

CPU를 이용한 ERP 시스템 벤치마크 성능 분석

나승배*

*고려대학교 컴퓨터과학기술 대학원 컴퓨터공학과
e-mail: nsbhit@korea.ac.kr

Analysis of Benchmark Performance for ERP System using CPU

Seung-Bae Na*

*Dept of Computer Science, Korea University

요 약

ERP 시스템 구축을 위해 기업의 환경에 맞는 최적의 성능을 발휘하는 시스템을 선정하여야 한다. 이를 위해 수많은 고려사항이 있을 수 있으나 본 논문에서는 ERP 구축 패키지인 SAP를 사용하는 것을 전제로 하여 SAP 표준 벤치마크 테스트를 통해 시스템 성능 향상에 상대적으로 많은 영향을 미치는 CPU의 증가가 성능향상에 비례하지 않음을 실험을 통해 검증한다.

1. 서론

ERP 시스템 벤치마크에 대한 이해 부족으로 인하여 ERP 시스템 선정 시 제안된 시스템에 대한 적절한 평가가 이루어 지지 않고 있으며 이로 인해 잘못된 벤치마크 결과를 가지고 시스템을 선정하는 경우가 빈번하게 발생하고 있다.

각 하드웨어 벤더는 신제품이 출시되었을 때 자사의 시스템이 타사의 시스템보다 ERP시스템 구축에 최상의 성능을 발휘함을 입증하기 위해 SAP 벤치마크위원회에서 제시된 방법론에 따라 공인된 벤치마크 테스트를 실시하여 제시하고 있다.

통상적으로 각 하드웨어 벤더는 벤치마크를 수행한 결과 중 최상의 결과만을 가지고 CPU의 용량 산정

을 하는 경향이 있다. 예를 들면 자사의 CPU 성능이 20CPU일 경우 20,000 SAPS(SAP Application Benchmark Performance Standard)의 결과와 30CPU일 경우 27,000 SAPS의 결과가 있을 때 대부분의 경우 최상의 기록을 토대로 제안 시스템의 용량 산정을 추정하고 있는 실정이다. 즉, 위의 경우에서 시스템 도입사의 요구가 40,000SAPS인 경우 20CPU일 때 1CPU당 산술적으로 1,000SAPS이므로 이를 기준으로 시스템을 제안하여 40,000 SAPS의 성능을 내기 위해서는 CPU의 수가 40개가 필요하다는 논리로 시스템을 제안하고 있다. 이러한 하드웨어 벤더들의 산정방식에는 ERP시스템을 도입하려는 고객들의 판단을 흐리게 할 수 있는 몇 가지 함정이 있다. 벤치마크 테스트 대상 시스템이 현재 고객이

도입하려는 시스템과 동일 기종 및 동일 환경이 아닐 가능성이 매우 높다. 그러나 하드웨어 벤더들은 최상의 기종과 최상의 구성으로 벤치마크 테스트된 결과를 가지고 다른 시스템 및 구성에도 동일하게 적용하고 있는 실정이다.

벤치마크 테스트 시스템과 실제 제안된 시스템과는 여러 소프트웨어 및 하드웨어 구성이 다르기 때문에 참고는 할 수 있을지라도 하드웨어의 성능을 판단하는 절대적인 기준이 될 수는 없다.

이에 본 논문에서는 SAP 시스템 구축 시 하드웨어 벤더의 제안된 시스템들에 대하여 SAP사의 표준 벤치마크 테스트 툴을 이용하는 것을 권고한다. 이는 공정한 평가 및 성능평가가 적절히 이루어 질 수 있는 바탕이 된다. 또한, CPU의 수량의 증가가 해당 성능을 100%반영하지 않음을 실험을 통해 검증한다. 본 논문의 구성은 다음과 같다. 2장은 관련 연구로서 기존 벤치마크 테스트에 대해 소개하며 3장에서는 SAP 표준 애플리케이션 벤치마크 프로그램에 대한 설명을 4장에서는 실험 환경 및 결과 분석을 마지막으로 5장에서는 본 논문을 요약하고 결론을 맺는다.

2. 관련 연구

벤치마크 테스트는 하드웨어의 성능 평가 및 비교에 사용되는데, 각 하드웨어 공급 업체나 데이터베이스 관리 시스템 벤더들의 입장에서는 경쟁 시장에 자사 제품의 성능을 공개하는 요소로 벤치마크 결과를 사용하기 때문에 그 결과는 매우 민감한 요소가 되어 있다.

