Research & Reading 1 – R4a

Comments

a. Which applications would benefit most from the VLIW architectures?

Embedded processors, designed to perform special-purpose functions, usually in real time, in some kind of hardware.

b. Is it suitable for desktop computers, or mobile devices?

It is not suitable for desktop computers and mobile devices

Research & Reading 2 – R4b

Comments

a. What assumptions were made about computer organization before adding SMT?

Simultaneous multithreading provides an efficient base technology that can be used in many ways to extract improved performance.

b. How does it compare to Wall's claims of ILP limits?

c. What changes were made to add SMT?

Simultaneous multithreading needs no special hardware to schedule instructions from the different threads onto the functional units. Dynamic scheduling hardware in current out-of-order superscalars is already functionally capable of simultaneous multithreaded scheduling. Register renaming eliminates register name conflicts both within and between threads by mapping thread-specific architectural registers onto the hardware registers; the processor then issues instructions withour regard to thread.

d. What performance advantages are claimed? For what workloads?

About 3-4 speedups for multiprogramming workload with different threads.

About 4-5 speedups for parallel workload.

Research & Reading 3

The Top 500 list categorizes the fastest scientific machines in the world accounting to their performance on the Linpack benchmarks. Visit the web site www.top500.org (also the page under the statistics tab).

a. Name the most dominant architecture and operating system.

The most dominant architecture is Cluster (453 / 500).

The most dominant operating system is Linux (243/500).

b. List the imperative characteristics of two machines (ranking among top ten) in a table format.

|  |  |  |
| --- | --- | --- |
| Name | Summit | Sunway TaihuLight |
| Rank | 1 | 3 |
| Manufacturer | IBM | NRCPC |
| Cores | 2,414,592 | 10,649,600 |
| Memory | 2,801,664 GB | 1,310,720 GB |
| Processor | IBM POWER9 22C 3.07GHz | Sunway SW26010 260C 1.45GHz |
| Interconnect | Dual-rail Mellanox EDR Infiniband | Sunway |
| Linpack Performance | 148,600 TFlop/s | 93,014.6 TFlop/s |
| Theoretical Peak | 200,795 TFlop/s | 125,436 TFlop/s |
| Nmax | 16,473,600 | 12,288,000 |
| HPCG [TFlop/s] | 2,925.75 | 480.848 |
| Power | 10,096.00 kW (Submitted) | 15,371.00 kW (Submitted) |
| Power Measurement Level | 3 | 2 |
| Operating System | RHEL 7.4 | Sunway RaiseOS 2.0.5 |

Exercise 1

a) Why might speculation and prediction be of less value in the embedded computer marketplace

than in the server or desktop arena? What are the market niches where they will be least valued?

Answer

b) Hardware multithreading is to increase the utilization of resources on a chip by allowing multiple threads to share the functional units of a single processor. Briefly describe the advantages and disadvantages of fine‐grain multithreading.

Answer

Advantages:

Hide both short and long stalls, since instructions from other threads execute when one thread stalls.

Disadvantages:

Slow down execution of individual threads, since a thread ready to execute without stalls will be delayed by instructions from other threads.

c) Briefly describe the advantages and disadvantages of simultaneous multithreading (SMT).

Answer

Advantages:

It gives better utilization of machine resources.

Disadvantages:

SMT can decrease performance if any of the shared resources are bottlenecks for performance.

Software developers need to test whether SMT is good or bad for various applications and insert extra logic to turn it off if it decreases performance.

Security cocern.

d) For each type of parallelisms: instruction-level parallelism, data-level parallelism, and thread-level parallelism, list at least 3 different kinds of applications that will benefit from.

Answer

Instruction-level Parallelism

VLIW

Statically scheduled superscalar processors

Dynamically scheduled superscalar provessors

Data-level Parallelism

SIMD

Vector machines

GPU

Thread-level Parallelism

Hardware multithreading

Symmetric multiprocessors

Chip multiprocessing

Exercise 2

a. No loop-carried dependency

First consider B[2\*i+4] and B[4\*i+5]

In this case, (a,b,c,d) = (2,4,4,5).

GCD (a,c) = GCD (2,4) = 2, d – b = 1

2 does not divide 1, no loop dependency on B.

Then, same index for A in S1 and S2, no loop dependency on A.

b. Dependences are as follows,

True dependences (RAW)

S1 to S2 on A

S3 to S4 on A

Output dependences (WAW)

S1 to S3 on A

Anti-dependences (WAR)

S1 to S2 on B

S2 to S3 on A

S3 to S4 on C

To eliminate the output dependences and output dependences, change the code as follows,

for (i = 0; i < 100; i++) {

X[i] = A1[i] \* B1[i];

B[i] = X[i] + c;

A[i] = C1[i] \* c;

C[i] = D[i] \* A[i];

}

c. Yes, dependency exists between S1 and S2 on B and the loop is not parallel.

To make it parallel, change the code as follows,

A[0] = A[0] + B[0];

for (i = 1; i < 100; i++) {

A[i] = A[i] + C[i - 1] + D[i - 1];

B[i] = C[i - 1] + D[i - 1];

}

B[100] = C[99] + D[99];