LẬP TRÌNH HỆ THỐNG

ThS. Đỗ Thị Thu Hiền (hiendtt@uit.edu.vn)



TRƯỜNG ĐH CÔNG NGHỆ THÔNG TIN - ĐHQG-HCM

KHOA MẠNG MÁY TÍNH & TRUYỀN THỐNG

FACULTY OF COMPUTER NETWORK AND COMMUNICATIONS

Tầng 8 - Tòa nhà E, trường ĐH Công nghệ Thông tin, ĐHQG-HCM Điện thoại: (08)3 725 1993 (122)

Machine-level programming: Procedure (Hàm/Thủ tục)



Cơ chế gọi hàm/thủ tục (procedure)

1. Chuyển luồng

- Bắt đầu thực thi hàm được gọi
- Trở về vị trí đã gọi hàm
- 2. Truyền dữ liệu
 - Truyền tham số (arguments) cho hàm
 - Nhận giá trị trả về của hàm
- 3. Quản lý bộ nhớ
 - Cấp phát bộ nhớ khi thực thi hàm
 - Thu hồi bộ nhớ khi thực thi xong
- Tất cả đều thực hiện được ở mức máy tính!
- Hàm ở IA32 và x86-64 sẽ có một số khác biệt.

```
P(...) {
    = Q(x);
  printf "%d", y
int Q(\int i)
  int t
         = 3*i;
  int v;
  return v;
```

Cơ chế gọi hàm/thủ tục (procedure)

```
int main()
{
    int result = func(5,6);
    return result;
}
```

```
main:
                %ebp
        pushl
                %esp, %ebp
        movl
                $16, %esp
        subl
        pushl
                $6
        pushl
                $5
        call
                func
                %eax, -4(%ebp)
        movl
                -4(%ebp), %eax
        movl
        leave
        ret
```

```
int func(int x, int y)
  int sum = 0;
  sum = x + y;
  return sum;
}
```

func:

```
%ebp
pushl
mov1
       %esp, %ebp
       $16, %esp
subl
movl
       $0, -4(%ebp)
       8(%ebp), %edx
movl
       12(%ebp), %eax
movl
       %edx, %eax
addl
       %eax, -4(%ebp)
movl
       -4(%ebp), %eax
movl
leave
ret
```

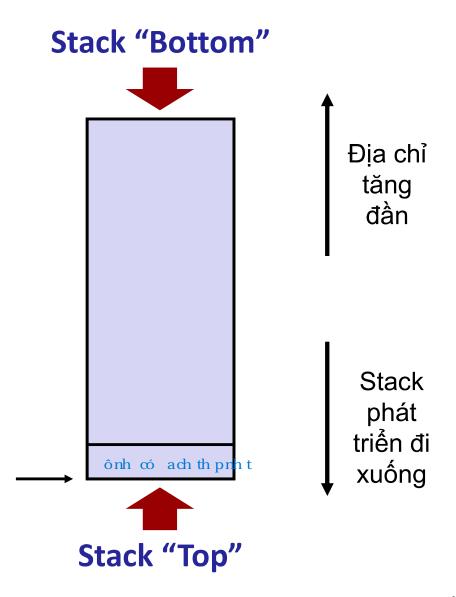
Nội dung

- Thủ tục (Procedures)
 - Cấu trúc stack
 - Gọi hàm trong IA32
 - Chuyển luồng
 - Truyền dữ liệu
 - Quản lý dữ liệu cục bộ
 - Gọi hàm trong x86-64
 - Minh hoạ hàm đệ quy
- Bài tập về hàm
- Dich ngược Reverse engineering

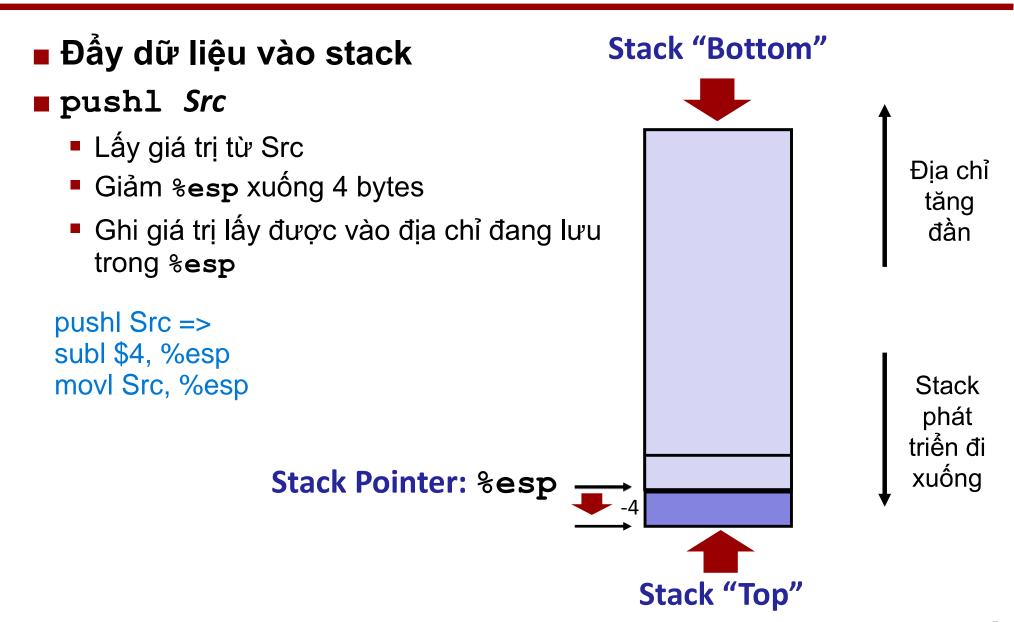
IA32 Stack

- Vùng nhớ được quản lý theo quy tắc ngăn xếp
 - First In Last Out
- Phát triển dần về phía địa chỉ thấp hơn
- Thanh ghi %esp chứa <u>địa chỉ</u> thấp nhất của stack
 - địa chỉ của "đỉnh" stack

Con tro stack (Stack Pointer): %esp



IA32 Stack: Push



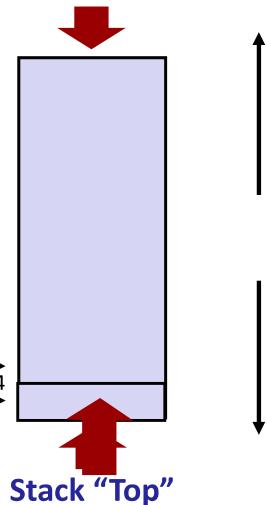
IA32 Stack: Pop

- Lấy dữ liệu từ stack
- popl Dest
 - Lấy giá trị ở địa chỉ lưu trong %esp
 - Đưa giá trị lấy được đưa vào Dest
 - Tăng %esp lên 4 bytes

popl Dest => movl (%esp), Dest addl \$4, %esp

Stack Pointer: %esp





Địa chỉ tăng đần

Stack phát triển đi xuống

IA32 Stack: Push and Pop - Ví du

- %esp = 0x108
- \bullet %eax = 0x1234
- %ebx = 0xABCD

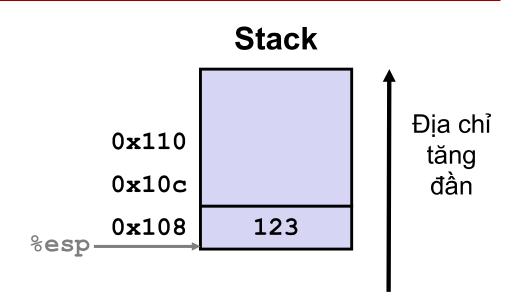
Các thanh ghi và stack thay đổi như thế nào khi thực hiện lần lượt các lệnh sau?

