

Contrôle de qualité de service optimal d'application multimédia

Loïc Strus

Laboratoire Verimag / Université Joseph Fourier - Grenoble I

19 septembre 2008



Contexte global

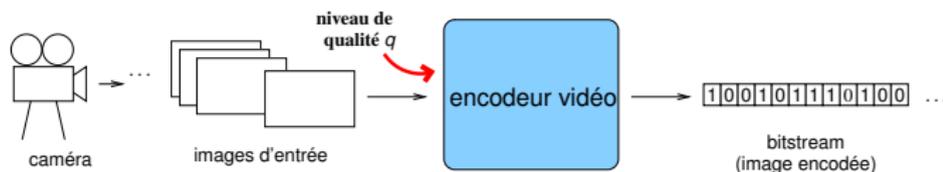
Systèmes embarqués/temps-réel

- Systèmes informatisés “non visibles”
- Utilisés dans des domaines différents
 - pilotage d'avion, appareils médicaux, ...
 - téléphone, lecteur multimédia,...
- Soumis à de nombreuses contraintes
 - ressources matérielles : miniaturisation, coût, ...
 - délai de mise sur le marché
- **Contrainte temps-réel**



Motivations et Contexte

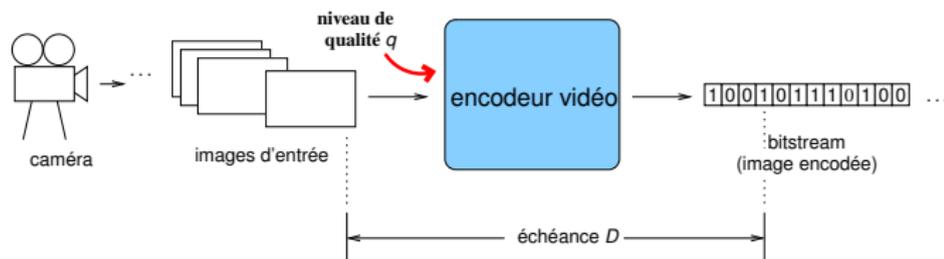
- En collaboration avec **STMicroelectronics** (projet **GaloGiC**)
- **Encodeur vidéo embarqué**



- Contraintes **temps-réel** (e.g. $D = 1/30$ s = 33 ms)
 - Ressources **limitées**
 - Calcul **intensif**
 - Durée d'exécution **incertaine**
- adapter le paramètre de niveau de qualité q des actions pour que les contraintes temps-réel soient respectées

Motivations et Contexte

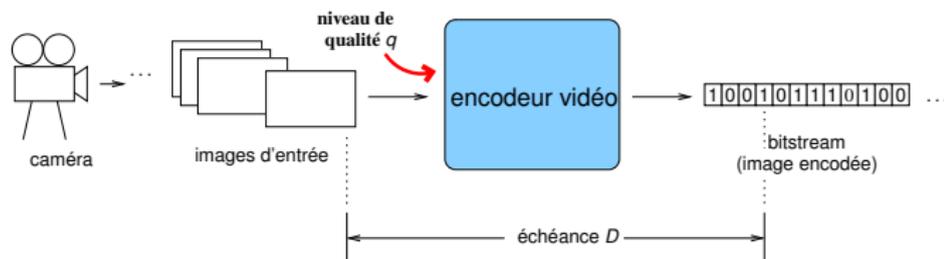
- En collaboration avec **STMicroelectronics** (projet **GaloGiC**)
- **Encodeur vidéo embarqué**



- Contraintes **temps-réel** (e.g. $D = 1/30 \text{ s} = 33 \text{ ms}$)
 - Ressources **limitées**
 - Calcul **intensif**
 - Durée d'exécution **incertaine**
- adapter le paramètre de niveau de qualité q des actions pour que les contraintes temps-réel soient respectées

Motivations et Contexte

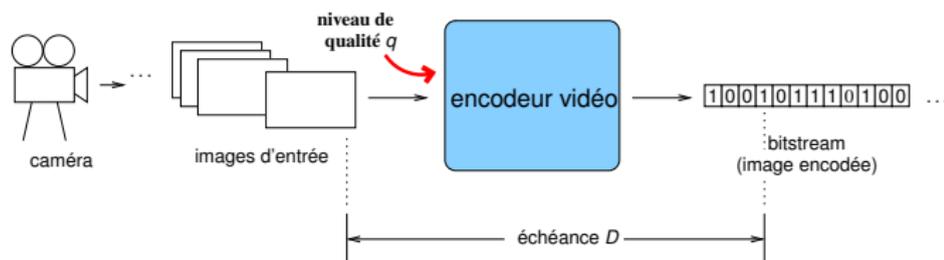
- En collaboration avec **STMicroelectronics** (projet **GaloGiC**)
- **Encodeur vidéo embarqué**



- Contraintes **temps-réel** (e.g. $D = 1/30 \text{ s} = 33 \text{ ms}$)
 - Ressources **limitées**
 - Calcul **intensif**
 - Durée d'exécution **incertaine**
- adapter le paramètre de niveau de qualité q des actions pour que les contraintes temps-réel soient respectées

Motivations et Contexte

- En collaboration avec **STMicroelectronics** (projet **GaloGiC**)
- **Encodeur vidéo embarqué**



- Contraintes **temps-réel** (e.g. $D = 1/30 \text{ s} = 33 \text{ ms}$)
 - Ressources **limitées**
 - Calcul **intensif**
 - Durée d'exécution **incertaine**
- adapter le paramètre de niveau de qualité q des actions pour que les contraintes temps-réel soient respectées

Motivations et Contexte

Problème

Problème (encodeur vidéo embarqué)

- Respecter les échéances
 - minimiser la taille du buffer d'entrée
 - conserver la fluidité de la vidéo d'entrée (frame-skipping)
- Utilisation optimale des ressources
 - maximisation de la qualité de service

Motivations et Contexte

Problème

Problème (encodeur vidéo embarqué)

- Respecter les échéances
 - minimiser la taille du buffer d'entrée
 - conserver la fluidité de la vidéo d'entrée (frame-skipping)
- Utilisation optimale des ressources
 - maximisation de la qualité de service

Approches existantes

- Temps-réel “dur” (système critique)
 - analyse pire-cas
 - satisfaire toutes les échéances
- Temps-réel “mou”
 - analyse moyen-cas → pas de garantie
 - utilisation efficace des ressources

Motivations et Contexte

Problème

Problème (encodeur vidéo embarqué)

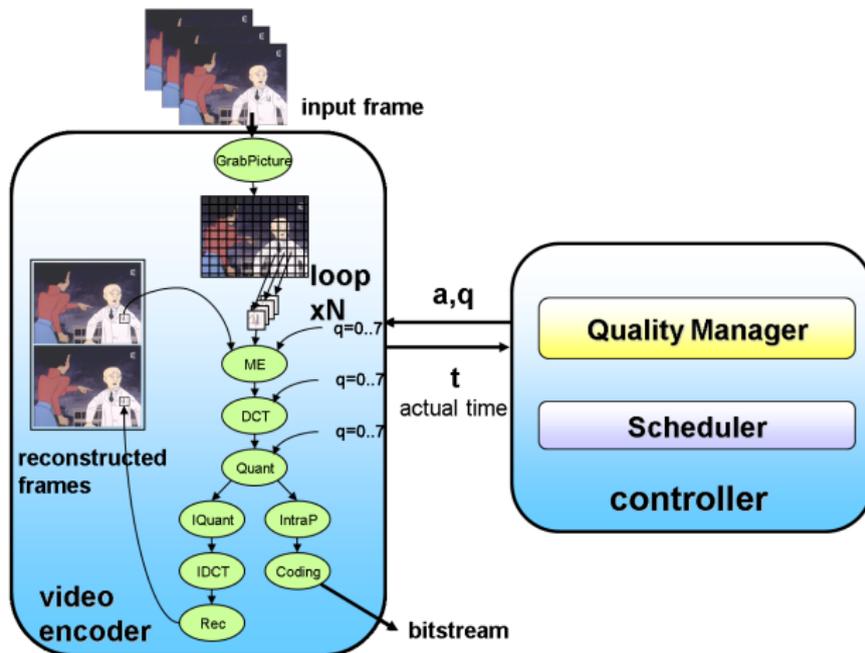
- Respecter les échéances
 - minimiser la taille du buffer d'entrée
 - conserver la fluidité de la vidéo d'entrée (frame-skipping)
- Utilisation optimale des ressources
 - maximisation de la qualité de service

