

# SZUPERSZÁMÍTÓGÉP A MISKOLCI EGYETEMEN

## SUPERCOMPUTER AT THE UNIVERSITY OF MISKOLC

Velezdi György\*, Csáki Tibor\*\*

### ABSTRACT

*This paper introduces an infrastructure development project at the University of Miskolc. The main goal of the project was to support research activities in computational mechanics. The main task was the building of a complete High Performance Computing (HPC) computer system. To achieve this goal we had to complete the following subtasks: procurement of HPC hardware, establishing complete HPC system, installing and optimizing operating software and creating operational environment.*

### 1. BEVEZETÉS

A Miskolci Egyetem Szerszámgépek Tanszékén hosszú idő óta folynak olyan kutatások, ahol nagymennyiségű számítás elvégzésére van szükség az alábbi szakmai területeken:

- a géptervezés elméletének és gyakorlatának összekapcsolása az informatika új módszereivel és eszközeivel
- szerszámgépek, robotok mechatronikai rendszereinek és részegységeinek szerkezeti fejlesztése, új szerkezeti megoldások kutatása
- hardver-szoftver eszközök fejlesztése bonyolult geometriájú felületek forgácsoló megmunkálásához
- gyártásautomatizálási módszerek és eszközök kutatása-fejlesztése

Minél pontosabban követi a valóságot egy vizsgálandó konstrukció modellje, azaz minél kevesebb elhanyagolással kell élnünk a geometria, illetve a fizikai és mechanikai peremfeltételek megadásánál, továbbá minél kisebb időszekvenciát választunk az elemzések, szimulációk elvégzéséhez, annál nagyobb erőforrást igényelnek az analízis és vizualizációs számítógépes alkalmazások. A trend mindenképpen a feladatok méretének, bonyolultságának növekedése irányában mutat, és ezzel párhuzamosan a vizsgált változatok számának növekedése is várható.

A már meglévő grafikus számítógépes munkaadásokon futó alkalmazások teljesítménynövekedését ugyan kisebb anyagi ráfordításokkal lehet megoldani

\* egyetemi adjunktus, Miskolci Egyetem, Szerszámgépek Tanszék

\*\* egyetemi docens, Miskolci Egyetem, Szerszámgépek Tanszék

rövid távon (pl. újabb processzorok vagy komplett munkaállomások beszerzése), de ez nem ad esélyt a valódi újdonságot jelentő vizsgálatok elvégzésére és valóban jelentős eredmények elérésére, mert sziget-szerű működéssel csak korlátozottan skálázható számítási kapacitások kiépítése lehetséges.

A Miskolci Egyetem Szerszámgépek Tanszékén ezért döntöttünk úgy, hogy - pályázati lehetőségek kihasználásával - egy olyan szuperszámítógép beszerzését valósítjuk meg, amely igen nagy számítási kapacitással, valódi ún. High Performance Computing (HPC) képességekkel rendelkezik.

### 2. A HPC RENDSZEREK SAJÁTÓSÁGAI

A teljesítmény valódi növekedése elsősorban a számítások „párhuzamosításából”, azok egyidejű elvégzéséből adódik. Egy szuperszámítógépes rendszer „lényege” akkor használható ki igazán, ha a használt vagy kifejlesztett alkalmazások minél inkább párhuzamos működésre képesek. Az ún.  *Soros (serial)* rendszereknél alaphelyzetben egy alkalmazás csak egy processzort illetve kiszolgálót képes használni. Az ún. *SMP* rendszereknél az alkalmazás ki tudja használni a kiszolgálóban lévő core-okat. Míg az *MPI* rendszerek esetén az alkalmazás párhuzamosan képes több kiszolgálón futni. A használt operációs rendszer alapvetően Unix/Linux vagy Windows verzió, több rendszer ezeket rugalmasan megosztva is képes kezelni.

Az elvégzendő tevékenység hatékonyságnövekedése alapvetően attól függ, hogy az adott alkalmazás (pl. a konkrét számításokat végrehajtó program) milyen verzióban áll rendelkezésre. A serial verziónál az egyetlen előny az, hogy nagyon sokat lehet belőle egyszerre futtatni több gépen (pl. más-más kiindulási feltételekkel vizsgálni ugyanazt a konstrukciót). Az *SMP* verzió már a core-ok számával arányos sebességnövekedést is ad, az *MPI* verzió nagyságrendekkel növeli a számítási sebességet. További jelentős hatékonyságnövekedés adódik az átviteli sebesség növekedésével (InfiniBand technológia akár 40-szer gyorsabb is lehet a Giga Bit Ethernethez képest), hiszen a programok általában nagyon sok adatot mozgatnak a node-ok (csomópontok) között.

