# Introduction au Développement en C sous UNIX

(4ème édition)

#### **Bruno Quoitin**

(bruno.quoitin@uclouvain.be)

#### Sébastien Barré

(sebastien.barre@uclouvain.be)

#### Virginie Van den Schrieck

(virginie.vandenschrieck@uclouvain.be)

### Introduction

#### Constatation

- Programmation en C/C++
  - → Toujours largement utilisée
  - Systèmes d'exploitation
  - Protocoles et applications réseau
  - Systèmes embarqués
  - Majorité des utilitaires/applications UNIX

#### Difficultés

- Gestion de la mémoire: segmentation faults, bus errors, memory leaks, ...
- Erreurs de syntaxe / type
- ◆Objets non trouvés: header (.h) / librairies / symbôles
- Organisation / maintenance du code

### Introduction

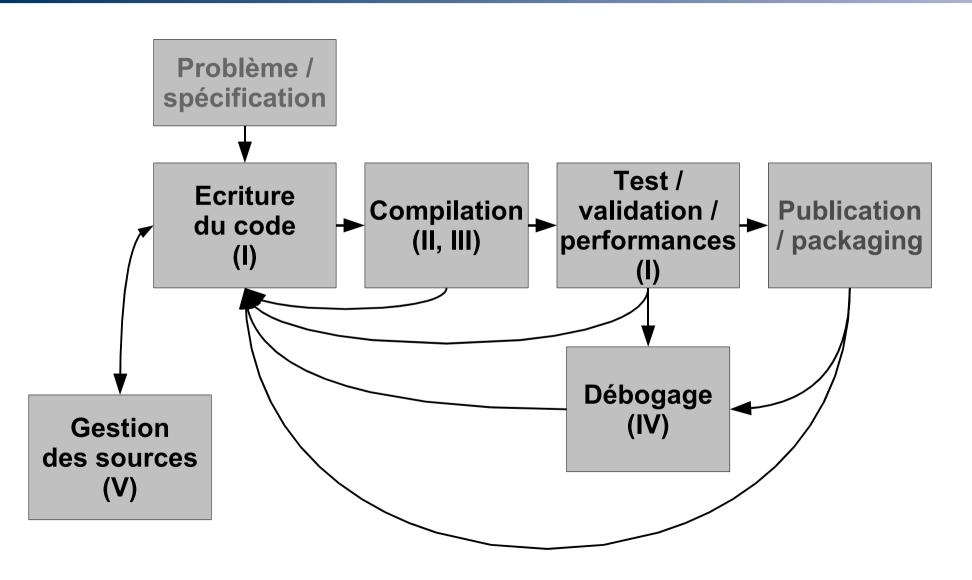
#### But du cours:

- Donner quelques <u>règles de bonne pratique</u>
  - Organisation d'un projet
  - Utilisation d'un système de contrôle de version
  - Conventions de codage
- Introduire les <u>outils disponibles</u> sous UNIX / Linux
  - ◆ Compilation C: gcc (et cpp)
  - → Débogueur: gdb (ddd), valgrind, strace
  - ◆ Automatisation de la compilation: make
  - Contrôle de version: cvs / svn
- Organisation
  - ◆ 1-2 heures de cours
  - ◆ Prérequis: notions de base de C

### Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
- V. Contrôle de version
- VI. Conclusion
- VII. Références

# Processus de développement



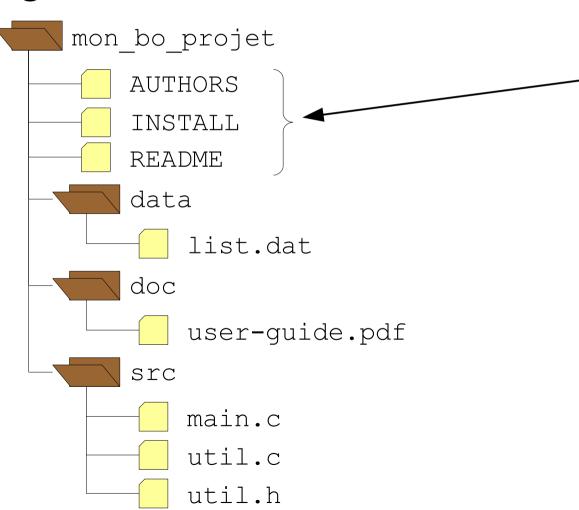
### Table des matières



#### I. Recommendations générales

- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
- V. Contrôle de version
- VI. Conclusion
- VII. Références

### Organisation des sources



#### **Bonne pratique:**

placez à la racine de votre projet les fichiers

- **README** décrivant ce que fait le projet
- INSTALL comment compiler et installer
- AUTHORS contenant la liste des auteurs

#### **Bonne pratique:**

séparez les fichiers selon leur fonction: sources (src), données (data), documentation (doc), ...

### Organisation des sources

```
* My very first C program
* (C) 1980, B. Gates
 * Intended license: public domain
  ----- * /
#include <stdio.h>
// ----[ main ]-----
* Entry point
main()
 printf("MS-DOS 0.1 starting...\n);
 return 0;
```

## Bonne pratique: au début du source

- décrivez le contenu
- copyright, licence, auteur(s)

#### **Bonne pratique:**

- limitez la longueur des lignes à 80 caractères
- limitez la longueur des fonctions

#### Gestion des erreurs

 Testez les arguments passés par l'utilisateur (nombre et valeurs)

- Codes de sortie standards (ou documentés)
  EXIT\_FAILURE / EXIT\_SUCCESS (stdlib.h)
- Reportez les erreurs sur la sortie d'erreur standard !

```
fprintf(stderr, "Erreur: qquechose a foiré !\n");
```

#### Gestion des erreurs

- Testez le code de retour des appels système !
  - ◆ Convention: 0 ⇒ succès, < 0 ⇒ erreur</p>
  - Vérifier dans les "man pages"

```
in_stream= fopen("list.txt", "r");
if (in_stream == NULL) {
   perror("Erreur: impossible d'ouvrir \"list.txt\"");
   exit(EXIT_FAILURE);
}
```

 Code d'erreur du dernier appel système disponible dans la variable globale erro (erro.h)

#### **Bonne pratique:**

perror() est une fonction standard qui affiche la raison pour laquelle un appel système a échoué (défini dans stdio.h).

- Gestion du tas (heap)
  - Même malloc() peut échouer s'il n'y a plus assez de mémoire

```
char * str= (char *) malloc(100*sizeof(char));
if (str == NULL) {
   perror("Erreur: pas possible d'allouer de la mémoire");
   exit(EXIT_FAILURE);
}
```

Ne pas oublier de libérer la mémoire

```
free(str);
```

### Lisez la documentation (man pages) !!!

→ Tapez "man <cmd>" dans une console pour obtenir de l'info sur la commande ou l'appel système <cmd>. man gdb -> donne de l'info sur le débogueur gdb man fopen -> donne de l'info sur la fonction d'ouverture de fichier fopen

# Exemple: la fonction gets() "BUGS

Never use gets(). Because it is impossible to tell without knowing the data in advance how many characters gets() will read, and because gets() will continue to store characters past the end of the buffer, it is extremely dangerous to use. It has been used to break computer security. Use fgets() instead."

#### Documentation

- Liste des fonctions C standard:
  - http://www.utas.edu.au/infosys/info/documentation/C/CStdLib.htm
  - http://www.acm.uiuc.edu/webmonkeys/book/c\_guide/index.html

### Tests

#### Pre-condition Tests

- Vérifier des conditions dans le programme
- Tests internes

#### Validation Tests

Vérifier les fonctionnalités d'un programme

#### Regression Tests

- Vérifier les fonctionnalités et l'interface d'un programme / librairie après modification
- Tests généralement externes (scripts, programme utilisateur, expect)

### Pre-condition Tests

#### Fonctions / macros utiles

assert(CONDITION)

```
ptr= malloc(...);
assert(ptr != NULL);
```

abort()

```
if (CONDITION == 0)
  abort();
```

→ macros LINE , FILE (, FCT )

### **Tests Unitaires**

### Unit Testing

- CUnit (~JUnit): http://cunit.sourceforge.net
- Dans libgds: http://libgds.info.ucl.ac.be

#### Guidelines

- Tests courts
- Une seule fonctionnalité par test
- Beaucoup de tests

### Tests Unitaires

#### Exemple

### Définissez des convention de codage !

- Raisons:
  - Lisibilité et vérifiabilité du code accrues
  - Travail en groupe facilité
- Que couvre une telle convention ?
  - Nomage des variables, fonctions, fichiers
  - ◆ Indentation, parenthèsage
  - Types, structures, librairies autorisées
  - ... et tout ce que vous estimez important
- Conventions de codage existantes:
  - ◆ Linux kernel: http://www.linuxjournal.com/article/5780
  - GNU standards: http://www.gnu.org/prep/standards
  - GNU indent: http://www.gnu.org/software/indent

### Table des matières

- I. Recommendations générales
- II. Compilation
  - III. Automatisation de la compilation
  - IV. Débogage
  - V. Contrôle de version
  - VI. Conclusion
  - VII. Références

