Expression et support de la QoS : les applications multimédias réparties

Département informatique et réseaux

Frank Singhoff

C-214

singhoff@enst.fr

Sommaire

Partie 1

Présentation, définitions

Quelques définitions

1. Médium = moyen par lequel l'information est perçue, stockée ou transmise (définition du MHEG[a] de l'ISO).
2. Multimédia = littéralement qui est constitué de plusieurs média (pluriel de médium).
3. Médium de représentation, de stockage, de transmission, etc.

• Application multimédia : qui manipule plusieurs média de représentation.

[a] Multimedia and Hypermedia Expert Group.

La qualité de Service

- Notion de Qualité de Service (QoS, *Quality of Service*) : décrit les conditions de réalisation d'un service donné et non sa fonction. Garantie ou non (déterministe ou probabiliste).
- Ex : restitution d'un médium de représentation $\Longrightarrow$ QoS dite "utilisateur".

1. Contraintes de fiabilité : taux d'erreurs, taux de pertes, probabilité de disponibilité.
2. Contraintes spatiales : débit, taille, volume. Variable ou fixe.

   Contraintes plus "nouvelles" :

3. Contraintes temporelles : gigue, délais de bout en bout, temps de réponse.

   Synchronisations : intra et inter-flux, entre objets, etc.

   Contraintes absolues et relatives.

Exemples de contraintes, besoins

- Contraintes spatiales (débit fixe ou variable) :
  - Audio qualité voie téléphonique : 64 Kbits/s.
  - Audio qualité CD : 1,4 Mbits/s.
  - Vidéo MPEG-1 : environ 1,5 Mbits/s, débit supérieur si MPEG 2.
  - Vidéo TVHD non compressé : 1,2 Gbits/s.
- Contraintes temporelles :
  - Intra-flux vidéo : cadence 40 ms ou 33 ms, gigue $\leq$ 120 ms.
  - Voix-lèvres : gigue $\leq\ +/-80$ ms.
  - Téléphonie : latence $\leq$ 250 ms.

Exemples d'applications multimédias

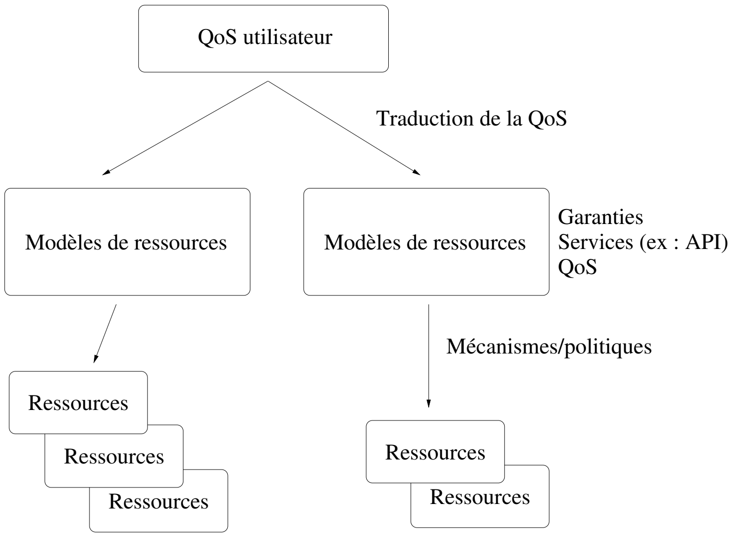
- Grande diversité : du temps réel fortement contraint aux applications partiellement déterministes.
- Dépend des : données, usages, environnements.
  - Exemple 1 : vidéo-conférence. Plate-forme dédiée. Opération chirurgicale $\Longrightarrow$ fort déterminisme. H.261. Débit fixe. RNIS.
  - Exemple 2 : télévision interactive sur le web.
    - Plate-forme généraliste. Non déterminisme temporel à cause du matériel et du logiciel (ex : PC + windows).
    - Application interactive.
    - Nombre de flots inconnu et QoS modifiable.
    - Débits variables et difficiles à estimer hors ligne.

