Моделирование центрального процессора с помощью интерпретации

Курс «Программное моделирование вычислительных систем»

Григорий Речистов grigory.rechistov@phystech.edu

21 февраля 2015 г.



- 1 Конвейер процессора
- 2 Fetch
- 3 Decode
- 4 Execute
- 5 Write Back
- 6 Исключения
- 7 Advance PC
- 8 Улучшенные схемы
- 9 Литература



На прошлой лекции

Требования к симуляторам:

- точность,
- скорость,
- совместимость/расширяемость



■ В чём измеряется скорость симуляции?



- В чём измеряется скорость симуляции?
- Как соотносятся скорости функционального и потактового симуляторов?



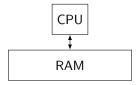
- В чём измеряется скорость симуляции?
- Как соотносятся скорости функционального и потактового симуляторов?
- Для чего необходимо предоставлять API симулятора?



- В чём измеряется скорость симуляции?
- Как соотносятся скорости функционального и потактового симуляторов?
- Для чего необходимо предоставлять API симулятора?
 Чтобы пользователи могли с ним поиграть.



Моделируемая система





Конвейер процессора





Переключаемый интерпретатор (switched)

```
while (run) {
    raw_code = fetch(PC);
    (opcode, operands) = decode(raw_code);
    switch (opcode) {
    case opcode1:
        func1(operands); PC++; break;
    case opcode2:
        func2(operands); PC++; break;
   /*...*/
}
```



Чтение инструкции из памяти



Чтение инструкции из памяти

```
data = mem[pc];
Скорее,
data = read_mem(pc);
```



Чтение инструкции из памяти

```
data = mem[pc];
Cкорее,
data = read_mem(pc);
И не забыть про преобразование адресов:
paddr = v2p(pc); // pc - vaddr
data = mem[paddr];
```

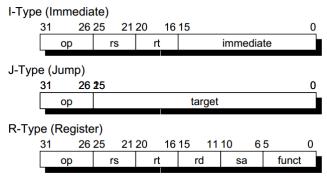


Декодирование

Перевод данных об инструкции из машинного представления (последовательность байт) во внутреннее (высокоуровневое), удобное для последующего использования



Пример: MIPS



MIPS Technologies Inc. MIPS32 4K™ Processor Core Family Software User's Manual - 2002



Пример: код 1

```
#define BITS(v, s, e) (v >> s) & ((1 << (e-s+1)) - 1)
typedef struct decode {
    uint32_t op;
    uint32_t rs;
    uint32_t rt;
    int32_t imm;
    int32_t tgt;
    /* ... */
} decode_t;
static inline
int32_t sign_extend(uint32_t v, int width)
    {/* ... */}:
```



Пример: код 2

```
decode_t decode(uint32_t raw) {
   uint32_t op = BITS(raw, 26, 31);
   uint32_t rs = BITS(raw, 21, 25);
   uint32_t rt = BITS(raw, 16, 20);
    int32_t imm = sign_extend(BITS(raw, 0, 15));
    int32_t tgt = sign_extend(BITS(raw, 0, 25));
   /* ... */
   uint32_t funct = BITS(raw, 0, 5);
   return (decode_t){op, rs, rt, /* ... */, funct};
```



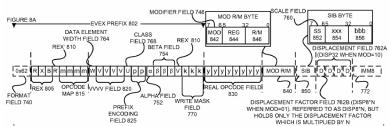
Более сложный пример: Intel IA-64 2.3

FP Arithmetic	F1	8 - D	X	sf		f ₄		f ₃	f ₂		f ₁			qp	
Fixed Multiply Add	F2	E	х	X ₂	П	f ₄	Г	f ₃	f ₂		f ₁			qp	1
FP Select	F3	Е	х			f ₄	Г	f ₃	f ₂		f ₁			qp	Ī
FP Compare	F4	4	r _b	sf	ra	P ₂	Г	f ₃	f ₂	ta	p	1		qp	
FP Class	F5	5	Г	If	c ₂	P ₂	Г	fclass _{7c}	f ₂	ta	p	1		qp	1
FP Recip Approx	F6	0 - 1	q	sf	X	P ₂	Г	f ₃	f ₂		f ₁			qp	Ī
FP Recip Sqrt App	F7	0 - 1	q	sf	X	P ₂	Г	f ₃			f ₁			qp	
FP Min/Max/Pcmp	F8	0 - 1		sf	X	x ₆	Г	f ₃	f ₂		f ₁			qp	Ī
FP Merge/Logical	F9	0 - 1	Г		×	x ₆	Г	f ₃	f ₂		f ₁			qp	
Convert FP to Fixed	F10	0 - 1		sf	X	x ₆	Г		f ₂		f ₁			qp	
Convert Fixed to FP	F11	0			X	x ₆			f ₂		f ₁			qp	
FP Set Controls	F12	0		sf	X	x ₆		omask _{7c}	amask _{7b}					qp	
FP Clear Flags	F13	0		sf	x	x ₆	Г							qp	
FP Check Flags	F14	0	S	sf	X	x ₆		imm _{20a}						qp	
Break	F15	0	İ		X	x ₆		imm _{20a}						qp	
Nop/Hint	F16	0	İ		X	x ₆	у							qp	
Break	X1	0	i	X	3	x ₆		imm _{20a}						qp	imm ₄₁
Move Imm ₆₄	X2	6	İ			imm _{9d}	imm _{5c} i _c v _c imm _{7b}			r ₁				qp	imm ₄₁
Long Branch	X3	С		d v	vh							btype		qp	imm ₃₉
Long Call	X4	D	i	ďν	vh	imm _{20b} p						b ₁		qp	imm ₃₉
Nop/Hint	X5	0	İ	X		x ₆ y imm _{20a}								qp	imm ₄₁
		40393837	736	3534	133	323130292827	26	252423222120	19181716151413	3121	1 10 9	8 7 6	5 4	3 2 1 0	

