

## "ELOSZTOTT" PÁRHUZAMOS SZÁMÍTÁSTECHNIKAI OKTATÁS

Vajda Ferenc, vajda@iif.kfki.hu

KFKI Mérés és Számítástechnikai Kutató Intézet és Budapesti Műszaki Egyetem

### Abstract

The paper outlines an approach to reshape the existing parallel computing curriculum. First as a background information a brief overview of the current parallel and distributed computing curricula at different universities is given. It includes the most important courses of a traditionally organised curriculum. The different faces of parallelism and concurrency are also described. Parallel computing topics as examples are suggested for basic courses of a typical curriculum. The paper concludes with the problems of distributed parallel computing education i.e. faculty members familiarity with the topic, contradictory requirements between future employers and graduate schools and harmonisation of the different subtopics in the individual courses.

### 1. Háttérinformáció

A párhuzamos feldolgozással kapcsolatos oktatás világhelyzetének vázlatos megismerésére úgy gondoltam, hogy a legkézenfekvőbb egy rövid "kirándulást" tenni a számítógépes világhálózaton, az Internet-en tekintve a legfontosabbnak ítélt intézmények könnyen elérhető információiba. Az a néhány óra amit erre fordítottam nagyon hasznosnak bizonyult. Szerencsével (és talán személyes ismereteim felhasználásával is) nagyon sok nagyon érdekes, (és hasznos!) adathoz lehetett hozzájutni. Bár talán ennek a "barangolásnak" az eredménye is megérne egy külön cikket, itt csupán röviden összefoglalom a "kinyert" információkat jellegüknek megfelelő csoportokba szedve. Ezen túlmenően csak a témánkkal közvetlenül összefüggő néhány megállapításra - és abból levonható tanulságra - koncentrálok. Talán még csak annyi előzetes megállapítás, hogy indulásként néhány általam ismert és a szakterületen tevékenykedő professzor egyetemének "homepage"-t kerestem fel és vagy az intézménnyel vagy a személlyel kapcsolatos információ adott lehetőséget a továbblépésre legtöbb esetben közvetlen "linkelés" segítségével a hypertext kiválasztás egyszerű módján.

A begyűjtött nagymennyiségű információ a párhuzamos feldolgozás (általánosan értelmezett) területéről a következő csoportokba sorolható:

#### 1.1. Különböző egyetemi tantárgyak tananyagainak ismertetése

Nagyon nagy számú egyetem oktatási anyaga a hálózaton keresztül hozzáférhető. (Ez összefügg a hallgatói adminisztráció kezelésének módszereivel és eszközeivel is, de ez a téma nem része ennek a cikknek.) A tantárgyak nagyobbik fele nem a törzsanyag (core curriculum) része, hanem szabadon választható illetve nem az alapképzéshez (undergraduate), hanem a graduális oktatáshoz tartozik.

Néhány kiragadott példa:

- CS 267: Application of Parallel Computers (Berkeley)
- Parallel Computer Architectures and Systems (University of Crete)
- Occam 2 and Transputer Engineering (Univ.of Kent)
- The Scalable Parallel System Model (George Mason University)

#### 1.2. A párhuzamos számítástechnika online bibliográfiái

Ismét csak néhány illusztratív példa:

- Bibliográfies on Parallel Processing (Három összehangolt rendszer Karlsruhe-ben, Arisonában és Manitobában)
- Index for PAPERS (Purdue)
- Parallel Computing Archive at HENSA (UK)

### 1.3. Párhuzamos számítástechnikai szabványok

Például az elosztott rendszerek (munkaállomás "farmok" ) kiemelten fontos szabványa a Message Passing Interface (Tutorial és Document)

### 1.4. A párhuzamos számítástechnika oktatásával kapcsolatos dokumentumok

Néhány érdekes kiemelt példa ismét:

- High Performance Computing Consortium (UK)
- High Performance Computing and Communications Glossary
- Russ Miller's Report on Parallel Computing Education (SUNY Buffalo)

Utóbbi report az Egyesült Államok legfontosabb 100 egyetemének adatait dolgozta fel megadva az alap és graduális oktatás keretében szereplő párhuzamos számítástechnikai tárgyakat (illetve az egyes területen oktatott tárgyak számát) [1].

