MAQAO
Hands-on exercises

Profiling bt-mz (incl. scalability)
Optimising a code

Setup (reminder)

Login to the cluster

```
> ssh <username>@login.devana.nscc.sk
```

Copy handson material to your workspace directory

```
> export MAQAO_TUTO=/home/projects/training-03/MAQAO
> export WORK=$HOME
Hint: copy in ~/.bash_profile or ~/.bashrc
> cd $WORK
> tar xvf $MAQAO_TUTO/MAQAO_HANDSON.tgz
> tar xvf $MAQAO_TUTO/NPB3.4-MZ-MPI.tgz
```

Load MAQAO environment

```
> module load maqao/2.17.4
```

Setup (bt-mz compilation with Intel compiler and MPI & debug symbols)

Go to the NPB directory provided with MAQAO handsons

```
> cd $WORK/NPB3.4-MZ-MPI
```

Load Intel compiler and environment

```
> module load impi
```

Compile and run

```
> make bt-mz CLASS=C
> cd bin
> cp $WORK/MAQAO_HANDSON/bt/bt.sbatch .
> sbatch bt.sbatch
```

Remark: with version 3.4 the generated executable supports any number of ranks (no need to generate one executable for 6 ranks, another for 8 etc.)

Profiling bt-mz with MAQAO

Setup ONE View for batch mode

The ONE View configuration file must contain all variables for executing the application. Retrieve the configuration file prepared for bt-mz in batch mode from the MAQAO_HANDSON directory

```
> cd $WORK/NPB3.4-MZ-MPI/bin
> cp $WORK/MAQAO_HANDSON/bt/bt_OV_sbatch.lua .
> less bt_OV_sbatch.lua
```

```
executable = "bt-mz.C.x"
...
batch_script = "bt_maqao.sbatch"
batch_command = "sbatch <batch_script>"
...
number_processes = 4
number_processes_per_node = 2
envv_OMP_NUM_THREADS = 16
...
mpi_command = "mpirun -n <number_processes>"
```

Review jobscript for use with ONE View

All variables in the jobscript defined in the configuration file must be replaced with their name from it.
Retrieve jobscript modified for ONE View from the MAQAO_HANDSON directory.

```
> cd $WORK/NPB3.4-MZ-MPI/bin #if current directory has changed
> cp $WORK/MAQAO_HANDSON/bt/bt_maqao.sbatch .
> less bt_maqao.sbatch
```

```
...
#SBATCH --ntasks-per-node=2<number_processes_per_node>
#SBATCH --cpus-per-task=16<OMP_NUM_THREADS>
...
export OMP_NUM_THREADS=16<OMP_NUM_THREADS>
...
mpirun -n ... $EXE
<mpi_command> <run_command>
...
```

Launch MAQAO ONE View on bt-mz (batch mode)

Launch ONE View

```
> cd $WORK/NPB3.4-MZ-MPI/bin #if current directory has changed
> maqao oneview -R1 --config=bt_OV_sbatch.lua -xp=ov_sbatch
```

The -xp parameter allows to set the path to the experiment directory, where ONE View stores the analysis results and where the reports will be generated.
If -xp is omitted, the experiment directory will be named maqao_<timestamp>.

WARNINGS:

- If the directory specified with -xp already exists, ONE View will reuse its content but not overwrite it.

Display MAQAO ONE View results

The HTML files are located in **<exp-dir>/RESULTS/<binary>_one_html**, where *<exp-dir>* is the path of he experiment directory (set with -xp) and *<binary>* the name of the executable.

To read them:

- Mount remote directory locally (requires sshfs installed locally)
- Compress and download the HTML directory to open locally
- Open the file remotely (if display redirect available)
- *Open a Jupyter Notebook Interactive App*

Display MAQAO ONE View results

- Mounting $WORK locally:

```
> mkdir devana_work
> sshfs <user>@login.devana.nscc.sk:\
/home/<user> devana_work
> firefox devana_work/NPB3.4-MZ-MPI/bin/ov_sbatch\
/RESULTS/bt-mz.C.x_one_html/index.html
```

- Compressing and downloading the results:

```
> tar czf $HOME/bt_html.tgz ov_sbatch/RESULTS/bt-mz.C.x_one_html
```

```
> scp <user>@login.devana.nscc.sk:bt_html.tgz .
> tar xf bt_html.tgz
> firefox ov_sbatch/RESULTS/bt-mz.C.x_one_html/index.html
```

- Use the Open OnDemand console to create a Jupyter Notebook Interactive session and open index.html from the file browser

sshfs & scp hints

- To install sshfs on Debian-based Linux distributions (like Ubuntu)

```
> sudo apt install sshfs
```

- Recommended to close a sshfs directory after use

```
> fusermount -u /path/to/sshfs/directory
```

- scp is slow to copy directories (especially when containing many small files), copy a .tgz archive of the directory

Display MAQAO ONE View results

Results can also be viewed directly on the console in text mode:

```
> maqao oneview -R1 -xp=ov_sbatch --output-format=text
```

Sample result directories are available in

/home/projects/training-03/MAQAO/MAQAO_offline.tgz

Scalability profiling of bt-mz with MAQAO

Setup ONE View for scalability analysis

Retrieve the configuration file prepared for bt-mz in batch mode from the MAQAO_HANDSON directory

```
> cd $WORK/NPB3.4-MZ-MPI/bin #if cur. dir. has changed
> cp $WORK/MAQAO_HANDSON/bt/bt_OV_scal.lua .
> less bt_OV_scal.lua
```

```
executable = "./bt-mz.C.x"
...
run_command = "<executable>"
...
batch_script = "bt_maqao.sbatch"
...
batch_command = "sbatch <batch_script>"
...
number_processes = 4
...
number_processes_per_node = 4
...
omp_num_threads = 1
...
mpi_command = "mpirun -n <number_processes>"
...
multiruns_params = {
 {number_processes = 1, envv_OMP_NUM_THREADS = 8, number_nodes = 1, number_processes_per_node = 1},
 {number_processes = 4, envv_OMP_NUM_THREADS = 1, number_nodes = 2, number_processes_per_node = 2},
 {number_processes = 4, envv_OMP_NUM_THREADS = 8, number_nodes = 2, number_processes_per_node = 2},
}
scalability_reference = "lowest-threads"
```

Launch MAQAO ONE View on bt-mz (scalability mode)

Launch ONE View (execution will be longer!)

```
> maqao oneview -R1 --with-scalability \
-c=bt_OV_scal.lua -xp=ov_scal
```

The results can then be accessed similarly to the analysis report.

```
> firefox devana_work/NPB3.4-MZ-MPI/bin/ov_scal/RESULTS/\
bt-mz.C.x_one_html/index.html
```

OR

```
> tar czf $HOME/bt_scal.tgz \
ov_scal/RESULTS/bt-mz.C.x_one_html
```

```
> scp <user>@login.devana.nscc.sk:ov_scal.tgz .
> tar xf ov_scal.tgz
> firefox ov_scal/RESULTS/bt-mz.C.x_one_html/index.html
```

OR

Use the Jupyter Notebook Interactive App

Optimising a code with MAQAO

Matrix Multiply code

```
void kernel0 (int n,
              float a[n][n],
              float b[n][n],
              float c[n][n]) {
  int i, j, k;

  for (i=0; i<n; i++)
    for (j=0; j<n; j++) {
      c[i][j] = 0.0f;
      for (k=0; k<n; k++)
        c[i][j] += a[i][k] * b[k][j];
    }
}
```

"Naïve" dense matrix multiply implementation in C

Compile with GNU compiler

Go to the handson directory

```
> cd $WORK/MAQAO_HANDSON/matmul
```

Compile all variants

```
> module load gcc/12.2
> make all
```

Load MAQAO environment (if necessary)

```
> module load maqao/2.17.4
```

Analysing matrix multiply with MAQAO

Parameters are: <size of matrix> <number of repetitions>

```
> cd $WORK/MAQAO_HANDSON/matmul #if cur. directory has changed
> srun -p ncpu --exclusive -t 1 ./matmul_orig/matmul 400 300
# 400x400 matrix, 300 repetitions
cycles per FMA: 2.77
```

Analyse matrix multiply with ONE View

```
> maqao oneview -R1 -c=ov_orig.lua -xp=ov_orig
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/matmul/ov_orig/RESULTS/\ matmul_orig_one_html/index.html`**

Global Metrics

Total Time (s)		39.01
Profiled Time (s)		39.00
Time in analyzed loops (%)		100.0
Time in analyzed innermost loops (%)		99.9
Time in user code (%)		100
Compilation Options Score (%)		50.0
Perfect Flow Complexity		1.00
Array Access Efficiency (%)		83.3
Perfect OpenMP + MPI + Pthread		1.00
Perfect OpenMP + MPI + Pthread + Perfect Load Distribution		1.00
No Scalar Integer	Potential Speedup	1.00
	Nb Loops to get 80%	1
FP Vectorised	Potential Speedup	2.81
	Nb Loops to get 80%	1
Fully Vectorised	Potential Speedup	16.0
	Nb Loops to get 80%	1
FP Arithmetic Only	Potential Speedup	1.00
	Nb Loops to get 80%	1

CQA output for the baseline kernel

Vectorization

Your loop is not vectorized. 8 data elements could be processed at once in vector registers. By vectorizing your loop, you can lower the cost of an iteration from 3.00 to 0.37 cycles (8.00x speedup).

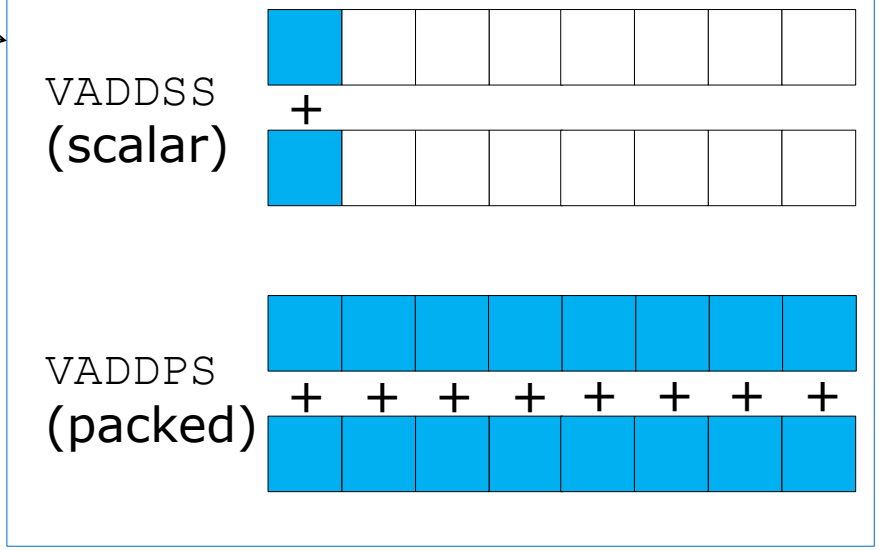
Details

All SSE/AVX instructions are used in scalar version (process only one data element in vector registers). Since your execution units are vector units, only a vectorized loop can use their full power.

Workaround

- Try another compiler or update/tune your current one:
  - recompile with fassociative-math (included in Ofast or ffast-math) to extend loop vectorization to FP reductions.
- Remove inter-iterations dependences from your loop and make it unit-stride:
  - If your arrays have 2 or more dimensions, check whether elements are accessed contiguously and, otherwise, try to permute loops accordingly: C storage order is row-major: for(i) for(j) a[j][i] = b[j][i]; (slow, non stride 1) => for(i) for(j) a[i][j] = b[i][j]; (fast, stride 1)
  - If your loop streams arrays of structures (AoS), try to use structures of arrays instead (SoA): for(i) a[i].x = b[i].x; (slow, non stride 1) => for(i) a.x[i] = b.x[i]; (fast, stride 1)

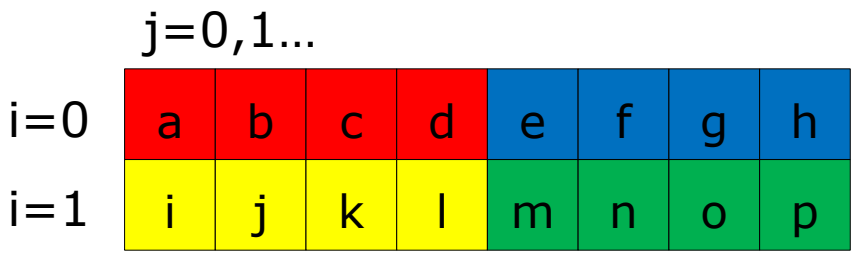
Vectorization (summing elements):



- Accesses are not contiguous => let's permute k and j loops
- No structures here…

Impact of loop permutation on data access

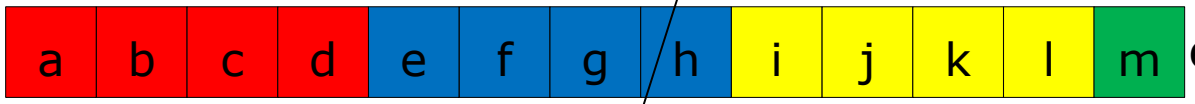
Logical mapping



Efficient vectorization + prefetching

Physical mapping

(C stor. order: row-major)