Notre approche

- Réduire la distance entre ces deux approches en utilisant des **techniques adaptatives**
 - Adapter le comportement de l'application en choisissant le **paramètre de niveau de qualité**
- ⇒ Utilisation optimale des ressources sans manquer les échéances

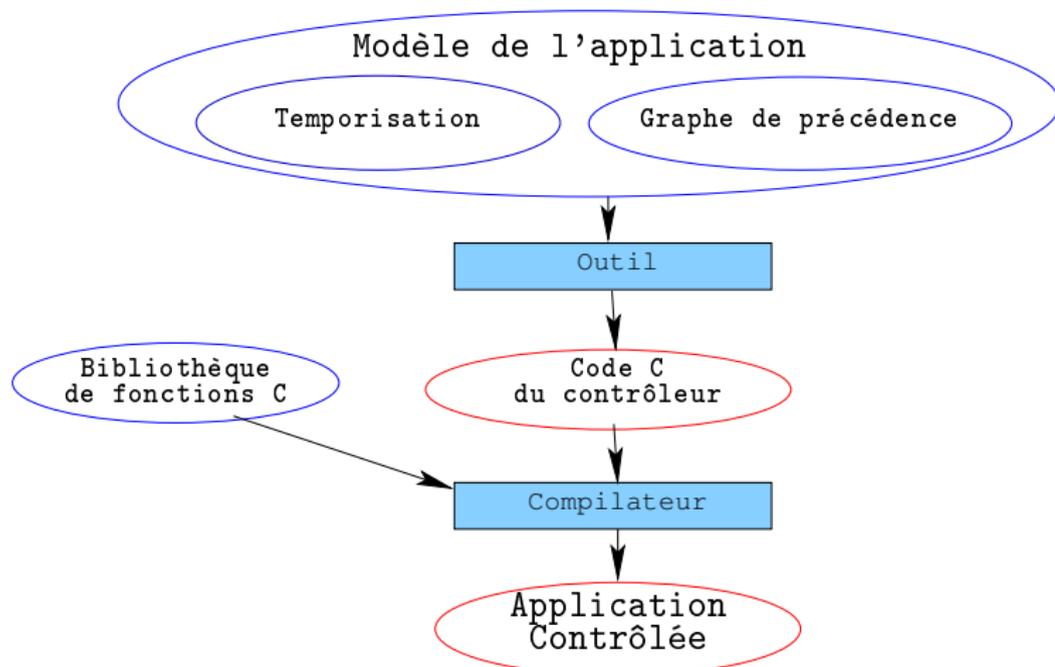
Notre approche

- Contrôle en ligne fait en deux étapes : ordonnancement et gestion de qualité
- Génération automatique d'une application contrôlée



Notre approche

- Contrôle en ligne fait en deux étapes : ordonnancement et gestion de qualité
- Génération automatique d'une application contrôlée



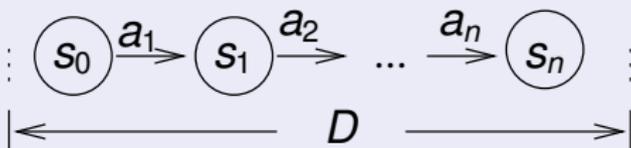
Plan

- 1 Introduction
- 2 **Contrôle d'application**
- 3 Approche Symbolique
- 4 Approche Stochastique
- 5 Conclusion et Perspectives

Gestionnaire de Qualité

Le modèle d'entrée

Logiciel applicatif



s_i : point de contrôle

a_i : action paramétrée par des niveaux de qualité entiers q_i

Estimations des durées d'exécution (croissantes avec la qualité)

$C^{av}(a_i, q)$: estimation moyenne de a_i pour le niveau de qualité q

$C^{wc}(a_i, q)$: estimation pire-cas de a_i pour le niveau de qualité q

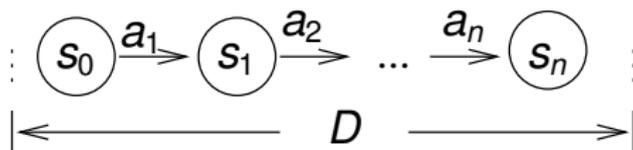
Contrainte temps-réel

D : échéance pour la terminaison de l'encodage d'une image

Gestionnaire de Qualité

Le problème

Logiciel applicatif :



Le Problème

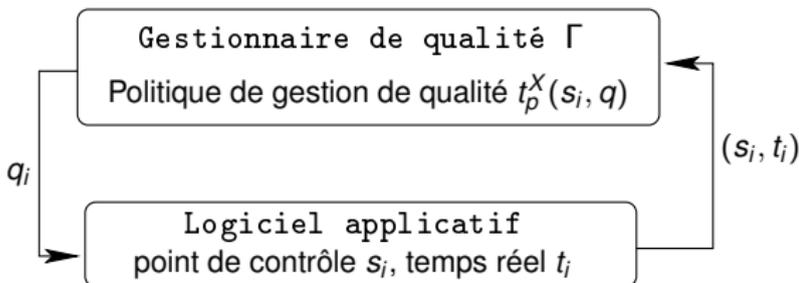
Trouver des paramètres de niveau de qualité q_i tels que :

- **Sûreté** : l'échéance est respectée
- **Optimalité** : maximisation de l'utilisation du budget de temps
- **Régularité** dans les choix de niveaux de qualité

Gestionnaire de qualité

→ Calcul **en ligne** des niveaux de qualité q_i

Conception du Gestionnaire de qualité



Politique de Gestion de qualité [DATE'05]

- Estimation des durées d'exécution : C^X
- Critère d'admissibilité : $D \geq C^X(s_i, q) + t_i$
- Nous définissons $t_p^X(s_i, q) = D - C^X(s_i, q)$

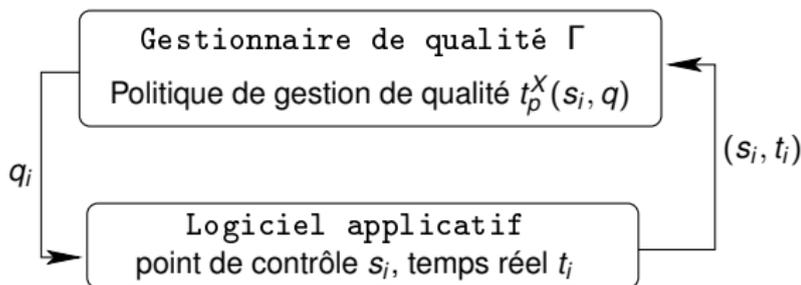
Choix du niveau de qualité

Le meilleur niveau de qualité satisfaisant le critère d'admissibilité :

$$q_i := \mathbf{max} \{ q \mid t_p^X(s_i, q) \geq t_i \}$$

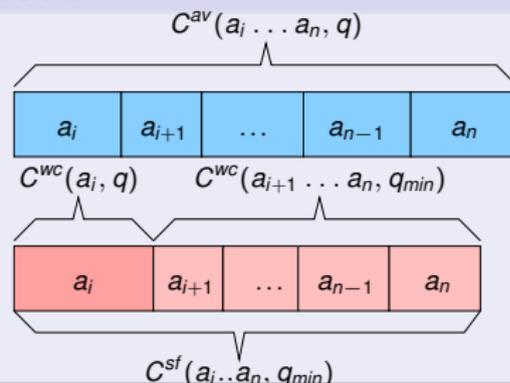
Politique de gestion de qualité mixte [EMSOFT'05]

