SC13 Poster @ Denver, CO, USA 17-22, Nov. 2013

Partially supported by

Joint Usage / Research Center for Interdisciplinary Large-scale Information Infrastructures

Efficient data Compression of time series of particles' positions for High-throughput Animated Visualization

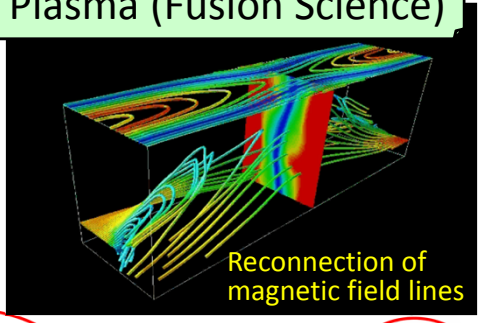
Katsumi Hagita(National Defense Academy), Takaaki Takeda (VASA Entertainment), Tsunehiko Kato (Hiroshima Univ.), Hiroaki Ohtani(National Institute for Fusion Science), Seiji Ishigro (NIFS)

Huge particle systems

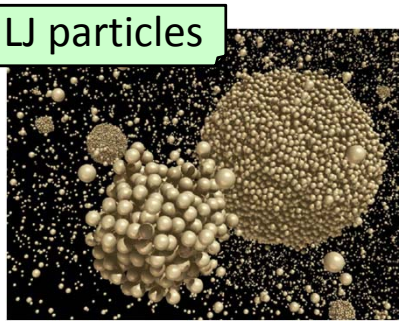
Today, over 1 billion particles!

N-body (Astro-physics)

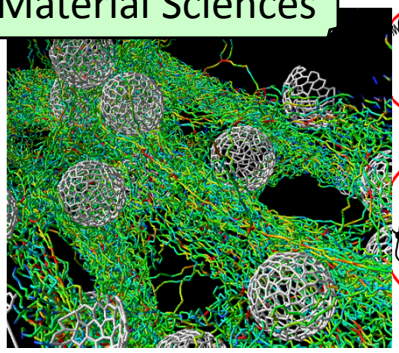
Plasma (Fusion Science)

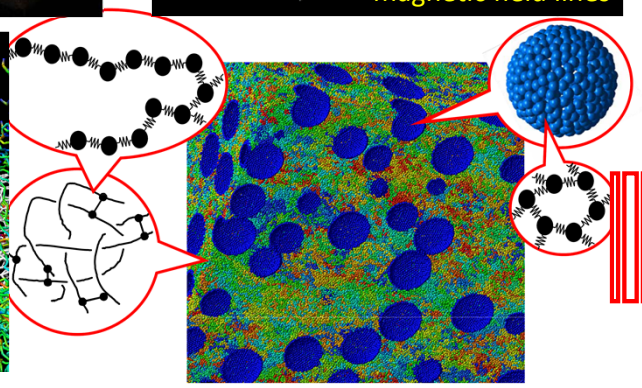


LJ particles



Material Sciences



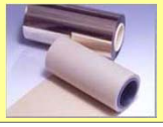


Industrial App.

Rubber & Elastomer

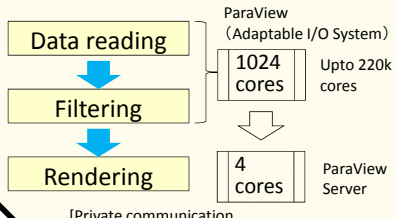
Molecular Level-CAD by simulation

High functional films



High-throughput -> High compression

High-throughput Animated Visualization of Large-scale Data



[Private communication
ParaView Tutorial Session @SC11]

Recently, Filtering and Rendering are highly parallelized.

In JAPAN, Prof. Kenji Ono performed parallel rendering using full-nodes (66kcores) of the K-computer.
[Private communication.]

Improvement of I/O

Parallel I/O

Choice 1 Distributed device

Choice 2 Single IO device

Slowness of I/O becomes main controlling factor of performance.

Speed of reading is controlled by speed of raid system and/or storage device (SSD, HDD)

Data Compression is essential factor.

Theory & Method

Consider smooth animated visualization of huge particle systems

- Short time interval is used. (Ex. in plasma physics, CFL condition)
- Compression of trajectory (sequence of positions in time order) of each particles is effective.
  1) Loss-less compression by HDF-5 with float precision.
  2) Lossy compression by XTC format.
  3) Lossy compression by our TOKI format.

TOKI (Time-Order, Kinetic and Irreversible) Compression

Lossy compression of trajectory of each particles with polynomial functions. Here, we defined allowable error under encoding to find coeff. of polynomial and length of encorded data. We used 3rd order for fast computation.

Test Data

Plasma particle simulation (explicit electro magnetic particles in cell method)

24000 frames with 150,000 electrons and 150,000 ions.