1. push %eax

%eax, %ebx không i %esp gi m xu ng 4 (bytes) Stack có thêm 1 giá tr là 0x1234 (giá tr c a %eax)

2. pop %ebx

%eax không i %espt ng lên 4 (bytes) Stack m t 1 giátr là 0x1234 %ebx c gán b ng giátr ãl yt stack là 0x1234



IA32 Stack: Push and Pop – Ví dụ 2

■ %esp =	0x108	3
----------	-------	---

Với các lệnh push dưới đây, giá trị bao nhiều được đưa vào stack?

Địa chỉ	Giá trị
0x108	0xF0
0x104	0xEF
0x100	0xAB

1. push \$0x100

Source là 1 h ngs, giátr Ox100s c push vào stack

2. push %eax

Source là 1 thanh ghi, giátr c a %eax s c push vào stack: 0x104

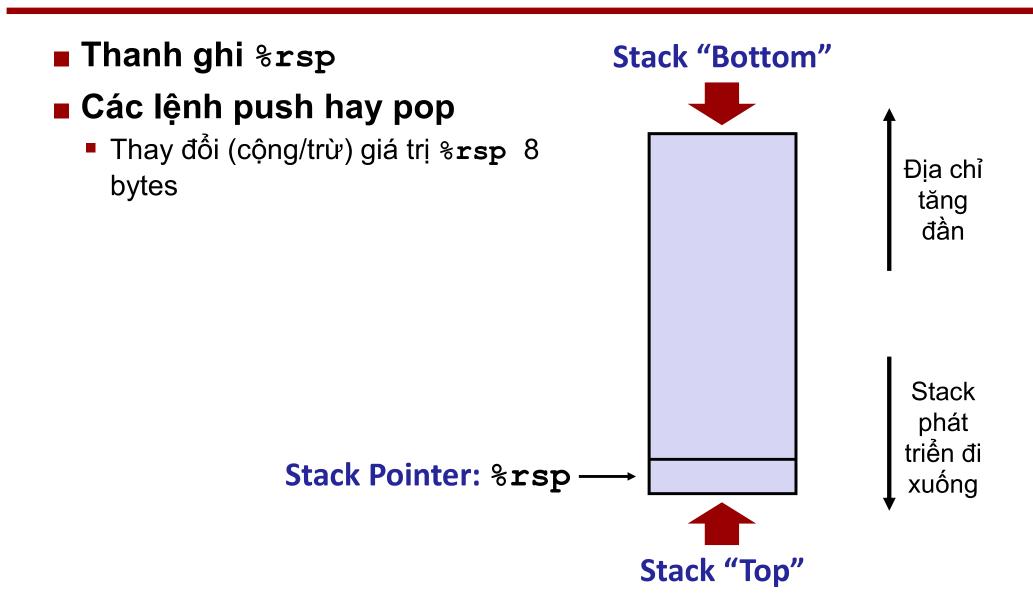
3. push (%eax)

Source là 1 ô nh , truy xu t ô nh có a ch trong %eax (ox 104) có giátr OxEF, giátr này c push vào stack

4. push 0x100

0xAB

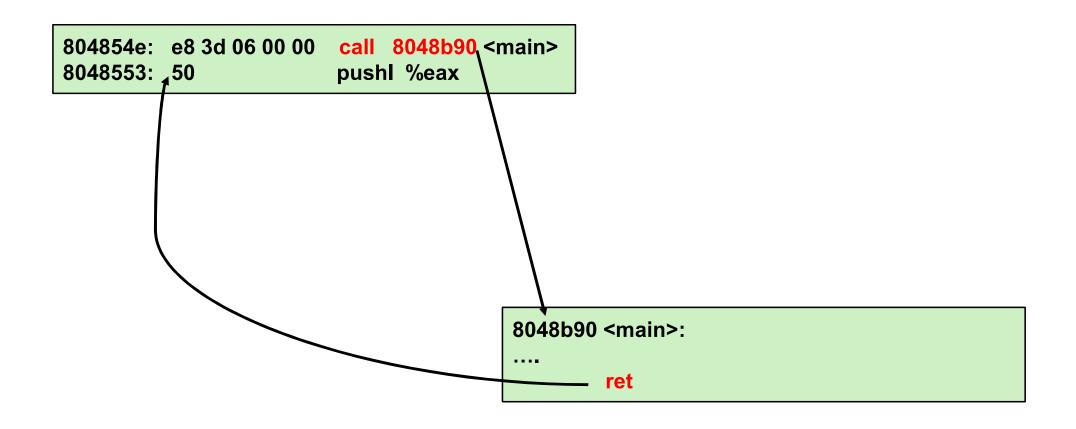
x86-64 Stack?



Nội dung

- Thủ tục (Procedures)
 - Cấu trúc stack
 - Gọi hàm trong IA32
 - Chuyển luồng
 - Truyền dữ liệu
 - Quản lý dữ liệu cục bộ
 - Gọi hàm trong x86-64
 - Minh hoạ hàm đệ quy
- Bài tập về hàm
- Dich ngược Reverse engineering

Chuyển luồng thực thi hàm



Chuyển luồng thực thi hàm

- Mỗi hàm đều có địa chỉ bắt đầu, thường được gán label
- Stack hỗ trợ gọi hàm và trở về từ hàm
 - Gọi 1 hàm con
 - Trở về hàm mẹ từ hàm con
- Gọi hàm: call label
 - Lưu địa chỉ trả về (return address) vào stack (push)
 - Nhảy đến label để thực thi
- Trở về từ hàm: ret
 - Lấy địa chỉ trả về ra từ stack (pop)
 - Nhảy đến địa chỉ lấy được để quay về hàm mẹ
- Địa chỉ trả về (Return address):
 - Địa chỉ câu lệnh assembly tiếp theo của hàm mẹ cần thực thi ngay phía sau lệnh call hàm con

8048553:

- Ví dụ trong mã assembly bên:
 - Địa chỉ trả về = 0x8048553

804854e: e8 3d 06 00 00 call 8048b90 <main>

Ví dụ: Gọi hàm và Trả về hàm

call 804854e: e8 3d 06 00 00 8048b90 <main> 8048553: 50 pushl %eax 0x8048553 — nằm trong %eip khi thực thi đến call Địa chỉ trả về? - Push địa chỉ trả về vào stack: push \$0x8048553 call main - Nhảy đến nhãn main: jmp 8048b90 - Push địa chỉ trả về vào stack: **push %eip** instruction pointer: con tr lnh(lu ach câu - Nhảy đến nhãn main: jmp 8048b90nhk ti pc acâu cth chi n) l nh 8048b90: main: 8048591: **c**3 ret - Pop địa chỉ trả về vào stack: pop %eip ret - Nhảy đến lệnh ở địa chỉ trả về: jmp *%eip

