HPC 大作业

刘涛

December 23, 2024

1 第一题

1.1 编译CloverLeaf

1.Serial 版本的CloverLeaf 观察Serial文件下的README.MD文件可以知道使用

make COMPILER=GNU MPI_COMPILER=gfortran C_MPI_COMPILER=gcc编译运行指令,如下所示:

```
advec_cell_kernel_c.o \
calc_dt_kernel_c.o \
field_summary_kernel_c.o \
update_halo_kernel_c.o \
timer_c.o \
pack_kernel_c.o \
generate_chunk_kernel_c.o \
initialise_chunk_kernel_c.o \
-o clover_leaf; echo
```

Figure 1: Serial 版本编译结果

直接运行的结果, 如下所示:

```
Wall clock 11.440356016159058

Average time per cell 1.4434396217093448E-007

Step time per cell 1.4226133417752053E-007

Step 87 time 0.4971244 control sound timestep 5.85E-03 1

, 1 x 5.21E-03 y 5.21E-03

Test problem 2 is within 0.1136868E-12% of the expected solution

This test is considered PASSED

Wall clock 11.578369140625000

First step overhead -3.1728744506835938E-003
```

Figure 2: Serial 版本运行结果

2.MPI 版本的CloverLeaf 同理,观察README.MD文件,由于我使用的是GNU编译器,可知采用

make COMPILER=GNU MPLCOMPILER=mpifort C_MPLCOMPILER=mpicc DEBUG=1 IEEE=1

编译运行指令,如下所示:

```
Warning: 'second_step' may be used uninitialized in this function [-Wmaybe-uninitialized]
hydro.f90:96:76: Warning: 'first_step' may be used uninitialized
d in this function [-Wmaybe-uninitialized]
visit.f90:69:77:

69 | ler%ideal_gas=profiler%ideal_gas+(timer()-kernel_time)

Warning: 'kernel_time' may be used uninitialized in this function [-Wmaybe-uninitialized]

(base) ubuntu@ubuntu-Precision-7920-Tower:~/github/learn/homework
```

Figure 3: MPI 版本编译结果

直接运行的结果,如图片4所示:

```
0.4912765 control
         86 time
                                            timestep
           1, 1 x 5.21E
10.945327043533325
.85E-03
                   1 x 5.21E-03 y 5.21E-03
Wall clock
Average time per cell
                     1.3809813877197044E-007
Step time per cell
                     1.3361762588222821E-007
Step
        87 time
                 0.4971244 control sound
                                           timestep
ution
This test is considered PASSED
            11.077787876129150
Wall clock
First step overhead -1.5370845794677734E-003
```

Figure 4: MPI 版本运行结果

3. OPENMP版本的ColverLeaf

继续观察openmp的CloverLeafOpenmp下README.md,同样得到采用相同的编译指令运行OPENMP

make COMPILER=GNU MPLCOMPILER=mpifort C_MPLCOMPILER=mpicc DEBUG=1 IEEE=1 编译运行指令,如图片5所示:

Figure 5: OPENMP 版本编译结果

直接运行的界面,如图片6所示

```
eally needed).

Clover Version 1.300
   MPI Version
   OpenMP Version
   Task Count 1
   Thread Count: 4

Clover Version 1.300
   MPI Version
   OpenMP Version
   Task Count 1
Thread Count: 4
```

Figure 6: OPENMP 版本运行

运行结果,如图片7所示

```
.85E-03 1, 1 x 5.2IE-03 y 5.2IE-03

Test problem 2 is within 0.4376943E-11% of the expected sol ution

This test is considered PASSED

Wall clock 44.396925926208496

First step overhead -9.5629692077636719E-003
(base) ubuntu@ubuntu-Precision-7920-Tower:~/github/learn/homewo
```