### GCC

#### Exemple classique "Hello World"

```
* Introduction au développement en C
* Programme d'exemple 1
  (c) 2006-2009, Université catholique de Louvain (UCLouvain)
              Ecole Polytechnique de Louvain (EPL)
              Département d'Ingénierie Informatique (INGI)
        _____ */
#include <stdio.h>
/* ----[ main ]----- */
int main()
 printf("Hello World !\n);
 return 0;
```

### GCC

### Compilation

- ◆ GNU C Compiler (gcc)
- Création d'un exécutable à partir d'un fichier source (vérification syntaxique, vérification de types, génération de code exécutable, ...)
- Exemple:

```
[toto@brol src] gcc -o prog1 prog1.c
prog1.c:14:10: missing terminating " character
prog1.c: In function `main':
prog1.c:15: error: syntax error before "return"
[toto@brol src]
```

En cas d'erreur, gcc indique

- 1). le fichier concerné
- 2). la ligne présumée de l'erreur
- 3). la cause de l'erreur

### GCC

### Compilation

- ◆ GNU C Compiler (gcc)
- Création d'un exécutable à partir d'un fichier source (vérification syntaxique, vérification de types, génération de code exécutable, ...)
- Exemple:

```
[toto@brol src] gcc -o prog1 prog1.c
[toto@brol src] ./prog1
Hello World !
[toto@brol src]
```

S'il n'y a ni erreur ni avertissement, gcc n'affiche pas de message.

# GCC: Options

#### -Wall

Affiche tous les avertissements (warnings)

Bonne pratique: toujours compiler avec -Wall et -Werror

#### -Werror

Considère les warnings comme des erreurs (empêche la compilation de se poursuivre en cas de *warning*).

#### -ansi

Vérifie que le code C est conforme à la norme ISO C90.

#### -pedantic

Affiche les avertissements recommandés par la norme ISO C.

#### **-o** < name >

Renomme le fichier de sortie en < name > .

Par défaut, le fichier généré est nommé "a.out".

#### 1). Pré-processing:

gcc -o progl.e -E progl.c

#### 2). Compilation:

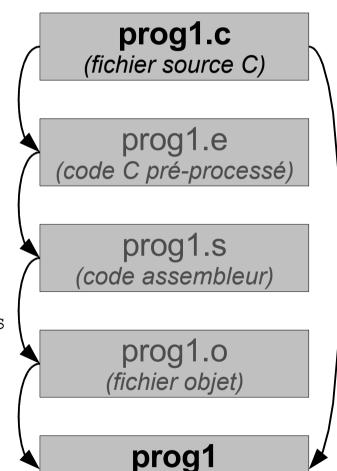
gcc -o prog1.s -S prog1.e

#### 3). Assemblage:

gcc -o prog1.o -c prog1.s

#### 4). Edition de liens (linkage):

gcc -o prog1 prog1.o



(fichier exécutable)

"Compilation":

gcc -o prog1 prog1.c

#### 1). Pré-processing:

gcc -o prog1.e -E prog1.c gcc -o util.e -E util.c

#### 2). Compilation:

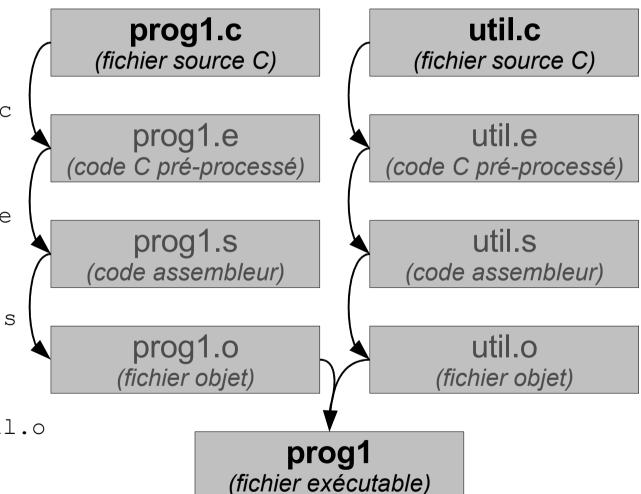
gcc -o prog1.s -S prog1.e gcc -o util.s -S util.e

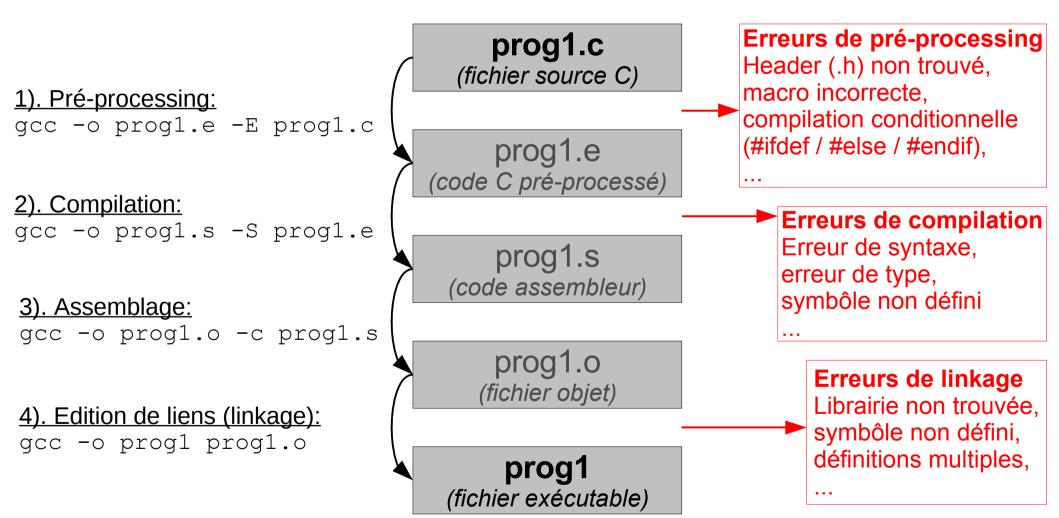
#### 3). Assemblage:

gcc -o prog1.o -c prog1.s
gcc -o util.o -c util.s

#### 4). Edition de liens (linkage):

gcc -o prog1 prog1.o util.o





### 1). Pre-Processing

- GNU C Pre-Processor (cpp)
- Transforme le programme avant la compilation
- Convertit les macros et directives:
  - →#define, #undef
  - ♦ #include
  - →#ifdef, #else, #endif
- Fichiers headers (.h) placés dans l'output
  - → option "-Idir" permet d'indiquer où trouver les headers
  - ◆Chemin standard: /usr/include (dépend du système)
- Supprime les commentaires

### 2). Compilation

◆ Génération de code, allocation des registres, optimisations ⇒ code d'assemblage

```
.file
             "prog1.c"
       .section
                      . rodata
.LC0:
       .string "Hello World !\n"
       .text
.qlobl main
       .type main, @function
main:
       pushl
              %ebp
       movl %esp, %ebp
       subl $8, %esp
       andl $-16, %esp
       movl $0, %eax
       subl %eax, %esp
       movl
              $.LC0, (%esp)
           printf
       call
       movl $0, %eax
```

### 3). Assemblage

- GNU Assembler (as)
- → Traduction du code d'assemblage en langage machine ⇒ fichier "objet" (binaire)

### Fichier "objet"

- Plusieurs types d'objets
  - → Relocatable (créé par l'assembleur)
  - ◆ Executable (créé par le linker)
  - ◆ Shared (créé par le linker: librairies)

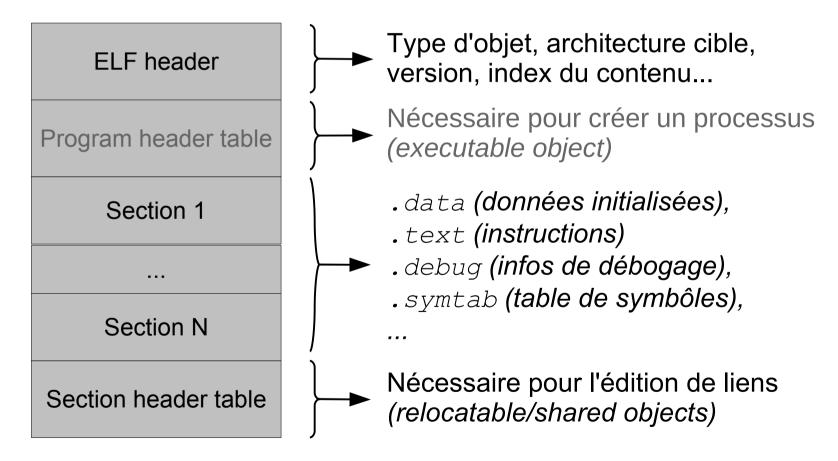
#### Format:

- ◆ Executable and Linking Format (ELF)
- ◆Autres formats: COM, a.out, COFF, ECOFF, EXE, Mach-O...

