

Паралелна производителност на високопроизводителни компютърни системи

1. Актуалност на проблема

Суперкомпютрите са изчислителни машини, достигащи днес скорост от порядъка на десетки и дори стотици трилиона изчисления с плаваща запетая за една секунда. Базиран на най-новите постижения на компютърните технологии, те се използват предимно в областта на военната промишленост, научните изследвания, космическите и геоложките проучвания, т.е. сфери, където са необходими изключително бързи и сложни изчисления за съставянето на математически модели и провеждането на комплексни симулационни процедури. Те се създават за решаване на задачи, които изискват огромни количества изчисления в секунда. Такива например са проблемите от квантовата механика, метеорологичните прогнози и изследванията на климата, молекулярното моделиране (структурите и свойствата на химичните съединения, биологичните макромолекули, полимерите и кристалите), криптоанализа, физичните симулации (например на самолет във въздушен тунел, на взривяване на ядрени оръжия, на ядрен синтез и др.). Основните потребители на суперкомпютрите са големите световни университети, научно-изследователските институти и военните ведомства и лаборатории.

Суперкомпютрите притежават огромна изчислителна мощност – многократно по-бързи са от най-големите Електронноизчислителни машини (ЕИМ). Основна характерна черта при тях е използването на паралелна обработка на данните. Множество процесори разделят изследвания проблем на отделни подзадачи, върху решаването на които те работят едновременно. В същото време традиционните ЕИМ използват последователна обработка, т.е. решаване на задачи една след друга. Работата на суперкомпютрите е изключително сложна и поради обстоятелството, че са необходими специално разработени за целта компютърни програми, които да осигурят разделянето на основния проблем на отделни подзадачи и тяхното разпределение и едновременно решаване на различни процесори.

Всички тези обстоятелства повишават значително финансовите разходи по създаването и прилагането на суперкомпютрите, но бързото развитие на компютърните технологии през последните години доведе до намаляване на тяхната цена. С огромното технологично развитие в способността да увеличават капацитета за обработка и изчисление, суперкомпютрите стават все по-мощни. Те променят начина, по който учените правят своите открития и изчислителни експерименти. Помагат за

ускоряване на времето за получаване на резултатите от година на часове или дори на минути.

2. Еталонни програми за оценка на производителността (Benchmarks)

В компютърните науки, бенчмаркът представлява процес на изпълнение на компютърна програма, набор от програми или други операции, с цел да се оцени относителното представяне на даден обект, като обикновено се изпълняват набор от стандартни тестове и опити на него. Терминът бенчмарк се използва най-често като събирателно понятие на всички софтуерни продукти, използвани за оценка и измерване на производителността. Използването на бенчмаркове най-често е насочено към оценка на характеристиките на производителността на компютърния хардуер, като например изчислителни операции с плаваща запетая за процесора. Тези тестове предоставят метод за сравняване на производителността на различни подсистеми в различни чипове или системни архитектури. Бенчмарковете са създадени да имитират определен вид натоварване върху даден компонент от системата. Синтетичните бенчмаркове например постигат това чрез специално създадени програми, които имитират натоварването върху дадения компонент, докато приложните бенчмаркове изпълняват истински програми на системата. Приложните бенчмаркове предоставят много по-добро измерване на реалната производителност на дадената система за разлика от синтетичните бенчмаркове, които пък са полезни и се използват за тестване на индивидуални компоненти, като например твърдия диск или мрежовото устройство.

Тестването с бенчмаркове не е лесна задача и често включва няколко цикъла на тестване, за да се достигне до правдоподобно заключение относно тествания обект. Интерпретацията на информацията, получена от бенчмарковете, също е много трудна задача. Съществуват редица предизвикателства, свързани с използването на бенчмаркове:

- Подправяне на резултати – често производителите на компютърен хардуер показват тестове с много по-високи стойности на резултатите, но в действителност стойностите са доста по-ниски при истинското работно натоварване;
- Много бенчмаркове са фокусирани изцяло върху определяне на скоростта на представяне, пренебрегвайки други важни аспекти на компютърните системи като:
 - Качеството на услугата – примери за това са неизмерените качества на услугите, включително защитата, достъпността, надеждността, цялостното изпълнение, и др.

