

Patch PREEMPT-RT : Impacts sur un cas d'utilisation

maxime.chevallier@smile.fr

Open Wide Ingénierie / Smile
Actia

14 juin 2016



- 1 PREEMPT_RT
- 2 Cas d'utilisation
 - Produit
 - Use case
- 3 Impacts
 - load average
 - context switching

- 1 PREEMPT_RT
- 2 Cas d'utilisation
- 3 Impacts

Déterminisme

Historique

- Patch PREEMPT_RT : Ingo Molnar, 2006
- High-resolution timer : Thomas Gleixner

Déterminisme

Historique

- Patch PREEMPT_RT : Ingo Molnar, 2006
- High-resolution timer : Thomas Gleixner

Kernel préemptible

- IRQ dans des threads
- Ne plus désactiver les IRQ
- Eliminer les sections critiques

https://rt.wiki.kernel.org/index.php/Main_Page

- 1 PREEMPT_RT
- 2 Cas d'utilisation
 - Produit
 - Use case
- 3 Impacts



Actia automotive



Actia automotive

Besoins RT

- 4 x Bus CAN
- Faible latence
- Déterminisme



Actia automotive

Besoins RT

- 4 x Bus CAN
- Faible latence
- Déterminisme

Autres besoins

- Wifi, Bluetooth, Ethernet, eMMC, etc.
- Forte charge CPU
- BSP : Adaptabilité

Conditions

- iMX6 Quad core
- linux 3.10.17

Conditions

- iMX6 Quad core
- linux 3.10.17

Benchmark

- CPU : **stress -c 1**
- Mem : **stress -c 1 -m 1**
- Extérieur : **ping -f**

- 1 PREEMPT_RT
- 2 Cas d'utilisation
- 3 **Impacts**
 - load average
 - context switching

load average

Définition

Nombre de process 'Runnable' ou 'Uninterruptible'

Définition complète : **man uptime**

load average

Définition

Nombre de process 'Runnable' ou 'Uninterruptible'

Définition complète : **man uptime**

Outils

- /proc/loadavg
- htop, top
- uptime, w

load average

Sans PREEMPT_RT

```
cat /proc/loadavg : 2.06 2.09 1.87 3/144 1064  
top : %id 91.4 %si 7.7
```

load average

Sans PREEMPT_RT

```
cat /proc/loadavg : 2.06 2.09 1.87 3/144 1064  
top : %id 91.4 %si 7.7
```

Avec PREEMPT_RT

```
cat /proc/loadavg : 2.45 2.62 1.96 4/176 1335  
top : %id 89.0 %si 8.7
```

PID	USER	S	%CPU	%MEM	TIME+	COMMAND
121	root	R	40	0.0	5:00.10	irq/150-2188000

context switching

Définition

- Changement d'espace d'adressage
- Mise à jour des registres
- Lié au scheduling

context switching

Définition

- Changement d'espace d'adressage
- Mise à jour des registres
- Lié au scheduling

outils

- vmstat
- pidstat

vmstat

Sans PREEMPT_RT

r	b	[...]	in	cs	us	sy	id	wa
0	0	[...]	1025	1243	0	0	100	0
0	0	[...]	1023	1242	0	1	100	0
0	0	[...]	1030	1255	0	0	100	0

Sans PREEMPT_RT, ping

r	b	[...]	in	cs	us	sy	id	wa
2	0	[...]	8685	1790	26	36	38	0
2	0	[...]	8776	1846	26	36	38	0
2	0	[...]	8705	1790	26	35	39	0

vmstat

Avec PREEMPT_RT

r	b	[...]	in	cs	us	sy	id	wa
2	0	[...]	1103	2168	25	25	49	0
2	0	[...]	1092	2162	26	25	49	0
3	0	[...]	1103	2170	26	25	49	0

Avec PREEMPT_RT, ping

r	b	[...]	in	cs	us	sy	id	wa
2	0	[...]	11459	10771	22	38	41	0
3	0	[...]	11444	10901	23	38	40	0
3	0	[...]	11456	10838	21	37	42	0

pidstat : Sans PREEMPT_RT

cswch/s	nvcswh/s	Command
0.15	0.02	systemd
76.08	0.00	ksoftirqd/0
0.01	0.00	ksoftirqd/1
111.42	0.88	kworker/0:1
1.10	0.02	kworker/2:2
119.36	0.38	rngd
14.30	0.08	kworker/1:2
1.81	0.40	dbus-daemon
0.50	0.01	watchdog
0.03	0.00	ofonod
0.00	214.69	stress
0.00	102.13	stress
0.22	15.46	pidstat

pidstat : Avec PREEMPT_RT

cswch/s	nvcschw/s	Command
0.14	0.22	systemd
100.96	0.00	ksoftirqd/0
100.93	8.85	ksoftirqd/1
110.24	3.82	ksoftirqd/2
100.00	0.94	ksoftirqd/3
36.40	5.05	kworker/0:1
41.04	2.67	kworker/1:1
4210.75	18.70	irq/150-2188000
118.84	0.37	rngd
0.02	0.00	ofonod
0.00	1588.59	stress
0.03	369.85	stress
0.17	4.78	pidstat

Questions ?