

HW2

Tao Chen

The CPU information I used is:

Architecture:	x86_64
[CPU op-mode(s):	32-bit, 64-bit
Byte Order:	Little Endian
CPU(s):	64
On-line CPU(s) list:	0-63
Thread(s) per core:	2
Core(s) per socket:	8
Socket(s):	4
NUMA node(s):	8
Vendor ID:	AuthenticAMD
CPU family:	21
Model:	2
Model name:	AMD Opteron(tm) Processor 6380
Stepping:	0
CPU MHz:	2499.940
BogoMIPS:	4999.88
Virtualization:	AMD-V
L1d cache:	16K
L1i cache:	64K
L2 cache:	2048K
L3 cache:	6144K
NUMA node0 CPU(s):	0-7
NUMA node1 CPU(s):	8-15
NUMA node2 CPU(s):	32-39
NUMA node3 CPU(s):	40-47
NUMA node4 CPU(s):	48-55
NUMA node5 CPU(s):	56-63
NUMA node6 CPU(s):	16-23
NUMA node7 CPU(s):	24-31

Problem 1:

1. x should be freed instead of deleted ;
exceeding array bound when visiting elements of x .
2. Not every elements of X is initialized.

From the above 6 orders, we can see that the most efficient order is "jpi" and "pji", that's because, under "jpi" and "pji", the innermost loop is column-wise iterations of A and C, and the data is stored column wise, so under this case it is the most efficient.

(ii) Blocking scheme under "jpi" order:
(with block size = 24)

Dimension	Time	Gflop/s	GB/s	Error
24	0.403922	4.951449	41.262073	0.000000e+00
72	0.434220	4.697368	37.370871	0.000000e+00
120	0.427332	4.682593	37.772918	0.000000e+00
168	0.438146	4.566898	36.752651	0.000000e+00
216	0.430559	4.681212	37.623071	0.000000e+00
264	0.432745	4.677052	37.558141	0.000000e+00
312	0.431582	4.644563	37.275597	0.000000e+00
360	0.442077	4.643682	37.252648	0.000000e+00
408	0.449663	4.531210	36.338525	0.000000e+00
456	0.472333	4.416403	35.408708	0.000000e+00
504	0.484933	4.231988	33.922435	0.000000e+00
552	0.472065	4.275594	34.266716	0.000000e+00
600	0.500275	4.317621	34.598539	0.000000e+00
648	0.499677	4.356377	34.984798	0.000000e+00
696	0.465294	4.347624	34.838963	0.000000e+00
744	0.570731	4.329511	34.682645	0.000000e+00
792	0.685272	4.349743	34.841885	0.000000e+00
840	0.546219	4.340416	34.764666	0.000000e+00
888	0.652836	4.290371	34.361622	0.000000e+00
936	0.766824	4.277520	34.256720	0.000000e+00
984	0.913340	4.172659	33.415194	0.000000e+00

We can see that with blocking, as the size of the matrix increases, flop and bandwidth won't decrease that much compared to the previous without blocking case.

Block size = 28

Block size = 32

Dimension	Time	Gflop/s	GB/s	Error
28	0.4800429	4.994651	41.384252	0.000000e+00
56	0.422823	4.730741	38.521750	0.000000e+00
84	0.446476	4.481696	36.288396	0.000000e+00
112	0.412862	4.845731	39.111969	0.000000e+00
140	0.428100	4.679091	37.700186	0.000000e+00
168	0.462100	4.329482	34.841375	0.000000e+00
196	0.422498	4.740515	38.117611	0.000000e+00
224	0.427434	4.688528	37.611387	0.000000e+00
252	0.466516	4.322211	34.714904	0.000000e+00
280	0.430145	4.695127	37.695165	0.000000e+00
308	0.526600	3.883912	31.172176	0.000000e+00
336	0.433849	4.721243	37.883800	0.000000e+00
364	0.433343	4.674359	37.497603	0.000000e+00
392	0.442498	4.628348	37.121243	0.000000e+00
420	0.455438	4.554876	36.525771	0.000000e+00
448	0.523285	4.123887	33.064737	0.000000e+00
476	0.481475	4.479994	35.915246	0.000000e+00
504	0.490529	4.175868	33.473225	0.000000e+00
532	0.483636	4.358574	34.934137	0.000000e+00
560	0.471475	4.469795	35.822173	0.000000e+00
588	0.466210	4.360639	34.944441	0.000000e+00
616	0.559350	4.178869	33.485226	0.000000e+00

