



**Nagy teljesítményű szuperszámítógépek a
tudomány szolgálatában**

Újfalussy Balázs

MTA Wigner Fizikai Kutatóközpont



Vázlat

- Mi is az a szuperszámítógép?
 - A Dongara féle 500-as lista
 - A lista fejlődése
 - A modern szuperszámítógépek jellemzői
- Masszívan parallel szuperszámítógépek
 - Parallel programozás
 - Párhuzamos végrehajtású-e a fizika? A párhuzamos programozás nehézségei
 - Képességi és mennyiségi számítástechnika
- Tudomány szuperszámítógépeken
 - Jelenlegi fő projektek
 - Klíma modellezés, robbanás (szupernóva robbanás, combustion)
 - Anyagtudomány: a természet rövidlátó





Mi is az a „Szuperszámítógép”



<http://www.top500.org/>

„Szuperszámítógép az, ami rajta van a listán”

A világ első szuperszámítógépe a CDC6600.
tervezője Seymour Cray



Hogyan mérjük egy számítógép teljesítményét?

A LINPACK Benchmark

Jack Dongara (University of Tennessee, Knoxville)

<http://www.netlib.org/benchmark/performance.ps>

<http://www.netlib.org/utk/people/JackDongarra/PAPERS/hpl.pdf>

Egy lineáris egyenletrendszer megoldó algoritmus és program műveletigénye

$$N_D = \frac{2}{3}n^3 + 2n^2$$

Duplapontos lebegőpontos művelet, $n \sim 100$ vagy 1000

Mértékegysége Mflop/s, Gflops/s, Tflop/s (néha Mflops, Gflops, etc)

Hogyan mérjük egy számítógép teljesítményét?

Elméleti csúc: (theoretical peak performance):

Egy órajelciklus alatt elvileg elvégezhető lebegőpontos műveletek (+,*) számából
Pl. Cray Y-MP/8 órajele 6ns-os, ezalatt 1 vektorprocesszor 1 összeadást és 1 szorzást végez:

$$\frac{2\text{muvelet}}{1\text{ ciklus}} * \frac{1}{6\text{ns}} = 333\text{Mflop/s}$$

A Cray Y-MP/8 elméletileg elérhető legnagyobb teljesítménye így 2667 Mflop/s

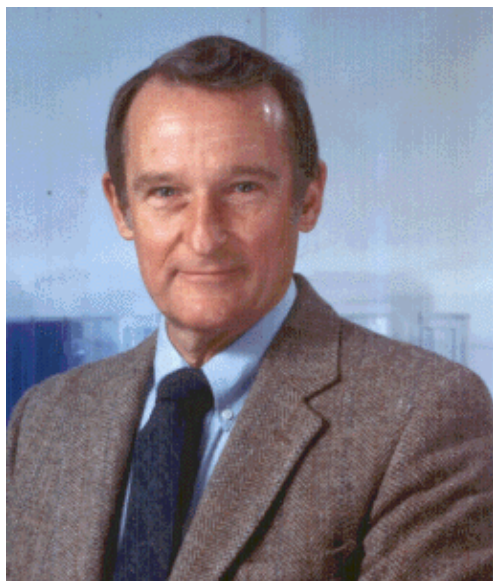
Mért LINPACK sebesség: LINPACK bechmark

A számolt műveletek N_b számából és a végrehajtási időből származtatott
Függhet az operációs rendszertől, fordítóprogramtól

Mért valós sebesség: Valamilyen alkalmazás esetén a mért műveletek számából és a végrehajtási időből származtatott mennyiség

Függhet az alkalmazástól, a fordítóprogramtól, parallelizációtól

A Cray supercomputerek



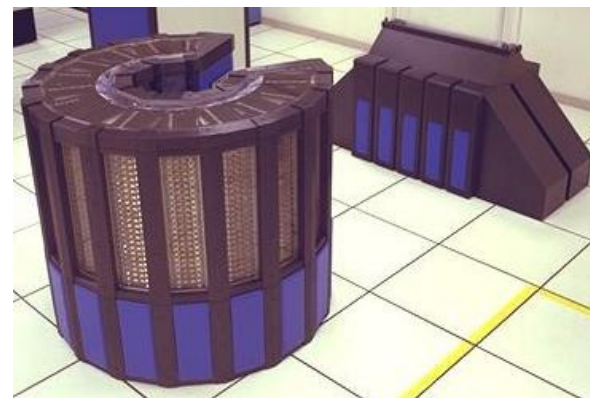
1960: CDC6600, CDC 6700 (Control Data Corporation)

1970: CDC8600 4 processzor, shared memory

1972: Cray Research Inc.

1976: Cray 1, Los Alamos, 8.8 millió dollár, 160 Mflops, 8 Mb

From Computer Desktop Encyclopedia
Reproduced with permission.
© 1996 Cray Research, Inc.



1985: Cray 2 az első vektor-szuperszámítógép, fluorinert hűtés, 1.2 Gflops

1988: Cray Y-MP 2.3 Gflops, 4 processzor

1993: Cray T3D az első masszívan párhuzamos (Cray)

1998: Cray T3E-1200E: 1.02 Tflops

Cray Research → SGI → CRAY Inc.