Често има неясни неща, относно тези качества на услугата, въпреки че всички те са от изключителна важност за работата на системите в бизнеса. *Transaction Processing Performance Council Benchmark* частично изследва тези аспекти като специфицира ACID свойствени тестове, правила за мащабиране на базите данни и ограничения за нивото на услугите;

- Бенчмарковете не измерват крайната цена – *Transaction Processing Performance Council Benchmark* частично засяга този проблем като определя, че зависимостта цена/производителност трябва да бъде съобщавана като допълнение на грубата производителност, използвайки опростена ТСО формула. Но разходите са само частични и разпространителите са известни с това, че поставят цени единствено въз основа на бенчмарка, като така създават специфична конфигурация от цена/производителност, която е изкуствено занижена от към цената. Дори малко несъответствие на бенчмарк резултатите на пакета рефлектират върху много по-висока цена в реалността;
- Ограничения на пространство, мощност и охлаждане – когато се използва по-голямо количество мощност, преносимото устройство ще има по-малък живот на батерията и ще налага по-честото ѝ презареждане. Сървър, консумиращ повече енергия и/или пространство, може да се окаже, че не е в състояние да се помести в съществуващите ресурсни ограничения, включително ограниченията за охлаждането;
- В някои вградени системи, където паметта е значителен разход, по-добрата плътност на кода би могла да намали разходите.
- Бенчмарковете изпитват трудности с адаптирането към широко дистрибутирани сървъри, особено с тези с по-голяма чувствителност към мрежовите топологии;
- Потребителите биха могли да имат доста различни възприятия за производителността от данните, предоставени от бенчмарка. Потребителите предпочитат предвидливостта – сървъри, които винаги отговарят или дори се представят по-добре от колкото са очаквали. Резултатите от бенчмарковете наблягат повече на осреднените стойности (от перспективата на IT), вместо на времената от резултатите при най-лошите възможни случаи (от перспективата на работата на системата в действителност) или ниските отклонения (от перспективата на потребителя);

- Много от бенчмарковете се фокусират само върху едно приложение или определена редица от такива. Повечето информационни центрове имплементират възможности за виртуализация по ред причини и бенчмарковете все още се опитват да се справят с реалността, в която множество приложения и редици от такива се изпълняват съгласувано на консолидирани сървъри;
- Институти, изследващи еталонни програми за оценка на производителността, често пренебрегват или изобщо не следват основните научни методи. Това включва липсата на контрол на променливите, ограничената повторяемост на резултатите, малките размери на входните данни, използвани в примерните тестове и др.

3. Видове еталонни програми за оценка на производителността

- Реални програми
 - Софтуер за текстова обработка
 - CAD системи
 - Потребителски софтуер
- Микроеталони
 - Измерва се производителността на точно определени компоненти чрез много малки специфични части от програмния код
- Ядра
 - Съдържат цифрови кодове
 - Обикновено взети от съществуваща програма
 - Подходящи за фокусиране върху точно определен аспект
 - Резултатите са в MFLOPS
- Синтетични
 - Процедура за програмиране на синтетичен еталон
 - Статистика за всички видове операции от множество потребителски програми
 - Определяне съотношението на всяка една операция
 - Създаване на програма, базирана на по-горното съотношение
- Входно-изходни
 - Измерват производителността на входно-изходната система
- Бази данни

- Измерват проводимостта и времената за отговор на управляващите системи
- Паралелни
 - Използва се на машини с множество процесори или системи, изградени от множество машини

4. Задачи за изпълнение

- 4.1. Анализирайте списъците със суперкомпютри TOP500 и GREEN500 и отговорете на следните въпроси:
 - Какви бенчмаркове се използват за оценка на производителността?
 - Каква е разликата между два списъка и колко често се актуализират?
- 4.2. Направете графика, която отразява развитието на най-мощните суперкомпютри в TOP500 за времето от 2006 г. до днес. Анализирайте получената графиката. Какви са тенденциите при архитектурите на тези машини?
- 4.3. Направете графика, която отразява развитието на най-енергоефективните суперкомпютри в GREEN500 за времето от 2014 г. до днес. Анализирайте получената графиката. Какви са тенденциите при архитектурите на тези машини?
- 4.4. Изтеглете от Интернет последната версия на бенчмарка Linpack Xtreme. Стартирайте бенчмарка и оценете производителността на компютърната система, която използвате в момента. Сравнете получения резултат с този на колегата до Вас. Има ли разлика в получените резултати и ако да, обяснете от къде идва тя.