



Anton应用——MD模拟

——国防科大2020年高性能评测与优化课程小组讨论

汇报人：陈凤婷（19023042）

组 员： 夏锦涛（19020135） 肖滔（19023089）

指导老师：龚春叶、甘新标、杨博

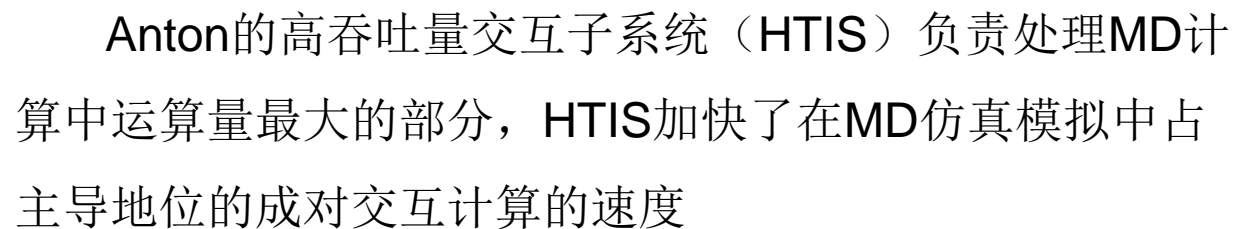


What is Anton and its application



Anton是一台大型并行专用超级计算机，主要用于执行毫秒级的分子动力学(MD)模拟，可以将MD模拟加速几个数量级。

Anton可以使科学家们和药物开发人员在实验室实验中无法观察到的重要现象可视化。



Anton进行MD模拟时的主要计算任务如图，其中HTIS
主要进行阴影部分的计算任务



HTIS（高吞吐量交互子系统）

MD模拟中的大部分计算都涉及到范围有限的成对粒子的相互作用，主要有HTIS来负责。

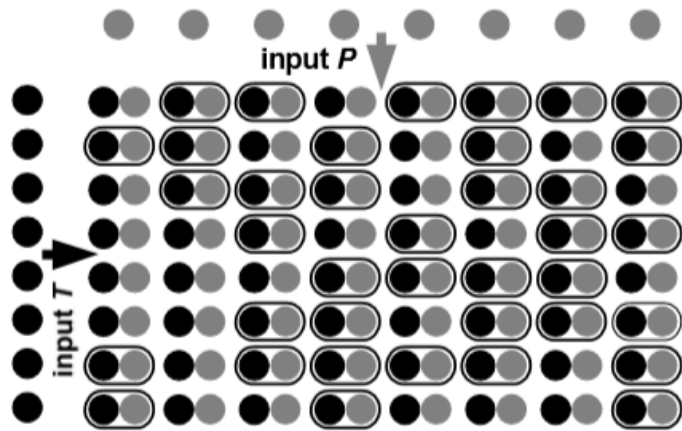
HTIS使用32个成对点交互流水线（**PPIPs**）阵列实现毫秒级仿真所需的巨大算术吞吐量，Anton的32个PPIPs可以支持超过10Tbit/s的数据总输入加输出速率。

简单的在HTIS和其他子系统之间传输流水线的输入输出有带宽限制，**DPSR**（**D**irect **p**roduct **s**election **r**eduction operation）解决了这个问题



DPSR (Direct product selection reduction)