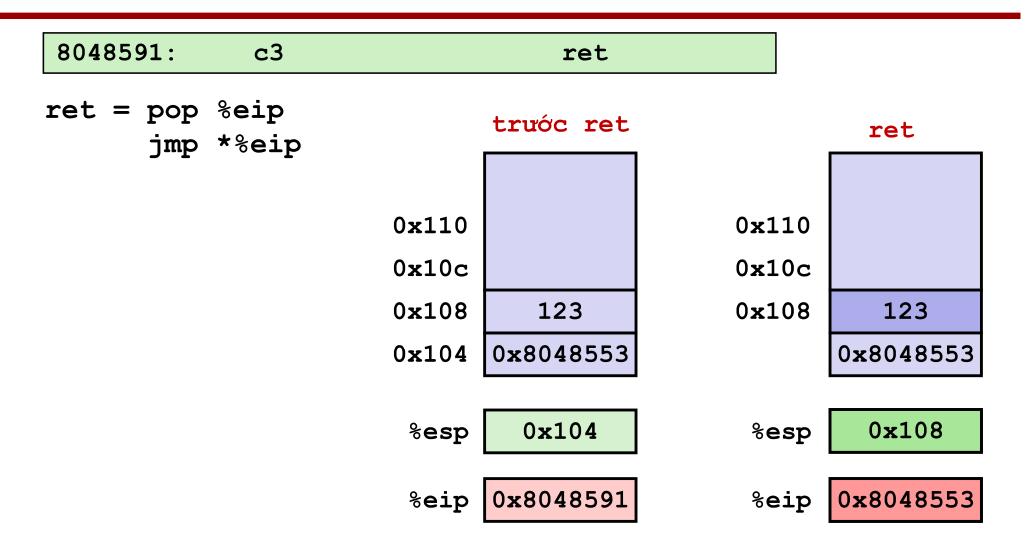
Ví dụ: Gọi hàm

%eip: program counter

804854e: e8 3d 06 00 00 call 8048b90 <main> 8048553: 50 pushl %eax call 0x8048b90 = push %eip th c thi l nh call có thay i stack (t ng kích th c stack, gi m%esp xu ng 4) jmp 0x8048b90 call 8048b90 trước call 0x1100x1100x10c0x10c123 123 0x1080x108%esp 0x1040x8048553%esp %esp %esp 0x108 0x1040x804854e 0x8048b90 %eip %eip

16

Ví dụ: Trả về hàm



%eip: program counter

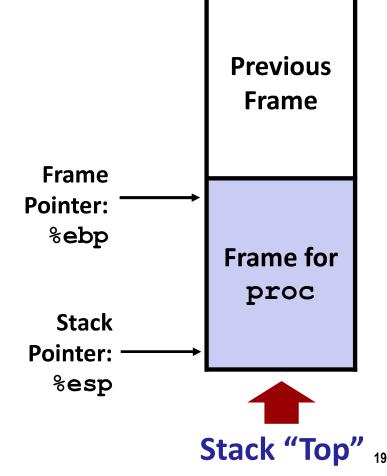
Gọi và trả về hàm – Ví dụ

```
int func(int x, int y)
int main()
                                          int sum = 0;
    int result = func(5,6);
                                          sum = x + y;
    return result;
                                          return sum;
                                 func:
main:
               %ebp
       pushl
                                         pushl
                                                %ebp
       movl %esp, %ebp
                                         movl
                                                %esp, %ebp
               $16, %esp
                                                 $16, %esp
        subl
                                         subl
                                                 $0, -4(%ebp)
       pushl
               $6
                                         movl
               $5
       pushl
                                                 8(%ebp), %edx
                                         movl
               func
                                                 12(%ebp), %eax
      → call
                                         movl
               %eax, -4(%ebp) 
                                                %edx, %eax
                                         addl
       movl
              -4(%ebp), %eax
                                                %eax, -4(%ebp)
       movl
                                         movl
 return
        leave
                                                 -4(%ebp), %eax
                                         movl
 addr
                                         leave
       ret
                                         ret
```

Hoạt động của hàm dựa trên stack

- Stack được cấp phát bằng Frames
 - 1 hàm (procedure) = 1 stack frame
 - Hỗ trợ lưu trữ các thông tin dùng để gọi và trả về hàm (procedure)
 - Địa chỉ trả về
 - Các tham số (arguments)
 - Các biến cục bộ (local variables)
- 1 Frame là vùng nhớ xác định bởi %ebp và %esp
 - %ebp trỏ đến vị trí cố định
 - %esp lưu động
 - Thường truy xuất các dữ liệu trên stack dựa trên %ebp

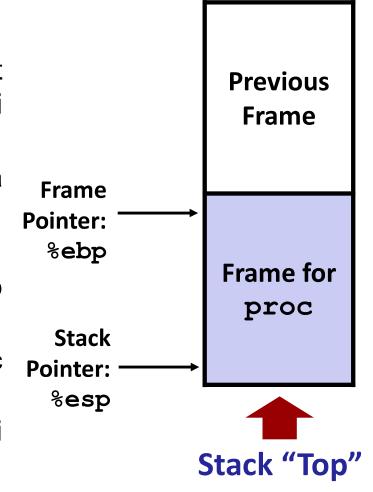
Ví dụ: -4(%ebp)



Stack Frames trong IA32

Quy tắc ngăn xếp

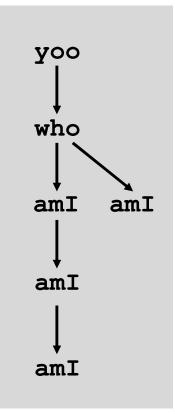
- Stack frame của 1 hàm tồn tại trong một khoảng thời gian từ lúc hàm được gọi đến lúc kết thúc.
 - Khi nào hàm được gọi thì stack frame của nó sẽ được tạo.
 - Khi kết thúc, stack frame sẽ được thu hồi.
- Hàm thực thi trước thì stack được cấp phát trước.
 - Stack frame cấp phát sau sẽ nằm ở các địa chỉ thấp hơn.
- Hàm kết thúc trước thì stack thu hồi trước.