### 1.5. Párhuzamos számítástechnikai könyvek

Ebbe a csoportba kétféle információ sorolható:

- Könyvek tartalomjegyzékei és ismertetései (Független forrásokban és kiadók propagandaanyagaiként)
- "Online" könyvek általában még kísérleti fázisban (például Caltech's Archetypes "Experimental Electronic Textbook on Sequential and Parallel Algorithms")

### 1.6. "Szabad" párhuzamos szoftver

Nagyon nagyszámú forrásból lehet értékes párhuzamos számítástechnikai szoftvert "letölteni". Ismét csak néhány kiragadott példa:

- Netlib Repository
- The Parallel Tools Consortium
- The C++ Programming Language (Caltech)
- DAPPLE - Data Parallel Programming Library for Education
- PIOUS (Parallel I/O System) for PVM

### 1.7. Párhuzamos számítástechnikai kutatás

Néhány példa a sok közül:

- CMU Supercomputing and Parallel Computing Research Group
- Caltech Center for Advanced Computing Research
- Cornell Theory Center
- Paderborn Center for Parallel Computing (Germany)

A különböző párhuzamos számítástechnikai központoknál elérhető információ nagyon sokrétű (konferenciák és egyéb újdonságok, munkatársak, tevékenységek, publikációk), amelyekben keresni lehet, és

számos új cikk vagy report közvetlenül lehívható. A központok egymásra is "mutogatnak" ami a továbbhaladást nagyon megkönnyíti. (Ez a megállapítás természetesen igaz az előző és következő információ csoportokra is.)

### 1.8. Párhuzamos számítástechnikai cégek

A szakterülethez tartozó gyártók mindegyike széleskörű információval van jelen a hálózaton. Itt talán példákra se volna szükség, de azért egy-két talán kevésbé ismert (?) cég nevét megemlítem:

- Convex Computer Corporation
- Parsytec
- Thinking Machines Corporation
- Yahoo

### 1.9. Párhuzamos számítástechnikai nyelvek

Ízelítőül néhány - a hálózatról megkapható - programozási nyelv (természetesen a szükséges dokumentáció is letölthető):

- Fortran 90 (CERN)
- KROC - Occam for all (Univ. of Kent)
- Sisal - High Performance, Portable, Parallel Programming Language (LLNL)

### 1.10. A párhuzamos számítástechnika vezető kutatói

Az egyetemi fakultások tagjairól rendelkezésre álló részletes információ könnyűvé teszi az egyes személyek megítélését, az ismertek elérését vagy új "ismerősök" megtalálását.

Ismét csak néhány fontos név "mutatóul":

- Russ Miller (SUNY Buffalo)
- H.J. Siegel (Purdue)
- Alan Edelman (MIT)
- Quention F. Stout (U. Michigan)

Végül néhány rövid tanulság:

- A hálózat ennek a szakterületnek is hatalmas tárháza már ma is.
- Érdemes "körülnézni" mielőtt bármit is teszünk (például cikkírás, témadefiniálás, erőforrás bővítés, szoftver beszerzés, laboratóriumi tevékenység, előadási tematika vagy új tárgy összeállítása, stb.)
- Különösen a graduális (doktorandusz) képzésben a hallgatók számára nagyszámú önálló feladat határozható meg a hálózati "keresés" bázisán.
- A párhuzamos számítástechnika különböző "fedőnevek" alatt (például Supercomputing, High Performance Computing, Scientific Computing, Distributed Computing, stb.) minden egyetem vonatkozó oktatási programjában jelen van különböző tárgyakkal és célkitűzésekkel. Ezzel kapcsolatos következtetésekre még a továbbiakban részletesebben visszatérünk.

## 2. Párhuzamosság a számítástechnikai oktatási tematikákban

A párhuzamossággal kapcsolatos problémakörök már hosszabb ideje különféle tárgyak formájában megjelentek az egyetemek elsősorban graduális oktatásában [3] [4]. A párhuzamos és elosztott számítástechnika jelentőségére felhívták a figyelmet az Európai Unió különféle programjai (például CaberNet

program, amely egy ESPRIT által támogatott alapkutatási program: "Basic Research Network of Excellence in Distributed Computing Systems Architectures [5]) vagy az Amerikai Egyesült Államok legmagasabb állami szinten indított programja (1992 US Presidential Initiative in High-Performance Computing and Communication) [8].

A világszerte ismert ACM/IEEE alap oktatási tematika javaslat (Computer Curricula 1991) [2] is már számos olyan "tudásegységet" tartalmaz, amelyek a párhuzamos és elosztott számítástechnikával kapcsolatosak. A gyakorlatban azonban az alap (undergraduate) oktatásban csak a szabadon választható (elective) tárgyak között szerepelnek ilyenek az oktatási ciklus utolsó negyedében.

