



# Measuring instruction latencies with llvm

**Guillaume Chatelet**

C. Courbet, B. De Backer, O. Sykora

Google Compiler Research

# Why?

- Scheduling needs latencies and  $\mu$ Op decomposition
  - This talk is about **latency measurement only**
- Vendors release some information
  - May be incomplete / not be in a machine readable format
- Updating LLVM td files
  - is tedious / requires careful guesswork and analysis.
- Consequences
  - scheduling information is [incomplete](#) for most X86 models

# How it works

∀ processor, ∀ instruction:

```
start_measure  
  .rept 10000  
  add rax, rax  
  .endr  
end_measure
```

# How it works - actually subtler than this...

∀ processor, ∀ instruction:

```
start_measure
```

```
.rept 10000
```

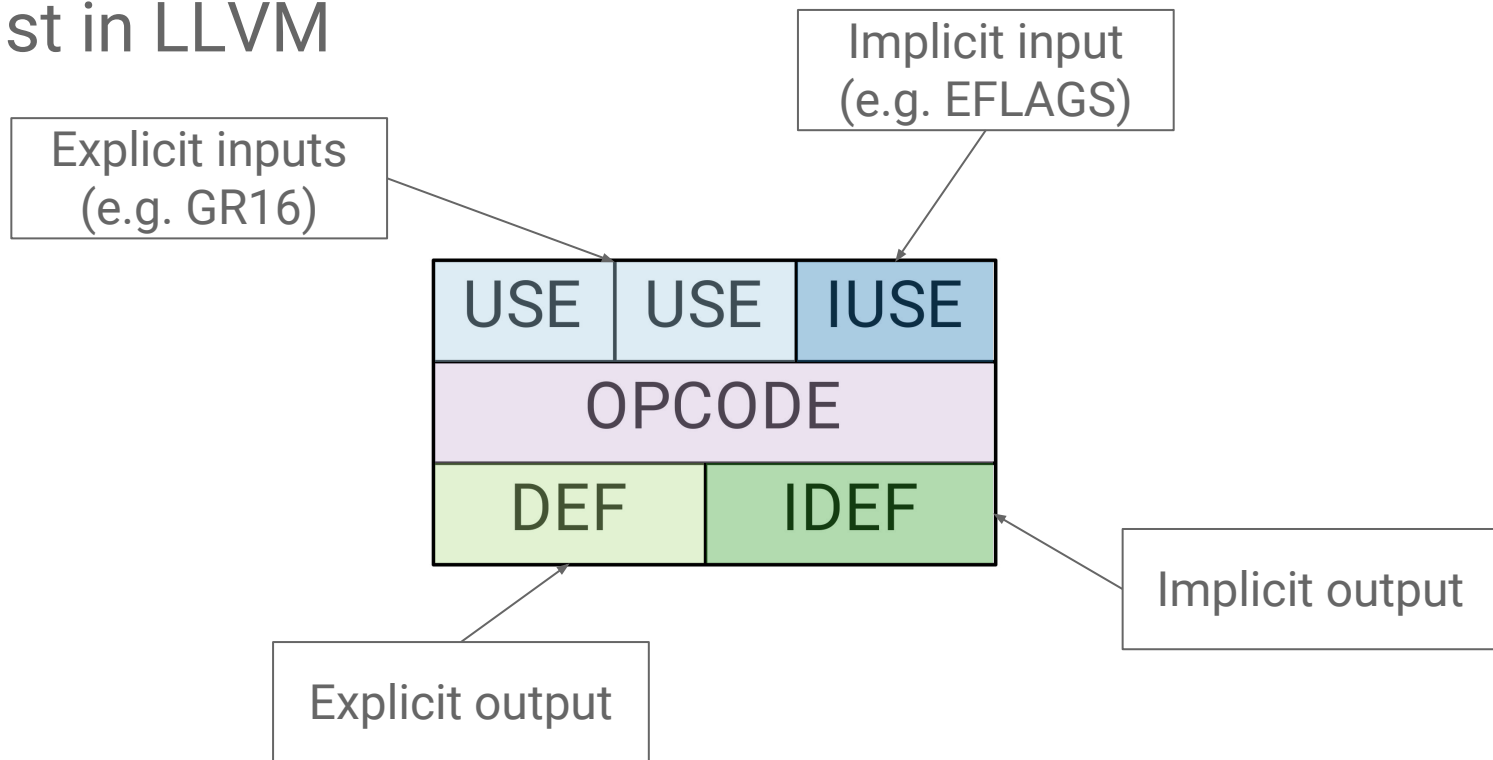
```
andn eax, ebx, edx # processor can execute these in parallel
```

```
.endr
```