A mai szuperszámítógépek architektúra szempontjából többfélék lehetnek. Az architektúra elnevezések és kategóriák meghatározása nem egyértelmű a szak-

mai szóhasználatban sem, de alapvetően *Cluster*, *MPP* és *Constellations* csoportokat különböztethetünk meg.

A Cluster legáltalánosabb meghatározása talán az, hogy „cluster” a nem külön HPC célokra gyártott, kereskedelmi forgalomban kapható, általában azonos kiépítésű 2, 4, 6, 8 core-t tartalmazó SMP szerverek vagy munkaállomások nagysebességű interconnect-el történő összekötéséből kialakult nagy kapacitású számítógép.

A legszívesebb kategória az MPP (Massive Parallel Processing) rendszerek családja. Ezen belül is több architektúrával találkozhatunk. Alapvetően két fontosabb alkategóriát érdemes kiemelni. Az *SMP* rendszer egy olyan architektúra, ahol a CPU-k mindegyike „látja” a teljes memóriát illetve I/O egységeket. Ezt az architektúrát nagyon nehéz 8-12 CPU-nál nagyobb méretben előállítani, mert a központi busz válik szűk keresztmetszetté. Ennek a feloldására születtek az ún. NUMA rendszerek. A NUMA rendszerekben a memória nem minden részét éri el egyforma sebességgel vagy késleltetéssel mindegyik CPU, de eléri. Ha több core-os CPU-kat használunk és ezeket kötjük SMP-szerű rendszerbe, akkor is gyakorlatilag NUMA architektúra jön létre, hiszen az egy CPU-ban levő core-ok a saját cache memóriájukat gyorsabban elérik mint a központi memóriát, azaz nem egyformán ér el minden core minden memória rekeszt. Ebben a kategóriában a nagyobb teljesítményű gépeknél gyakorlatilag ezt az utóbbi megoldást találjuk meg mindenhol. A Microkernel architektúra (IBM BlueGene, CRAY XT) viszonylag ismert a nagyközönség számára is, első-sorban az érdekes technikai hírek világából. A „Microkernel” egy olyan lebutított Unix kernel, amely csak a processz és thread kreálás, megszüntetés és az IPC (Inter Process Communication) kódot tartalmazza, de mást nem. Minden egyéb szolgáltatás, pl. a file system-ek, - ami pl. egy monolitikus Linuxban a kernel része - itt egy szolgáltatásként (service) fut. A CRAY architektúrában ez annyival egészül ki, hogy az IPC a SeaStar2+ interconnect tetején is működik. Azaz a compute node-ok futatják a számológépes thread-eket, és ha bármilyen más szolgáltatásra van szükségük, (pl. file I/O vagy hálózat) akkor ezt IPC-vel megkérlik a service node-ok valamelyikétől. Fontos, hogy a mikrokernel alapú Unix-ok nem 100%-ig kód kompatibilisek a monolitikus kernel alapú Unix/Linux rendszerekkel, ami azt jelenti, hogy az alkalmazásokat újra kell fordítani a speciális programozási környezetek segítségével, ami a gyakorlatban nem mindig valósítható meg egyszerűen.

Constellations architektúrájú gépen olyan feladatok futnak, amelyek egy csomópont kapacitását nem haladják meg (ami lehet egy „nagy” SMP gép) és ilyen alkalmazásból futtatnak sokat.

A TOP500 ([www.top500.org](http://www.top500.org)) - ez a világ ötszáz leggyorsabb számítógépét tartalmazó lista - alapján megállapítható, hogy a mai penge (blade) szerverekből összeállított nagy kapacitású gépek, azaz a *Cluster* architektúrájúak uralják a TOP500 listát (82%).

### 3. A MEGVALÓSÍTOTT HPC RENDSZER TULAJDONSÁGAI

Az előzőek alapján teljesen logikus, hogy a projekt keretében a ma meglévő szakmai kompetenciához illeszkedő „cluster” architektúrát választottunk úgy,

- hogy a beszerzés hatékony és gazdaságos legyen,
- teljesítménye legalább 5-8 évre oldja meg a számítási feladatok növekedéséből adódó igényeket,
- valóban új kutatási eredmények eléréséhez segítsen hozzá,
- a nemzetközi kapcsolatok kiszélesítésének is hatékony eszköze lehessen.