#### Outils utiles:

- objdump : display content of object file
- nm : display list of symbols in object file

### Structure générale d'un fichier ELF



Contenu de l'objet

-x: show all sections

```
[toto@brol src] objdump -x prog1.o
         file format elf32-i386
progl.o:
prog1.o
architecture: i386, flags 0x00000011:
HAS RELOC, HAS SYMS
start address 0x00000000
                                                     Position, taille et attributs
                                                     de chaque section
Sections:
                  Size
                                                 File off
Tdx Name
                            AMV
                                      TIMA
                                                           Alan
                  00000023
                           0000000
                                      0000000
                                                 00000034
                                                           2**2
  0 .text
                  CONTENTS, ALLOC, LOAD, RELOC, READONLY, CODE
                            00000000
                                      00000000
                                                 00000058 2**2
  1 .data
                  00000000
                  CONTENTS, ALLOC, LOAD, DATA
  2 .bss
                  0000000 0000000 00000000
                                                 00000058
                                                           2**2
                  ATITIOC
                            00000000
                                      00000000
  3 .rodata
                  0000000f
                                                 00000058
                                                           2**0
                  CONTENTS, ALLOC, LOAD, READONLY, DATA
```

Contenu de l'objet

-s: show content of sections

```
[toto@brol src] objdump -s prog1.o
progl.o:
         file format elf32-i386
Contents of section .text:
 0000 5589e583 ec0883e4 f0b80000 000029c4
                                                                 code compilé
 0010 c7042400 000000e8 fcffffff b8000000
                                                                 (instructions)
 0020 00c9c3
Contents of section .rodata:
                                                                  données
 0000 48656c6c 6f20576f 726c6420 210a00
                                            Hello World !..
                                                                 constantes
Contents of section .comment:
 0000 00474343 3a202847 4e552920 332e332e
                                            .GCC: (GNU) 3.3.
 0010 35202844 65626961 6e20313a 332e332e
                                            5 (Debian 1:3.3.
 0020 352d3133 2900
                                            5-13).
```

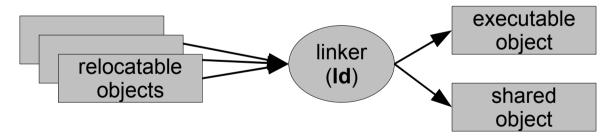
#### Contenu de l'objet

-d: disassemble

```
[toto@brol src] objdump -d prog1.o
             file format elf32-i386
progl.o:
Disassembly of section .text:
00000000 <main>:
        55
   0:
                                push
                                       %ebp
   1: 89 e5
                                       %esp, %ebp
                                mov
   3: 83 ec 08
                                sub
                                       $0x8, %esp
   6: 83 e4 f0
                                       $0xfffffff0, %esp
                                and
   9:
      b8 00 00 00 00
                                       $0x0, %eax
                                mov
   e:
      29 c4
                                sub
                                       %eax, %esp
 10: c7 04 24 00 00 00 00
                                movl
                                       $0x0, (%esp)
 17: e8 fc ff ff
                                       18 < main + 0 \times 18 >
                                call
  1c: b8 00 00 00 00
                                       $0x0, %eax
                                mov
  21: c9
                                leave
  22: c3
                                ret
```

Le contenu doit être similaire à prog1.s

- 4). Edition de liens
  - GNU linker (Id)
  - Production d'un <u>exécutable</u> ou d'une <u>librairie</u> dynamique à partir d'un ensemble d'objets (relocatable)



- Lier avec les symbôles non-définis (statiques):
  - utilise les sections .reloc et .symtab pour modifier le contenu de la section .text
  - "relocations"

### Création d'une librairie

### Librairie statique (static library)

Archive de fichiers objets (.o)

```
[toto@brol src] gcc -Wall -c hello.c
[toto@brol src] ar cru libhello.a hello.o

[toto@brol src] ranlib libhello.a

[toto@brol src]

[toto@brol src]

[toto@brol src]

[toto@brol src]

[toto@brol src]

[toto@brol src]
```

#### Exemple

Crée une archive avec

# Création d'une librairie

## Librairie dynamique (shared library)

- Archive de fichier objets "position-independent" (PIC)
- Tous les systèmes ne supportent pas les librairies dynamiques.

```
[toto@brol src] gcc -fPIC -Wall hello.c
[toto@brol src] gcc -shared -Wl,-soname,libhello.so.1 \
    -o libhello.so.1.0.1 hello.o -lc
```

Sous d'autres systèmes, la méthode diffère. Voir http://www.fortran-2000.com/ArnaudRecipes/sharedlib.htm

- Installation nécessaire
  - ◆ Modifier /etc/ld.so.conf et lancer /sbin/ldconfig
  - ◆Ou utiliser ld library path En général

En général lorsqu'on n'a pas les droits nécessaires pour modifier ld.so.conf

# Utiliser une librairie

#### API de la librairie

- ◆ <u>API</u>: types et fonctions fournis par la librairie généralement décrit dans des fichiers . ħ
- Type: statique ou dynamique ?

### Lier à une librairie statique

Exemple

```
[toto@brol src] gcc -o main.o -c main.c
[toto@brol src] gcc -o hello main.o libhello.a
```

 Si la librairie est situé ailleurs, par exemple dans /usr/local/lib

```
[toto@brol src] gcc -L/usr/local/lib -o hello main.o libhello.a
```

## Utiliser une librairie

- Lier à une librairie dynamique
  - Exemple

```
[toto@brol src] gcc -o main.o -c main.c
[toto@brol src] gcc -o hello -lhello main.o
```

## Table des matières

- I. Recommendations générales
- II. Compilation
- | III. Automatisation de la compilation
  - IV. Débogage
  - V. Contrôle de version
  - VI. Conclusion
  - VII. Références

# Compilation automatisée

## Compiler un gros projet manuellement

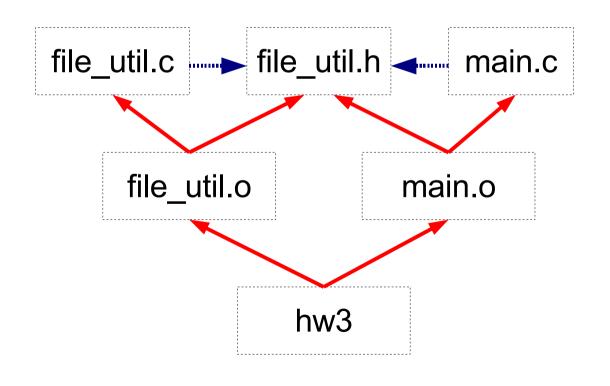
- peut causer des erreurs et être fastidieux
- Solution 1: créer un script shell

```
#!/bin/bash
gcc -c list_util.c -o list_util.o
gcc -c main.c -o main.o
gcc -o hw3 list_util.o main.o
```

- Solution 2: utiliser l'utilitaire make
  - Make détermine automatiquement quelles parties d'un programme doivent être recompilées
  - ◆ Se base sur un fichier de configuration: le Makefile
  - ◆ Définit des règles de compilation et des dépendances

## Make

### Graphe de dépendances



#### Exemple:

Si file\_util.c est modifié, alors file\_util.o et hw3 doivent être reconstruits.

#### Exemple:

Si **file\_util.h** est modifié, alors tous les fichiers objets ainsi que **hw3** doivent être reconstruits.



### Le fichier Makefile sert à définir

- les dépendances entre fichiers:
  - → 1 cible dépend de 0-n dépendances
  - → par exemple: main.o dépend de main.c et file\_util.h
- Les règles de transformation
  - ◆ Comment générer une cible à partir de ses dépendances
  - → par exemple: comment passer de main.c à main.o
  - → générique: permet de fonctionner avec gcc, javac, latex, ...

### Syntaxe

```
cible : [dépendance1 [, dépendance2...]]
<tab>commande1
<tab>commande2
...
```

#### Exemple:

```
hw3: main.o list_util.o
gcc -o hw3 main.o list_util.o

main.o: main.c list_util.h
gcc -c main.c -o main.o

list_util.o: list_util.c list_util.h
gcc -c list_util.c -o list_util.o
```

Ce Makefile représente le graphe de dépendances de hw3.

## Make et Makefile

#### Invocation de make

La cible par défaut est la première qui apparaît dans le Makefile.

```
[toto@brol src] make
gcc -c main.c -o main.o
gcc -c list_util.c -o list_util.o
gcc -o hw3 main.o list_util.o
```

Possibilité d'appeler une autre règle

```
[toto@brol src] make main.o
```

- Utiliser un autre fichier Makefile
  - → Par défaut: fichier Makefile du répertoire courant

```
[toto@brol src] make -f Makefile.solaris
```

## Make

## Recompilation partielle par make

- ne recompile que le nécessaire
- se base sur la date de modification des fichiers

```
[toto@brol srcl make
                                            Première compilation:
gcc -c main.c -o main.o
                                            tout est compilé.
gcc -c list_util.c -o list_util.o
gcc -o hw3 main.o list_util.o
                                          Suppression de hw3:
[toto@brol srcl rm hw3
                                         il faut seulement relinker.
[toto@brol srcl make
gcc -o hw3 main.o list_util.o
                                                Edition de
[toto@brol src] vi list_util.c
                                                list util.c:
[toto@brol srcl make
                                                il faut recompiler
gcc -c list_util.c -o list_util.o
                                                list util.o
gcc -o hw3 main.o list_util.o
                                               et relinker.
```

## Clean (phony target)

- → Règle conventionnelle pour "nettoyer" le projet
  - Supprimer tous les fichiers objets
  - Supprimer d'autres fichiers générés durant la compilation

#### Comment ?

.PHONY: clean ◀
clean:
rm -f main.o list\_util.o hw3

**.PHONY** spécifie une liste de cibles qui ne sont pas des fichiers. C'est nécessaire en cas de présence d'un fichier nommé "clean".

