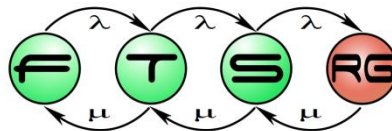


Benchmarking

**Budapest University of Technology and Economics
Fault Tolerant Systems Research Group**



Benchmarking

- Célok: szoftver/hardver eszközök teljesítményének összehasonlítása
 - Döntéstámogatás
 - melyiket előnyösebb megvenni/telepíteni
 - mekkora terhelésre elég a meglévő rendszer
 - Teljesítménytesztelés
 - kell-e még teljesítményen javítani és hol (fejlesztésnél)
 - optimális-e egy konkrét beállítás
 - van-e egy beállításnak teljesítményre gyakorolt hatása

Elvárások

- Ismételhetőség
 - Repeatability
- Reprodukálhatóság
 - Reproducibility
- Relevancia
- Szabványok/megállapodások betartása
- Általánosított felhasználói eset
 - Átlag felhasználó számára értelmezhető legyen az eredmény

Bechmark környezet

■ Specifikáció

- hardver
- szoftver
- üzemviszonyok
- ütemezés
- dokumentáció

■ Alapelvek

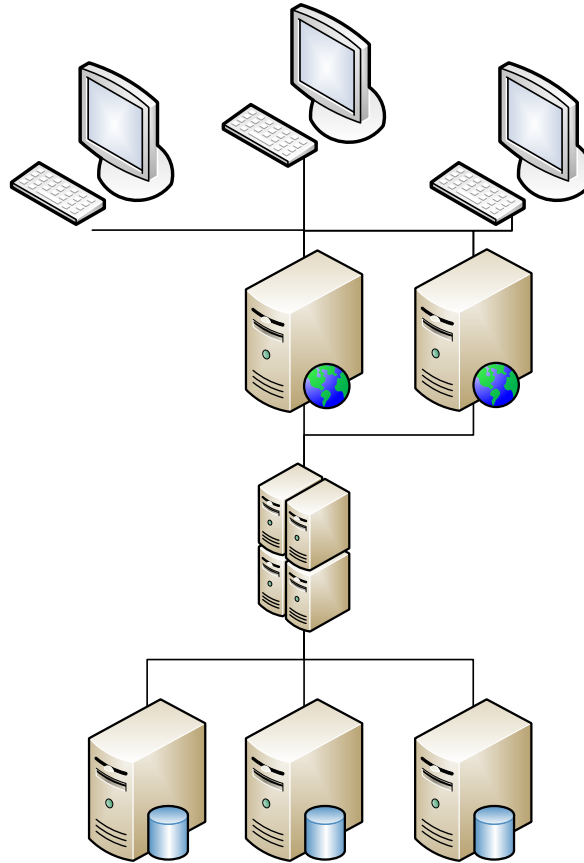
- kölcsönös zavarás minimalizálása (üzem közben)
- Pareto elv (80/20) kihasználása
- prekonceptió (mit mérünk)
- terhelés közelítse a valós mintát (profilok)

Benchmark terhelési modellek

- Tudományos/műszaki rendszerek
 - nagy mennyiségű adat feldolgozása (number crunching)
 - párhuzamos módszerek
- Tranzakciókezelés (OLTP)
 - kliens-szerver környezet
 - sok gyors, párhuzamos tranzakció
- Batch jellegű adatfeldolgozás
 - riport készítés nagy mennyiségű adatból
- Döntéstámogatás
 - kevés, bonyolult lekérdezés
 - ad hoc műveletek
 - sok adat (pl. OLAP)
- Virtualizáció

Benchmark terhelési modellek példa

- OLTP architektúra példa



Mérendő paraméterek (Metrikák)

- Futási idő
 - kezdet, vég?
 - eloszlás
 - CPU, I/O, hálózat,...
- Tranzakciósebesség
 - rendszer reakcióideje
 - akár egymásba ágyazott tranzakciók
- Áteresztőképesség
 - **feldolgozott** adatmennyiség / futási idő
 - terhelés függvényében

Mérendő paraméterek (2)

