

## PRODUCTION A FLUX TIRES DANS UNE CHAINE LOGISTIQUE

Vincent GIARD\* & Gisèle MENDY BILEK\*\*

---

Résumé. – Le choix d'un pilotage de la production entre une approche à flux tirés et une approche à flux poussés est moins simple qu'il ne le paraît lorsque l'on prend en compte le périmètre des chaînes logistiques. En fait, il s'avère en général nécessaire de combiner ces deux modes de pilotage pour respecter au mieux les délais clients, de plus en plus courts, sans augmenter les coûts. Ces couplages ne sont pas indépendants des caractéristiques de la chaîne logistique, des produits à fabriquer et de la régularité de la demande, et font appel à différentes formes de pilotage. On examinera ici une des approches les plus connues, le système kanban, utilisée avec succès sous certaines conditions, en particulier celle de la stabilité de la demande. On analysera l'usage du système kanban dans des environnements violant les hypothèses classiquement faites dans ce domaine, ce qui permettra de mettre en évidence les conditions de robustesse de ce type de pilotage. La remise en cause du régime de croisière dans le système kanban complique un peu l'analyse des modalités de pilotage des flux sur la chaîne logistique et pousse à imaginer de nouvelles règles de pilotage pour permettre la synchronisation des productions dans la chaîne logistique.

Mots-clés : Pilotage des flux, Chaîne logistique, Production à flux tirés, Production à flux poussés, Système Kanban.

---

\* Professeur des Universités, Université Paris Dauphine, Lamsade, Place de Lattre de Tassigny, 750016 PARIS, vincent.giard@dauphine.fr.

\*\* Doctorante, Renault Direction de la Logistique et Lamsade, Université Paris Dauphine, giselemendy@numericable.fr.

## 1. Introduction

On s'intéresse ici aux chaînes logistiques orientées vers la fabrication et l'assemblage et travaillant pour satisfaire les commandes de clients finaux ou celles d'intermédiaires du réseau de distribution, formulées en réponse à des demandes connues ou anticipées.

Dans ce réseau, on peut distinguer la chaîne logistique principale, constituée d'un enchaînement de processus convergeant vers le processus d'assemblage, qui délivre des produits demandés par les clients, des chaînes logistiques secondaires, qui alimentent en composants les processus de la chaîne logistique principale. Le pilotage de la production de la chaîne logistique principale et celui des chaînes logistiques secondaires doivent être coordonnés pour que la demande finale soit satisfaite (critère d'efficacité) dans des conditions économiques satisfaisantes (critère d'efficience). Le raccourcissement des délais commercialement admissibles, l'accroissement de la variété des produits finis et la tension sur les coûts obligent les acteurs de la chaîne logistique à adapter les pilotages existants, ce qui pose quelques problèmes méthodologiques que l'on examinera ici, ou à en imaginer de nouveaux .

On commencera par examiner comment mieux utiliser la distinction entre production à flux tirés et production à flux poussés dans le pilotage de la chaîne logistique, avant de présenter les principales approches disponibles (§ 2).

Dans le reste de cet article, on se focalisera sur les modalités d'utilisation du système kanban en régime de croisière lorsque certaines hypothèses implicitement retenues ne sont pas respectées (§ 3), avant d'examiner sous quelles conditions on peut conserver ce mode de pilotage des flux lorsque l'hypothèse de stabilité de la demande devient intenable (§ 4).

## 2. Production à flux tirés et à flux poussés dans la chaîne logistique

L'utilisation de la distinction classique entre flux tirés et flux poussés dans le pilotage d'une chaîne logistique conduit à préciser ce qu'il faut entendre par flux tirés (§ 2.1), pour pouvoir apprécier les différents mécanismes de pilotage des flux dans la chaîne logistique. On examinera d'abord celui de la production à la commande (§ 2.2), puis celui de la production pour stock (§ 2.3), le système kanban faisant l'objet de la suite de cet article.

### 2.1 Flux tirés et chaîne logistique

Dans une perspective locale, la qualification de flux tirés ne vaut que pour le processus-fournisseur dont la production n'est déclenchée que par la commande de l'un des processus-clients avec lesquels il est en relation directe. Dans la perspective globale, qui est normalement celle de la chaîne logistique, on peut réserver cette qualification au pilotage des flux d'un

enchaînement de processus allant jusqu'à celui satisfaisant la demande d'un produit fini, considérée comme le fait générateur de la production dans cette chaîne logistique. Dans ce cas, la production d'un processus n'est pas nécessairement déclenchée par la commande de l'un des processus-clients avec lequel il est en relation directe. Il est alors nécessaire de préciser si c'est le point de vue local ou le point de vue global que l'on retient en parlant de flux tirés.

On parle de flux poussés lorsque la production d'un processus est décidée sur la base d'une anticipation, et non en réponse à une commande passée par l'un de ses processus-clients. Cette anticipation peut être le fruit d'une prévision portant sur des demandes non encore formulées (demande potentielle), auquel cas, on est clairement en production pour stock. Elle peut aussi résulter d'un calcul de mise à disposition à temps de composants à un processus-client, pour rendre possible la livraison de produits à une date ultérieure conformément à une commande effective déjà passée par un client du dernier processus de la chaîne logistique considérée ; ce dernier processus ne pouvant être le processus-client du processus considéré ici. On est alors localement dans une logique de flux poussés et, globalement, dans une logique de flux tirés.

Lorsque la demande ne peut être satisfaite par un stock de produits disponibles et que le délai commercial  $C$  est supérieur au délai de production  $P$ , la production peut s'effectuer entièrement à la commande (§ 2.2). Cependant, si l'on est en présence d'une production très fortement diversifiée s'appuyant sur la différenciation retardée, on peut avoir intérêt, pour limiter le nombre de lancements en fabrication, à produire pour stock les composants, communs à beaucoup de produits finis, dont le niveau dans la nomenclature est éloigné du niveau 0 des produits finis. Lorsque la diversité est plus faible et que la demande de chaque référence porte sur des volumes suffisants, il peut également être économiquement préférable de travailler pour stock (§ 2.3).

Ajoutons enfin que, dans certains secteurs, les chaînes logistiques principales ont tendance à transmettre plus tôt leurs besoins aux chaînes logistiques secondaires pour raccourcir leurs délais de réponse. Le déplacement du point de pénétration de commande qui en résulte permet d'accroître le poids relatif de la production à la commande par rapport à celle pour stock et permet d'améliorer l'efficacité et l'efficience sous certaines conditions (Giard et Mendy, 2006a).

## ***2.2 Cas de demandes satisfaites par une production à la commande***

On est alors en présence d'un pilotage localement à flux poussés car les composants de niveau  $k$  ne sont pas mis en production à la suite d'une réquisition du centre qui les consommera pour fabriquer des composants de niveau  $k-1$ , mais sur celle d'une anticipation (OF). On est également en présence d'une logique globalement à flux tirés car c'est la livraison

au client du produit demandé (référence de niveau 0) qui déclenche l'ordonnancement de la production. Dans la figure 1 :

- La commande est lancée au plus tard en production à la date  $t$  pour être livrée à la date  $t + P$  et, d'une manière générale, le lancement en production des références de niveau  $k$  est déclenché pour permettre ensuite de lancer en production des références de niveau  $k-1$ .
- La séquence de processus visualisée correspond implicitement à celle du chemin critique de la chaîne logistique secondaire (qui définit le délai  $P$ ). Le composant de niveau  $k$  est implicitement intégré dans toutes les références de la chaîne, du niveau  $k$  au niveau 0 ; cette convention sera toujours utilisée. La chaîne logistique secondaire est souvent une arborescence de processus convergeant vers le processus de rang 0 ; la visualisation d'un seul de ces chemins sur la figure 1 ne restreint pas la généralité des observations faites ici.