Figure 7: OPENMP 版本运行结果

4. MPI+OPENMP版本的ColverLeaf 到这里,结合已经学习的知识和做过的实验,不难验证编译指令如下所示。

make COMPILER=GNU MPLCOMPILER=mpifort C_MPLCOMPILER=mpicc DEBUG=1 IEEE=1

OMP_NUM_THREAD=4 mpirun -n 4 ./ clover_leaf 测试结果如图片8所示:

```
Step time per cell 5.6258351024654174E-008
Step 87 time 0.1242811 control sound timestep 1
.46E-03 1, 1 x 1.30E-03 y 1.30E-03
Test problem 4 is within 0.1124079E-10% of the expected sol ution
This test is considered PASSED
Wall clock 74.375893831253052
First step overhead -1.2004852294921875E-002
```

Figure 8: MPI+openmp 版本运行结果

1.2 性能分析

性能分析工具采用intel OneAPI vtune进行处理,这是一款性能优良适合分析的工具。

1.Serial 版本的CloverLeaf性能分析 使用性能分析参数

OMPNUM_THREAD=1 mpirun -n 1 ./clover_leaf

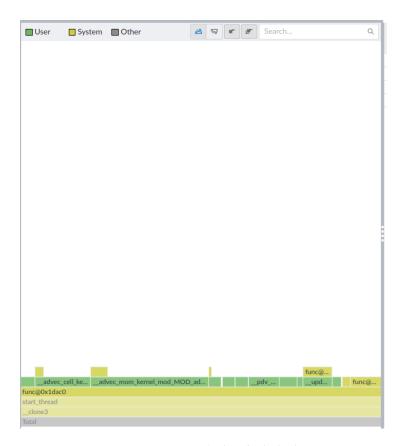


Figure 9: Serial 版本运行火焰图

采用Bottomup的分析模式可以知道程序中的热点位置在于

Function / Call Stack	CPU Time ▼ 🕒	Module
advec_mom_kernel_mod_MOD_advec_mom_kernelomp_fn.0	1586.649s	clover_leaf
▶ func@0x20760	1397.037s	libgomp.so.1
advec_cell_kernel_module_MOD_advec_cell_kernelomp_fn.0	739.200s	clover_leaf
pdv_kernel_module_MOD_pdv_kernelomp_fn.0	483.194s	clover_leaf
reset_field_kernel_module_MOD_reset_field_kernelomp_fn.0	265.736s	clover_leaf
accelerate_kernel_module_MOD_accelerate_kernelomp_fn.0	236.022s	clover_leaf
ideal_gas_kernel_module_MOD_ideal_gas_kernelomp_fn.0	205.320s	clover_leaf
▶flux_calc_kernel_module_MOD_flux_calc_kernelomp_fn.0	196.240s	clover_leaf
calc_dt_kernel_module_MOD_calc_dt_kernelomp_fn.0	150.438s	clover_leaf
viscosity_kernel_module_MOD_viscosity_kernelomp_fn.0	140.019s	clover_leaf
▶ func@0x205c0	124.816s	libgomp.so.1
revert_kernel_module_MOD_revert_kernelomp_fn.0	90.899s	clover_leaf
field_summary_kernel_module_MOD_field_summary_kernelomp_fn.0	16.254s	clover_leaf
▶ build_fieldomp_fn.0	15.038s	clover_leaf
▶ func@0x1f230	5.326s	libgomp.so.1
update_halo_kernel_module_MOD_update_halo_kernelomp_fn.0	3.890s	clover_leaf
generate_chunk_kernel_module_MOD_generate_chunk_kernelomp_fn.0	2.685s	clover_leaf
▶ GOMP_parallel	2.356s	libgomp.so.1
initialise_chunk_kernel_module_MOD_initialise_chunk_kernelomp_fn.0	1.133s	clover_leaf
▶ memset	0.511s	libc-dynamic.so
▶libc_start_main_impl	0.440s	libc.so.6
▶ GOMP_barrier	0.377s	libgomp.so.1
▶ PMPI_Init_f08	0.330s	libmpi_mpifh.so.4
▶ func@0x208e0	0.317s	libgomp.so.1
▶ OS_BARESYSCALL_DoCallAsmIntel64Linux	0.164s	libc-dynamic.so
▶ memcmp	0.160s	libc-dynamic.so
▶ memmove	0.160s	libc-dynamic.so
▶ _GLOBALsub_l_acl_platform.cpp	0.160s	libalteracl.so
▶ func@0x20910	0.149s	libgomp.so.1
▶ operator new	0.105s	libc++abi.so
cxa_demangle	0.090s	libc++abi.so
▶ func@0x1dac0	0.069s	libgomp.so.1

Figure 10: Serial 版本运行热点图

由于Serial版本的代码不支持重定位到具体的代码细节,详细的分析在后续写下.

