Programación a Nivel-Máquina IV: Procedimientos x86-64, Datos

Estructura de Computadores Semana 5

Bibliografía:

[BRY11] Cap.3 Computer Systems: A Programmer's Perspective. Bryant, O'Hallaron. Pearson, 2011

Signatura ESIIT/C.1 BRY com

Transparencias del libro CS:APP, Cap.3

Introduction to Computer Systems: a Programmer's Perspective

Autores: Randal E. Bryant y David R. O'Hallaron

Guía de trabajo autónomo (4h/s)

- **Lectura:** del Cap.3 CS:APP (Bryant/O'Hallaron)
 - x86-64, Procedures, Data Structures, Observations.
 - 3.13.4 .13.6 pp.316-325
 - Array Allocation & Access, Heterogeneous Data Structures (hasta Unions)
 - 3.8 3.9.1 pp.266-278
- **Ejercicios:** del Cap.3 CS:APP (Bryant/O'Hallaron)
 - Probl. 3.50 3.53 pp.318, 323-25
 - Probl. 3.35 3.39
 pp.267-68, 270, 272, 277

Bibliografía:

[BRY11] Cap.3

Computer Systems: A Programmer's Perspective. Bryant, O'Hallaron. Pearson, 2011

Signatura ESIIT/C.1 BRY com

Progr. Máquina IV: Procs. x86_64 / Datos

- Procedimientos (x86-64)
- Matrices*
 - Uni-dimensionales
 - Multi-dimensionales (anidadas)
 - Multi-nivel
- Estructuras
 - Ubicación
 - Acceso

x86-64, registros enteros

%rax	%eax	% r8	%r8d
%rbx	%ebx	%r9	%r9d
%rcx	%ecx	%r10	%r10d
%rdx	%edx	%r11	%r11d
%rsi	%esi	%r12	%r12d
%rdi	%edi	%r13	%r13d
%rsp	%esp	%r14	%r14d
%rbp	%ebp	%r15	%r15d

- El doble de registros
- Accesibles como 8, 16, 32, 64 bits

x86-64, registros enteros : Convenios de uso

%rax	Valor retorno
%rbx	Salva invocado
%rcx	Argumento #4
%rdx	Argumento #3
%rsi	Argumento #2
%rdi	Argumento #1
%rsp	Puntero de pila
%rbp	Salva invocado

%r8	Argumento #5
%r9	Argumento #6
%r10	Salva invocante
%r11	Salva invocante
%r12	Salva invocado
%r13	Salva invocado
%r14	Salva invocado
%r15	Salva invocado

x86-64, registros

Argumentos pasados a funciones vía registros

- Si más de 6 parámetros enteros, pasar resto en pila
- Estos registros pueden usarse también como salva-invocante

■ Todas las referencias a marco pila vía puntero pila

Elimina la necesidad de actualizar %ebp/%rbp

Otros registros

- 6 salva invocado
- 2 salva invocante
- 1 valor de retorno (también usable como salva invocante)
- 1 especial (puntero pila)

x86-64, swap* long int

```
void swap_l(long *xp, long *yp)
{
  long t0 = *xp;
  long t1 = *yp;
  *xp = t1;
  *yp = t0;
}
```

```
swap:
    movq (%rdi), %rdx
    movq (%rsi), %rax
    movq %rax, (%rdi)
    movq %rdx, (%rsi)
    ret
```

Dir. ret.

- Operandos pasados en registros
 - Primero (xp) en %rdi, 2º (yp) en %rsi
 - Punteros de 64-bit
- No requiere operaciones en pila (salvo ret)
- Evita marco pila
 - Puede mantener toda la información local en registros

%rsp

de pila

Sin marco

x86-64, Locales en la zona roja*

```
/* Swap, using local array */
void swap_a(long *xp, long *yp)
{
    volatile long loc[2];
    loc[0] = *xp;
    loc[1] = *yp;
    *xp = loc[1];
    *yp = loc[0];
}
```

```
swap_a:
    movq (%rdi), %rax
    movq %rax, -24(%rsp)
    movq (%rsi), %rax
    movq %rax, -16(%rsp)
    movq -16(%rsp), %rax
    movq %rax, (%rdi)
    movq -24(%rsp), %rax
    movq %rax, (%rsi)
    ret
```