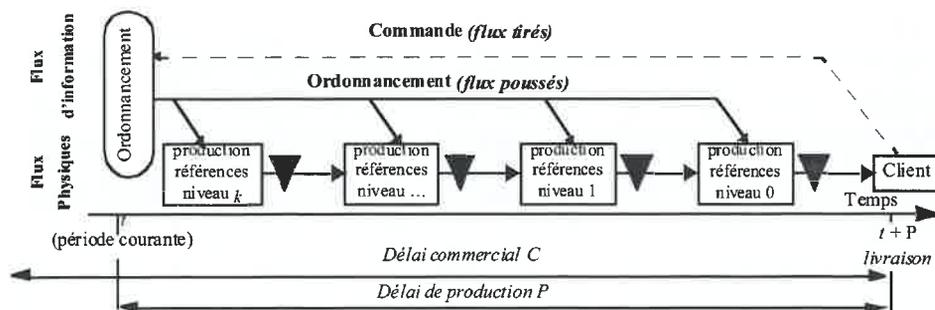


Figure 1 : Pilotage des flux – cas du délai commercial  $>$  délai de production- flux poussés

Pour satisfaire la demande, il faut définir une programmation cohérente de la production des différents composants rentrant dans la fabrication du produit commandé (pilotage à flux poussés). Des ordres de fabrication (OF) de ces composants sont créés par une entité chargée de l'ordonnancement, en tenant compte de contraintes d'antécedence découlant des nomenclatures arborescentes, des gammes de production et des ressources disponibles dans les différents processus.

- Si le délai de production  $P$  est long et que le système productif de la chaîne logistique concernée ne traite simultanément qu'un nombre faible de produits, les techniques d'ordonnancement de projets sont mobilisables.
- Si le nombre de produits fabriqués est plus important, les techniques de MRP peuvent être appropriées, le plan directeur de production correspondant au carnet de commandes fermes. Dans ce cas, l'approche est hiérarchisée avec une planification glissante de la production basée sur une maille temporelle lâche et un niveau de

détail assez grossier dans l'ajustement charge-capacité, suivie d'un ordonnancement détaillé pour la mise en œuvre des décisions prises par la MRP (livraisons programmées pour la fin de la première période du programme de production), dans chacun des centres de ressources concernés.

- Si le délai de production est assez court et le nombre de références important, on retrouve une problématique d'ordonnancement classique. Le plus souvent, le problème d'ordonnancement posé est trop complexe pour qu'une solution puisse être directement trouvée. Lorsque les processus sont fiables, deux approches sont mobilisables. La première consiste à procéder à une décomposition judicieuse du problème initial en plusieurs problèmes hiérarchisés (la solution d'un problème devenant une contrainte pour d'autres problèmes) en tenant compte, notamment, des configurations de production (*job shop, flow shop, ligne...*) à mobiliser, des enchaînements de processus, de la rapidité d'obtention d'une solution satisfaisante, etc. La seconde approche consiste à procéder à un ordonnancement collaboratif sur une chaîne logistique avec une politique de partage total de l'information de la part des différents acteurs. Différents travaux mettent en évidence le gain global réalisable en adoptant un comportement collaboratif (Cachon et Zipkin, 1999). Si les processus mobilisés ne sont pas fiables, le problème devient plus compliqué avec un besoin d'adaptation des techniques d'ordonnancement et des approches nouvelles comme l'approvisionnement synchrone et la production synchrone (Giard et Mendy, 2006c).

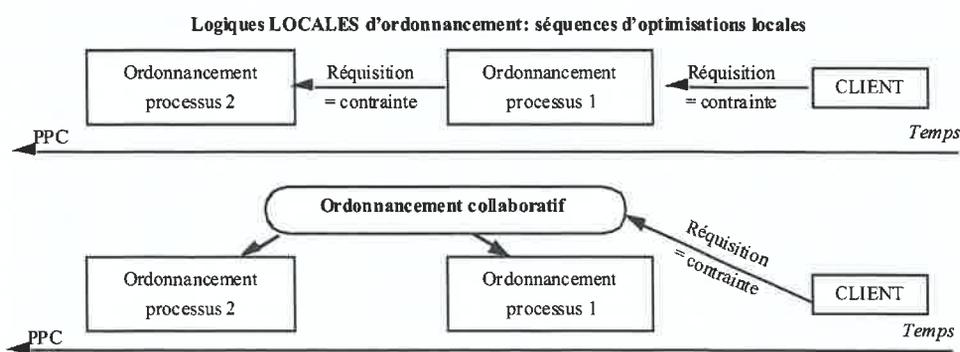


Figure 2. Ordonnancement en univers certain : deux approches possibles

La production envisagée ici peut être exécutée dans des chaînes logistiques secondaires faisant intervenir des entreprises différentes, ce qui entraîne parfois des problèmes de cohérence décisionnelle. Prenons le cas d'un appel aux techniques de MRP. Si c'est la même entreprise qui traite l'ensemble de la production, la mise en œuvre d'une démarche de type MRP-II permet de procéder, si nécessaire, à des ajustements charge-capacité. Dans le cas contraire, cet ajustement peut être difficilement envisageable, provoquant l'incohérence des décisions de programmation

de la production et une absence d'un mécanisme centralisé d'ajustement charge-capacité. Cette situation est observable dans des entreprises ayant choisi d'externaliser des activités de production pour se focaliser sur celles de conception et d'assemblage et qui sont organisées par centre de responsabilité par famille de produits. Ces centres peuvent alors passer des commandes aux mêmes fournisseurs sans aucune concertation préalable. Si ces fournisseurs ne disposent pas de marges de négociation, le délai associé à certaines commandes peut être impossible à tenir en raison d'une surcharge de travail. Les retards de production qui en résultent sont alors généralement imputés à une non-fiabilité des fournisseurs au lieu de l'être à une incohérence décisionnelle doublée d'une forte asymétrie dans les rapports clients-fournisseurs.

### 2.3 Cas de demandes satisfaites par une production pour stock

L'information de prélèvements dans un stock peut être exploitée par l'une des approches à flux tirés par des gestions d'approvisionnement à reapprovisionnement périodique ou à point de commande (Giard, 2003) mais aussi en s'appuyant sur un système-kanban (§ 2). L'information relative à la demande finale peut aussi être immédiatement répercutée vers tous les processus-amonts sur une chaîne logistique de distribution et être exploitée localement dans une approche à flux poussés, approche que nous présentons dans cette partie.

Parallèlement à l'émergence du système kanban au Japon, un système de gestion multi-échelons connu sous le nom de *base stock system*, d'abord utilisé dans les réseaux de distribution, est apparu aux États-Unis. Il partait du constat que la gestion indépendante de chaque nœud du réseau conduit à transformer des demandes relativement régulières des clients en des demandes de plus en plus irrégulières au fur et à mesure que l'on remonte vers le centre de production, en raison de réapprovisionnements par lots déclenchés à chaque nœud du réseau. Le système de production doit alors se dimensionner pour absorber des à-coups importants sans pouvoir éviter à certains moments des ruptures de stock. Cet effet amplificateur d'irrégularités lié aux pratiques de lotissement était dénoncé en production par Orliky, l'un des fondateurs de la MRP (Orliky, 1975). Le système du *base stock* s'appuie sur deux principes (Vollmann et al., 1997) :

- Le système d'information de la chaîne logistique est conçu pour propager immédiatement de l'aval vers l'amont toute information relative aux prélèvements dans le stock d'une référence d'un produit fini, pour en faciliter le réapprovisionnement ; on est donc dans un système globalement à flux tirés. À chaque nœud du réseau de distribution, sont agrégées des informations de consommation des nœuds qu'il alimente directement ou indirectement ; la perception des besoins à satisfaire n'est alors pas biaisée par les règles de lotissement. Arrivée au nœud correspondant au dernier processus de l'usine de production, cette information permet une programmation plus efficiente de la production de cette référence. Ce pilotage

reste compatible avec le système « classique » de réapprovisionnement tiré par des commandes remontant dans le réseau de l'aval vers l'amont. Dans ce cas, seule la production tire pleinement parti de cette remontée d'informations en temps réel. Il rend possible, également, une meilleure organisation de la logistique d'approvisionnement des différents nœuds du réseau de distribution en s'appuyant sur une logique de recomplètement des stocks ; c'est le second principe de ce système.