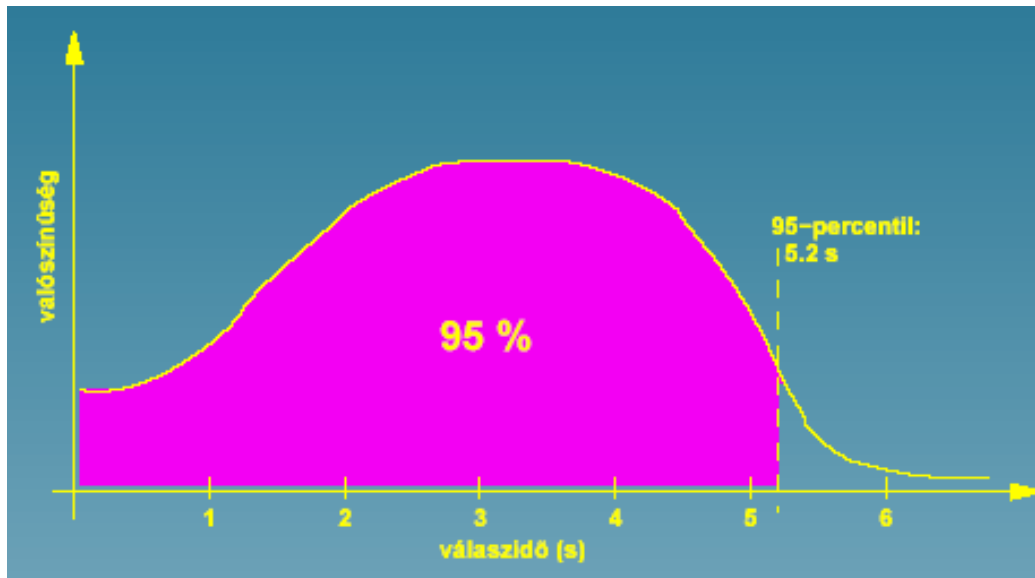
■ Válaszidő

○ terhelés függvényében

- felhasználók
- tranzakciók száma, stb.

■ X-Percentil

- Egy adott halmaz X százaléka ez alatt az érték alatt van



Eredmények összehasonlítása

- Pl. referenciarendszer alapján
 - Több paraméter alapján ellentmondást kaphatunk
 - Absztrakt referencia („standard benchmark”)
 - Adatbányászati módszerek
-
- Nincs egy univerzális teljesítménymérő metrika még egy-egy szűkebb területen belül sem!

Tipikus problémák

- Túl kicsi problémaméret
- A felhasználási terület számára releváns eredmény?
- Elavult referenciák
- „Rejtett” paraméterek
 - konfigurációs beállítások
 - adott környezet specifikus tulajdonságai
- Elfogultság
- Hiányos specifikáció

Benchmark típusok

■ Tudományos terminológia

- Macrobenchmark – nagy összetett alkalmazás mérése
 - relevancia biztosításával közvetlenül használható eredményeket ad
- Microbenchmark – alkalmazás kis részének kiemelt, analitikus mérése
 - analitikus felhasználás, profiling, teljesítmény előrejelzés
- Nanobenchmark – egy-egy atomi művelet teljesítményének elkülönített mérése
 - főleg hibakeresés, profiling célra hasznos

Benchmark elvégzése

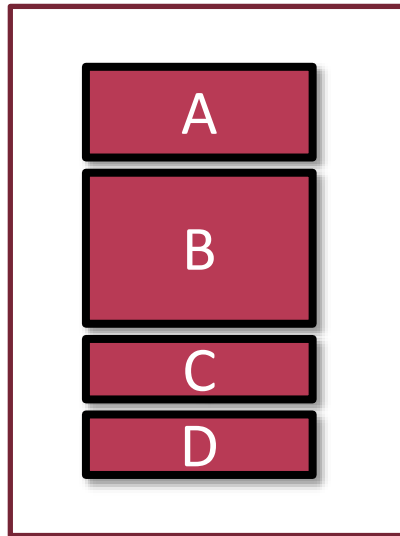
■ Relevancia biztosítása

- Tényleg azt az alkalmazást mérjük, amit kell
 - Különböző macrobenchmarkok eredményei többnyire nem vihetők át egymás között.
- Terhelésgenerálás jellege közelítse a valódi terhelést
 - Főleg macrobenchmarkoknál fontos, időben szétosztott valódi terhelést félrevezető lehet egy összefüggő batch terheléssel helyettesíteni.
 - Ügyeljünk a terhelésgenerátorra ható visszacsatolásokra
- Minimalizáljuk a zavaró tényezőket
 - Főleg microbenchmarknál fontos, Pl.: ne mérjük bele véletlenül diszk I/O-t a memória áteresztőképességébe, ne fusson más alkalmazás közben
 - Futási eredmények szórását kezelni kell

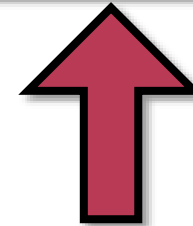
Eredmény analitikus felhasználása

- Cél: következtetünk jövőbeli (tehát nem mérhető) terhelésnél kapott eredményre

Összetett alkalmazás



Becsült teljesítmény



Eredmények részfeladatonkénti microbenchmarkra:



Egyes részfeladatok aránya ismert

(vagy legalább jól becsülhető) a terhelés során

SPEC benchmarkok

- <http://www.spec.org/benchmarks.html>
- Standard Performance Evaluation Corp.
- Erőforrás és alkalmazás szintű benchmarkok
 - CPU
 - Alkalmazások
 - Levelező szerverek
 - Web szerverek
 - Network File System, stb.
- Benchmark: megrendelhető szolgáltatás
 - Forráskódot adnak, nekünk kell fordítani
 - Licenz díjas!