Evita cambios del puntero pila

 Puede almacenar toda la info. dentro de una pequeña ventana más allá del puntero de pila (debajo)

```
Dir. ret. %rsp
-8 sin usar
-16 loc[1]
-24 loc[0]
```

x86-64, Procs. padre* sin marco pila

```
/* Swap a[i] & a[i+1] */
void swap_ele(long a[], int i)
    swap(&a[i], &a[i+1]);
```

- No hay valores que mantener mientras swap es invocado
- No necesita regs. salva-invocado
- Instr. rep insertada como no-op
 - Basado en recomendación de AMD++

```
swap ele:
                                  # Sign extend i
   movslq %esi,%rsi
          8(%rdi,%rsi,8), %rax # &a[i+1]
   leaq
   leaq (%rdi,%rsi,8), %rdi # &a[i] (1st arg)
                                  # (2<sup>nd</sup> arg)
          %rax, %rsi
   mova
   call
          swap
                                  # No-op
††
   rep
   ret
```

† "Move with Sign-extend Long to Quad", mnemotécnico MOVSXD según Intel † † problema predicción saltos Opteron y Athlon 64 (2003-2005) en RET 1B tras flowctrl. Software Optimization Guide for AMD64 Family 10-12h (2010-2011) recomienda RET 0 Software Optimization Guide for AMD64 Family 15h (2011) no menciona ningún problema 9

x86-64, Ejemplo de marco de pila

```
long sum = 0;
/* Swap a[i] & a[i+1] */
void swap_ele_su
   (long a[], int i)
{
    swap(&a[i], &a[i+1]);
    sum += (a[i]*a[i+1]);
}
```

- Mantiene los valores de &a[i] y &a[i+1] en registros salva-invocado
- Debe ajustar marco de pila para salvar estos registros

```
swap ele su:
          %rbx, -16(%rsp)
  movq
          %rbp, -8(%rsp)
  movq
   subq $16, %rsp
  movslq %esi,%rax
  leaq
          8(%rdi,%rax,8), %rbx
   leaq (%rdi, %rax, 8), %rbp
          %rbx, %rsi
  movq
          %rbp, %rdi
  movq
   call
          swap
          (%rbx), %rax
  movq
   imulq (%rbp), %rax
  addq
          %rax, sum(%rip)
  movq (%rsp), %rbx
  movq
          8(%rsp), %rbp
  addq
          $16, %rsp
   ret
```

Comprendiendo marco pila x86-64

```
swap ele su:
   movq %rbx, -16(%rsp)
                               # Salvar %rbx
   movq %rbp, -8(%rsp)
                               # Salvar %rbp
   subq $16, %rsp
                                # Reservar* marco pila
   movslq %esi,%rax
                               # Extender i (signo)
   leaq 8(%rdi,%rax,8), %rbx # &a[i+1] (salva invocado)
   leaq (%rdi,%rax,8), %rbp
                                # &a[i] (salva invocado)
   movq %rbx, %rsi
                                # 2° argumento
                                # 1er argumento
   movq
          %rbp, %rdi
   call
          swap
   movq (%rbx), %rax
                                # Traerse a[i+1]
   imulq (%rbp), %rax
                                # Multiplicar por a[i]
                                # Sumar a sum **
   addq
          %rax, sum(%rip)
                               # Restaurar %rbx
   movq (%rsp), %rbx
   movq 8(%rsp), %rbp
                               # Restaurar %rbp
   addq $16, %rsp
                                # Liberar* marco
   ret
```

%rbx

%rsp

Comprendiendo marco pila x86-64

```
# Salva%rbx
       %rbx, -16(%rsp)
movq
                                            %rsp
                                                     dir. ret.
                               # Salva%rbp
       %rbp, -8(%rsp)
movq
                                                  -8
                                                      %rbp
                                                      %rbx
                                                 -16
subq $16, %rsp
                               # Reservar marco pila
                                                      dir. ret.
                                                      %rbp
```

```
movq (%rsp), %rbx  # Restaurar %rbx
movq 8(%rsp), %rbp  # Restaurar %rbp
addq $16, %rsp  # Liberar marco
```

Características interesantes del marco pila

Reserva el marco entero de una vez

- Todos los accesos a pila pueden hacerse relativos a %rsp
- Hacerlo decrementando puntero pila
- Reserva puede posponerse, ya que es seguro usar temporalmente la zona roja

Liberación sencilla

- Incrementar puntero pila
- No se necesita puntero marco/base

Resumen Procedimientos (x86-64)

