Lezione R8

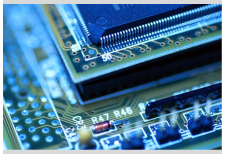
Algoritmi a conservazione di banda

Sistemi embedded e real-time

27 novembre 2012

Marco Cesati

Dipartimento di Ingegneria Civile e Ingegneria Informatica
Università degli Studi di Roma Tor Vergata

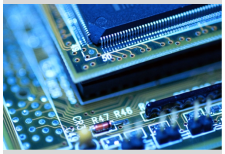


Di cosa parliamo in questa lezione?

In questa lezione si discutono alcuni algoritmi a conservazione di banda utilizzati per integrare la gestione dei job aperiodici con gli schedulatori priority-driven

1. I server periodici
2. Il server procastinabile
3. Il server sporadico
4. Server sporadico e schedulazione EDF
5. Algoritmo GPS e job aperiodici hard RT

Algoritmi a conservazione di banda
Marco Cesati



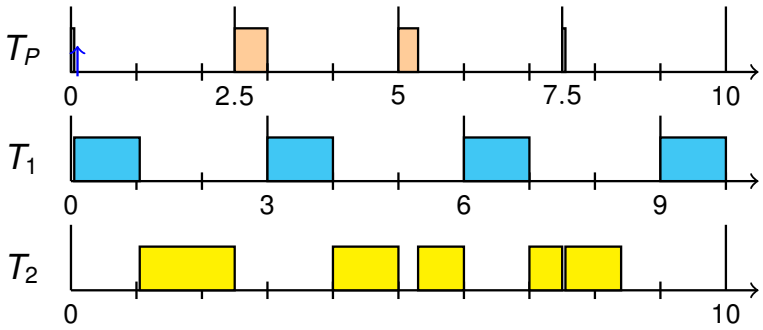
Schedulazione di job aperiodici con polling

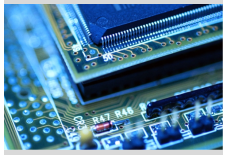
L'algoritmo di *schedulazione con polling* è basato su un task periodico (*server di polling* o *poller*) con fase 0, periodo p_s, tempo d'esecuzione e_s, e priorità massima

Il server di polling controlla la coda di job aperiodici: se è vuota, si auto-sospende fino al prossimo periodo, altrimenti esegue il job in cima alla coda per max e_s unità di tempo

- Se i parametri del poller sono corretti, i job aperiodici non influiscono sulla schedulabilità dei task periodici
- Se il job aperiodico arriva subito dopo l'inizio del periodo del poller, non sarà eseguito fino al periodo successivo (i tempi di risposta non sono minimizzati)

T_P=(2.5, 0.5)
T_1=(3, 1)
T_2=(10, 4)
Job aper. A:
rilascio a 0.1
durata 0.8





Server periodici

I *server periodici* sono una classe di task periodici aventi:

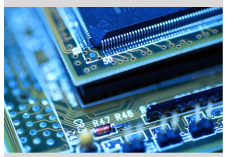
- Periodo p_s, budget e_s, e dimensione $u_s = e_s/p_s$
- Regola di consumo: come il budget viene consumato
- Regola di rifornimento: come il budget viene ripristinato

Si dice che il server periodico è:

- *impegnato* quando ha lavoro da svolgere
- *idle* quando non ha lavoro da svolgere
- *eleggibile*, *pronto* o *schedulabile*: impegnato e con budget positivo

Esempio: il poller è assimilabile ad un server periodico

- impegnato quando la coda di job aperiodici è non vuota
- regola di consumo: sottrae il tempo impiegato ad eseguire un job aperiodico dal budget; azzera il budget se la coda è vuota
- regola di rifornimento: il budget è impostato a e_s all'inizio di ogni periodo



Algoritmi a conservazione di banda

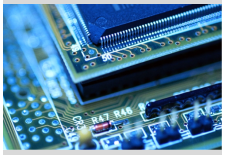
Il problema del server di polling è che il budget è perso non appena la coda di job aperiodici si svuota

Gli algoritmi basati su server periodici che non hanno questo problema sono definiti a *conservazione di banda*

Idea: preservare il budget quando il server periodico è idle per migliorare i tempi di risposta dei job aperiodici

Esistono molti tipi di algoritmi a conservazione di banda

- Server procastinabile
- Server sporadico
- Server a utilizzazione costante
- Server a banda totale
- Algoritmo WFQ



Server procastinabile

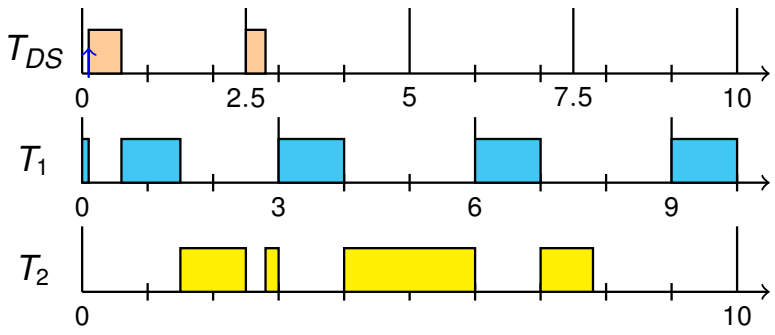
Il *server procastinabile* (o *deferrable server*) è il più semplice algoritmo a conservazione di banda