Dimension	Time	Gflop/s	GB/s	Error
32	0.529963	3.773903	31.134696	0.000000e+00
64	0.561299	3.563504	28.953474	0.000000e+00
96	0.537866	3.720762	30.076159	0.000000e+00
128	0.694161	2.882160	23.237412	0.000000e+00
160	0.540266	3.714912	29.985041	0.000000e+00
192	0.579923	3.466182	27.873883	0.000000e+00
224	0.540709	3.699988	29.732043	0.000000e+00
256	0.688011	2.926214	23.501155	0.000000e+00
288	0.543434	3.692412	29.641863	0.000000e+00
320	0.583714	3.488049	27.931002	0.000000e+00
352	0.547562	3.663974	29.395064	0.000000e+00
384	0.668857	3.047635	24.444573	0.000000e+00
416	0.557730	3.614214	28.983217	0.000000e+00
448	0.626428	3.444879	27.620541	0.000000e+00
480	0.617207	3.583630	28.728761	0.000000e+00
512	0.766717	2.800882	22.458821	0.000000e+00
544	0.631076	3.571472	28.624297	0.000000e+00
576	0.668392	3.431435	27.499140	0.000000e+00
608	0.627330	3.582736	28.709025	0.000000e+00
640	0.699033	3.000078	24.038121	0.000000e+00
672	0.675384	3.594571	28.799359	0.000000e+00
704	0.610793	3.427480	27.458788	0.000000e+00

We can see that the optimal block size is 24 or 28.

civ) Implementing openmp. (set number of threads = 8
block size = 24)

without omp for

With omp for collapse(3)

Dimension	Time	Gflop/s	GB/s	Error
24	0.400901	4.988765	41.573043	0.000000e+00
72	0.435086	4.598192	37.296447	0.000000e+00
120	0.435086	4.711058	38.002424	0.000000e+00
168	0.436187	4.587413	36.917755	0.000000e+00
216	0.431582	4.670120	37.533924	0.000000e+00
264	0.431330	4.692398	37.681376	0.000000e+00
312	0.433158	4.627658	37.139919	0.000000e+00
360	0.442907	4.634979	37.182834	0.000000e+00
408	0.454868	4.479443	35.923375	0.000000e+00
456	0.471901	4.420449	35.441147	0.000000e+00
504	0.483028	4.240728	33.993070	0.000000e+00
552	0.461748	4.371126	35.032361	0.000000e+00
600	0.499930	4.320603	34.622433	0.000000e+00
648	0.504098	4.318170	34.598670	0.000000e+00
696	0.465944	4.341555	34.782346	0.000000e+00
744	0.567700	4.352626	34.867818	0.000000e+00
792	0.691479	4.310781	34.529153	0.000000e+00
840	0.546936	4.334728	34.719042	0.000000e+00
888	0.658916	4.303025	34.462965	0.000000e+00
936	0.774787	4.233555	33.904627	0.000000e+00
984	0.923411	4.127148	33.050742	0.000000e+00

Dimension	Time	Gflop/s	GB/s	Error
24	1.094056	1.828062	15.233848	0.000000e+00
72	0.386443	16.556798	41.991061	1.403188e+04
120	0.120859	16.556718	133.557523	5.820613e+02
168	0.080243	24.895654	280.356740	3.425747e+01
216	0.069709	28.913489	232.378143	1.546141e-11
264	0.069011	29.328438	235.516244	1.22343e-11
312	0.065447	30.628053	245.809761	9.549694e-12
360	0.065117	31.525765	252.906889	8.440200e-12
408	0.064930	31.388225	251.657899	4.774847e-12
456	0.067290	30.999593	248.549597	5.46968e-12
504	0.068529	29.890788	239.600763	5.911716e-12
552	0.067368	29.968604	248.114537	3.183231e-12
600	0.071861	30.058056	240.865654	3.410605e-12
648	0.074573	29.198067	233.889093	3.524292e-12
696	0.068014	29.742884	238.284307	3.183231e-12
744	0.083028	29.766862	238.405464	3.63779e-12
792	0.184082	28.638676	229.398685	4.433787e-12
840	0.081026	29.260851	234.359073	1.705303e-12
888	0.097926	28.602259	229.075749	1.875833e-12
936	0.118295	27.728159	222.062265	1.705303e-12
984	0.148534	25.657850	205.471404	1.932676e-12

with omp, the peak flop rate increases by 58%

(v) without omp, using -O2 and -O3 respectively.
 (block size = 24)