Uso intensivo de registros

- Paso de parámetros
- Más temporales ya que hay más registros

Uso mínimo de la pila

- A veces ninguno
- Reservar/liberar el bloque entero

Muchas optimizaciones complicadas

- Qué tipo de marco de pila usar
- Diversas técnicas de reserva

Progr. Máquina IV: Procs. x86_64 / Datos

- Procedimientos (x86-64)
- Matrices
 - Uni-dimensionales
 - Multi-dimensionales (anidadas)
 - Multi-nivel
- Estructuras

Tipos de datos básicos

Enteros

- Almacenados & manipulados en registros (enteros) propósito general
- Con/sin signo depende de las instrucciones usadas*

Intel	ASM**	Bytes	C
byte	b	1	[unsigned] char
word	W	2	[unsigned] short
double word	1	4	[unsigned] int
quad word	q	8	[unsigned] long int (x86-64)

■ Coma flotante

Almacenados & manipulados en registros coma flotante

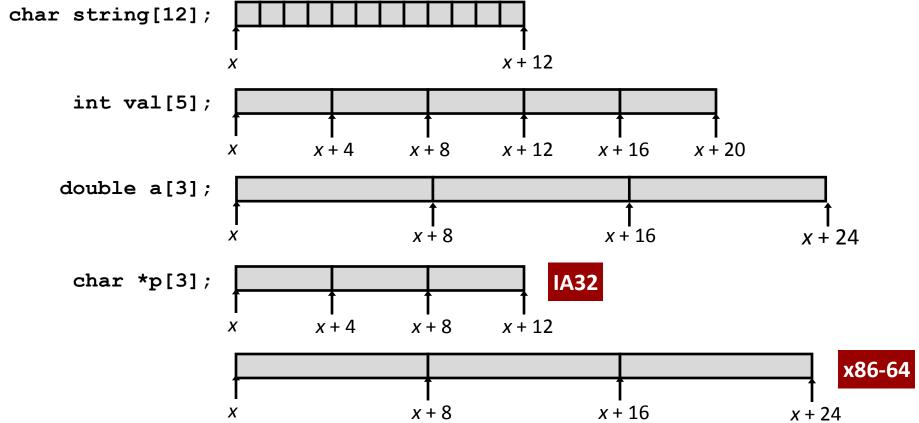
Intel	ASM	Bytes	С
Single	s	4	float
Double	1	8	double
Extended	t	10/12/16	long double

Ubicación* de matrices

Principio básico

```
T \mathbf{A}[L];
```

- Matriz de tipo T y longitud L
- Se reserva * región contigua de L * sizeof(T) bytes



Acceso a matrices

■ Principio básico

```
T A[L];
```

- Matriz de tipo T y longitud L
- El identificador **A** (tipo T^*) puede usarse como puntero al elemento 0

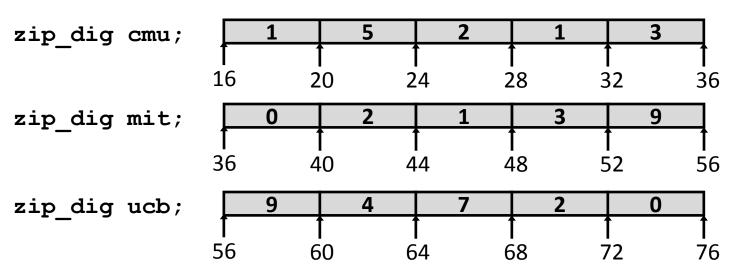
■ Referencia*	Tipo	Valor				
val[4]	int	3				
val	int *	X				
val+1	int *	x + 4	(aritmética de punteros)			
&val[2]	int *	<i>x</i> + 8				
val [5]	int	.				
* (val+1)	int	5	(aritmética de punteros)			
val + <i>i</i>	int *	x + 4i	(aritmética de punteros)			

^{*} otros autores usan "(de)reference" en sentido mucho más estricto, para indicar el tipo puntero (o la operación de seguir el puntero)18

Ejemplo de matrices

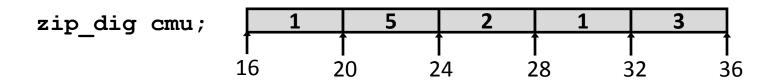
```
#define ZLEN 5
typedef int zip_dig[ZLEN];

zip_dig cmu = { 1, 5, 2, 1, 3 };
zip_dig mit = { 0, 2, 1, 3, 9 };
zip_dig ucb = { 9, 4, 7, 2, 0 };
```