Système d'exploitation multimédia

- Objectif : gestion des ressources et prise en compte des contraintes de QoS. Tâche complexe car :
  - Nécessite de nouvelles abstractions + traduction (ex : flots continus).
  - Ressources nombreuses et services de QoS très différents. QoS de bout en bout.
  - Cohabitation avec les applications temps partagé.
- Et dans l'idéal :
  - Masquer la complexité au concepteur $\Longrightarrow$ sinon nécessite compétences étendues.
  - Fonctions d'intégration (logiciels).
  - Portable, maintenable.

$\Longrightarrow$ Utilisation d'un modèle global.

Des besoins aux services (1)



Des besoins aux services (2)

- Le système doit donc offrir les services suivants :

  1. Expression des contraintes de QoS, des besoins. Traduction automatique.
  2. Réservation des ressources, contrôle d'admission.
  3. Allocation des ressources.
  4. Supervision : estimation, rétroaction et pilotage, "policing".

- Implantation "répartie" : matériel, système d'exploitation. Intégration de services $\Longrightarrow$ espace utilisateur, middleware.

Partie 2

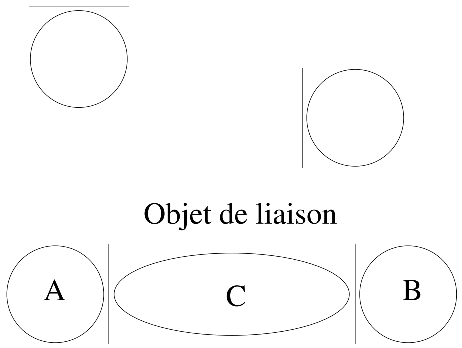
Spécification de la QoS

Expression de la QoS

- Expression de la QoS utilisateur et de la QoS des éléments du système.
- Formalimes et puissance d'expression dépendent de l'élément concerné. Compromis nécessaires pour la traduction.
- Exemples :

1. API (ex : couche transport).
2. Structure syntaxique (FLEX [LIN 88], SMART [NIE 97]).
3. Méthodes formelles : possibilité de preuves (Réseaux de Petri, logiques, etc). Nouveau compromis. Principalement deux approches : événementielle et intervalle.

Exemple : le modèle ODP



- Objets, objets de liaison [BLA 98a].
- Contrat de QoS, QoS requise et QoS offerte.

L'approche "événements"

- Expression de contraintes déterministes et probabilistes avec la logique temporelle QL [BLA 98a].
- Flux de données continues, événements, horloges.
- Exemples de contraintes déterministes :

1. Synchronisation inter-flux. Dérive (rafales non bornées).

$$\forall\ n, k\ :\ x \leq \tau(e, n+k) - \tau(e', n) \leq y$$

2. Synchronisation intra-flux. Ne dérive pas.

$$\forall\ n\ :\ x \leq \tau(e, n) - \tau(H_p, n) \leq y$$

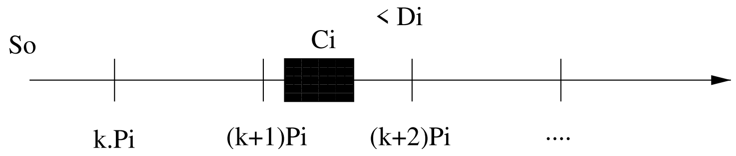
3. Flux de données "strictement périodique".

$$\forall\ n\ :\ \tau(e, n) = n * p$$

Partie 3

La ressource processeur : modèles de tâches et ordonnancement

Ordonnancement, définitions



- Modèles de tâches :
  - Activation, tâches périodiques, tâches sporadiques, tâches apériodiques.
  - Modèle de Lui et Layland [LIU 73] : $D_i = P_i$. Généralisation D_i et P_i quelconques, avec ou sans contrainte de dépendance (précédences, ressources).
- Algorithmes : hors ligne/en ligne, priorités statiques/dynamiques, préemptif ou non, oisif ou non $\Longrightarrow$ propriétés.