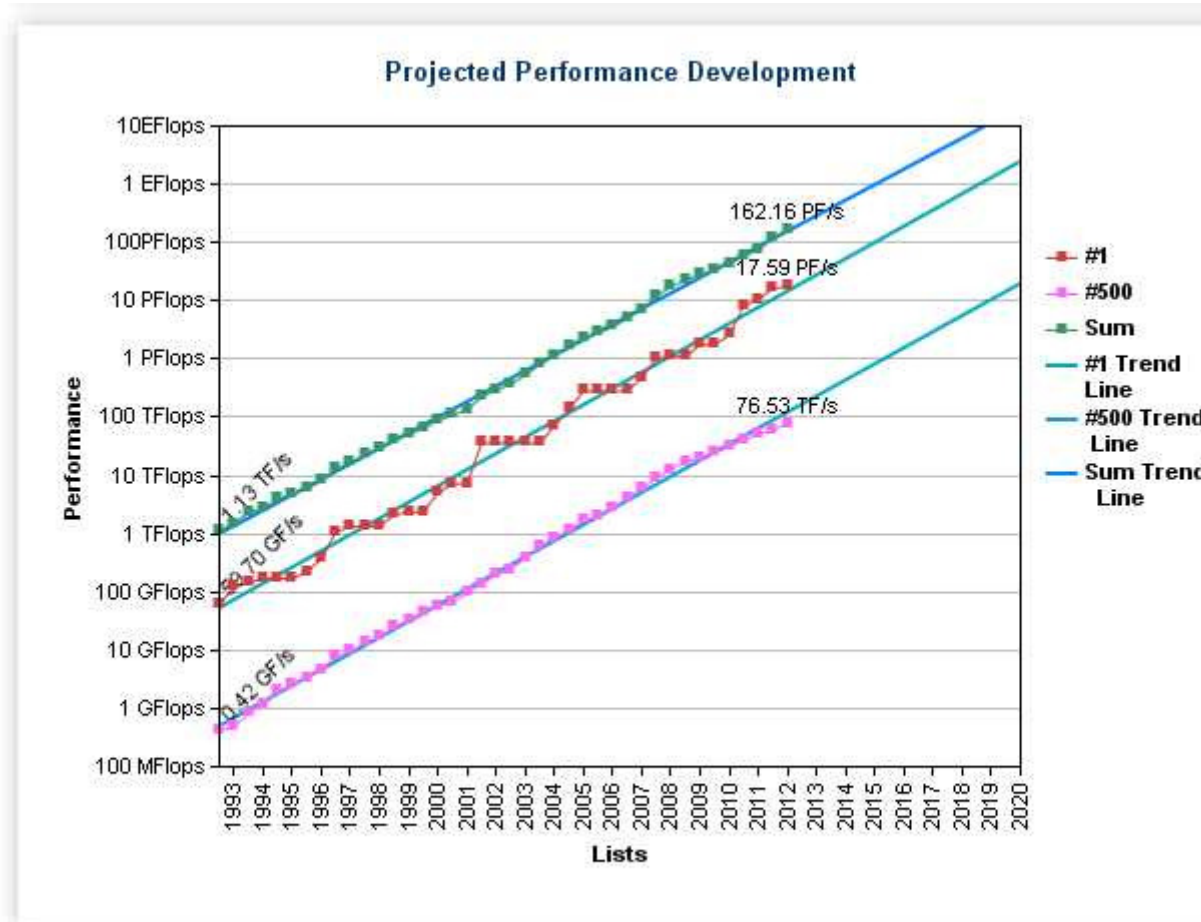
A szuperszámítógépek fejlődése

1993: 7th Mannheim Supercomputer Seminar: TOP500 Lista első kiadása

Év	Computer	Core-ok	Linpack	Elméleti	Mért Valós
1993	Thinking Machines CM5	1024	59,70	131,00	<u>1988 1 Gflop/s Cray Y-MP</u>
1993-1995	Fujitsu	140	124,20	235,80	
1994	Intel Paragon	3680	143,40	184,00	
1996	Hitachi SR2201	1024	232,40	307,20	
1997	Hitachi CP-PACS	2048	368,20	614,40	
1997-2000	ASCI RED (Intel)	9632	1068-2379	3207,00	<u>1998 1Tflop/s Cray T3E</u>
2000-2001	ASCI WHITE (IBM)	8192	7226,00	12288,00	
2002-2004	Earth Simulator (NEC)	5120	35860,00	40960,00	
2008-2009	IBM Roadrunner	122400	1105,00	1456,70	<u>2009 1.03 Pflop/s Cray XT5</u>
2009-2010	Cray XT5 "Jaguar"	224162	1759,00	2331,00	
2010-2011	Tianhe-1A	186368	2566,00	4701,00	
2011-2012	K-computer (Fujitsu)	548352	8162,00	8773,60	
2012	IBM Blue Gene "Sequoia"	1572864	16324,00	20132,70	
2012	Cray XK7 "Titan"	552960	17590,00	27112,50	

Exaflop?

A szuperszámítógépek fejlődése



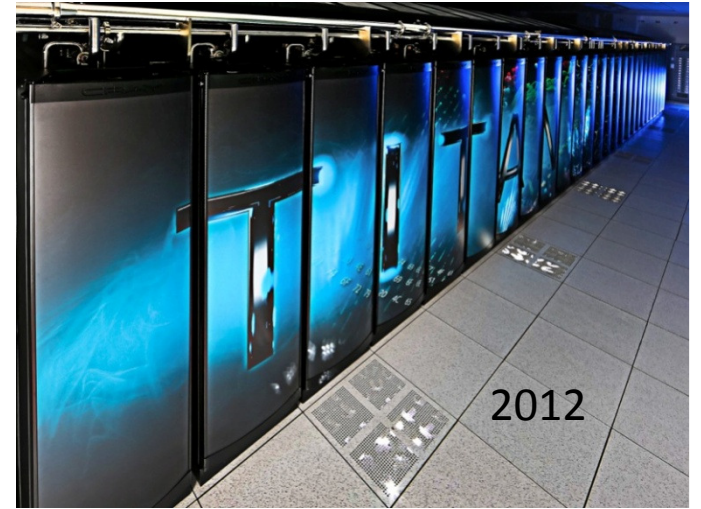
<http://www.top500.org>



1993



1994



2012



1996



1997-2000



2000-2001



2009



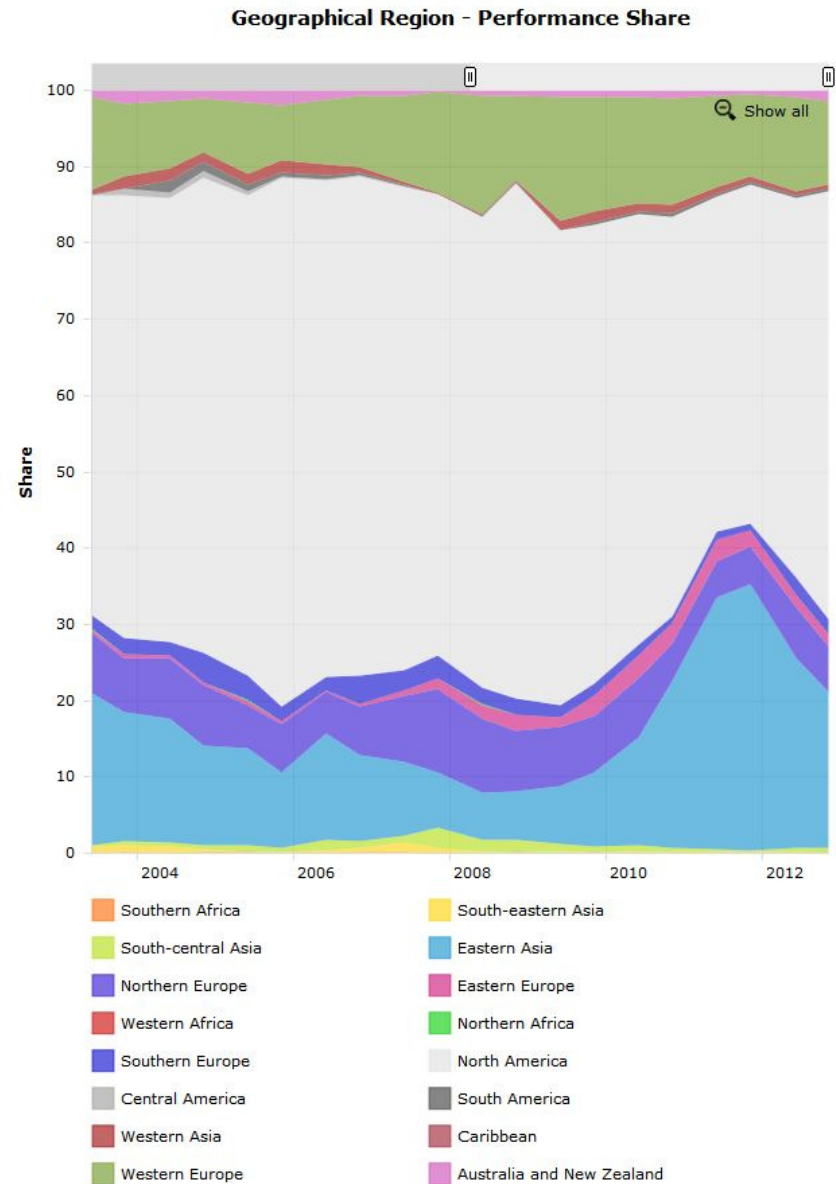
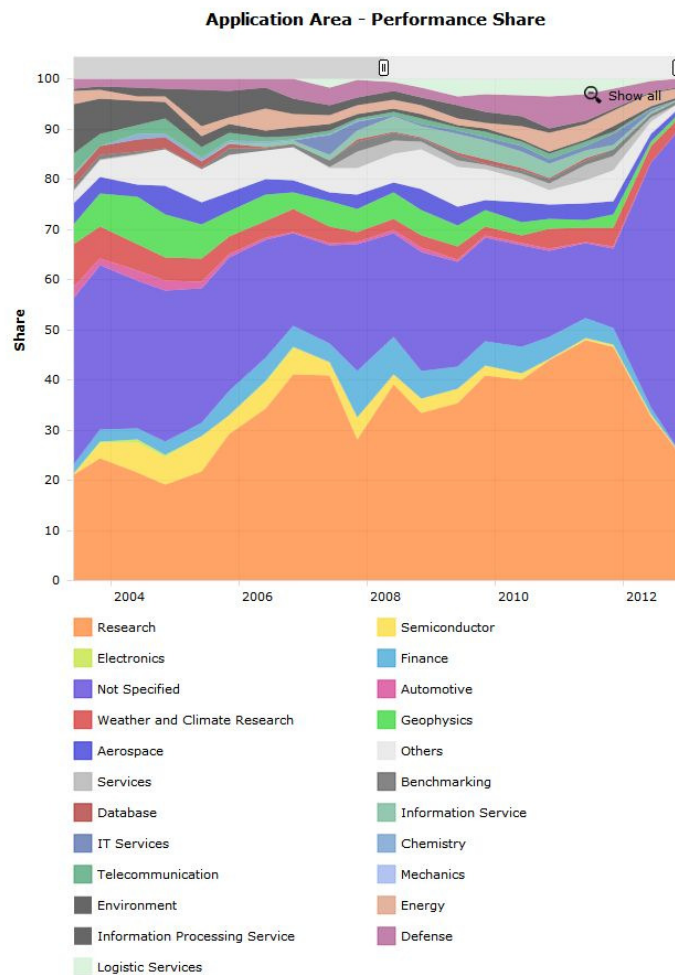
2011