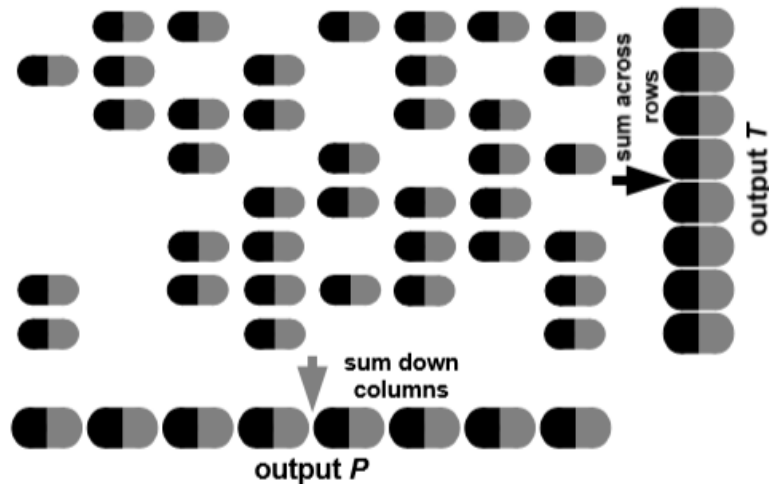
Step 1. Direct product generation and selection



第1步，应用选择条件，选择集合P的一个元素和集合T的一个元素组成对。满足选择条件的对被圈起来。

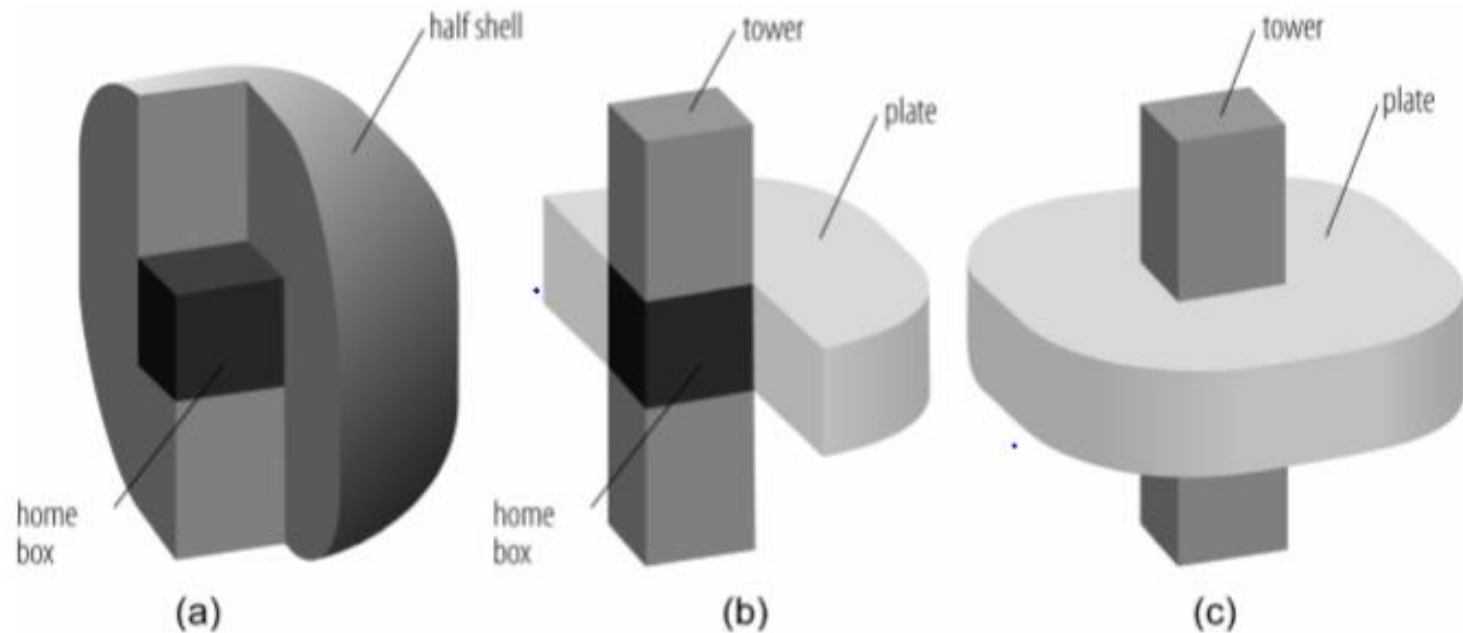
第2步，应用函数，对列和行求和，产生不同的输出 $|P|+|T|$ 。

Step 2. Function evaluation and reduction





DPSR (Direct product selection reduction)