- Le pilotage de l'approvisionnement d'une référence  $i$  à l'un des nœuds  $k$  du réseau se fonde sur un stock de base  $S_{ik}$  (le *base stock*) qui correspond à un niveau de recomplètement permettant de satisfaire la demande entre deux réapprovisionnements avec un certain niveau de service (probabilité de non-rupture de stock ou pourcentage de demandes immédiatement satisfaites). Il est alors possible d'organiser les réapprovisionnements pour recompléter en temps utile les stocks, ce qui oblige à tenir compte de nombreux paramètres (coûts des tournées d'approvisionnement, organisation de ces tournées, possibilité de partage d'un acheminement par plusieurs références et plusieurs destinations, encombrements et poids des références...). Le paramétrage des stocks de base est donc une opération délicate, même si les caractéristiques des demandes aléatoires sont stables. Une gestion calendaire des réapprovisionnements conduit à acheminer périodiquement des quantités variables destinées à recompléter les stocks, ce qui peut compliquer le transport (risque d'insuffisance de la capacité du véhicule). D'autres politiques d'approvisionnement sont possibles, comme celles consistant à déclencher un réapprovisionnement dès que la position de stock d'une référence descend en dessous d'un certain seuil (point de commande) et à compléter le chargement du véhicule avec d'autres références dont la position s'approche du point de commande ; leur analyse est scientifiquement plus complexe.

La demande des différentes références ainsi gérées peut être aléatoire à condition que les caractéristiques en soient stables. Dans le cas contraire, se pose le délicat problème de la révision périodique du paramétrage des stocks de base, problème que l'on retrouvera dans le système kanban.

En production, l'application du second principe du système *base stock* se heurte aux critiques classiques relatives à l'inadéquation des modèles de gestion des stocks pour piloter la production de produits complexes diversifiés s'appuyant sur une nomenclature à plusieurs niveaux, critiques à l'origine du courant de la MRP. Cette approche est utilisable à condition que la demande des différentes références de produits finis soit suffisamment régulière pour permettre de stabiliser les stocks de base des différents composants. Fort de ce constat, Karaesmen *et al.* (2002) proposent, dans le cas de systèmes capacitaires, une politique de pilotage utilisant l'information anticipée de la demande qui est une extension de la politique du Base

stock. Le modèle proposé est un système de production à étages ; la variable temps est discrète et la visibilité de la demande est de  $H$  périodes. La mise à disposition de la demande anticipée permet d'ajuster le déclenchement de la production grâce à un paramètre noté  $L$ , lié au temps d'écoulement du système. Cette politique présente l'avantage de distinguer clairement la visibilité de la demande (qui peut être exogène) et le paramètre de déclenchement  $L$  qui est choisi en fonction des caractéristiques du système et de la demande. Il permet également de prendre en compte des commandes dont les délais sont variables. En contrepartie les limites de ce modèle sont liées au problème de variabilité des temps d'écoulement qui complique nettement la relation avec  $L$ .

### 3. Le système kanban en régime de croisière

Le système kanban peut être utilisé pour piloter aussi bien tout ou partie d'une chaîne logistique principale que de chaînes logistiques secondaires. Après avoir rappelé les principales caractéristiques du système kanban (§ 3.1), on formulera une série de remarques sur son usage dans la chaîne logistique (§ 3.2).

#### 3.1 Les caractéristiques du système kanban

Le kanban est un ordre de fabrication ouvert portant sur une référence ; il se trouve soit attaché à un conteneur plein ou en cours de remplissage, soit « autonome » lorsqu'il revient au centre de production ou qu'il s'y trouve en attente d'utilisation. Son usage s'est progressivement généralisé dans le courant des années quatre-vingt (Hall, 1983).

En régime de croisière, le nombre d'unités correspondant au nombre de kanbans d'une référence mis en circulation est normalement supérieur à celui du stock observable de cette référence à un instant donné. C'est seulement en cas d'arrêt prolongé de la production par le centre de consommation de cette référence que ce stock pourra atteindre le stock maximal défini par l'ensemble des kanbans de cette référence, sans pouvoir le dépasser. Le système kanban permet une bonne coordination entre les sous-systèmes et n'autorise pas la production à chaque étage d'une pièce tant que celle-ci n'a pas été effectivement consommée en aval de ce sous-système. Une conséquence intéressante de ce mécanisme est que le nombre de pièces en cours de fabrication est limité par le nombre de cartes de chaque étage. Cependant, la demande ne pouvant pas être immédiatement transmise à toute la chaîne, les délais de livraison des pièces peuvent être plus longs que pour la politique du *Base Stock*. Le cas particulier de la politique kanban est aussi connu sous le nom de CONWIP (CONstant Work In Process). La « myopie » de ce mode de pilotage implique le respect d'un certain nombre de conditions sur l'environnement et le paramétrage du système, comme on va le voir. On s'intéressera ici au cas de centres de production fabriquant un composant optionnel (toit ouvrant...) ou un ensemble de modules

substituables parmi lesquels on doit en choisir un pour le monter dans le produit fini (boucliers de couleurs différentes à monter sur un véhicule automobile...). Par ailleurs, l'analyse du pilotage des flux par le système kanban que l'on va effectuer ici repose sur l'hypothèse du refus d'une rupture de stock. La remise en cause du régime de croisière oblige alors à prendre en compte la variation de niveau de la demande et/ou celle de sa structure. La mise en avant de caractéristiques du système kanban en régime de croisière (§ 3.1) permet d'aborder son adaptation lorsque le régime de croisière change de manière occasionnelle ou systématique (§ 4.1).

Lorsqu'on est « très proche » d'un régime de croisière, la stabilité des flux permet de piloter la chaîne logistique à flux tirés en s'appuyant sur le système kanban. Le paramétrage de ce système de pilotage des flux, qui relève d'une « fonction ordonnancement », consiste à déterminer le nombre d'étiquettes à mettre en circulation à chaque niveau. Il se base principalement (tableau 1) sur l'anticipation de la demande exprimée par un plan directeur de production et relève donc d'une logique de flux poussés. Ce système hybride est illustré à la figure 3 dans laquelle les commandes fermes sont celles du jour suivant.

Paramétrage (détermination du nombre de kanbans)	Utilisation du système Kanban
- Demande : demande anticipée (moyenne quotidienne de la demande du PDP) - <i>flux poussés</i> - Autres déterminants : <ul style="list-style-type: none"> <li>. durée du cycle,</li> <li>. niveau de qualité de la production,</li> <li>. nombre de références différentes fabriquées dans le centre de production,</li> <li>. taille du lot associé au kanban.</li> </ul>	- Demande effective (pilotage par kanbans) - <i>flux tirés</i>

Tableau 1 : Caractéristiques du système kanban classique.

Le nombre de kanbans  $N$  est proportionnel à la demande quotidienne  $D$  et à la durée  $\theta$  du cycle d'un kanban et inversement proportionnel à la taille  $q$  du lot du kanban ( $N = \theta.D/q$ ). La diminution de la durée du cycle  $\theta$  est liée à des actions d'amélioration continue portant, notamment, sur l'organisation logistique des flux (circulation des conteneurs et des kanbans). La taille  $q$  du lot d'un kanban est le résultat d'un compromis entre des coûts de lancement en fabrication et des coûts logistiques (possession du stock et transport) en tenant compte, pour les centres fabriquant plusieurs références, de l'incidence du lotissement (le lancement portant alors sur un nombre minimal de kanbans) sur la flexibilité et la réactivité de ces centres.

