



NKUER4

时浩铭 严诗慧 林坤 梅骏逸



设计架构

- 使用C++开发
- 前端：基于Flex&Yacc，完成词法分析、语法分析，构建抽象语法树
- 中端：使用与LLVM IR兼容的中间代码。由AST生成load/store形式的SSA，mem2reg后得到phi形式SSA，并在此基础上优化，最终由SSA Destruction得到非SSA形式IR
- 后端：生成arm汇编代码





成员分工

- 梅骏逸：浮点部分实现、窥孔优化
- 林坤：循环部分优化
- 严诗慧：常量部分优化、常数除法取模优化、公共子表达式删除
- 时浩铭：基础框架实现及其他优化



浮点数支持

- 词法分析中加入对浮点字面量和浮点变量声明等的支持
- 浮点与整型的隐式类型转换
- 汇编代码上使用 vadd, vsub 等一系列指令以及浮点数寄存器实现浮点操作
- 浮点数传参时使用 s0-s3, 选择性使用 s4 以保证栈对齐



中间代码优化

- 全局变量局部化
- Mem2Reg
- 函数内联
- 常量传播
- 稀有条件常量传播
- 公共子表达式删除
- 死代码删除
- 全局值编号
- 循环优化
- 指令调度
- 控制流优化



中间代码优化

■ 全局变量局部化

- 对于全局int/float类型变量，转换为函数内的局部变量。
- 在函数内发生函数调用前，store被调函数需要load的全局变量，调用后load被调函数store的全局变量
- 返回前store所有发生修改的全局变量
- 对于全局从未发生store的全局变量，将其视为常数处理。

■ MEM2REG

- 将load/store形式的SSA转换成含有Phi的SSA
- 减少多余load/store，并便于后续优化



中间代码优化

- 常数传播：在编译时计算操作数均为已知常数的表达式
- 拷贝传播：给定变量 **x**, **y**, **y <- x**, **y** 替代后面出现的 **x** 的引用
- 代数化简和重结合：利用代数性质简化表达式
- 稀有条件常数传播：给定变量 **x**, 常数 **c**, **x <- c**, **c** 替代后面出现的 **x** 的引用
- 公共子表达式删除：删除对公共子表达式的重复计算，用保存值来替代

```
b <- a
c <- b + 1
d <- b
```

```
b <- a
c <- a + 1
d <- a
```

```
b <- a + 3
c <- a + 3
```

```
d <- a + 3
b <- d
c <- d
```



中间代码优化

■ 函数内联

- 对于非递归的函数，如果函数参数过多或函数代码较短，则选择将函数内联。
- 减少由于函数调用带来的额外开销。

■ 死代码删除

- 标识所有定义必要值的指令，如部分函数调用、返回语句和store语句等。
- 迭代标识所有对定义必要值有贡献的指令。
- 到达不动点时，未被标记的指令均为死代码，从而进行删除。

■ 全局值编号

- 基于支配的全局值编号，遵循支配树上先根遍历顺序
- 根据指令哈希值判断两指令是否等价并删除冗余指令
- 同时完成了常数表达式的化简与代数恒等式化简



中间代码优化

- 指令调度
 - 调整基本块内指令顺序，尽可能缩短变量的生存期，从而减少寄存器分配中的溢出
- 控制流优化
 - 无用控制流删除
 - 不可到达代码删除
 - 删除常数cmp以及对应的条件判断等
 - 伸直化
 - 合并单前驱单后继基本块
 - 删除只有一句无条件跳转的基本块
 - ...
- 循环倒置
 - 将while循环转换为repeat循环，从而减少循环中跳转语句的执行次数



中间代码优化

■ 循环优化

■ 循环不变代码外提

- 减少循环中重复代码的计算

■ 循环合并

- 针对相邻的简单循环，当循环初值、边界、步长均相同，且循环的指令间不存在反依赖则合并两个循环

■ 循环展开

■ 针对简单循环

- 若循环次数确定，则全部展开；否则以4为单位展开，减少跳转语句，并为冗余删除创造条件



机器码优化

- 死代码删除
- 窥孔优化
- 局部值编号
- 控制流优化
- 乘除法优化
- 寄存器分配



机器码优化

- 死代码删除
 - 基于活跃变量分析完成死代码删除
- 局部值编号
 - 基本块内值编号，主要删除数组相关的冗余
 - 同时完成数组中乘法计算转换为移位
- 控制流优化
 - 单前驱与单后继块的合并
 - 只有一句无条件跳转的基本块的删除
 - 删除顺序连接基本块间的无条件跳转
 - 包含对条件跳转的优化