为Anton使用一种新颖的并行化方法来处理范围有限的两两交互作用，即NT方法

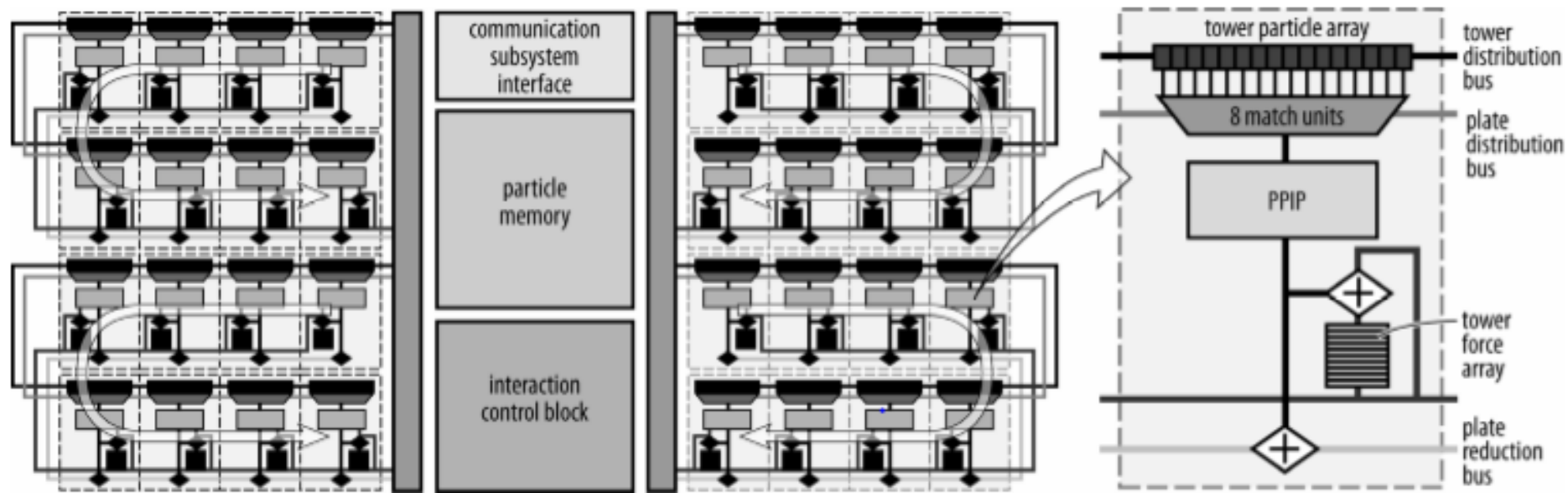
(a)在传统的空间分解方法中，每个节点导入半壳区域内的粒子，使其与主盒内的粒子相互作用。

(b)在NT方法中，每个节点计算塔区粒子与板区粒子之间的相互作用。

(c)在用于网格点-粒子相互作用的NT方法的变体中，平板区域更大。



Architecture



高通量交互子系统 (HTIS) 和单对点交互模块 (PPIM) 的细节。箭头表示粒子分布和减力网络中数据流的方向。



1. Particle Memory and Interaction Control Block

1、Particle Memory

交错、多端口的RAM结构，使HTIS操作和芯片其他部分分离

2、ICB (Interaction Control Block)