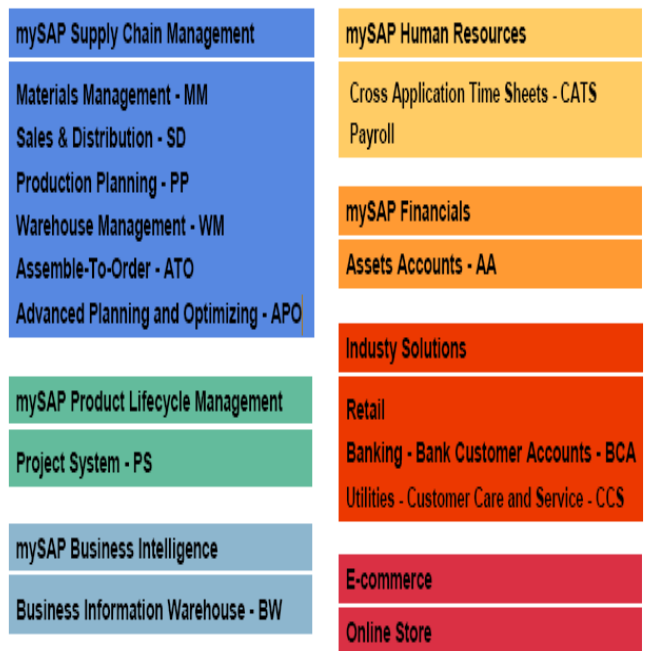
지속적으로 벤치마크에 대한 연구 활동이 이루어지고 있는데, TPC(Transaction Processing Performance Council), SPEC(Standard Performance Evaluation Corporation), 그리고 BAPco(Business Application Performance Consortium)등이 표준 벤치마크를 주도적으로 이끌고 있는 기구들이다[1]. TPC는 주로 트랜잭션 처리 부문과 데이터베이스 부문에 관련한 벤치마크 방법을, SPEC은 과학 분야와 워크스테이션 컴퓨터 위주의 벤치마크 방법을, 그리고 BAPco는 클라이언트-서버 애플리케이션을, 이용한 PC위주의 벤치마크 방법을 이끌고 있다[2].

3. SAP 표준 애플리케이션 벤치마크

SAP 표준 벤치마크는 SAP 제품에 대한 테스트와 확장성에 대한 안정성을 입증하기 위해 고안되었으며 표준 비즈니스 시나리오를 반영하고 성능과 관련된 파라미터들을 측정하기 위해 사용되고 있다.

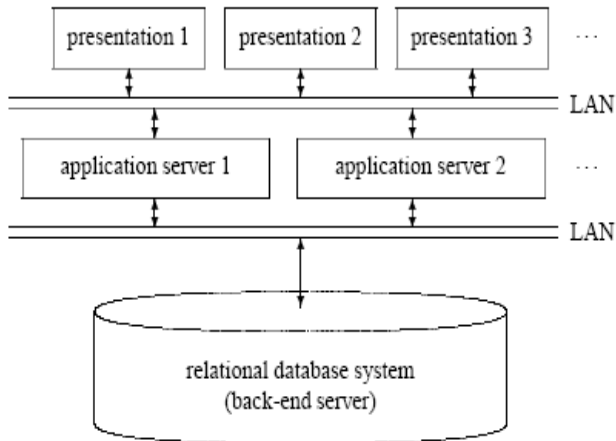
SAP 표준 애플리케이션 벤치마크는 가장 전형적인 트랜잭션들을 시뮬레이션 할 수 있는 수많은 스크립트 파일로 구성되어 있다. 벤치마크 실행 중에 SAP 시스템의 모든 영역 예를 들면, CPU, 메모리, I/O 시스템, 네트워크 부하, 기능상의 문제점 및 시스템의 가용성 등이 테스트된다. 더 자세히 살펴보면 벤치마크의 주요한 판단기준은 데이터베이스에 대한 응답시간, 대기 시간, CPU 사용률, 특정수의 벤치마크 사용자에게 의한 평균 다이얼로그 응답시간 및 처리량이라고 할 수 있다. SAP 표준 벤치마크를 통해 IT 솔루션에 적합한 하드웨어 구성을 할 수 있는 장점이 있다. [3]

SAP 표준 벤치마크의 종류로는 그림 1과 같이 OLTP (Online Transaction Processing)인 SAP R/3 시스템에 대한 테스트로서 MM, SD, PP, WM, ATO 등이 있으며 OLAP (Online Analysis Processing)중에 대표적인 제품인 BW (Business Information Warehouse)시스템에 대한 벤치마크가 있다.