Les bases



Estimations des durées d'exécution

- Comportement **moyen-cas**
→ régularité
- Comportement **pire-cas**
sous contrôle



Politique de gestion de qualité mixte [EMSOFT'05]

Exemple

Comportement sûr



$$C^{sf}(a_1 \dots a_n, q)$$

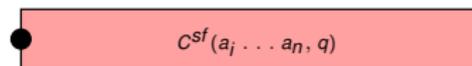
Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

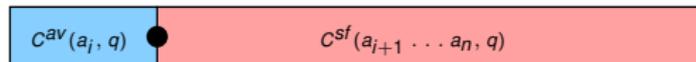
Politique de gestion de qualité mixte [EMSOFT'05]

Exemple

Comportement sûr



Comportement mixte
au pas i

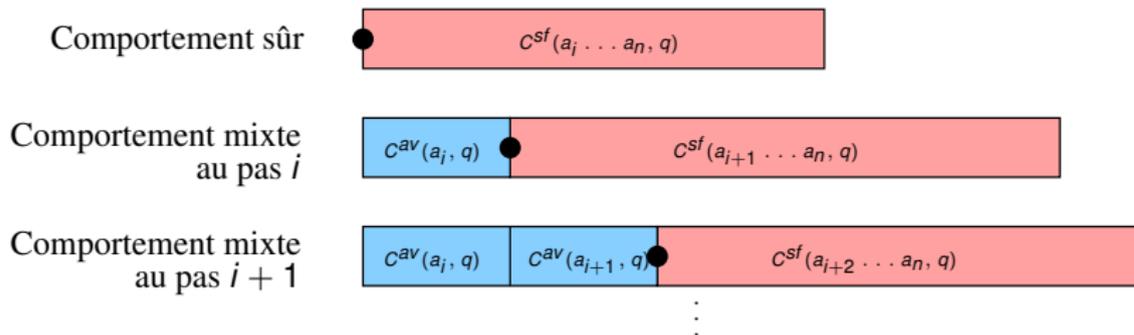


Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

Politique de gestion de qualité mixte [EMSOFT'05]

Exemple

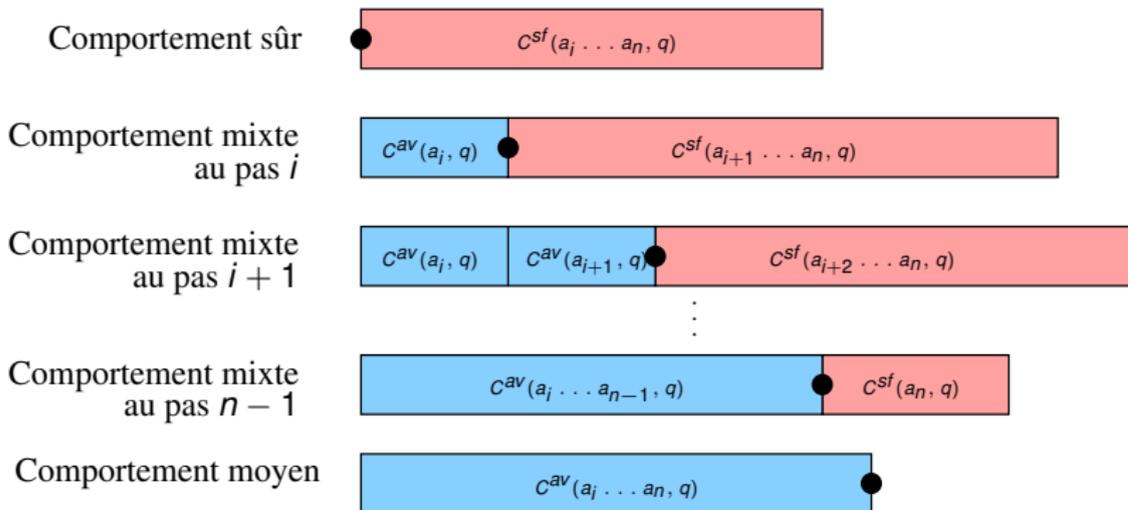


Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

Politique de gestion de qualité mixte [EMSOFT'05]

Exemple

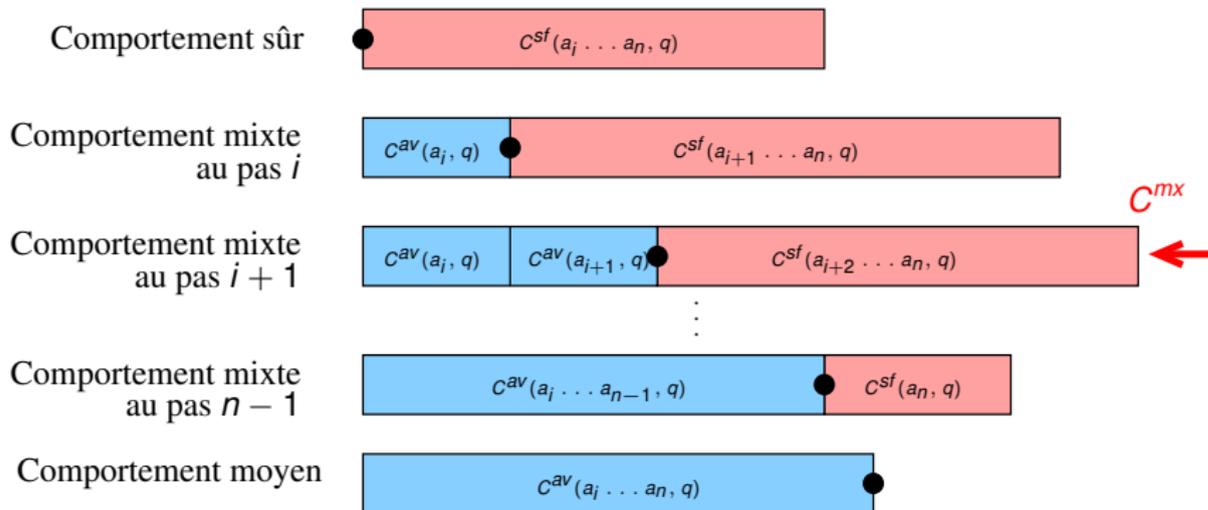


Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

Politique de gestion de qualité mixte [EMSOFT'05]

Exemple

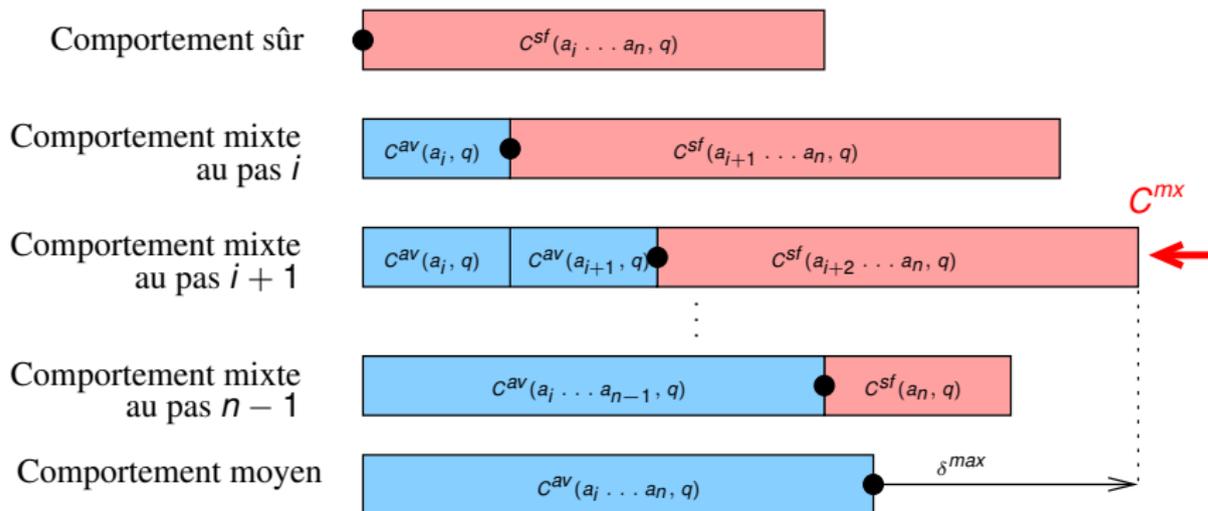


Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

Politique de gestion de qualité mixte [EMSOFT'05]