增强了HTIS特定指令通用处理器，控制HTIS中的所有配置和数据移动。

指令发布到硬件控制的队列中，该队列控制着与粒子存储器缓冲区之间的数据传输。

ICB可以在粒子分配网络之前并且在关键计算路径之外进行操作。



2. PPIM Array and Particle Distribution Network

1、PPIM (pairwise point interaction module) Array

由32道流水线和支持逻辑组成，可以计算MD所需的选择、交互和缩减。

塔点在流水线之间循环分布，板点广播到每个流水线，在流水线里与塔点配对

选择条件筛选这些对，然后对筛选出来的对进行计算交互



2. PPIM Array and Particle Distribution Network

2、粒子分布网络

将PPIMs组织成4个链，每个链8个PPIMs，实现塔点分布和板点广播。

板点广播到每个链的顶端，然后通过链流，每个时钟周期访问一个PPIM

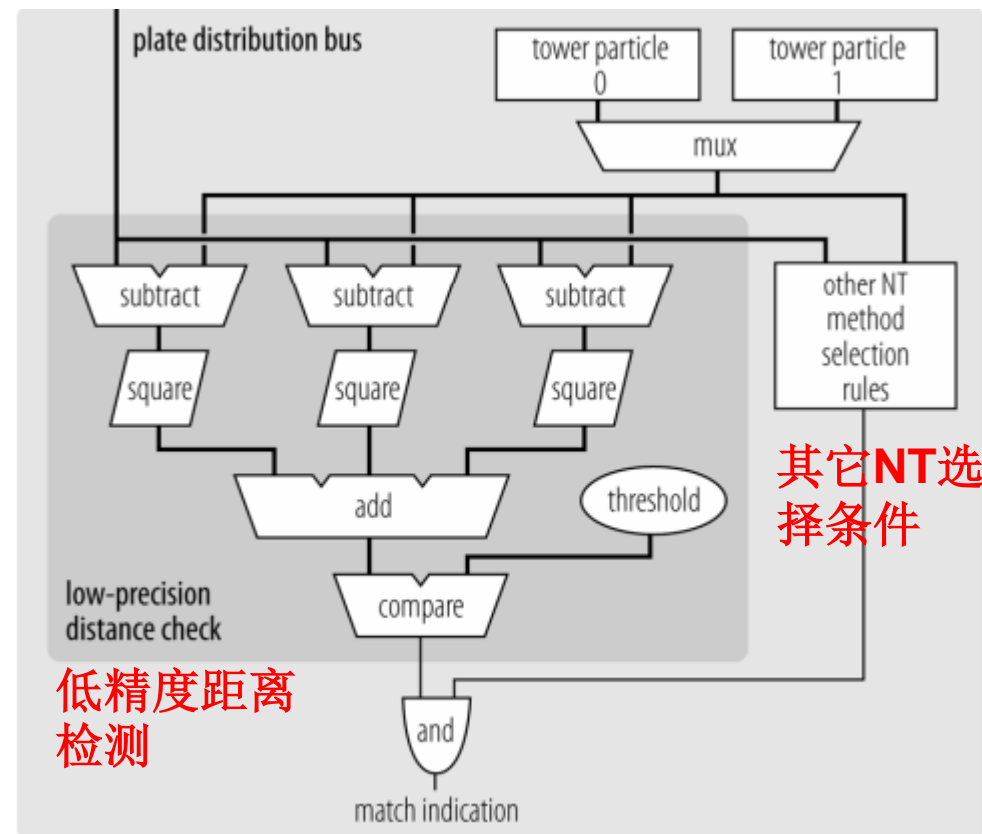
塔粒子阵列的有限容量限制了DPSR大小上限，该上限是HTIS可以一次性计算的数值

将DPSR分解为更小的DSRs，DSRs通过HTIS进行流水线操作



3. March Units

- 输入PPIP路径数据之前，先对数据进行筛选。（如右图）
- 减少乘法器的数量，减小了占March units的面积。
- 可以避免粒子间相互作用重复计数。





4. Pairwise Point Interaction Pipelines(PPIPs)

- 1.PPIP实际上就是流水线式执行，工作频率800Mhz，每个周期可以产生一个结果。
- 2.它的输入是两个矢量（点的空间位置），还有六个相关的参数。
- 3.它的输出可以是矢量，也可以是标量。计算过程就是March units一样。
- 4.所有的数字都是定点，左侧隐含小数点。