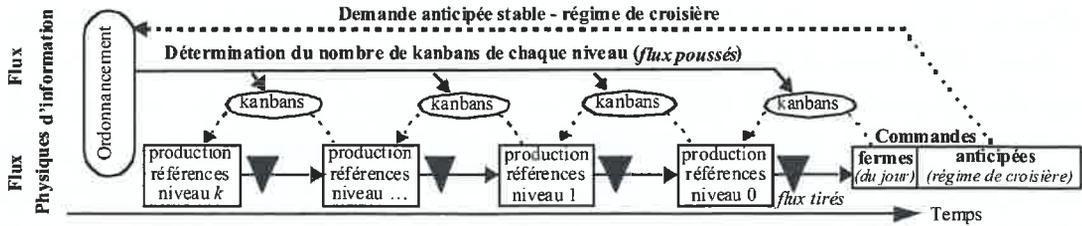


Figure 3 : Pilotage des flux – Cas du délai commercial > délai de production – flux poussés.

### 3.2 Remarques sur les conditions de fonctionnement du système kanban « classique »

Cinq remarques doivent être formulées.

Tout d’abord, la formule de calcul est un peu plus compliquée lorsqu’un centre de production alimente plusieurs centres de consommation (adaptation du système avec une dissociation entre kanbans de production et kanbans de transfert) et/ou que le centre de production fabrique plusieurs références. Sur le fond, le nombre de kanbans dépend toujours des demandes, des temps de cycle mais s’y ajoute, en cas de production multiple, des kanbans additionnels correspondant à des stocks de sécurité. Ce point est bien connu, contrairement aux quatre suivants.

En cas d’incident prolongé au centre de consommation de la référence, le stock de cette référence dans ce centre devient égal à  $N$  conteneurs, en raison des modalités de fonctionnement du système kanban. Ce stock maximal possible découlant de la relation de calcul rappelée ci-dessus peut ne pas être acceptable en raison de la place disponible au centre de consommation de la référence. Dans ce cas, le nombre d’étiquettes est prédéterminé par l’espace disponible et, la demande étant exogène, on peut agir sur le temps de cycle  $\theta$ , qu’il convient de réduire. On peut également agir sur  $q$ , si le conteneur peut accueillir plus de  $q$  unités et si cette variation permet de diminuer le nombre maximal de conteneurs présents en cas d’incidents prolongés. On est alors en présence d’un changement de variable de commande.

Troisième remarque, on peut utiliser le système kanban pour approvisionner un centre de consommation en références différentes, intervenant dans la diversification d’un produit (couleurs d’un même composant, par exemple). Dans ce cas, le calcul des nombres  $N_i$  de kanbans associés à ces références  $i$  repose souvent sur l’hypothèse d’un lissage de la consommation de chacune de ces références. Cette hypothèse n’est pas toujours réaliste, notamment lorsque le poste de consommation se trouve sur une ligne d’assemblage, pour deux types de raisons. Tout d’abord, la prise en compte de certaines contraintes dans l’ordonnancement des produits à traiter sur certains postes de la ligne est difficilement compatible avec l’objectif d’une régularité dans la consommation de certains composants (Giard & Jeunet, 2005 et 2006c). Par ailleurs, l’ordonnancement trouvé, qu’il assure ou non ce lissage,

sera perturbé par des problèmes d'approvisionnement ou de qualité, impliquant dans ces deux cas un retrait momentané de la ligne avant réinsertion (Danjou, Giard & Le Roy, 2000 ; Boctor, Danjou & Giard, 2001). Pour prévenir toute rupture d'approvisionnement, le taux de demande à utiliser dans le calcul ne doit pas reposer sur l'hypothèse de lissage de la consommation de ces références mais plutôt sur celle d'une concentration de la demande. Supposons, par exemple, que le taux de consommation horaire soit de 20 unités/heures et que 25 % des références consommées correspondent à la référence  $i$  ; l'hypothèse de lissage conduira à calculer  $N_i$  à partir d'une demande horaire  $D_i = 5$  tandis que le scénario le plus défavorable se produit lorsque toute la consommation de la référence  $i$  est concentrée en une séquence ininterrompue, ce qui implique un calcul de  $N_i$  à partir d'une demande horaire  $D_i = 20$ . En conséquence, si l'on veut se prémunir contre tout risque de rupture de stock, il faut, dans le calcul du nombre de kanbans, remplacer la demande horaire  $D_i$  par la demande horaire totale<sup>1</sup>  $D = \sum_h D_h$ .

Quatrième remarque, on peut également utiliser le système kanban sur une chaîne logistique avec l'introduction de kanbans multiples liés à une prise en compte d'enchaînements de processus différents. Dallery et al. (2005) s'intéressent à ce type de problème pour un système de kanban fonctionnant en régime de croisière sur une ligne d'assemblage avec différentes modalités de contrôle, dépendant de la façon dont les kanbans des composants en attente d'assemblage sur un poste sont libérés, la taille du lot de ces kanbans étant unitaire, la rupture de stock étant acceptée et la demande finale étant liée à l'état du système. Dans un premier scénario, on libère simultanément les kanbans lorsque tous les composants nécessaires pour l'assemblage sont disponibles et que le poste d'assemblage est prêt. Dans le second scénario, on libère les kanbans en prélevant les composants sans se préoccuper du fait que tous les composants soient disponibles ou non. Les résultats analytiques mettent en évidence la supériorité d'une gestion indépendante de libération des kanbans lorsque la demande tend à saturer le système.

Considérer la durée du cycle  $\theta$  comme certaine est une hypothèse a priori peu réaliste ; c'est notre cinquième remarque. Trois sources d'aléa doivent être distinguées :

- La première est liée à l'irrégularité de la consommation qui se traduit par un aléa sur l'intervalle de temps séparant deux libérations successives de kanbans par un même centre de consommation de la référence. Cette irrégularité est amplifiée par le nombre

<sup>1</sup> Ceci revient à multiplier le nombre de kanbans d'une référence  $i$  calculé sous l'hypothèse d'un lissage de la consommation par l'inverse du pourcentage que représente la référence dans la consommation totale. Si ces  $n$  différentes références ont le même temps de cycle  $\theta$ , le nombre de kanbans lié à la diversité est alors  $n$  fois celui que l'on obtient en l'absence de diversité si la demande globale est stable, dès lors que le lissage de chacune des demandes ne peut être garanti et que toute rupture de stock est interdite. Cette observation permet de calculer l'espace de stockage qu'il faut avoir en amont d'un poste approvisionné par le système kanban lorsque la stabilité de la structure de la demande (réputée constante) n'est pas garantie et que l'on refuse toute rupture de stock.

limité de kanbans disponibles à chaque station. Pour analyser de tels systèmes, certains auteurs développent des méthodes exactes reposant sur l'analyse d'une chaîne de Markov, soit analytiquement, soit numériquement, tandis que d'autres développent des méthodes approximatives (Di Mascolo et al, 1996). La plupart de ces approches analytiques se limitent au cas de stations en série assurant la production d'un même produit.

Pour l'ajustement du nombre de kanbans, Tardif et Maaseidvaag (2001) proposent un mécanisme adaptatif de contrôle. Leur procédure est basée sur l'arbitrage entre niveau de stock de produits finis et importance des ruptures de stocks, l'objectif étant de minimiser la somme des coûts des encours et des désamorçages. Dans leur procédure, si une demande arrive et si le niveau de stock est en dessous d'un seuil de libération prédéterminé, un kanban supplémentaire est introduit dans le système ; dès que le stock redescend à ce seuil, les kanbans additionnels sont retirés. Il a été démontré que ce mécanisme est plus performant qu'un système kanban classique pour une demande suivant une loi de Poisson et pour des temps de cycles exponentiels. Di Mascolo (1989) utilise comme indicateurs de performance le temps d'attente moyen d'une demande et la proportion de demandes non satisfaites (ruptures de stock) ; les temps de service de chaque station et les lois d'arrivées des demandes sont représentés par des distributions de type phase (combinaison de distributions exponentielles) qui permettent de caractériser le comportement du système par une chaîne de Markov. L'approche Markovienne est liée au régime stationnaire et son usage repose, notamment, sur la possibilité de détecter immédiatement sa modification.