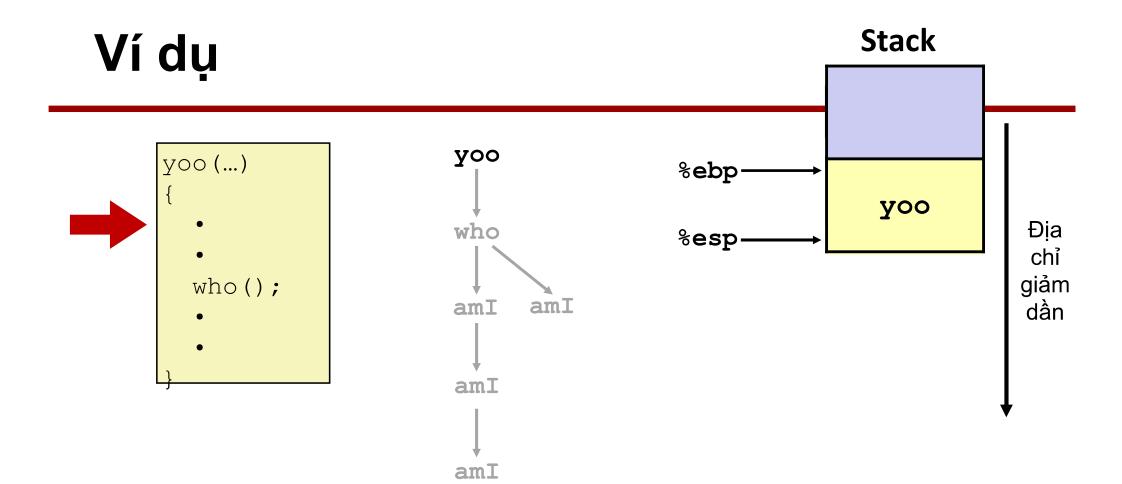
Ví dụ chuỗi gọi hàm

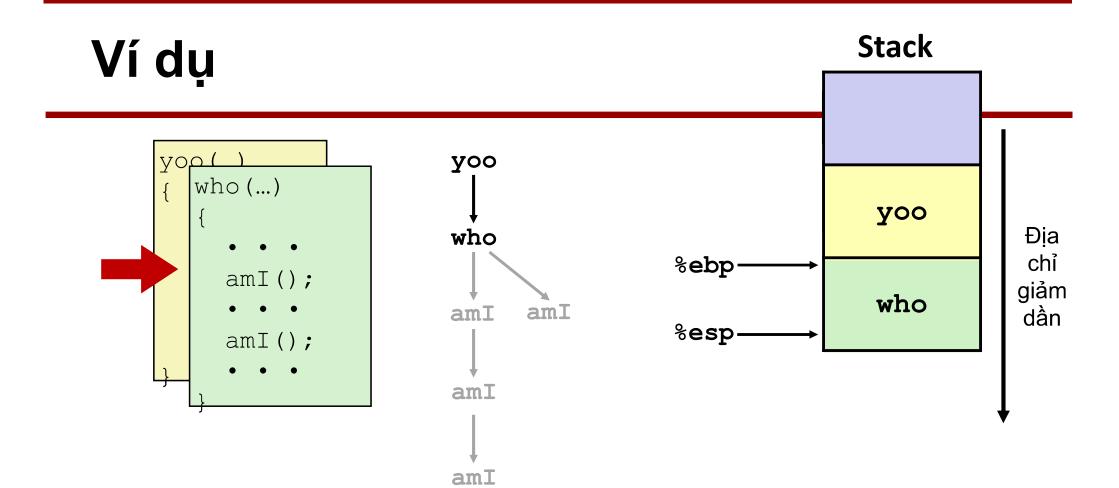
```
who(...)
{
    amI();
    am
```

Example Call Chain



Procedure amI() is recursive





Stack Ví dụ you yoo who (...) yoo Địa amI(...) who chỉ giảm who amI amI dần %ebpamI(); amI amI%esp

amI

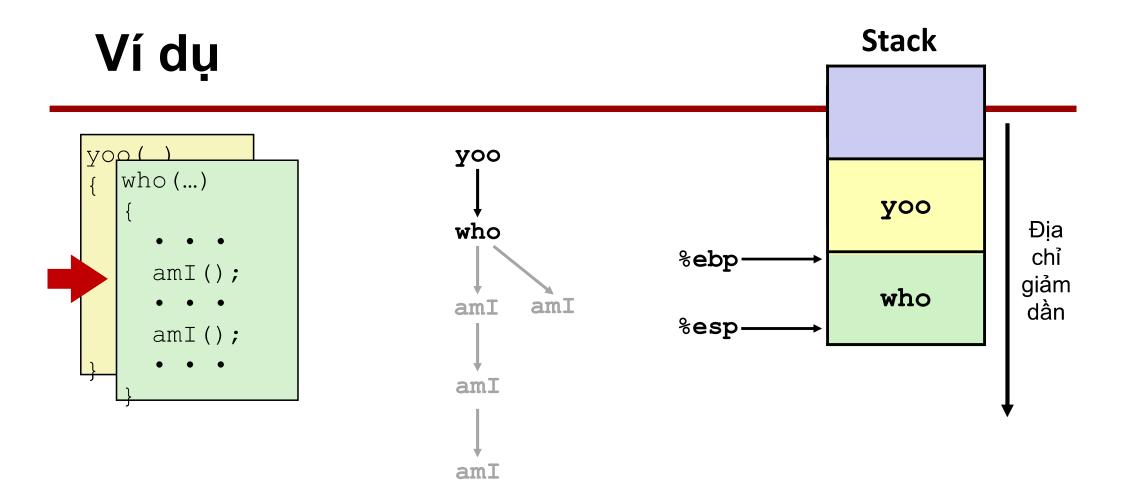
Stack Ví dụ YOPL yoo who (...) yoo amI (...) who Địa chỉ amI(...) giảm who amI amI dần amIamIamI(); %ebp amIamI %esp

Stack Ví dụ YOPL yoo who (...) yoo amI (...) who Địa chỉ • amI (...) giảm who amI amI dần • amI (...) amIamIamI(); amIamI %ebp amI%esp

Stack Ví dụ YOUL yoo who (...) yoo amI (...) who Địa chỉ amI (...) giảm who amI amI dần amIamIamI(); %ebp amIamI %esp

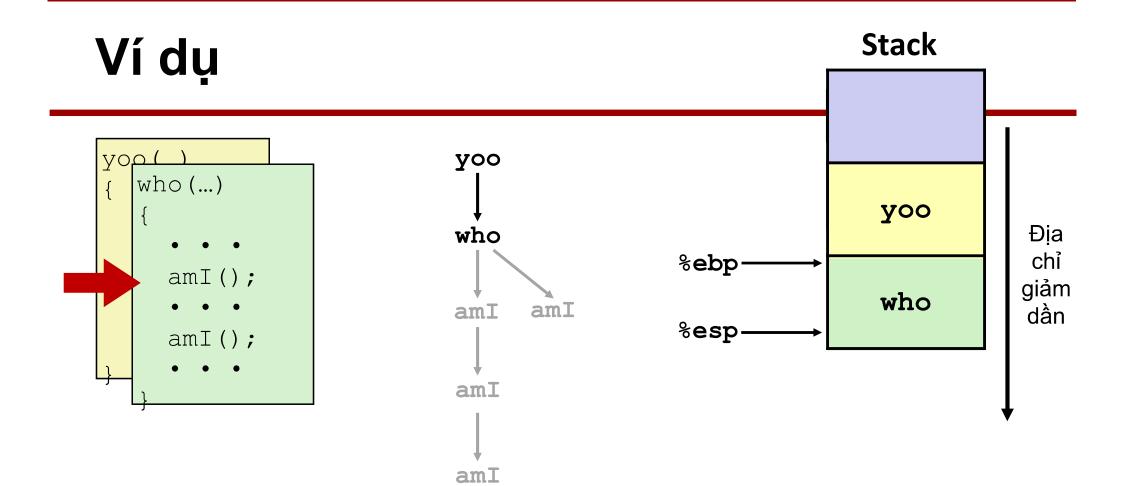
Stack Ví dụ YOOL yoo who (...) yoo Địa amI (...) who chỉ giảm who amI amI dần %ebpamI(); amI amI%esp