L'algorithme Rate Monotonic (1)

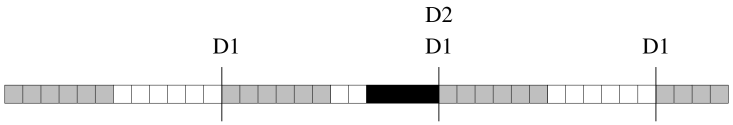
- Priorités fixes. Priorités = inverse de la périodicité. Election de la plus forte priorité.
- Tâches périodiques uniquement.
- Analyse hors ligne $\Longrightarrow$ applications statiques.
- Cas préemptif $\Longrightarrow$ réservation =

$$\sum_{i=1}^{n} \frac{C_i}{P_i} \leq n(2^{\frac{1}{n}} - 1)$$

L'algorithme Rate Monotonic (2)

- Exemple (cf. [DEM 99]) :

T1 : C1=6 ; P1=12 (gris)
T2 : C2=8 ; P2=24 (blanc)
Noir = libre



L'algorithme EDF (1)

– Priorités dynamiques. Election de la plus courte échéance d'abord.

– Tâches périodiques et apériodiques + en ligne $\Longrightarrow$ mieux adapté que RM aux applications dynamiques.

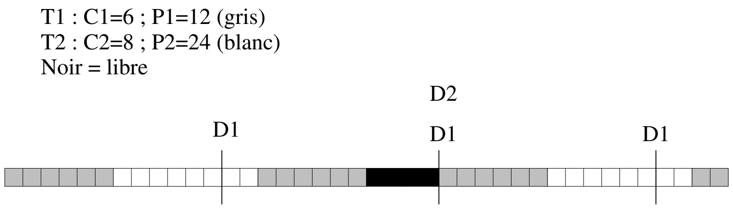
– Cas préemptif $\Longrightarrow$ optimal, réservation =

$$\sum_{i=1}^{n} \frac{C_i}{P_i} \leq 1$$

– Instable en sur-charge.

L'algorithme EDF (2)

- Exemple (cf. [DEM 99]) :



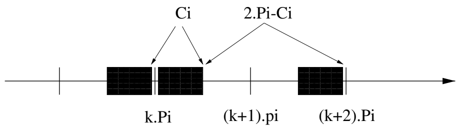
Et le multimédia dans tout ça?

– Contraintes exprimées.

– Dynamicité.

– Répartition.

Modèles pour le multimédia



• Le modèle périodique est il adapté aux applications multimédias? Gigue = $2.P_i$.

1. Contrainte de distance [SAK 95, HAN 96] :

$$\epsilon_1 \leq \tau(i, n+1) - \tau(i, n) \leq \epsilon_2$$

   Dimensionnement de la période, algorithme oisif, ajout d'une contrainte artificielle.

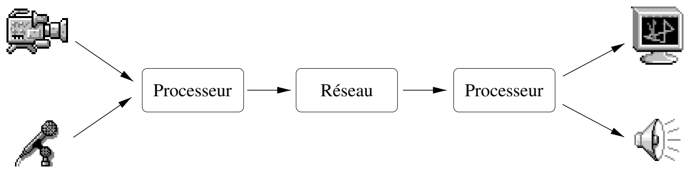
$$P_i = \frac{\epsilon_2 + C_i}{2}$$

2. Cadence d'activation [JEF 95] : plusieurs activations dans une période. Echéances relatives. Adaptation par le nombre d'activations en fonction des ressources.

Prise en compte de la dynamicité

- Arrivée de nouvelles tâches, modification de la QoS, non conformité aux spécifications (systèmes avec "policing"). Exemples de mécanismes :
  - Choix de l'algorithme (RM, EDF ?).
  - Technique du serveur sporadique [SPR 89].
  - Renégociation : modification des besoins de l'application (YARTOS [JEF 92]).
  - Basculement "passif". Critère d'importance. Niveaux de QoS. Avec ou sans réservation (ex : Pegasus [SIJ 96] et Brandt et al. [BRA 98]).
  - Supervision + adaptation. Avec réservation préalable (Jones et al. [JON 97]) ou sans. Application adaptative.