- À l'irrégularité de la demande s'ajoute souvent celle des temps de transport des conteneurs ou des kanbans isolés et celle des temps opératoires. Pour se ramener à l'univers certain, on peut prendre les bornes supérieures de ces temps de transport et temps opératoires ayant une probabilité considérée comme négligeable d'être dépassées, ce qui revient à multiplier le temps de cycle moyen par un coefficient de sécurité, et donc à introduire dans le système des « kanbans de sécurité ». Cela étant, le problème majeur est celui de l'aléa sur les temps opératoires car il joue sur l'ajustement « charge-capacité » des centres de production ; dans le cas d'un centre spécialisé dans la production d'un seul produit, il est facile de calculer le nombre d'unités produites pendant une période de production de plusieurs heures ou dizaine d'heures, qui a la probabilité minimale d'être dépassée, si le centre ne peut pas être en rupture de travail (solution analytique ou utilisation du théorème de la limite centrale). Plusieurs auteurs se sont intéressés à ce problème.

Gupta et Al Turki (1997) proposent un système kanban flexible (Flexible Kanban System) qui s'appuie sur un algorithme d'ajustement dynamique du nombre de kanbans de centres de production travaillant en série et dédiés à la fabrication d'un

produit unique, avec un temps de cycle global de fabrication inférieur à la journée. Dans ce système, les temps opératoires sont aléatoires (lois Normales), les temps de transfert sont certains, la demande, connue à l'avance, varie dans une plage de valeurs prédéterminée et la durée quotidienne de travail est fixe. Pour éviter toute rupture de stock, on adapte à la hausse le nombre de kanbans de telle sorte que la capacité de n'importe quelle séquence de jours à venir reste supérieure à la durée de production de la demande totale sur cette séquence ayant une probabilité négligeable d'être dépassée. L'algorithme proposé est intéressant, mais il suppose une anticipation suffisante de la demande pour pouvoir piloter la variation du nombre de kanbans sans avoir à faire varier la capacité disponible et, par ailleurs, il élude les problèmes de la propagation de ces variations le long de la chaîne logistique, avec l'hypothèse implicite d'un temps de cycle très court, et de la diversité.

Autre exemple, Philipoom *et al.* (1987b) proposent une méthodologie pour un ajustement dynamique du nombre de kanbans en cas de variations de la demande pour un centre de production travaillant en série et pouvant traiter différents types de produits. Le principe sur lequel repose la méthodologie exploite l'équation classique de Monden ( $N = \theta.D/q$ ). Les temps de cycle prennent en compte les temps d'utilisation du centre de production, la variabilité des temps de production, le nombre d'opérations et les temps de transfert. Périodiquement (tous les mois), le nombre de kanbans pour un centre donné est ajusté sur la base de la demande prévisionnelle pour le mois suivant et sur les observations collectées des temps de cycles obtenus le mois précédent. Les temps de cycle ne peuvent pas être statistiquement estimés sur des données empiriques car les paramètres de production pour un centre donné évoluent avec le temps. En conséquence il est nécessaire dans un premier temps, d'établir selon la méthode de Box et Jenkins (1976) la fonction d'autocorrélation des temps de cycle sur la période passée et d'interpréter les résultats de corrélogramme. Ces résultats permettent de déduire le nombre de périodes à partir duquel les observations sont statistiquement indépendantes ; ceci leur permet de construire leur échantillon sur une centaine d'observations. Dans un second temps, la démarche consiste à estimer une fonction de densité des temps de cycles à partir d'histogrammes. L'objectif final est de déterminer le nombre de kanbans ajusté à partir de la demande prévisionnelle et de la fonction de densité. La variation de la demande repose sur l'hypothèse de fluctuations mineures avec une demande quotidienne fluctuante mais un taux horaire de la demande constant (ajustement possible du temps de travail sur une journée). Pour déterminer le nombre de kanbans sur la période suivante, la fonction de densité associée à la demande prévisionnelle va permettre de définir un nombre de kanbans Mini et Maxi. Un arbitrage économique entre coûts de possession et coûts de ruptures de stock permettra de retenir le nombre optimal de kanbans. Cette méthodologie est intéressante mais elle suppose une connaissance des

données de temps de cycle sur une période suffisamment importante (de l'ordre de 25 jours dans l'étude de cas). Pour des centres de productions connaissant de fortes perturbations internes (qualités, pannes) avec des changements rapides, l'ajustement du nombre de kanbans mensuel n'est pas suffisant. L'usage des approches de Box et Jenkins suppose la possibilité d'exploiter des historiques longs caractérisés par une certaine stabilité des caractéristiques des composantes des chroniques étudiées, ce qui peut limiter l'intérêt opérationnel de cette approche.

- La troisième source d'aléa est liée à la qualité de la production, problème que l'on supposera détecté par un contrôle effectué au centre de production juste après la fabrication du lot. Pour un lot de  $k$  composants à fabriquer (taille du lot du kanban), chaque composant ayant une probabilité  $p$  d'être défectueux, il faut lancer en production un nombre  $Q \geq k$  de composants, sachant que le nombre  $X$  de composants acceptés au contrôle suit la loi binomiale  $B\{Q; (1-p)\}$  sous l'hypothèse d'indépendance de l'occurrence de ces problèmes de qualité. On notera  $Y = Q - X$  le nombre de composants rejetés au contrôle. La distribution de probabilité du nombre  $Q$  de composants à fabriquer pour ne pas avoir moins de  $k$  composants corrects, et donc plus de  $Q - k$  composants défectueux, suit une loi Binomiale Négative, ce qui permet facilement de calculer la valeur la plus faible de  $Q$  permettant de ne pas dépasser un risque prédéterminé de ne pas disposer de  $k$  composants corrects. Le tableau 2 illustre cette gestion du risque, en explicitant l'usage simultané de lois Binomiales et de la loi Binomiale Négative.

Taille $Q$ du lot lancé en production	5	6	7	8	9
Nombre maximal admissible de composants défectueux $Q - \kappa = Q - 5$	0	1	2	3	4
Probabilité de disposer d'au moins $\kappa = 5$ composants corrects dans un lot de $Q$ composants lancé en production $P(X \geq \kappa) = P(Y \geq 5) = P(Y \leq Q - \kappa) = P(Y \leq Q - 5) =$ $Q - \kappa$ $\sum_{i=0}^{Q-\kappa} \frac{Q!}{(Q-i)! \cdot i!} p^i \cdot (1-p)^{Q-i}$ (loi Binomiale $\mathcal{B}(Q; (1-p))$ )	59,05%	88,57% <sup>1</sup>	97,43%	99,50%	99,91%
Probabilité de ne pas disposer d'au moins $\kappa = 5$ composants $P(X < \kappa) = P(X < 5) = P(Y > Q - \kappa) = P(Y > Q - 5)$	40,95%	11,43%	2,57%	0,50%	0,09%
Probabilité $\alpha$ de devoir lancer exactement $Z = Q$ composants pour en avoir $\kappa = 5$ conformes <sup>2</sup> (loi Binomiale Négative)	59,05%	29,52% <sup>3</sup>	8,86%	2,07%	0,41%

$$1. = \{6!/(6! \cdot 0!) \cdot 0,1^0 \cdot 0,9^6 + \{6!/(5! \cdot 1!) \cdot 0,1^1 \cdot 0,9^5$$

2. Pour  $Q = \kappa$ ,  $\alpha = P(X = \kappa)$ ; pour  $Q > \kappa$ ,  $\alpha = P(Y \leq Q - \kappa) - P(Y \leq Q - \kappa - 1)$ ; on est en présence de la loi Binomiale Négative, le calcul de  $\alpha$  étant directement donné par la formule  $P(Z = Q) = \frac{(Q-1)!}{(n-1)! \cdot (Q-n)!} (1-p)^n p^{Q-n}$

$$3. 29,52\% = 88,57\% - 59,05\% = \{5!/(4! \cdot 1!) \cdot 0,9^5 \cdot 0,1^1$$

Tableau 2 : Évolution du risque de ne pas disposer de  $k = 5$  composants corrects dans un lot de  $Q$  composants.