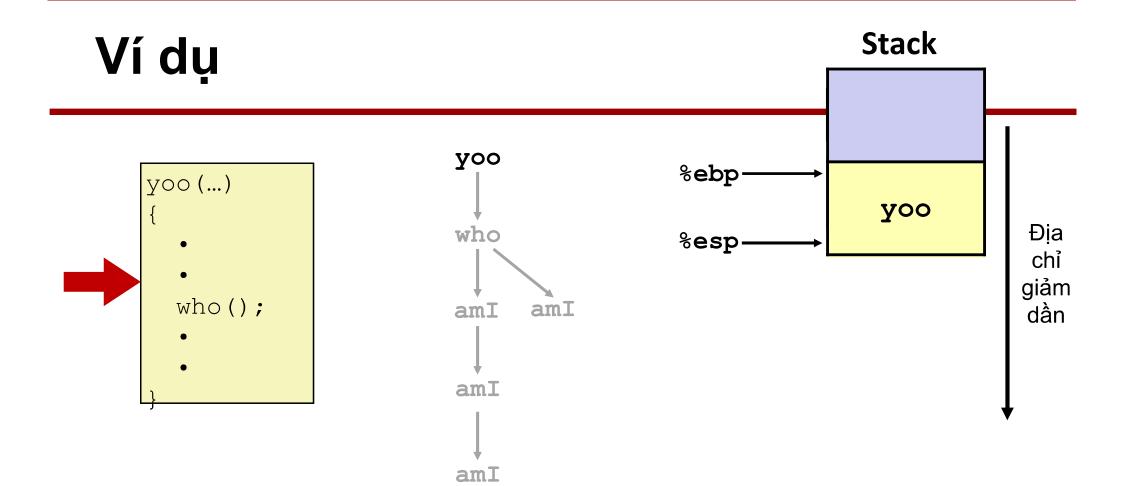
amI



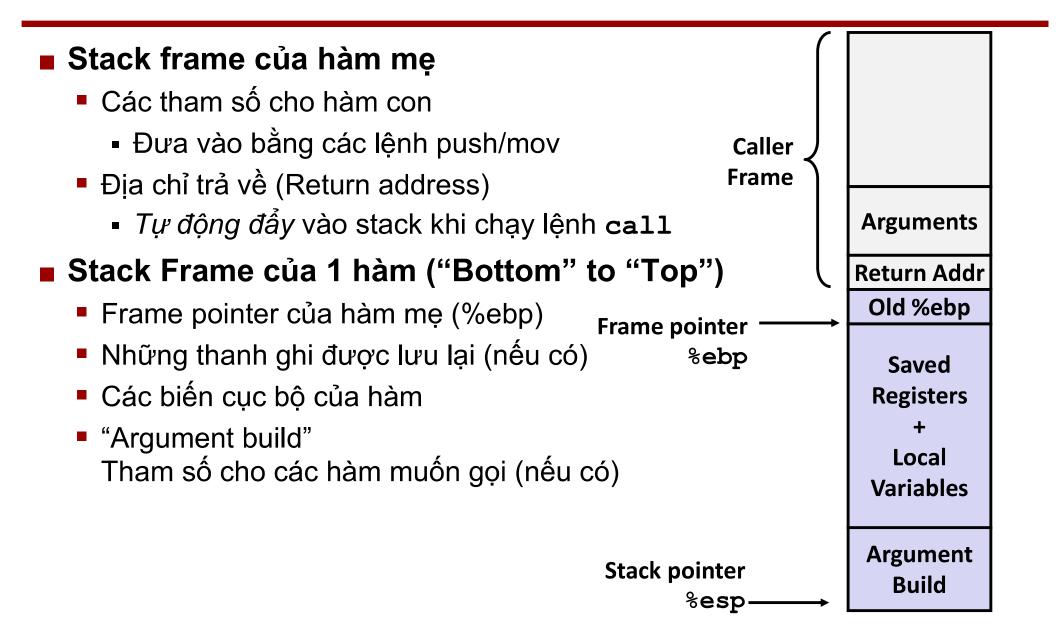
Stack Ví dụ YOOL yoo who (...) yoo Địa amI (...) who chỉ giảm who amI amI dần %ebp amI(); amI amI%esp

amI

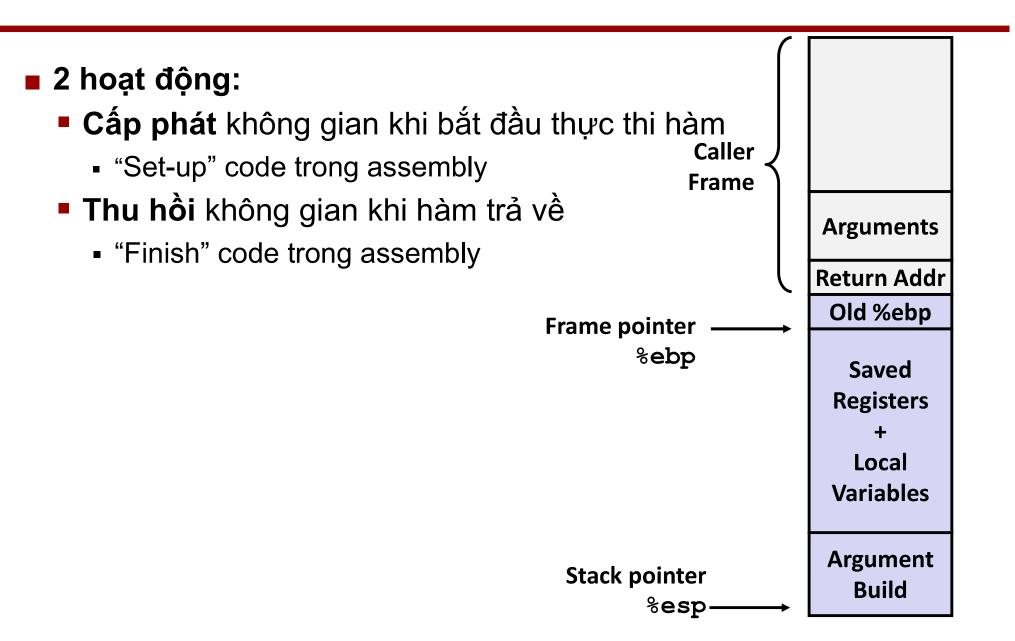




IA32 Stack Frame chứa thông tin gì?



Quản lý IA32 Stack Frame



IA32 Stack frame - Set up & Finish

■ Stack Frame – Set up

- Thực hiện khi 1 hàm bắt đầu thực thi
- Lưu lại %ebp của hàm trước
- Thiết lập %ebp cho stack frame của nó
- Lưu lại các thanh ghi sẽ sử dụng trong hàm (nếu có)