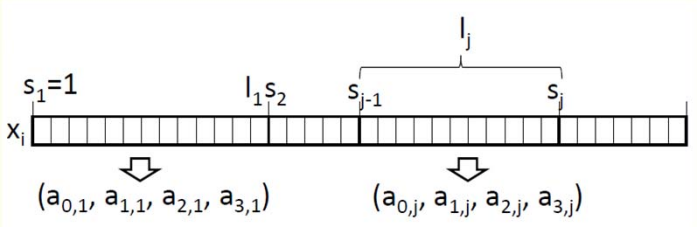
The minmum time step is $0.016/\omega_{ce}$, which is given by CFL condition.

Acknowledgement

This work is partially supported by "Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures" in Japan.
This work is partially supported by NIFS/NINS under the project of Formation of International Scientific Base and Network.
This work is partially supported by HPC project of Nagoya University.

Definition of TOKI compression

- Allowed error (input parameter):ε
- Particle position at time i : x_i
- Data division number: j



1.Find candidate of fitting parameters $a_{n,j}$ by least squares method (for speed)

$$\sum_{i=s_j}^{s_j+l_j}\{x_i - \sum_{n=0}^{3} a_{n,j}t_i^n\}^2 \rightarrow min$$

2.Check the difference

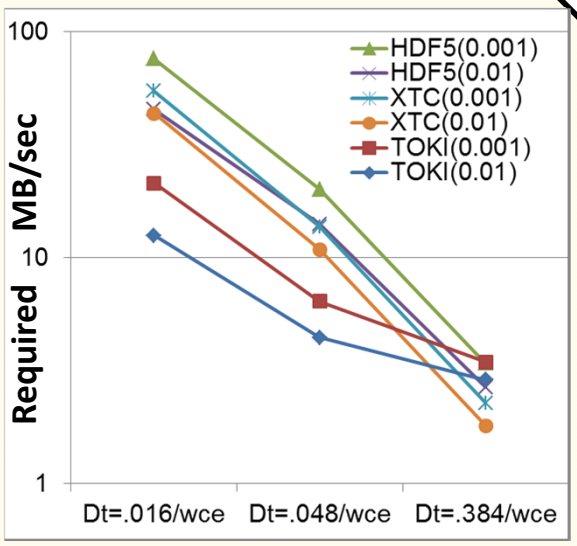
$$\max_{i=s_j}^{s_j+l_j}|x_i - \sum_{n=0}^{3} a_{n,j}t_i^n| \leq \varepsilon$$

3.Finding the longest length l_j and storing $a_{n,j}$. If not finding, raw data is stored.

File size of compressed files (Speed of reading)

Table 1. File size and required speed of compressed files

Filesize (GB) / Speed (MB/sec)	24000frames Dt=0.016/ω_{ce}	6000frames Dt=0.048/ω_{ce}	1000frames Dt=0.384/ω_{ce}
Text (double)	525.6 / 657	131.4 / 657	21.9 / 657
Binary (double)	172.9 / 216	43.2 / 216	7.20 / 216
HDF5(double)	163.2 / 204	41.4 / 207	6.92 / 208
HDF5(single)	78.4 / 98.0	19.7 / 98.5	3.30 / 98.9
HDF5(0.001)	60.7 / 75.8	16.1 / 80.5	2.73 / 82.0
XTC(0.001)	43.7 / 54.7	11.0 / 54.8	1.82 / 54.5
TOKI(0.001)	17.1 / 21.4	5.13 / 25.6	2.77 / 83.1
HDF5(0.01)	36.2 / 45.3	11.3 / 56.5	2.14 / 64.3
XTC(0.01)	34.8 / 43.7	8.69 / 43.5	1.45 / 43.5
TOKI(0.01)	10.0 / 12.5	3.55 / 17.7	2.31 / 69.2



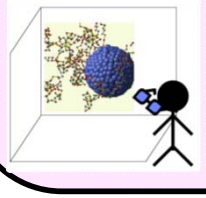
For all formats except for our TOKI format, data size is almost proportional to number of frames.
It is found that increase of data size of the TOKI format is weak for increasing of number of frames.
We considered TOKI is good for smooth animation in precision required for visualization.

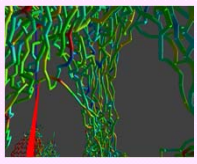
Data compression using TOKI compression leads to speedup of reading from storage.

Conclusion and Future

We implemented data compression by HDF-5, XTC, and out TOKI format in order to improve performance of data reading from storage device and/or file system. We confirmed TOKI is good for smooth animation in precision required for visualization. It is confirmed that data compression is good solution to resolve the problem of slow reading of files.

Immersive vis. by CAVE will be accelerated.





AVS/Express

FiatLux EasyVR

Recent env. for high-throughput animated vis. (Nagoya Univ., JAPAN)

SMP-supercom.
1024> cores
16TB> mem
Nvidia-Quadro

Direct Connection!

NHK super HiVision
8K display (145 inch)
7680 x 4320 /60p 10bit
Bitrate 340Mbps

We have a plan of interactive MD simulations of huge particle systems.