```
for (j=0; j<n; j++)
  for (i=0; i<n; i++)
    f(a[i][j]);
```


```
for (i=0; i<n; i++)
  for (j=0; j<n; j++)
    f(a[i][j]);
```

Removing inter-iteration dependences and getting stride 1 by permuting loops on j and k

```
void kernel1 (int n,
              float a[n][n],
              float b[n][n],
              float c[n][n]) {
  int i, j, k;

  for (i=0; i<n; i++) {
    for (j=0; j<n; j++)
      c[i][j] = 0.0f;

    for (k=0; k<n; k++)
      for (j=0; j<n; j++)
        c[i][j] += a[i][k] * b[k][j];
  }
}
```

Analyse matrix multiply with permuted loops

Run permuted loops version of matrix multiply

```
> cd $WORK/MAQAO_HANDSON/matmul #if cur. directory has changed
> srun -p ncpu --exclusive -t 1 ./matmul_perm/matmul 400 300
cycles per FMA: 0.40
```

Analyse matrix multiply with ONE View

```
> maqao oneview -R1 -c=ov_perm.lua -xp=ov_perm
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/matmul/ov_perm/RESULTS/\ matmul_perm_one_html/index.html`**

Global Metrics		
Total Time (s)		5.71
Profiled Time (s)		5.70
Time in analyzed loops (%)		99.8
Time in analyzed innermost loops (%)		98.3
Time in user code (%)		99.8
Compilation Options Score (%)		50.0
Perfect Flow Complexity		1.00
Array Access Efficiency (%)		100
Perfect OpenMP + MPI + Pthread		1.00
Perfect OpenMP + MPI + Pthread + Perfect Load Distribution		1.00
No Scalar Integer	Potential Speedup	1.01
	Nb Loops to get 80%	1
FP Vectorised	Potential Speedup	1.41
	Nb Loops to get 80%	1
Fully Vectorised	Potential Speedup	4.02
	Nb Loops to get 80%	1
FP Arithmetic Only	Potential Speedup	1.41
	Nb Loops to get 80%	1

Faster (was 39.01)

More efficient vectorization (was 16.00)

CQA output after loop permutation

gain | potential | hint | expert

Vectorization

Your loop is vectorized, but using only 128 out of 512 bits (SSE/AVX-128 instructions on AVX-512 processors). By fully vectorizing your loop, you can lower the cost of an iteration from 1.40 to 0.35 cycles (4.00x speedup).

Details

All SSE/AVX instructions are used in vector version (process two or more data elements in vector registers). Since your execution units are vector units, only a fully vectorized loop can use their full power.

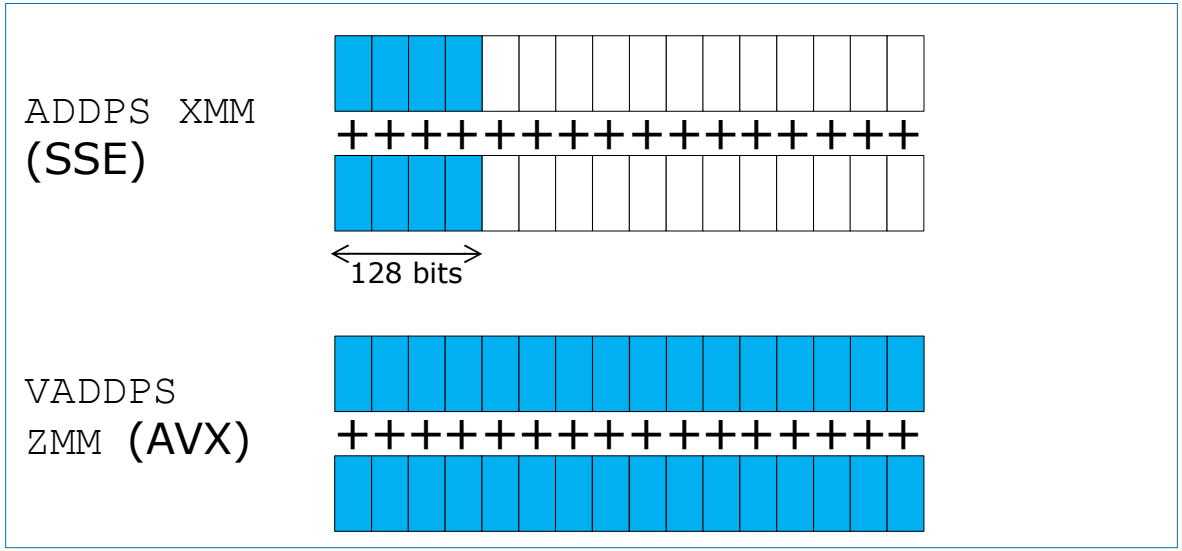
Workaround

- Recompile with march=icelake-server. CQA targ… eon(R) Icelake SP) but specialization flags are -march=x86-64
- Use vector aligned instructions:
  1. align your arrays on 64 bytes boundaries
  2. inform your compiler that your arrays are vector aligned: if array 'foo' is 64 bytes-aligned, define a pointer 'p_foo' as __builtin_assume_aligned (foo, 64) and use it instead of 'foo' in the loop.

Let's try this

Impacts of architecture specialization: vectorization

- Vectorization
  - SSE instructions (SIMD 128 bits) used on a processor supporting AVX512 ones (SIMD 512 bits)
  - => 75% efficiency loss



Analyse matrix multiply with microarchitecture-specialization and array alignment

Run array-aligned version of matrix multiply

```
> cd $WORK/MAQAO_HANDSON/matmul #if cur. directory has changed
> srun -p ncpu --exclusive -t 1 ./matmul_align/matmul 400 300
# remark: size%8 has to equal 0
driver.c: Using posix_memalign instead of malloc
cycles per FMA: 0.23
```

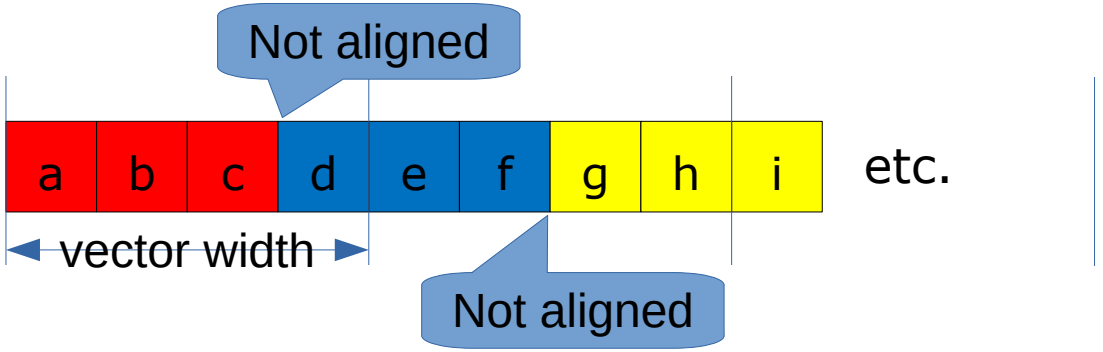
Analyse matrix multiply with ONE View

```
> maqao oneview -R1 -c=ov_align.lua -xp=ov_align
```

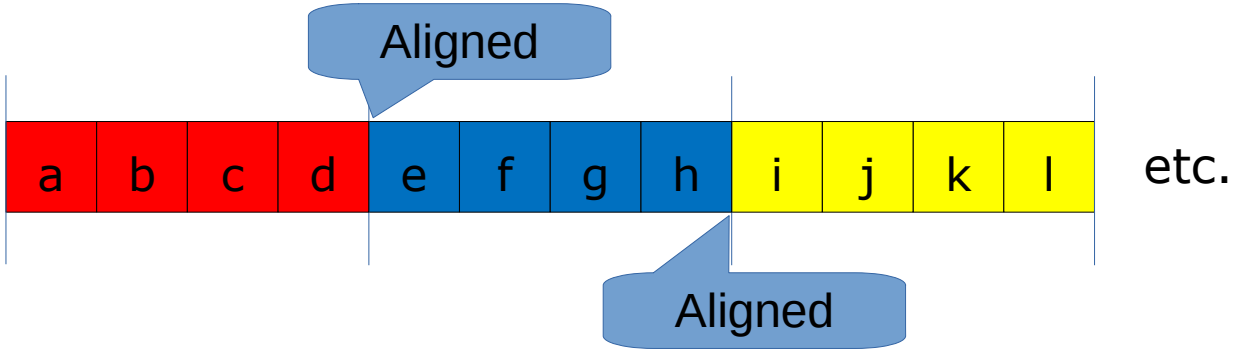
Multidimensional array alignment

Data organized as a 2D array: n lines of 3 columns
Each vector can hold 4 consecutive elements