- Declaración "zip dig cmu" equivalente a "int cmu[5]"
- Las matrices del ejemplo se ubicaron en bloques sucesivos de 20 bytes
 - En general no está garantizado que suceda

Ejemplo de acceso a matrices



```
int get_digit
  (zip_dig z, int dig)
{
  return z[dig];
}
```

IA32

```
# %edx = z
# %eax = dig
movl (%edx,%eax,4),%eax # z[dig]
```

- El registro %edx contiene dirección inicio de matriz
- El registro %eax contiene el índice a la matriz
- El dígito deseado está en 4*%eax + %edx
- Usar referencia a memoria (%edx, %eax, 4) *

Ejemplo de bucle sobre matriz (IA32)

```
void zincr(zip_dig z) {
  int i;
  for (i = 0; i < ZLEN; i++)
    z[i]++;
}</pre>
```

```
# edx = z
movl $0, %eax  # %eax = i
.L4:  # loop:
addl $1, (%edx,%eax,4) # z[i]++
addl $1, %eax  # i++
cmpl $5, %eax  # i:5
jne .L4  # if !=, goto loop
```

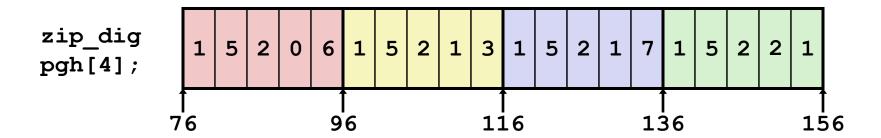
Ejemplo de bucle con puntero (IA32)

```
void zincr_p(zip_dig z) {
  int *zend = z+ZLEN;
  do {
    (*z)++;
    z++;
  } while (z != zend);
}
void zincr_v(zip_dig z) {
  void *vz = z;
  int i = 0;
  do {
    (*((int *) (vz+i)))++;
    i += ISIZE;
  } while (i != ISIZE*ZLEN);
}
```

```
# edx = z = vz
movl $0, %eax # i = 0
.L8: # loop:
addl $1, (%edx,%eax) # Increment vz+i
addl $4, %eax # i += 4
cmpl $20, %eax # Compare i:20
jne .L8 # if !=, goto loop
```

Ejemplo de matriz anidada

```
#define PCOUNT 4
zip_dig pgh[PCOUNT] =
   {{1, 5, 2, 0, 6},
    {1, 5, 2, 1, 3},
    {1, 5, 2, 1, 7},
    {1, 5, 2, 2, 1 }};
```



- "zip dig pgh[4]" equivalente a "int pgh[4][5]"
 - Variable pgh: matriz de 4 elementos, ubicados contiguamente
 - Cada elemento es una matriz de 5 ints, ubicados contiguamente
- Garantizado almacenamiento por filas ("row-major order")

Matrices multidimensionales (anidadas)

Declaración

 $T \mathbf{A}[R][C];$

- Matriz 2D de (elems. de) tipo T
- R filas (rows), C columnas
- Elems. tipo T requieren K bytes

Tamaño matriz

• *R* * *C* * *K* bytes

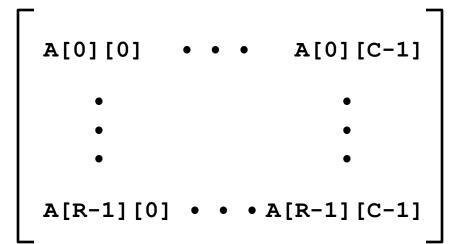
Disposición

Almacenamiento por filas

int A[R][C];

A [0] [0]	• • •	A [0] [C-1]		• • •	A [1] [C-1]	•	•	•	A [R-1] [0]	• • •	A [R-1] [C-1]
-----------------	-------	-------------------	--	-------	-------------------	---	---	---	-------------------	-------	---------------------