Exemple



Marge de sécurité

- Différence entre comportement moyen et mixte
- Récupération du système si le comportement moyen n'est pas suivi

Politique de gestion de qualité mixte[EMSOFT'05]

Définition

Marge de sécurité

$$\delta^{max} = \mathbf{max}_{i \leq j \leq n} [C^{sf}(a_j \dots a_n, q) - C^{av}(a_j \dots a_n, q)]$$

Politique de Gestion de qualité mixte t_p^{mx}

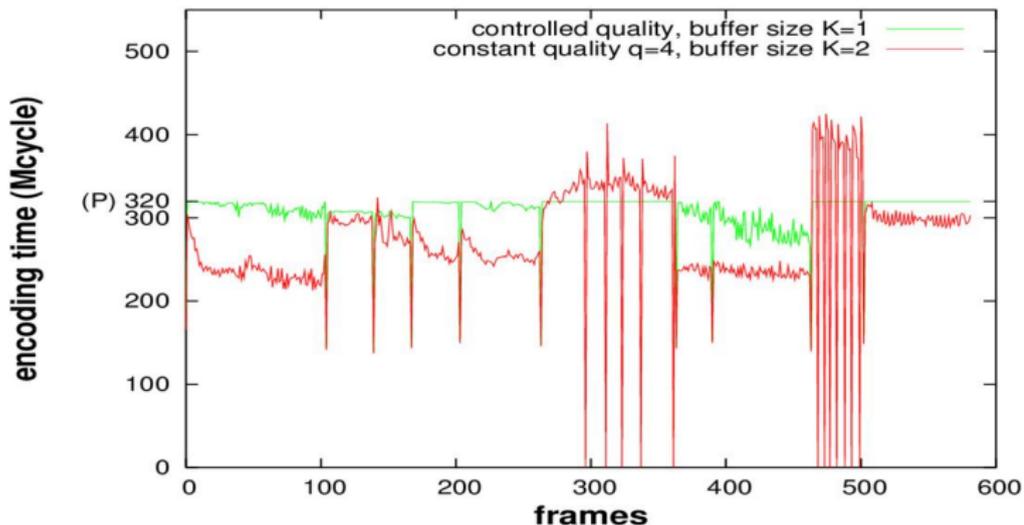
$$C^{mx}(a_i \dots a_n, q) = C^{av}(a_i \dots a_n, q) + \delta^{max}(a_i \dots a_n, q)$$

⇒ politique respectant les échéances tout en conservant une régularité des niveaux de qualité

Résultats expérimentaux

Plateforme de tests

- Board STm8010 avec processeurs ST231(400MHZ)
- Implémentation sur machine nue
- Temps d'exécution précis



Plan

- 1 Introduction
- 2 Contrôle d'application
- 3 Approche Symbolique**
- 4 Approche Stochastique
- 5 Conclusion et Perspectives

Limitations de la politique de gestion de qualité mixte

Compréhension

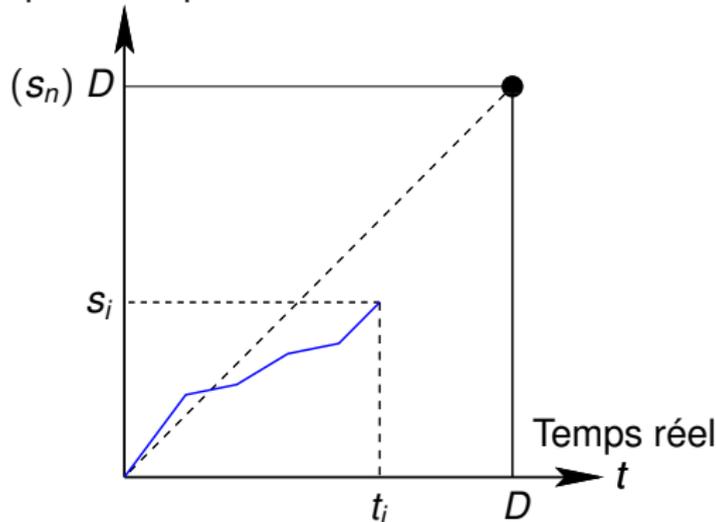
- Comportement de l'application contrôlée au cours du temps
- Influence de la marge de sécurité
- Interprétation de la politique de gestion de qualité mixte

Surcoût en temps d'exécution

- Le calcul de t_p^{mx} à chaque point de contrôle est très coûteux
- La qualité choisie peut être constante pour une séquence d'actions

Définition du diagramme des vitesses

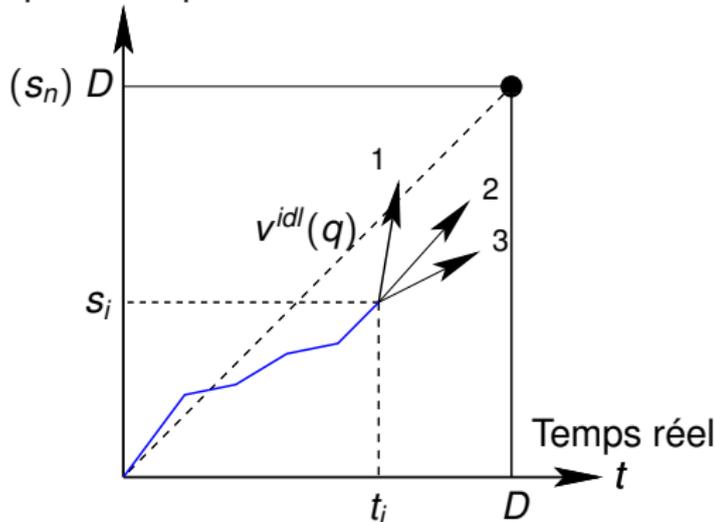
Temps théorique



- niveau de qualité \rightarrow vecteur de vitesse
- δ^{max} : marge de sécurité garantissant l'échéance
- v^{opt} : sûreté et optimalité
- $t_p^{mx} \geq t \Leftrightarrow v^{idl}(q) \geq v^{opt}(q)$

Définition du diagramme des vitesses

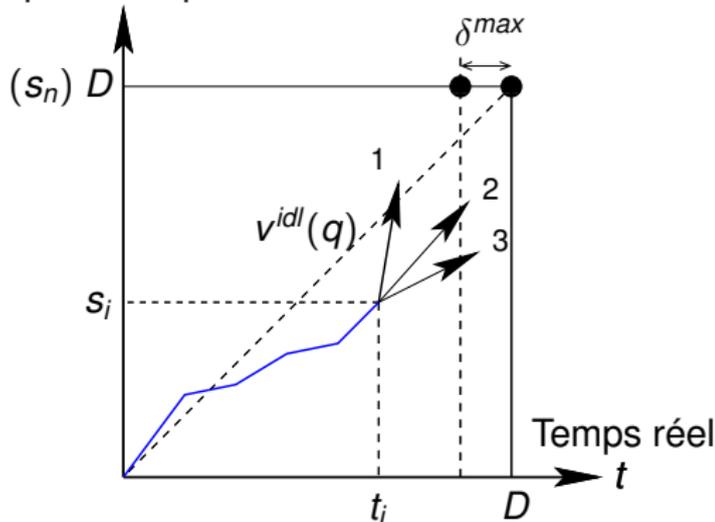
Temps théorique



- niveau de qualité \rightarrow vecteur de vitesse
- δ^{max} : marge de sécurité garantissant l'échéance
- v^{opt} : sûreté et optimalité
- $t_p^{mx} \geq t \Leftrightarrow v^{idl}(q) \geq v^{opt}(q)$