a[n][3], only 1st element is aligned

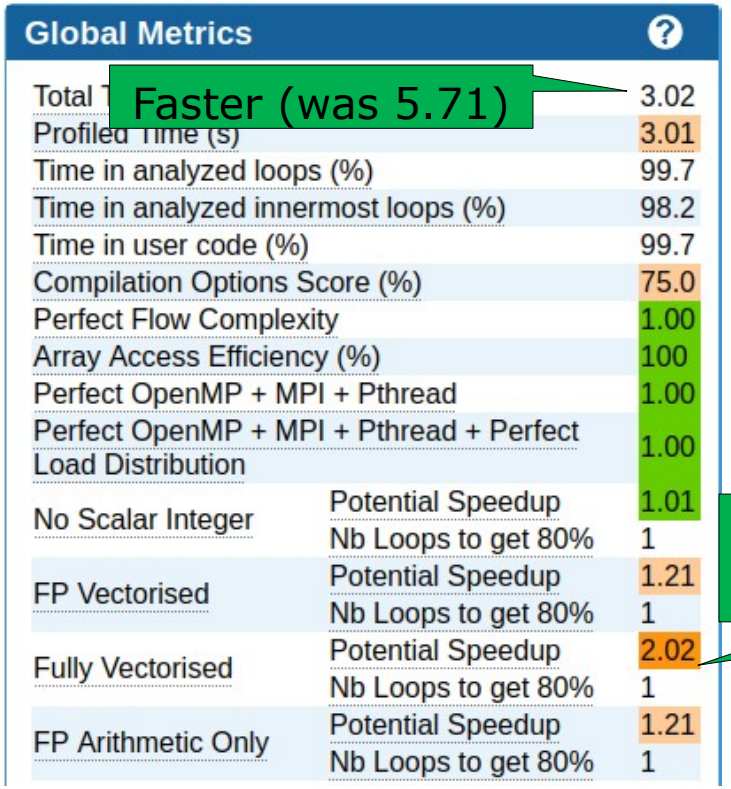


a[n][4], 1st element of each line are aligned

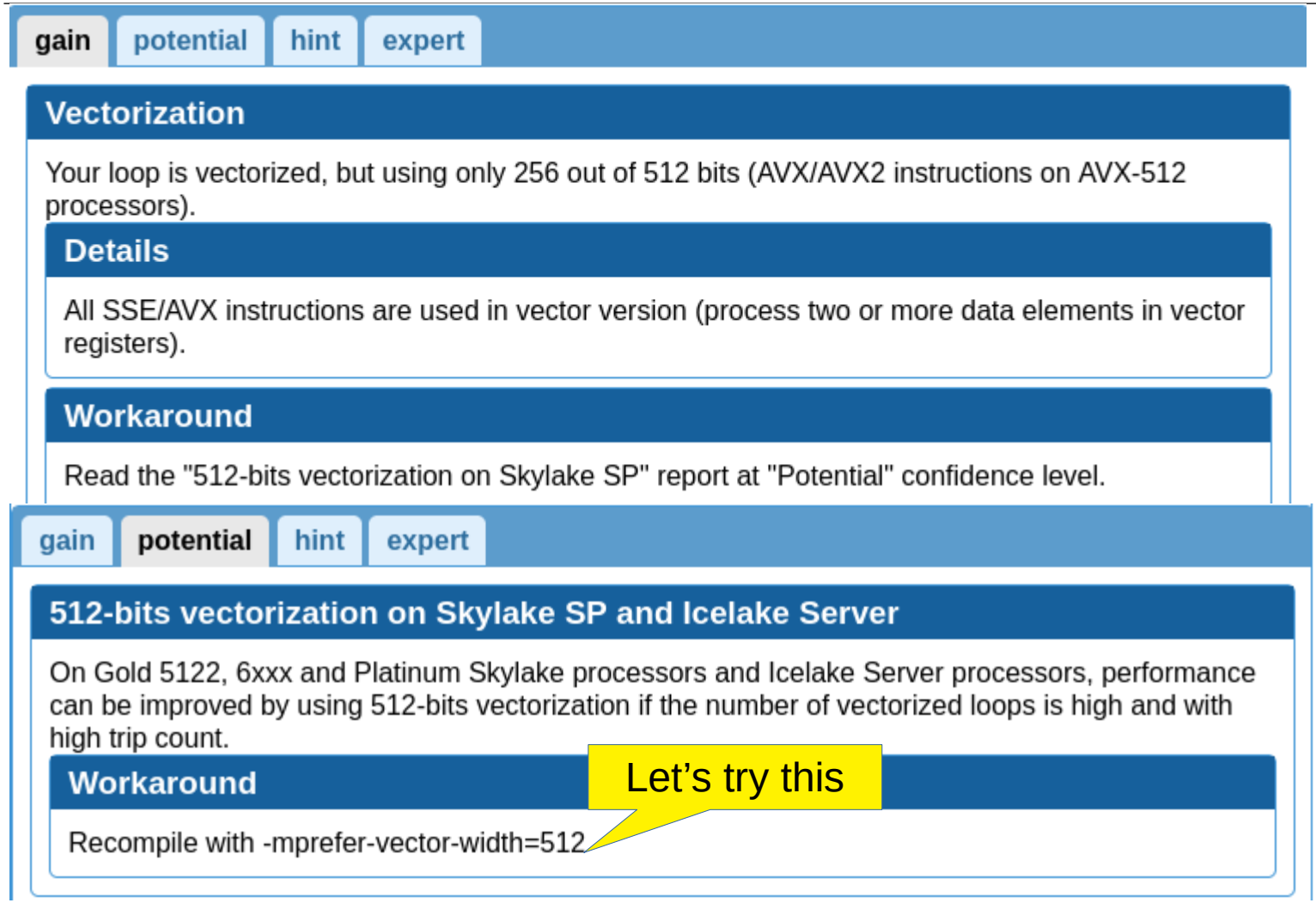


Viewing results (HTML)

Open file **`MAQAO_HANDSON/matmul/ov_align/RESULTS/\ matmul_align_one_html/index.html`**



Viewing results (HTML)



Analyse matrix multiply with enforcing 512 bits vectorization

Run the 512 bits vectorized version of matrix multiply

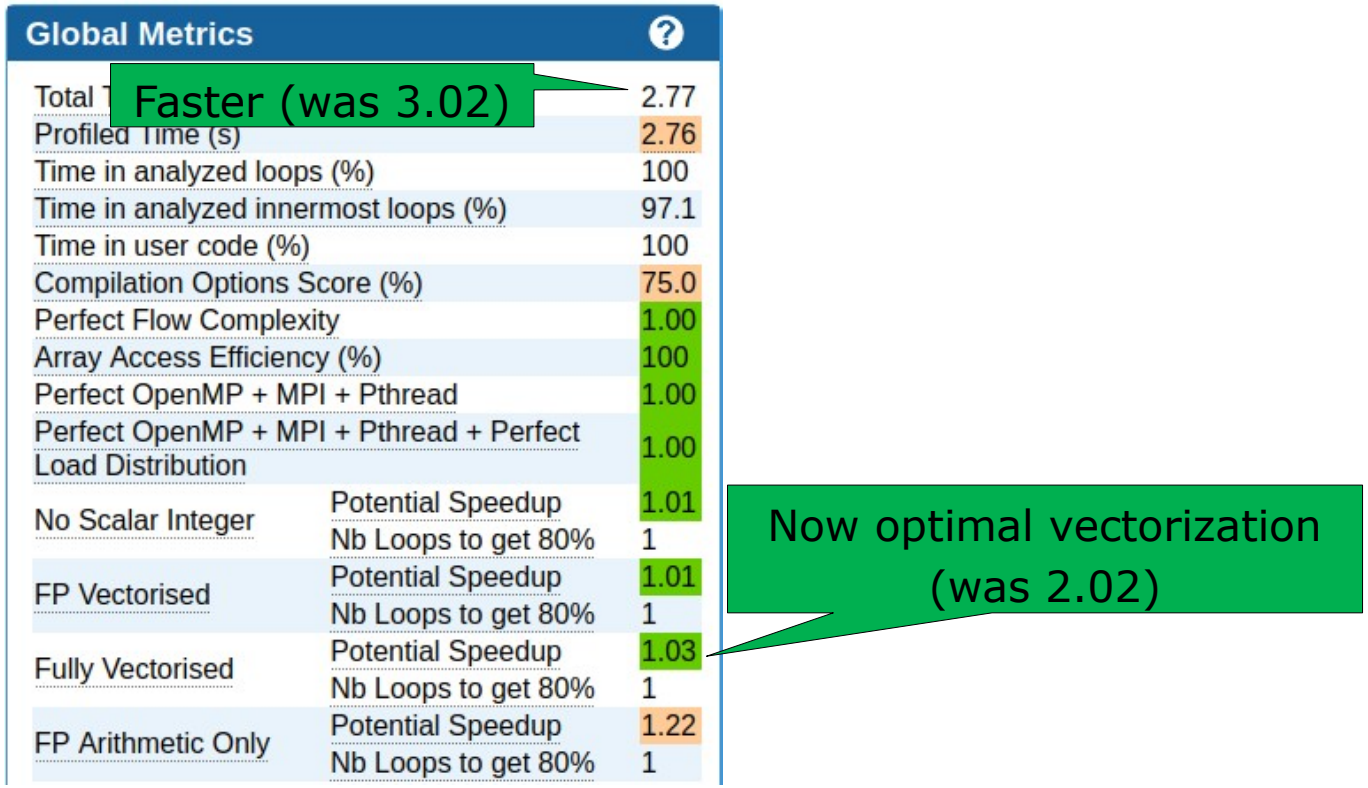
```
> cd $WORK/MAQAO_HANDSON/matmul #if cur. directory has changed
> srun -p ncpu --exclusive -t 1 ./matmul_align_512/matmul 400 300
# remark: size%8 has to equal 0
driver.c: Using posix_memalign instead of malloc
cycles per FMA: 0.22
```

Analyse matrix multiply with ONE View

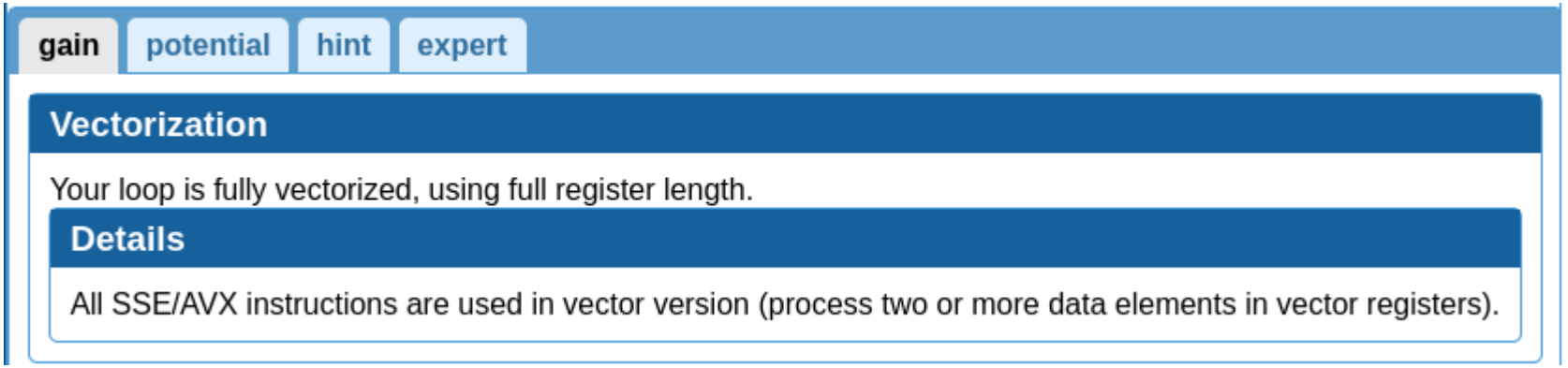
```
> maqao oneview -R1 -c=ov_align_512.lua -xp=ov_align_512
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/matmul/ov_align_512/RESULTS/\ matmul_align_512_one_html/index.html`**



Viewing results (HTML)



Viewing results (HTML)

Global Metrics

Total Time (s)		2.77
Profiled Time (s)		2.76
Time in analyzed loops (%)		100
Time in analyzed innermost loops (%)		97.1
Time in user code (%)		100
Compilation Options Score (%)		75.0
Perfect Flow Complexity		1.00
Array Access Efficiency (%)		100
Perfect OpenMP + MPI + Pthread		1.00
Perfect OpenMP + MPI + Pthread + Perfect Load Distribution		1.00
No Scalar Integer	Potential Speedup	1.01
	Nb Loops to get 80%	1
FP Vectorised	Potential Speedup	1.01
	Nb Loops to get 80%	1
Fully Vectorised	Potential Speedup	1.03
	Nb Loops to get 80%	1
FP Arithmetic Only	Potential Speedup	1.22
	Nb Loops to get 80%	1

Compilation Options

Source Object	Issue
▼ matmul	
○ kernel.c	-funroll-loops is missing.

Let's try this

Analyse matrix multiply with loop unrolling

Run unrolled version of matrix multiply

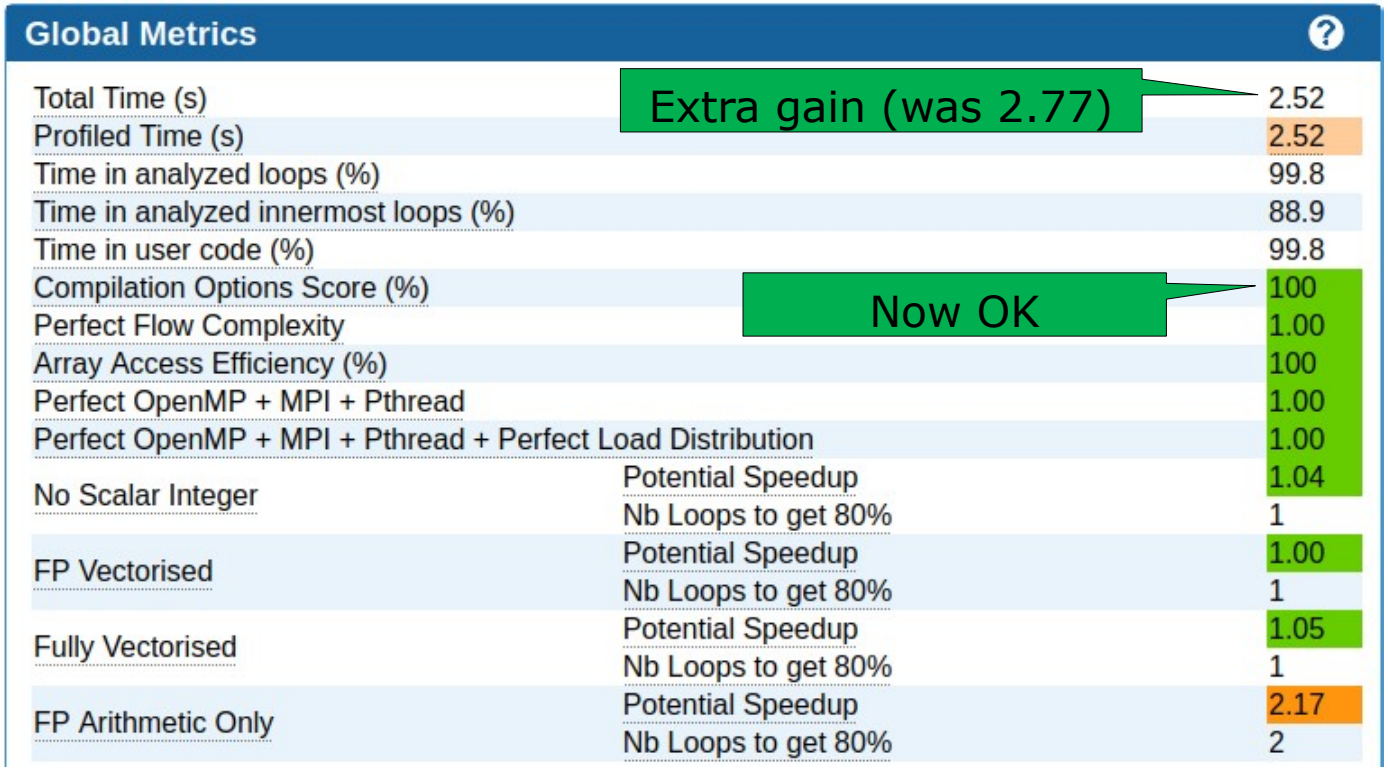
```
> cd $WORK/MAQAO_HANDSON/matmul #if cur. directory has changed
> srun -p ncpu --exclusive -t 1 ./matmul_unroll/matmul 400 300
driver.c: Using posix_memalign instead of malloc
cycles per FMA: 0.20
```

Analyse matrix multiply with ONE View