Avec  $p = 10\%$  et  $k = 5$ , il n'y a que 0,5 % de chances d'avoir besoin de lancer en production un lot de plus de  $Q = 8$  composants pour obtenir au moins 5 composants corrects. Le nombre de composants corrects d'un lot de taille  $Q \geq k$  pouvant être supérieur à la taille requise pour le lot, les composants excédentaires doivent être stockés provisoirement dans un conteneur qu'il faudra compléter lors du prochain lancement en production. Dans l'exemple numérique précédent, les calculs ont été effectués sous l'hypothèse d'un stock nul de composants au moment du lancement. Supposons qu'après le précédent lancement en production, un excédent de 2 unités ait été obtenu, il faut alors disposer de  $k = 5 - 2 = 3$  nouvelles pièces après le prochain lancement. Le lancement de  $Q = 6$  composants conduit alors à un risque de 0,86 % de produire moins de 3 pièces bonnes et donc à ne pas pouvoir compléter le conteneur à 5 unités. Le nombre de pièces à lancer pour compléter un conteneur à  $k$  composants est aléatoire ; il en est alors de même pour le temps de production associé au remplissage d'un conteneur. Comme pour les deux sources d'aléas précédemment identifiées, le risque est pris en charge par une majoration du nombre de kanbans, étant entendu que cette majoration doit être calculée en prenant en compte globalement ces trois sources d'aléas.

Le rejet de l'hypothèse du régime de croisière complique un peu l'analyse des modalités du pilotage des flux dans la chaîne logistique, en particulier lorsque celle-ci traite une production diversifiée et que le temps de cycle global de fabrication d'un produit est long

(plusieurs dizaines d'heures au moins). Dans ce qui suit, on supposera que la capacité est suffisante pour absorber les variations de demande mais aussi celles induites par la variabilité des temps opératoires. Cette hypothèse, pour être levée, impliquerait une analyse plus poussée.

#### **4. Remise en cause du régime de croisière dans le système Kanban**

L'étude des conséquences d'une modification du régime de croisière sur les modalités de pilotage de la production par système kanban (§ 4.1) fournit les bases permettant de voir comment travailler avec ce système dans un environnement instable n'excluant pas une anticipation minimale de la demande (§ 4.2).

##### **4.1 Le système kanban en cas de modification du régime de croisière**

Le changement de régime de croisière ( $D \rightarrow D'$ ) nécessite une adaptation du pilotage des flux, sachant que les impacts induits par une diminution de la demande diffèrent de ceux occasionnés par son accroissement.

Dans le premier cas ( $D' < D$ ), le maintien du nombre de kanbans défini antérieurement pour chaque étage est possible mais il conduit mécaniquement à la création de stocks dormants. On peut aussi diminuer immédiatement le nombre de kanbans de chaque étage, auquel cas l'excédent momentané de stock sera résorbé par une diminution temporaire de la production en dessous du niveau associé au nouveau régime de croisière.

Une variation à la hausse de la demande ( $D' > D$ ) suppose tout d'abord que l'on puisse adapter la production de chaque étage. Elle nécessite ensuite de modifier le paramétrage du pilotage des flux puisque le maintien du nombre  $N$  de kanbans, supposé calibré initialement au plus juste à chaque étage, conduit mécaniquement à des ruptures de stock et donc à un désamorçage des flux dans la chaîne logistique. La diminution du temps de cycle implique une réorganisation souvent difficile, ce qui fait que l'accroissement du nombre de kanbans est normalement le levier d'action retenu. La myopie du système kanban classique ne permet pas de déceler à temps le changement de niveau de production. En effet, si l'on se place en univers certain, le signal d'un changement du niveau de demande est donné par l'accélération des retours de kanbans, qui ne peut être immédiatement perceptible que si le kanban libéré au centre de consommation est immédiatement acheminé au centre de production. L'usage de tournées régulières d'acheminement de kanbans ne permet pas de repérer immédiatement ce changement de rythme de consommation qui ne sera connu que lorsque l'on acheminera un kanban de plus que prévu. En univers aléatoire, la variabilité des dates de libération de kanban et celle de leur durée d'acheminement viennent compliquer cette détection, la question se posant toujours de savoir si l'écart entre ce que l'on observe et ce que l'on attend est imputable à un aléa

ou à un changement de niveau. Le système kanban classique est donc dans l'incapacité de récupérer à temps l'information du changement de régime de croisière. Ce système n'est capable de s'adapter sans provoquer de désamorçage de la chaîne logistique que si une « fonction ordonnancement » récupère à temps l'information pour modifier le nombre de kanbans. La vitesse de propagation de la modification de la demande le long de la chaîne logistique joue dans ce cadre un rôle important.

Cette vitesse de propagation est liée aux temps de cycles  $\theta_k$  des kanbans des processus  $k$  de la chaîne logistique. Pour déterminer le niveau d'anticipation nécessaire au système kanban, on commencera par supposer que l'information du changement de niveau de demande se propage progressivement vers l'amont par le biais d'une modification du nombre de kanbans, réputée possible à tout moment : dès qu'un centre de consommation a connaissance d'un changement de la demande à la hausse ou à la baisse, il adapte le nombre de kanbans<sup>2</sup>, immédiatement en cas de hausse et progressivement en cas de baisse (on ne s'intéressera pas à la baisse dans ce qui suit). Pour mieux comprendre ce mécanisme de propagation, décomposons le temps de cycle  $\theta_k$ , en deux intervalles de temps  $\theta_{1k}$  et  $\theta_{2k}$ , où :

- $\theta_{1k}$  est l'intervalle de temps séparant la libération d'un kanban par le centre de consommation de son arrivée au centre de production  $k$
- $\theta_{2k}$  est l'intervalle de temps séparant l'arrivée de ce kanban au centre de production, de sa libération par le centre de consommation.

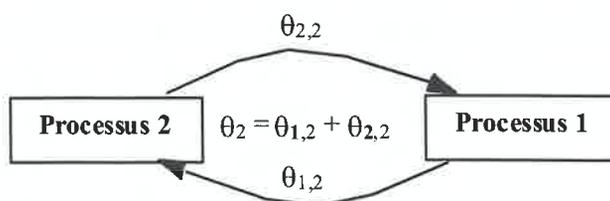


Figure 3 : Décomposition du temps de cycle  $\theta_k$  d'un kanban.

Ce mécanisme de propagation conduit à ce que l'information partant du processus livrant le client final arrive au centre de production du processus  $k$  après l'intervalle de temps  $\Theta_{1k} = \sum_{i=0}^k \theta_{1i}$ . L'augmentation de la production du processus  $k$  va permettre celui du processus  $k - 1$  après le délai  $\theta_{2k}$ . En cascade, l'augmentation de la production du processus 0 s'effectuera

<sup>2</sup> Le nombre de kanbans étant l'arrondi entier supérieur de  $\theta_k \cdot D / q_k$ , noté  $\lceil \theta_k \cdot D / q_k \rceil$ , l'accroissement du nombre de kanbans est  $\lceil \theta_k \cdot D' / q_k \rceil - \lceil \theta_k \cdot D / q_k \rceil$ .

après le délai  $\Theta_{2k} = \sum_{i=k}^i \theta_{2i}$ . En définitive, un système kanban myope réagit avec un retard  $\Theta_k = \sum_{i=0}^i \theta_i$  pendant lequel on ne peut satisfaire la demande ; il faut disposer nécessairement d'une anticipation du changement de régime de croisière au centre de production du processus  $k$  avec un décalage  $\Theta_{2k}$  si on est mesuré de transmettre instantanément l'information le long de la chaîne logistique (ce court-circuit dans la transmission d'information revenant à admettre d'avoir ponctuellement  $\Theta_{1k} = 0$ ). La figure 3 doit être adaptée pour tenir compte de l'anticipation minimale permettant d'éviter tout désamorçage (figure 4).