SPEC CPU2006

- CPU intenzív
- CINT2006
 - Számításigényes egész számos
- CFP2006
 - Lebegőpontos

- Példa:



cpu2006.html

CINT2006 és CFP2006 terhelésgenerátorok

■ CINT2006 :

400.perlbench	C	Programming Language
401.bzip2	C	Compression
403.gcc	C	C Compiler
429.mcf	C	Combinatorial Optimization
445.gobmk	C	Artificial Intelligence
456.hmmer	C	Search Gene Sequence
458.sjeng	C	Artificial Intelligence
462.libquantum	C	Physics / Quantum Computing
464.h264ref	C	Video Compression
471.omnetpp	C++	Discrete Event Simulation
473.astar	C++	Path-finding Algorithms
483.xalancbmk	C++	XML Processing

■ CFP2006:



[SPEC CFP2006 Benchmark Descriptions.htm](#)

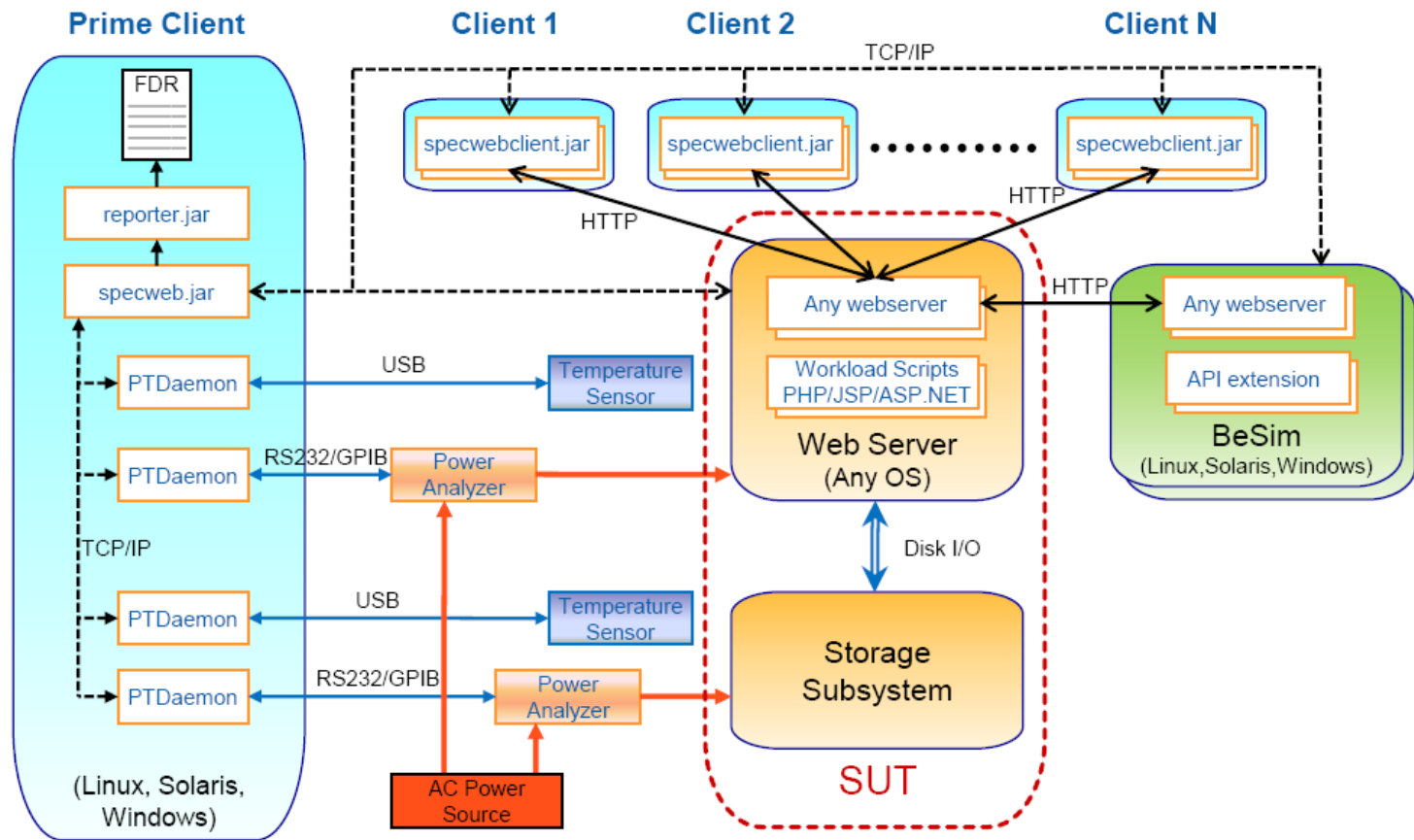
- Terhelés:
 - Biztonságos bankolás (SSL)
 - E-kereskedelem (SSL+ http letöltés pl)
 - Áteresztőképesség
 - Teljesítmény / energia metrikák
- Architektúra (ábra)
- Terhelésgenerálás (ábra)
- Példa:



web2009.html

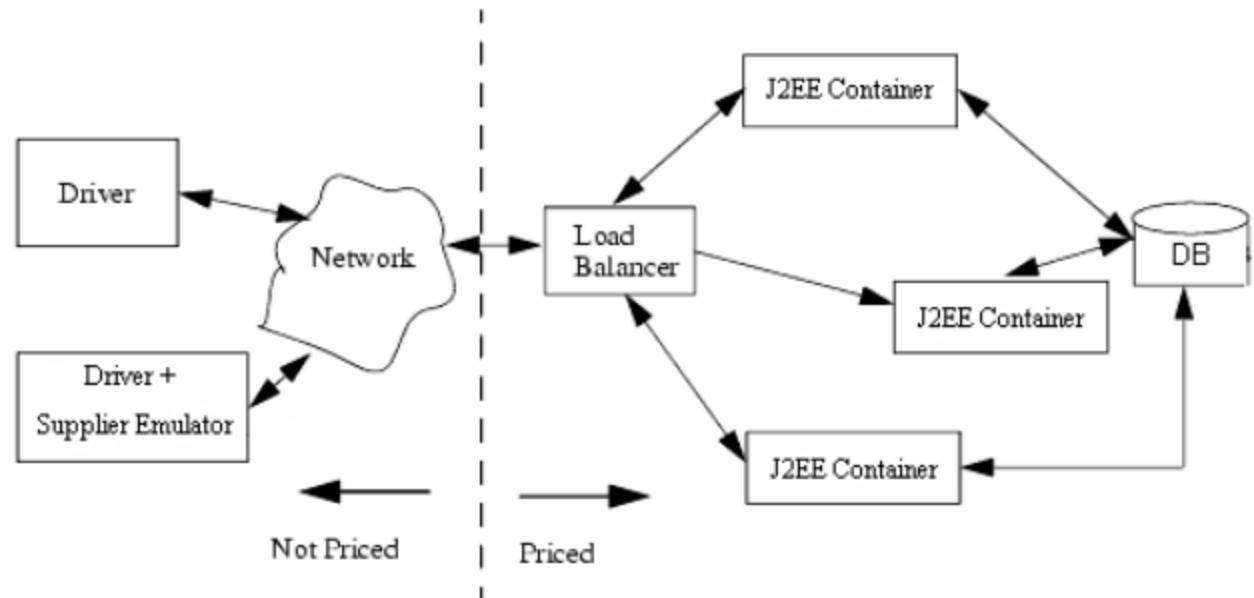
SPECweb2009

SPECweb2009 Setup (External Storage)



SPECjAppServer2004

- [Html](#) oldal elemzése előadáson
- J2EE Benchmark
 - Alkalmazáserver skálázhatósága és teljesítménye
- Architektúra (több szerveres, kliens - szerver arch.)
- Metrika:
 - Áteresztő-képesség
- Terhelés:
 - Autógyártás
 - Gyár – megrendelők
 - Autókatalógus