```
> maqao oneview -R1 -c=ov_unroll.lua -xp=ov_unroll
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/matmul/ov_unroll/RESULTS/\ matmul_unroll_one_html/index.html`**



Using comparison mode: global level

```
> maqao oneview --compare-reports -xp=ov_matmul_cmp \
-inputs=ov_orig,ov_perm,ov_align,ov_align_512,ov_unroll
```

Remark: open `ov_matmul_cmp/RESULTS/ov_matmul_cmp/index.html`

▼ **Compared Reports**

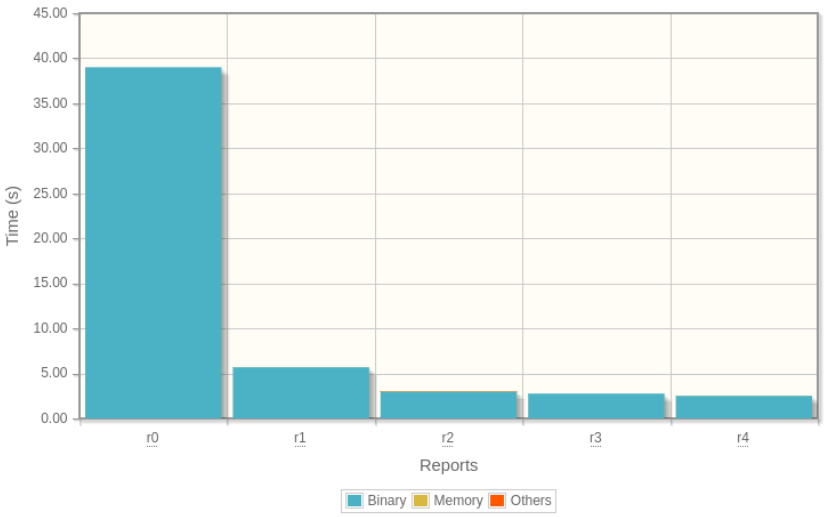
- r0: ov_orig
- r1: ov_perm
- r2: ov_align
- r3: ov_align_512
- r4: ov_unroll

Global Metrics

Metric		r0	r1	r2	r3	r4
Total Time (s)		39.01	5.71	3.02	2.77	2.52
Profiled Time (s)		39.00	5.70	3.01	2.76	2.52
Time in analyzed loops (%)		100.0	99.8	99.7	100	99.8
Time in analyzed innermost loops (%)		99.9	98.3	98.2	97.1	88.9
Time in user code (%)		100	99.8	99.7	100	99.8
Compilation Options Score (%)		50.0	50.0	75.0	75.0	100
Perfect Flow Complexity		1.00	1.00	1.00	1.00	1.00
Array Access Efficiency (%)		83.3	100	100	100	100
Perfect OpenMP + MPI + Pthread		1.00	1.00	1.00	1.00	1.00
Perfect OpenMP + MPI + Pthread + Perfect Load Distribution		1.00	1.00	1.00	1.00	1.00
No Scalar Integer	Potential Speedup	1.00	1.01	1.01	1.01	1.04
	Nb Loops to get 80%	1	1	1	1	1
FP Vectorised	Potential Speedup	2.81	1.41	1.21	1.01	1.00
	Nb Loops to get 80%	1	1	1	1	1
Fully Vectorised	Potential Speedup	16.0	4.02	2.02	1.03	1.05
	Nb Loops to get 80%	1	1	1	1	1
Only FP Arithmetic	Potential Speedup	1.00	1.41	1.21	1.22	2.17
	Nb Loops to get 80%	1	1	1	1	2

Application Categorization

Time



Using comparison mode: experiment summaries

Experiment Summaries

	r0	r1	r2	r3	r4
Application	./matmul_orig/matmul	./matmul_perm/matmul	./matmul_align/matmul	./matmul_align_512/matmul	./matmul_unroll/matmul
Timestamp	2023-06-01 15:00:28	2023-06-01 15:01:46	2023-06-01 16:29:59	2023-06-01 17:01:43	2023-06-01 16:16:57
Experiment Type	MPI;	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Machine	n052	n036	*same as r0*	*same as r0*	*same as r0*
Architecture	x86_64	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Micro Architecture	ICELAKE_SP	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Model Name	Intel(R) Xeon(R) Gold 6338 CPU @ 2.00GHz	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Cache Size	49152 KB	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Number of Cores	32	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Maximal Frequency	2.001 GHz	*same as r0*	*same as r0*	*same as r0*	*same as r0*
OS Version	Linux 3.10.0-1160.71.1.el7.x86_64 #1 SMP Tue Jun 28 15:37:28 UTC 2022	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Architecture used during static analysis	x86_64	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Micro Architecture used during static analysis	ICELAKE_SP	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Compilation Options	**matmul**: GNU C17 12.2.0 -mtune=generic -march=x86-64 -g -O3 -fno-omit-frame-pointer	*same as r0*	**matmul**: GNU C17 12.2.0 -march=icelake-server -g -O3 -fno-omit-frame-pointer	**matmul**: GNU C17 12.2.0 -march=icelake-server -mprefer-vector-width=512 -g -O3 -fno-omit-frame-pointer	**matmul**: GNU C17 12.2.0 -march=icelake-server -mprefer-vector-width=512 -g -O3 -funroll-loops -fno-omit-frame-pointer
Number of processes observed	1	*same as r0*	*same as r0*	*same as r0*	*same as r0*
Number of threads observed	1	*same as r0*	*same as r0*	*same as r0*	*same as r0*

Using comparison mode: function & loop level

Functions

Name	Module	Coverage (%)					Time (s)				
		ov_orig	ov_perm	ov_align	ov_align_512	ov_unroll	ov_orig	ov_perm	ov_align	ov_align_512	ov_unroll
kernel	matmul	99.99	99.82	99.67	99.82	99.6	39	5.69	3.01	2.75	2.51
__GI_memset	libc-2.17.so	NA	0.18	0.17	NA	0.2	NA	0.01	0	NA	0
init_mat	matmul	NA	NA	NA	0.18	0.2	NA	NA	NA	0	0
__random_r	libc-2.17.so	NA	NA	0.17	NA	NA	NA	NA	0	NA	NA
Unknown function	matmul	0.01	NA	NA	NA	NA	0	NA	NA	NA	NA

Loops

▼ **kernel.c: 24 - 482.04%**

Run ov_orig

Loop Source Regions: • /home/trainer7/MAQAO_HANDSON/matmul/matmul_orig/kernel.c: 24-25

ASM Loop ID	Max Time Over Threads (s)	Time w.r.t. Wall Time (s)	Cov (%)	Vect. Ratio (%)	Vector Length Use (%)
1	38.98	38.98	99.94	0	6.25

Run ov_perm

Loop Source Regions: • /home/trainer7/MAQAO_HANDSON/matmul/matmul_perm/kernel.c: 24-25

Assembly Loop ID	Max Time Over Threads (s)	Time w.r.t. Wall Time (s)	Cov (%)	Vect. Ratio (%)	Vector Length Use (%)
4	5.61	5.61	98.33	100	25

Run ov_align

Loop Source Regions: • /home/trainer7/MAQAO_HANDSON/matmul/matmul_align/kernel.c: 24-25

Assembly Loop ID	Max Time Over Threads (s)	Time w.r.t. Wall Time (s)	Cov (%)	Vect. Ratio (%)	Vector Length Use (%)
4	2.96	2.96	98.18	100	50

Run ov_align_512

Loop Source Regions: • /home/trainer7/MAQAO_HANDSON/matmul/matmul_align/kernel.c: 24-25

