constant et d'arêtes plus nombreuses et plus longues par rapport à une grille classique, ce qui implique un câblage plus complexe et coûteux. Cela reste néanmoins une architecture de choix à l'heure actuelle.

4.2 Template de communication

On trouve dans [Fos95] un *template* de communication classiquement utilisé dans de nombreux problèmes SIMD implémentés sur un hypercube :

Procédure 1 hypercube($task_id, n, input, output$)

Input: $task_id, n, input, output$

$state \leftarrow input$

for $i = 0$ **to** $\log_2 n - 1$ **do**

$partner \leftarrow task_id \oplus 2^i$

 send($partner, state$)

$message \leftarrow receive(partner)$

$state \leftarrow OP(state, message)$

end for

$output \leftarrow state$

Cette procédure est exécutée par chacune des n tâches. La tâche $task_id$, dont l'état initial vaut $input$, communique avec ses $\log_2 n$ voisins qu'elle identifie simplement avec une opération XOR sur chaque dimension. Elle envoie son état courant à son partenaire et reçoit de ce dernier le sien. La tâche met à jour son état en effectuant une opération « OP » combinant son état courant et l'information reçue de son voisin.

4.3 Tri bitonique

Dans ce paragraphe, nous décrivons l'algorithme de tri par fusion bitonique, particulièrement adapté aux réseaux de tri et au calcul parallèle. Pour simplifier les explications, on considère un tableau de taille multiple de n , notée $l = Mn = M2^d$ avec M entier. On commence par distribuer le tableau sur les n nœuds d'un hypercube de dimension d , chaque nœud ayant pour état initial un sous-tableau de taille M .

Prenons par exemple le tableau de 24 éléments suivant : [15, 10, 2, 11, 6, 4, 0, 3, 2, 13, 14, 7, 9, 9, 4, 13, 6, 5, 13, 0, 5, 2, 13, 1]. On distribue ce tableau en 8 sous-tableaux de trois éléments, de manière arbitraire sur un hypercube de dimension 3 (figure 6a) puis chaque tâche trie ensuite, de manière séquentielle, son sous-tableau (figure 6b).

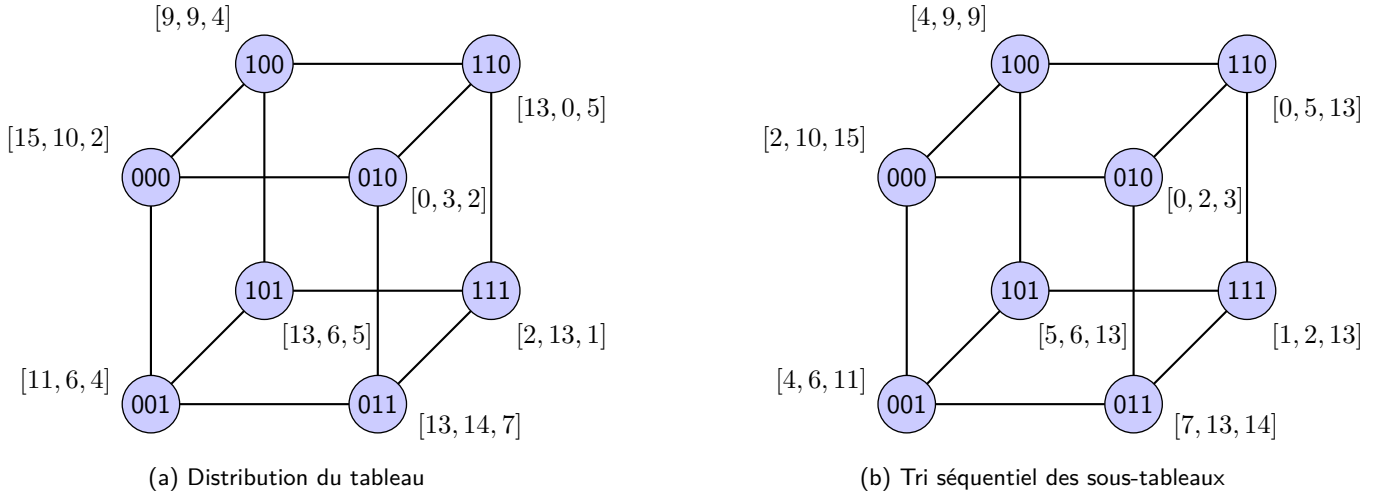


FIGURE 6 – Initialisation du tri bitonique sur un hypercube ($M = 3$)