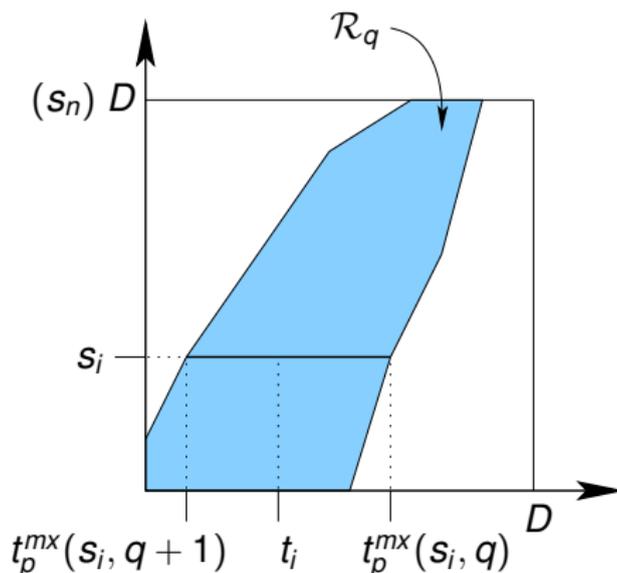
Définition du diagramme des vitesses

Temps théorique



- niveau de qualité \rightarrow vecteur de vitesse
- δ^{max} : marge de sécurité garantissant l'échéance
- v^{opt} : sûreté et optimalité
- $t_p^{mx} \geq t \Leftrightarrow v^{idl}(q) \geq v^{opt}(q)$

Régions de qualité



- **Région de qualité** $\mathcal{R}_q = \{ \text{états } (s_i, t_i) \mid \text{Gestionnaire de qualité choisit la qualité } q \}$
- A l'état i : $t_p^{mx}(s_i, q) \geq t_i > t_p^{mx}(s_i, q+1)$

Amélioration du gestionnaire de qualité

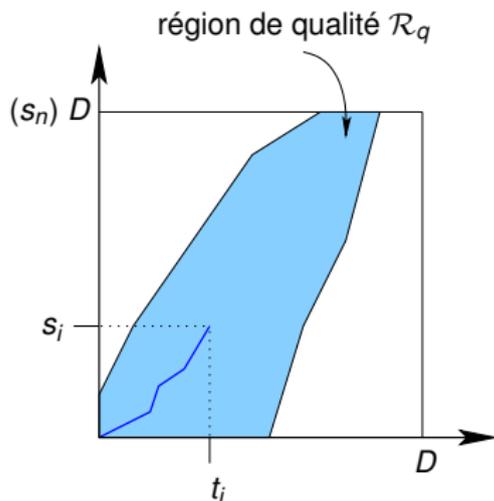
Surcoût en temps d'exécution

- Le calcul de t_p^{mx} à chaque point de contrôle est très coûteux
- La qualité choisie peut être constante pour une séquence d'actions

Méthode Symbolique

- Réduire le surcoût en temps d'exécution
- Préserver les propriétés des politiques de gestion de qualité
- Adapter dynamiquement la granularité de contrôle

Régions de relâchement du contrôleur [WPDRTS07, RT08]

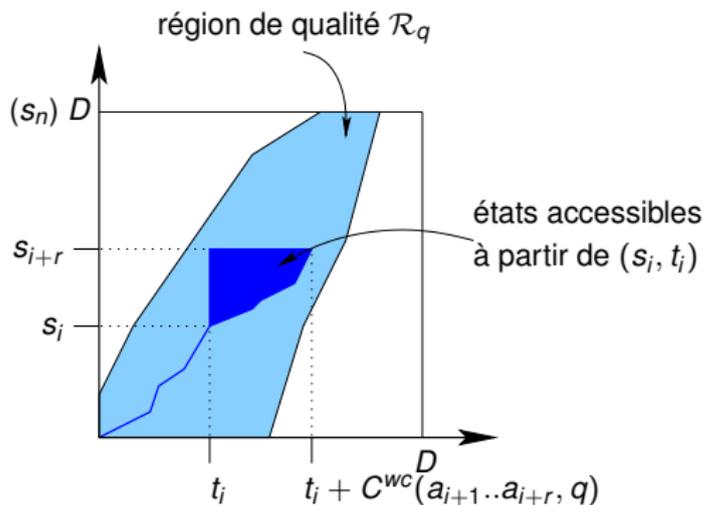


- Région de relâchement du contrôleur $\mathcal{R}_q^r = \{ \text{états } (s_i, t_i) \mid \text{Gestionnaire de qualité peut relâcher de } r \text{ pas} \}$

- Au pas i :

$$\min_{1 \leq j \leq r} t_p^{mx}(s_{i+j}, q) - C^{wc}(a_{i+1}..a_{i+j}, q) \geq t_i > t_p^{mx}(s_{i+r}, q + 1)$$

Régions de relâchement du contrôleur [WPDRTS07, RT08]

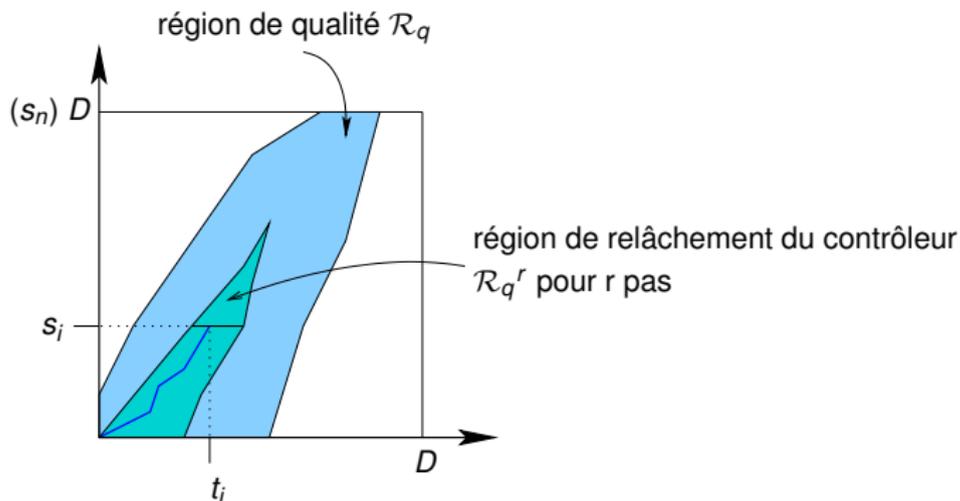


- Région de relâchement du contrôleur $\mathcal{R}_q^r = \{ \text{états } (s_i, t_i) \mid \text{Gestionnaire de qualité peut relâcher de } r \text{ pas} \}$

- Au pas i :

$$\min_{1 \leq j \leq r} t_p^{mx}(s_{i+j}, q) - C^{wc}(a_{i+1}..a_{i+j}, q) \geq t_i > t_p^{mx}(s_{i+r}, q + 1)$$

Régions de relâchement du contrôleur [WPDRTS07, RT08]



- Région de relâchement du contrôleur $\mathcal{R}_q^r = \{ \text{états } (s_i, t_i) \mid \text{Gestionnaire de qualité peut relâcher de } r \text{ pas} \}$

- Au pas i :

$$\min_{1 \leq j \leq r} t_p^{mx}(s_{i+j}, q) - C^{wc}(a_{i+1}..a_{i+j}, q) \geq t_i > t_p^{mx}(s_{i+r}, q + 1)$$

Résultats expérimentaux pour le symbolique

Implémentation

Nous avons comparé trois méthodes :

- **Approche Énumérative**

- Calcul direct de t_p^{mx}

- **Approche Symbolique – régions de qualité**

- Régions de qualité pré-calculées \mathcal{R}_q
- Surcoût en allocation mémoire : 20 Ko