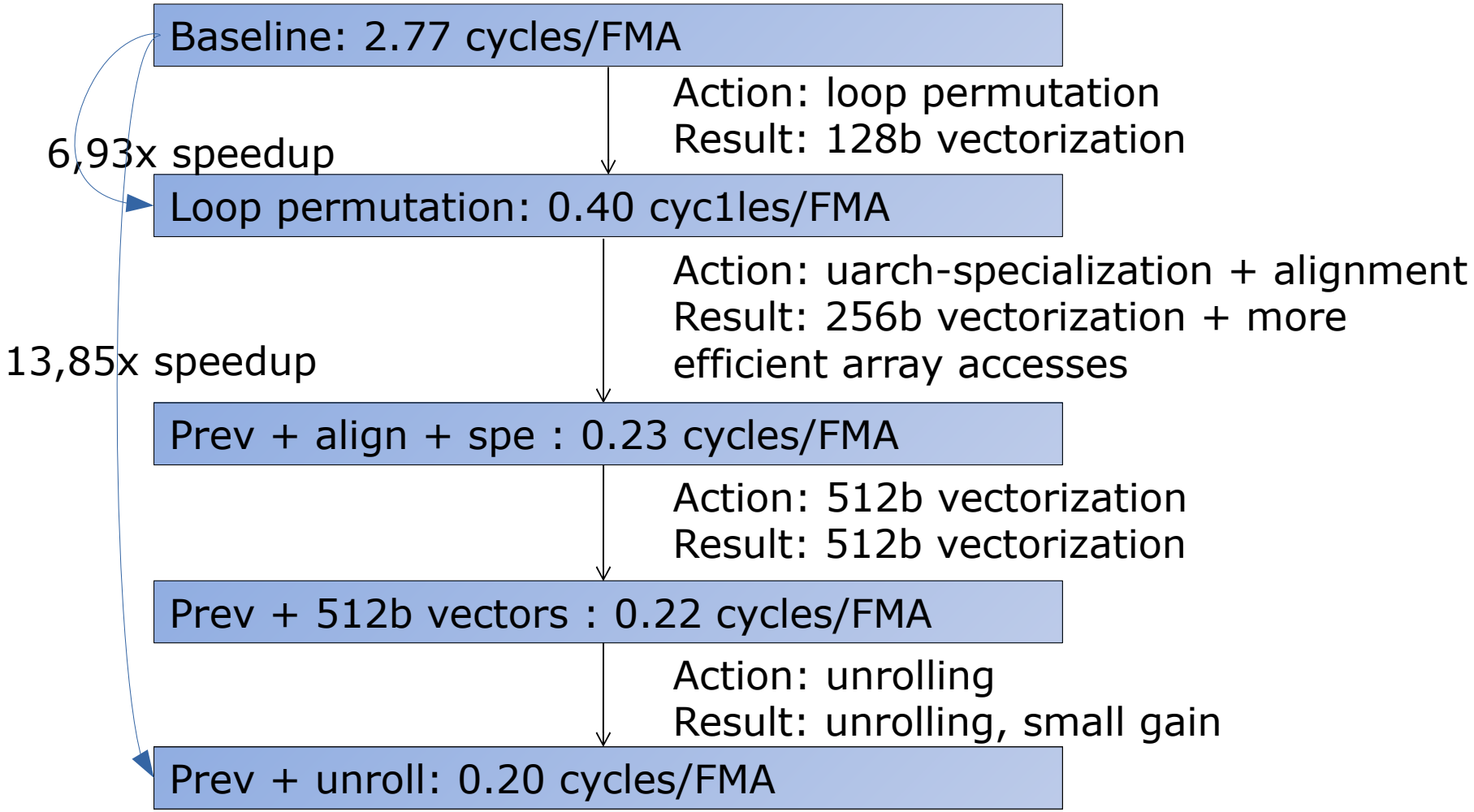
Assembly Loop ID	Max Time Over Threads (s)	Time w.r.t. Wall Time (s)	Cov (%)	Vect. Ratio (%)	Vector Length Use (%)
4	2.67	2.67	96.92	100	100

Run ov_unroll

Loop Source Regions: • /home/trainer7/MAQAO_HANDSON/matmul/matmul_align/kernel.c: 24-25

Assembly Loop ID	Max Time Over Threads (s)	Time w.r.t. Wall Time (s)	Cov (%)	Vect. Ratio (%)	Vector Length Use (%)
4	2.23	2.23	88.67	100	100

Summary of optimizations and gains



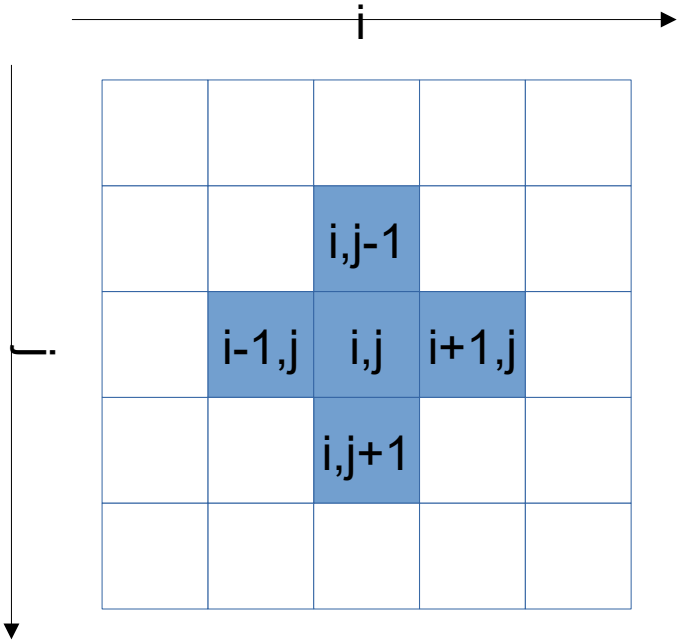
Hydro code

```
int build_index (int i, int j, int grid_size)
{
  return (i + (grid_size + 2) * j);
}

void linearSolver0 (...) {
  int i, j, k;

  for (k=0; k<20; k++)
    for (i=1; i<=grid_size; i++)
      for (j=1; j<=grid_size; j++)
        x[build_index(i, j, grid_size)] =
  (a * ( x[build_index(i-1, j, grid_size)] +
         x[build_index(i+1, j, grid_size)] +
         x[build_index(i, j-1, grid_size)] +
         x[build_index(i, j+1, grid_size)]
       ) + x0[build_index(i, j, grid_size)]
  ) / c;
}
```

Iterative linear system solver using the Gauss-Siedel relaxation technique.
« Stencil » code



Compile and run with Intel compiler

Switch to the hydro handson folder

```
> cd $WORK/MAQAO_HANDSON/hydro
```

Load MAQAO (if necessary)

```
> module load maqao/2.17.4
```

Load latest Intel Compiler (icx 22.)

```
> module load intel-compilers
```

Compile

```
> make
```

Running and analyzing original kernel

The ONE View configuration file must contain all variables for executing the application.

```
> cd $WORK/MAQAO_HANDSON/hydro #if cur. directory has changed
> less ov_orig.lua
```

```
executable = "./hydro_orig"
run_command = "<executable> 300 200" -- <size of matrix> <number of repetitions>
...
number_processes_per_node = 1
mpi_command = "srun -p ncpu --exclusive -t 1"
...
```

Running and analyzing original kernel

Run

```
> srun -p ncpu --exclusive -t 1 \
./hydro_orig 300 200 # 300x300 mesh, 200 repetitions
Cycles per element for solvers: 1248.44
```

Profile with MAQAO

```
> maqao oneview -R1 -xp=ov_orig -c=ov_orig.lua
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/hydro/ov_orig/RESULTS/\ hydro_orig_one_html/index.html`**

Global Metrics

Total Time (s)		11.15
Profiled Time (s)		11.13
Time in analyzed loops (%)		99.9
Time in analyzed innermost loops (%)		99.9
Time in user code (%)		100.0
Compilation Options Score (%)		100.0
Perfect Flow Complexity		1.00
Array Access Efficiency (%)		48.8
Perfect OpenMP + MPI + Pthread		1.00
Perfect OpenMP + MPI + Pthread + Perfect Load Distribution		1.00
No Scalar Integer	Potential Speedup	1.14
	Nb Loops to get 80%	1
FP Vectorised	Potential Speedup	1.21
	Nb Loops to get 80%	2
Fully Vectorised	Potential Speedup	13.9
	Nb Loops to get 80%	5
FP Arithmetic Only	Potential Speedup	1.21
	Nb Loops to get 80%	3

CQA output for original kernel

Workaround

- Try another compiler or update/tune your current one:
  - recompile with fassociative-math (included in Ofast or ffast-math) to extend loop vectorization to FP reductions.
- Remove inter-iterations dependences from your loop and make it unit-stride:
  - If your arrays have 2 or more dimensions, check whether elements are accessed contiguously and, otherwise, try to permute loops accordingly: C storage order is row-major: for(i) for(j) a[j][i] = b[j][i]; (slow, non stride 1) => for(i) for(j) a[i][j] = b[i][j]; (fast, stride 1)
  - If your loop streams arrays of structures (AoS), try to use structures of arrays instead (SoA): for(i) a[i].x = b[i].x; (slow, non stride 1) => for(i) a.x[i] = b.x[i]; (fast, stride 1)

As for matmul, loops should be permuted. CF build_index

Unroll opportunity

Loop is data access bound.

Workaround

Unroll your loop if trip count is significantly higher than target unroll factor and if some data references are common to consecutive iterations. This can be done manually. Or by recompiling with -funroll-loops and/or -floop-unroll-and-jam.

Consider loop unrolling

Running and analyzing kernel with loop permutation

Run

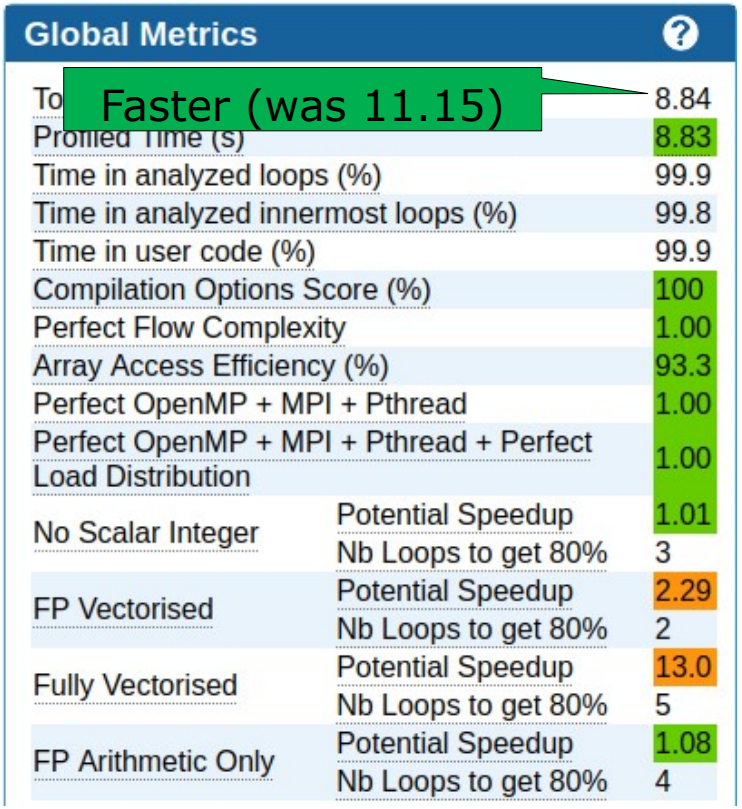
```
> srun -p ncpu --exclusive -t 1 ./hydro_perm 300 200
# 300x300 mesh, 200 repetitions
Cycles per element for solvers: 1007.27
```

Profile with MAQAO

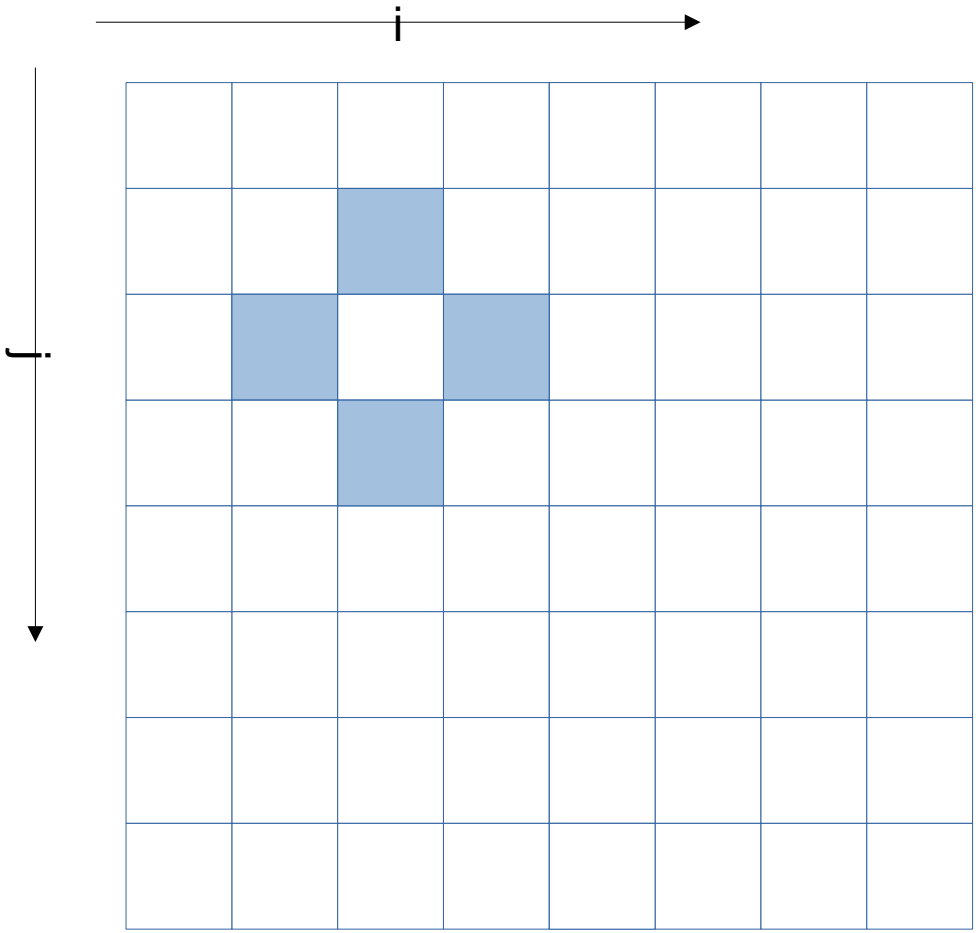
```
> maqao oneview -R1 -xp=ov_perm -c=ov_perm.lua
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/hydro/ov_perm/RESULTS/\ hydro_perm_one_html/index.html`**