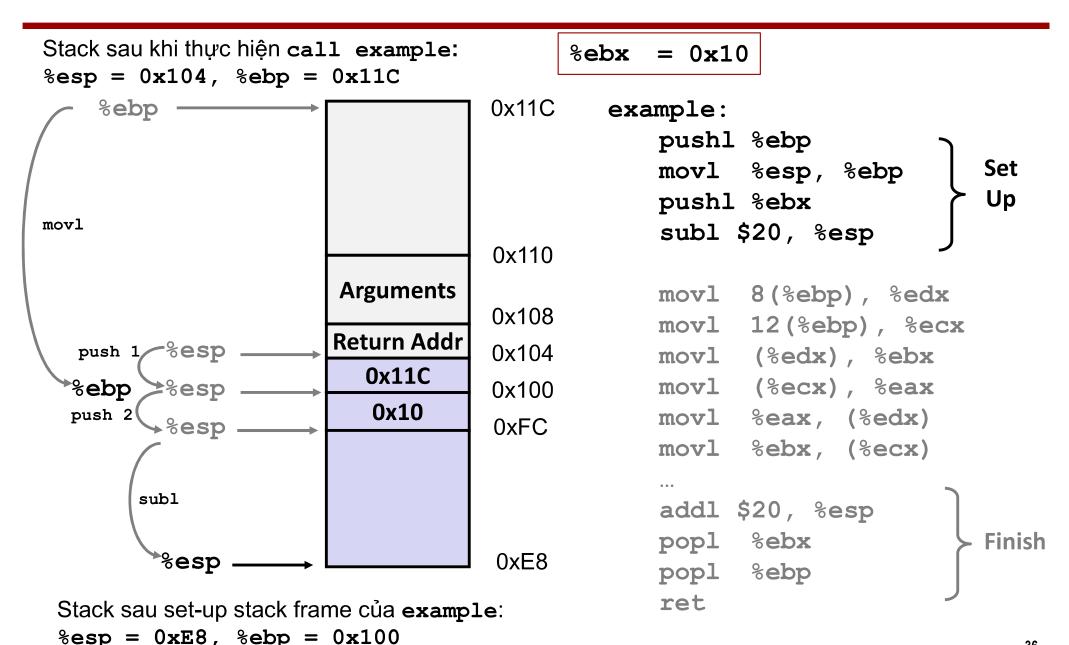
Stack frame - Finish

- Thực hiện khi 1 hàm chuẩn bị trả về
- Khôi phục giá trị cũ của các thanh ghi đã sử dụng (nếu có)
- Khôi phục %ebp của hàm trước

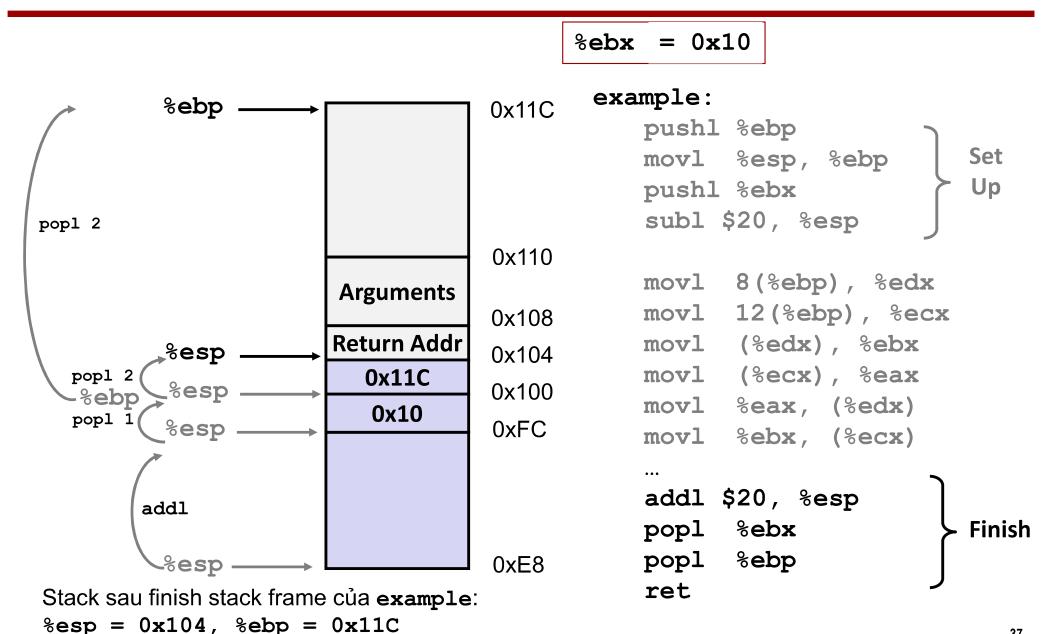
swap:

```
pushl %ebp
movl %esp, %ebp
pushl %ebx
      8(%ebp), %edx
movl
      12(%ebp), %ecx
movl
      (%edx), %ebx
movl
movl (%ecx), %eax
      %eax, (%edx)
movl
movl
      %ebx, (%ecx)
popl
      %ebx
      %ebp
popl
                     Finish
ret
```

Stack frame set up – Ví du



Stack frame Finish – Ví du



Stack frame set up & Finish – Ví dụ

```
int func(int x, int y)
    int sum = 0;
                     func:
    sum = x + y;
                             push1
                                     %ebp
                                                        Set
    return sum;
                                     %esp, %ebp
                             movl
                                                        Up
                                     $16, %esp
                             subl
                                     $0, -4(%ebp)
                             movl
                                     8(%ebp), %edx
                             movl
                                     12(%ebp), %eax
                             movl
                                     %edx, %eax
                             addl
                                     %eax, -4(%ebp)
                             movl
                                     -4(%ebp), %eax
                             movl
   Gán %esp = %ebp
                             leave
  - Pop %ebp từ stack
                                                        Finish
                             ret
```

Nội dung

- Thủ tục (Procedures)
 - Cấu trúc stack
 - Gọi hàm trong IA32
 - Chuyển luồng
 - Truyền dữ liệu
 - Quản lý dữ liệu cục bộ
 - Gọi hàm trong x86-64
 - Minh hoạ hàm đệ quy
- Bài tập về hàm
- Dich ngược Reverse engineering

Truyền tham số trong Stack frame IA32

- Hàm mẹ (caller) đưa tham số vào stack cho hàm con (callee)
 - Trước khi thực thi call label
 - Lệnh push/mov
 - Nàm ngay phía trước địa chỉ trả về (return address) trong stack
 - Thứ tự: reverse order
- Hàm con (callee) truy xuất tham số
 - Dựa trên vị trí so với %ebp của hàm con
 - %ebp sau khi hoàn thành "set up" code

```
Thams 1: v trí %ebp + 8
Thams 2 v trí %ebp + 12
Thams n: v trí %ebp + 4* (n+1)
```

Stack pointer %esp →

Frame pointer

%ebp

Caller

Frame

Argument n

• • •

Argument 3

Argument 2

Argument 1

Return Addr

Old %ebp

Saved Registers

+

Local

Variables

Argument Build

Truyền tham số cho hàm – Ví dụ 1

```
int func(int x, int y)
int main()
                                           int sum = 0;
    int result = func(5,6);
                                           sum = x + y;
    return result;
                                           return sum;
main:
                                 func:
                %ebp
        pushl
                                          pushl
                                                 %ebp
                %esp, %ebp
        movl
                                         movl
                                                 %esp, %ebp
                $16, %esp
                                                  $16, %esp
        subl
                                          subl
        pushl
                $6 ←
                                         movl
                                                  $0, -4(%ebp)
                $5 ←
                                                  8(%ebp), %edx
        pushl
                                         movl
                                                 12(%ebp), %eax
        call
                func
                                         movi
                $8, %esp
                                                  %edx, %eax
        addl
                                          addl
                %eax, -4(%ebp)
                                                  %eax, -4(%ebp)
        movl
                                         movl
                -4(%ebp), %eax
                                                  -4(%ebp), %eax
        mov1
                                         movl
                                         leave
        leave
        ret
                                         ret
```