©RIKEN

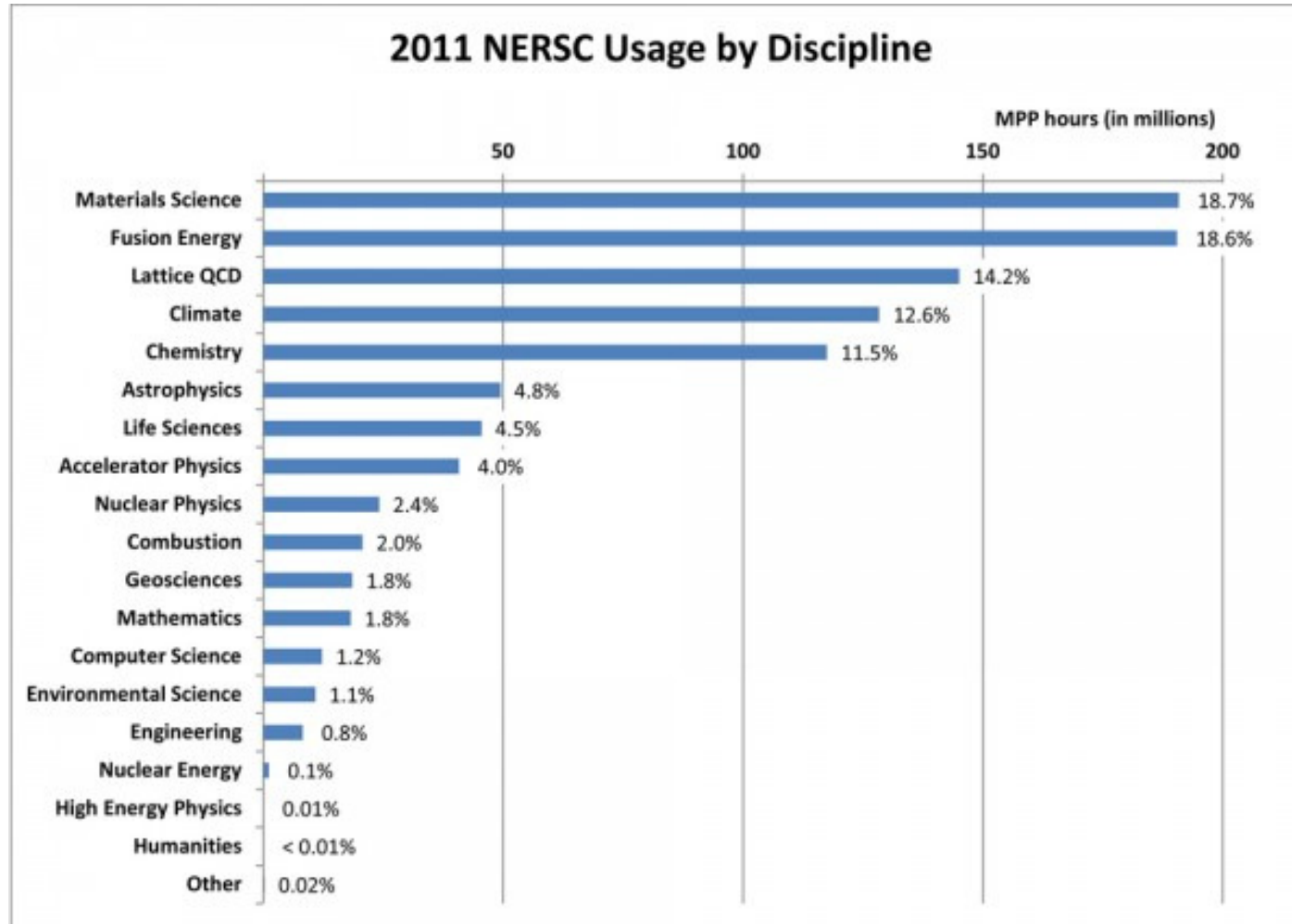


1993-1995

Hol vannak szuperszámítógépek? Mire használják őket?



Projectek megoszlása a NERSC-i szuperszámítógép központban



A jelenlegi bajnok: a „Titán”



„Hibrid” architektúra:

18 688 nodus

1 nodus: 16 core (AMD Opteron)

+ 32Gb RAM

+ 1 db NVIDIA Kepler K20 GPU (6Gb RAM)

+ Gemini interconnect

Összesen 299,008 CPU core

Elméleti csúcs 27.112 Pflop/s

LINPACK csúcs 17.590 Pflops

Teljesítményfelvétel 8.029 MW

Alapterület: 404 m²

Futottak még....

Elméleti csúcs

helyezés	név	Core-ok száma	hely	Teljesítmény Tflops/s	Áramfelvétel kW
2	Sequoia	1572864	USA (LLANL)	20132.7	7890.00
3	K-computer	705024	Japán	11280.4	12659.89
4	Mira	786432	USA	10066.3	3945
5	JUQUEEN	393216	Németország	5033.2	1970
6	SuperMUC	147456	Németország	2897	3422.67
7	Stampede	204900	USA	2660.3	NA
8	Tianhe-1A	186368	Kína	2566.0	4040

LINPACK csúcs

Szoftver környezet

- Unix operációs rendszer (Linux variánsok)
- Parallel könyvtárak és környezet
 - SMP
 - MPI
 - Cuda
- Kötegelt feldolgozás
- Queue rendszer (PBS)

Hibrid architektúrákon akár egyszerre is

```
> cat script.pbs
#PBS -A ABC123
#PBS -l 24:00:00
#PBS -l size=124000

cd /tmp/work/$USER
date
aprun -n 124000 a.out

> qsub script.pbs
99955051.nid00004
>
```

jobnév

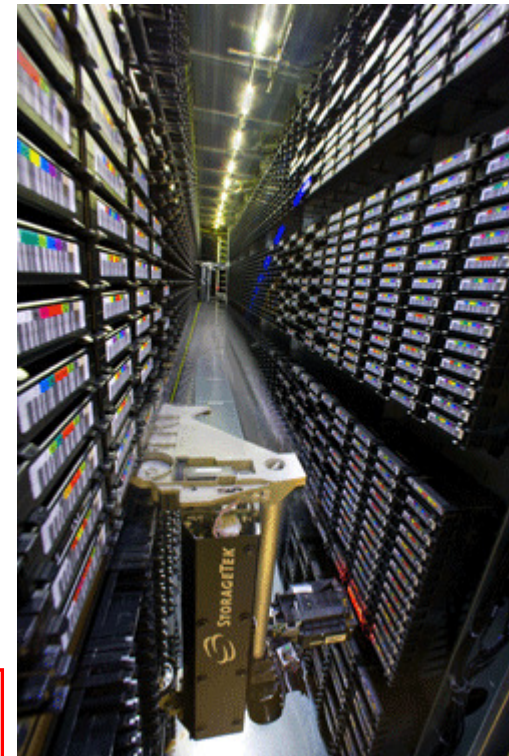
Max idő

Max memória

Temporáris könyvtár, nagy tárterület

Futás!