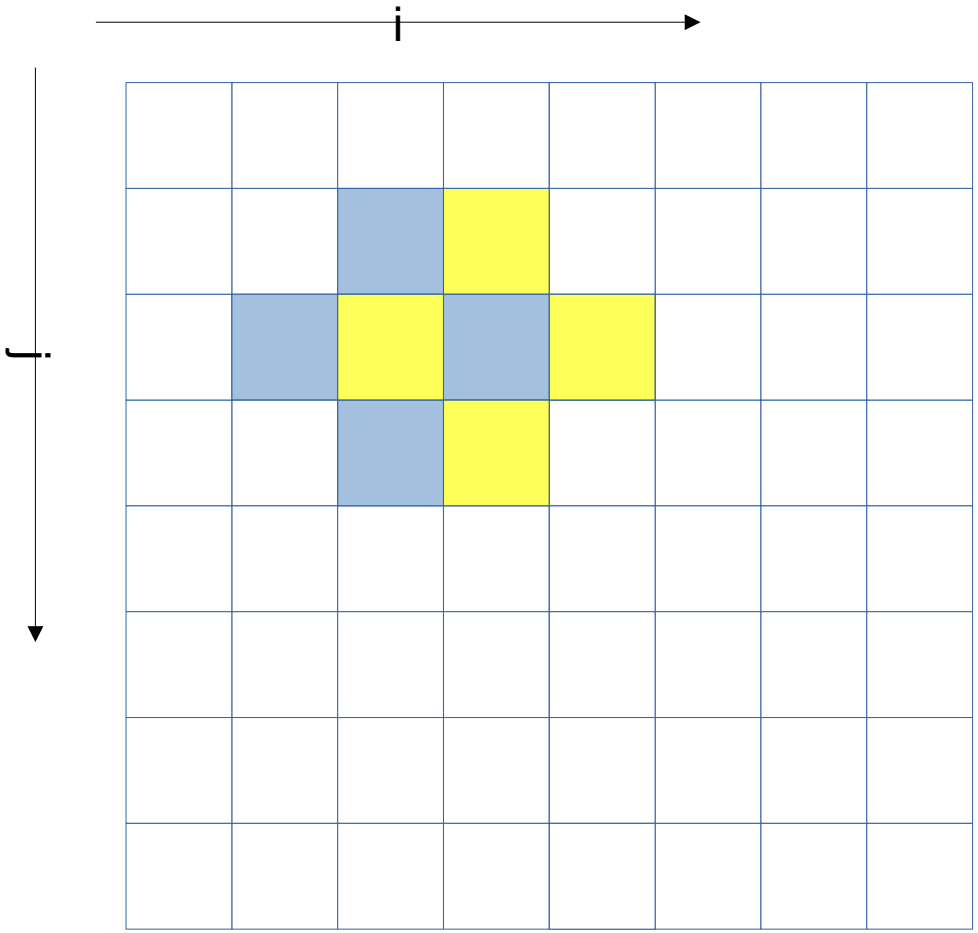


Memory references reuse : 4x4 unroll footprint on loads



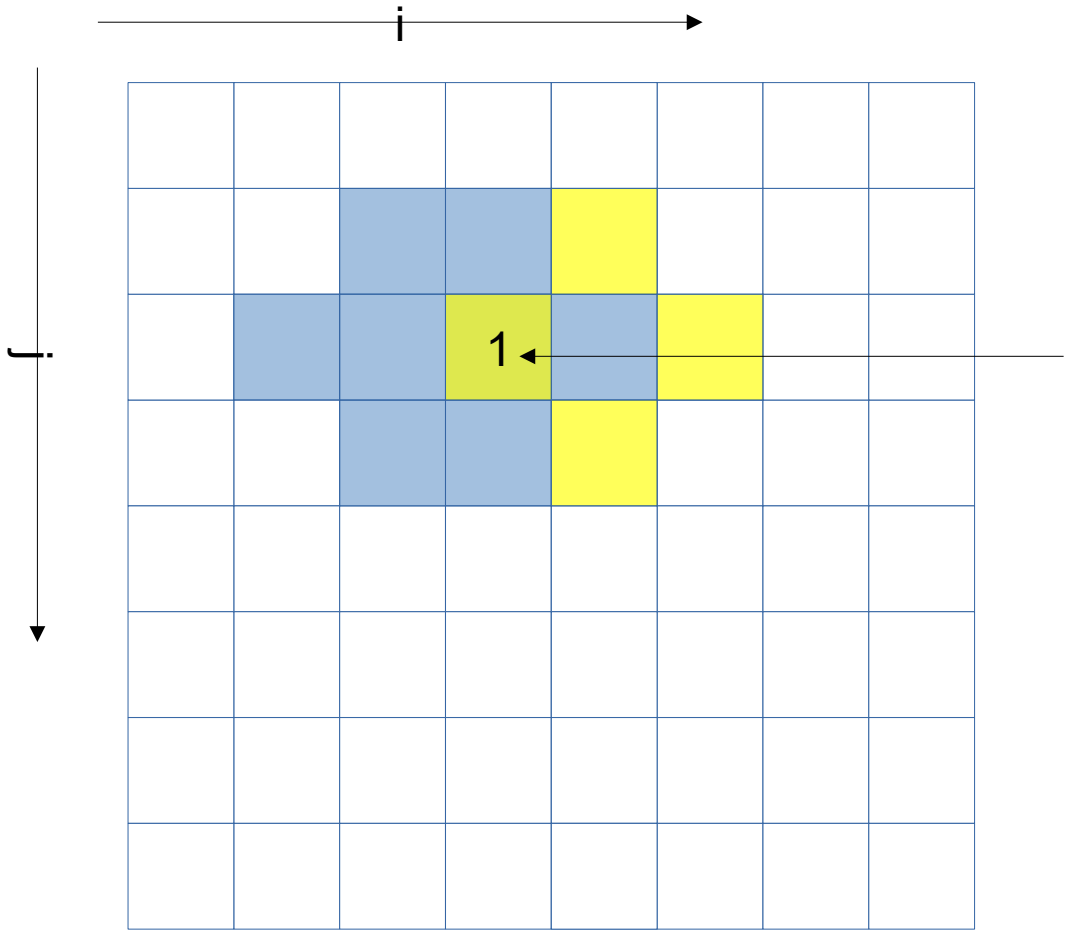
LINEAR_SOLVER(i+0,j+0)

Memory references reuse : 4x4 unroll footprint on loads



LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)

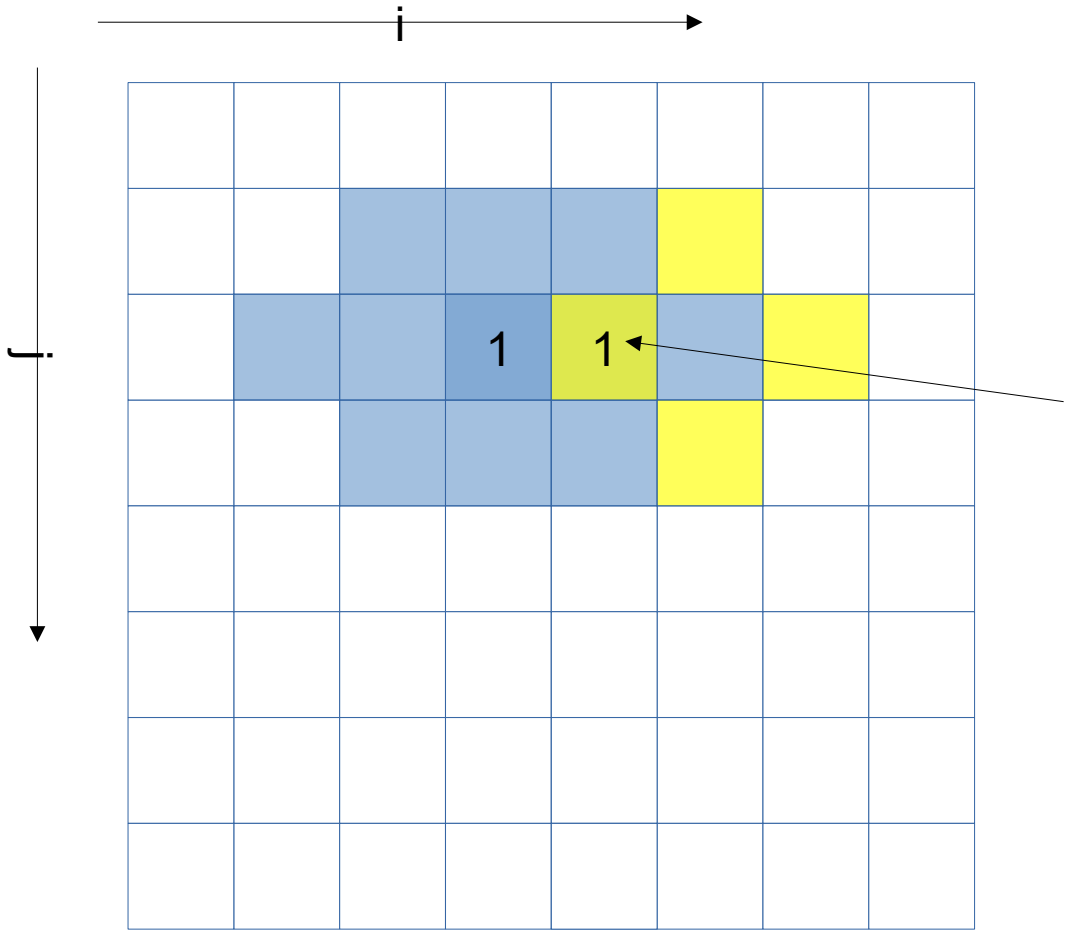
Memory references reuse : 4x4 unroll footprint on loads



LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)

