

Telepek fejlődési lehetőségei

A Linux szerepe a géptelepek jövőjében.

A telepek építése tíz évvel ezelőtt, mindössze egy dolgot jelentett, kis számú és nagy teljesítményű kiszolgálók összekapcsolását és megosztott lemezeket. Ezáltal az állandó hálózati hozzáférést igénylő ügyletek feldolgozását végző alkalmazások használhatóbbakká és – néhány esetben – méretezhetőkké váltak. Azóta láthatunk az Internet-alapú rendszerek, a vékony áruki-szolgáltatók (thin commodity servers) felemelkedését és a Linux-alapú nyílt, méretezhető, megosztott feldolgozás teljesen új modelljét. A linuxos világ számos forgalmazója megpróbál újra olyan telepialakítási eljárásokat kidolgozni, amelyek a hagyományos Unix OLTP rendszerekhez már tíz évvel ezelőtt léteztek. Nem biztos, hogy ez a legjobb megoldás. A telepkészítés módszereinek az olyan típusú alkalmazások és szerkezetek irányába kell fordulnunk, amelyek internetes adatközpontokba települnek, és linuxos vékony kiszolgálók csoportjaira épülnek. Az internetes adatközpontban – legyen az egy telephely vagy egy cég – megtalálható alkalmazásszerkezeteknek van néhány olyan közös jellemzőjük, amelyek alapvetően különböznek a régi üvegpalaták szerkezetétől. Bővebben kifejtve az internetes adatközpont-szerkezetekben többnyire fellelhető néhány közös minta:

- A szerkezetek többrétegűek (multitiered), különösen igaz ez a webkiszolgálókra, alkalmazáskiszolgálókra és adatbáziskiszolgálókra.
- Minden réteg kiszolgálók tucatjait tartalmazhatja.
- Minden réteg egyedi követelményeket támaszt a megosztott és a lemásolt adatokkal szemben.

A nagy forgalmú e-kereskedéssel foglalkozók vagy tartalomszolgáltatók esetében a webes szolgáltatás alapját képező, megosztott erőforráskészlet egészének kezelése sokkal összetettebb feladat, mint egyszerű kiszolgálók hibátűrő működését biztosítani.

Míg néhány alkalmazás, mint például az OLTP vagy más hasonló e-kereskedelmi hibajavító rendszer továbbra is megosztott lemeztérleten alapuló megoldásokat igényel, a korszerű alkalmazások sok esetben képesek kihasználni a középréteg által nyújtott programalapú többszörözés (replication) előnyeit.

Hiányzik azonban belőlük az a háttér, amely egyetlen nagy számítási környezetté olvasztaná egybe ezeket a megoldásokat. Az alkalmazásvezérlés – beleértve a hibajavítást – továbbra is olyan kulcsfeladat marad, amivel az üzemeltetőnek manapság mindenképpen szembe kell néznie. Szerencsére azonban a lehetőségek köre egyre bővül. Az alkalmazások, amelyeknek megosztott adatelérésre van szükségük, manapság gyakran gépek tucatjait használják, többnyire több megjelenési ponton (Points of Presence, POPs). Ez azt jelenti, hogy többé már nem elegendő a hagyományos fürtkiszítés megosztott lemezmegoldásai, amelyek többnyire fizikailag megosztott SCSI-kapcsolatokon alapulnak. E megoldások lehetőségeit behatárolja az alkalmazott fizikai vezetékelés és a SCSI-címkorlát, így általában csak két-három rendszert képesek kiszolgálni.