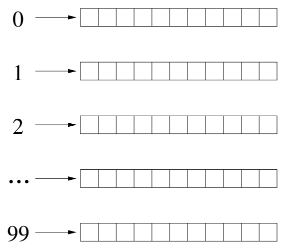
Prise en compte de la répartition



  - Généralement pas de migration ; pas d'ordonnancement global donc. Placement statique.
  - Problème traité : ordonnancement conjoint des communications et des tâches.
- Principales solutions :
  - Analyse hors ligne : calcul holistique de Tindell [TIN 94]. Temps de réponse au pire cas. Tâche fictive. Temps de service et de blocage.
  - Report des contraintes. Généralement pour les systèmes sans garantie.

Systèmes embarqués

1. Système mono-programmé.
2. Système multi-programmé "à la POSIX".



- Priorités fixes, préemptif $\Longrightarrow$ RM facile.
- Une file d'attente par priorité + politiques de gestion de la file (*SCHED_OTHERS*, *SCHED_FIFO*, *SCHED_RR*).

Systèmes Unix temps réel

- Nouvelles contraintes :
  - Non déterminisme temporel (matériel et logiciel).
  - Cohabitation avec des applications temps partagé.
- Généralement à la POSIX.
- Construction d'ordonnanceurs "utilisateurs" peu aisée :
  - Abstraction sans coopération. Ressources systèmes cachées.
  - Surcoût important. Ordonnancement contextuel.
  - Préemptivité réduite. Inversion de priorité possible.

Partie 4

Services de communication et QoS

Présentation

- Définition : réseaux isochrores, synchrones et asychrones.
- Deux approches selon la QoS offerte :

1. Réservation de ressources pour les réseaux avec garantie de service (bande passante, comportement temporel). Orientés connexions.
2. Supervision et rétroaction pour les réseaux “best effort” (non déterministe).

- Mécanismes présents de l’application en passant par la couche transport, la couche liaison, etc...

Ex. de LAN déterministes (1)

- Solution principale aujourd'hui : ATM
  - Maturité naissante : portabilité, support parfois partiel (paramètres, services).
  - Essentiellement services CBR et VBR-rt :
    - CBR : (Constant Bit Rate). Débit et comportement temporel garantis. Trafic et QoS (CBR, PCR, maxCTD, ppCDV, CLR).
    - Gigue d'une grande précision $\Longrightarrow$ synchronisation.
    - VBR rarement offert (débit moyen + multiplexage statistique).
  - Prise en compte de la QoS soit en ATM natif (AAL5), soit IP "à la sauce QoS", éventuellement avec Classical-IP.

Ex. de LAN déterministes (2)

1. L'outsider : FDDI.
   - Trafics synchrone et asynchrone. Traduction de la QoS simple. Partage de la bande synchrone.
   - Borne déterministe (2*TTRT).
   - Compromis efficacité/valeur du TTRT. TTRT grand = gigue grande.
   - Backbone, donc rarement de bout en bout.
   - En perte de vitesse.
2. Challenger : le bus 1394.
   - Connexion de périphériques variés et "grand public".
   - Haut débit (jusqu'à 400 Mbits/s).
   - Services isochrones et asynchrones $\Longrightarrow$ logiciels pour les services isochrones rarement disponibles aujourd'hui.

Réseaux non déterministes (1)

- Les protocoles RTP/RTCP [SCH 96] :

1. RTP (Real Time Protocol) : Protocole utilisateur. N'offre pas de garantie de QoS. Utilisation de couche transport à faible coût (ex : UDP). Unicast et multicast. Données comprenant des contraintes temporelles $\Longrightarrow$ estampille horloge logique ou horloge globale (synchronisation, prise en compte de la gigue).
2. RTCP (Real Time Control Protocol) : Transfert de données de supervision (ex : perte de paquet (congestion), estimation de délais de bout en bout, etc).

