

A barátságos keselyű

A számítógéptelemek és a Grid egyre nagyobb jelentőséggel bír az informatika világában. A Condor ennek az egyik megvalósítása.

Elsősorban arra szeretném felhívni a figyelmet, hogy mi a projekt célja és kik tudják a legjobban hasznosítani, illetve mire lehet és mire nem ajánlatos használni.

A Condor különálló számítógépek erőforrásainak vezérlésére szolgál, elsősorban nagy számítás- és kis I/O-igényű feladatok feldolgozására. Támogatja a feladatok sorba állítását (queue), ütemezését, fontosság- (priority) megadását, az erőforrás-figyelést és -vezérlést. Egyaránt lehet alkalmazni kinevezett telepeket (dedicated clusters), de más célra is használt számítógépek erőforrásainak az összekapcsolására.

Nem alkalmas rövid, gyors lefutású, illetve nagy I/O-igényű programok futtatására. Olyan kinevezett telep esetén, amely közös hálózati fájlrendszert használ, a második feltétel nem olyan szoros megkötés – alkalmatlan például a webkiszolgáló feladatainak több gépre történő szétosztására.

Elsődleges célja, hogy nagy számításigényű feladatok futását irányítsa eltérő és gyakran változó erőforrásokon. Azt nem teszi lehetővé, hogy egy folyamatot valamely erőforráson azonnal elindítson – több perc is eltelhet a tényleges futtatásig –, de ha nagyszámú erőforrás áll rendelkezésre, amin számos olyan feladatot (job) kell lefuttatni, amelyek kellően párhuzamosíthatók, ennek a vezérlését viszonylag jól elvégzi. Nem is biztos, hogy megtalálja a leghatékonyabb sorrendet és erőforrás-kiosztást, de általában közelít hozzá. A szakirodalomban ezt összefoglalóan nagy feldolgozási képességű környezetnek (High-Throughput Computing) nevezik.

Az ötlet onnan származott, hogy az egyetemi tanszékeken viszonylag sok gép található, kihasználtságuk azonban csekély, ugyanakkor általában egész nap üzemelnek és hálózatba vannak kötve. Ezt az „elpocsékolt” számítási teljesítmény akarták hasznosítani. Például valaki le szeretne futtatni néhány feladatot (job), ami a saját számítógépén akár hónapokig is eltarthatna. Mindez a tanszék vagy az egyetem gépein pár nap alatt elvégezhető, ha Condorral össze vannak kapcsolva. Egy feladatot csak el kell indítani, és amennyiben a Condor talál hozzá megfelelő erőforrást, azonnal oda is költözteti, majd amikor a program lefutott, értesíti a tulajdonost.

Másik jelentős felhasználási területe a Grid – egy világméretű telep létrehozása, amelyben az erőforrások teljesen eltérőek és minduntalan változnak.

A Condor az utóbbi pár évben a Grid kapcsán egyre inkább előtérbe került. Bár a projekt még akkor indult (1988), amikor a mai „Grid” fogalom nem is alakult ki teljesen, mégis nagyon alkalmas eszköz a Grid megvalósítására, ugyanis mindent tud, ami a Gridhez szükséges: képes teljesen eltérő, állandóan változó környezetben erőforrásokat nyilvántartani és kiosztani. Egyetlen hiányossága a felügyelet (monitoring) területén tapasztalható, de ezt például a magyar fejlesztésű (MTA SZTAKI) P-GRADE kiválóan kiegészíti.

Járulékos előnyök

Lehetőség nyílik arra, hogy a Condor a feladatok futtatására csak a szabad gépidőt használja fel. Az is meg van oldva, hogy-

ha valaki hosszú idő után tér vissza a saját számítógépéhez, amin más feladat (job) fut, akkor amint egyetlen billentyűt lenyom, a Condor önműködően más gépre költözteti az ott futó programot.

Lehetőség van Glide-in módszer alkalmazására, amelynek az az előnye, hogy olyan gépek is használhatók Condor-feladatok (Condor-job) futtatására, amelyekre nem telepítettek Condort. A ClassAd módszer segítségével megadható, hogy milyen típusú erőforrások szükségesek a feladat futtatásához: milyen a kiépítés (architect), a gépek száma, a gép processzorainak a gyorsasága, a processzorok száma és a memória mérete.

Használat

A telepítésről a hivatalos weblapon egy csomó leírás található (☞ <http://www.cs.wisc.edu/condor/>). Lépésről lépésre végigvezet bennünket az egyes szakaszokon. Itt erre nem térek ki, csak néhány lényeges pontra hívom fel a figyelmet.

Az első fontos tényező: a tűzfalak beállításai. A Condor nagyszámú különböző kaput (port) használ a folyamatok költöztetésére és a kapcsolattartásra, ezért az összekapcsolt gépek között legalább az 1024 feletti kapuk használatát engedélyezni kell. Ez tanszékek esetében általában nem gond, mert egyetlen külső tűzfalal bírnak és a belső hálózaton szinte minden megengedett.

Nehézség támadhat azonban az egyetemek, az egymástól távol lévő szervezetek összekapcsolásakor. E gondokra is született néhány megoldás, ilyen például a Condornak a Globusszal történő kombinálása.

A Globus – a Condorhoz hasonlóan – folyamatok más gépen való futtatását teszi lehetővé, de csak alacsonyabb szintű szolgáltatásokat nyújt. Biztonságos, titkosított csatornákat tartalmaz a folyamatok költöztetésére, a költöztetés viszont nem önműködő. A Globus esetén valóban le kell írni, hogy melyik folyamatot hol szeretnénk futtatni és hogy a folyamatok hogyan tartsanak kapcsolatot egymással. Ez nagymértékben megnehezítheti a programozók, illetve a kutatók munkáját, ezért kifejlesztették a Condor-G változatot, amely a Globust a folyamatirányítás képességével bővíti ki.

A másik megoldás, hogy a különálló Condor-telepeket (Condor-Pool) VPN-nel kapcsoljuk össze. Ilyen megoldást alkalmaz a már korábban említett KlaszterGrid-projekt is. Használata egyszerű: össze kell állítani a feladatot, le kell írni, hogy milyen erőforrások szükségesek, és a Condor nyújtotta programkönyvtárral újra kell építeni a programot. Ez a folyamat függhet attól, hogy a feladatot milyen Condor-univerzumban futtatjuk. (Általános univerzum esetén a fenti lépéseket kell végrehajtani.)

Azondor egyszerű választás lehet kutatóintézetek, egyetem számára, ahol nincs pénz szuperszámítógépek beszerzésére, így minimális ráfordítással is viszonylag nagy számítási teljesítményt lehet összefogni.

Kolcza Péter (kpeter@sysconfig.hu)