Az ilyen típusú - koncentráltam a párhuzamosság és elosztott rendszerek problémáival foglalkozó - tárgyak "hagyományos" módon bontják fel a tématerületet például a következő tárgyak szerint [6]:

- Bevezetés a párhuzamos számítástechnikába
- Párhuzamos és elosztott számítógép architektúrák
- Párhuzamos programozási nyelvek
- Párhuzamos algoritmusok
- Párhuzamos rendszerek
- Párhuzamos alkalmazások
- Elosztott adatbázisrendszerek
- Korszerű operációs rendszerek

Arra vonatkozóan, hogy kinek kell a párhuzamos programozás témaköreit tanítani elég eltérő vélemények vannak forgalomban. (Itt csak röviden utalunk a számítástechnika elterjedésének kezdeti korszakára. Akkor olyan extrém megállapításokat is lehetett hallani, hogy például már a középiskolában mindenkinek oktatni kell a Boole algebrát, mert ez a számítástechnika használatának a szükséges feltétele.)

Az informatikusok szívesen képviselnek olyan álláspontot, hogy olyan érdekes (és hasznos) ismereteket mind a párhuzamos és elosztott számítástechnika mindenkinek (minden mérnök és fizikus hallgatónak, tanárjelöltnek, stb.) oktatni kell, mert ez a "jövő útja". Ugyanakkor a szakterület egész fejlődése azt igazolja, hogy a viszonylag bonyolult megoldások igazi elterjedésének a feltétele a megfelelő "felhasználói kapcsolat" amikor is a bonyolultság már eltűnik illetve "átlátszóvá" válik a felhasználók számára. Igazat kell adni egy nemrégiben megjelent cikk [7] megállapításának, hogy a nem közvetlen szakterület hallgatóit meg kell "kímélni" a párhuzamos számítástechnika részleteitől. (A cikk ezt egy régi közmondással vezeti be: "Never try to teach a horse to whistle - it wastes your time and annoys the horse" azaz sohase próbálj egy lovat fütyülni tanítani - ezzel csak az idődet vesztegeted és a lovat is bosszantja.)

Az informatika (a számítástechnika és számítástudomány) esetén azonban más a helyzet. Itt a megközelítés többféle lehet, a számítástechnikai tantervek tárgy szemléleteinek megfelelően vagyis

- eszköz
- szemlélet
- jelenség

lehet a tárgyalás alapvető rendezője.

A párhuzamos számítástechnika tárgyai is definiálhatók (legalább is a speciális, graduális tárgyak esetén) ezeken az alapokon a tárgy témájának és jellegének függvényében. A számítástechnika - számítástudomány, (másképpen fogalmazva az informatika gyakorlat és elmélet orientált területei) nem tárgyalható(k) korszerű szinten a párhuzamos és elosztott rendszerek gyakorlati és elméleti problémái nélkül.

### 3. "Elosztott" párhuzamos számítástechnikai oktatás

Szinte lehetetlen számba venni, hogy a párhuzamosság mennyi különböző formában jelenik meg a számítástechnikában/informatikában. Álljon itt csupán néhány példa csak illusztrációként:

Párhuzamosság a hardver architektúrában:

- "Alacsony szintű" alrendszerek: közvetlen tárhozzáférés (DMA), csővonal kiképzésű szorzók, átvitel előképző összeadók (carry-lookahead adders), versenyhelyzet a sorrendi áramköröknél

- "Magas szintű" alrendszerek: társprocesszorok, utasítás szintű párhuzamosság (szuperskaláris processzorok), hálózatok, VLIW architektúrák
- Multiprocesszálás: aszinkron kommunikáló folyamatok, durva és finom "szemcsézettségű" párhuzamosság, adatfolyam paradigma, processzor tömbök, munkaállomás farmok

Párhuzamosság a nyelvekben és fordítóprogramjaikban:

- Explicit párhuzamosság: üzenet átadás, szinkronizációs primitívek, megosztott és lokális címmező
- Implicit párhuzamosság: nem eljárás-alapú programozás, funkcionális programozás, logikai programozás, adatfolyam nyelvek

Párhuzamosság az algoritmusokban:

- Kommunikációval kapcsolatos: kommunikációs primitívek, üzenet átadás
- Algoritmikus paradigmák: szimmetriával kapcsolatos, párhuzamos előtag, permutációs elérés

Az itteni kiragadott példák is mutatják, hogy a párhuzamosság a hagyományos tárgyak spektrumához tartozó nagyszámú fogalmat, megoldást, új alapelvet hozott be. Ezek legkézenfekvőbbben nem koncentrált formában, hanem a megfelelő tárgyakban "elosztva" tárgyalhatók. A teljesség igénye nélkül ennek bemutatására néhány kiragadott tárgy keretében adunk példákat a tárgyak tematikájához kapcsolódó párhuzamossággal kapcsolatos témákról.