#### Utilisation

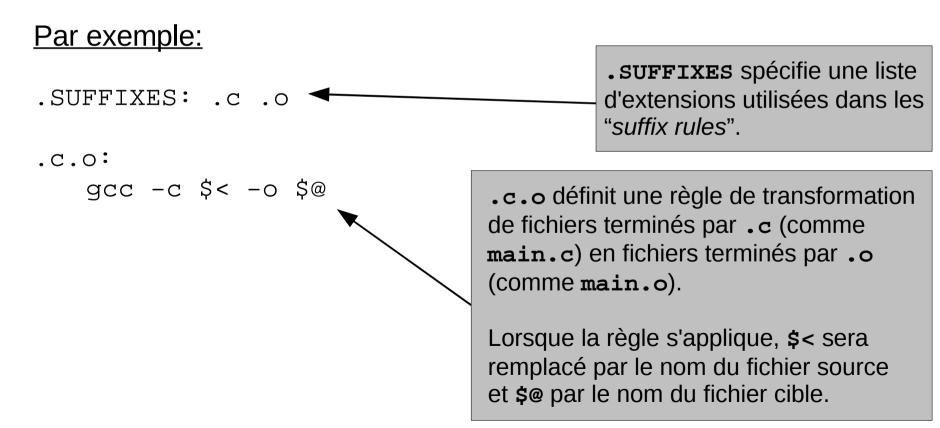
[toto@brol src] make clean

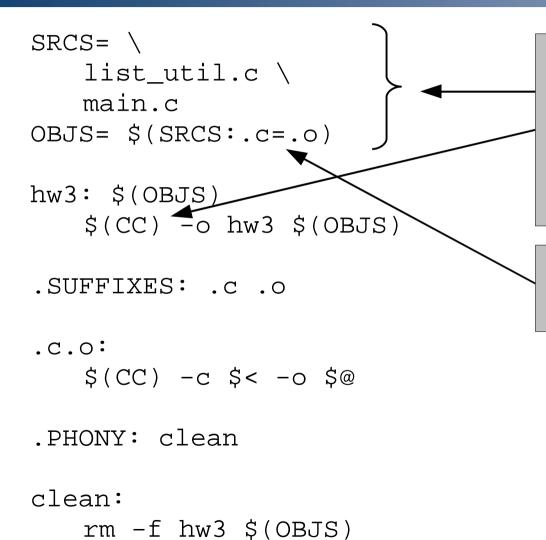
#### **ATTENTION!**

Il faut faire attention à ne pas mettre de fichier "source" (.c, .h) dans la règle clean

## Règles génériques (suffix rules)

 Type de règle spécial s'appliquant aux fichiers qui correspondent aux extensions spécifiées.





SRCS et OBJS sont des variables définies par l'utilisateur. CC est une variable standard de make qui contient le nom du compilateur C par défaut. En général: cc ou gcc.

Le contenu de **OBJ**s est obtenu en transformant le contenu de **SRC**s.

#### **Conseil:**

lisez les **Makefiles** écrits par des développeurs expérimentés. C'est instructif!!

# Utilisation récursive

### Grands projets

- Organisés en une hiérarchie de répertoires
- Pratique courante:
  - ◆ 1 Makefile par sous-répertoire
  - Chaque Makefile appelle les Makefiles des ss-répertoires

```
SUBDIRS=src1 src2
all-recursive:
  for subdir in $(SUBDIRS); do \
    (cd $$subdir && $(MAKE)) \
    done;
```

# Bon à savoir...

#### automake / autoconf

- permettent de générer des fichiers Makefile
- processus de compilation "standard"

```
[toto@brol src] ./configure
[toto@brol src] make
[toto@brol src] make install
```

- processus de compilation + portable
- relativement complexe à apprendre
- bonne documentation disponible chez RedHat http://sourceware.org/autobook/autobook/autobook\_toc.html

### makedepend

 Génère automatiquement des règles de dépendance (et les ajoute au fichier Makefile)

## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- **IV. Débogage** 
  - Principes
  - GDB
  - VALGRIND
  - V. Contrôle de version
  - VI. Conclusion
  - VII. Références

## Erreurs fréquentes en C

- Accès en dehors des bornes d'un tableau
- Utilisation d'une variable non initialisée
- Utilisation d'un pointeur non initialisé
- Pas de libération des ressources

## Symptômes

- Message "Segmentation fault"
- Message "Bus error"
- Fonctionnement non déterministe
- Plus de mémoire disponible

### Exemple:

 programme qui prend un nom de fichier en ligne de commande et imprime son contenu (ligne par ligne)

```
#include <stdio.h>
01
02
    #include <stdlib.h>
03
04
    #define BUFFER SIZE 100
0.5
06
    main(int argc, char * argv[])
07
08
    FILE * in stream;
09
    char * buffer= NULL;
10
11 /* Vérifie la ligne de commande */
12
    if (argc != 2) {
13
        fprintf(stderr, "Error: missing argument\n");
14
        exit (EXIT FAILURE);
15
16
```

◆ Exemple: suite...

```
17
     /* Ouvre le fichier en lecture... */
18
      if ((in stream= fopen(argv[1], "r")) == NULL) {
        fprintf(stderr, "Error: could not open \"%s\" ",
19
20
                argv[1]);
21
        exit (EXIT FAILURE);
22
      } else
        printf("Fichier \"%s\" ouvert :-)\n", argv[1]);
23
24
25
     /* ... et lit les utilisateurs */
2.6
     while (!feof(in stream)) {
27
        if (fgets(buffer, BUFFER SIZE, in stream) == NULL)
2.8
         break;
29
       printf("-> %s", buffer);
30
31
32
    fclose(in stream);
33
      exit (EXIT SUCCESS);
34 }
```

### Compilation et exécution

```
[toto@brol src] gcc -o prog3 prog3.c
[toto@brol src] cat list.txt

Eddard Stark
Jon Snow
Robert Baratheon
[toto@brol src] ./prog3 list.txt
Fichier "list.txt" ouvert :-)
Segmentation fault
```

Une erreur de segmentation s'est produite. Il y a probablement une erreur ...

... Mais où?

## Débogage

- Re-lecture des sources
  - → ne suffit pas toujours à localiser l'erreur
  - programmes longs / complexes
- L'utilisation d'un débogueur (debugger) permet de
  - ◆suivre l'exécution pas-à-pas
  - définir des points d'arrêt (breakpoints)
  - définir des watchpoints
  - tracer la pile d'exécution (stack trace)
- ◆ Sous UNIX (Linux): gdb

### Breakpoints

- Arrêt du programme à une adresse de programme donnée
  - ◆Lors de l'instruction fetching

### Hardware

- → registres spéciaux du CPU (debug registers sur i386)
- ◆limite sur le nombre de breakpoints (4 sur i386)
- → génération d'une exception (trap)

### Software

- ◆changements dynamiques du code (int3 sur i386)
- ◆ fonctionne même si pas de support du CPU
- utile si beaucoup de breakpoints
- plus lent/complexe qu'en hardware