- RFC mais philosophie application ALF (Application Level Framing).

Réseaux non déterministes (2)

- Réservation (protocole RSVP). Contrôle d'admission.
- Services temps réel, services élastiques.

1. L'approche intserv [SHE 97, WRO 97] :
   - Mise à l'échelle difficile : état/routeur + signalisation, mécanisme de bout en bout $\Longrightarrow$ LAN ?
2. L'approche diffserv [BLA 98b] : différenciation de services.
   - Classification et agrégation de trafics. Champ DS dans le paquet IP.
   - Pas (ou peu) de signalisation.
   - Mise à l'échelle plus aisée.
   - Notion de domaine.

Partie 5

Mémoire et disques

La gestion des supports de masse (1)

- Principalement magnétique/optique [DEM 99].
- Caractéristiques :
  - Progrès significatif en terme de volume.
  - Débit toujours insuffisant.
- Critères de performance :
  - Délai de rotation (fixe pour le magnétique, variable pour l'optique).
  - Temps de transfert (lecture/écriture).
  - Temps d'accès (10 ms pour le magnétique, 200 ms pour l'optique) $\Longrightarrow$ paramètre principalement ciblé.

La gestion des supports de masse (2)

- Problématique :

1. Débit ; systèmes de disques parallèles (RAID [PAT 88]).
2. Déterminisme temporel : minimisation de la gigue ou de la latence.
   - Principal indéterminisme : le temps d'accès.
   - Ordonnancement des requêtes (EDF, SCAN, SSTF).
   - Placement des données (contigues pour les flots continus) $\Longrightarrow$ fragmentation interne.

La gestion de la mémoire

- Parent pauvre dans les systèmes multimédias.

Deux problèmes traités :

1. Dimensionnement : dans les extrémités : compensation de gigue (t=2*gigue*débit), partage d'informations et traitement (pipe-line, taille fonction du format de données traité plutôt que de la QoS).
2. Allocation. Actuellement statique. Peu de solutions pour allocation dynamique et ramasse-miettes (ex : avec RM [HEN 97]). Problèmes principaux $\Longrightarrow$ fragmentation, concurrence avec le ramasse-miettes.

   $\Longrightarrow$ et pourtant, applications dynamiques, Java temps réel.

Partie 6

Conclusion, synthèse et perspectives

Conclusion, synthèse et perspectives

- Système multimédia : système manipulant plusieurs médias différents $\Longrightarrow$ besoins spécifiques (expression + gestion).
- Difficultés liées :
  - Eléments hétérogènes ; nombreuses ressources très différentes.
  - Cohabitation applications temps partagé.
  - Systèmes partiellement déterministes.
- Bilan :
  - Nombreux mécanismes et politiques proposés. Systèmes fermés bien maîtrisés.
  - Couverture de (presque) toutes ressources et catégories d'applications.
  - Manque d'intégration $\Longrightarrow$ système complexe.
  - Tendance middleware, systèmes non dédiés.

Conclusion, synthèse et perspectives

- Perspectives :
  - Ressources non/mal traitées (mémoire, disques) $\Longrightarrow$ applications de + en + dynamiques.
  - Intégration $\Longrightarrow$ complexité = frein à la diffusion de la technologie.
  - Comportements stochastiques (ex : Tenet [BAN 96]) $\Longrightarrow$ problèmes non résolus, méfiance des utilisateurs.
  - Applications nouvelles :
    - Voix sur IP.
    - Environnements virtuels, jeux répartis.
    - Applications mobiles (télécoms).

Partie 7

Références

[BAN 96] A. Banerjea, D. Ferrari, B. Mah, M. Moran, D. Verma, et H. Zhang. « The Tenet Real-Time Protocol Suite : Design, Implementation, and Experiences ». *IEEE/ACM Transactions on Networking*, 4(1):1–11, February 1996.

[BLA 98a] G. Blair et J. B. Stefani. *Open Distributed Processing and Multimedia*. Addison-Wesley, 1998.

