

Mini-projet système : développement d'un noyau de système d'exploitation

Responsable : Christophe RIPPERT
Christophe.Rippert@Grenoble-INP.fr



Mise en place de l'environnement de travail

Introduction

Avant de commencer le mini-projet, il faut mettre en place l'environnement de travail que l'on va utiliser pendant les séances.

Un système d'exploitation s'exécute sur machine nue. Mais pour simplifier le développement et la mise au point des prototypes, nous allons utiliser un émulateur gratuit très populaire appelé QEmu disponible sur la plupart des systèmes. L'intérêt de cet environnement d'exécution est qu'il est portable (vous pouvez l'installer sur vos machines personnelles et travailler en dehors des salles PC) et parfaitement transparent pour votre prototype : il n'y aurait rien à changer dans votre code pour que votre système s'exécute sur une machine nue.

Vous devez tout d'abord récupérer un ensemble de sources de départ que vous devez décompresser dans votre répertoire de travail. Les sources distribuées sont dans le répertoire `src_de_base`, qui contient lui-même les sources du noyau et une mini-bibliothèque C qui vous aidera à développer votre prototype. La compilation d'un noyau se fait simplement en se plaçant dans le répertoire `src_de_base` et en tapant la commande `make` : si tout se passe bien, le binaire `kernel.bin` est produit, il s'agit d'un exécutable un peu différent de ceux que l'on produit habituellement quand on compile un programme C ou assembleur (vous ne pouvez donc pas l'exécuter directement dans un terminal).

Prise en main de l'environnement de développement

Lorsque vous lancez l'exécution du noyau dans QEmu, un certain nombre d'opérations d'initialisation sont effectuées puis la fonction `kernel_start` localisée dans le fichier `start.c` s'exécute : il s'agit du point d'entrée de votre noyau (comme la fonction `main` dans un programme C classique). Dans les sources fournies, cette fonction commence par un appel à la fonction `fact` qui calcule $5!$: il s'agit d'un simple exemple pour vous entraîner à utiliser GDB, vous pourrez supprimer `fact` une fois que vous commencerez le développement de votre noyau.

Pour lancer l'exécution du noyau, vous devez taper la commande ci-dessous dans un 1^{er} terminal :

```
/usr/libexec/qemu-kvm -no-kvm -net nic -net user,tftp='pwd',bootfile=boot.pxe -boot n  
-cpu pentium -rtc base=localtime -m 64M -S -s
```

Comme cette commande devient vite fatigante à taper, le `Makefile` contient une cible `run` qui l'exécute : vous n'avez donc qu'à taper `make run` pour lancer l'exécution du noyau.

Cette commande lance l'émulateur QEmu, et lui dit d'exécuter le *stub* de démarrage `boot.pxe` (fourni et que vous ne devez surtout pas effacer), qui va lui-même charger le noyau `kernel.bin`. L'option `-s` met QEmu en pause juste avant de démarrer le noyau, pour nous permettre d'y connecter un logiciel de visualisation.

En effet, QEmu offre plusieurs façons de gérer un écran virtuel (une console) pour le système émulé. Nous allons utiliser sous Linux le protocole VNC pour lequel on trouve des logiciels de visualisation sur tous les systèmes. Nous allons sous CentOS utiliser le client Vinagre.

Ouvrez un 2^e terminal et taper la commande ci-dessous :

```
vinagre localhost:5900
```

Cette commande connecte le client VNC à l'écran virtuel de QEmu. Initialement, on se retrouve avec un écran noir puisque le système est en pause.

Notez que si Vinagre s'arrête avec une erreur (ou si vous le laissez tourner en arrière plan), le port 5900 ne sera peut-être plus disponible pour les exécutions suivantes : essayez alors les ports 5901, 5902, ...

Mise au point

Pour mettre au point le système, il est souvent utile d'utiliser un logiciel comme GDB. QEmu fournit tout ce qu'il faut pour permettre l'exécution pas à pas du système émulé depuis GDB (par contre, il n'y a pas de support pour Valgrind).

Ouvrez un 3^e terminal et tapez la commande :

```
gdb kernel.bin
```

Cela lance GDB sur le binaire du noyau. Pour connecter le GDB à QEmu, on doit taper dans GDB la commande

```
target remote :1234
```

1234 est ici le numéro du port via lequel QEmu communique avec GDB.

On peut ensuite mettre un point d'arrêt au début du noyau en tapant **b kernel_start** puis lancer l'exécution avec la commande **c (continue)**.

Le noyau démarre alors et s'arrête au début de **kernel_start**. Vous pouvez utiliser les commandes classiques **s (step)**, **n (next)** et **display** pour afficher le contenu de **x** et exécuter pas à pas la fonction factorielle.

Lors de vos développements, vous ferez vraisemblablement des erreurs d'accès mémoire (déréférencement d'un pointeur nul, accès à une zone interdite, ...). Lorsque cela arrivera, la page d'information ci-dessous s'affichera dans l'écran de Vinagre :

Exception caught			
[SPACE] View Screen		[I] Ignore	
		[D] Connect to debugger	
TRAP : 03		ERROR CODE : 00000000	
TSS address : 00020000		Back link (previous TSS) : 0000	
ESP : 0011F400	SS : 0018	ESP0 : 0011F420	SS0 : 0018
ESP1 : 00000000	SS1 : 0000	ESP2 : 00000000	SS2 : 0000
EIP : 001113CB		CS : 0010	EFLAGS : 00000006
EAX : 001000BA	EBX : 0001FFB0	ECX : 00000000	EDX : 00000000
ESI : 00000000	EDI : 03000007	EBP : 0011F418	LDT : 0000
DS : 0018	ES : 0018	FS : 0000	GS : 0000
CR3 (page table) : 00101000		Debug Trap : 0	

Cette page d'information contient notamment :

- le numéro de l'exception (TRAP) et le code d'erreur qui servent à identifier la cause du problème (vous pouvez consulter la liste des exceptions du x86 pour plus d'information) ;
- le contenu des registres du processeur : les plus importants sont le pointeur de pile (ESP), le pointeur d'instruction (EIP) et les registres de segment (principalement CS, DS et SS), leurs valeurs peuvent vous donner une indication sur le problème (e.g. : une valeur de 0 dans un registre-pointeur est rarement une bonne chose!).

Lorsqu'on arrive sur cette page d'information, on peut aussi appuyer sur la touche espace qui permet de basculer entre l'affichage de la page d'information et de l'écran normal du système, pour voir d'éventuelles traces.

Les informations affichées sur cette page d'information sont bien sûr accessibles dans GDB, si vous connaissez les commandes ésotériques appropriées !