## Interruption de débogage (i386)

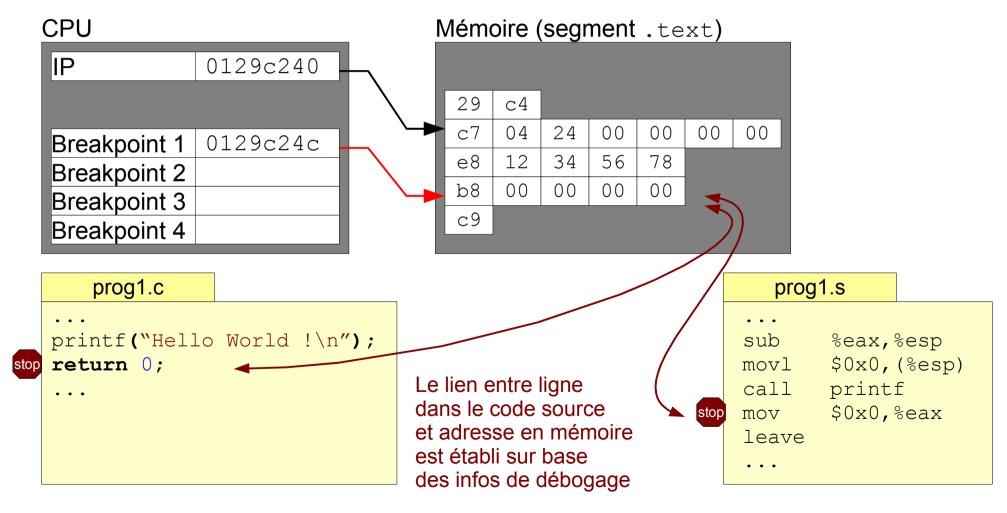
```
.text
.global main

main:

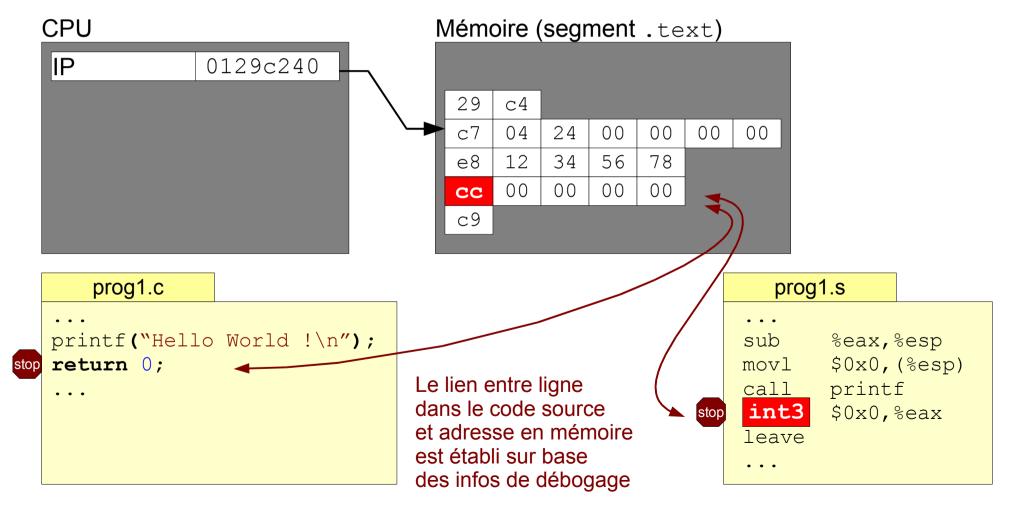
int3
.end
```

```
[toto@brol src] gcc -o int3 int3.s
[toto@brol src] ./int3
Trace/breakpoint trap
```

#### Points d'arrêt matériels



### Points d'arrêt logiciels



### Watchpoints

 Arrêt du programme lors de l'accès à une zone de mémoire

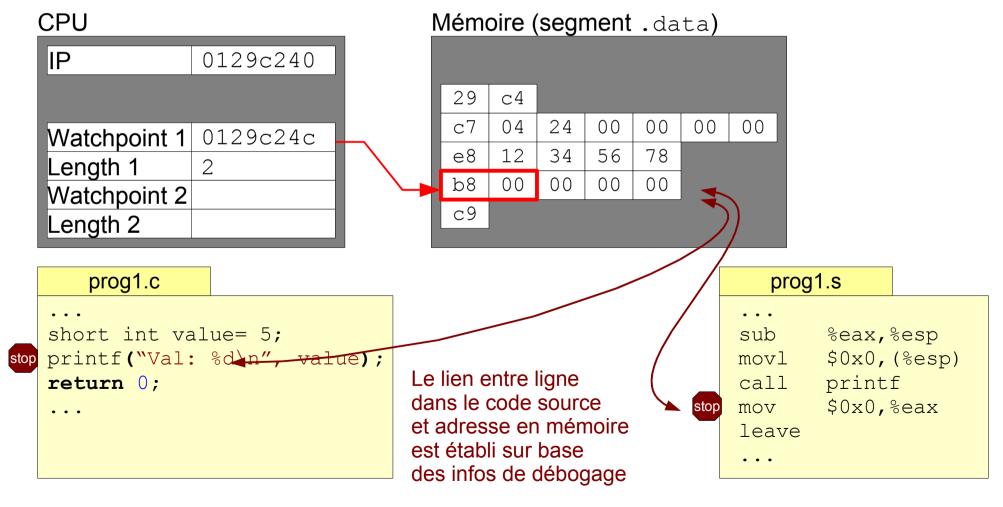
### Hardware

- → registres spéciaux (adresse + longueur)
- ◆limite sur nombre de watchpoints (i386: 4)
- ◆limite sur taille des watchpoints (i386: 1/2/4 octets)
- ◆ en lecture/écriture

### Software

- ◆ single-stepping + comparaison avant/après de la zone surveillée
- plus lent/complexe qu'en hardware
- en écriture uniquement !

### Watchpoints matériels



## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
  - Principes
  - GDB
    - VALGRIND
- V. Contrôle de version
- VI. Conclusion
- VII. Références

# Débogage: GDB

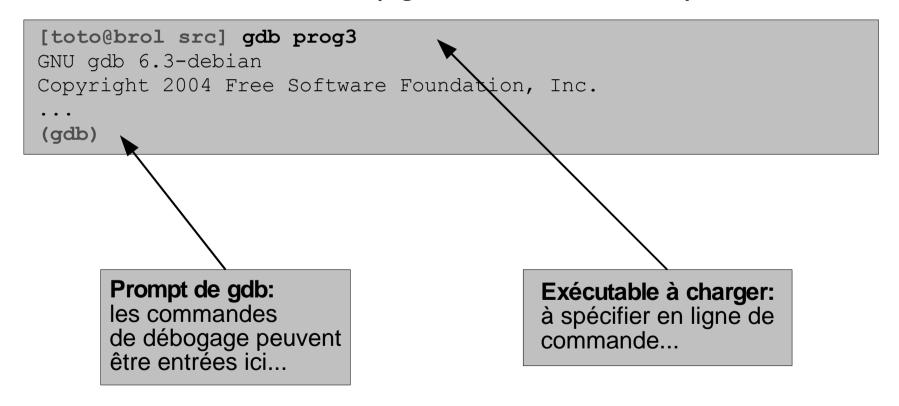
## Préparation du programme

 Compilation du programme avec les symboles de débogage: option "-g"

```
[toto@brol src] gcc -g -o prog3 prog3.c
```

- Résultat:
  - ◆L'exécutable contient des sections .debug xxx
  - Noms des symbôles (fonctions, variables)
  - ◆Lien entre le code source (instruction C) et le code binaire (adresse dans le segment .text)

- Lancement du débogueur
  - Interface utilisateur (ligne de commande)



#### Utilisation de base

- Quitter le débogueur: <ctrl-D> ou "quit"
- Lancer l'éxécution du programme

```
[toto@brol src] gdb prog3
GNU gdb 6.3-debian
Copyright 2004 Free Software Foundation, Inc.
...
(gdb) run list.txt
Starting program:.../prog3 list.txt
Fichier "list.txt" ouvert :-)
-> Eddard Stark

Program received signal SIGSEGV, Segmentation fault.
0x4007a295 in _IO_getline_info () from /lib/tls/libc.so.6
(gdb)
```

Une erreur s'est produite. Le programme a causé une erreur de segmentation (SEGV).

### Examiner les sources

list <file>:<line>; list; ...

```
[toto@brol src] gdb prog3
GNU qdb 6.3-debian
Copyright 2004 Free Software Foundation, Inc.
(qdb) list prog3.c:1
01 #include <stdio.h>
02
  #include <stdlib.h>
03
04
     #define BUFFER SIZE 100
0.5
                                                     10 lignes (max)
06
      main(int argc, char * argv[])
07
08
     FILE * pIn;
09
        char * pcBuffer= NULL;
10
(gdb) list
11 /* Vérifie la ligne de commande */
                                                     10 lignes de plus
12 if (argc != 2) {
                                                     (max)
```

### Définir un breakpoint

break <line>; break <fct-name>; break \*<address>

```
[toto@brol src] gdb prog3
GNU gdb 6.3-debian
Copyright 2004 Free Software Foundation, Inc.
...
(gdb) break 26
Breakpoint 1 at 0x804855f: file prog3.c, line 26.
(gdb) run list.txt
Starting program:.../prog3 list.txt
Fichier "list.txt" ouvert :-)

Breakpoint 1, main (argc=2, argv=0xbffff5f4) at prog3.c:26
26     while (!feof(pIn)) {
     (gdb)
```

### Breakpoints

- Définir un breakpoint
  - ◆break <line>
- Définir une condition
  - → condition <bp-#> <expression> (sur bp existant)
  - →break <line> if <expression> (nouveau bp)
- Désactiver
  - →disable <bp-#>
- Supprimer
  - ◆clear <line>
  - →delete <bp-#>

(tous les bps de la ligne) (par numéro de bp)

### Continuer l'exécution...

```
[toto@brol src] qdb proq3
GNU qdb 6.3-debian
Copyright 2004 Free Software Foundation, Inc.
Breakpoint 1, main (argc=2, argv=0xbffff5f4) at prog3.c:26
2.6
          while (!feof(pIn)) {
(qdb) continue
Continuing.
-> Eddard Stark
Breakpoint 1, main (argc=2, argv=0xbffff5f4) at prog3.c:26
2.6
          while (!feof(pIn)) {
(qdb) continue
Continuing.
Program received signal SIGSEGV, Segmentation fault.
0x4007a295 in IO getline info () from /lib/tls/libc.so.6
(qdb)
```

### Exécution pas-à-pas

L'erreur s'est produite lors de l'exécution de la ligne 27.