Példa, kötegelt feldolgozású állomány



HPSS

Tudomány a szuperszámítógépeken

Biológia: Emberi bőrfelület modellezése (65 millió cpu óra), Biomassza (78),
Biomolekuláris rendszerek (110)

Kémia: Belső égés (100), QED (30), Funkcionális soktestprobléma (27)

Számítástechnika: Exascale (21), PEAC (Performance Evaluation and Analysis Consortium) (45)

Mérnöki tudományok: Robbanás (45), Véráramlás szimuláció (51), Áramlás (34)

Földtudományok: Klíma (60) (+Gaia), Földrengéskutatás (45), (100)

Anyagtudomány: Nano-transzport (25), Mágneses anyagok (45), Nem-kovalens kötések (55),
Energiaipari anyagok(45), Hidrogén és jég (45), Li-ion elemek (45),
Véges hőmérsékletű mágneses anyagok (110)

Fizika: Szupernóva (35), Plazma (30), Kozmológia (35), Szupernóva II (55), Magszerkezet (74),
Rács QCD (140)

Sok-e 65 millió CPU óra?

Soros számítógépen igen! 65 millió CPU óra ~ 7420 év

Ugyanakkor 500 000 processzoron ~ 5.4 nap

Képességi alkalmazások: a futáshoz szükség van a teljes számítógépre, vagy annak szignifikáns részére

Mennyiségi alkalmazások : Minél több felhasználó (relative kis) igényeinek kielégítése

Éppen mit futtat?

19	DOE/SC/LBNL/NERSC United States	Hopper - Cray XE6, Opteron 6172 12C 2.10GHz, Custom Cray Inc.	153408	1054.0	1288.6	2910
----	------------------------------------	---	--------	--------	--------	------

Cray XE6 „Hopper”
6384 Node
153216 core
1287 Tflop/s elméleti csúcs
211.5 TB memória
(1.42Gb/core)
2.2MW

www.nersc.gov

NOW COMPUTING 2013.02.19 22:03

A small sample of massively parallel scientific computing jobs running right now at NERSC.

PROJECT	MACHINE	CPU CORES	CPU CORE HOURS USED
Investigations of advanced ignition physics and extreme states of matter PI: Warren Mori, University of California Los Angeles	Hopper	20,016	117,401.1
Simulations of Conformational Changes in Biomolecular Complexes PI: Jeff Wereszczynski, University of California San Diego	Hopper	19,152	89,249.9
Data Analysis for EXO - the Enriched Xenon Observatory PI: Giorgio Gratta, Stanford University	Hopper	16,416	33,964.2
Multiscale Simulations of Particle-, Molecule-Surface Interactions, simulations of nanowires: Structure, dynamics, and quantum transport, and structure and electronic structure of high Tc materials PI: Hai-Ping Cheng, University of Florida	Carver	256	127.8
Spin-lattice Coupling in Magnetic Phase Transition PI: Yi Wang, Pennsylvania State University	Carver	256	1,161.8

Nézzünk egy alkalmazást!

Többszörös szórás algoritmus

- Többszörös szórás elmélete (MST)
 - *J. Koringa, Physica 13, 392, (1947); W. Kohn, N. Rostoker, PR, 94, 1111, (1954)*
- MST Green függvény MST-ben
 - *B. Gyorffy, and M. J. Stott, "Band Structure Spectroscopy of Metals and Alloys", Ed. D.J. Fabian and L. M. Watson (Academic 1972)*
 - *J. S. Faulkner and G. M. Stocks, PR B 21, 3222, (1980)*

- Green függvény egy „adag” szóróra

$$G(r, r'; \varepsilon) = \sum_{LL'} \left[Z_L^n(\mathbf{r}_n; \varepsilon) \tau_{LL'}^{nn'} Z_{L'}^{n'}(\mathbf{r}_{n'}; \varepsilon) \right]$$

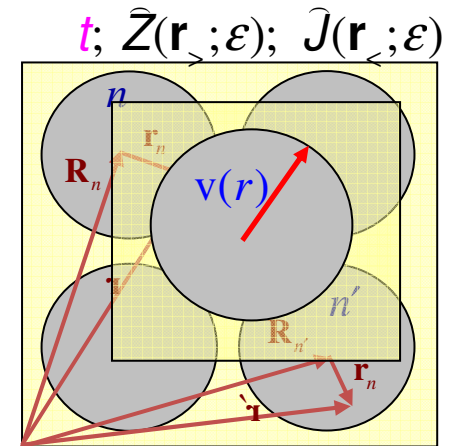
$$\tau^{nn'} = \mathbf{M}^{-1} |^{nn'} \quad M_{LL'}^{nn'} = t_{LL'}^{-1} \delta_{nn'} - G_{LL'}^{nn'}$$

- Töltés- és Mágneszettségsűrűség, Állapotsűrűség:

$$n(\mathbf{r}) = -\frac{1}{\pi} \text{Im} \int d\varepsilon f(\varepsilon - \mu) \text{Tr} G(\mathbf{r}, \mathbf{r}; \varepsilon)$$