1 reuse

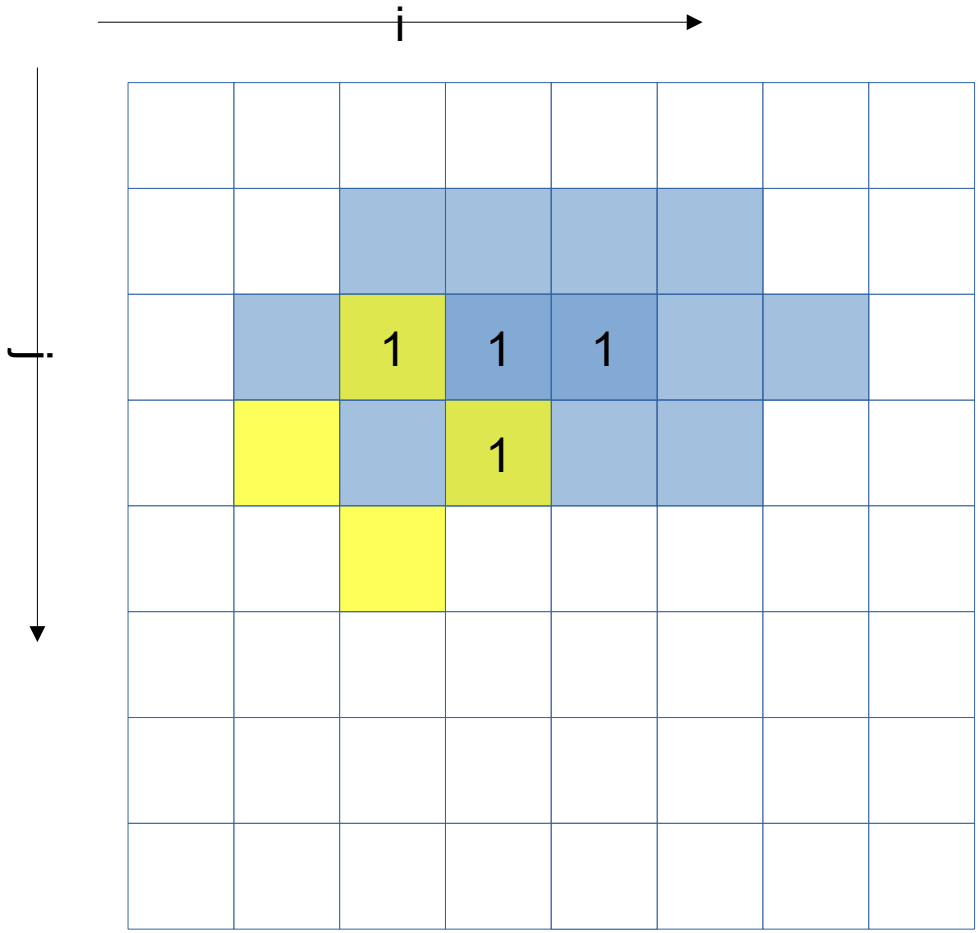
Memory references reuse : 4x4 unroll footprint on loads



LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)
LINEAR_SOLVER(i+3,j+0)

2 reuses

Memory references reuse : 4x4 unroll footprint on loads

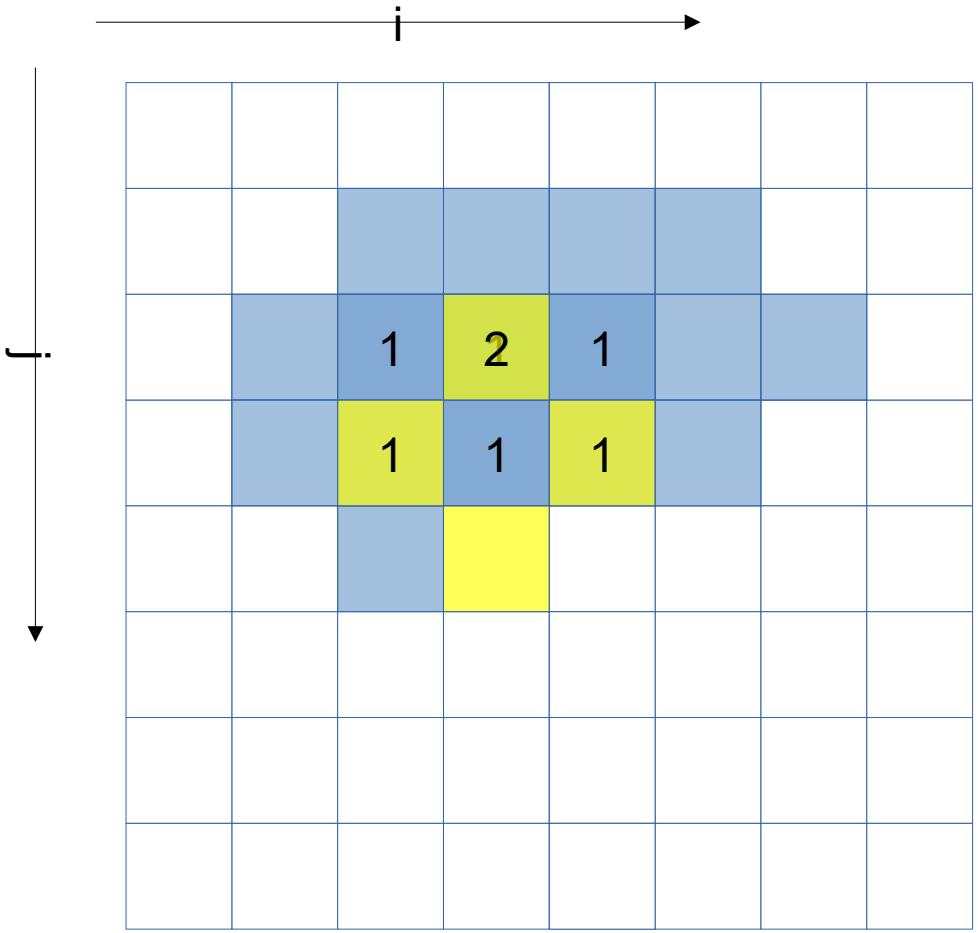


LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)
LINEAR_SOLVER(i+3,j+0)

LINEAR_SOLVER(i+0,j+1)

4 reuses

Memory references reuse : 4x4 unroll footprint on loads

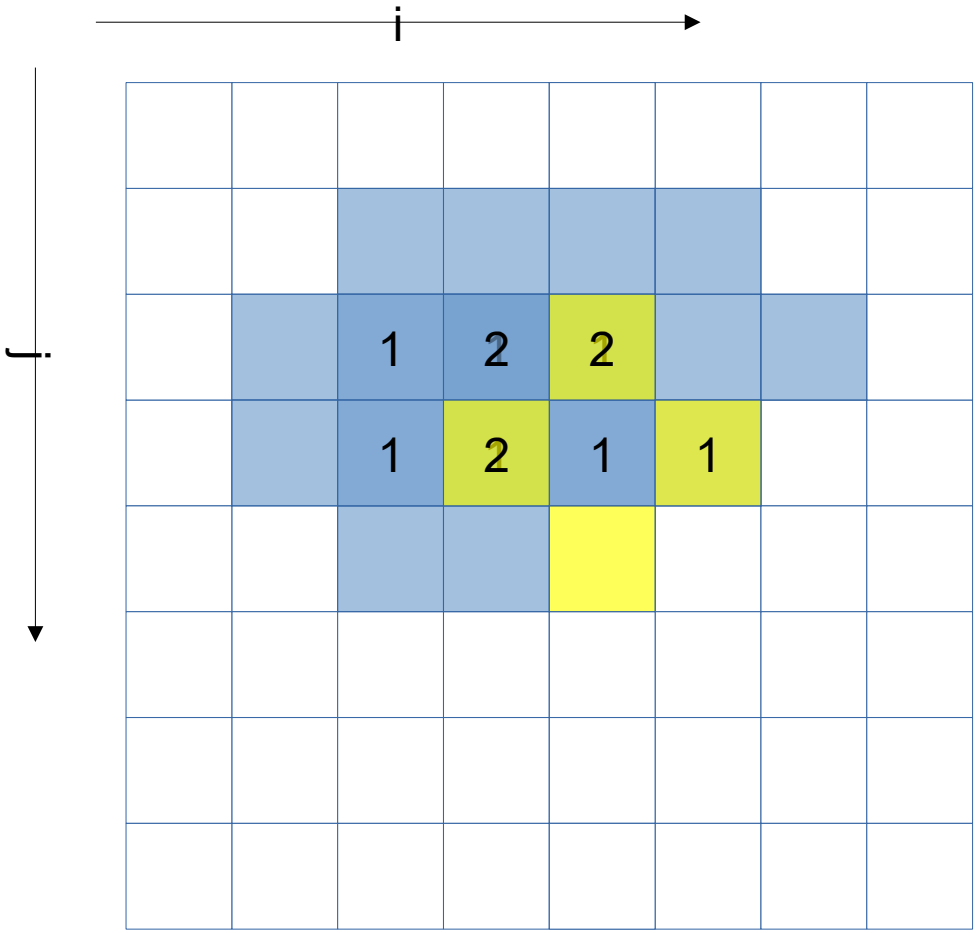


LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)
LINEAR_SOLVER(i+3,j+0)

LINEAR_SOLVER(i+0,j+1)
LINEAR_SOLVER(i+1,j+1)

7 reuses

Memory references reuse : 4x4 unroll footprint on loads

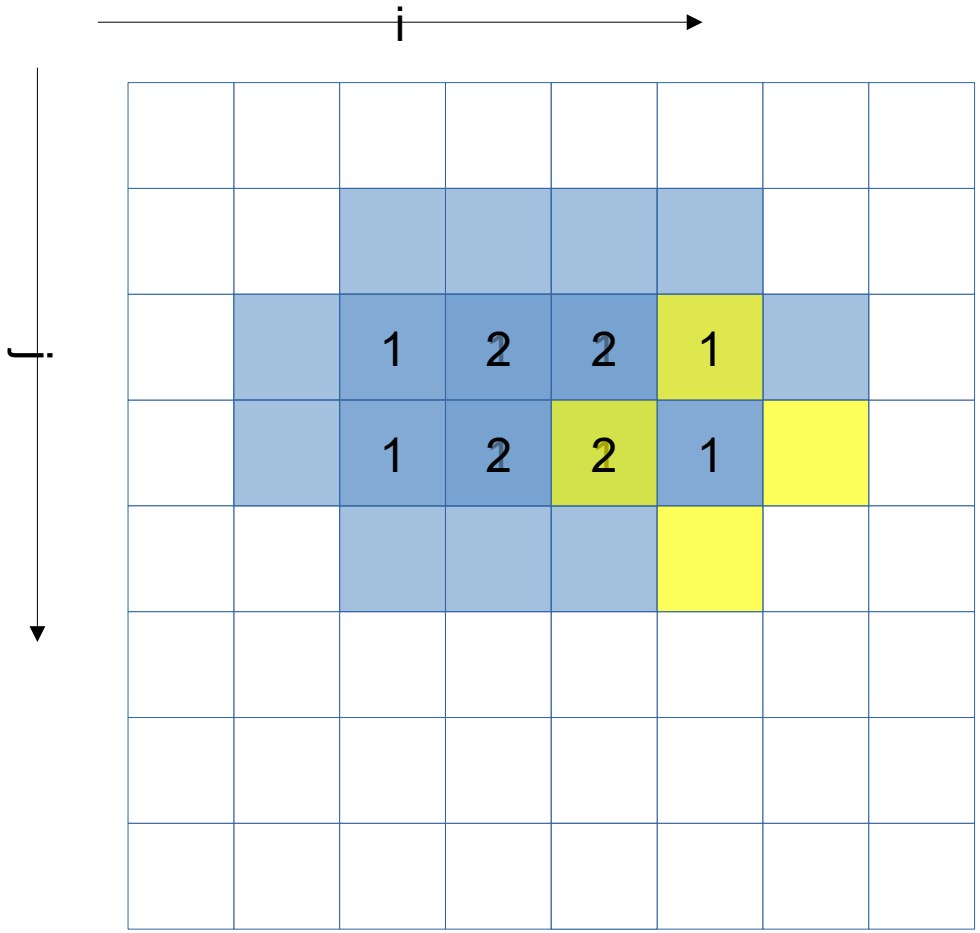


LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)
LINEAR_SOLVER(i+3,j+0)

LINEAR_SOLVER(i+0,j+1)
LINEAR_SOLVER(i+1,j+1)
LINEAR_SOLVER(i+2,j+1)