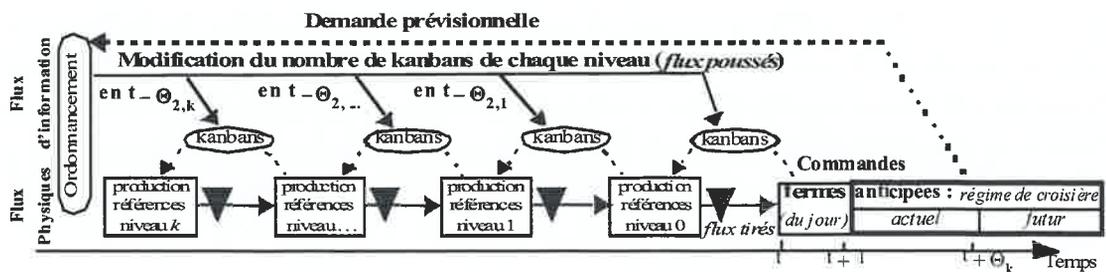


Figure 4 : Pilotage des flux par kanban (sans désamorçage) - changement de régime stationnaire.

Plusieurs remarques complémentaires peuvent être faites.

- Les kanbans additionnels d'un processus ne sont réellement utiles (dimensionnement au plus juste des kanbans) qu'à partir du moment où son processus-fournisseur s'est adapté au nouveau régime de croisière et donc est en mesure de fournir les composants additionnels requis. Le dernier processus en mesure d'utiliser ces kanbans additionnels est le processus final à l'origine de cette propagation. Cet accroissement de kanbans ne génère donc pas de stocks dormants durant le régime transitoire séparant les deux régimes de croisière.
- Le délai d'anticipation  $\Theta_{2k}$  nécessaire pour changer de régime de croisière doit être inférieur au délai commercialement admissible  $C$ . On retrouve ici l'intérêt de la mise en évidence du point de pénétration de commande (Giard et Mendy, 2006a). Bien évidemment, la demande doit être connue sur cet horizon  $\Theta_{2k}$ .
- À la différence du système du *base stock*, l'information répercutée le long de la chaîne logistique correspond à une demande anticipée et non à une demande satisfaite.
- Cette remontée de l'information le long de la chaîne logistique implique une collaboration des acteurs concernés qui ne va pas toujours de soi pour des raisons techniques et organisationnelles.

## 4.2 Le système kanban en régime non stationnaire

L'analyse faite au paragraphe précédent donne les conditions d'utilisation du système kanban dans un régime non stationnaire, si les différents processus sont en mesure de faire varier leur capacité de production. Il faut adapter le nombre de kanbans du processus  $k$  à la période  $t$  (découpage temporel horaire) pour tenir compte de la demande finale qu'il faudra satisfaire à partir de la période  $t + \Theta_{2k}$ . Si le rôle des kanbans reste celui d'un pilotage fin des flux ne nécessitant pas d'ordonnancement, la fonction « ordonnancement » doit cependant être exercée sous une forme particulière consistant à faire varier le nombre de kanbans en anticipation de variations de la demande finale, pour limiter au maximum les stocks sans provoquer de désamorçage dans la chaîne logistique. La figure 4 ne peut plus être utilisée sans convention graphique supplémentaire puisque, à la période  $t$ , le nombre de kanbans du processus  $k - j$  se fonde sur la demande prévisionnelle de la période  $t + \Theta_{2k-j}$ , ce qui implique que le nombre de kanbans de chaque étage s'appuie sur une demande anticipée différente. La limite de ce mode de pilotage est celui de la profondeur de l'anticipation de la demande, ce qui ramène à la problématique du point de pénétration de commande puisque, si l'anticipation est inférieure à  $\Theta_{2k}$ , le processus  $k$  ne peut travailler à la commande et doit, a priori, produire pour stock. Une analyse plus fine du problème posé montre que ce propos peut être nuancé. Si la variabilité de la demande s'exerce sur la structure de la demande et non sur son niveau, le processus permettant de répondre à cette variété (fabrication de boucliers de couleurs différentes, pour reprendre notre exemple) peut faire face à une assez forte variabilité en structure, tout simplement en dimensionnant le nombre de kanbans de chacune des références à des niveaux supérieurs à ce qu'impliquerait une structure stable de la demande. On a montré au § 3.1 que, si l'on veut se prémunir contre tout risque de rupture de stock, il convient de multiplier le nombre de kanbans d'une référence  $i$  calculé sous l'hypothèse d'un lissage de la consommation par l'inverse du pourcentage que représente la référence dans la consommation totale. Cette stratégie permet de faire face à toute modification de structure de la demande et, pendant un certain temps, à une variation du niveau de la demande si les demandes des références  $i$  sont suffisamment lissées : dans ce cas en effet, on dispose du stock de sécurité constitué pour chaque référence pour pouvoir faire face à la variabilité de la structure de la demande.

Des variantes du système kanban ont été proposées<sup>3</sup>. Différents auteurs comme Liberopoulos et Dallery, 2000, se sont basés sur la théorie des files d'attente pour évaluer les performances de la politique kanban dans des systèmes mono-étages et multi-étages. Outre ces politiques, ils ont proposé des extensions de politiques hybrides tirant leurs caractéristiques à la fois de la méthode du base stock et de la méthode kanban. Elles ne considèrent plus les kanbans

<sup>3</sup> Une synthèse de ces approches peut être trouvée dans Baynat *et al.* (2002). Ces approches se fondent explicitement sur la possibilité d'une rupture de stock, hypothèse non retenue ici.

de production comme des ordres de fabrications ouverts : l'activation d'un kanban est liée à l'arrivée d'une information transmise par un processus aval, client direct (approche du kanban généralisé mise au point par Buzacott, 1989) ou indirect (approche du kanban étendu développé par Dallery et al., 1996) du centre de consommation, indiquant que cette référence, intégrée à une autre référence consommée en aval, vient d'être consommée. Implicitement, on est en présence d'un système à double kanbans : les « kanbans d'autorisation » et les kanbans de production dont le cycle est raccourci<sup>4</sup> passant de  $\theta_k$  à  $\theta_{2k}$  pour le processus  $k$ .

L'analyse du fonctionnement du kanban étendu (l'adaptation au kanban généralisé étant immédiat) est aisée lorsque le kanban de la référence étudiée porte sur un lot unitaire, que les enlèvements se font au plus tôt et que les kanbans de production sont dimensionnés au plus juste. Dans ce cas, si, à l'instant  $t$  le centre de production du processus  $k-n$  prélève un conteneur dans lequel se trouve un composant intégrant la référence  $R$  produite par le centre de production  $k$ , un kanban d'autorisation est immédiatement envoyé aux centres de production  $k-(n-1)$  à  $k$ . Le kanban d'autorisation arrivant dans le centre de production du processus  $k-i$  déclenche la production d'un composant intégrant la référence  $R$  qui arrivera dans le centre de production du processus  $k-j$  ( $j < i \leq n$ ) à la date  $t + \sum_{h=i}^{h=j} \theta_{2h}$ . Par ailleurs, à l'instant  $t$ , le centre de production du processus  $k-n$  prélève un conteneur dans lequel se trouve un composant intégrant la référence  $R$  produite par le centre de production  $k$  à l'instant  $t - \sum_{h=k-n+1}^{h=k} \theta_{2h}$ . Comme dans le kanban classique, le pilotage de la production se fonde sur une logique de remplacement des références consommées et cette anticipation du déclenchement ne modifie pas les risques de rupture de stock encourus : ils ne sont supprimés que si le nombre de kanbans  $N_k$  reste défini pour chaque référence produite par le centre de production  $k$  en considérant que la demande est concentrée exclusivement sur la référence considérée ( $D_R = D \Rightarrow N_k = N$ ). En effet, une séquence de  $H$  produits finis demandés intégrant tous la référence  $R$  déclenche nécessairement une rupture de stock si  $N_k < H \leq N$ . Dans le cas contraire, on disposera, pour chaque référence, d'un stock de sécurité suffisant pour faire face à toute variation de la structure de la demande. L'hypothèse de lancement en production des références pour des quantités supérieures à 1 complique un peu l'analyse de ces variantes du système kanban mais ne change pas la conclusion selon laquelle l'anticipation de l'information ne modifie pas les risques encourus.