Raison: pcBuffer non alloué.

# Débogage: GDB

#### Exécution pas à pas

- Avance à la ligne suivante (ou N lignes)
  - ◆step [N]
  - → next [N] (comme step, n'entre pas dans les fonctions)
- Avance à l'instruction suivante (ou N instructions)
  - ♦stepi [N]
  - → nexti [N] (comme stepi, n'entre pas dans les fonctions, i.e. call)
- Affiche la valeur de variables / expressions
  - →print <expression>
- Termine l'exécution d'une fonction (même stack frame)
  - finish

#### GDB

#### Autres commandes

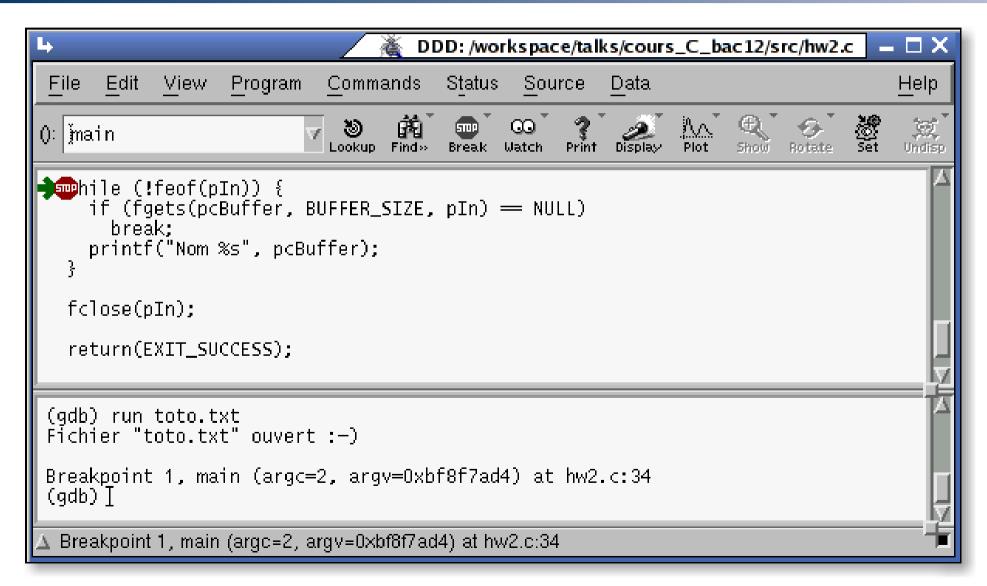
- Watchpoints
  - →watch <expression>

- Back trace
  - ◆backtrace [full]

```
(gdb) break 5
...
(gdb) run
...
(gdb) backtrace
#0 fct_b () at prog7.c:5
#1 0x080483a3 in fct_a () at prog7.c:10
#2 0x080483ba in main () at prog7.c:15
```

```
#include <stdio.h>
void fct b(char * msg)
 printf(msg);
void fct a(char * msq)
  fct b(msg);
int main()
  fct a("Hello World !\n");
  return 0;
```

# DDD: un GUI pour GDB



## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
  - Principes
  - ◆ GDB
  - **VALGRIND**
- V. Contrôle de version
- VI. Conclusion
- VII. Références

# Débogage: memory leaks

#### Traçage des accès en mémoire

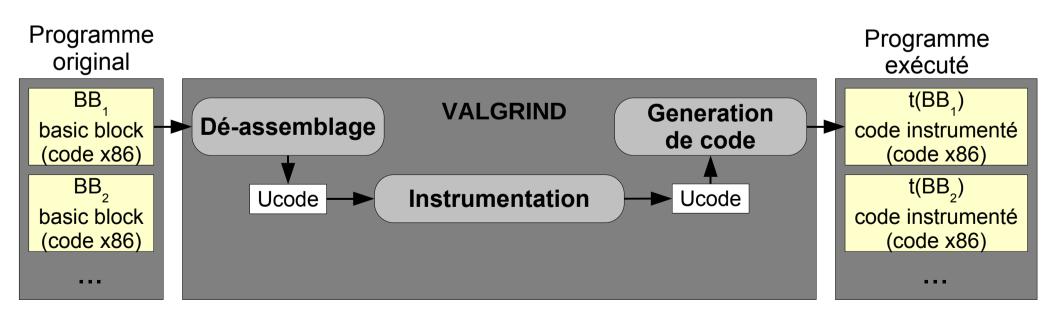
- Permet de détecter
  - l'accès à des zones non initialisées
  - ◆ l'accès hors des bornes d'un tableau,
  - ◆ les zones de mémoire non libérées
  - Les tentatives de libération de zones non allouées
- Traçage non exhaustif
  - → uniquement le long du chemin d'exécution

- Plate-forme d'analyse de programmes
  - Analyse par simulation DBI (Dynamic Binary Instrumentation)
  - → Memcheck: détection d'erreurs de mémoire
  - Cachegrind: "cache misses" analyser
  - Massif: space profiler
  - **\***
- Utilisé par de grands projets
  - OpenOffice, MySQL, Gimp, Python, ...

# Compilateur JIT / Instrumentation

#### Principes

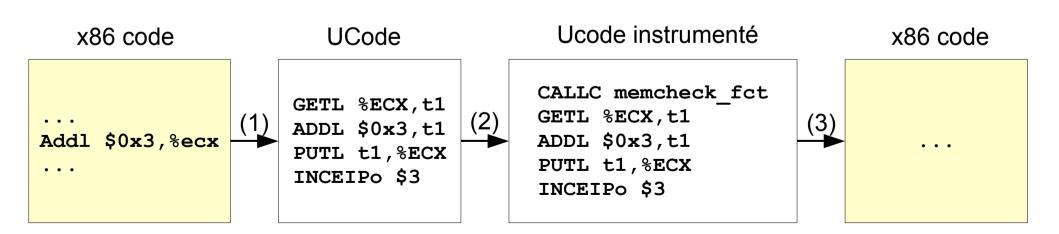
- Code x86: centaines d'instructions différentes
- Traduction en UCode, de type RISC
  - ⇒ Moins d'instructions
- Instrumentation de l'UCode



# Compilateur JIT / Instrumentation

#### Exemple

- 1) Instruction x86 traduite en instructions UCode
- 2) Instructions UCode ajoutées
  Par exemple, appel de la fonction memcheck\_fct
- 3) Code x86 généré

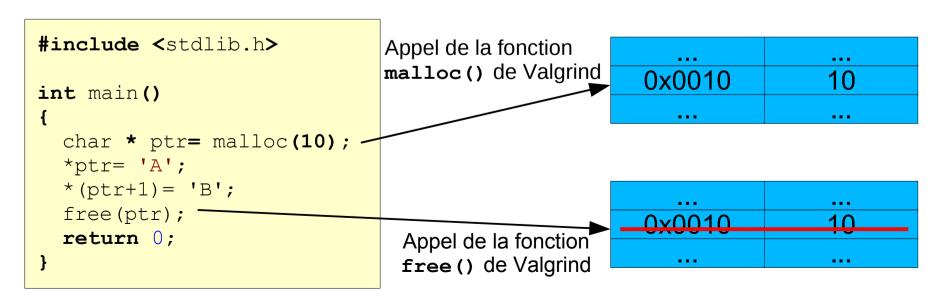


#### MemCheck

- Détecte
  - Les pertes de mémoire (memory leaks)
  - Accès incorrects à la mémoire (heap)
  - Zones de mémoire utilisées sans être initialisées
- Grâce à l'instrumentation du code
  - Espionne les opérations UCode de lecture et d'écriture
  - Maintient des metadatas ("shadow values") pour chaque registre et zone de mémoire utilisée
  - Chaque zone de mémoire allouée / libérée est pistée

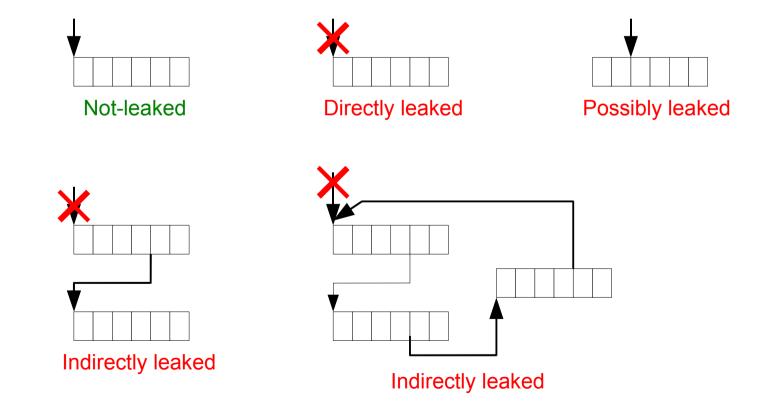
#### MemCheck

- → Remplace les fonctions malloc, realloc, free, ...
  - garde une trace des allocations (emplacement + taille)
  - → place des "red-zones" autour des zones de mémoire