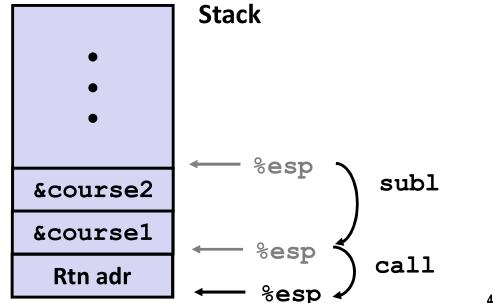
Truyền tham số cho hàm: Ví dụ 2 - swap

```
int course1 = 15213;
int course2 = 18243;
void call_swap() {
  swap(&course1, &course2);
```

Gọi swap từ hàm call swap

```
call swap:
   subl
         $8, %esp
  movl
         $course2, 4(%esp)
         $course1, (%esp)
  movl
   call
          swap
```

```
void swap(int *xp, int *yp)
  int t0 = *xp;
  int t1 = *yp;
  *xp = t1;
  *yp = t0;
```



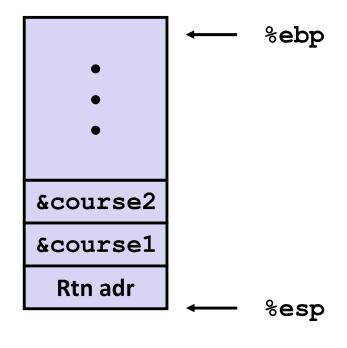
Truyền tham số: Ví dụ swap

```
void swap(int *xp, int *yp)
{
  int t0 = *xp;
  int t1 = *yp;
  *xp = t1;
  *yp = t0;
}
```

```
swap:
   pushl %ebp
                          Set
   movl %esp, %ebp
                          Up
   pushl %ebx
   movl 8(%ebp), %edx
         12(%ebp), %ecx
   movl
   movl (%edx), %ebx
                          Body
   movl (%ecx), %eax
   movl %eax, (%edx)
         %ebx, (%ecx)
   movl
   popl
         %ebx
                          Finish
         %ebp
   popl
   ret
```

swap Setup #1

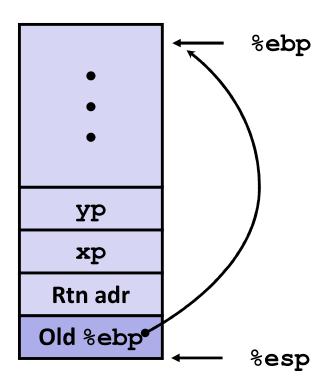
Entering Stack



swap:

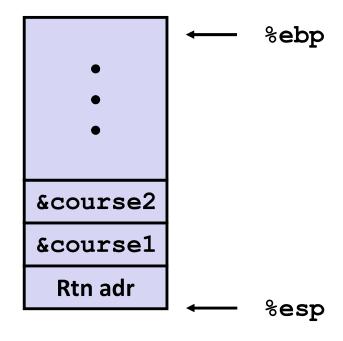
pushl %ebp
movl %esp,%ebp
pushl %ebx

Resulting Stack



swap Setup #2

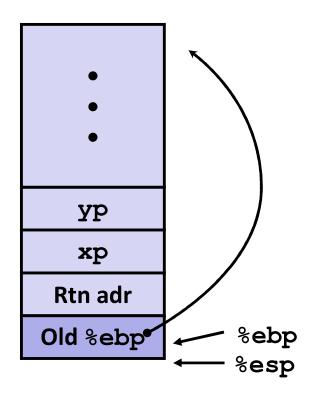
Entering Stack



swap:

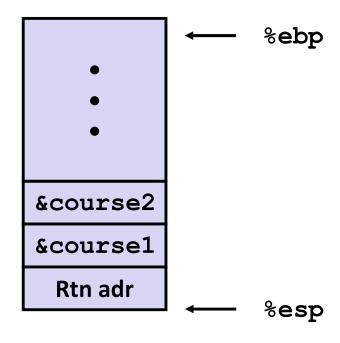
pushl %ebp
movl %esp,%ebp
pushl %ebx

Resulting Stack



swap Setup #3

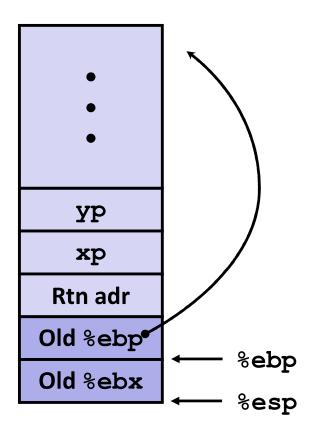
Entering Stack



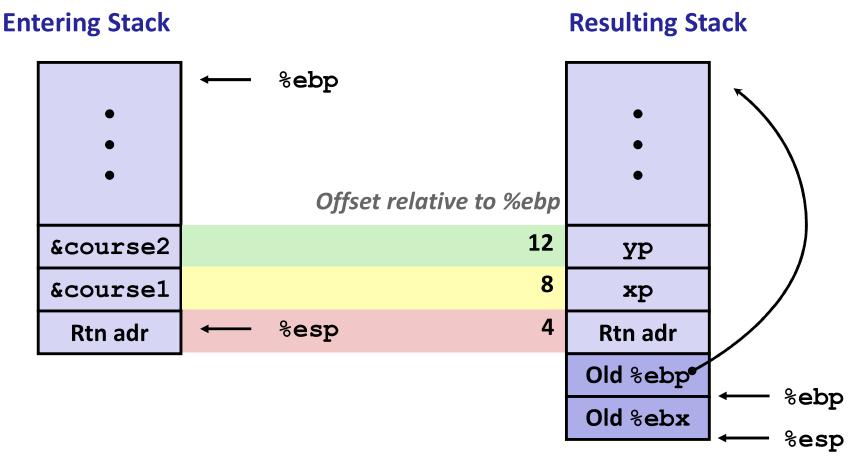
swap:

pushl %ebp
movl %esp,%ebp
pushl %ebx

Resulting Stack



swap: Lấy các tham số



```
movl 8(%ebp),%edx # get xp
movl 12(%ebp),%ecx # get yp
```

• • •

Giá trị trả về từ hàm

- Hàm có trả về giá trị
 - Trong C: qua lệnh return x.