È caratterizzato da un periodo p_s, da un budget massimo e_s, e dalle seguenti regole:

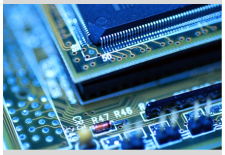
- Regola di consumo: il budget è decrementato di uno per ogni unità di tempo in cui il server è in esecuzione
- Regola di rifornimento: il budget è impostato al valore e_s agli istanti $k \cdot p_s$, per $k = 0, 1, 2, \ldots$

Nota: il budget non si accumula (quello non speso alla fine del periodo viene perso)

T_{DS}=(2.5, 0.5)
T_1=(3, 1)
T_2=(10, 4)
Job aper. A:
rilascio a 0.1
durata 0.8



Algoritmi a conservazione di banda
Marco Cesati

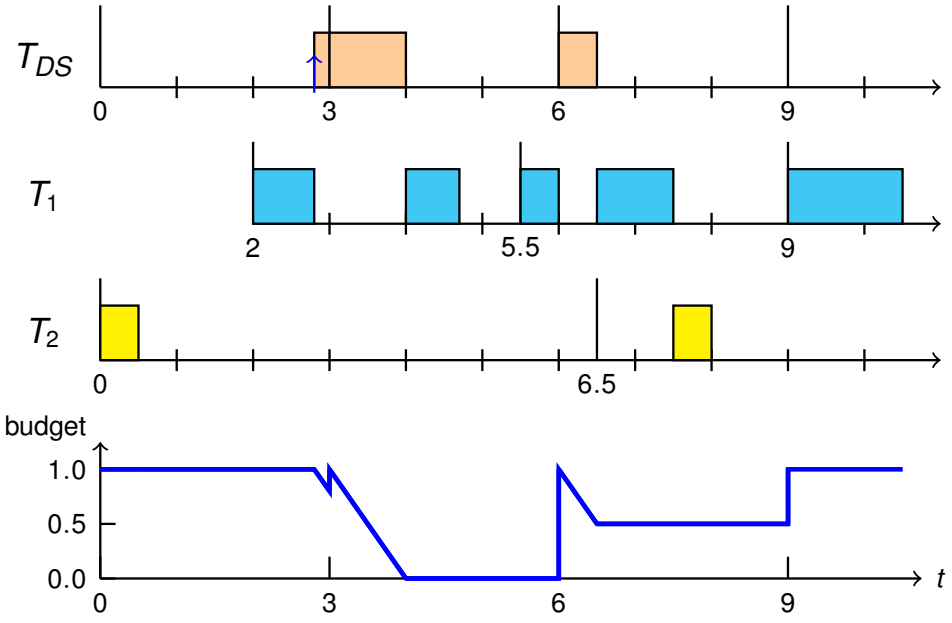


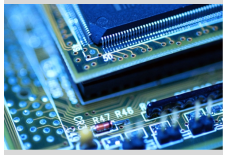
Schema della lezione
Server periodici
Server procastinabili
Server sporadici
Server sporadici+EDF
Algoritmo GPS

Schedulazione a priorità fissa con server procastinabile

Sistema: T_{DS}=(3, 1), T_1=(2.0, 3.5, 1.5, 3.5), T_2=(6.5, 0.5)