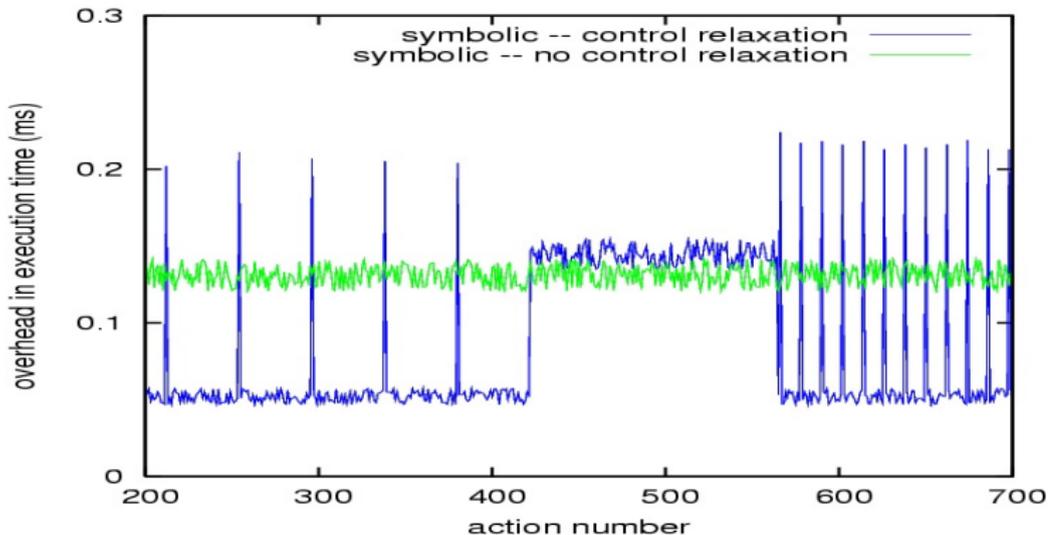
- **Approche Symbolique – Régions de relâchement du contrôleur**

- Régions de relâchement du contrôleur pré-calculées \mathcal{R}_q^r
- Pas de relâchement : 1, 10, 20, 30, 40, 50
- Surcoût en allocation mémoire : 350 Ko

Résultats expérimentaux pour le symbolique

Surcoût en temps d'exécution

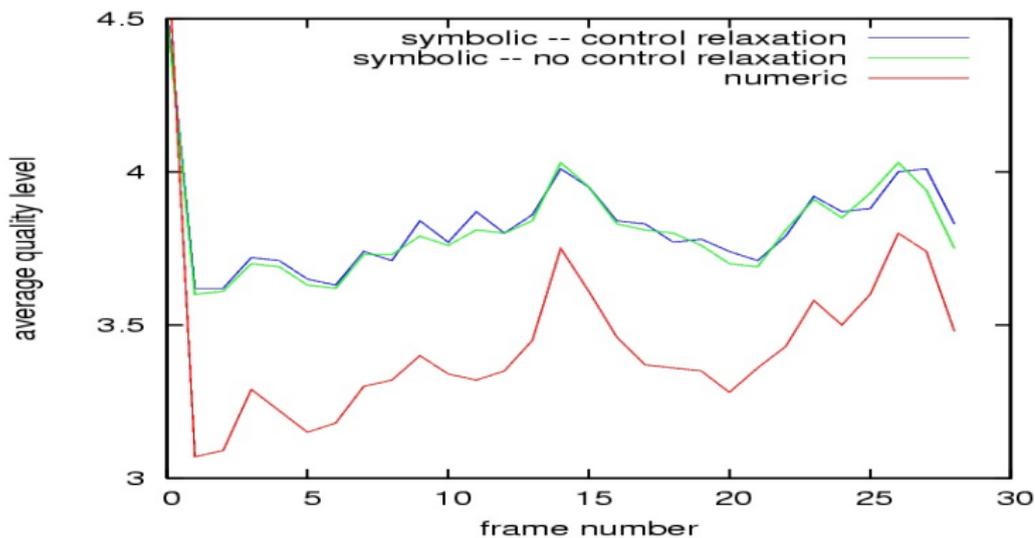
	Énumérative	Régions de qualité	Régions de relâchement
Surcoût moyen	4%	1.2%	< 1 %



Résultats expérimentaux pour le symbolique

Niveau de qualité moyen

► Surcoût du contrôleur réduit \Rightarrow Gain dans le choix des niveaux de qualité



Plan

- 1 Introduction
- 2 Contrôle d'application
- 3 Approche Symbolique
- 4 Approche Stochastique**
- 5 Conclusion et Perspectives

Autres limitations de l'approche

Limitations (estimations pire-cas)

- Plate-formes matérielles complexes — calcul des durées d'exécution pire-cas
 - Pas de flexibilité — respect des échéances garanti à 100 %
- ⇒ Utilisation des ressources non maximale, malgré l'utilisation du contrôle à grain fin

Maîtrise de l'influence des durées d'exécution pire-cas

- Besoin de méthodologies temps-réel mou qui apportent des garanties
- Fournir à l'utilisateur des possibilités d'avoir une application dont le comportement est plus flexible

Autres limitations de l'approche

Limitations (estimations pire-cas)

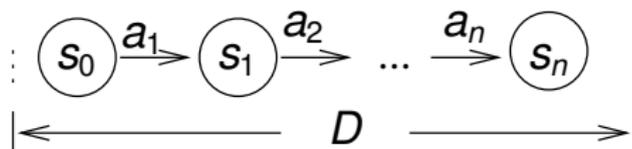
- Plate-formes matérielles complexes — calcul des durées d'exécution pire-cas
 - Pas de flexibilité — respect des échéances garanti à 100 %
- ⇒ Utilisation des ressources non maximale, malgré l'utilisation du contrôle à grain fin

Maîtrise de l'influence des durées d'exécution pire-cas

- Besoin de méthodologies temps-réel mou qui apportent des garanties
- Fournir à l'utilisateur des possibilités d'avoir une application dont le comportement est plus flexible

Modèle

Logiciel applicatif :



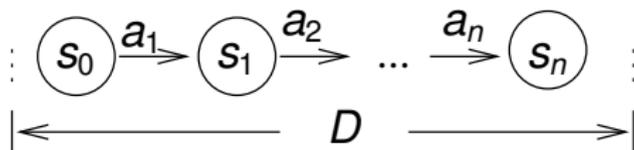
Estimation des durées d'exécution

$C^{av}(a_i, q)$: durée d'exécution moyenne de a_i pour le niveau de qualité q

$C^{wc}(a_i, q)$: durée d'exécution pire-cas de a_i pour le niveau de qualité q

Modèle

Logiciel applicatif :



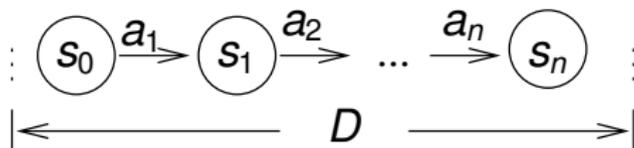
Estimation des durées d'exécution

Fonction de distribution des probabilités d_i^q :

- $d_i^q : \mathbb{N} \rightarrow [0, 1]$
- $d_i^q(k) = P["\text{durée d'exécution de } a_i \text{ pour la qualité } q \text{ est } k"]$
- $\sum_{k=0}^{+\infty} d_i^q(k) = 1$

Problème

Logiciel applicatif :



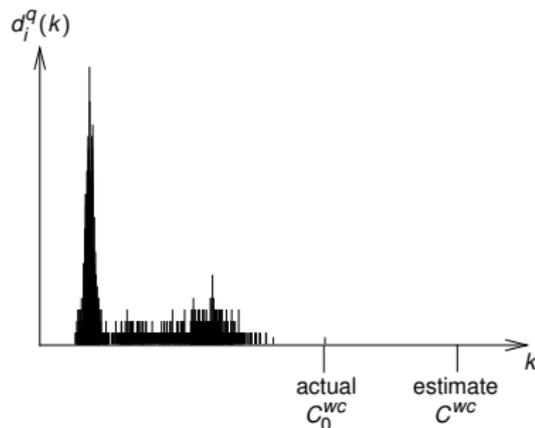
Le Problème

Trouver un Gestionnaire de qualité Γ tel que :