#### MemCheck: Memory leaks

Classification



(voir présentation de Julian Seward, FOSDEM'06)

# Memory Leaks

#### Exemple

```
#include <stdlib.h>
int main()
{
   char * ptrChars= malloc(6 * sizeof(char));
   ptrChars[0]= 'H';
   //free(ptrChars);
   return 0;
}
```

# Memory Leaks

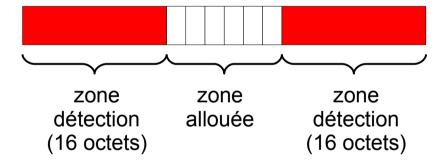
#### Exemple

```
[toto@brol src] valgrind -leak-check=yes prog6
==13082== ERROR SUMMARY: 0 errors from 0 contexts (suppressed: 14 from 1)
==13082== malloc/free: in use at exit: 6 bytes in 1 blocks.
==13082== malloc/free: 1 allocs, 0 frees, 6 bytes allocated.
==13082== For counts of detected errors, rerun with: -v
==13082== searching for pointers to 1 not-freed blocks.
==13082== checked 76,196 bytes.
==13082==
==13082== 6 bytes in 1 blocks are definitely lost in loss record 1 of 1
==13082==
            at 0x401C867: malloc (vg replace malloc.c:149)
==13082==
          by 0x804839F: main (prog6.c:5)
==13082==
==13082== LEAK SUMMARY:
==13082==
            definitely lost: 6 bytes in 1 blocks.
==13082==
              possibly lost: 0 bytes in 0 blocks.
==13082==
            still reachable: 0 bytes in 0 blocks.
==13082==
                  suppressed: 0 bytes in 0 blocks.
```

# Invalid Heap Access

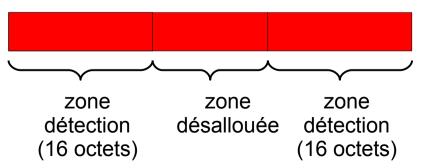
- MemCheck: Accès invalides au tas (heap)
  - ◆ En écriture et lecture
  - Autres outils (ElectricFence, ...) ne détectent que les écritures incorrectes







## Accès à des zones libérées:



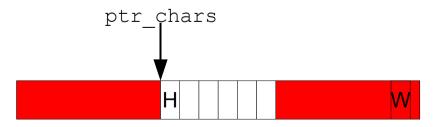
Attention: une fois désallouée, une zone peut être ré-allouée plus tard. Conséquence, des accès incorrects peuvent ne pas être détectés!

(voir présentation de Julian Seward, FOSDEM'06)

# Invalid Heap Access

#### Exemple

```
#include <stdlib.h>
int main()
{
   char * ptr_chars= malloc(6 * sizeof(char));
   ptr_chars[0]= 'H';
   ptr_chars[12]= 'W';
   free(ptr_chars);
   ptr_chars[1]= 'e';
   return 0;
}
```



# Invalid Heap Access

#### Exemple

```
[toto@brol src] qcc -q -o proq4 proq4.c
[toto@brol src] ./proq4
[toto@brol src] valgrind -leak-check=yes prog4
==7643== Memcheck, a memory error detector.
==7643== Copyright (C) 2002-2007, and GNU GPL'd, by Julian Seward et al.
==7643== Invalid write of size 1
==7643== at 0x80483EF: main (prog4.c:7)
==7643== Address 0x415602E is 0 bytes after a block of size 6 alloc'd
           at 0x401C867: malloc (vg replace malloc.c:149)
==7643==
          by 0x80483DF: main (proq4.c:5)
==7643==
==7643==
==7643== Invalid write of size 1
==7643==
           at 0x8048401: main (proq4.c:9)
==7643== Address 0x4156029 is 1 bytes inside a block of size 6 free'd
==7643==
           at 0x401D4FE: free (vg replace malloc.c:233)
==7643== by 0x80483FC: main (prog4.c:8)
```

#### Invalid Stack Access

#### Attention!

Les erreurs d'accès à la pile (stack) ne sont pas détectées!

```
#include <stdlib.h>

int main()
{
    char array_chars[6];
    array_chars[0]= 'H';
    array_chars[12]= 'W';
    return 0;
}

| Ci, le tableau array_chars[]
    est alloué sur la pile.
| Valgrind ne détecte pas l'écriture invalide de la valeur 'W'.
| Valgrind ne détecte pas l'écriture invalide de la valeur 'W'.
| Ci, le tableau array_chars[]
| Est alloué sur la pile.
| Valgrind ne détecte pas l'écriture invalide de la valeur 'W'.
```

#### MemCheck: Zones non initialisées

- Chaque bit dans les zones de mémoire et registres est associé à un "shadow bit" qui indique si la zone a été initialisée (affectée)
- Propage les shadow bits avec les opérations en Ucode + avertit lors de leur utilisation

```
if (...GARBAGE...)
branchement dépend de valeur non-initialisée
```

```
* ( . . . GARBAGE . . . )

adresse mémoire dépend de valeur non-initialisée (dé-
référencement)
```

```
syscall (...GARBAGE...)

paramètre d'appel système dépend de valeur non-initialisée
```

## Résumé

#### • Quel outil utiliser ?

- Segmentation fault: gdb, valgrind
- Memory leaks: valgrind

#### Attention !!!

Les outils tels que gdb et valgrind ne tracent que le chemin d'exécution. Ils ne peuvent détecter des erreurs dans des parties de programme non exécutées!

## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
- V. Contrôle de version
  - VI. Conclusion
  - VII. Références

## Contrôle de version

- Un contrôle de version, mais...
  - À quoi ça sert ?
    - Conserver un historique des modifications
      - ◆ Ca "marchait", vous faites une modification et "plouf!"
        - possibilité de retourner chercher l'ancienne version
      - Gestion des versions successives du projet
        - possibilité de récupérer une ancienne version, d'y apporter des corrections et de la fournir aux utilisateurs.
    - Travail collaboratif
      - Permet de travailler à plusieurs sur un projet

## Contrôle de version

#### Un contrôle de version, mais...

- Comment ça marche ?
  - → "Repository central"
    - contient le projet et son historique
    - ◆ typiquement sur un serveur dédié
  - **→** "Working copy"
    - ◆ Copie locale du repository
    - ◆ Un développeur travaille uniquement sur sa copie. Il ne gêne donc pas les autres développeurs.
    - lorsqu'il a fini ses modifications, il les transfère (commit) dans le repository.
    - plusieurs working copies sont possibles sur plusieurs workstations en même temps
- Exemples: CVS ou SVN (subversion)

#### SVN en INGI

Inscription via https://scm.info.ucl.ac.be/cgi-bin/inscription.sh

Gestion des repositories pour TP de groupes	Server info: scm  Tue Feb 10 2009 09h33:16  HTTPD: Apache/2.2.3 (CentOS)  OS: Linux 2.6.18-92.1.22.el5xen kerne Arch: i686	
Veuillez compléter les paramètres du nouveau proje		
Informations générales		
Nom du groupe ou du projet :		
☐ Intégration dans un repository monitoré de travaux de	e groupes.	_
Cours concerné : faites votres choix ‡		
Adresse(s) email du(des) membre(s) du projet		
@student.uclouvain.be	pa	us recevrez une confirmation r e-mail accompagnée de
Envoyer le formulaire (Caractères autorisés : alphanumériques, "-", ".", "_"; les champs doiven		re mot de passe.

Utilisation décrite sur le Wiki étudiants http://wiki.student.info.ucl.ac.be/index.php/Subversion

#### Création d'une Working Copy

Commande svn checkout <url>

Nom du projet.

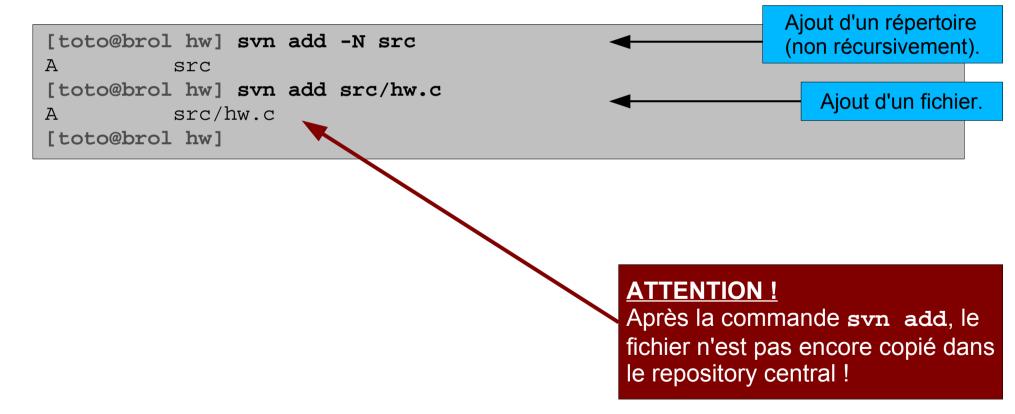
```
[toto@brol ~] svn --username toto checkout
  https://scm.info.ucl.ac.be/studentsvn/hw
Password for 'toto': *********
Checked out revision 0.
[toto@brol ~] cd hw
[toto@brol hw] ls -a
. . . .svn
[toto@brol hw]
```

URL complète du projet à récuperer.