Job aperiodico A: arrivo a 2.8, durata 1.7

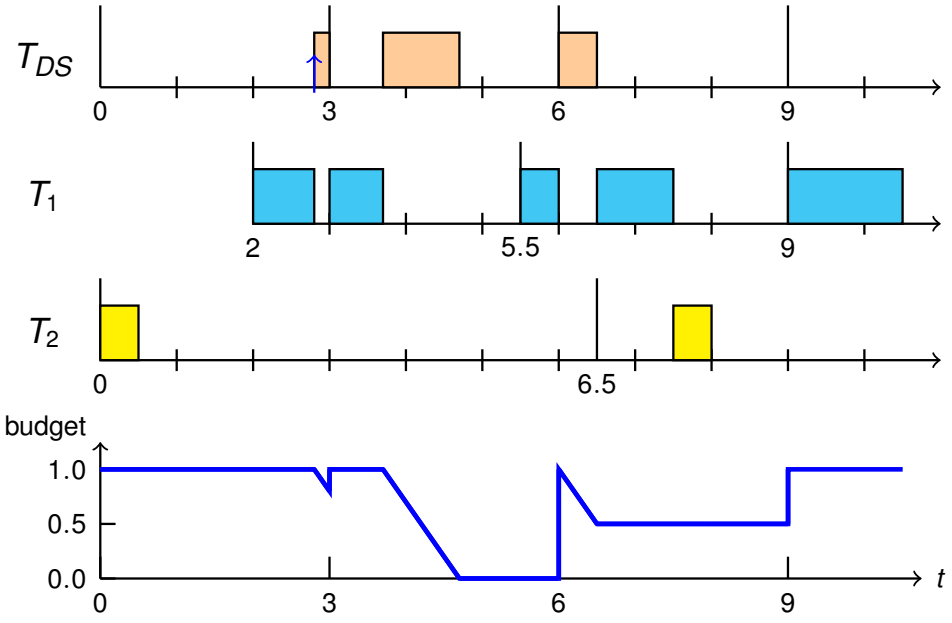


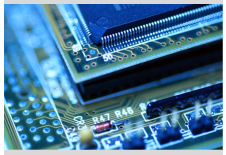


Schedulazione EDF con server procastinabile

Sistema: T_{DS}=(3, 1), T_1=(2.0, 3.5, 1.5, 3.5), T_2=(6.5, 0.5)

Job aperiodico A: arrivo a 2.8, durata 1.7





Schedulabilità per priorità fissa con server procastinabile

È possibile applicare il test o le condizioni di schedulabilità per sistemi a priorità fissa con server procastinabile? ***Sì, ma...***

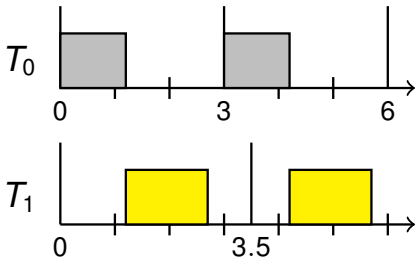
Il server procastinabile non è identico agli altri task periodici:

- Se il server è eleggibile e nessun task a priorità maggiore è in esecuzione, viene subito attivato dallo scheduler
- Un server con budget > 0 può diventare eleggibile in qualunque istante (dipende dagli arrivi dei job aperiodici)

T_0=(3, 1.2), T_1=(3.5, 1.5)

$w_1(t) = 1.5 + \lceil t/3 \rceil 1.2$

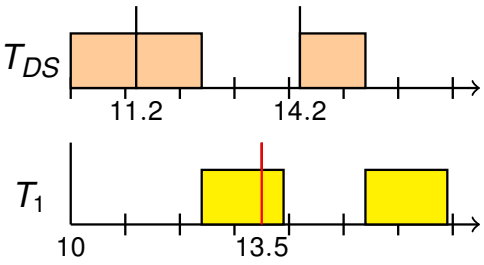
$w_1(1.5)$=2.7=$w_1(2.7) \leq 3.5$

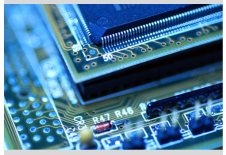


T_{DS}=(3, 1.2), T_1=(3.5, 1.5)

$r_{1,c}$=10, r_A=10, $e_A > 3$

budget(10)=1.2, fase=1.2





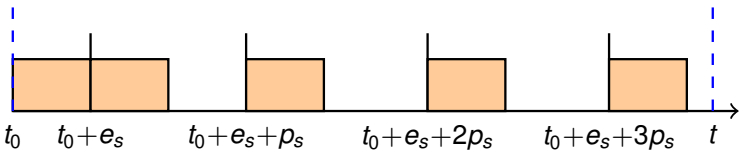
Istanti critici per sistemi con server procastinabile

Lemma (Lehoczky, Sha, Strosnider, 1987, 2000)

In un sistema di task periodici indipendenti e, interrompibili a priorità fissa con $D_i \leq p_i$, e con un server procastinabile (p_s, e_s) con priorità massima, un istante critico di un task T_i si verifica all'istante t_0 se

- a t_0 è rilasciato un job di tutti i task $T_1, \ldots, T_i$
- a t_0 il budget del server è e_s
- a t_0 è rilasciato almeno un job aperiodico che impegna il server da t_0 in avanti
- l'inizio del successivo periodo del server è a $t_0 + e_s$

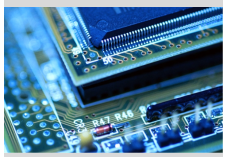
Nelle ipotesi del lemma, quanto tempo di processore occupa al massimo il server procastinabile nell'intervallo $(t_0, t]$?



$$e_s + \left\lceil \frac{t - t_0 - e_s}{p_s} \right\rceil \cdot e_s$$

Algoritmi a conservazione di banda

Marco Cesati



Schema della lezione
Server periodici
Server procastinabili
Server sporadici
Server sporadici+EDF
Algoritmo GPS

Test di schedulabilità con server procastinabile

Con priorità fissate ed un server procastinabile di massima priorità, la funzione di tempo richiesto è:

$$w_i(t) = e_i + b_i + e_s + \left\lceil \frac{t - e_s}{p_s} \right\rceil e_s + \sum_{k=1}^{i-1} \left\lceil \frac{t}{p_k} \right\rceil e_k \quad \text{per } 0 < t \le p_i$$

Il test controlla se $w_i(t) \le t$ per i valori di $t \le D_i$ tali che $t = h \cdot p_k$, oppure $t = e_s + h \cdot p_s$, oppure $t = D_i$ $(h = 0, 1, \ldots)$