- **Sûreté** — le taux d'échéances manquées est \leq au taux de violation des échéances fixé $\lambda \in [0, 1]$
- **Optimalité** — maximisation des niveaux de qualité
- **Régularité** des niveaux de qualité

Politique de gestion de qualité mixte stochastique

Modèle de temps



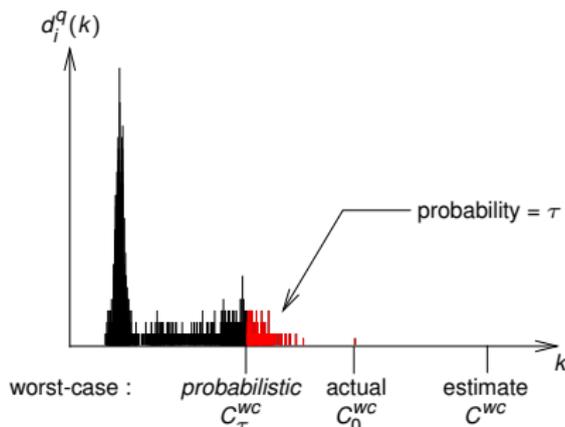
- Seuil de tolérance $\tau \in [0, 1]$
- Durée d'exécution pire-cas *stochastique* $C_\tau^{wc}(a_i, q)$ telle que :

$$C_\tau^{wc}(a_i, q) = \min \left\{ l \mid \sum_{k=l}^{+\infty} d_i^q(k) \leq \tau \right\}$$

- Durée d'exécution moyenne : $C^{av}(a_i, q) = \sum_{k=0}^{+\infty} k d_i^q(k)$

Politique de gestion de qualité mixte stochastique

Modèle de temps



- Seuil de tolérance $\tau \in [0, 1]$
- Durée d'exécution pire-cas *stochastique* $C_\tau^{wc}(a_i, q)$ telle que :

$$C_\tau^{wc}(a_i, q) = \min \left\{ l \mid \sum_{k=l}^{+\infty} d_i^q(k) \leq \tau \right\}$$

- Durée d'exécution moyenne : $C^{av}(a_i, q) = \sum_{k=0}^{+\infty} k d_i^q(k)$

Politique de gestion de qualité mixte stochastique

Influence du seuil de tolérance

Définition de la politique

$$C_{\tau}^{mx} = C^{av} + \delta_{\tau}^{max}$$

où

$$\delta_{\tau}^{max}(a_{j..a_n}, q) = \mathbf{max} \{ 0 \} \cup \{ C_{\tau}^{sf}(a_{j..a_n}, q) - C^{av}(a_{j..a_n}, q) \mid i \leq j \leq n \}$$

Seuil de tolérance

τ définit la dureté des échéances :



- $\tau = 0$: utilisation de la politique mixte
- $\tau = 1$: politique de gestion de qualité moyenne

Analyse de performance

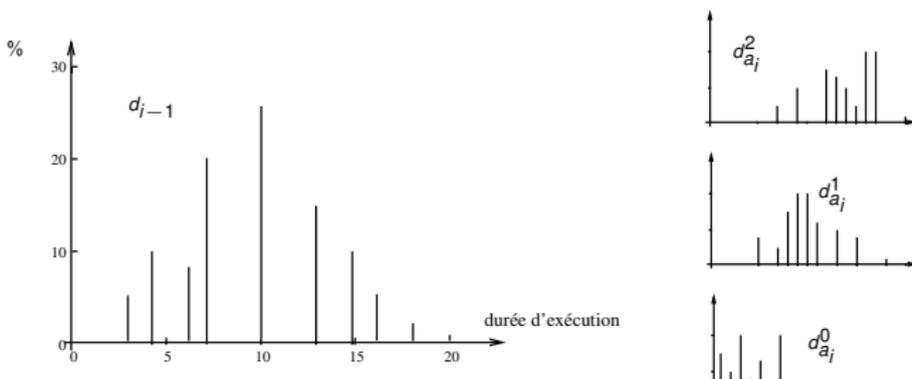
Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_i : distribution de probabilités pour t_i
- Calcul récursif de d_i :

$$d_i(k) = \sum_{l=0}^{+\infty} d_{i-1}(l) d_i^{q_i}(k-l) \quad \text{où} \quad q_i = \Gamma(s_{i-1}, l)$$



Analyse de performance

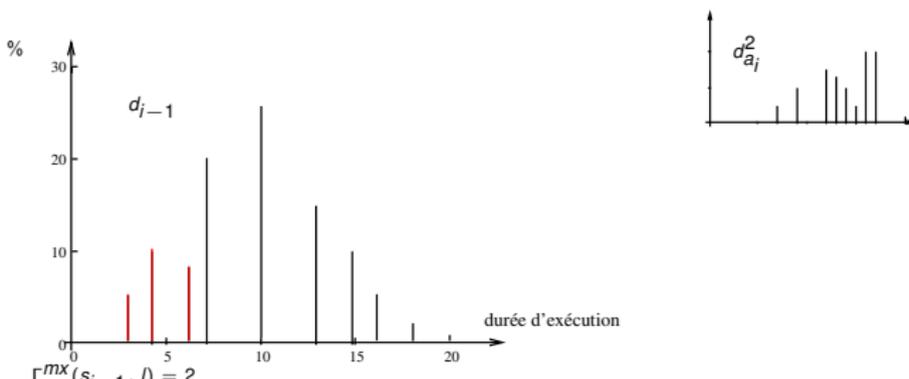
Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_i : distribution de probabilités pour t_i
- Calcul récursif de d_i :

$$d_i(k) = \sum_{l=0}^{+\infty} d_{i-1}(l) d_i^{q_i}(k-l) \quad \text{où} \quad q_i = \Gamma(s_{i-1}, l)$$



Analyse de performance

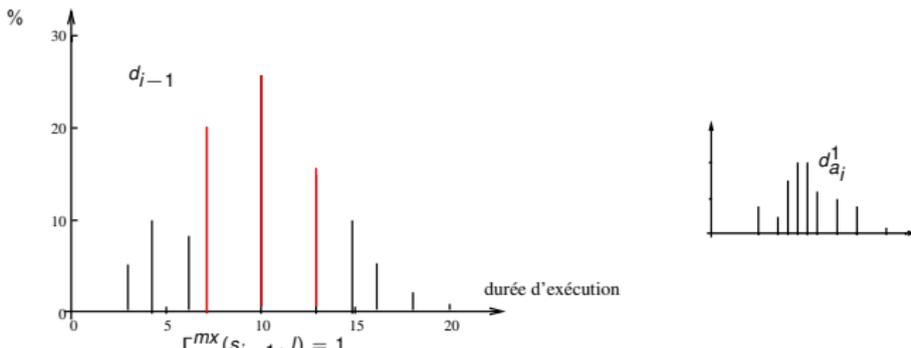
Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_i : distribution de probabilités pour t_i
- Calcul récursif de d_i :

$$d_i(k) = \sum_{l=0}^{+\infty} d_{i-1}(l) d_i^{q_i}(k-l) \quad \text{où} \quad q_i = \Gamma(s_{i-1}, l)$$



Analyse de performance

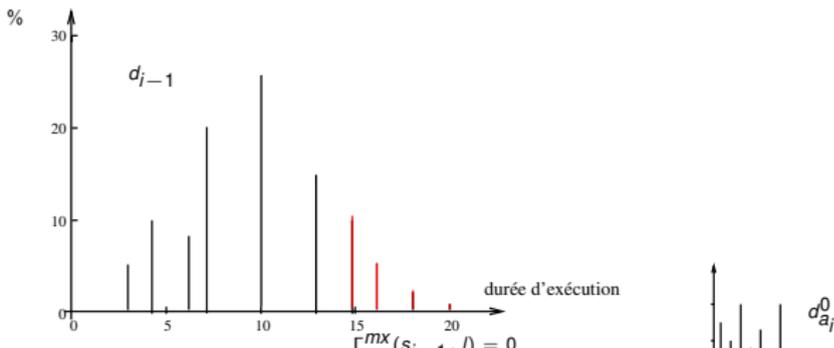
Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_j : distribution de probabilités pour t_j
- Calcul récursif de d_j :

$$d_i(k) = \sum_{l=0}^{+\infty} d_{i-1}(l) d_i^{q_i}(k-l) \quad \text{où} \quad q_i = \Gamma(s_{i-1}, l)$$