Intel Corporation. Intel® Itanium® Architecture Software Developer's Manual, p 3:296



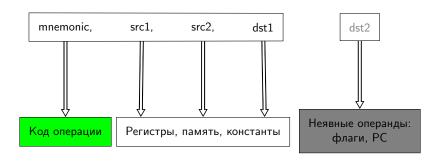
Пример посложнее: Intel IA-32



J.C.S. Adrian et al. Systems, Apparatuses, and Methods for Blending Two Source Operands into a Single Destination
Using a Writemask. US Patent Application Publication. № 2012/0254588 A1



Что извлекать из машинного кода инструкции



На выходе декодера:

- Успех, неуспех, недостаточно данных
- Для успеха: длина инструкции



Декодирование

- Код декодера редко пишется вручную, чаще он генерируется по описанию.
- A5 YX OZ OO \Rightarrow MOD RX, RY, RZ
- В общем случае: классическая задача построения синтаксического анализатора.
- На практике: специализированные инструменты и языки.



Декодирование: суровая реальность

- Переменная длина инструкций. IA-32: от 8 до 120 бит. Сколько байт пытаться декодировать за один раз?
- Зависимость смысла от префикса, режима работы процессора. Пример: 0x40 0x4f в IA-32/Intel 64/AMD64
- Полное несоответствие какому-либо здравому смыслу



Дизассемблирование

- Дизассемблирование перевод инструкций из машинного представление в понятный человеку вид (мнемонику)
- (За)кодирование (encoding) перевод инструкций из мнемонической записи в машинный код



Исполнение

- Базовая единица функция-эмулятор одной инструкции (service routine).
- Функция-эмуляторы пишутся на языке высокого уровня переносимость кода между хозяйскими платформами, компиляторами.



Пример: ADD reg reg reg

```
typedef struct {uint32_t regs[16];} cpu_t;

void add32_rrr(cpu_t *cpu, int src1, int src2, int dst) {
   cpu->regs[dst] = cpu->regs[src1] + cpu->regs[src2];
```

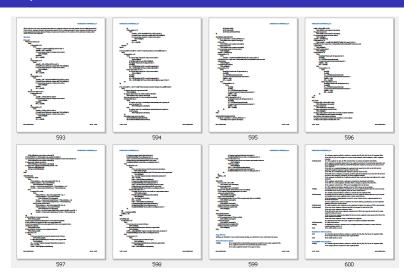


Пример: ADD reg reg reg

```
typedef struct {uint32_t regs[16];} cpu_t;
void add32_rrr(cpu_t *cpu, int src1, int src2, int dst) {
   cpu->regs[dst] = cpu->regs[src1] + cpu->regs[src2];
    cpu->z_flag = cpu->regs[dst] == 0;
    cpu-n_flag = cpu-regs[dst] & (1 << 31);
    cpu->o_flag = cpu->regs[dst] <</pre>
            MAX(cpu->regs[src1], cpu->regs[src2]);
    cpu->c_flag = calc_c_flag(cpu->regs[src1],
                               cpu->regs[src2]);
```



Пример посложнее: IA-32 CALL





Запись результата в память

```
write_mem(cpu, dst_addr, data);
```



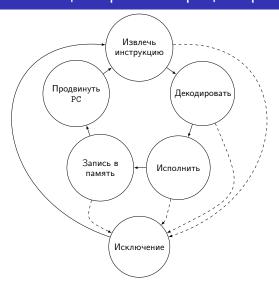
Запись результата в память

write_mem(cpu, dst_addr, data);

- Невыровненный адрес
- Граница страниц
- Запись в память «только для чтения»
- Исключения для всего или части записи/чтения