(그림 1) SAP 표준 벤치마크 테스트의 종류

표준 벤치마크 환경에는 2-Tier 와 3-Tier환경으로 구성된다. 2-Tier는 데이터베이스와 SAP 애플리케이션이 같은 하드웨어 상에 존재하는 구성이며 그림 2는 3-Tier구조로서 데이터베이스와 애플리케이션이 별도의 하드웨어 상에 존재하는 구성을 말한다.



(그림 2) 3-Tier 클라이언트/서버 구조

SAP 벤치마크의 가정은 다음과 같다. 벤치마크 유저의 Think time은 10초, 최대 응답시간은 2초 이내로 정의한다. Think time이란 다음 화면전환(다이얼로그 스텝)까지의 대기시간을 말한다. 벤치마크 시뮬레이션 사용자는 1분에 5번의 화면전환(다이얼로그 스텝)을 수행한다. 한명의 벤치마크 유저는 5SAPS의 부하를 가진다.

SAP R/3 시스템의 성능은 SAPS 라는 단위로 정의 되는데 100 SAPS는 SD 벤치마크에서 시간 당 6,000번의 화면 전환(다이얼로그 스텝)과 2,000개의 주문항목전기 (Order Line Item Posting)를 수행하는 정도의 수치이다.

SD 벤치마크는 SAP의 표준 벤치마크이며, 현재까지 가장 많이 사용되고 있는데, 그 이유는 SD 벤치마크 테스트가 CPU에 가장 많은 부하를 주어 가장 정확한 시스템 성능 결과를 얻을 수 있기 때문이다.

그림 3은 SD 벤치마크테스트의 수행 순서를 나타낸다. 총 18개의 다이얼로그 스텝(화면 전환)으로 이루어져 있으며 다이얼로그 스텝 2번에서 16번까지 1루프에 해당된다. 즉 1루프는 총 15번의 화면전환(다이얼로그 스텝)과 5번의 주문항목전기로 구성된

다.[3]

TA	Explanation	#DS	#Postings
.	Logon	2	
VA01	Create Customer Order	4	1
VL01	Create a delivery	3	1
VA03	Display Customer Order	2	
VL02	Change delivery	2	2
VA05	List orders	2	
VF01	Create invoice	2	1
END	Logoff	2	

(그림 3) SD 벤치마크 테스트 시나리오

4. 실험환경 및 결과 분석

실험은 CPU의 수량 증가에 따른 전체 성능의 변화를 알아보기 위하여, 다른 시스템 구성요소는 일정하게 유지하면서 1,000MHz CPU를 4CPU, 8CPU, 12CPU로 장착한 각각의 동일기종의 서버에서의 CPU수량 증가에 따른 성능 변화를 측정한다. 또한 SAP 표준 벤치마크 테스트 시나리오 중 범용적으로 사용하고 있는 SD벤치마크 테스트를 실시한다.

실험 시스템들은 모두 데이터베이스 오라클 9.2.0.6 버전을 사용하고 SAP R/3 버전은 Enterprise 4.71이며 2-Tier 테스트 환경으로서 데이터베이스와 애플리케이션이 한 서버에 구성된다. 표준 벤치마크 테스트 툴로는 각 클라이언트에 100명의 사용자만 테스트 가능하므로 테스트 대상 클라이언트를 900번부터 909번 클라이언트까지 총 10개를 생성하여 동시에 최대 1,000명의 유저가 실제로 SD벤치마크 테스트를 수행할 수 있도록 구성한다. 루프는 10회 실시한다. 각 환경에서 벤치마크 유저의 수를 증가시키면서 표준 벤치마크 테스트의 제약 사항인 평균 응답 시간 2초미만을 만족하고 CPU의 사용률이 최대가 되도록 각 환경에서 벤치마크 유저수를 적절히 조절하여 시간당 최대 다이얼로그 스텝을 처리한 결과를 비교대상으로 한다.