Analogamente per il test di schedulabilità generale:

$$w_{i,j}(t) = j\, e_i + b_i + e_s + \left\lceil \frac{t - e_s}{p_s} \right\rceil e_s + \sum_{k=1}^{i-1} \left\lceil \frac{t}{p_k} \right\rceil e_k$$

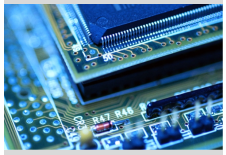
$$\text{per } (j-1)\, p_i < t \le w_{i,j}(t)$$

Esempio: T_{DS}=(3, 1.2), T_1=(3.5, 1.5)

$w_1(t) = 2.7 + \lceil (t - 1.2)/3 \rceil\, 1.2$

$w_1(1.5) = 3.9 = w_1(3.9) > 3.5 \quad \Rightarrow \quad T_1$ non schedulabile!

Se il server non ha priorità massima, il test fornisce una condizione solo sufficiente (può dare falsi negativi)



Condizione di schedulabilità RM con server procastinabile

Teorema (Lehoczky, Sha, Strosnider, 1987, 2000)

Un server procastinabile (p_s, e_s) ed n task periodici indipendenti e interrompibili con $p_i = D_i$ tali che

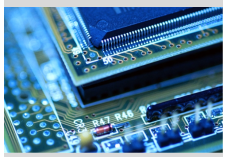
$$p_s < p_1 < \cdots < p_n < 2\, p_s \quad \text{e} \quad p_n > p_s + e_s$$

sono schedulabili con RM se l'utilizzazione totale dei task periodici e del server è minore o uguale a

$$U_{RM/DS}(n) = \frac{e_s}{p_s} + n \left[\left(\frac{e_s + 2\, p_s}{p_s + 2\, e_s} \right)^{1/n} - 1 \right]$$

- Se $e_s = 0$, $U_{RM/DS}(n) = U_{RM}(n)$
- $\lim_{n \to \infty} U_{RM/DS}(n) = \frac{e_s}{p_s} + \ln \left(\frac{e_s + 2\, p_s}{p_s + 2\, e_s} \right)$
- per ogni n, $U_{RM/DS}(n) \ge 0.652$

Algoritmi a conservazione di banda
Marco Cesati



Schema della lezione
Server periodici
Server procastinabili
Server sporadici
Server sporadici+EDF
Algoritmo GPS

Condizione di schedulabilità RM con server procastinabile (2)

Se $p_s, p_1, \ldots, p_n$ non verificano le condizioni del teorema?

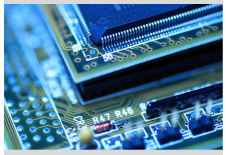
Applichiamo la condizione di schedulabilità task per task:

- Il server non ha alcuna influenza sui task aventi periodo minore di p_s
- Il server è schedulabile se lo è il corrispondente task periodico
- Per ogni task T_i con $p_i > p_s$, il server si comporta come un task periodico, tranne che può eseguire per un tempo e_s in più (tempo di blocco aggiuntivo):

$$\sum_{k=1}^{i} \frac{e_k}{p_k} + \frac{e_s}{p_s} + \frac{e_s + b_i}{p_i} \leq U_{RM}(i+1)$$

Esempio: T_{DS}=(3, 1.2), T_1=(3.5, 1.5)

$$\frac{1.5}{3.5} + \frac{1.2}{3} + \frac{1.2}{3.5} > 1 > U_{RM}(2) \Rightarrow T_1 \text{ forse non schedulabile!}$$



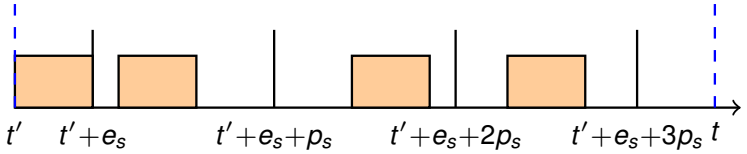
Condizione di schedulabilità EDF con server procastinabile

Teorema (Ghazalie, Baker 1995)

Un task periodico T_i in un sistema di n task indipendenti e interrompibili è schedulabile con EDF insieme ad un server procastinabile (p_s, e_s) se

$$\sum_{k=1}^{n} \frac{e_k}{\min(D_k, p_k)} + \frac{e_s}{p_s}\left(1 + \frac{p_s - e_s}{D_i}\right) \leq 1$$

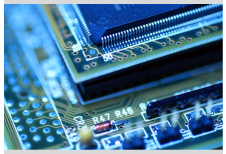
Dim. (sketch per $D_k \geq p_k$) Un job di T_i rilasciato a r_i manca la scadenza a t; $t' < t$ è l'ultimo istante in cui il processore è idle o esegue un job con scadenza $> t \Rightarrow r_i \geq t' \Rightarrow 1/(t-t') \leq 1/D_i$



$$e_s + \left\lfloor \frac{t-t'-e_s}{p_s} \right\rfloor \cdot e_s$$

$$t - t' < \sum_{k=1}^{n} \frac{e_k}{p_k}(t-t') + \frac{e_s}{p_s}(t - t' + p_s - e_s)$$