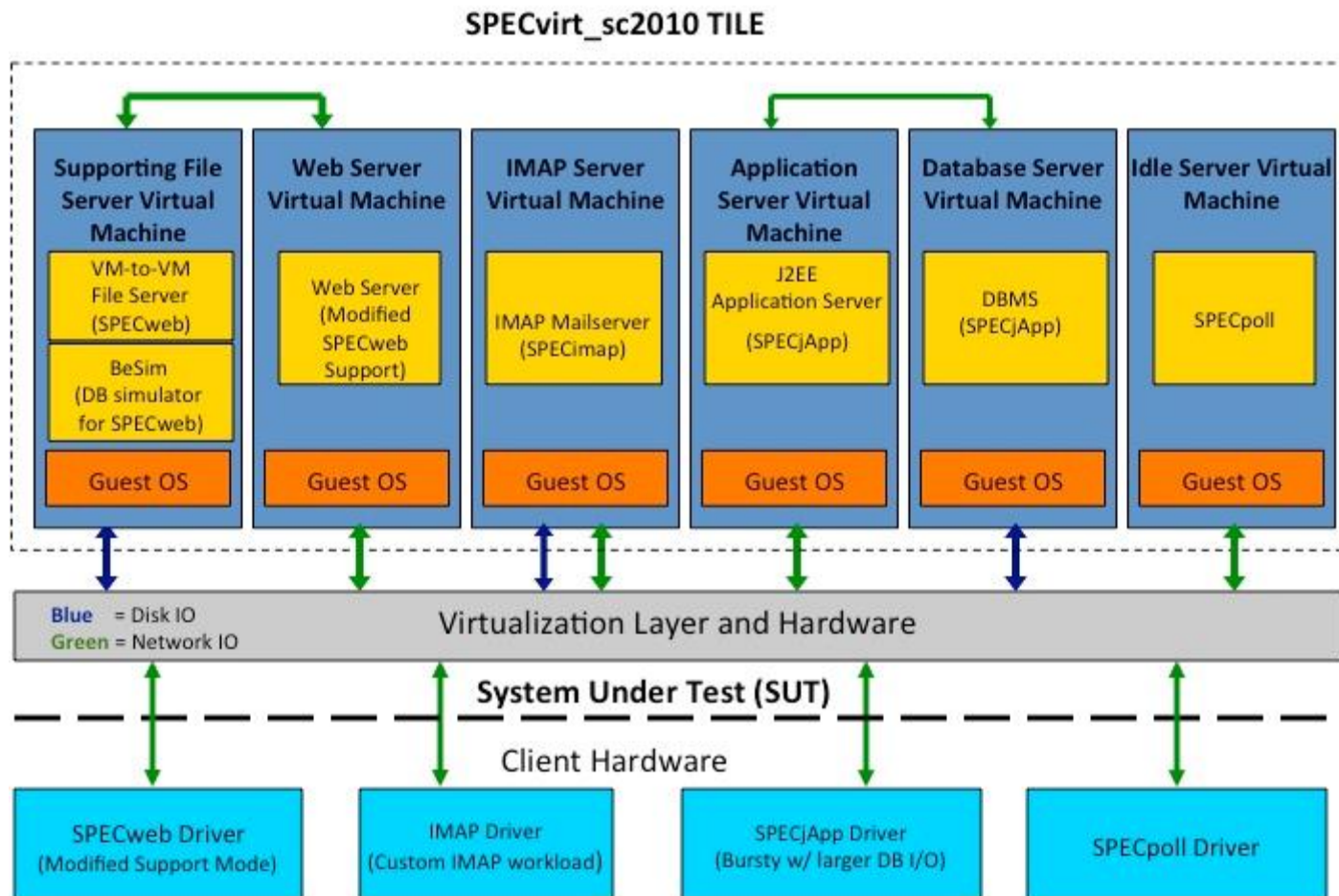
SPECjvm2008

- Ingyenes!
- JRE teljesítmény
 - CPU / memória
 - Kevés fájl I/O, nincs hálózati I/O
- Alkalmazások
 - Javac
 - LZW
 - Startup
 - MPEGaudio
 - Xml (transzformáció és validáció)
 - Crypto
 - Aes, rsa, signverify

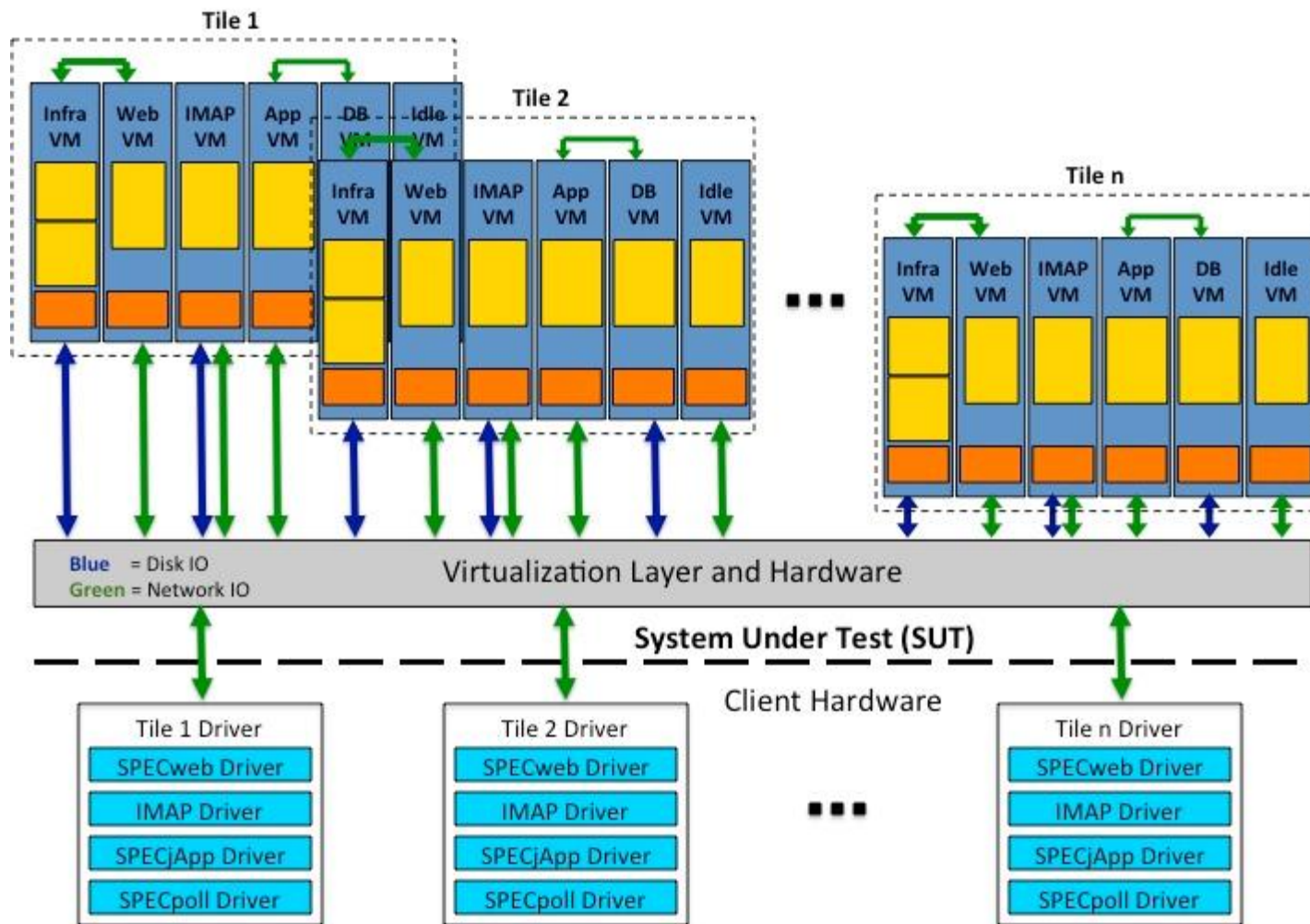
SPECvirt_sc2010

- Szerverek 18%-a volt virtualizált 2009q4-ben
- Cél: skálázhatóság / energiahatékonyság
 - Hány VM konszolidálható egy HW-re
- Architektúra (ld. ábra)
 - 1 „tile” 6 VM
 - Függőségek!
- Terhelés
 - Tile-ok számát növeli
 - Amíg QoS megvan, vagy amíg a metrikák már nem nőnek
- Metrika
 - Szubkomponensek normalizált metrikái
 - PI kérés/s a webserveren, alkalmazáserver burstössége