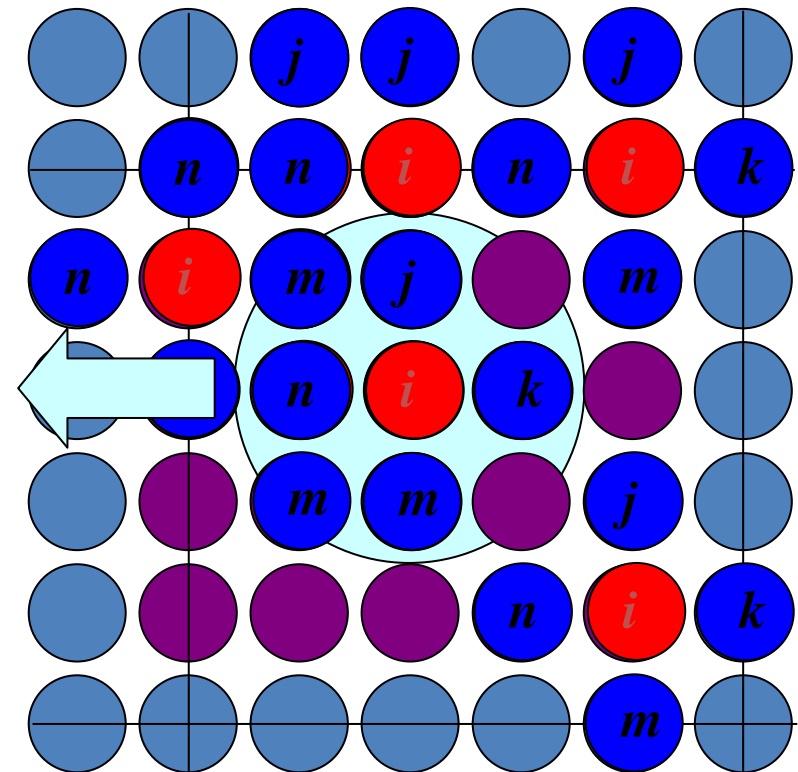
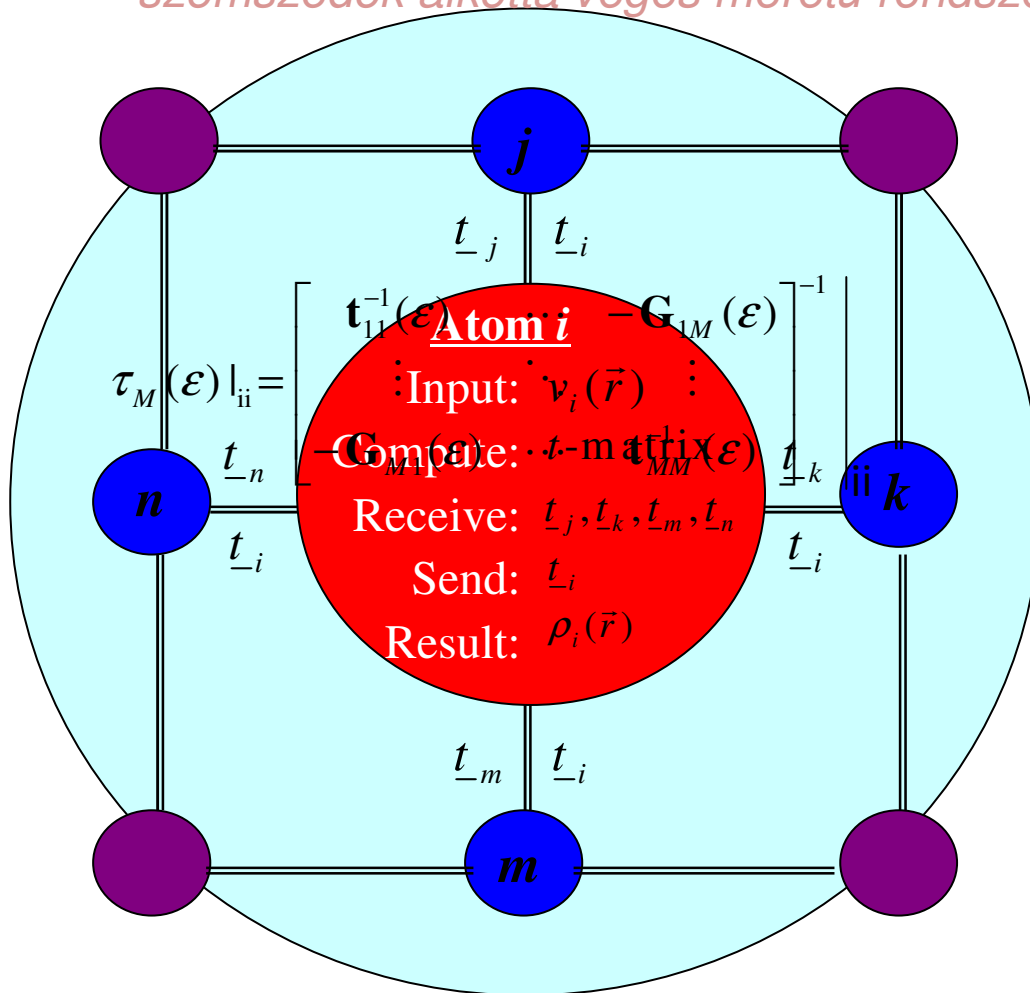
$$m(\mathbf{r}) = -\frac{1}{\pi} \text{Im} \int d\varepsilon f(\varepsilon - \mu) \tilde{\sigma} \text{Tr} G(\mathbf{r}, \mathbf{r}; \varepsilon)$$

$$n(\varepsilon) = -\frac{1}{\pi} \text{Im} \int d\mathbf{r} f \text{Tr} G(\mathbf{r}, \mathbf{r}; \varepsilon)$$



Lokálisan önkonzisztens többszörös szórás (LSMS)

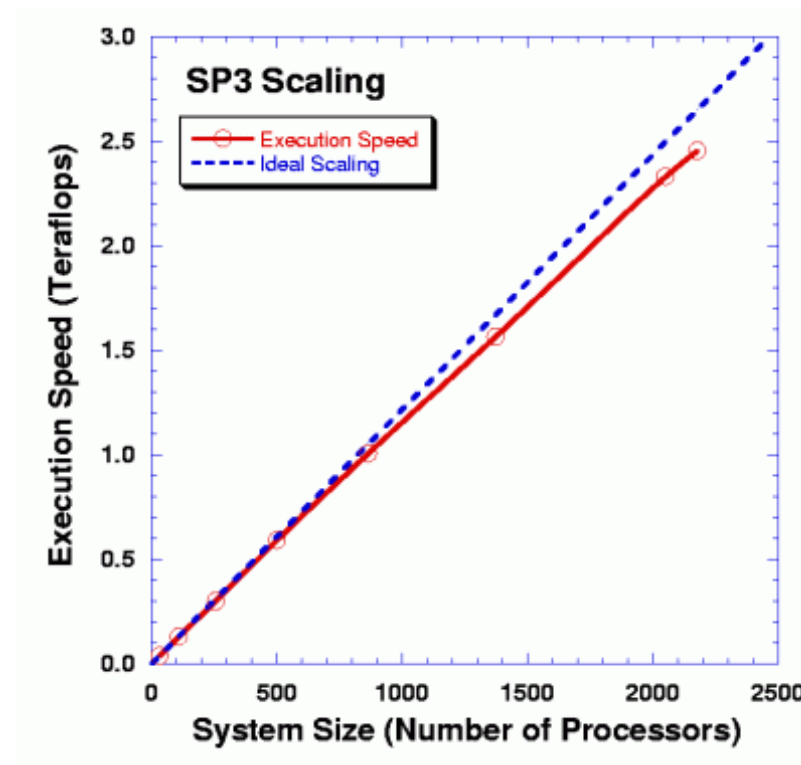
- Masszívan parallel $O[N]$ LSMS megközelítés: a természet rövidlátó
 - Közelítsük a teljes elektron sűrűséget helyileg meghatározott sűrűségekkel
 - A i -edik rácshelyen közelítsük a végtelen rendszer tau mátrixát az n -edik szomszédok alkotta véges méretű rendszer tau mátrixával



LSMS - gyenge skálázás a processzorok számával

LSMS: 1 atom 1 processzor, akkor **gyenge** skálázás:
Ha több a processzor, akkor a rendszer is nagyobb