A webkiszolgáló réteg

A webkiszolgáló réteg a Linux erőssége, csak olvasható vagy program által készülő adatok jellemzik. Ezt a típusú adatot meg lehet osztani (és gyakran meg is teszik) több kiszolgáló közt, hálózati csatlakozású, NFS-t használó tárolóeszközök segítségével. Minden

kiszolgáló csatolja az NFS-eszközt, és máris elérheti a megfelelő adatot. Ehhez a sorhoz a bonyolult megosztott SCSI vezetékezési sémák teljesen szükségtelenek, ugyanis a TCP/IP hálózat az összes szükséges kapcsolatot biztosítja a kiszolgálók és a tárolóegység között. Emellett fontos megemlíteni a telepvezérlő (clustering) programot, ez folyamatosan felügyelheti a webkiszolgálók teljesítményét és épségét, valamint megfelelő lépéseket tehet az így nyert adatok alapján.

Az alkalmazási réteg

A hagyományos munkaterületeken az alkalmazások jellemzően kiszolgálóalapúak voltak, és relációs adatbázisokon alapultak. A legtöbb esetben az adat nem az alkalmazási rétegben helyezkedett el, hanem minden az adatbázisba került. Ez a kép sokkal részletgazdagabb az Internet és a Linux világában. A Java és az Enterprise JavaBeans felemelkedése arra ösztönözte a fejlesztőket, hogy az alkalmazásréteg elemeit az adatbázison kívül, a fájlrendszerben helyezték el. Ráadásul az alkalmazások manapság sokkal változatosabb adatokkal dolgoznak, ideértve a folyamatos műsor-szórását, a szöveges adatokat és sok más olyan adattípust, amelyeket általában nem szoktak SQL-háttérben tárolni.

Emiatt az alkalmazási réteg olyan hibajavító és adatmegosztó követelményeket támaszt, amelyeket sem a hálózatsatolt tárolók, sem a megosztott SCSI-eszközök nem képesek kielégíteni. A hálózatsatolt tárolók általában azért nem működnek, mert az NFS protokoll nem képes megfelelően megvalósítani az összefüggőséget egy olyan esetben, ahol egy időben több kiszolgáló is írhatja ugyanazt az adatot. A megosztott SCSI azért nem megfelelő, mert többszörözést, és nem megosztást kellene alkalmazni több rendszer vagy több POP között. Sok esetben éppen ezért, az alkalmazás inkább saját magának készít másolatot. Vegyünk például egy valutaváltó kereskedelmi alkalmazást, amit egy tucat nagy banknál használnak. Ebben az esetben a valós idejű kereskedelmi adatok többszörözése alkalmazásszinten folya, a hibajavító eljárást pedig telepszinten lehetne megoldani.

Az adatbázis réteg

Az adatbázisréteg a hagyományos fürtöző megoldások birodalma, ezt a nagy Unix-terjesztők – Sun, HP és az IBM – fejlesztették ki. Az eredeti elképzelés szerint két kiszolgálót kell csatlakoztatni néhány SCSI-meghajtóhoz. Az egyik közülük a tevékeny mester, a másik pedig tétlen várakozó. A korszerűbb megvalósításokban több kiszolgáló használ egyetlen nagy teljesítményű háttértárat (SCSI-eszköz vagy fiber channel), és egy párhuzamos adatbázis-kezelőt (például az Oracle Parallel Server – OPS) alkalmaznak, hogy kiszolgálják a gépek által kezdeményezett egyidejű eléréseket. Csakhogy a jó minőségű fiber channel SAN-ok igen drágák, csakúgy, mint az OPS-hez hasonló programok. Ezeknek a megoldásoknak ár/teljesítmény aránya gyakran elfogadhatatlan a linuxos vékony piac vagy egy internetes adatközpont számára. A tevékeny/tétlen SCSI-megoldás erőforrás-pazarló és csak korlátozott méretezhetőséget tesz lehetővé. Szerencsére vannak más megközelítések is, amelyek jobban illeszkednek a Linuxon futó alkalmazásokhoz.