2.MPI 版本的CloverLeaf性能分析 采用如下版本进行分析

OMP_NUM_THREAD=1 mpirun $-n\ 4\ ./\ {\tt clover_leaf}$

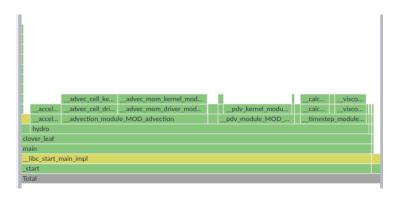


Figure 11: MPI 版本运行火焰图

采用Bottomup的分析模式可以知道程序中的热点位置在于

Function / Call Stack	CPU Time ▼ 🍱	Modu
advec_mom_kernel_mod_MOD_advec_mom_kernel	9.921s	clover_leaf
pdv_kernel_module_MOD_pdv_kernel	7.561s	clover_leaf
advec_cell_kernel_module_MOD_advec_cell_kernel	6.260s	clover_leaf
viscosity_kernel_module_MOD_viscosity_kernel	3.470s	clover_leaf
accelerate_kernel_module_MOD_accelerate_kernel	3.429s	clover_leaf
calc_dt_kernel_module_MOD_calc_dt_kernel	3.210s	clover_leaf
ideal_gas_kernel_module_MOD_ideal_gas_kernel	1.280s	clover_leaf
flux_calc_kernel_module_MOD_flux_calc_kernel	1.119s	clover_leaf
reset_field_kernel_module_MOD_reset_field_kernel	0.650s	clover_leaf
memset	0.482s	libc-dynamic
revert_kernel_module_MOD_revert_kernel	0.300s	clover_leaf
libc_start_main_impl	0.299s	libc.so.6
▶ PMPI_Init_f08	0.298s	libmpi_mpifh
field_summary_kernel_module_MOD_field_summary_kernel	0.270s	clover_leaf
▶ build_field	0.220s	clover_leaf
operator new	0.190s	libc++abi.so
OS_BARESYSCALL_DoCallAsmIntel64Linux	0.162s	libc-dynamic
■ _GLOBALsub_I_acl_platform.cpp	0.120s	libalteracl.so
memmove	0.120s	libc-dynamic
memcmp	0.110s	libc-dynamic
update_halo_kernel_module_MOD_update_halo_kernel	0.080s	clover_leaf
cxa_demangle	0.060s	libc++abi.so
Unknown stack frame(s)]	0.040s	
▶ strlen	0.040s	libc-dynamic
▶ tbb::detail::r1::TBB_InitOnce::TBB_InitOnce	0.029s	libtbb.so.12
memset_evex_unaligned_erms	0.020s	libc.so.6
svfscanf	0.010s	libc-dynamic
▶ AllocateNewTlsPerThreadFunc	0.010s	libc-dynamic
▶ func@0xe4a0	0.010s	libc++abi.so
initialise_chunk_kernel_module_MOD_initialise_chunk_kernel	0.010s	clover_leaf
_dwarf_skip_leb128	0.010s	libdwarf.so
▶tcf_0	0.010s	libalteracl.so
	0.010-	115-11-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-1-