A rendszer összeállításánál arra törekedtünk, hogy a beszerzésre kerülő berendezés, kompakt, könnyen telepíthető és rugalmasan áthelyezhető legyen, és ne igényeljen különleges hűtési beruházást. Olyan szállító partnert választottunk, amely magyar képvisellel rendelkezik, és ahol magyar és angol nyelven folyamatos támogató kompetencia áll rendelkezésre.

Ezek a számítógépes rendszerek nem kaphatók dobozban, hanem a részegységeket kell az adott igények szerint összeszerelni. Ezért célszerűen a gyártó céget, a Hewlett-Packard-t bíztuk meg ezzel a feladattal, mert nálunk ilyen kompetencia jelenleg nem állt rendelkezésre (ehhez többéves HPC gyakorlat szükséges).

A specifikált, majd megrendelt és végül beüzemelt architektúrában 3 teraflops (!) körüli számítási teljesítmény áll rendelkezésre. A kiszolgálók (work node) InfiniBand (40Gb/s) sávszélességgel vannak összekötve, és amelyekhez jelenleg 100Mb/s és 1 GB/s switch-en keresztül tudnak a külső előkészítő munkahelyek csatlakozni. Minden kiszolgáló saját biztonságos tárolót használ, illetve a vezérkiszolgálóhoz (head node) RAID5 képes tárolót csatlakozik.

A rendszerkonfiguráció:

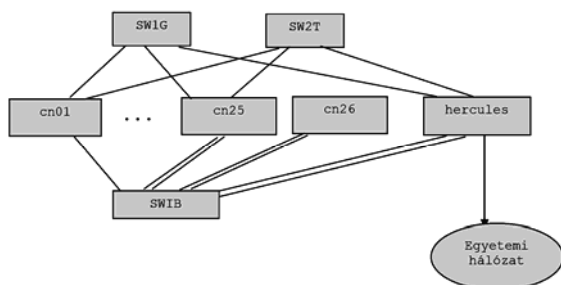
- Egy menedzser gép un. head node (Hercules)
- Két grafikus processzorral felszerelt gép (NVIDIA Quadro FX5800) grafikus alkalmazások és GPU-t használó párhuzamos számítások számára) (cn25 és cn26)
- Huszonnégy munkaállomás un. work node, vagy kiszolgáló (cn01 - cn24)

SW1G	
SW2T	
Head node	
cn25 (grafikus)	cn26 (grafikus)
cn12	cn24
cn11	cn23
cn10	cn22
cn09	cn21
cn08	cn20

cn07	cn19
cn06	cn18
cn05	cn17
cn04	cn16
cn03	cn15
cn02	cn14
cn01	cn13

1. ábra. A fizikai elrendezés

Mind a 24 gép 2 db hatmagos processzort tartalmaz (összesen 288 db 2.93 GHz mag), gépenként 160 Gb helyi tárolókapacitással. A Head node által kezelt teljes tárolókapacitást 12 db 2 TB-os merevlemez képezi, RAID5-be szervezve.



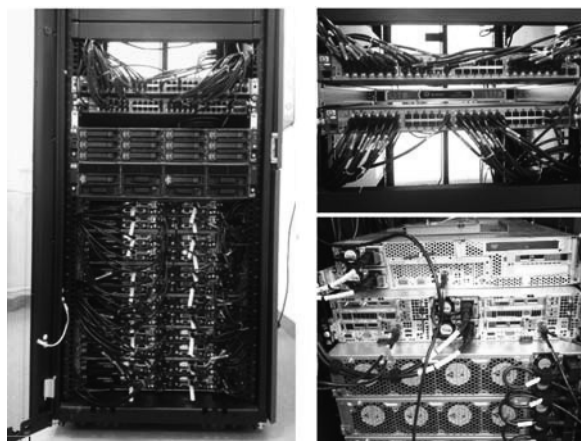
2. ábra. A hálózati topológia vázlatja

A leszállított szuperszámítógépre a hónapokig tartó utószerelési, átkábelezési munkálatok után *Red Hat Enterprise Linux 5.7 rhel5u7 ib* operációs rendszer lett installálva, amely időközönként frissítésre kerül.

