The procedure abstraction

Separate compilation:
- allows us to build large programs
- keeps compile times reasonable
- requires independent procedures

The linkage convention:
- a social contract
- machine dependent
- division of responsibility

The linkage convention ensures that procedures inherit a valid run-time environment and that they restore one for their parents.

Linkages execute at run time
Code to make the linkage is generated at compile time

Each system has a standard linkage
**Procedure linkages**

Assume that each procedure activation has an associated activation record or frame (at runtime).

**Assumptions:**
- RISC architecture
- Can always expand an allocated block
- Locals stored in frame

**Run-time storage organization**

To maintain the illusion of procedures, the compiler can adopt some conventions to govern memory use.

**Code space**
- Fixed size
- Statically allocated

**Data space**
- Fixed-sized data may be statically allocated
- Variable-sized data must be dynamically allocated

**Control stack**
- Dynamic slice of activation tree
- Return addresses
- May be implemented in hardware

**Run-time storage organization**

**Typical memory layout**

The classical scheme
- Allows both stack and heap maximal freedom
- Code and static data may be separate or intermingled

**Procedure linkages**

The linkage divides responsibility between caller and callee.

**Call**
- Pre-call
  1. Allocate basic frame
  2. Evaluate & store params.
  3. Store return address
  4. Jump to child

**Return**
- Post-call
  1. Copy return value
  2. Deallocate basic frame
  3. Restore parameters
  (if copy out)

At compile time, generate the code to do this. At run time, that code manipulates the frame & data areas.
Run-time storage organization

Where do local variables go?
When can we allocate them on a stack?
Key issue is lifetime of local names

Downward exposure:
- called procedures may reference my variables
- dynamic scoping
- lexical scoping

Upward exposure:
- can I return a reference to my variables?
- functions that return functions
- continuation-passing style

With only downward exposure, the compiler can allocate the frames on the run-time call stack

Storage classes

Each variable must be assigned a storage class

Static variables:
- addresses compiled into code
- (usually) allocated at compile-time
- limited to fixed size objects
- control access with naming scheme

Global variables:
- almost identical to static variables
- layout may be important
- naming scheme ensures universal access
- Link editor must handle duplicate definitions

Access to non-local data

How does the code find non-local data at run-time?

Real globals
- visible everywhere
- naming convention gives an address
- initialization requires cooperation

Lexical nesting
- view variables as (level,offset) pairs
- chain of non-local access links
- more expensive to find

Procedure local variables
Put them on the stack
- if sizes are fixed
- if lifetimes are limited
- if values are not preserved

Dynamically allocated variables
Must be treated differently
- call-by-reference, pointers, lead to non-local lifetimes
- (usually) an explicit allocation
- explicit or implicit deallocation
Access to non-local data

Two important problems arise
- How do we map a name into a (level, offset) pair?
  Use a block-structured symbol table (remember previous lecture?)
  - look up a name, want its most recent declaration
  - declaration may be at current level or any lower level
- Given a (level, offset) pair, what's the address?
  Two classic approaches
  - access links
  - displays (or static links)

Maintaining access links: (static links)
- calling level \( k + 1 \) procedure
  1. pass my FP as access link
  2. my backward chain will work for lower levels
- calling procedure at level \( l < k \)
  1. find link to level \( l - 1 \) and call it
  2. its access link will work for lower levels

Display management

Single global display: complex, obsolete method bogus idea, do not use

Call from level \( k \) to level \( l \)
- if \( l = k + 1 \), add a new display entry for level \( k \)
- if \( l = k \), no change to display is required
- if \( l < k \), preserve entries for levels \( l \) through \( k - 1 \) in the local frame

On return (back in calling procedure)
- if \( l < k \), restore preserved display entries

A single display ties up another register
Display management

Single global display: simple method

Key insight: overallocate the display by 1 slot

On entry to a procedure at level $i$
- save the level $i$ display value
- push FP into level $i$ display slot

On return
- restore the level $i$ display value

Quick, simple, and foolproof!

Individual frame-based displays:

Call from level $k$ to level $l$

if $l \leq k$
  copy $l - 1$ display entries into child's frame
if $l > k$ ($l = k + 1$)
  copy $k - 1$ entries into child's frame
  copy own FP into $k$th slot in child's frame

No work required on return
- display is deallocated with frame

Display accessed by offset from FP
⇒ one less register required

Display versus access links

How to make the trade-off?

The cost differences are somewhat subtle
- frequency of non-local access
- average lexical nesting depth
- ratio of calls to non-local access

(Sort of) Conventional wisdom
- tight on registers ⇒ use access links
- lots of registers ⇒ use global display
- shallow average nesting ⇒ frame-based display

Your mileage will vary
Making the decision requires understanding reality

Parameter passing

What about parameters?

Call-by-value
- store values, not addresses
- never restore on return
- arrays, structures, strings are a problem

Call-by-reference
- pass address
- access to formal is indirect reference to actual

Call-by-value-result
- store values, not addresses
- always restore on return
- arrays, structures, strings are a problem

Call-by-name
- build and pass thunk
- access to parameter invokes thunk
- all parameters are same size in frame
Parameter passing

What about variable length argument lists?