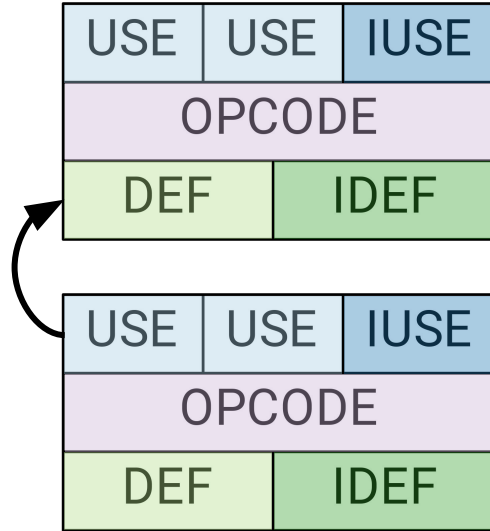
```
end_measure
```

- We need a way to make the **execution sequential**

# MCIInst in LLVM



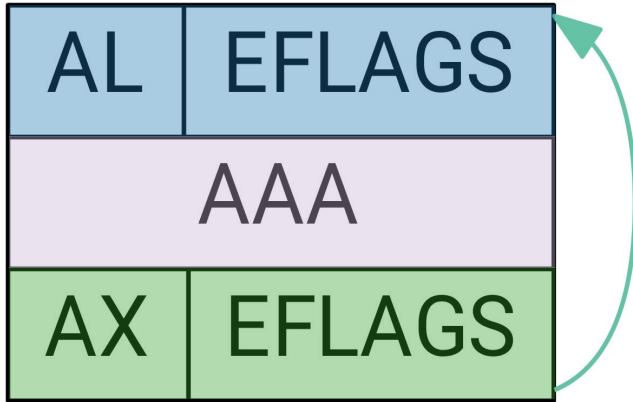
# Sequential execution: Create Dependency



Current instruction must use an output of previous instruction

# Implicit self cycle

Possible cycle:

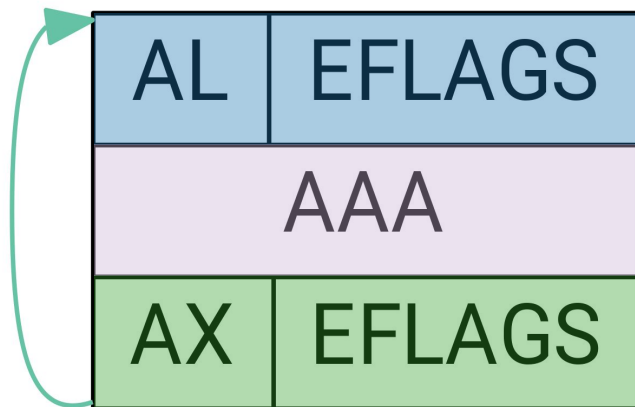


Possible instance:

**AAA**

# Implicit self cycle - through register aliasing

Possible cycle:



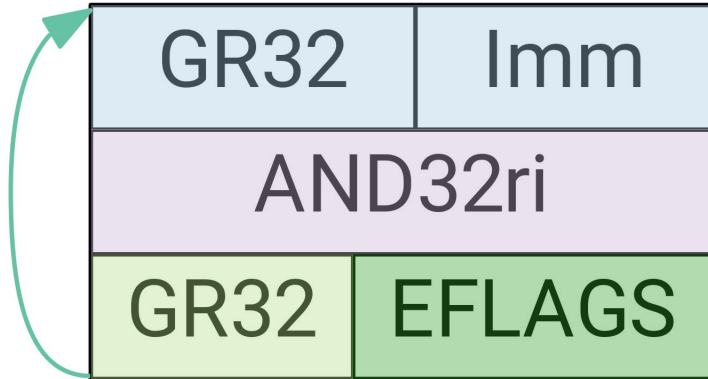
Possible instance:

**AAA**



# Possible explicit self cycle

Possible cycle:

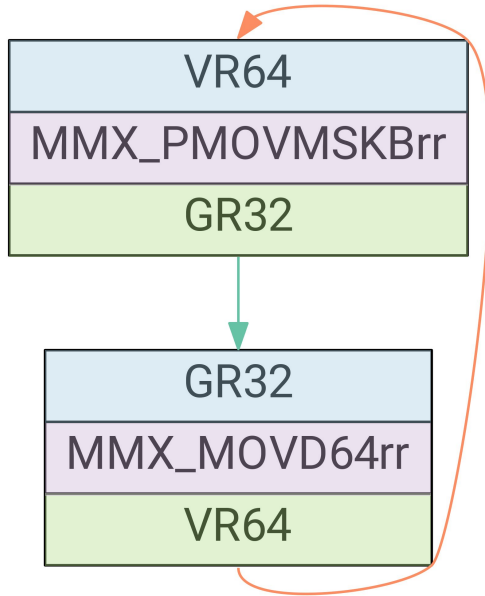


Possible instance:

```
AND32ri EAX, EAX, 1
```

# Possible cycle through another instruction

Possible cycle:

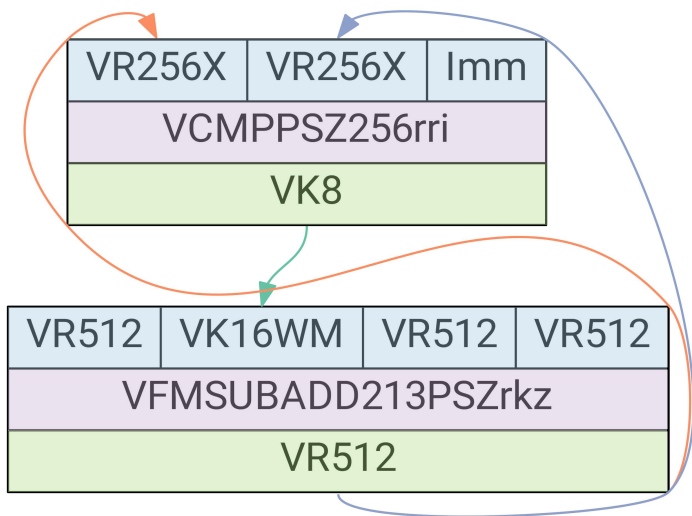


Possible instance:

`MMX_PMOVMSKBrr R10D, MM1`  
`MMX_MOVD64rr MM1, R10D`

# Possible cycle through another instruction

Possible cycle:



Possible instance:

VCMPPSZ256rri

K5, YMM31, YMM31, 1

VFMSUBADD213PDZrk

ZMM31, ZMM25, K5, ZMM29, ZMM9

**Keep in mind:**

**This process is fully automated**

# Results

```
> llvm-exegesis -opcode-name IMUL16rri8 -benchmark-mode latency
---
asm_template:
name:          latency IMUL16rri8
cpu_name:      sandybridge
llvm_triple:   x86_64-grteev4-linux-gnu
num_repetitions: 10000
measurements:
  - { key: latency, value: 4.0115, debug_string: '' }
error:        ''
...
```

- Identified [discrepancies](#) between TD files and measurements

# What's next?

- Extend to memory operands
- Automate fixing of TD files
- Measure the effect of
  - immediate:  $\pm 0$ , 1,  $\sim 1$ ,  $2^{8,16,32,64}$ ,  $\pm\infty$ , nan, denorm
  - register values: SUB EAX, EAX, EAX vs SUB EAX, EAX, EBX
- Make it work on other CPUs (ARM under way, Power?)

# Try It Out!

<https://llvm.org/docs/CommandGuide/llvm-exegesis.html>