

Rendszermodellezés

Teljesítménymodellezés gyakorló példák

1. Szálkészslet

Legalább hány aktív szálát kell engedélyeznünk egy webserveren alkalmazásunknak, ha az egyenletes terhelés melletti teljesítményét nem szeretnénk visszafogni? Szálkorlát nélküli mérésekkel megállapítottuk, hogy egy kérés átlagosan 120 ezredmásodpercet tölt a rendszerben, és a szerver másodpercenként 50 felhasználót szolgál ki.

2. Tudásbázis

Vállalatunk nyilvános szakmai tudástára egymásra is hivatkozó szócikkeket kínál a cég termékeit világszerte használó ügyfeleknek. Egy szócikk lekérés kiszolgálásának szerveroldali igénybevételének ideje normális eloszlással modellezhető, 60ms várható értékkel és 45ms szórással. A szócikk megtekintése után az olvasó csak 30% valószínűséggel hagyja el az oldalt, az esetek többségében ugyanis egy újabb szócikkre mutató hivatkozásra kattint.

- Egy olvasó tudásszomjának kielégítéséhez összesen átlagosan mekkora szerveridő szükséges?
- Tekintsük úgy, hogy az egyes kérések a szerveren nem párhuzamosíthatóak. Óránként hány egyedi látogatót képes kiszolgálni a szerver?

3. Terheléselosztás

Adott egy webserverver (WS) és két fürtözött adatbázisszerver (DB1, DB2). A két adatbázis szerver közt súlyozott RR terheléselosztás alapján választunk, 1:2 arányban. Minden felhasználói kérés kiszolgálása során mindkét fajta erőforrást használjuk. A csúcsideőszakban 30 percig monitorozzuk a rendszert, ezalatt 9000 kérést szolgál ki. A szervereken mért foglaltsági idők: WS - 1350s CPU idő; DB1 - 810s, DB2 - 1320s diszk IO idő.

- Mekkora az egyes szerverek jelenlegi átbecsátása?
- Mekkora a rendszer maximális átteresztőképessége?
- Mennyi időt töltenek egy-egy kérés kiszolgálásával a szerverek?
- Miért nem egyféle foglaltsági időt vettünk figyelembe a két erőforrástípusnál?
- Hol csal még így is a modell?

4. Közösségi oldal

Internetes közösségi oldalt működtetünk. Az utóbbi időben számottevően megnőtt a népszerűsége, de ezáltal a válaszidő is kellemetlenül megnyúlt. Az üzleti cél, hogy csúcsideőszakban egyszerre 1500 felhasználót átlagosan négy másodperces válaszidővel szolgáljon ki a honlap.

- Minimálisan mekkorára kell tervezni a kiszolgáló infrastruktúra átbecsátóképességét, ha az azon kívüli késleltetés (hálózati forgalom, HTML megjelenítés a kliensoldalon) egy másodpercnek becsülhető?
- Az újratervezett weboldalon a mérések szerint egyetlen kérés kiszolgálása átlagosan 20ms CPU-időt igényel a webserveren, és 12.5ms erejéig foglal le egy adatbázisszerveret. Jelenleg 15 webserverver fogadja a kéréseket és az adatbázis 5 kiszolgálóra van replikálva. Lineáris skálázhatóságot feltételezve, milyen számítógépből és mennyit kell még legalább venni, hogy a fenti cél teljesülhessen?
- A kibővített rendszerben mekkora lesz az egyes szerver típusok kihasználtsági aránya? Ha az a cél, hogy még a csúcsideőszakban is legfeljebb 50%-os legyen a kihasználtság, meddig kellene még bővíteni a rendszert?