4*R*C Bytes -

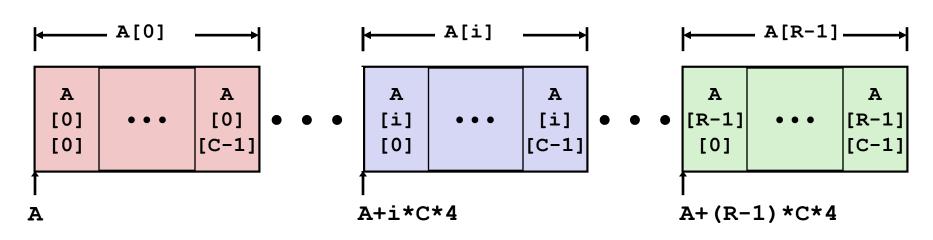


Acceso a filas en matrices anidadas

Vectores fila

- A[i] es una matriz de C elementos
- Cada elemento de tipo T requiere K bytes
- Dirección de comienzo A + i * (C * K)

int A[R][C];



Código acceso filas matrices anidadas

```
int *get_pgh_zip(int index)
{
  return pgh[index];
}
```

```
#define PCOUNT 4
zip_dig pgh[PCOUNT] =
   {{1, 5, 2, 0, 6},
    {1, 5, 2, 1, 3},
    {1, 5, 2, 1, 7},
    {1, 5, 2, 2, 1 }};
```

```
# %eax = index
leal (%eax,%eax,4),%eax # 5 * index
leal pgh(,%eax,4),%eax # pgh + (20 * index)
```

Vector fila

- pgh[index] es array de 5 ints
- Dirección de comienzo pgh+20*index

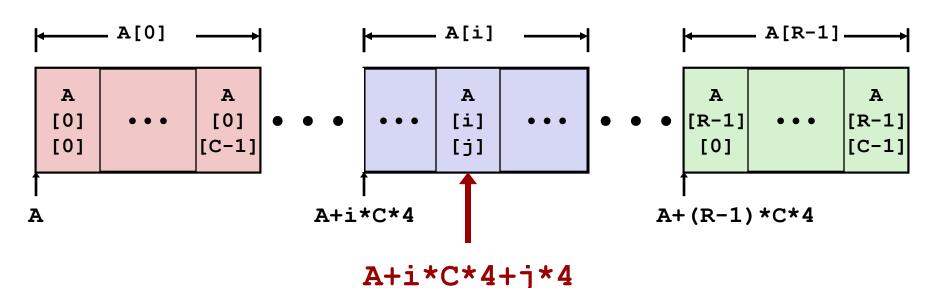
Código IA32

- Calcula y devuelve dirección
- La calcula como pgh + 4* (index+4*index)

Acceso a elementos en matrices anidadas

- Elementos de la matriz
 - A[i][j] es elemento de tipo T, que requiere K bytes
 - Dirección **A** + i * (C * K) + j * K = A + (i * C + j) * K

int A[R][C];



Código acceso elems. matrices anidadas

```
int get_pgh_digit
  (int index, int dig)
{
  return pgh[index][dig];
}
```

```
movl 8(%ebp), %eax  # index
leal (%eax,%eax,4), %eax # 5*index
addl 12(%ebp), %eax # 5*index+dig
movl pgh(,%eax,4), %eax # offset 4*(5*index+dig)
```

Elementos de la matriz

- pgh[index][dig] es int
- Dirección: pgh + 20*index + 4*dig
 - = pgh + 4*(5*index + dig)

Código IA32

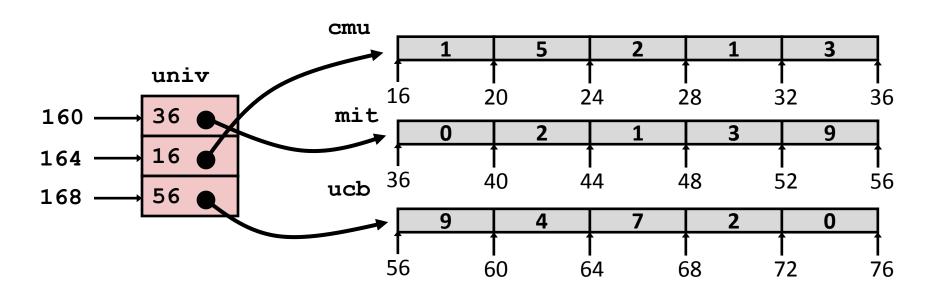
Calcula la dirección pgh + 4*((index+4*index)+dig)

Ejemplo de matriz multi-nivel*

```
zip dig cmu = \{ 1, 5, 2, 1, 3 \};
zip dig mit = \{0, 2, 1, 3, 9\};
zip dig ucb = { 9, 4, 7, 2, 0 };
```

```
#define UCOUNT 3
int *univ[UCOUNT] = {mit, cmu, ucb};
```