机器码优化

- 常量整数除法和取模优化
 - 除形如 2^k 的常量时, 可以替换成移位asr
 - 除非 2^k 形式的常量时, 可以替换成smull,asr
 - 模形如 2^k 的常量时, 可以替换成and
 - 模非 2^k 形式的常量时, 可以替换成div, mul, sub
- 寄存器分配
 - 实现了线性扫描和图着色寄存器分配



机器码优化

■ 窥孔优化

- 前后衔接的 (v)mul 和 (v)add/(v)sub 指令合并为 (v)mmla/(v)mmls

<code>mul v0, v1, v2</code>	<code>vmul.f32 v0, v1, v2</code>
<code>add v3, v4, v0</code>	<code>vadd.f32 v3, v4, v0</code>
-----	-----
<code>mmla v3, v1, v2, v4</code>	<code>vmmmla.f32 v4, v1, v2</code>
	<code>vmmov.f32 v3, v4</code>

■ ldr 指令转为 mov

<code>str v355, [v11]</code>	<code>ldr r0, [fp, #-12]</code>
<code>ldr v227, [v11]</code>	<code>ldr r1, [fp, #-12]</code>
-----	-----
<code>str v355, [v11]</code>	<code>ldr r0, [fp, #-12]</code>
<code>mov v227, v355</code>	<code>mov r1, r0</code>

■ 浮点数取负

```
ldr r4, =0
vmov s5, r4
vsub.f32 s7, s5, s6
-----
vneg.f32 s7, s6
```

■ ldr/str 寻址

```
add v11, fp, #-12
ldr v227, [v11]
-----
ldr v227, [fp, #-12]
mov v227, v355
```



测试工具

■ 使用Gitlab CI/CD 搭建自动测试工具，并完成结果数据的网页呈现

NKUER4 > compile > Pipelines

All 62 Pending 0 Running 0 Finished 62 Branches Tags

Run PipelineClear Runner CachesCI Lint

Status	Pipeline	Triggerer	Commit	Stages	
passed	#6074 latest		main 24614072 Merge branch 'main' into 'main'		00:07:41
passed	#6065		main 6d2d7007 Merge branch 'main' into 'main'		
passed	#6040		main 2960a32c remove ...		
passed	#6039		main 26612d8c Merge branch 'loopop' into 'main'		

time	00_bitset1	00_bitset2	00_bitset3	02_mv1	02_mv2	02_mv3	03_sort1	03_sort2	03_sort3	04_spmv1
2099-TEDA	1382333	2771953	4154154	4941648	2099617	3498906	1022954	14346074	2865799	2686449
2022-08-21 17:49:27	389243	778351	1166189	3459830	1308290	2424161	677213	9148730	1701828	2649916
2022-08-21 17:42:14	389066	779939	1166885	3456686	1310922	2420514	675324	8983405	1674823	2658274
2022-08-21 16:52:41	389066	779939	1166885	3457894	1309634	2416745	674265	8948679	1680442	2865281
2022-08-21 15:58:23	389066	779939	1166885	3456686	1310922	2420514	675324	8983405	1674823	2658274
2022-08-21	397172	718939	1071981	3455124	1306971	2416989	594233	7856060	1426783	2657792



决赛成绩

比赛提交到排行榜更新有20秒左右的延迟

详情

#	用户名	队伍	提交次数(ASC)	最后提交时间(ASC)	正确分	性能分	总分
1	Cydiater	赫露艾斯塔/ 清华大学	34	2022-08-21 17:02:36	100	69.7506	76.4720
2	ksitta	啊对对队/ 清华大学	52	2022-08-21 17:30:09	100	67.6604	74.8462
3	18373046	喵喵队仰卧起坐/ 北京航空航天大学	43	2022-08-21 09:56:44	100	67.3488	74.6039
4	HammerWang	从容应队/ 西北工业大学	19	2022-08-21 16:08:01	100	50.3050	61.3472
5	18377149	编不出来不起床/ 北京航空航天大学	44	2022-08-21 16:40:03	100	49.9070	61.0377
6	1911463	NKUER4/ 南开大学	37	2022-08-21 15:50:29	100	44.4575	56.7990
7	RaVincent	嘉然今天偷着乐/ 国防科技大学	25	2022-08-21 17:31:12	100	43.2724	55.8773
8	Origami404	萝杨空队/ 哈尔滨工业大学 (深圳)	29	2022-08-21 17:32:13	100	39.6914	53.0920
9	ChenLD3124	邯郸路企鹅编译器/ 复旦大学	48	2022-08-21 12:10:45	100	38.5712	52.2207
10	CoolColoury	LoveLive!Virtual Magic!/ 北京航空航天大学	15	2022-08-21 17:03:13	100	37.3946	51.3055



问答环节

感谢评委老师指导