-O3

-O2

Dimension	Time	Gflop/s	GB/s	Error
24	0.400981	4.988765	41.573043	0.000000e+00
72	0.435086	4.598192	37.296447	0.000000e+00
120	0.424751	4.711050	38.002473	0.000000e+00
168	0.436187	4.587413	36.917755	0.000000e+00
216	0.431582	4.670120	37.533924	0.000000e+00
264	0.431330	4.692398	37.681376	0.000000e+00
312	0.433158	4.627658	37.139919	0.000000e+00
360	0.442987	4.634979	37.182834	0.000000e+00
408	0.454860	4.479443	35.923375	0.000000e+00
456	0.471981	4.420449	35.441147	0.000000e+00
504	0.483028	4.240720	33.993070	0.000000e+00
552	0.461748	4.371126	35.032361	0.000000e+00
600	0.499930	4.320603	34.622443	0.000000e+00
648	0.504098	4.318170	34.598670	0.000000e+00
696	0.465944	4.341555	34.782346	0.000000e+00
744	0.567700	4.352626	34.867810	0.000000e+00
792	0.691479	4.310701	34.529153	0.000000e+00
840	0.546936	4.334720	34.719842	0.000000e+00
888	0.650916	4.303025	34.462965	0.000000e+00
936	0.774787	4.233555	33.984627	0.000000e+00
984	0.923411	4.127148	33.050742	0.000000e+00

Dimension	Time	Gflop/s	GB/s	Error
24	1.630712	1.226459	10.220489	0.000000e+00
72	1.640428	1.219565	9.892028	0.000000e+00
120	1.710477	1.169863	9.436894	0.000000e+00
168	1.677966	1.192539	9.597101	0.000000e+00
216	1.678814	1.200573	9.649052	0.000000e+00
264	1.652624	1.224782	9.834726	0.000000e+00
312	1.639021	1.222991	9.815283	0.000000e+00
360	1.731750	1.185428	9.509765	0.000000e+00
408	1.692924	1.203550	9.652001	0.000000e+00
456	1.783226	1.169798	9.378904	0.000000e+00
504	1.752592	1.168775	9.368749	0.000000e+00
552	1.679987	1.201414	9.628721	0.000000e+00
600	1.833338	1.178179	9.441137	0.000000e+00
648	1.799218	1.209849	9.693728	0.000000e+00
696	1.679199	1.204694	9.651399	0.000000e+00
744	2.077281	1.189528	9.529017	0.000000e+00
792	2.493566	1.195469	9.575344	0.000000e+00
840	1.972866	1.201712	9.625137	0.000000e+00
888	2.329389	1.202422	9.630208	0.000000e+00
936	2.746289	1.194432	9.565581	0.000000e+00
984	3.347292	1.138549	9.117648	0.000000e+00

In general, -O3 is better than -O2.

Problem 3.

2. Add reduction (t ; total) after pragma for
3. Delete the last "/* pragma omp barrier"
4. define N to be smaller
5. In the second section,
First lock a then lock b;
First unlock b then unlock a.
6. define 'sum' in the outer most scope.

Problem 4.

jacobi:

# threads time	N	10	50	100	200
4		0.010	0.372	0.989	3.504
8		0.013	0.351	0.727	2.383
16		0.019	0.404	0.692	1.856

A-S:

# threads time	N	10	50	100	200
4		0.007	0.324	0.996	3.354
8		0.009	0.321	0.813	2.338
16		0.014	0.439	0.917	2.031

(Note: I set `max_iterations = 5000` and break the iteration when $\text{error} < 1e-6$. So only when $N=10$, the error can be reduced to $1e-6$ within 5000 iterations; for $N=50, 100$, and 200, it implements 5000 iterations.)