- Variable univ denota una matriz de 3 elementos
- Cada elemento un puntero
 - 4 bytes
- Cada puntero apunta a una matriz de ints



Acceso a elementos en matriz multi-nivel

```
int get_univ_digit
  (int index, int dig)
{
  return univ[index][dig];
}
```

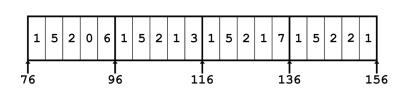
Cuentas (IA32)

- Acceso a elemento Mem [Mem [univ+4*index]+4*dig]
- Debe hacer dos lecturas de memoria
 - Primero obtener puntero a la matriz* fila
 - Entonces acceder elemento dentro de la matriz*

Acceso a elementos en matrices

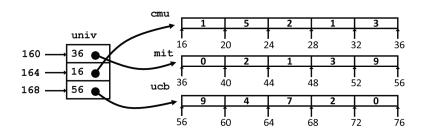
Matriz anidada

```
int get_pgh_digit
  (int index, int dig)
{
  return pgh[index][dig];
}
```



Matriz multi-nivel*

```
int get_univ_digit
  (int index, int dig)
{
  return univ[index][dig];
}
```



Accesos parecen similares en C, pero cuentas muy diferentes:

Mem[pgh+20*index+4*dig]

Mem[Mem[univ+4*index]+4*diq]

Cód. matriz N X N

Dimensiones fijas

 Se conoce valor de N en tiempo de compilación

Dimensiones variables, indexado explícito

 Forma tradicional de implementar matrices dinámicas

Dimensiones variables, indexado implícito

Soportado ahora* por gcc

```
#define N 16
typedef int fix_matrix[N][N];
/* Get element a[i][j] */
int fix_ele
   (fix_matrix a, int i, int j)
{
   return a[i][j];
}
```

```
#define IDX(n, i, j) ((i)*(n)+(j))
/* Get element a[i][j] */
int vec_ele
  (int n, int *a, int i, int j)
{
   return a[IDX(n,i,j)];
}
```

```
/* Get element a[i][j] */
int var_ele
  (int n, int a[n][n], int i, int j)
{
   return a[i][j];
}
```

Acceso a matriz 16 X 16

Elementos de la matriz

- Dirección **A** + i * (C * K) + j * K
- C = 16, K = 4

```
/* Get element a[i][j] */
int fix_ele(fix_matrix a, int i, int j) {
  return a[i][j];
}
```

```
movl 12(%ebp), %edx # i
sall $6, %edx # i*64
movl 16(%ebp), %eax # j
sall $2, %eax # j*4
addl 8(%ebp), %eax # a + j*4
movl (%eax,%edx), %eax # *(a + j*4 + i*64)
```

Acceso a matriz n X n

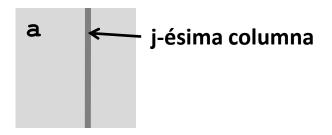
Elementos de la matriz

- Dirección **A** + i * (C * K) + j * K
- C = n, K = 4

```
/* Get element a[i][j] */
int var_ele(int n, int a[n][n], int i, int j) {
  return a[i][j];
}
```

```
movl 8(%ebp), %eax # n
sall $2, %eax # n*4
movl %eax, %edx # n*4
imull 16(%ebp), %edx # i*n*4
movl 20(%ebp), %eax # j
sall $2, %eax # j*4
addl 12(%ebp), %eax # a + j*4
movl (%eax,%edx), %eax # *(a + j*4 + i*n*4)
```

Optimizando acceso a matrices dims. fijas



Operación

 Pasar por todos los elementos de la columna j

Optimización

 Direccionando los sucesivos elementos de una misma columna (separados C*K)

```
#define N 16
typedef int fix_matrix[N][N];
```

```
/* Retrieve column j from array */
void fix_column
  (fix_matrix a, int j, int *dest)
{
  int i;
  for (i = 0; i < N; i++)
    dest[i] = a[i][j];
}</pre>
```

Optimizando acceso a matrices dims. fijas

Optimización

- Calcular ajp = &a[i][j]
 - Inicialmente = a + 4*j
 - Incrementos de 4*N

Registro	Valor
%ecx	ajp
%ebx	dest
%edx	i

```
/* Retrieve column j from array */
void fix_column
  (fix_matrix a, int j, int *dest)
{
  int i;
  for (i = 0; i < N; i++)
    dest[i] = a[i][j];
}</pre>
```

```
.L8:  # loop:
  movl (%ecx), %eax  # Read *ajp
  movl %eax, (%ebx,%edx,4) # Save in dest[i]
  addl $1, %edx  # i++
  addl $64, %ecx  # ajp += 4*N
  cmpl $16, %edx  # i:N
  jne .L8  # if !=, goto loop
```

Optimizando acceso a matrices dims. variabl.