SPECvirt_sc2010



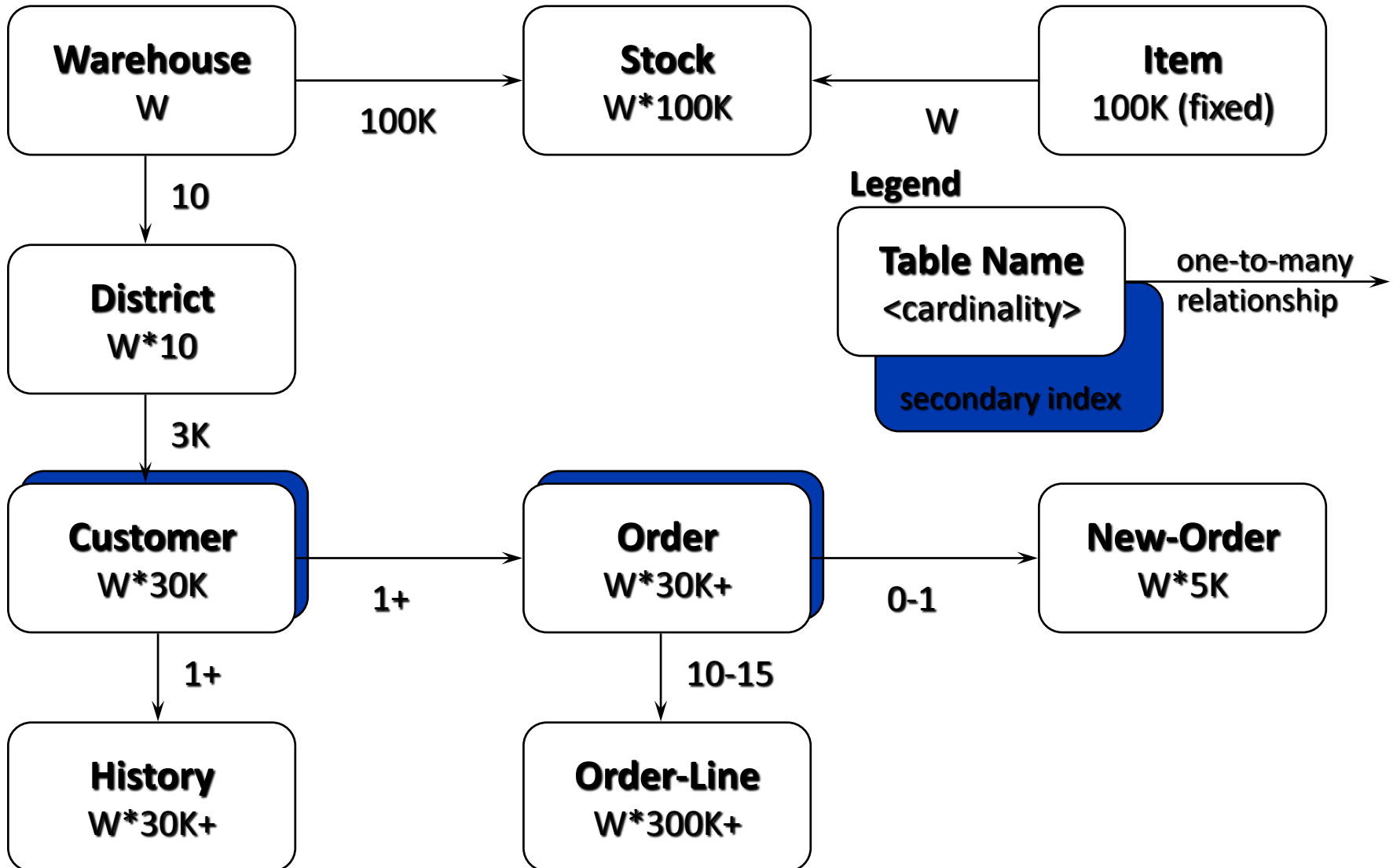
SPECvirt_sc2010



TPC benchmarkok

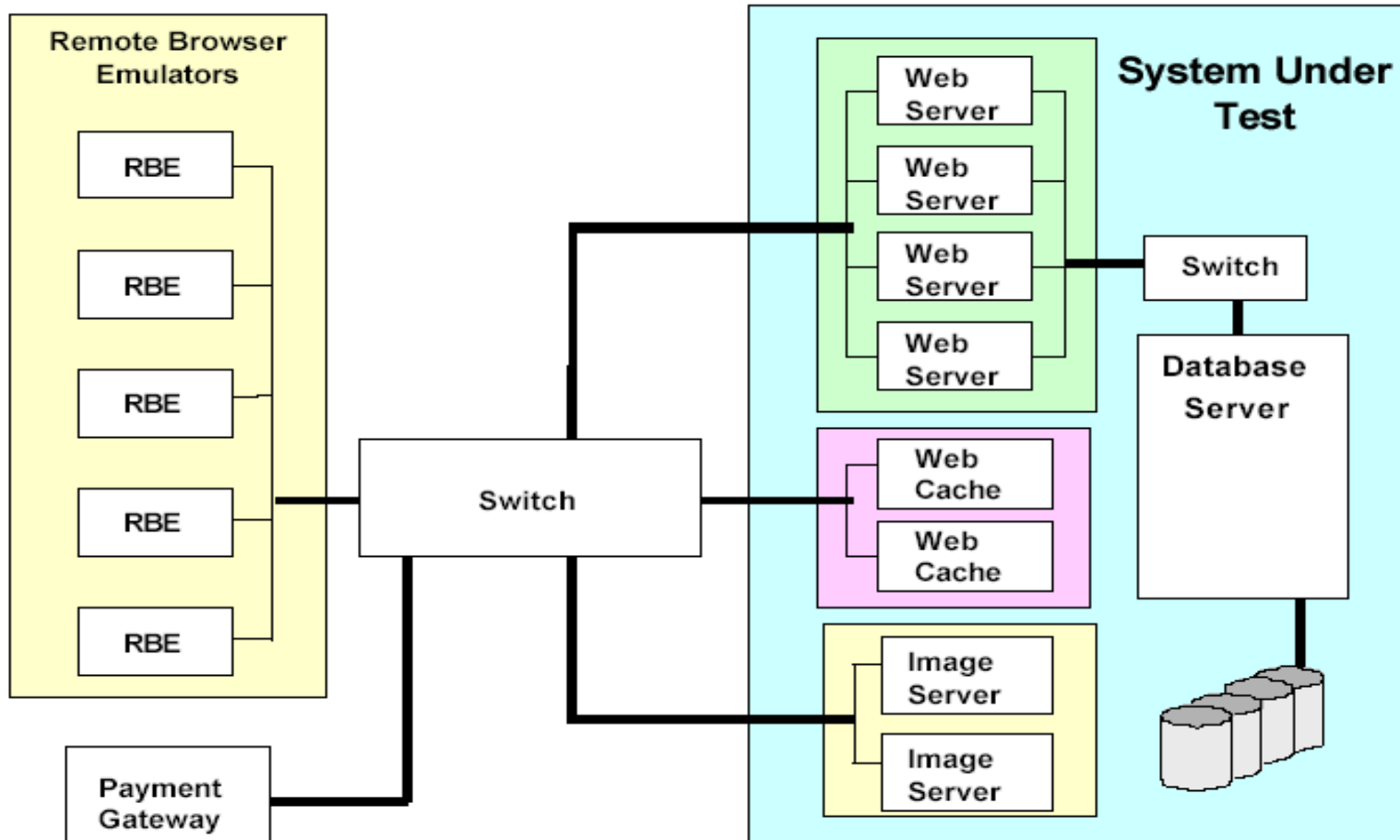
- <http://www.tpc.org/information/benchmarks.asp>
- Transaction Processing Council
- TPC-C (elektronikus kereskedelem, banki rendszer):
 - felhasználók tranzakciókat hajtanak végre
 - rendelés/lemondás, lekérdezés, stb.
 - szervereket hasonlít össze
 - HW
 - OS
 - DBMS
 - egyéb paraméterek:
 - tervezett rendelkezésre állás, elérhetőség (pl. 24/7 vs. 8/5)
 - OLTP rendszerek mérőszámai:
 - tranzakciós ráta (tpmC): 5 különböző fajta tranzakció alapján
 - ár / tranzakció (\$/tpmC): fenntartási költségek / tranzakciók
 - Bővebben: info [itt](#), pdf [itt](#).

TPC-C



- TPC-W (Web e-Commerce rendszerek):
 - komplex rendszereket hasonlít össze
 - különféle szerverek összekapcsolása
 - dinamikus oldalak
 - 3 különböző profil vizsgálata
 - rendelés gyakorisága különböző
 - szimulált terhelés
 - könyvesbolt a mintarendszer
 - bejelentkezés, bevásárlókocsi, online rendelés
 - hitelkártya információ külső szolgáltatótól (Payment Gateway)
 - böngészőből elérhető

TPC-W konfiguráció



TPC-W mérés

- Emulált kliensek
- Konfiguráció
 - „gondolkodási idő”, átlag 7 sec, max. 70 sec
 - döntési valószínűségek (Web Interaction Mix)
 - válaszidő követelmények
 - új / régi felhasználók (regisztrálás, cache)
 - felhasználók száma (Number of Users)
- Scale factor:
 - könyvek száma az adatbázisban, 1000...10,000,000
- Átlagos vagy Worst Case értékek

TPC-W tulajdonságai

■ Előnyök

- komplex rendszer tesztelés
- valós viszonyok
 - terhelés kiegyenlítés a Web szerverek közt
 - külön Image szerver
 - Web cache használata