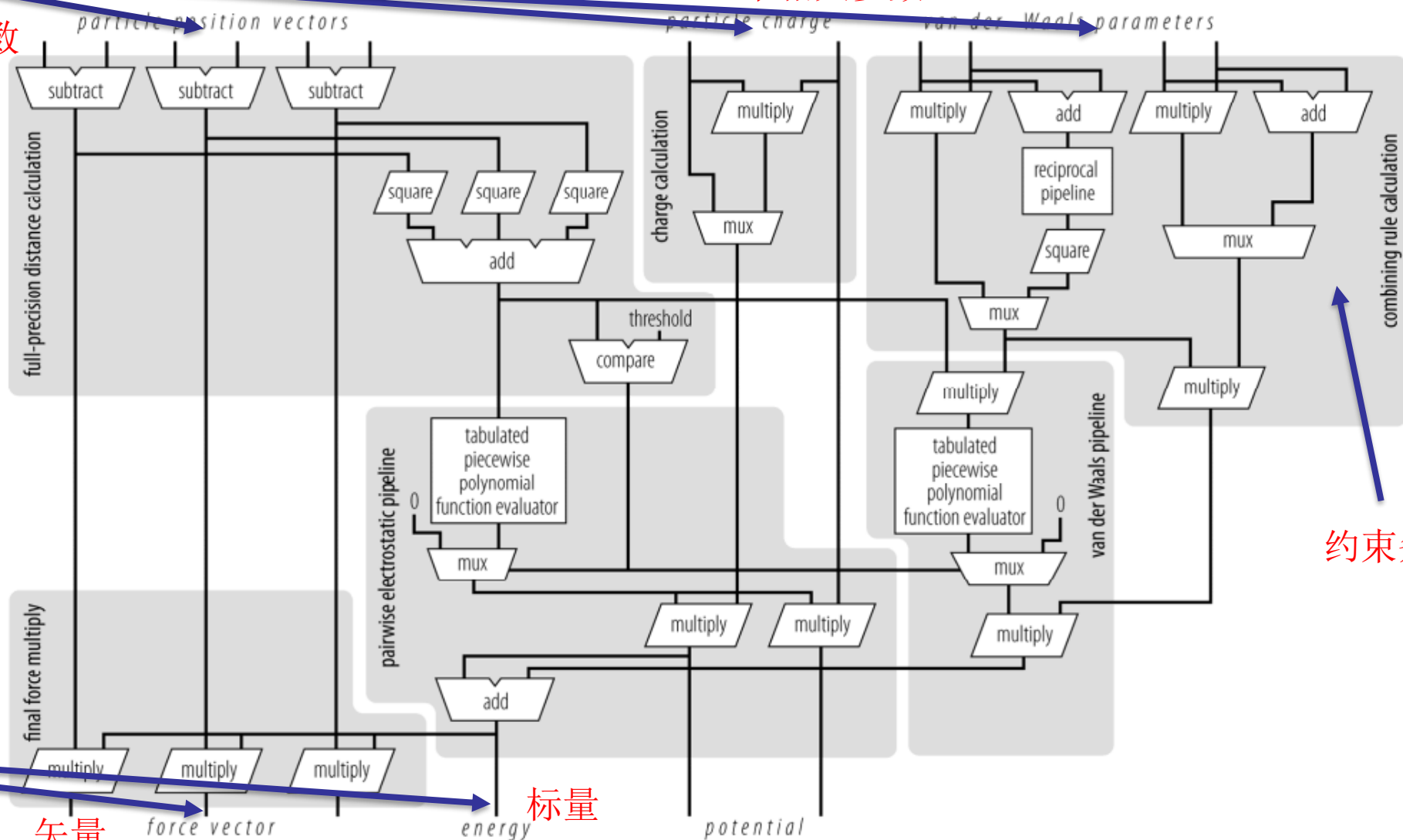


4. Pairwise Point Interaction Pipelines(PPIPs)

输入:

位置参数

6个相关参数



全精度举例计算

约束条件计算

输出:

矢量

force vector

标量

energy

potential



5. Force Reduction Network and Tower Force Array

1. PPIP数组作为一个整体，为每个输入的点产生一个缩减的输出。
2. 减少的塔力累积在塔力数组中，其条目与保存塔数据的塔粒子数组一一对应。由PPIP计算出的每一个力都累积到塔力数组中与参与相互作用的塔粒子相对应的一个槽中。
3. 板力的减小发生在力的缩减网络上。每个流经PPIM的板粒子都会产生一个结果，即通过与PPIM的塔粒子数组中的粒子相互作用而产生的力的总和。这些结果进一步累积在力减少网络中，为每个板粒子产生单个总力。



6. Performance

实验将Anton的性能与化学体系中其他许多MD平台的性能进行了比较。

对比对象：

1.NAMD被广泛认为是高并行度商品集群中速度最快的MD软件

2.Desmond是本文团队开发的一个软件包，它包含了在设计Anton的过程中发现的算法

NAMD和Desmond都运行在一个2.4GHz的Opteron处理器集群上，该集群由无限带宽网络连接。

NAMD在256个处理器核上实现了最大性能，而Desmond在512个处理器核上实现了最大性能。

3.Blue Matter MD代码的Blue Gene/L超级计算机在这种规模的化学系统中可以扩展到多达16384个处理器核

4.MDGRAPE-3（MDGRAPE项目中MD的最新专用ASIC）

5.GROMACS被广泛认为是可用的最快单处理器MD代码之一。

由于本文的重点是HTIS，因此本文试图量化它相对于其他MD平台上的等效交互计算的性能。但是由于几个原因，没法直接比较。为了进行比较，不得不使用来自不同机器配置、不同化学系统和不同模拟参数的近似值。



6. Performance

Table 2. Performance of various MD platforms on the DHFR system. All simulations use the benchmark parameters specified in Table 2 of Shaw et al. [23]. Data in the third column are restated from Figure 6 of Shaw et al. [23], except for the GROMACS entry, which reflects a more recent measurement on a 3.2 GHz Xeon core. Many of the entries in the fourth column represent rough estimates, as described in the text. Dates reflect when the measurement was made, which may differ from the date the machine commenced operation.

System	Configuration	Time to complete a time step (μ s)	Time to complete particle-particle force computation (μ s)
Anton (2008 estimate)	512 ASICs	19	2.4
Desmond on Cluster (2006)	512 cores	1,400	361
Blue Matter on Blue Gene/L (2006)	16,384 cores	1,700	850
NAMD on Cluster (2006)	256 cores	6,300	N/A
MDGRAPE-3 (2003)	12 ASICs	26,000	7,900
GROMACS (2007)	1 core	181,000	111,000



6. Performance

System	Peak performance (interactions/ μ s/chip)
Anton (single ASIC)	25,600
MDGRAPE-3 (single ASIC)	1,500
Dual core Xeon	122

表3列出了三种不同平台的峰值性能（以每个芯片每微秒的粒子-粒子相互作用测量）。

专用硬件支持可以解决各种深度（每个Anton PPIP 28级深，每时钟周期执行大约50个算术运算）和宽度（Anton每个ASIC有32个PPIPs）。这导致了片上并行性远远超出了通用可编程系统的能力。两种专用解决方案（Anton和MDGRAPE-3）的峰值性能都比运行在通用处理器上的软件高出几个数量级。