Sok adatbázis forgalmazó fektetett munkát például a megosztás nélküli többszöröző programmegoldásokba (Informix Replicator, DB/2 Replication, Oracle Multi-master and hot standby). Ezeknek az

eszközöknek a használatával megoldható, hogy több kiszolgáló osztozzon az adatbázisrétegen, közös, megosztott tárolóeszköz használata nélkül. Maguk az adatbázisok azonban gyakran híján vannak azoknak a képességeknek, amelyek segítségével felderíthetnék a hibákat, megkezdhetnék a hibamentesítést, és szavatolhatnák az alkalmazás épségét. Ennek következtében a legjobb az, ha egy olyan megoldással társulnak, amely biztosítja ezeket a szolgáltatásokat. Egy életből ellesett példa egy lapkakészítő műhely, amely egy önműködő gyártási alkalmazást futtat Intel-alapú vékony kiszolgálókon. Az alkalmazás nagy fontossága végett a lapkagyártó három többszörözést is kér. Így két másik gép áll készen arra, hogy az esetlegesen leállt kiszolgáló feladatait átvegye. A választott megoldás az Informix adatbázis-többszöröző szolgáltatás volt, amely fizikailag elkülönített gépeken futott, megosztott tárhely igénybevétele nélkül, géptelepkezelő motorral összekapcsolva.

A megosztott tárolók jövője

Néhány más alkalmazásnak fizikailag megosztott tárhelyre van szüksége. Mivel a megosztott SCSI-eszközök továbbra is ritkaságszámba mennek, egy ilyen rendszer összerakása nagy figyelmet, valamint teljes körű – mind a gépjártóktól, mind a programterjesztőktől – minőségbiztosítást igényel. (Mikor ellenőrizted utoljára a gépi program szintjén a merevlemezdet?) És mikor mindezt nagy fáradsággal összeaktuk, amint láttuk, elég komoly korlátokba ütközünk mind a csomópontok számát illetően, mind pedig a lehetséges fizikai vezetékezés terén. A fiber channel megold ugyan néhányat a gondjaink közül, de rögtön okoz is helyettük néhányat, ideértve a drágaságát, a több terjesztő közötti működési nehézségeket, illetve a kezelési módot, mely sok linuxos vékony kiszolgáló felhasználójának szokatlan lehet.

A következő évben nagy változások várhatók a megosztott tárolók területén. Az ipar vezető cégei két új tároló-összekapcsolási módszert mutatnak be, így a megosztott tárolók ára a vékony kiszolgálók számára is elfogadhatóra mérséklődhet. Így is képesek létrehozni az internetes adatközpontoknál megszokott, nagyszámú csomópont-mennyiség szerinti megosztást is képesek lesznek létrehozni.

A Cisco és más, a piac változásaira érzékeny cégek mint például a 3ware, már támogatja a SCSI-over-IP módszert, amely lehetővé teszi, hogy hálózati kapcsolókkal ellátott Ethernet hálózaton SCSI-protokollt futassunk.

Az Intel és a kiszolgálóforgalmazók az Infiniband nevű új I/O szabvány mögött sorakoztak fel. Ez olyan kapcsolt I/O rendszert képes szolgáltatni, amelyet akár a kereskedelmi kiszolgálók alaplapján megtalálható lapkakészletekben is alkalmazni lehet. Igazság szerint ezek a fejlesztések kiegészítik egymást – a Gigabit Ethernet kártyák ülhetnek Infiniband szerkezeteken, és futtathatják a SCSI-over-IP protokollokat. Következésképpen a közeljövőben várhatóan rendszerek tucatjai vagy akár százai is osztozhatnak egyazon tárhelyen, ennek a haladó kapcsolatrendszernek köszönhetően.

És hogy mi határozza meg a jövő adatközpontjának szerkezeti alapjait? A telepészítési módszerek, amelyek segítségével a megismert kapcsolatok révén valóban használható és méretezhető megosztott megoldásokat készíthetünk a linuxos gépekből.

Ken Dove

jelenleg a PolyServe Inc. vezető mérnöke. Tizenkét éven át a Sequent Computer Systems vezető programmérnöke volt, majd az IBM alkalmazásában állt.