- Giá trị trả về của hàm trong assembly
 - Thường lưu trong thanh ghi %eax

```
int Q(int i)
{
   int t = 3*i;
   int v[10];
   .
   return v[t];
}
```

Giá trị trả về từ hàm – Ví dụ

```
int func(int x, int y)
int main()
                                           int sum = 0;
    int result = func(5,6);
                                           sum = x + y;
    return result;
                                           return sum;
main:
                                 func:
        pushl
               %ebp
                                         pushl
                                                 %ebp
        movl %esp, %ebp
                                         movl
                                                 %esp, %ebp
                                                 $16, %esp
        subl
              $16, %esp
                                         subl
        pushl
                $6
                                         movl
                                                 $0, -4(%ebp)
                $5
                                         movl
                                                 8(%ebp), %edx
        pushl
                                                 12(%ebp), %eax
        call
                func
                                         movl
                $8, %esp
                                                 %edx, %eax
        addl
                                         addl
                                                 %eax, -4(%ebp)
                %eax <u>-4(%ebp)</u>
       movl
                                         movl
                -4(%ebp), %eax
                                                 -4(%ebp), %eax
        movl
                                         movl
                                         leave
        leave
        ret
                                         ret
```

Nội dung

- Thủ tục (Procedures)
 - Cấu trúc stack
 - Gọi hàm trong IA32
 - Chuyển luồng
 - Truyền dữ liệu
 - Quản lý dữ liệu cục bộ
 - Gọi hàm trong x86-64
 - Minh hoạ hàm đệ quy
- Bài tập về hàm
- Dich ngược Reverse engineering

Sử dụng thanh ghi cho trong hàm

- Giả sử yoo là hàm mẹ, gọi hàm who
- Có thể dùng thanh ghi để lưu trữ tạm?

```
yoo:

movl $15213, %edx
call who
addl %edx, %eax

ret
```

- Giá trị của thanh ghi %edx bị ghi đè trong hàm who
- Có thể gây ra vấn đề → cần lưu lại!

Quy ước lưu các thanh ghi

■ Giả sử yoo gọi who:

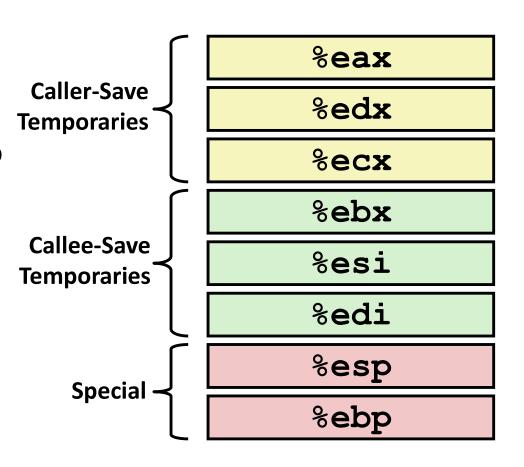
- yoo là hàm mẹ (caller)
- who là hàm con (callee)

■ Quy ước

- "Caller Save"
 - Hàm mẹ lưu lại các giá trị tạm thời trong stack frame của nó trước khi gọi hàm con
- "Callee Save"
 - Hàm con lưu lại các giá trị tạm thời trong stack của nó trước khi sử dụng

Sử dụng các thanh ghi IA32/Linux + Windows

- %eax, %edx, %ecx
 - Hàm mẹ lưu trước khi gọi nếu giá trị sẽ được sử dụng tiếp
- %eax
 - được sử dụng để trả về giá trị số nguyên
- %ebx, %esi, %edi
 - Hàm con sẽ lưu nếu muốn sử dụng
- %esp, %ebp
 - Trường hợp đặc biệt cần hàm con lưu
 - Khôi phục lại giá trị ban đầu trước khi thoát hàm



Khởi tạo biến cục bộ: Ví dụ

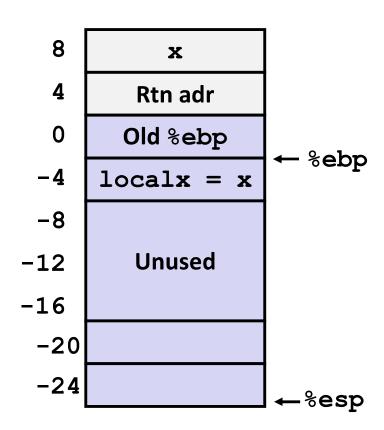
■ Biến cục bộ

- Cấp phát vùng nhớ trong stack để lưu các biến cục bộ của hàm
- Truy xuất dựa trên %ebp
 - Địa chỉ thấp hơn so với %ebp

```
int add3(int x) {
  int localx = x;
  incrk(&localx, 3);
  return localx;
}
```

First part of add3

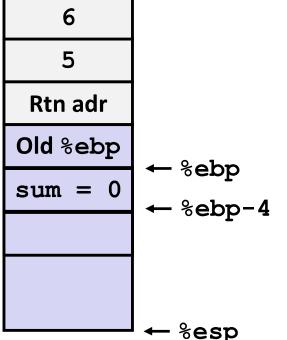
```
add3:
   pushl%ebp
   movl %esp, %ebp
   subl $24, %esp # Alloc. 24 bytes
   movl 8(%ebp), %eax
   movl %eax, -4(%ebp)# Set localx to x
```



Biến cục bộ – Ví dụ

```
int main()
{
   int result = func(5,6);
   return result;
}

func(int x, int y)
{
   int sum = 0;
   sum = x + y;
   return sum;
}
```



```
pushl
       %ebp
movl
       %esp, %ebp
       $16, %esp
subl
movl
       $0, -4(%ebp)
       8(%ebp), %edx
movl
       12(%ebp), %eax
movl
addl
       %edx, %eax
       %eax, -4(%ebp)
movl
       -4(%ebp), %eax
movl
leave
ret
```

Gọi hàm (IA32): Tổng kết

- Stack đóng vai trò quan trọng trong gọi/trả về hàm
 - Lưu trữ địa chỉ trả về

x = 2:

y = 1;

x + a = 12

Các tham số (trong stack frame hàm mẹ)

1. caller: Maco thế lưu các giá trị trong stack frame hoặc các thanh ghi callee: functi Giá trị trả về ở thanh ghi %eax

```
2. Hàm main:
Bi nc cb:-4(%ebp), -8(%ebp), -12(%ebp)
Giátr: 1, 2, 0
Hàm function:
Bi nc cb:-4(%ebp), -8(%ebp)
Giátr: 10, Giátr c a %eax (tính toán)

3 Function nh n 2 tham s
Hàm m push bao nhiều tham s tr ckhi g i hàm con / Hàm con có bao nhiều v trí liên quan n tham s 8(%ebp), 12(%ebp)

4. function(2, 1){}

5 Ch cn ng hàm funtion:
a = 10; %edx = a
```

Bài tập gọi hàm 1

```
main:
              %ebp
       pushl
       movl
              %esp, %ebp
              $16, %esp
       subl
       movl $1, -4(%ebp)
       mov1 $2, -8(%ebp)
              $0, -12(%ebp)
       movl
       pushl
              -4(%ebp)
       pushl
              -8(%ebp)
              function
       call
       addl $8, %esp
              %eax, -12(%ebp)
       mov1
              $0. %eax
       mov1
       leave
       ret
```

- 1. Hàm nào là caller/callee?
- 2. Mỗi hàm có bao nhiều biến cục bộ? Giá trị như thế nào?

```
function:
              %ebp
       pushl
              %esp, %ebp
       mov1
              $16, %esp
       subl
              $10, -4(%ebp)
       mov1
                              # a
       movl -4(\%ebp), \%edx
       movl 8(%ebp), %eax
                              # X
       addl %eax, %edx
       mov1
              12(%ebp), %eax
                              # y
       imull
              %edx, %eax
       mov1
              %eax, -8(%ebp)
              -8(%ebp), %eax
       mov1
                              # result
       leave
       ret
```

- 3. Hàm function nhận bao nhiều tham số?
- **4.** Hàm main đã truyền các tham số có giá trị cho function?
- **5.** Hàm function làm gì? Với các giá trị tham số đã tìm thấy ở Câu 4, tìm giá trị được function trả về cho main?