**体系结构 第十一章**

**陈彦帆 2018K8009918002**

1. UCA (Uniform Cache Access 一致cache访问) 是集中式共享结构，多个核之间通过总线或者交叉开关连接核间共享的Cache。特点是，所有处理器核对共享cache的访问延迟相同。但是可扩展性受限。

NUCA （Non-Uniform Cache Access非一致cache访问）是分布式共享结构。每一个处理器核都拥有独立Cache和共享Cache。处理器核可以通过片上互连访问其他处理器核的共享cache，有不同的访问延迟。特点是可扩展性良好，支持的核数大。

2.

P2

3: Y=1

4: print X

P1

1: X=1

2: print Y

顺序一致性：有6种正确的执行顺序，结果X=1, Y=1。对于不同的顺序：

1234 P1,P2分别打印 0 1

1324 P1,P2分别打印 1 1

1342 P1,P2分别打印 1 1

3412 P1,P2分别打印 1 0

3142 P1,P2分别打印 1 1

3124 P1,P2分别打印 1 1

弱一致性：每个处理器的各自的赋值和打印没有相关，故1234的全排列共24种执行顺序全部正确。

结果X=1,Y=1, 可能的打印结果为0,0; 0,1; 1,0; 1,1

3.

Modified状态表示当前Cache行被当前处理器核独占并且已经被修改过了。如果没有M状态，处理器想替换独占状态的cache行，需要将该cache行写回内存。如果有M状态，只需要把M状态的cache行写回内存即可，减少了写回次数。

4.

Fetch\_and\_Increment：

void acquire()

{

while(Fetch\_and\_Increment(lock)>1);

}

void release()

{

lock = 0;

}

Compare\_and\_Swap：

void acquire()

{

while(compare\_and\_swap (lock,0,1)==0);

}

void release()

{

lock = 0;

}

5. 多线程之间不需要共享的变量的地址应保持至少1个cache行的间隔。

如两个线程想分别访问两个int类型的全局变量。

不应写成 int a[2]; //访问a[0]和a[1]

而应写成 int a[2][16]; //访问a[0][0]和a[1][0]

6.

一个物理通道可包含多个虚通道，每个虚通道包含独立的缓冲区和状态标识，物理通道选择一个虚通道的数据进行传输。当某个虚通道阻塞时，可切换到其它虚通道进行传输，提高整体吞吐率。片上互联网络中，虚通道分配用来分配流控单元在缓冲队列的位置。虚通道的设置，可以避免死锁。

7. **Fermi:**

**Registers** (Not shared)

Each SM has 32K of 32-bit registers. Each SP has its own registers.

Latency: Maybe 1-2 cycle

Bandwidth: about 8,000 GB/s.

**L1+Shared Memory** (Shared)

On-chip memory that can be used either to cache data for individual threads.

Total 64 KB.

Latency: 45 cycles

Bandwidth 1,600 GB/s

**L2 Cache** (Shared)

768 KB unified L2 cache, shared among the 16 SMs.

Latency: 310 cycles

Bandwidth: Unknown

**Global memory** (Shared)

Accessible by all threads as well as host (CPU).

Latency 680 cycles

Bandwidth: 192GB/s