Figure 12: MPI 版本运行热点图

对于在Bottomup图中的消耗CPU时间最多的函数,其代码如下所示



Figure 13: 源码所在函数

0x5311f		Block 364:	
0x5311f	238	mov %r15, %rcx	
0x53122	238	lea 0x2a997(%rip), %rsi	
0x53129	238	lea 0x2ad90(%rip), %rdi	
0x53130	238	mov \$0x0, %eax	
0x53135	238	callq 0x2250	
0x5313a		Block 365:	
0x5313a	238	movq 0x40(%rsp), %rcx	
0x5313f	238	lea 0x2a9f2(%rip), %rsi	
0x53146	238	lea 0x2ad73(%rip), %rdi	
0x5314d	238	mov \$0x0, %eax	
0x53152	238	callq 0x2250	
0x53157		Block 366:	
0x53157	238	mov %rax, %rcx	
0x5315a	238	lea 0x2a83f(%rip), %rsi	
0x53161	238	lea 0x2ad58(%rip), %rdi	
0x53168	238	mov \$0x0, %eax	
0x5316d	238	callq 0x2250	
0x53172		Block 367:	
0x53172	238	mov %r15, %rcx	
0x53175	238	mov %rdi, %rdx	
0x53178	238	lea 0x2ad71(%rip), %rsi	
0x5317f	238	lea 0x2ad3a(%rip), %rdi	
0x53186	238	mov \$0x0, %eax	
0x5318b	238	<u>callq 0x2250</u>	
0x53190		Block 368:	
0x53190	238	movq 0x40(%rsp), %rcx	
0x53195	238	mov %rdi, %rdx	
0x53198	238	lea 0x2ad99(%rip), %rsi	
0x5319f	238	lea 0x2ad1a(%rip), %rdi	
0x531a6	238	mov \$0x0, %eax	
0x531ab	238	<u>callq 0x2250</u>	

Figure 14: 反汇编图片

3.OPENMP 版本的CloverLeaf性能分析 OMP_NUM_THREAD=4 mpirun -n 1 ./clover_leaf



Figure 15: OPENMP 版本运行火焰图

采用Bottomup的分析模式可以知道程序中的热点位置在于

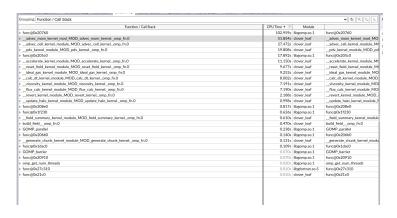


Figure 16: OPENMP 版本运行热点图

此时再进行分析可知出现了系统函数调用func@0x20760,这说明在这个位置调用了OPENMP的系统函数, 反汇编可知这部分的操作没有对应在程序中的代码, 在openmp动态链接库文件中被调用,观察其内容,可能 是对应的资源访问互斥锁。

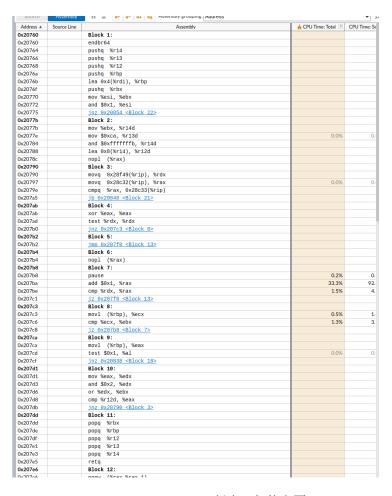


Figure 17: OPENMP 版本运行热点图

4.MPI+OPENMP 版本的CloverLeaf性能分析

OMPNUM_THREAD=4 mpirun -n 4 ./clover_leaf



Figure 18: OPENMP 版本运行热点图

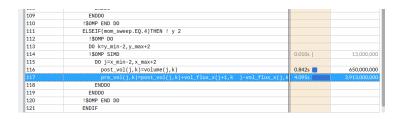


Figure 19: OPENMP+MPI运行反汇编程序

1.3 效率对比

1.不同MPI+OPENMP的对比 x_cell=960,y_cell=960

OMP_NUM_THREADS	MPI ranks	time consumption
1	1	47.35715s
1	2	24.93825s
1	4	12.38018s
2	1	43.900086s
2	2	22.80418s
2	4	6.75957
4	1	44.56657s
4	2	23.21618s
4	4	4.126149s