Уточнённый цикл работы процессора





Классификация

- Interruptions (термин из документации IA-64) вмешательство, перерыв, приостановка
- Exception синхронное исключение, без повторения текущей инструкции
- Fault синхронное, с повторением текущей инструкции
- Trap синхронное, без повторения, намеренно вызванное
- Interrupt внешнее асинхронное прерывание
- Abort внешние асинхронное с отсутствием информации о точке возврата



Продвижение РС

 Для большинства команд: увеличение счетчика на длину обработанной инструкции.

```
cpu->pc += instr_length;
```



Продвижение РС

 Для большинства команд: увеличение счетчика на длину обработанной инструкции.

```
cpu->pc += instr_length;
Бывают исключения: REP MOVS в IA-32.
```



Продвижение РС

 Для большинства команд: увеличение счетчика на длину обработанной инструкции.

```
cpu->pc += instr_length;
Бывают исключения: REP MOVS в IA-32.
```

- Явное изменение РС команды управления исполнением:
 - (Без)условные (не)прямой прыжок/переход
 - Вызов/возврат из процедуры
 - Системный вызов



Преимущества и недостатки интерпретации

- Пишется на языках высокого уровня: код переносим
- Простая структура: надёжность, расширяемость, переиспользование



Преимущества и недостатки интерпретации

- Пишется на языках высокого уровня: код переносим
- Простая структура: надёжность, расширяемость, переиспользование
- (Очень) низкая скорость работы



Куда тратится время?

```
start: interruption = false;
while (!interruption) {
    raw code = fetch(PC):
   (opcode, operands) = decode(raw_code); // <-- здесь
    switch (opcode) { // <-- здесь
    case opcode1:
        func1(operands); PC++; break;
    case opcode2:
        func2(operands); PC++; break;
    /*...*/
handle_interruption();
goto start;
```



Сцепленная интерпретация (threaded interpretation)

Вместо возвращения к началу цикла «прыгаем» прямо на исполнение следующей инструкции

```
func0: /* simulate instr0 */; PC++;
  next_opcode = decode(fetch(PC));
  goto func_ptr[next_opcode];
func1: /* simulate instr1 */; PC++;
  next_opcode = decode(fetch(PC));
  goto func_ptr[next_opcode];
func2: /* simulate instr2 */; PC++;
  next_opcode = decode(fetch(PC));
  goto func_ptr[next_opcode];
```

http://stackoverflow.com/questions/11227809/

why-is-processing-a-sorted-array-faster-than-an-unsorted-array



Кэширующая интерпретация

- В большинстве случаев в код гостевого приложения неизменен
- Велика вероятность того, что инструкции с некоторыми
 РС будут исполнены много раз (задача)
- Зачем каждый раз их декодировать?
- Заводим таблицу соответствия «адрес инструкции \to декодированный результат»



Кэширующая интерпретация

```
while (!interruption) {
  if (operation = cache[PC]); // короткий путь
  else { // не в кэше, длинный путь
  operation = decode(fetch(PC));
  cache[PC] = operation; // на будущее
  }
  switch (operation) {
    /* ... */
  }
}
```



Кэширующая интерпретация

- Ёмкость любого кэша ограничена, старые данные надо выбрасывать
- Необходимо следить за неизменностью исходного кода, иначе сохранённое соответствие будет неверным



Итоги

- Фазы исполнения: F, D, E, W, A
- Decoder, disassembler, encoder
- Переключаемый (switched) И.
- Сцепленный (threaded) И.
- Кэшируюший И.
- Ситуации: interrupt, trap, exception, fault, abort



Литература I

- D. Mihoka, S. Shwartsman. Virtualization Without Direct Execution or Jitting: Designing a Portable Virtual Machine
- Fredrik Larsson, Peter Magnusson, Bengt Werner. SimGen: Development of Efficient Instruction Set Simulators ftp://ftp.sics.se/pub/SICS-reports/Reports/SICS-R--97-03--SE.ps.Z

Infrastructure http://bochs.sourceforge.net/

Yair Lifshitz, Robert Cohn, Inbal Livni, Omer Tabach, Mark Charney, Kim Hazelwood. Zsim: A Fast Architectural Simulator for ISA Design-Space Exploration http://www.cs.virginia.edu/kim/docs/wish11zsim.pdf



Литература II

Префиксы в системе команд IA-32 http://habrahabr.ru/company/intel/blog/200598/

 Программная симуляция микропроцессора. Коробка передач

http://habrahabr.ru/company/intel/blog/202926/



На следующей лекции

Моделирование архитектурного состояния



Спасибо за внимание!

Замечание: все торговые марки и логотипы, использованные в данном материале, являются собственностью их владельцев. Представленная здесь точка эрения отражает личное мнение автора, не выступающего от лица какой-либо ооганизации.