Az operációs rendszer telepítése után következett a rendszer finomhangolása. A három hónapig tartó optimalizálási munka végén sikerült elérni - a kezdetben nem is remélt - igen figyelemre méltó, **3 teraflops** ( $3 \times 10^{12}$  lebegőpontos művelet másodpercenként) számítási teljesítményt, amellyel igen előkelő helyet foglalunk el a Magyarországon található szuperszámítógépet üzemeltetők elit klubjában.



3. ábra. A berendezés a szerver helyiségben



4. ábra. A berendezés nyitott első és hátsó ajtókkal:

- felül a kommunikációs szekció a switchekkel,
- közepén a tároló szekció a merevlemezekkel,
- alul a Head node és a 24 db munkaállomás

A működőképes számítógépes rendszer önmagában csak a lehetőséget teremtette meg a párhuzamos feldolgozás kihasználásához. A nagyságrendekkel magasabb számítási potenciál kiaknázásához az adott feladatokhoz szoftverek, az adott feladatok megoldását lehetővé tevő alkalmazások is szükségesek. Mivel az ilyen szoftverrendszerekből a világon viszonylag kevés van (a személyi számítógéphez képest), a megvásárolható alkalmazások köre is szűk (alig egy-két tucatnyi). Az általunk használni kívánt szoftvercsomagokat a konkrét kutatásokhoz kell megválasztani, illetve speciális igények esetén egyedileg fejleszteni. Ehhez jelentős anyagi erőforrások szükségesek, amelyek előteremtése - pl. pályázatok segítségével - jelenti az előttünk álló évek - a projekt folytatása szempontjából - legfontosabb feladatát, amely egyben a széleskörű alkalmazás, a megfelelő kapacitás kihasználás elérésének elengedhetetlen feltétele is.

Jelenleg az alábbi telepített szoftvereink vannak:

- o CMU (Cluster Management Utility)
- o NVIDIA CUDA SDK (párhuzamos számítások, a GPU kihasználása)
- o CONDOR (HTC) (feladatütemező és elosztó)
- o Modules Software Environment Manager
- o gcc, gfortran (fordító programok)

A számítógépünk, amelynek a **Hercules** nevet adtuk, az alábbi módokon érhető el:

- o Egyetemen belülről:
  - <http://193.6.3.130/>
  - ssh
  - sftp
- o Egyetemen kívülről:
  - ssh
  - 80-as és 1099-es port

Tekintettel arra a tényre, hogy a projekt eredeti célkitűzése egy szuperszámítógép beszerzése és üzembe

helyezése volt - beleértve bizonyos terhelési, skálázhatósági tesztek is -, amely beüzemelés 2011. végén fejeződött be, valamint a felhasználói igényfelmérések éppen, hogy csak elkezdődtek, korai lenne konkrét gyakorlati eredményekről beszámolni.

Az új architektúrájú rendszerünkön a számítási feladatok paralel futtatási képességeit kihasználni képes alkalmazások, szoftverek száma jelenleg még eléggé korlátozott világszerte, és ezen speciális programrendszerek beszerzési ára a hasonló célokra alkalmas szokásos szoftverekéhez képest lényegesen magasabb. További figyelembe veendő tényező az ilyen típusú szoftverrendszerek használatához szükséges ismeretek, tapasztalatok hiánya. E miatt könnyen belátható, hogy különös körültekintést és komoly források előteremtését igényli egy-egy ilyen eszköz megvásárlása. A kompetenciák, tapasztalatok hiánya fokozottan igaz a különböző feladatok megoldására alkalmas saját fejlesztésű szoftverek esetén. Ezen alkalmazások megléte pedig elengedhetetlenül szükséges ahhoz, hogy az új eszköz igénybe vehető legyen valódi feladatok megoldásra, új kutatási projekt beindításra.

Ezért a közeljövő legfontosabb feladata számunkra, hogy minél szélesebb szakmai körben mutassuk be ezen új eszközünk képességeit, lehetőségeit, keltsük fel ilyen szakirányú képzettséggel rendelkező kutatók, szakemberek, fiatal tehetséges hallgatók érdeklődését, valamint a hasonló berendezéssel - és remélhetőleg már tapasztalatokkal is - rendelkező intézményekkel történő kapcsolatfelvételt, aminek révén elindulhat a szuperszámítógép használatához leginkább szükséges szoftverek beszerzésének előkészítése, új alkalmazások fejlesztése, illetve közös kutatási projektek kialakítása. Ez irányban tett első lépéseink egyikeként a Miskolci Akadémiai Bizottság egyetemünkön tartott ülésén előadás keretében mutattuk be a régió tudományos életének legrangosabb képviselői előtt az új berendezést ill. annak használata lévén megnyíló minőségileg új lehetőségeket.