Analyse de performance

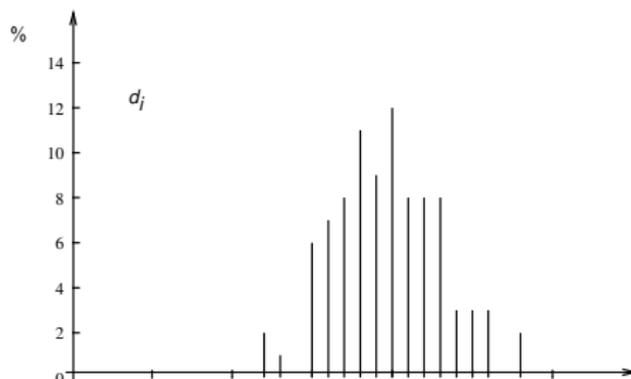
Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_i : distribution de probabilités pour t_i
- Calcul récursif de d_i :

$$d_i(k) = \sum_{l=0}^{+\infty} d_{i-1}(l) d_i^{q_i}(k-l) \quad \text{où} \quad q_i = \Gamma(s_{i-1}, l)$$



Analyse de performance

Probabilité de dépasser les échéances

Méthode calculant le nombre d'échéances manquées en fonction de

τ

- d_j : distribution de probabilités pour t_j
- Calcul récursif de d_j :

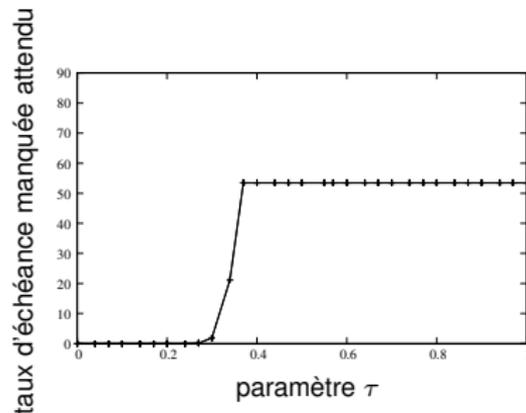
$$d_j(k) = \sum_{l=0}^{+\infty} d_{j-1}(l) d_j^{q_j}(k-l) \quad \text{où} \quad q_j = \Gamma(s_{j-1}, l)$$

Performance du gestionnaire de qualité Γ

- Taux d'échéances manquées attendu : $P[t_n > D] = \sum_{k>D} d_n(k)$
- Utilisation du budget de temps attendue : $E(t_n) = \sum_{k=0}^{+\infty} k d_n(k)$

Résultats expérimentaux

Analyse de performance

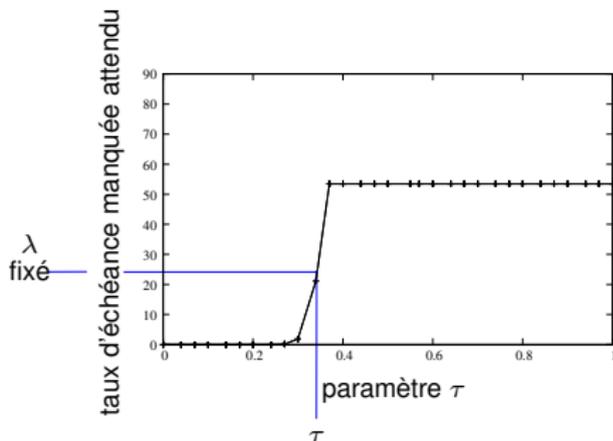


Utilisation de l'analyse de performance

- Adapter le seuil de tolérance en fonction de l'estimation de λ
- Vérifier l'utilisation maximale du budget de temps
- Gérer le compromis entre la sûreté et l'optimalité

Résultats expérimentaux

Analyse de performance

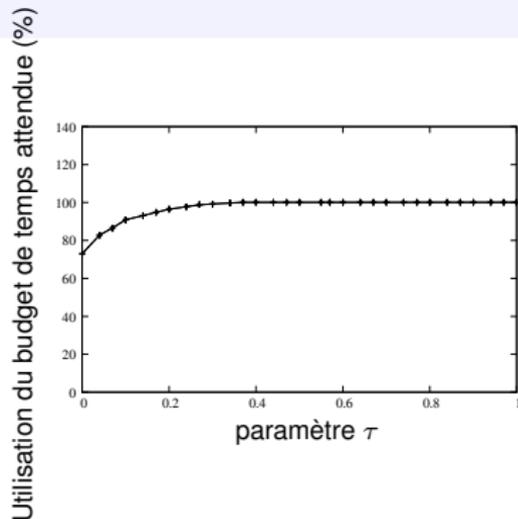
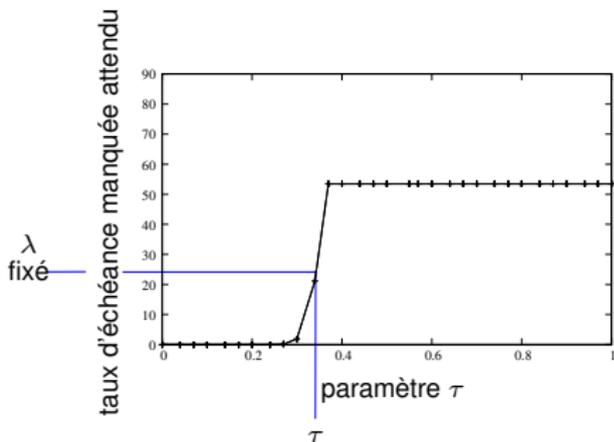


Utilisation de l'analyse de performance

- Adapter le seuil de tolérance en fonction de l'estimation de λ
- Vérifier l'utilisation maximale du budget de temps
- Gérer le compromis entre la sûreté et l'optimalité

Résultats expérimentaux

Analyse de performance



Utilisation de l'analyse de performance

- Adapter le seuil de tolérance en fonction de l'estimation de λ
- Vérifier l'utilisation maximale du budget de temps
- Gérer le compromis entre la sûreté et l'optimalité

Résultats expérimentaux

Exemple de vidéo

- Comparaison entre deux vidéos pour des seuils de tolérance différents
- Un seuil de tolérance élevé entraîne un manque de fluidité dans la vidéo

$$\tau = 0.35$$

$$\tau = 0.1$$

Plan

- 1 Introduction
- 2 Contrôle d'application
- 3 Approche Symbolique
- 4 Approche Stochastique
- 5 Conclusion et Perspectives

Conclusion

Basée sur une thèse déjà existante (J. Combaz)

- Définition du contrôleur d'application
- Politique de gestion de qualité mixte

Approche symbolique [WPDRTS'07, RT'08]

- Modélisation du comportement de l'application contrôlée
- Relâchement des appels au contrôleur
- Réduction du surcoût en temps d'exécution \Rightarrow Augmentation des choix de niveaux de qualité

Approche stochastique [ESTIMedia'08]

- Méthodologie du temps-réel mou mais maîtrisé
- Gestion du compromis entre la sûreté et la maximisation du budget de temps
- Outil d'analyse de performance (1000 lignes de code)

Perspectives

- Amélioration de l'approche symbolique
 - réduction de l'espace mémoire utilisé (enveloppes convexes)
- Amélioration de l'approche stochastique
 - prise en compte de la dépendance entre les actions
- Ajout de techniques d'apprentissage [ETFA'08]
- D'autres applications : Moteur graphique 3D
- D'autres critères :
 - Consommation mémoire, consommation d'énergie
 - Contrôleur d'application multi-critères

Merci de votre attention

Des questions ?