Algoritmi a conservazione di banda
Marco Cesati



Schema della lezione
Server periodici
Server procastinabili
Server sporadici
Server sporadici+EDF
Algoritmo GPS

Server sporadici

Un server procastinabile può ritardare i task di priorità minore più di un task periodico con identici parametri

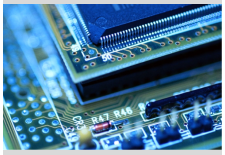
I *server sporadici* sono una classe di server periodici completamente assimilabili come schedulabilità a task periodici con medesimi parametri

Un sistema con task periodici e server sporadici può essere analizzato tramite le condizioni ed il test di schedulabilità *generale* dei sistemi per task periodici

Esistono diversi tipi di server sporadici: la differenza è tutta nelle due regole di consumo e di rifornimento del budget

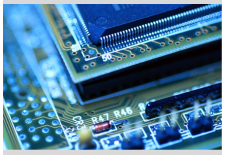
Regole più sofisticate:

- preservano il budget più a lungo o lo riforniscono più velocemente
- sono più difficili e costose da implementare



Definizioni per server sporadici in sistemi a priorità fissa

- Sistema $\mathcal{T}$ di task periodici a priorità fissa
- Server sporadico $T_s=(p_s, e_s)$ con priorità π_s
- $\mathcal{T}_H$: insieme di task di $\mathcal{T}$ con priorità maggiore di π_s
- *Intervallo totalmente occupato* di un insieme di task:
  (1) prima dell'intervallo tutti i job sono stati completati,
  (2) all'inizio viene rilasciato almeno un job, e
  (3) la fine dell'intervallo è il primo istante in cui tutti i job rilasciati entro l'intervallo sono completati
- t_r: ultimo istante in cui è stato aumentato il budget
- t_f: primo istante dopo t_r in cui il server è in esecuzione
- t_e: istante che determina il momento del prossimo rifornimento (generalmente sarà a t_e+p_s)
- BEGIN: per ogni t, considerare l'ultima sequenza di intervalli totalmente occupati contigui dei task $\mathcal{T}_H$ iniziata prima di t; BEGIN è l'istante di inizio del primo intervallo totalmente occupato di questa sequenza
- END: l'istante finale della sequenza, se precedente a t, altrimenti ∞



Server sporadico semplice

Regola di consumo

In ogni istante $t > t_r$, il budget è decrementato di uno per ogni unità di tempo se una delle due condizioni C1 e C2 è vera:

C1 Il server è in esecuzione

C2 Il server è stato in esecuzione dopo t_r e inoltre END $< t$

Altrimenti (se C1 e C2 sono false) il budget è conservato

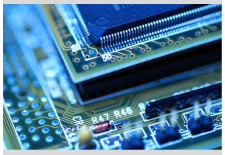
Regola di rifornimento

R1 Ad ogni rifornimento: budget $\leftarrow e_s$, $t_r \leftarrow$ istante corrente

R2 All'istante t_f: se END $= t_f$, $t_e \leftarrow \max(t_r, \text{BEGIN})$; se END $< t_f$, $t_e \leftarrow t_f$

R3 Il prossimo rifornimento sarà a t_e+p_s, con due eccezioni:

- **(a)** se $t_e+p_s < t_f$, il budget sarà rifornito non appena esaurito
- **(b)** il budget sarà rifornito a $t_b < t_e+p_s$ se esiste un intervallo $[t_i, t_b)$ in cui nessun task di $\mathcal{T}$ è eseguibile, ed un task di $\mathcal{T}$ comincia l'esecuzione a t_b



Server sporadico semplice (2)

Significato di C1: nessun job del server esegue per un tempo maggiore di e_s in un periodo p_s

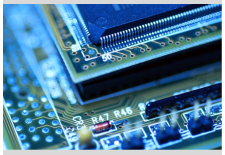
Significato di C2: il server conserva il budget se un task di $\mathcal{T}_H$ è eseguibile oppure il server non ha mai eseguito da t_r; altrimenti il budget è sempre consumato

Significato di R2:

- se nell'intervallo (t_r, t_f) sono stati sempre in esecuzione task di $\mathcal{T}_H$, il prossimo rifornimento sarà a t_r+p_s
- altrimenti il prossimo rifornimento sarà a t_e+p_s ove t_e è l'ultimo istante di $(t_r, t_f]$ in cui *non* esegue un task di $\mathcal{T}_H$

Significato di R3a: il job del server ha atteso per più di p_s unità di tempo prima di iniziare l'esecuzione, quindi il job continua nel prossimo periodo (è richiesto il test di schedulabilità *generale*)

