Syscally

- Linux má několik set syscallů
  - kernel 2.4-190 sc, 2.6.0-271 sc, 2.6.18-316 sc.
- V linuxu je většina syscallu realizovana přes knihovní obalovací funkce je to vhodné pro linker
- Knihovny glibc (glibc1)
- pokud potřebujeme syscall z nějakého důvodu volat sami, můžeme použít makro _syscallX
  - X je počet parametrů
-

Syscally

- Příklad
  -

```
#include <linux/unistd.h>
    #include <stdio.h>

    /* this macro expands to a function definition returning "int" called
     * "getuid" ( (int)(*getuid)() ) */
    _syscall0(int,getuid)

    int main(void)
    {
       printf( "my uid is %d\n", getuid() );
       return 0;
    }
```

Syscally

- Příklad, jak to dopadne na starším intelu (asi do P2)
  - ```
    int getuid (void)
      {
        long __res;
        __asm__ volatile ( "int $0x80"
              : "=a" (__res)
              : "0" (__NR_getuid) );
        do {
            if( __res >= -125 )
            {
              errno = -__res;
              __res  = -1;
            }
            return  __res;
        } while (0);
      }
    ```
    ```

Syscally

- Jak je realizován syscall na Intelu
  - dvě možnosti
    - interrupt 0x80
    - Instrukce sysenter a sysexit
  - instrukce sysenter a sysexit byly zavedeny kvůli efektivitě
    - zatímco mov vyžadoval na P3 dva tiky, přechod do kernel módu asi 300 tiků (jen vstoupit do kernel módu bylo drahé)

Syscally

- Mimo jiného (viz dále) kvůli těmto dvou možnostem je generována dynamicky knihovna linux-gate, příklad implementace pro sysenter a sysexit a pro int $0x80
- Kernel si při bootu vybere co bude používat a podle toho vygeneruje danou knihovnu

–

```
__kernel_vsyscall: push %ecx
push %edx
push %ebp
__resume: mov %esp,%ebp
sysenter
jmp __resume
__return: pop %ebp
pop %edx
pop %ecx
ret
```

```
__kernel_vsyscall: int $0x80
retl
```

Syscally

- Dalším důvodem proč generovat tuto knihovnu je přenesení některých syscallů které se volají často z kenrl módu do userlandu, což je zlevní.
  - Klasickým podobným voláním je gettimeofday(), které volá např X-server skoro pořád

Syscally

- Knihovna linux-gate je linkovana k mnoha programům

```
tom@tom-XX:~> ldd /bin/sh
        linux-gate.so.1 =>  (0xffffe000)
        libreadline.so.5 => /lib/libreadline.so.5 (0xb7f51000)
        libhistory.so.5 => /lib/libhistory.so.5 (0xb7f4a000)
        libncurses.so.5 => /lib/libncurses.so.5 (0xb7f03000)
        libdl.so.2 => /lib/libdl.so.2 (0xb7eff000)
        libc.so.6 => /lib/libc.so.6 (0xb7dde000)
        /lib/ld-linux.so.2 (0xb7f9d000)
```

Syscally

- Parametry
  - paramentry do určitého počtu (myslím 6 ??) jsou předávány v registrech
  - Může mít syscall víc parametrů ??
  - Zkuste si

```
ldd /bin/sh
cat /proc/self/maps
dd if=/proc/self/mem of=linux-gate.dso bs=4096 skip=1048574 count=1
objdump -T linux-gate.dso
```