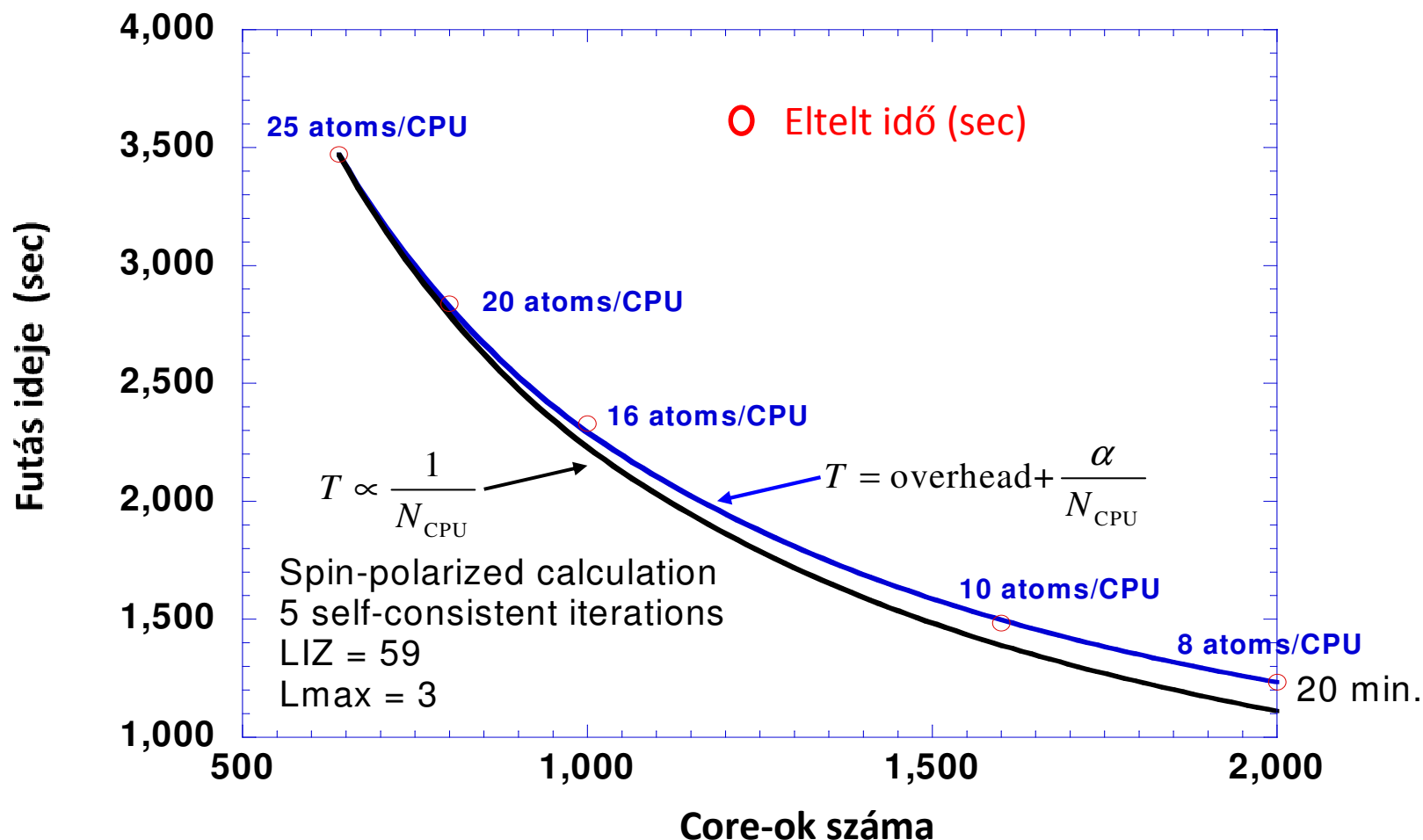
IBM SP, NERSC



- IBM SP RS600
 - Power 3 375 Mhz
 - 136 16-Processor nodes.

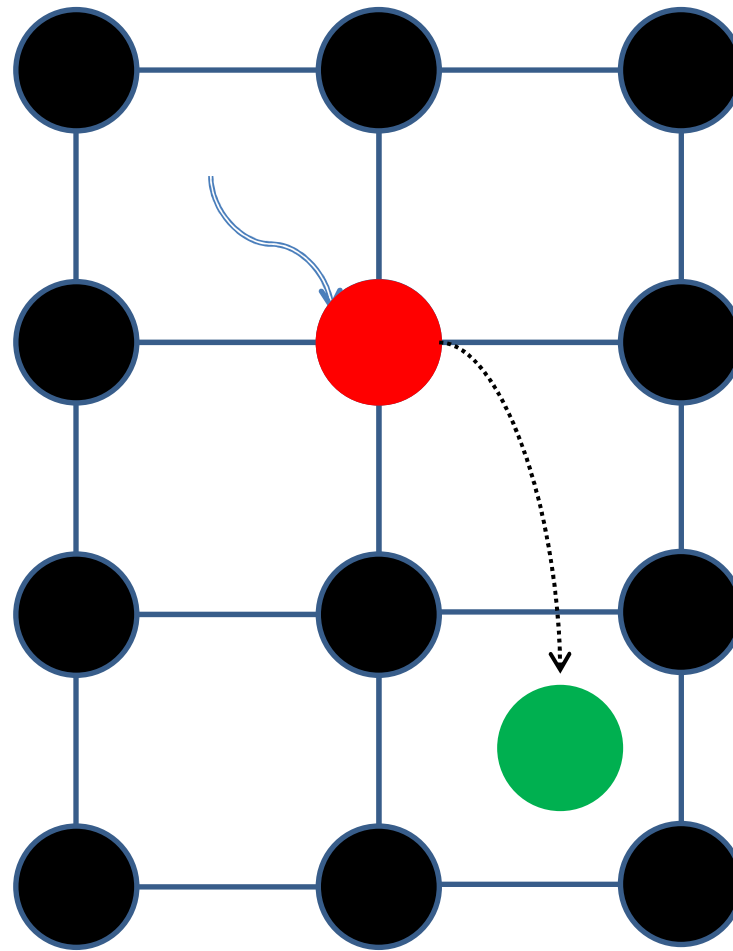
LSMS – (erős)skálázás a processzorok számával

- Fe nanorészecske FeAl kristályban
 - 16,000 Fe és Al atom elemi cellánként
 - LSMS-2 (Full potential; Multi-atom/core)