6. Performance

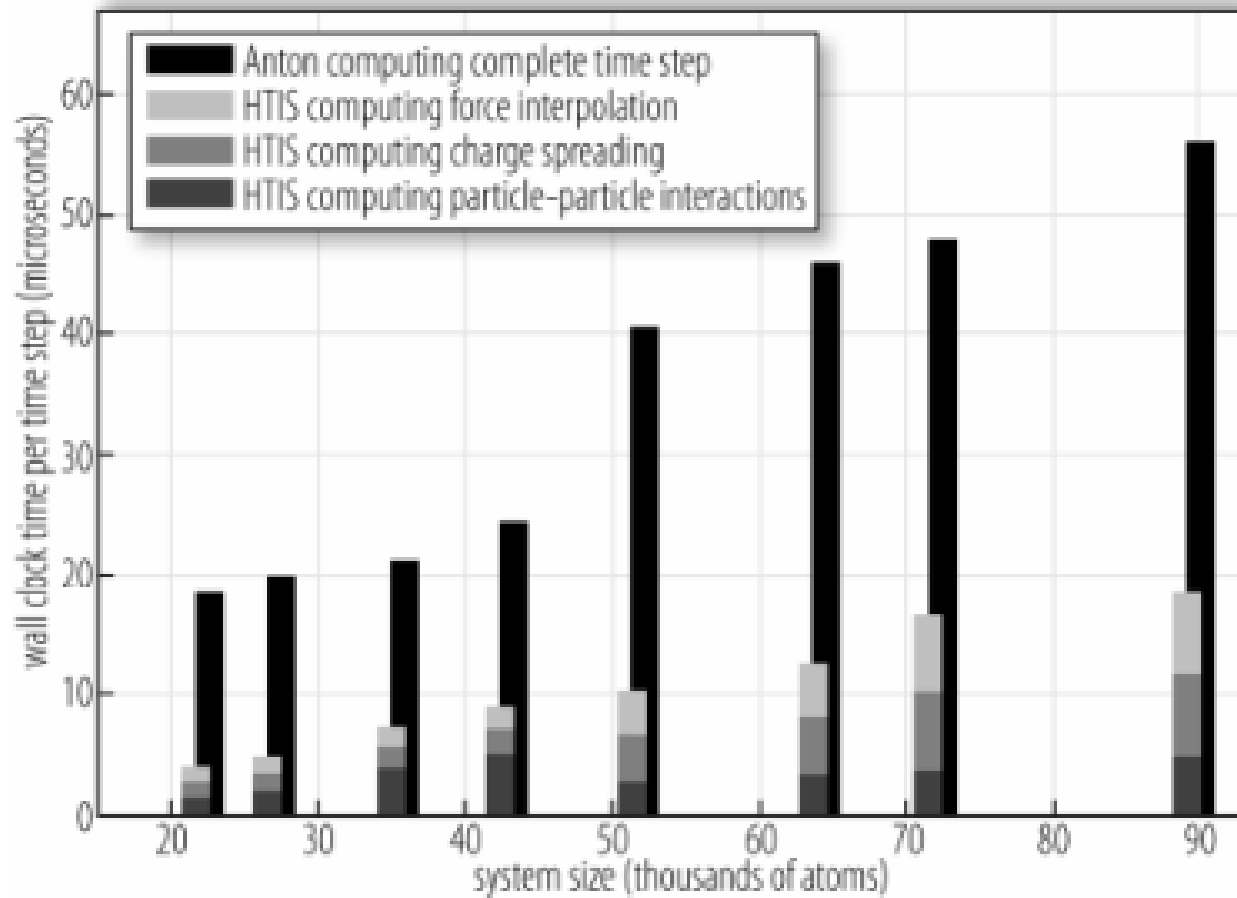


图7显示了从20000到100000个原子的化学系统的单个时间步的时钟时间的分解。将整个时间步长的时钟时间和各种HTIS任务绘制为化学系统中原子数量的函数。很明显，Anton的性能在这个系统大小范围内实现了线性缩放。系统平衡在很大程度上也与问题大小无关；



Question?

HTIS有多少个PPIP?



End, 感谢观看!