10 reuses

Memory references reuse : 4x4 unroll footprint on loads

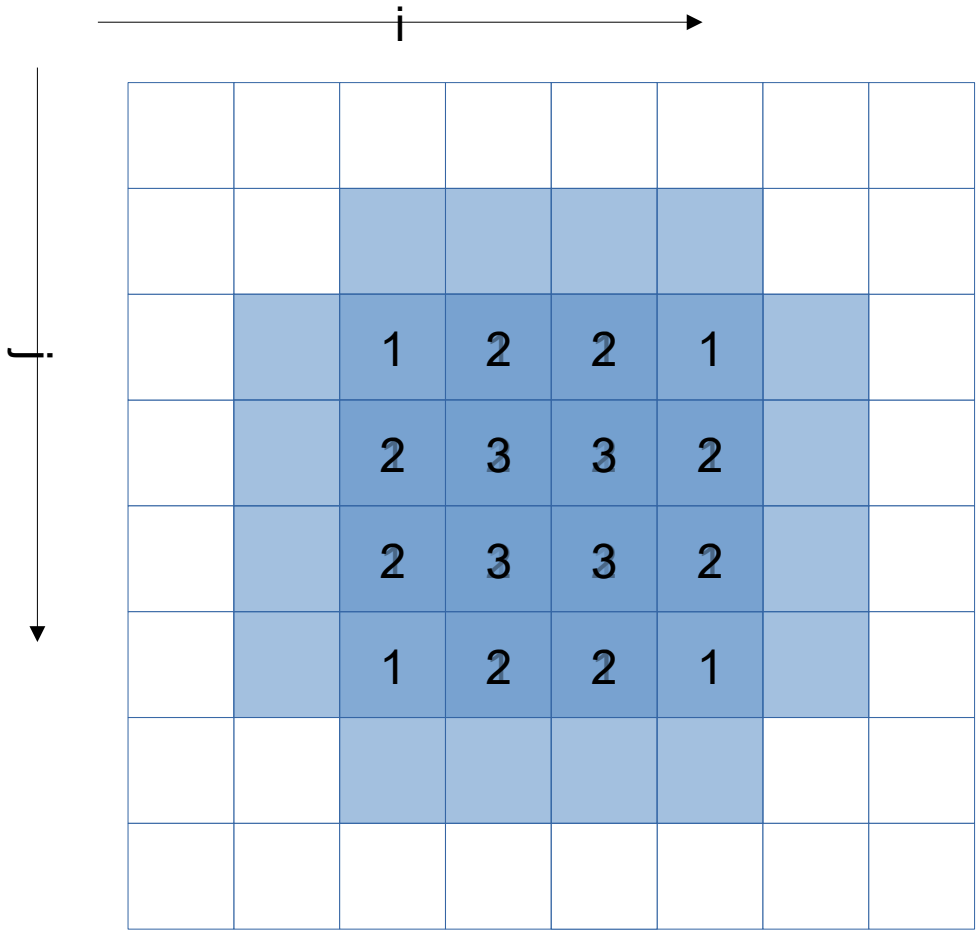


LINEAR_SOLVER(i+0,j+0)
LINEAR_SOLVER(i+1,j+0)
LINEAR_SOLVER(i+2,j+0)
LINEAR_SOLVER(i+3,j+0)

LINEAR_SOLVER(i+0,j+1)
LINEAR_SOLVER(i+1,j+1)
LINEAR_SOLVER(i+2,j+1)
LINEAR_SOLVER(i+3,j+1)

12 reuses

Memory references reuse : 4x4 unroll footprint on loads



LINEAR_SOLVER(i+0-3,j+0)

LINEAR_SOLVER(i+0-3,j+1)

LINEAR_SOLVER(i+0-3,j+2)

LINEAR_SOLVER(i+0-3,j+3)

32 reuses

Impacts of memory reuse

- For the x array, instead of 4x4x4 = 64 loads, now only 32 (32 loads avoided by reuse)
- For the x0 array no reuse possible : 16 loads
- Total loads : 48 instead of 80

4x4 unroll

```
#define LINEARSOLVER(...) x[build_index(i, j, grid_size)] = …

void linearSolver2 (...) {
  (...)

  for (k=0; k<20; k++)
    for (j=1; j<=grid_size-3; j+=4)
      for (i=1; i<=grid_size-3; i+=4) {
        LINEARSOLVER (…, i+0, j+0);
        LINEARSOLVER (…, i+1, j+0);
        LINEARSOLVER (…, i+2, j+0);
        LINEARSOLVER (…, i+3, j+0);

        LINEARSOLVER (…, i+0, j+1);
        LINEARSOLVER (…, i+1, j+1);
        LINEARSOLVER (…, i+2, j+1);
        LINEARSOLVER (…, i+3, j+1);

        LINEARSOLVER (…, i+0, j+2);
        LINEARSOLVER (…, i+1, j+2);
        LINEARSOLVER (…, i+2, j+2);
        LINEARSOLVER (…, i+3, j+2);

        LINEARSOLVER (…, i+0, j+3);
        LINEARSOLVER (…, i+1, j+3);
        LINEARSOLVER (…, i+2, j+3);
        LINEARSOLVER (…, i+3, j+3);
      }
}
```

grid_size must now be multiple of 4. Or loop control must be adapted (much less readable) to handle leftover iterations

Running and analyzing kernel with manual 4x4 unroll

Run

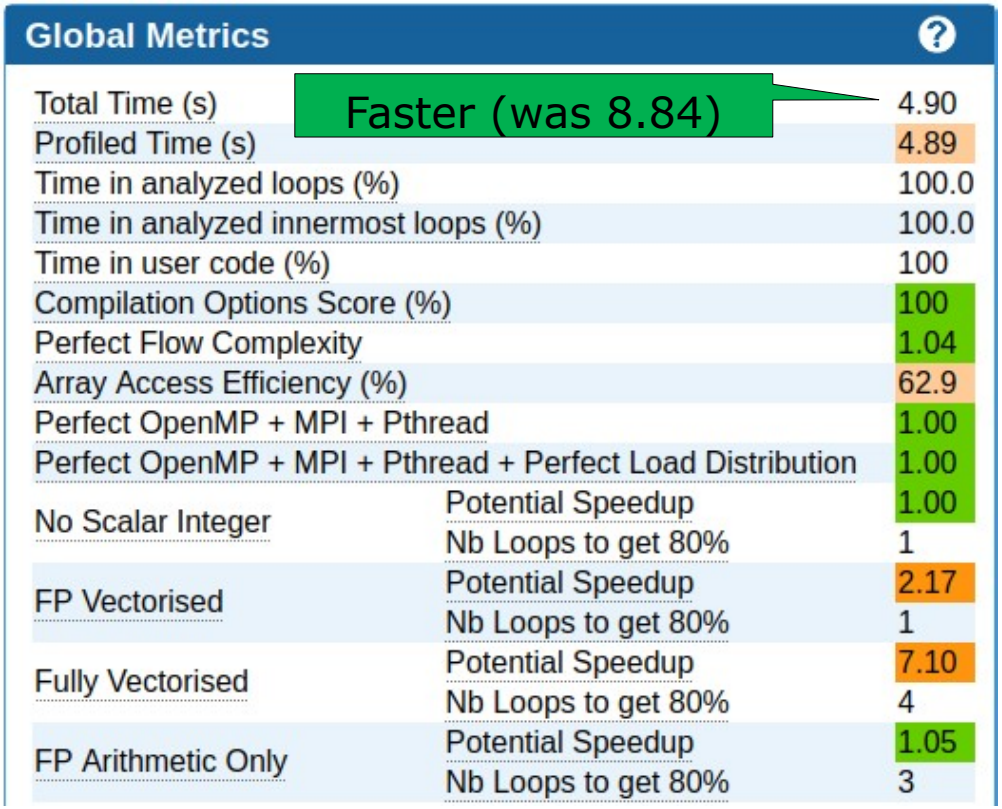
```
> srun -p ncpu --exclusive -t 1 ./hydro_unroll 300 200
# 300x300 mesh, 200 repetitions
Cycles per element for solvers: 433.13
```

Profile with MAQAO

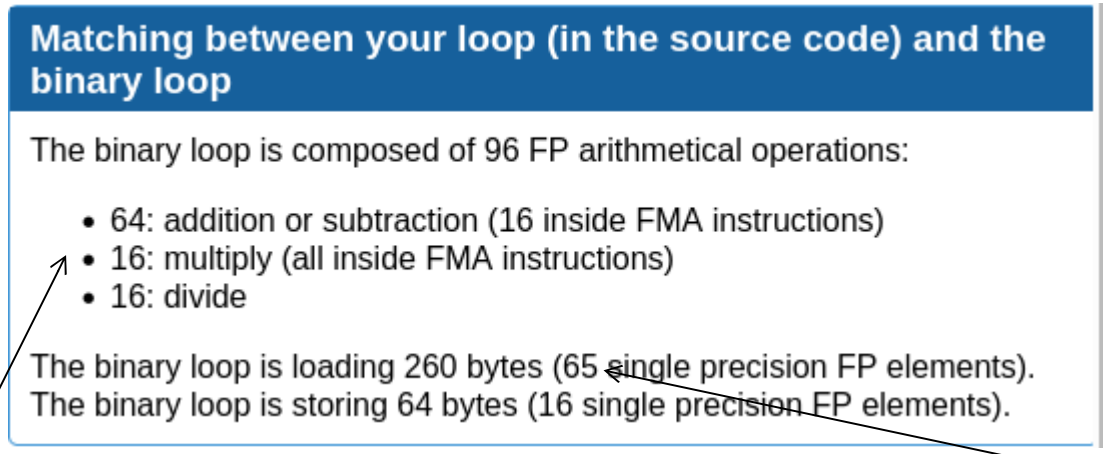
```
> maqao oneview -R1 -xp=ov_unroll -c=ov_unroll.lua
```

Viewing results (HTML)

Open file **`MAQAO_HANDSON/hydro/ov_unroll/RESULTS/\ hydro_unroll_one_html/index.html`**



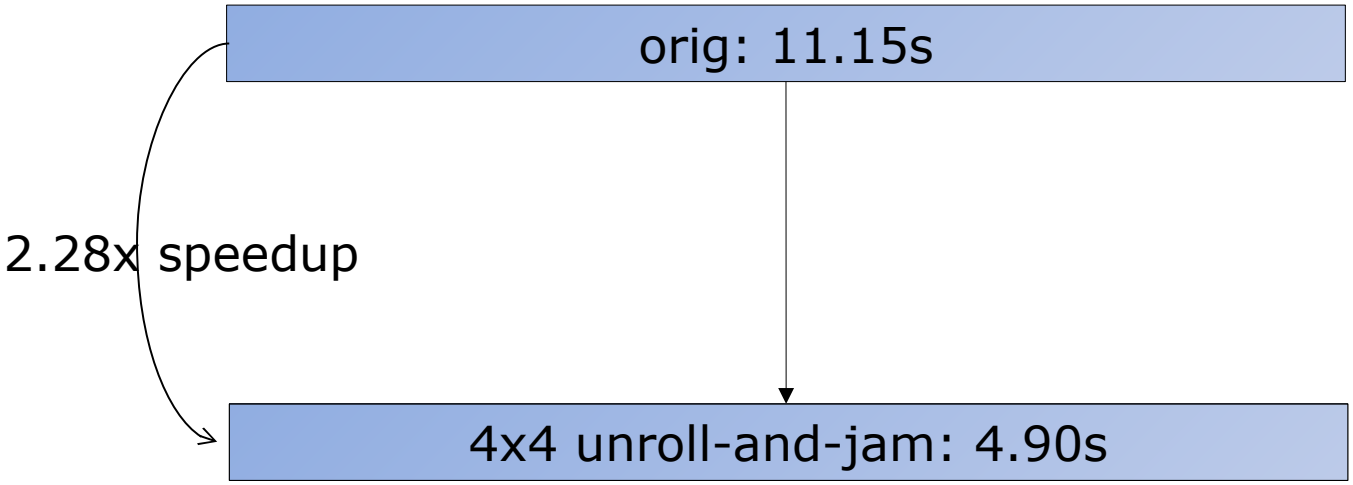
CQA output for unrolled kernel



4x4 Unrolling were applied

Lower than 80: 64 (from x) + 16 (from x0)

Summary of optimizations and gains



More sample codes

More codes to study with MAQAO in

```
$WORK/MAQAO_HANDSON/loop_optim_tutorial.tgz
```