[BLA 98b] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, et W. Weiss. « RFC2475 : An architecture for Differentiated Services ». Network Working Group, pages 1-33, December 1998.

[BRA 98] S. Brandt, G. Nutt, T. Berk, et M. Humphrey. « Soft Real-Time Application Execution with Dynamic Quality of Service Assurance ». International Workshop on Quality of Service (IWQOS'98), Napa - CA, May 1998.

[DEM 99] I. Demeure et C. Bonnet. *Introduction aux systèmes temps réel*. Collection pédagogique de télécommunications, Hermes, September 1999.

[HAN 96] C. C. Han, K. J. Lin, et C. J. Hou. « Distance-constrained Scheduling and Its Applications to Real-time Systems ». *IEEE Transactions on computers*, 45(7):814–826, July 1996.

[HEN 97] R. Henriksson. « Predictable Automatic Memory Management for Embedded Systems ». Dans Peter Dickman et Paul R. Wilson, éditeurs, *OOPSLA '97 Workshop on Garbage Collection and Memory Management*, October 1997.

[JEF 92] K. Jeffay, D. L. Stone, et F. D. Smith. « Kernel Support for Live Digital Audio and Video ». *Computer Communications*, 15(6):388–395, August 1992.

[JEF 95] K. Jeffay et D. Bennett. « A Rate-Based Execution Abstraction For Multimedia Computing ». *In Lectures Notes in Computing Science, T. D. C. Little and R. Gusella, Springer-Verlag, Heidelberg*, 1018:64–75, April 1995.

[JON 97] M. Jones, D. Rosu, et M. Rosu. « CPU Reservations and Time Constraints: Efficient, Predictable Scheduling of Independent Activities ». 16th ACM Symposium on Operating Systems Principles in Saint-Malo (SOSP'97) - France, October 1997.

[LIN 88] K. J. Lin et S. Natarajan. « Expressing and Maintaining Timing Constraints in Flex ». pages 96–105. In Real Time Systems Symposium, December 1988.

[LIU 73] C. L. Liu et J. W. Layland. « Scheduling Algorithms for Multiprogramming in a Hard Real-Time Environnment ». *Journal of the Association for Computing Machinery*, 20(1):46–61, January 1973.

[NIE 97] J. Nieh et M. Lam. « The Design, Implementation and Evaluation of SMART: A Scheduler for Multimedia Applications ». 16th ACM Symposium on Operating Systems Principles in Saint-Malo (SOSP'97) - France, October 1997.

[PAT 88] D. A. Patterson, G. Gibson, et R. H. Katz. A Case for Redundant Arrays of Inexpensive Disks (RAID). pages 109–116. June 1988.

[SAK 95] M. Saksena, R. Gerber, et W. Pugh. « Parametric Dispatching of Hard Real-time Tasks ». *IEEE Trans. on Computers*, 44(3):471–479, 1995.

[SCH 96] H. Schulzrinne, S. Casner, R. Frederick, et V. Jacobson. « RFC1889 : RTP : A Transport Protocol for Real-Time Applications». Network Working Group, pages 1-75, January 1996.

[SHE 97] S. Shenker, C. Partridge, et R. Guerin. « RFC2212 : Specification of Guarantedd Quality of Service ». Network Working Group, pages 1-19, September 1997.

[SIJ 96] P. Sijben et S. J. Mullender. « Quality of Service in Distributed Multimedia Systems ». in Trends in distributed systems, Springer Lectures Notes on computing systems, TREVS 1161, 1996.

[SPR 89] B. Sprunt, L. Sha, et J.P. Lehoczky. « Aperiodic task scheduling for hard-real-time systems ». *The Journal of Real Time Systems*, 1:27–60, 1989.

[TIN 94] K. W. Tindell et J. Clark. « Holistic schedulability analysis for distributed hard real-time systems ». *Microprocessing and Microprogramming*, 40(2-3):117–134, April 1994.

[WRO 97] J. Wroclawski. « RFC2210 : The Use of RSVP with IETF Integrated Services ». Network Working Group, pages 1-31, September 1997.