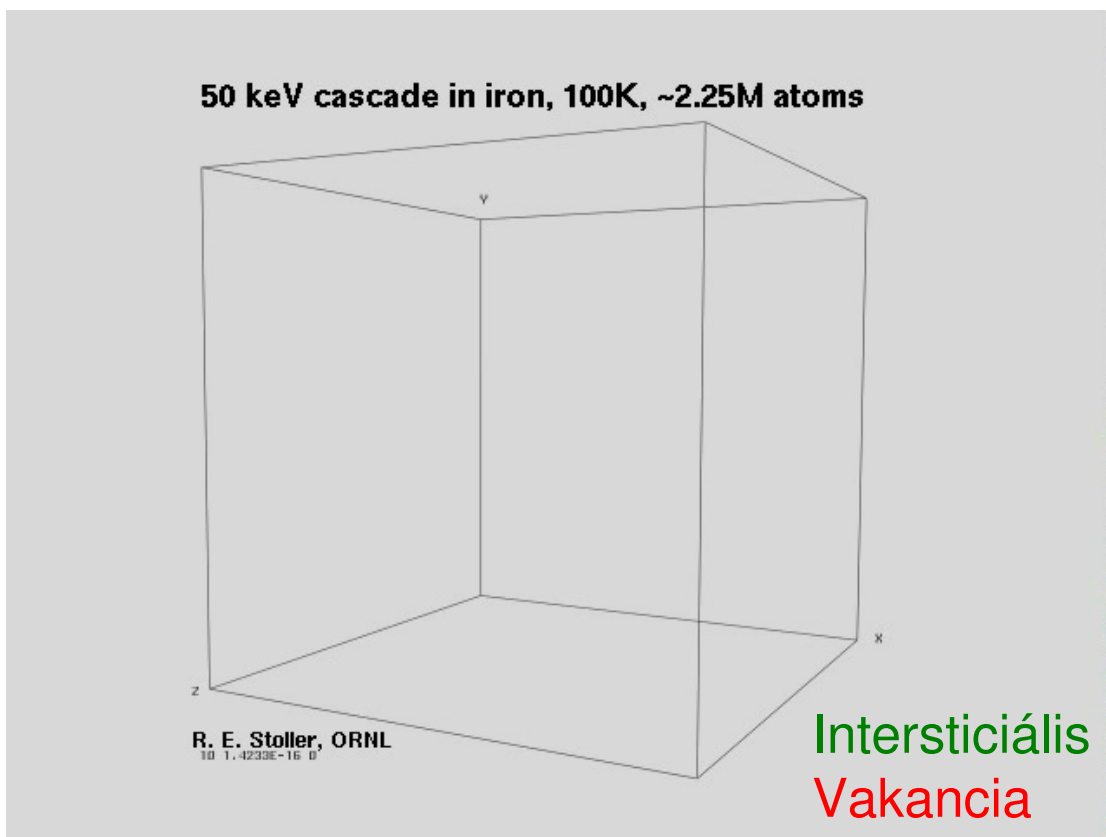


Alkalmazás: Sugárzási kaskád

Kaskád: (rács)hiba lavina



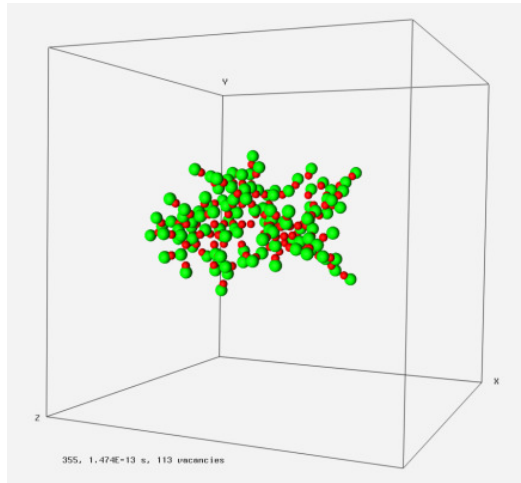
Sugárzási kaszkád szimulációja klasszikus molekuladinamikával



Fe
9826-atom
50 keV Kaszkád T = 100 K

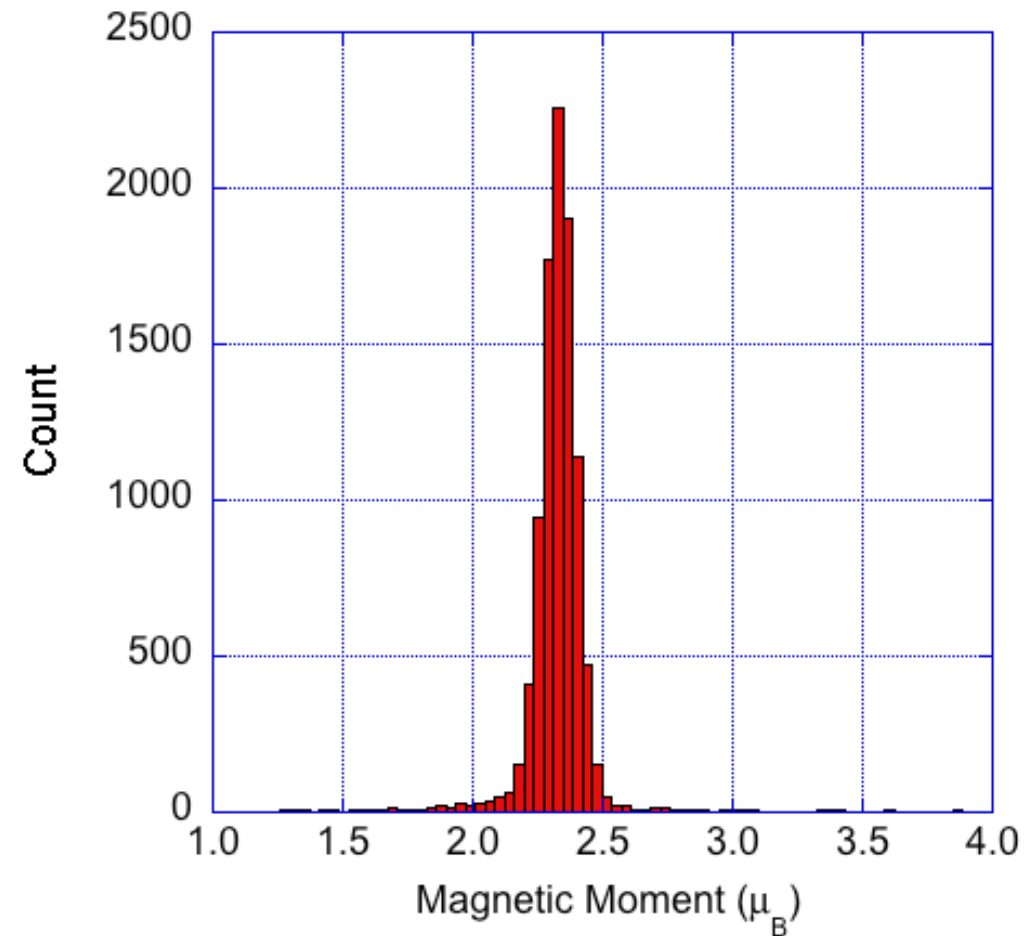
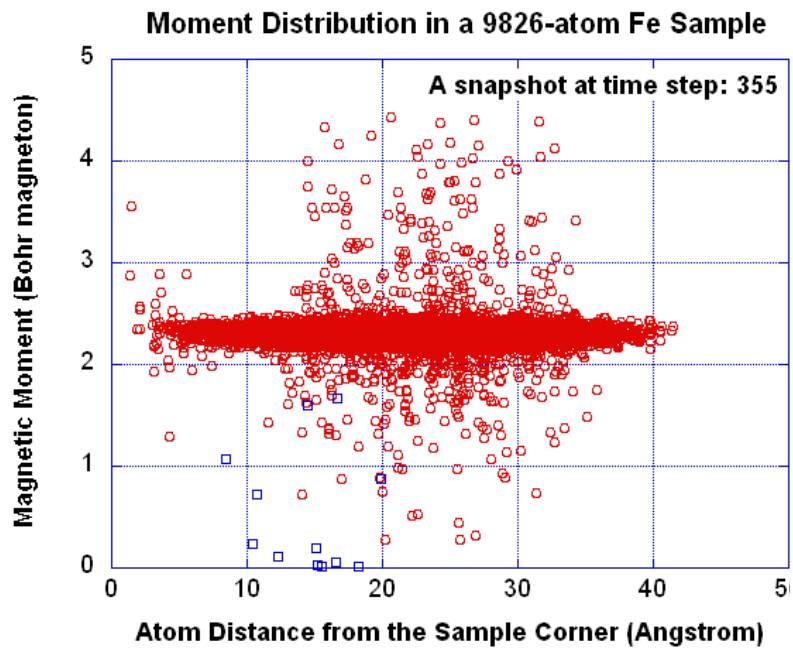
Vakanciák, Intersticiális atomok,
Termalizált

A mágneses állapot szimulációja: az LSMS munkában

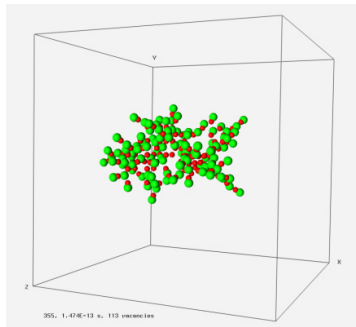


Fe: 50keV Kaszkád- 9286-atom/elemi cella

Idő: 355 [0.04 ps]