표 1은 실험 결과를 나타낸다. 4CPU일 때 동시 유

저 500명을 10번의 루프 처리한 결과 시간당 다이얼로그 스텝은 164,558이었으며 평균 응답시간은 약 2초의 결과가 나타났다. 이때 CPU 사용률은 100%였다. CPU당 SAPS를 산출해보면 4CPU일 때 164,558 다이얼로그 스텝을 처리하였고 100 SAPS는 시간당 6,000 다이얼로그 스텝이므로 164,558다이얼로그 스텝을 6,000다이얼로그 스텝으로 나누면 27.43다이얼로그 스텝이며 여기에 100 SAPS를 곱한 결과 값은 2,745 SAPS이다. 이와 같은 방식으로 CPU당 SAPS를 산정해보면 4CPU는 686 SAPS, 8CPU일 때 570 SAPS, 12 CPU는 483 SAPS이다. 결과적으로 CPU의 수가 배로 증가한다고 해서 성능이 100% 반영되지 않음을 실험을 통해서 확인할 수 있다.

<표 1> CPU 수량변화에 따른 시간당 다이얼로그스텝 비교

유저 수	4CPU		8CPU		12CPU	
	DS/h	Avg.	DS/h	Avg.	DS/h	Avg.
375	143,626	0.583	146,591	0.389	133,660	0.834
500	164,558	2.112	140,700	0.195	197,622	0.221
750			273,586	1.040	297,510	0.252
950					347,826	0.936

실험 결과에서 벤치마크 유저의 수가 375명일 때는 CPU의 수에 상관없이 시간당 처리량이 CPU수가 많다고 해서 CPU의 차이만큼 늘어나지 않음을 알 수 있다. 이때 12CPU의 CPU사용률은 27%로 아주 낮은 수치를 보였다. 8CPU이고 750명의 벤치마크 유저를 사용했을 때 시간당 다이얼로그 스텝이 최대이고 12CPU이고 950명일 때 최대임을 알 수 있다. 각각의 경우 CPU 사용률은 90%이상을 보였다.

표 2에서 실험 결과 SAPS와 하드웨어 벤더가 적용하는 SAPS를 비교해보면 많은 차이가 있음을 알 수 있다.

<표 2> 실험 결과 및 기존방식의 적용 SAPS

CPU 수량	실험결과 SAPS	적용 SAPS
4개	2,745	2,745
8개	4,560	5,490
12개	5,796	8,235

5. 결론

Amdahl의 법칙에 따르면 전체 시스템의 성능 향상은 개선되는 일부 시스템의 구성요소가 얼마나 전체 성능에 기여하느냐에 따라서 결정된다. 예를 들어, 임의의 일부 시스템 구성요소가 50%의 성능향상을 보이더라도 전체 시스템에 있어서는 50%의 성능향상을 보장할 수 없다. 즉, 임의의 구성요소의 성능이 무한정 증가한다 하더라도 해당 구성요소의 전체 시스템 성능에 대한 기여도가 100% 보다 작은 경우 해당 구성요소의 개선에 따른 전체 시스템의 성능향상은 어느 한계를 갖게 된다.[4]

본 논문에서는 동일한 조건하에 CPU의 수량 증가에 따른 성능의 변화를 분석하였으며 CPU의 수량의 증가가 시스템 성능 향상에 비례하지 않음을 실험을 통해 증명하였다. 이를 통해 ERP 시스템 CPU 용량 산정 시 공인된 벤치마크 테스트 결과를 잘못 해석, 적용하는 오류를 줄일 수 있는 계기를 마련하였으며 또한 자사에 맞는 최적의 ERP시스템 선정을 위해 각 하드웨어 벤더가 제안한 시스템에 대해 본 논문에서와 같은 방식으로 벤치마크 테스트를 실시함으로써 제안 시스템들이 요구하는 성능을 충족하는지, 어느 시스템이 우수한지를 공정하게 평가하여 최상의 시스템을 선정하는데 도움이 될 것이다.

참고문헌

- [1] j. Gray, "The Benchmark Handbook for database and transaction ", 3rd. Morgan Kaufmann,1993
- [2] ETRI, "주기억 상주 실시간 DBMS의 고성능 실시간 트랜잭션 관리자에 관한 연구", 최종연구 보고서,1998
- [3] SAP AG SAP Standard Application Benchmarks <http://service.sap.com/benchmark>
- [4] Henning, J.L, " SPEC CPU2000: measuring CPU performance in the New Millennium" IEEE Computer , Volume: 33 Issue: 7 , July 2000, p28-35
- [5] Jochen Doppelhammer,Thomas H'oppler, Alfons Kemper, Donald Kossmann "Database Performance In Real World-TPC-D and SAP R/3" 1995