- Calcular ajp = &a[i][j]
 - Inicialmente = a + 4*j
 - Incrementos de 4*n

Registro	Valor
%есх	ajp
%edi	dest
%edx	i
%ebx	4*n
%esi	n

```
/* Retrieve column j from array */
void var_column
  (int n, int a[n][n],
   int j, int *dest)
{
  int i;
  for (i = 0; i < n; i++)
    dest[i] = a[i][j];
}</pre>
```

Progr. Máquina IV: Procs. x86_64 / Datos

- Procedimientos (x86-64)
- Matrices
 - Uni-dimensionales
 - Multi-dimensionales (anidadas)
 - Multi-nivel

Estructuras

- Ubicación
- Acceso

Ubicación* de estructuras

```
struct rec {
  int a[3];
  int i;
  struct rec *n;
};
```

```
Disposic<sup>†</sup> Memoria

a i n

0 12 16 20
```

Concepto

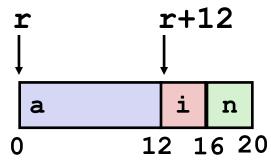
- Se reserva* región contigua de memoria
- Referencia a miembros de la estructura mediante sus nombres.
- Los miembros pueden ser de tipos diferentes

^{* &}quot;allocat(ion)/(ed)" en inglés

^{† &}quot;Memory Layout" = Disposición en Memoria. Para tamaño también se dice:

Acceso a estructuras

```
struct rec {
  int a[3];
  int i;
  struct rec *n;
};
```



Accediendo a un miembro de la estructura

- Puntero indica primer byte de la estructura*
- Acceder a los elementos mediante sus desplazamientos

Ensamblador IA32

```
# %edx = val
# %eax = r
movl %edx, 12(%eax) # Mem[r+12] = val
```

Generando puntero a miembro estructura

```
struct rec {
  int a[3];
  int i;
  struct rec *n;
};
```

```
r r+idx*4

a i n

12 16 20
```

Generando puntero a un elemento de la matriz

- Desplaz. de cada miembro struct queda determinado en tiempo compilación
- Argumentos
 - Mem[%ebp+8]: **r**
 - Mem[%ebp+12]: idx

```
int *get_ap
  (struct rec *r, int idx)
{
   return &r->a[idx];
}
```

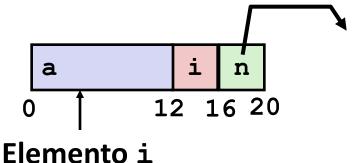
```
movl 12(%ebp), %eax # Get idx
sall $2, %eax # idx*4
addl 8(%ebp), %eax # r+idx*4
```

Siguiendo lista encadenada

Código C

```
void set_val
  (struct rec *r, int val)
{
  while (r) {
    int i = r->i;
    r->a[i] = val;
    r = r->n;
  }
}
```

```
struct rec {
  int a[3];
  int i;
  struct rec *n;
};
```



Registro Valor
%edx r
%ecx val

Resumen

Procedimientos en x86-64

- Marco pila es relativo a puntero pila
- Parámetros pasados en registros

Matrices

- Uni-dimensionales
- Multi-dimensionales (anidadas)
- Multi-nivel

Estructuras

- Ubicación
- Acceso

Guía de trabajo autónomo (4h/s)

- **Estudio:** del Cap.3 CS:APP (Bryant/O'Hallaron)
 - x86-64, Procedures, Data Structures, Observations.
 - 3.13.4 .13.6 pp.316-325
 - Probl. 3.50 3.53 pp.318, 323-25
 - Array Allocation & Access, Heterogeneous Data Structures (hasta Unions)
 - 3.8 3.9.1 pp.266-278
 - Probl. 3.35 3.39 pp.267-68, 270, 272, 277

Bibliografía:

[BRY11] Cap.3

Computer Systems: A Programmer's Perspective. Bryant, O'Hallaron. Pearson, 2011

Signatura ESIIT/C.1 BRY com