■ Hátrányok:

- nem valós alkalmazást használ
 - alacsony szinten van kódolva
- túl egyszerű lekérdezéseket használ
- kevés kép/oldal van a mintarendszerben
 - ezért tiltja a cache-t az emulált böngészőben

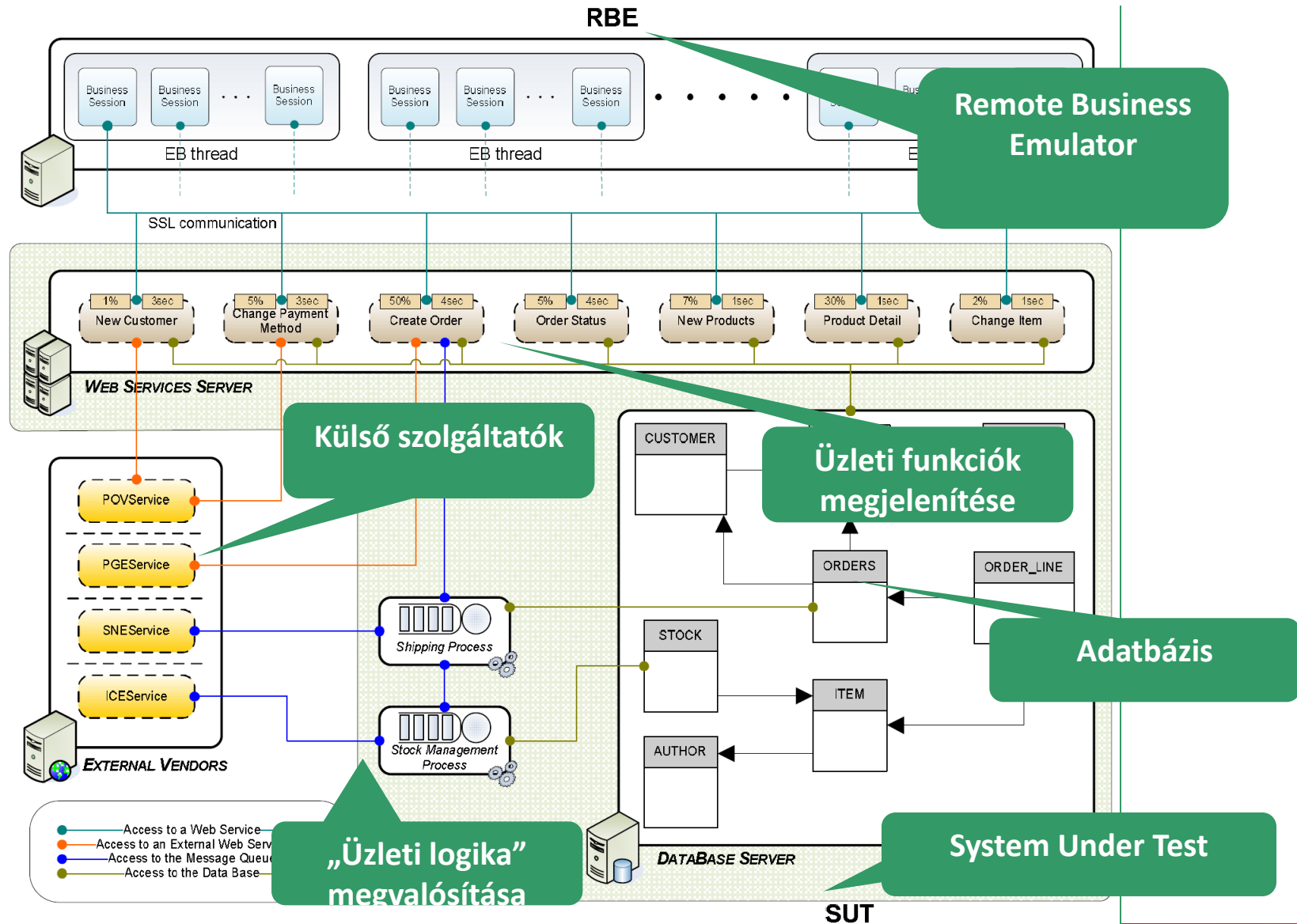
forrás: Wayne D. Smith: TPC-W: Benchmarking An Ecommerce Solution

http://www.tpc.org/tpcw/TPC-W_Wh.pdf

TPC-App (TPC-W 2.0)

- Alkalmazás szerver és Web szerver együttes mérése
- TPC-W utóda
- Kereskedelmi alkalmazások
 - B2B is
 - Webszolgáltatások
 - Sokféle tranzakció egyszerre
- Alap metrika: SIPS
 - Web Service Interactions Per Second
- SIRT: Web Service Interaciton Response Time

TPC-App konfiguráció



TPC-App környezet

■ Külső szolgáltatások

- Purchase Order Validation Emulator
- Payment Gateway Emulator
- Inventory Control Emulator
- Shipment Notification Emulator

■ Tesztelési paraméterek

- ACID adatbáziskezelés
- Configured EBs (felhasználók száma)
- Active EBs (párhuzamosan belépett felhasználók száma)
 - $(0.9...1) * (\text{Configured EBs})$