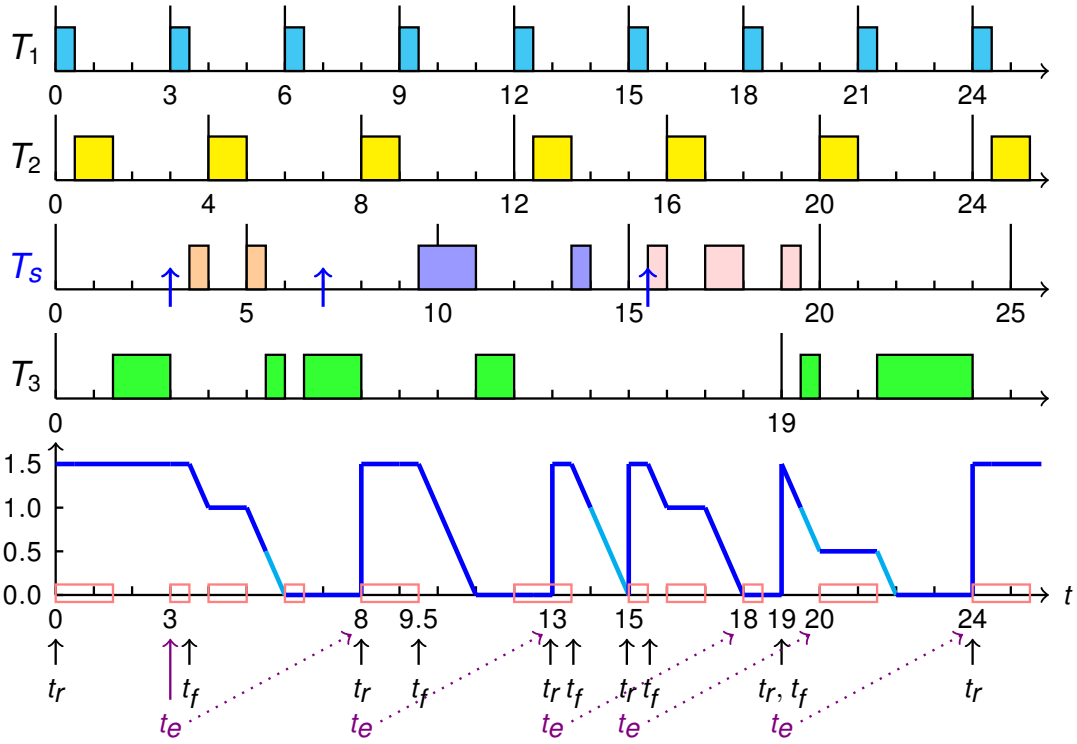
Significato di R3b: il budget è rifornito nell'istante iniziale di ogni intervallo totalmente occupato di $\mathcal{T}$

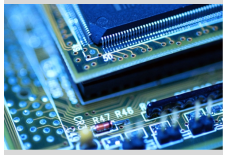


Schedulazione RM con server sporadico semplice

Sistema: T_1=(3, 0.5), T_2=(4, 1), T_s=(5, 1.5), T_3=(19, 4.5)

Aperiodici: A_1(r=3, e=1), A_2(r=7, e=2), A_3(r=15.5, e=2)





Server sporadico/background

Esistono molte varianti di server sporadico, con regole sempre più complesse (e costose da implementare)

Variante più utile e diffusa: *server sporadico/background*

Differenza rispetto al server sporadico semplice: esegue sempre job aperiodici se nessun task periodico è esegubile

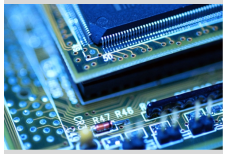
Regola di consumo

Identica a quella del server sporadico semplice, tranne che se nessun task periodico è eseguibile il budget è uguale a e_s

Regola di rifornimento

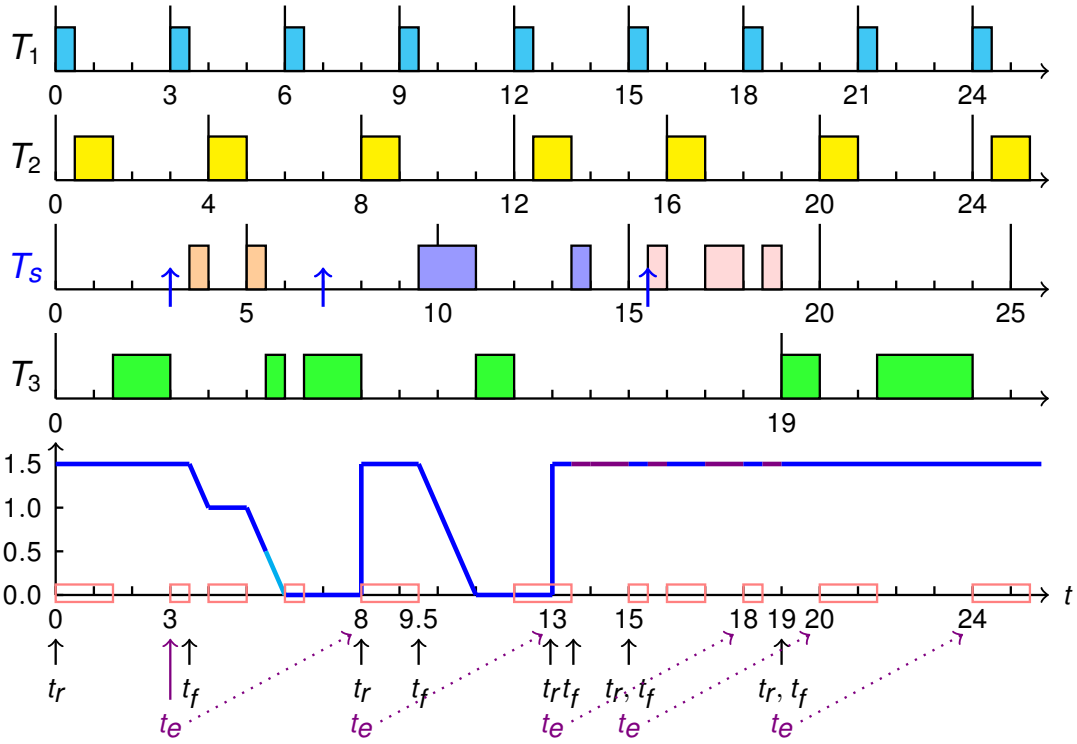
Identica a quella del server sporadico semplice, tranne R3b: il budget è ripristinato all'inizio di ogni intervallo in cui nessun task periodico è eseguibile; t_r (e ev. t_f) è la fine dell'intervallo