Számítógép szervezés és architektúra:

- Alapvető párhuzamos architektúrák: SIMD és MIMD
- Rejtett tár koherencia
- Üzenet átadás

Operációs rendszerek:

- Szinkronizáció
- Holtpont
- Folyamatok közötti kommunikáció
- Megosztott és elosztott tár modell

Programozási nyelvek:

- Parallel C és C++
- Linda
- PVM
- Ada
- Occam 2
- Indeterminizmus

A számítás elmélete:

- Párhuzamos számítás elméleti modelljei
- CSP
- Petri háló

Algoritmusok elmélete:

- Komplexitás analízis
- Párhuzamos algoritmusok
- Például párhuzamos FFT
- Párhuzamos válogatás
- Sebességnövekedés analízis

Természetesen a sort lehetne folytatni olyan - gyakran csak szabadon választható - tárgyakkal is mint a numerikus analízis vagy diszkrét események szimulációja.

#### 4. Következtetések

A párhuzamos megoldások mind a gyakorlati számítástechnikai rendszerben mind az azokat alátámasztó elvekben és eljárásokban egyre nagyobb jelentőséget kaptak az elmúlt időszakban. Ennek a fejlődésnek az

informatika oktatásra is közvetlen hatása van. A régebbi időszak gyakorlata az volt, hogy - elsősorban választható formában - koncentráltan párhuzamos számítás-orientált tárgyak képezték ennek a területnek az oktatási hátterét.

A javasolt megoldás - a szükséges speciális tárgyak megtartása mellett - "elosztani" az ismereteket a vonatkozó alaptárgyak témaköreinek megfelelően és így oktatni a vonatkozó ismereteket.

A tárgyak tematikájába "beépítendő" párhuzamos számítástechnikai ismereteknél a következő szempontokat kell figyelembe venni:

- A gyors integrálásnak akadálya lehet, hogy az egyes oktatóknak meg kell ismerkedni a tantárgyukhoz kapcsolódó új anyaggal.
- A párhuzamos számítástechnikai ismereteknek "elosztott" elhelyezése a tárgyak tematikán belüli összehangolását (definíciók, jelölések, értelmezések, stb.) még jobban megnehezíti.
- Az anyagmennyiség optimális meghatározásánál azt a szokásos ellentmondó kívánalmat is ki kell elégíteni, amely az alapképzésnél a végzősök közvetlen munkavállalása (a munkahelyek kívánalmai) illetve graduális képzésben való továbbtanulása (doktori iskolák kívánalmai) között van.

### Hivatkozások

1. R. Miller: The Status of Parallel Processing Education  
Computer, Vol.27. No.8. Aug.1994, 40-43.
2. A. Tucker et al.: Computing Curricula 1991 (Report of the ACM/IEEE - es Joint Curriculum Task Force), ACM Press, New York and IEEE CS Press, Los Alamitos, Calif. 1991
3. C. H. Nevison: Parallel Computing in the Undergraduate Curriculum  
Computer, December 1995, 51-56.
4. A. L. Rosenberg: Thoughts on Parallelism and Concurrency in Computing Curricula  
Proc.of the Wellesley Forum on Parallel Computing Curricula (April 1995). Eds.: T. Metaxas and M. Merzbacher (Internet paper)
5. Cabernet (Computing Architectures for Basic European Research)  
ESPRIT Basic Research Network of Excellence in Distributed Computing Systems Architectures, June 1992
6. M. Paprzycki and J. Zalewski: Focusing Undergraduate Curriculum on Parallel Computing  
Proc.of the Wellesley Forum on Parallel Computing Curricula (April 1995) Eds.: T. Metaxas and M. Merzbacher (Internet paper)
7. G.V. Wilson: Teaching Horses to Whistle: An Apostate's View of Parallel Computing and the Undergraduate Computational Science Curriculum. Proc.of the Wellesley Forum on Parallel Computing Curricula (April 1995) Eds.: T. Metaxas and M. Merzbacher (Internet paper)
8. From Desktop to Teraflop: Exploiting the U.S. Lead in High Performance Computing (NSF Blue Ribbon Panel on High Performance Computing), NSF 1993