4.3.1 Compare-split

Nous aurons besoin de faire communiquer les nœuds pour fusionner les sous-tableaux de manière à pouvoir obtenir la séquence triée attendue. L'idée est de maintenir des relations d'ordre sur les états des nœuds : pour des sous-tableaux T_i et T_j vivant respectivement sur les tâches P_i et P_j , on écrira $T_i \leq T_j$ si chaque élément de T_i est inférieur ou égal à chacun des éléments de T_j . On cherchera alors, pour deux nœuds voisins P_i et P_j , à obtenir des sous-tableaux $T'_i \leq T'_j$ en combinant leurs sous-tableaux initiaux T_i et T_j . L'algorithme permettant de réaliser cette opération porte le nom de *compare-exchange* ou *compare-split* et est notamment décrit dans [GGKK02] :

1. Chaque tâche envoie à l'autre tâche son sous-tableau de M éléments, si bien que les deux nœuds disposent d'une copie locale de T_i et T_j .
2. Par un critère bien défini, chacune des tâches sait si elle joue le rôle de P_i (resp. P_j) et fusionnera les deux tableaux T_i et T_j de manière à sélectionner les M plus petits (resp. plus grands) éléments de $T_i \cup T_j$ (*compare-split-high* ou *compare-exchange-low*).
3. À l'issue, les deux tâches contiendront l'un des sous-tableaux T'_i et T'_j avec $T'_i \leq T'_j$ (et bien entendu $T'_i \cup T'_j = T_i \cup T_j$).

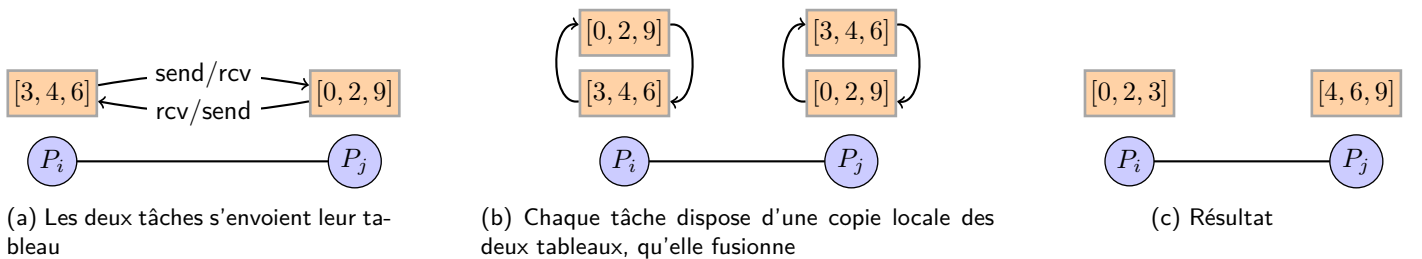


FIGURE 7 – *Compare-split* entre deux tâches P_i et P_j

Les sous-tableaux initiaux étant triés, l'opération de fusion ne présente pas de difficulté et sa complexité est en $O(M)$. Pour cette même raison, il n'est pas toujours nécessaire de disposer des M éléments de l'autre sous-tableau pour obtenir le nouveau sous-tableau et on pourrait imaginer que les tâches s'échangent à la volée les éléments nécessaires à la décision, à condition que les coûts engendrés par le surplus de messages ne soient pas supérieurs aux gains liés à la réduction du volume des échanges.

4.3.2 Parallel merge

4.4 Quelques autres exemples d'algorithmes

Références

- [Fos95] Ian Foster. *Designing and Building Parallel Programs : Concepts and Tools for Parallel Software Engineering*. Pearson Addison-Wesley, 1995. <https://www.mcs.anl.gov/~itf/dbpp/>.
- [GGKK02] Ananth Grama, Anshul Gupta, George Karypis, and Vipin Kumar. *Introduction to Parallel Computing*. Pearson Addison-Wesley, 2nd edition, 2002. <http://parallelcomp.uw.hu/index.html>.