A mágneses állapot szimulációja: az LSMS munkában

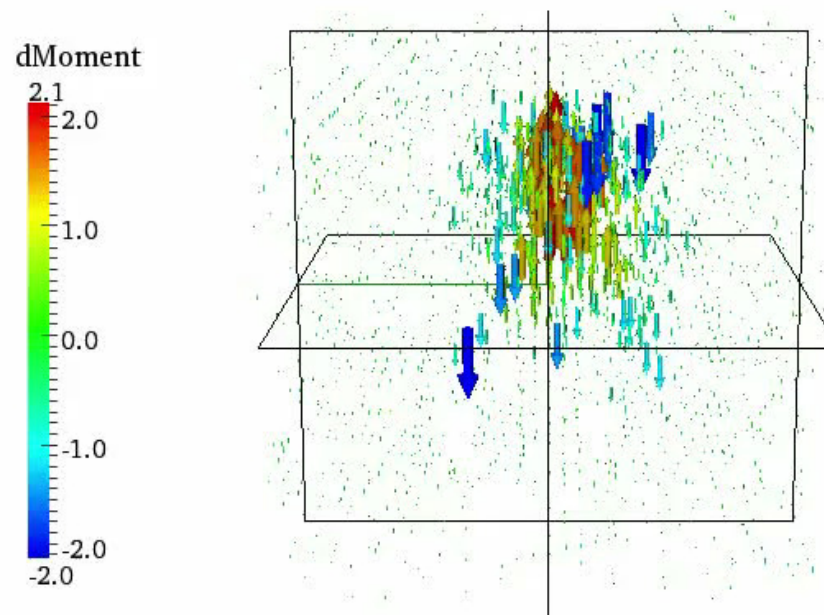


Fe: 50keV Kaszkád - 9286-atom/elemi cella:

Time Step -355 [0.04 ps]

Fe 500eV Cascade: 9826-atom unit cell
at Time Step 355 [0.04 ps]

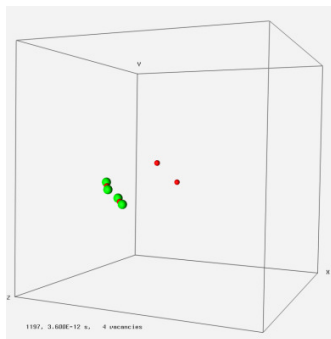
$d\text{Moment} = m - \langle m \rangle$



BCC 17x17x17 Supercell

$m - \langle m \rangle$

A mágneses állapot szimulációja: az LSMS munkában

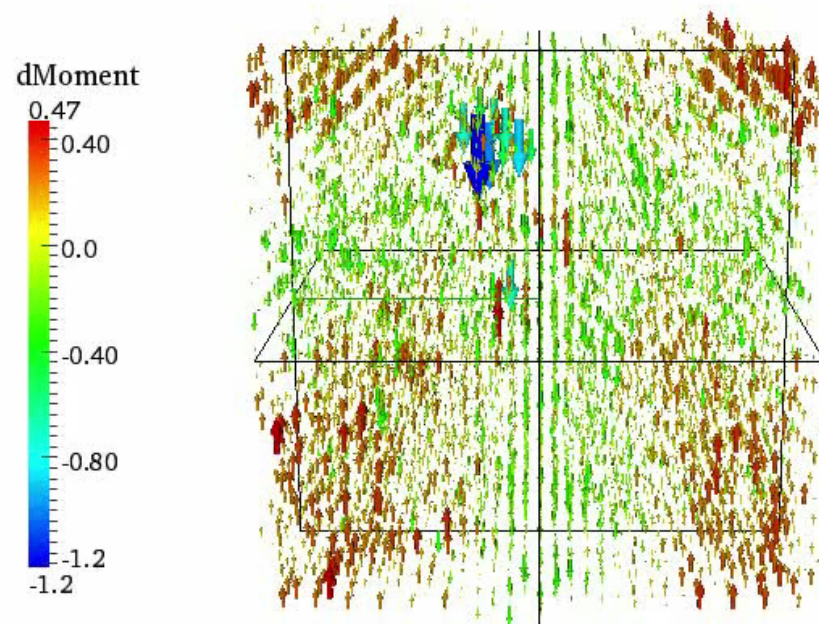


Fe: 50keV Cascade - 9286-atom/elemi cella

Time Step: 1197 – [0.21ps]

Fe 500eV Cascade: 9826-atom unit cell
at Time Step 1197 [0.21 ps]

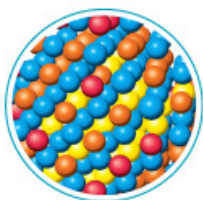
dMoment=m-<m>



BCC 17x17x17 Supercell

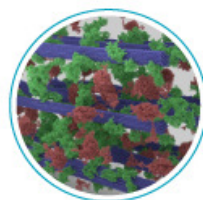
m- <m>

Óriás szimulációk



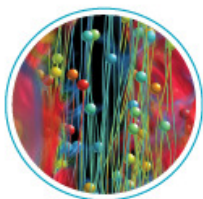
WL-LSMS

Rendezetlenség, fluktuációk
nanoszerkezetekben



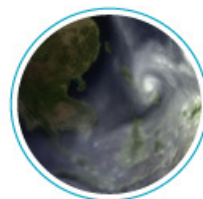
LAMMPS

Membrándiffúzió molekuláris
leírása



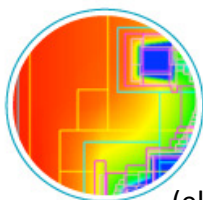
S3D

Anyagspecifikus turbulens égés



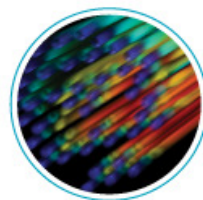
CAM-SE

Klímamodellek



NRDF

(elektromágneses) sugárzási transzport
alkalmazás az asztrofizikában, lézer fúzióban,
orvosi képalkotásban



Denovo

(részecske)sugárzási transzport energiaipari
és technológiai alkalmazások

Összefoglalás

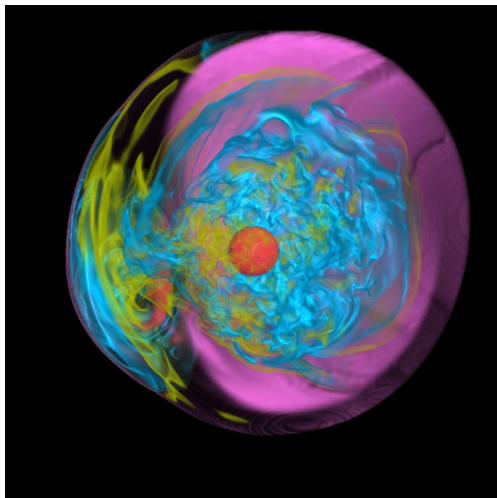
A szuperszámítógépek nélkülözhetetlen eszközei az élvonalbeli tudományos kutatásnak
Szuperszámítógépeken gyakorlati közeli „szupertudomány”

A programok egyre jobban támogatják a masszívan parallel architektúrákat

A szuperszámítógépek energiafogyasztása fokozatosan növekszik, külön erőmű kell hozzá

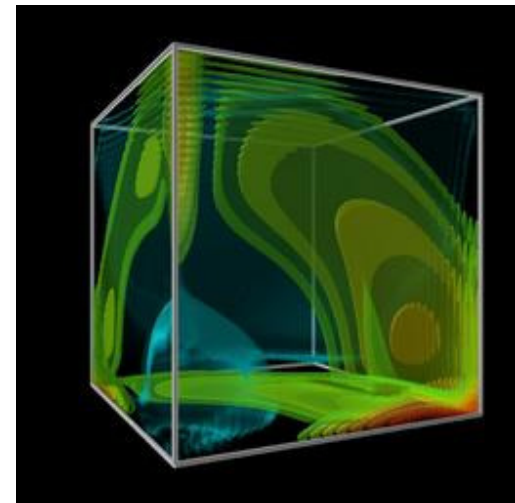
A cél a közeljövőben: 1 Exaflop/s (10^{18} lebegőpontos művelet másodpercenként)

Titán 20, 100, 200 petaflop/s, 1 exaflop/s (roadmap)



Szupernóvarobbanás szimulációja

Köszönöm
a
figyelmet



Elektron hullámfüggvénye többtest szóródás esetén