Table 1: ref下不同OMP_NUM_THREADS and MPI 的时间消耗

 $x_cell{=}1920, y_cell{=}1920$

OMP_NUM_THREADS	MPI ranks	time consumption
1	1	188.6974799s
1	2	96.022246s
1	4	49.126179s
2	2	22.80418s
2	4	26.75957
4	4	19.46512s

Table 2: ref下不同OMP_NUM_THREADS and MPI 的时间消耗

2.不同MPI的对比 xcell=960,ycell=960

OMP_NUM_THREADS	MPI ranks	time consumption
1	1	37.21615s
1	2	19.731973s
1	4	9.932440s
1	8	5.461988s
1	16	3.7412819s

Table 3: ref下不同OMP_NUM_THREADS and MPI 的时间消耗

3.OPEMP版本的对比 xcell=960,ycell=960

OMP_NUM_THREADS	MPI ranks	time consumption
1	1	46.32359s
2	1	24.176760s
4	1	13.139265s
8	1	7.1625220s
16	1	4.6675159s

Table 4: OPENMP下不同OMP_NUM_THREADS and MPI 的时间消耗

1.4 CUDA 编译

由于CUDA版本的编译需要适配,做如下的修改,使用对应的GNU编译器

Figure 20: 编译命令修改

由于我的fortran版本的是gcc11对应的fortran版本,所以使用对应的编译器会出现早期编译器对浮点数和int类型坚持不严格,而在新编译器上的检查严格造成的错误.

```
FLAGS=$(FLAGS_$(COMPILER)) $(OMP) $(I3E) $(OPTIONS) -fallow-argument-mismatch

CFLAGS=$(CFLAGS_$(COMPILER)) $(OMP) $(I3E) $(C_OPTIONS) -c

MPI_COMPILER=mpif90

C_MPI_COMPILER=mpicc

NV_FLAGS+=-D MANUALLY_CHOOSE_GPU
```

Figure 21: 编译命令修改

Figure 22: 编译命令修改

由于我的电脑是Ampere架构的显卡,修改Makefile文件获得计算结果

```
CODE_GEN_MAXWELL=-gencode arch=compute_5
CODE_GEN_PASCAL=-gencode arch=compute_60
CODE_GEN_AMPERE=-gencode arch=compute_86
CUDA_LIB_PATH=/usr/local/cuda-12.4/lib64
LDLIBS+=-lstdc++ -lcudart
```

Figure 23: 修改为Ampere架构

CUDA是12.4类型的CUDA,所以修改对应的CUDA编译器

Figure 24: 指定CUDA和NVCC位置

采用如下的编译指令

make COMPILER=GNU MPLCOMPILER=mpifort C_MPLCOMPILER=mpicc DEBUG=1 IEEE=1 成功编译对应的程序,其结果如下所示

apfort. Ob. -fopomp. Indiana-grament-assanth pack-kernel. Colorer data definitions. Or report o timer o parse o read imput. O initialise think kernel. O initialise chunk be build field. O update halo kernel. O generate chunk control to the color of the definitions of the color of the color

Figure 25: 编译完成

测试结果,如下所示:

```
Step Lime per cell 8.6091717408170007E-008
Step 87 time 0.4091244 control sound timestep 5.85E-03 1, 1 x 5.21E-03 y 5.21E-03
Test problem 2 is within a .082294E-07% of the expected solution
This test is considered PASSED
Wall clock 7.439551838720703
First step overhead 9.3552794647216797E-003
```

Figure 26: 测试通过

1.5 CUDA版本程序测试

GPU=RTX 3090 CPU=gold 6132 Mpi_num=1 openmp_num=1 可以得出对比

x_cell	y_cell	time consumption
960	960	6.87124s
1920	960	22.105352s
3840	960	38.15225s
1920	1920	27.509817s
3840	1920	54.6727240s
3840	3840	114.67717s

Table 5: CUDA 下不同网格大小的时间消耗

我的GPU和CPU,3090GPU约为15个左右的CPU加速效果

1.6 CUDA代码技巧分析