- if caller knows that callee expects a variable number
  - caller can pass number as 0th parameter
  - callee can find the number directly
- if caller doesn’t know anything about it
  - callee must be able to determine number
  - first parameter must be closest to FP

Consider printf:
- number of parameters determined by the format string
- it assumes the numbers match

Call/return

Assuming callee saves:
- caller pushes space for return value
- caller pushes SP
- caller pushes space for: return address, static chain, saved registers
- caller evaluates and pushes actuals onto stack
- caller sets return address, callee’s static chain, performs call
- callee saves registers in register-save area
- callee copies by-value arrays/records using addresses passed as actuals
- callee allocates dynamic arrays as needed
- on return, callee restores saved registers
- jumps to return address

Caller must allocate much of stack frame, because it computes the actual parameters
Alternative is to put actuals below callee’s stack frame in caller’s: common when hardware supports stack management (e.g., VAX)

Calls: Saving and restoring registers

<table>
<thead>
<tr>
<th>callee saves</th>
<th>caller’s registers</th>
<th>callee’s registers</th>
<th>all registers</th>
</tr>
</thead>
<tbody>
<tr>
<td>callee saves</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>caller saves</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

- Call includes bitmap of caller’s registers to save/restore (best with save/restore instructions to interpret bitmap)
- Caller saves and restores its own registers
- Unstructured returns (e.g., non-local gotos, exceptions) create some problems, since code to restore must be located and executed
- Backpatch code to save regs used in callee on entry, restore on exit e.g., VAX places bitmap in callee’s stack frame for use on call/return/non-local goto/exception
- Non-local gotos and exceptions must unwind dynamic chain restoring callee-saved registers
- Bitmap in callee’s stack frame is used by caller to save/restore (best with save/restore instructions to interpret bitmap directly) Unwind dynamic chain as for 3
- Easy: Non-local gotos and exceptions must restore all registers from “outermost callee”
- Easy (use utility routine to keep calls compact): Non-local gotos and exceptions need only restore original registers from caller

MIPS procedure call convention

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>zero</td>
<td>Constant 0</td>
</tr>
<tr>
<td>1</td>
<td>at</td>
<td>Reserved for assembler</td>
</tr>
<tr>
<td>2, 3</td>
<td>v0, v1</td>
<td>Expression evaluation, scalar function results</td>
</tr>
<tr>
<td>4–7</td>
<td>a0–a3</td>
<td>first 4 scalar arguments</td>
</tr>
<tr>
<td>8–15</td>
<td>t0–t7</td>
<td>Temporaries, callee-saved; caller must save to preserve across calls</td>
</tr>
<tr>
<td>16–23</td>
<td>s0–s7</td>
<td>Callee-saved; must be preserved across calls</td>
</tr>
<tr>
<td>24, 25</td>
<td>t8, t9</td>
<td>Temporaries, caller-saved; caller must save to preserve across calls</td>
</tr>
<tr>
<td>26, 27</td>
<td>k0, k1</td>
<td>Reserved for OS kernel</td>
</tr>
<tr>
<td>28</td>
<td>gp</td>
<td>Pointer to global area</td>
</tr>
<tr>
<td>29</td>
<td>sp</td>
<td>Stack pointer</td>
</tr>
<tr>
<td>30</td>
<td>s8 (fp)</td>
<td>Callee-saved; must be preserved across calls</td>
</tr>
<tr>
<td>31</td>
<td>ra</td>
<td>Expression evaluation, pass return address for calls</td>
</tr>
</tbody>
</table>
MIPS procedure call convention

Philosophy:

*Use full, general calling sequence only when necessary; omit portions of it where possible (e.g., avoid using fp register whenever possible)*

Classify routines as:
- non-leaf routines: routines that call other routines
- leaf routines: routines that do not themselves call other routines
  - leaf routines that require stack storage for locals
  - leaf routines that do not require stack storage for locals

Pre-call:
- Pass arguments: use registers $a0 . . . $a3; remaining arguments are pushed on the stack along with save space for $a0 . . . $a3
- Save caller-saved registers if necessary
- Execute a jal instruction: jumps to target address (callee’s first instruction), saves return address in register $ra

Prologue:
- Leaf procedures that use the stack and non-leaf procedures:
  - Allocate all stack space needed by routine:
    - local variables
    - saved registers
    - sufficient space for arguments to routines called by this routine
    - subu $sp, $sp, framesize
  - Save registers ($ra, etc.):
    - sw $31, framesize+frameoffset($sp)
    - sw $17, framesize+frameoffset-4($sp)
    - sw $16, framesize+frameoffset-8($sp)
  - where framesize and frameoffset (usually negative) are compile-time constants
  - Emit code for routine
MIPS procedure call convention

Epilogue:
- Copy return values into result registers (if not already there)
- Restore saved registers
  
  \[ \text{lw } \text{reg,framesize+frameoffset-N(sp)} \]
- Get return address
  
  \[ \text{lw } \$31,\text{framesize+frameoffset(sp)} \]
- Clean up stack
  
  \[ \text{addu } \text{sp,framesize} \]
- Return
  
  \[ \text{j } \$31 \]