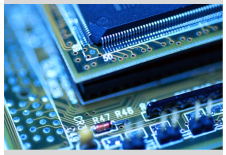
In effetti l'unico caso in cui *non* conviene usare un server sporadico/background al posto di uno semplice è quando si utilizzano più server sporadici per differenti tipi di job aperiodici



Schedulazione RM con server sporadico/background

Sistema: T_1=(3, 0.5), T_2=(4, 1), T_s=(5, 1.5), T_3=(19, 4.5)
Aperiodici: A_1(r=3, e=1), A_2(r=7, e=2), A_3(r=15.5, e=2)





Server sporadico semplice per EDF

Per utilizzare un server sporadico semplice in un sistema schedulato con EDF è necessario modificare le regole di consumo e di rifornimento

In generale, il server è schedulabile se

- È impegnato (esistono job aperiodici in coda)
- È stata definita la sua scadenza d
  (e di conseguenza la sua priorità)

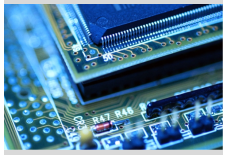
Regola di consumo

Il budget è decrementato di uno per ogni unità di tempo finché almeno una delle due condizioni C1 e C2 è vera:

C1 Il server è in esecuzione

C2 La scadenza del server d è definita, il server non è impegnato, e non c'è alcun job eseguibile con scadenza precedente a d

Altrimenti il budget è conservato



Server sporadico semplice per EDF (2)

Regola di rifornimento

R1 Ad ogni rifornimento: budget $\leftarrow e_s$, $t_r \leftarrow$ istante corrente

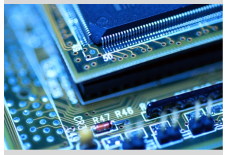
R2 Se t_e è definito, $d = t_e + p_s$; altrimenti d rimane indefinito.
Se t_e non è definito:

- **(a)** Quando all'istante t arriva un job aperiodico e la coda è vuota: se solo job con scadenze precedenti a $t_r + p_s$ sono stati eseguiti in (t_r, t), allora $t_e \leftarrow t_r$; altrimenti $t_e \leftarrow t$
- **(b)** Nell'istante di rifornimento t_r: se il server è impegnato allora $t_e \leftarrow t_r$, altrimenti t_e e d diventano indefiniti

R3 Il prossimo rifornimento sarà a $t_e + p_s$, con due eccezioni:

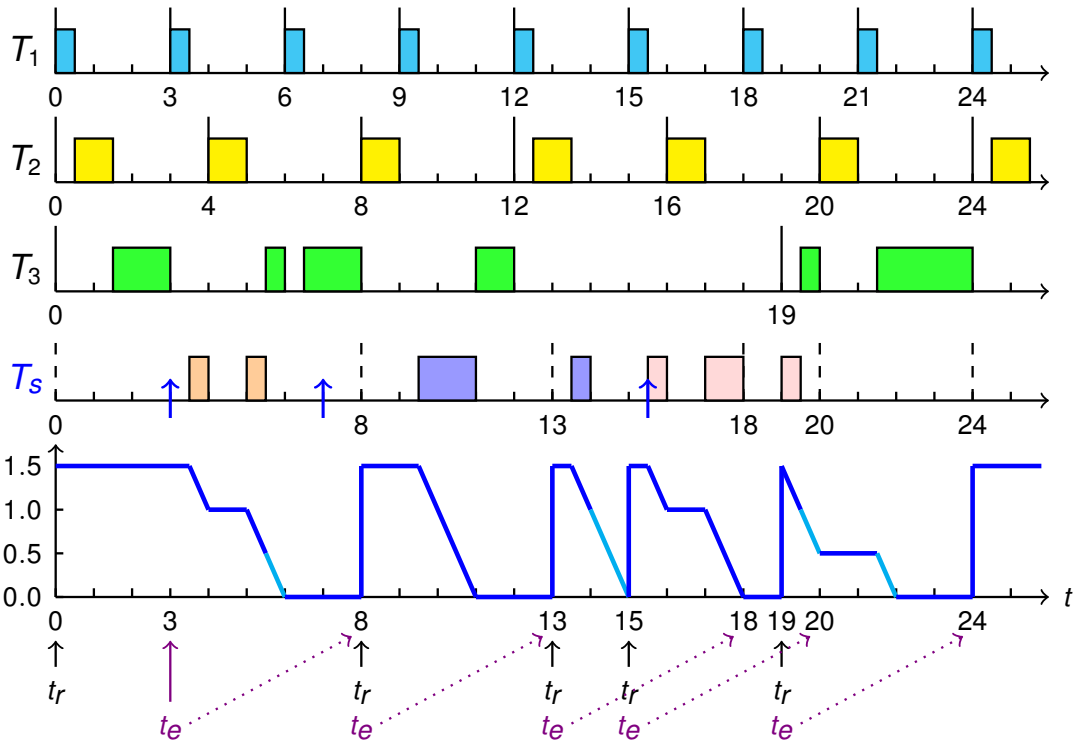
- **(a)** Se $t_e + p_s$ è precedente all'istante successivo a t_r in cui il server è diventato impegnato, il budget è rifornito non appena esaurito
- **(b)** Il budget è rifornito alla fine di ogni intervallo in cui nessun task periodico è eseguibile