Un répertoire **.svn** est créé dans chaque répertoire de la *working copy*. Ce répertoire maintient des informations sur l'état de la *working copy*.

Il ne faut pas le supprimer ni en modifier le contenu !!!

- Ajout d'un fichier dans le projet
  - Commande svn add <path\_to\_file>



- Envoyer un fichier vers le repository central
  - Commande svn commit <path\_to\_file>

```
[toto@brol hw] svn commit -N src

Adding src

Committed revision 1. 

[toto@brol hw] svn commit src/hw.c

Adding src/hw.c

Committed revision 1.
```

Note: il est possible d'ajouter un commentaire lors du commit du fichier avec l'option  $-\mathbf{m}$  < msg >

```
[toto@brol hw] svn commit -m "Correction bug" src/hw.c
```

Bonne pratique: utiliser -m permet de se souvenir plus tard de la modification qui avait été apportée au fichier.

#### Mettre à jour une copie locale

- Récupère les modifications apportées par les autres
- Commande svn update <path\_to\_file>

```
[toto@brol hw] svn update src/hw.c

A src/hw.c

Updated to revision 2.

[toto@brol hw]
```

#### Notes:

 Les commandes add, commit et update peuvent être utilisées avec plusieurs fichiers en même temps

```
[toto@brol hw] svn add src/main.c src/list_util.c src/list_util.h
```

ou bien

```
[toto@brol hwl svn add src/*.c
```

- → Il est recommandé de NE PAS placer dans le projet les fichiers compilés (objets et exécutables)
  - ◆ Les autres développeurs ne travaillent peut-être pas sur la même plateforme que vous.

#### Résolution de conflits

→ modifications simultanées ⇒ conflit ?

```
[bob@truc src] vi hw.c
[bob@truc src] svn commit hw.c
Sending hw.c
Committed to revision 3.

Bob a modifié hw.c et l'a envoyé dans le repository
```

```
[toto@brol src] vi hw.c
[toto@brol src] svn commit hw.c
svn: Commit failed (details follow)
svn: Out of date: 'hw.c'

c src/hw.c
synchw.c
```

CONFLIT

#### Résolution de conflits

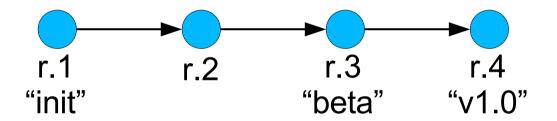
Il faut décider manuellement quelle version garder !

```
#include <stdio.h>
                main()
                                                               Modification de Toto
                                                               (locale).
                <<<<< .mine
Les symboles
>>>>, <<<<
                   /* This is a comment */
et ==== sont
ajoutés par
                  /* Ceci est un commentaire
svn pour
                délimiter le
                                                                Modification de Bob
                  printf("Hello World!\n");
conflit.
                                                                (repository).
                  return(0);
```

La plupart du temps, svn parvient à résoudre les conflits automatiquement.

#### Tagging

 Donner un nom à une révision particulière de l'ensemble d'un projet

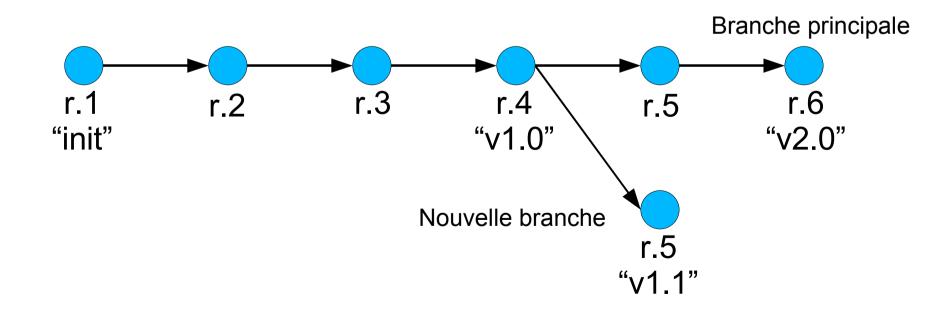


#### Tagging

```
[toto@brol src] svn copy
  https://scm.info.ucl.ac.be/studentsvn/hw/trunk
  https://scm.info.ucl.ac.be/studentsvn/hw/tags/release-1.0
```

#### Branching

- Modifications isolées pour une version particulière
- Maintenance d'une version dans une branche
- Développement expérimental

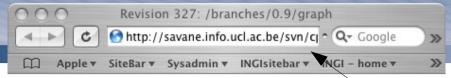


#### Branching

```
[toto@brol src] svn copy
  https://scm.info.ucl.ac.be/studentsvn/hw/tags/release-1.0
  https://scm.info.ucl.ac.be/studentsvn/hw/branches/release-1
  .0
```

```
[toto@brol src] cd /tmp
[toto@brol tmp] svn checkout
  https://scm.info.ucl.ac.be/studentsvn/hw/branches/release-1
    .0 hw
[toto@brol tmp] cd hw/src
[toto@brol src] vi hw.c
[toto@brol src] svn commit hw.c
```

# SVN par le web



#### Revision 327: /branches/0.9/graph

- <u>.</u>
- INSTALL
- LICENSE
- Makefile
- Makefile.in.in
- TODO
- binarysimple.hh
- binarysimple.icc
- branch/
- configure
- · configure.ac
- doxygen.conf
- doxygen.hh
- examples/
- graph.hh
- graphutils.h
- graphutils.icc
- misc/
- path/
- path.hh
- path.icc
- shortdesc.ac
- stlutility.icc
- var.icc
- view/
- view.icc

**URL** complète du projet à visualiser. C'est la même URL que celle utilisée avec la commande svn.

Powered by Subversion version 1.1.4 (r13838).

## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
- V. Contrôle de version
- VI. Conclusion
  - VII. Références

#### Recommendations



# 1). Lisez le code de programmeurs expérimentés afin de comprendre comment il est organisé!

Il existe un grand nombre de projets Open-Source. Préférez la lecture de projets de <u>grande taille</u> et qui <u>sont réellement utilisés</u> (par exemple le kernel Linux)

# 2). Programmez, essayez, faites des erreurs!

C'est ainsi que vous ferez évoluer votre technique de programmation.

## Table des matières

- I. Recommendations générales
- II. Compilation
- III. Automatisation de la compilation
- IV. Débogage
- V. Contrôle de version
- VI. Conclusion
- VII. Références

# Références (1)

- \* "Advanced Programming in the UNIX(R) Environment (2<sup>nd</sup> edition)", de W. Richard Stevens et Stephen A. Rago. Addison-Wesley, 2005.
- "The Practice of Programming", de Brian W. Kernighan et Rob Pike. Addison-Wesley, 1999.
- "Write Portable Code: An Introduction to Developing Software for Multiple Platforms", de Brian Hook. NO STARCH PRESS, 2005.
- "Managing Projects with GNU Make, Third Edition", de Robert Mecklenburg. O'Reilly, 2004.
- "Pragmatic Version Control Using Subversion", de Mike Mason. O'Reilly, 2005.
- "Version Control with Subversion", de Ben Collins-Sussman, Brian W. Fitzpatrick et C. Michael Pilato. O'Reilly. Accessible gratuitement http://svnbook.red-bean.com/
- \* "The C Programming Language (2<sup>nd</sup> edition)", par Brian W. Kernighan et Dennis M. Ritchie. Prentice Hall, Inc., 1988.
- "Intel 80386 Reference Programmer's Manual", Intel Corporation (unofficial), 1986. Accessible à partir de http://pdos.csail.mit.edu/6.828/2005/readings/i386/toc.htm

# Références (2)

- \* "Executable and Linkable Format (ELF) Specification (version 1.2)", Tool Interface Standard (TIS) Committee, 1995.
- \* "Debugging with GDB: The gnu Source-Level Debugger (9<sup>th</sup> Edition)", Richard Stallman, Roland Pesch, Stan Shebs, et al., 2004.
- \* "GDB Internals: A guide to the internals of the GNU debugger (2<sup>nd</sup> Edition)", John Gilmore and Stan Shebs, 2004.
- "Dynamic Binary Analysis and Instrumentation", Nicholas Nethercote, PhD Dissertation, University of Cambridge, November 2004.
- \* "Using Valgrind to detect undefined value errors with bit-precision". Julian Seward and Nicholas Nethercote, Proceedings of the USENIX'05 Annual Technical Conference, Anaheim, California, USA, April 2005.
- "Good Practices in Library Design, Implementation, and Maintenance (Version 0.1)".
  Ulrich Drepper, Red Hat Inc. March 7, 2002
- Ce cours est partiellement inspiré d'une formation RedHat http://www.redhat.com/docs/wp/intro\_dev/index.html