<sup>4</sup> Il en est de même lorsque le kanban est électronique, cette dématérialisation neutralisant la partie du cycle.

## 5. Conclusion

La coordination de la production entre la chaîne logistique principale et les chaînes logistiques secondaires peut s'appuyer sur le système kanban si les caractéristiques de la demande en volume et en structure sont suffisamment stables, si la diversité des références montées par le processus d'assemblage vers lequel converge une chaîne logistique secondaire est compatible avec l'espace de stockage disponible (voir la deuxième remarque du § 3.1) et si l'insuffisance de fiabilité des processus concernés est correctement prise en compte (voir la quatrième remarque du § 3.1).

Le problème d'espace de stockage en bord de chaîne posé par une forte diversité en assemblage a conduit les entreprises du secteur automobile à passer en approvisionnement synchrone dans un contexte caractérisé par une faible anticipation de la demande à satisfaire. On ne peut guère jouer dans ce cas que sur une différenciation retardée réalisée par l'un des derniers processus de la chaîne logistique secondaire à partir de stocks préexistants de composants différents (le montage des housses des sièges automobiles constituant l'exemple classiquement retenu). Une augmentation de cette anticipation permet de passer d'un approvisionnement synchrone à la production synchrone où les composants assurant la différenciation peuvent être fabriqués à la commande au lieu de l'être pour stock. Si la fiabilité des processus de la chaîne logistique secondaire est bonne, on est ramené à un problème d'ordonnancement du type de ceux évoqués au § 2.2, à ceci près que l'on est dans une logique de programmation glissante sur un horizon de quelques dizaines d'heures. Le problème est nettement plus compliqué lorsque les processus de production de la chaîne logistique secondaire n'ont pas une fiabilité suffisante par rapport aux normes de qualité (dont la sévérité a tendance à s'accroître). De nouvelles approches du pilotage doivent être alors imaginées. L'application d'une production synchrone (Giard et Mendy, 2006c) a une capacité d'adaptation immédiate aux variations de structure de la demande, ce que ne permet une approche kanban qu'au prix de stocks de sécurité importants comme on l'a montré ci-dessus.

## 6. Bibliographie

- Anderson, D.M., and J. Pine II, 1997, *Agile Product Development for Mass Customization : How to Develop and Deliver Products for Mass Customization, Niche Markets, JIT, Build-to-Order and Flexible Manufacturing*, McGraw-Hill, 1997.
- Baynat, B., J. A. Buzacott and Y. Dallery, 2002, Multiproduct Kanban-like Control Systems, *International Journal of Production Research*, vol. 40, n° 16, p. 4225-4255.
- Box, G.E.P, and Jenkins, G.M., *Time Series Analysis : Forecasting and Control* (Revised Edition), Holden-Day, San Francisco, 1976.

- Buzacott, J.A., 1989, Queuing models of kanban and MRP controlled manufacturing Systems, *Engineering Cost and Production Economics*, vol.17, p. 17.
- Boctor, F., F. Danjou et V. Giard, 2001, Analyse théorique des décyclages sur lignes de production, *Journal Européen des Systèmes Automatisés (APII-JESA)*, vol. 35, n° 5, p. 623-645.
- Cachon, G.P., and P. Zipkin, 1999, Competitive and cooperative inventory policies in a two stage supply chain, *Management Science*, vol. 45, p. 936-953.
- Dallery, Y., Di Mascolo, M., and A. Matta, 2005, Analysis of assembly systems controlled with kanbans, *European Journal of Operational Research*, n° 166, p. 310-336.
- Dallery, Y., Di Mascolo, M., and Y. Frein, 1996, An analytical method for performance evaluation of Kanban controlled production system, *Operations Research, Special Issue on New Directions for Operations Management Research*, 44, p. 50-64.
- Danjou, F., V. Giard et E. Le Roy, 2000, Analyse de la robustesse des ordonnancements / réordonnements sur ligne de production et d'assemblage dans l'industrie automobile, *Revue Française de Gestion Industrielle*, vol. 19, n° 1.
- Di Mascolo, M., 1989, Analysis of a synchronisation station for the performance evaluation of a kanban system with a general arrival process of demands, *European Journal of Operational Research*, p 147-163.
- Giard, V., 2003, *Gestion de la production et des flux*, Economica, Paris.
- Giard, V., 2004, Cohérence décisionnelle et réseau, *Les réseaux : dimensions stratégiques et organisationnelles*, C. Voisin, S. Ben Mahmoud-Jouini & S. Edouard éditeurs, Economica.
- Giard, V., et J. Jeunet, 2005, A formulation of the mixed model assembly line sequencing problem to minimize utility workers and setups and to smooth out part utilization, *International Conference on Industrial Engineering and System Management (IESM'05)*, Marrakech.
- Giard, V., et G. Mendy, 2006a, Le passage de l'approvisionnement synchrone à la production synchrone dans la chaîne logistique, *Revue Française de Gestion*, vol 25, n° 1, p. 63-82.
- Giard, V., et G. Mendy, 2006b, Exploitation of the knowledge of the final request in the piloting of a supply chain, *International Conference on Information Systems, Logistics and Supply Chain*, Lyon.
- Giard, V., et G. Mendy, 2006c, Amélioration de la synchronisation de la production sur une chaîne logistique, *Revue Française de Gestion Industrielle*, vol. 25, n° 1, p. 63-82.
- Giard, V., et J. Jeunet, 2006d, Modélisation du problème général d'ordonnement de véhicules sur une ligne de production et d'assemblage, *Journal Européen des Systèmes Automatisés (APII-JESA)*, vol. 40, n° 4-5/2006, p. 463-496.
- Gupta, S.M., and Y.A.Y Al Turki, 1997, An algorithm to dynamically adjust the number of kanbans in stochastic processing times and variable demand environment, *Production Planning & Control*, vol. 8, n° 2, p. 133-142.
- Hall, R.W., 1983, *Zero inventories*, Dow Jones Irwin, Homewood, IL.
- Karaesmen, F., Buzacott, J. and Y. Dallery, 2002, Integrating Advance Order information in Make to stock Production Systems, *IIE Transactions*, vol. 34, p. 649-662.

- Kleijnen, P.C., and E. Gaurry, 2003, Short-term robustness of production management systems : a case study, *European Journal of Operational Research*, n° 148, p. 452-465.
- Monden, Y., 1983, *Toyota Production System*, London, Institute Of Industrial Engineering
- Orlicky, J., 1975, *Material Requirement Planning*, McGraw-Hill.
- Philipoom, P., L. Rees, B. Taylor III and P. Huang, 1987a, An investigation of the factors influencing the number of kanbans required in the implementation of the JIT technique with kanban, *International Journal of Production Research*, n°25, p. 457-472.
- Philipoom, P., L. Rees, B. Taylor III and P. Huang, 1987b, Dynamically adjusting the number of kanbans in a just-in-time production system using estimated values of lead time, *IIE Transactions*, n° 19.
- Takahashi, K., and N. Nakamura, 1998, Ordering alternatives in JIT production systems, *Production Planning & Control*, vol. 9, n° 8, p 784-794.
- Tardif, V., and L. Maaseidvaag, 2001, An adaptative approach to controlling kanban systems, *European Journal of Operational Research*, 132, p. 411-424.
- Vollmann, T.E., W. L. Berry and D. C. Whybark, 1997, *Manufacturing planning and control system*, McGraw-Hill, 4e édition.