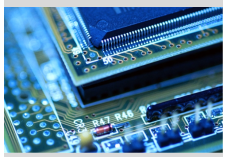
Le regole sono molto simili ed hanno le stesse finalità di quelle del server sporadico semplice per sistemi a priorità fissa



Schedulazione EDF con server sporadico semplice

Sistema: T_1=(3, 0.5), T_2=(4, 1), T_3=(19, 4.5), T_s=(5, 1.5)
Aperiodici: A_1(r=3, e=1), A_2(r=7, e=2), A_3(r=15.5, e=2)





Algoritmo GPS

GPS (Generalized Processor Sharing) è un algoritmo ideale che implementa un meccanismo di round-robin pesato

Idea: ogni server impegnato in ciascun round ottiene una porzione di tempo infinitesimamente piccola ma proporzionale alla propria dimensione e_s/p_s

Quale è il punto di forza di questo algoritmo?

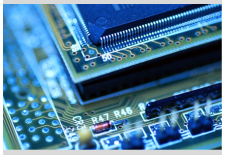
È l'isolamento temporale dei task: il massimo tempo di risposta di un task non dipende dai tempi di esecuzione degli altri task

Quale è lo svantaggio di questo algoritmo?

È impossibile da implementare in pratica!

Si implementano invece delle approssimazioni di GPS:

- Server a utilizzazione costante
- Server a banda totale
- Server WFQ



Schedulabilità EDF di job aperiodici hard real-time

La *densità di un job aperiodico* avente istante di rilascio r, massimo tempo di esecuzione e e scadenza d è $\dfrac{e}{d-r}$

Teorema

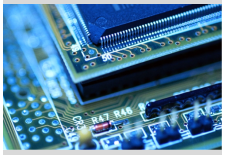
Un sistema di job aperiodici indipendenti e interrompibili è schedulabile con EDF se la densità totale di tutti i job attivi (nell'intervallo tra rilascio e scadenza) è in ogni istante ≤ 1

Dim. (sketch) Un job manca la scadenza a t; sia $t' < t$ l'ultimo istante in cui il processore *non* ha eseguito un job con scadenza $\leq t \quad \Rightarrow \sum_i e_i > t - t'$

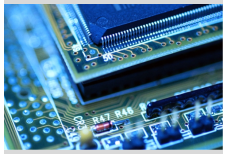
L'intervallo $(t', t\,]$ è partizionato in $(t'{=}t_1, t_2\,]$, $(t_2, t_3\,]$, $\ldots$ $(t_\ell, t_{\ell+1}{=}t\,]$ ove t_k è l'istante di rilascio o scadenza per qualche job

Sia $\mathcal{X}_k$ l'insieme di job attivi in $(t_k, t_{k+1}\,]$ e sia Δ_k la loro densità

$$\sum_i e_i = \sum_{j=1}^{\ell} (t_{j+1} - t_j) \sum_{J_k \in \mathcal{X}_j} \frac{e_k}{d_k - r_k} = \sum_{j=1}^{\ell} \Delta_j \, (t_{j+1} - t_j) \leq t - t'$$



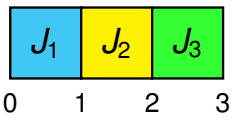
Schedulabilità EDF di job aperiodici hard real-time (2)



Consideriamo i job aperiodici J_1:(r=0, e=1, d= 2), J_2:(r=0.5, e= 1, d=2.5), e J_3:(r=1, e=1, d=3)

Intervalli:	$(0, 0.5]$	$(0.5, 1]$	$(1, 2]$	$(2, 2.5]$	$(2.5, 3]$
Job attivi:	J_1	J_1 J_2	J_1 J_2 J_3	J_2 J_3	J_3
Densità:	0.5	1.0	1.5	1.0	0.5

Sono schedulabili con EDF? ***Sì!***



La condizione del teorema è solo sufficiente!