A kapcsolatépítés első eredménye a Budapesti Műszaki Egyetem Irányítástechnika és Informatika Tanszékével történő kapcsolatfelvétel, ahol számos kutatási projekt indult már HPC technológiák alkalmazása terén, és az azokon szerzett tapasztalatok révén sikerült a berendezésünk finomhangolási munkálatai során elérni a 3 teraflops számítási teljesítményt.

#### 4. ÖSSZEFOGLALÁS

Magyarországon a HPC területen csak szórványosan folynak kutatások és csak néhány kutató használhat ilyen rendszereket. Az alkalmazott kutatásoknál is csak néhány olyan alkalmazási területet ismerünk, ahol szuperszámítógép segítségét tudják igénybe venni (pl. autóiipari tervezési és szimulációs számítások). A hazai helyzet javítására központi erőfeszítések is zajlanak, melyek az EU-s források segítségével három nagy egyetemet (Debrecen, Szeged, Pécs) szereltek fel egy-egy szuperszámítógépes kategóriába sorolt beren-

dezéssel és egy országos tárolókapacitás-hálózat megvalósítása is folyik. Sajnálatos módon a mi régióink nem volt érintett ebben a központi programban, de a sikeresen megvalósított projektünk eredményeként a közeljövőben mi is csatlakozni tudunk ehhez a folyamathoz.

A szuperszámítógép révén a kutatók – közülük is elsősorban a fiatal kutatók – bekapcsolódhatnak egy olyan nemzetközi akadémiai és felsőoktatási hálózatba, ahol versenyképes kompetenciákkal új, világszínvonalú kutatásokhoz való kapcsolódás is lehetővé válik.

A pályázat segítségével beszerzett szuperszámítógéppel és a hozzá kapcsolódó magas képzettségű munkaerővel olyan informatikai infrastruktúrát szeretnénk létrehozni, amely a régióban alapját képezheti ezen szakterületnek világszínvonalon való művelését.

A létrehozott rendszerre alapozva, a jövőben a Miskolci Egyetemen széles körben kívánjuk elterjeszteni a szuperszámítógépes ismereteket, alkalmat szeretnénk keríteni a gyakorlati képzés lehetőségeinek bővítéséhez, és indukálni szándékozunk olyan K+F+I tevékenységeket, amelyek révén a kutatók, az oktatók és hallgatók hatékonyan kapcsolódhatnak be mind az EU, mind a szélesebb nemzetközi szakmai közösségek munkájába.

A K+F+I tevékenységek a régióban minőségileg is új szakaszba léphetnek a rendszer üzembe állításával, mert olyan projektek megvalósítására is lehet gondolni, amelyek ezen kapacitás hiánya miatt eddig szóba sem kerülhettek (pl. nagy méretű rendszerek analízise és vizualizációja). A tevékenységek társadalmi hatásaként a Miskolci Egyetem K+F+I tevékenységeiben részt vevő hallgatóinak tudásszintje is minőségileg javul, és a projekt megvalósításának végén már nagyon jó eséllyel helyezkedhetnek el a regionális és nemzetközi munkaerőpiacon.

"A cikk a Miskolci Egyetemen megvalósult fejlesztési eredményeket mutatta be, amely a Nemzeti Fejlesztési Ügynökség támogatásával meghirdetett Baross Gábor Program, Kutatás-fejlesztési Infrastruktúra Fejlesztése alprogram REG\_EM-09-1-2009-0023 számú támogatási szerződés alapján folyt, továbbá részét képezte a TÁMOP-4.2.1.B-10/2/KONV-2010-0001 számú Kiválósági Központok Projektnek".

"*This paper presented the results of a development project at the University of Miskolc based on the REG\_EM-09-1-2009-0023 contract, under the Baross Gábor Project, Development of R&D Infrastructure sub-project supported by the National Development Agency, and also takes part of the TÁMOP-4.2.1.B-10/2/KONV-2010-0001 Centre of Excellence Project*"

#### 5. IRODALOM

- [1] VELEZDI, GY.: Projekt zárójelentés, *Baross Gábor Program, Kutatás-fejlesztési Infrastruktúra Fejlesztése alprogram, 2012.*