

Dynamic I/O Budget Reallocation For In Situ Wavelet Compression

Abstract

In situ wavelet compression is a potential solution for enabling post hoc visualization on supercomputers with slow I/O systems. While this in situ compression is typically accomplished by allocating an equal storage budget to each compute node, we propose an adaptive approach. With our approach, we introduce an assessment step prior to compression, where each node characterizes the variation in its portion of the data, and then dynamically adapts storage budgets to the nodes with the most variation. We conducted experiments comparing our adaptive approach with the traditional, non-adaptive approach, on two different simulation codes with concurrencies of 512 cores and mesh resolutions of one billion cells. Our findings show that our adaptive approach yields three orders of magnitude of improvement for one simulation and is not harmful for the other.

1. Introduction

The increasing gap between compute and I/O capabilities on supercomputers has made it difficult for simulation codes to save their state at sufficient temporal frequency for visualization. This gap motivates the use of in situ processing. That said, in situ processing can be used to enable multiple usage paradigms. In one usage paradigm, in situ techniques are used to generate the desired visualizations, resulting in images. This paradigm is effective when a user knows the visualization they want to see prior to the simulation. In another usage paradigm, in situ techniques are used to transform and reduce the data. This reduced data can be stored to disk (i.e., it can be made small enough to fit within a supercomputer's I/O constraints) and then explored post hoc. This paradigm can be effective when a user does not know the visualizations they want to see prior to the simulation. That said, the data reduction can compromise accuracy, and so it is critical that the reduction process preserve accuracy as best as possible. The work described in this paper focuses on the latter usage paradigm, and specifically on increasing accuracy.

In situ wavelet compression is an important technique for reducing the size of simulation output [GGRE13, SSEM15, BXH*17]. In a typical workflow, wavelet compression concentrates the vast majority of information into a small amount of coefficients. Further, wavelets can operate with a specific I/O budget, choosing the coefficients with the most information content.

Typically, during a large-scale, parallel simulation, each

domain is allocated the same amount of resources, including the desired I/O budget. In practice, however, some domains may contain data of little consequence, and thus, their resources may be better utilized elsewhere. This work researches the effectiveness of resource reallocation for compression, i.e., adapting the I/O budget for each compute node as the simulation saves its state.

We consider two reallocation strategies. The first reallocation strategy is specific to wavelet compression. Wavelet compression inherently prioritizes data within a domain. We use this information to calculate the global importance of a compute node's data, and reallocate its I/O budget accordingly. The second reallocation strategy incorporates Shannon Entropy, a calculation that has become common in information science and determining the information content of data. The Shannon Entropy is calculated for each compute node, and then compared globally; a compute node's I/O budget is then determined by the global entropy calculation.

In terms of findings, this research shows that dynamically reallocating the I/O budget can lead to increased storage savings and more accurate output in some cases.

2. Related Work

Related work is broken into three subsections. The first subsection reviews work that reallocates resources to achieve better efficiency. The second subsection covers work that uses entropy in scientific visualization. The third subsection briefs how wavelet compression works and reviews its use in scientific simulation and visualization.

2.1. Resource Reallocation

Understanding workflow execution and scheduling resources accordingly is an important consideration when trying to find the right resource allocation strategy. If one task is given insufficient resources, it could become a bottleneck for the entire workflow. Thus, adapting resource usage dynamically, and reallocating resources according to runtime needs will speed up the overall workflow [sem18]. Labasan et al. [LLCR17] studied the adaptation of different resource allocations for different visualization routines in a power-constrained environment. They showed that adapting power resources on a per-node need basis performed better than a uniform resource distribution strategy.

Dynamically adapting I/O resources and options according to user specifications has been shown to be a useful technique for visualization tools. Dorier et al. [DSP*13] presented Damaris/Viz, an in situ visualization framework to support I/O middleware that adapts to the specific needs of simulations by using a dynamic plugin-loading architecture. ADIOS [BLZ*14], an in situ data transformation framework, also provides a level of adaptability by providing different I/O options, thereby transparently changing how the data is processed. [REMINDER: SAM WILL LOOK INTO THAT KENNY GRUCHALLA PAPER.]

2.2. Entropy

Information theory has been used as a viable way to distinguish significance in parts of the data [WS11]. Using information theory concepts can help scientists determine which portions of the data contain important features [PAJKW]. Entropy is one such information theory technique. According to Shannon [Sha01], entropy can be calculated for a variable, x , to measure the average uncertainty in it. The higher the entropy, the more information content this variable has. Entropy is always non-negative [CT06] and describes the unit of information on average required to describe this variable [LMG*18].

Dorier et al. [DSG*16] compare a number of different approaches, including entropy, to determine the saliency of the data, and then uses these metrics for load balancing in a time-constrained environment. This strategy is particularly useful when visualization resources are limited and should be allocated to the most important data. Additionally, Wang et al. [WK08, WYM08] demonstrate how information theory, and entropy in particular, can highlight certain areas of time-varying data that are particularly important, but their work has not been evaluated for in situ capabilities.

2.3. Wavelet Compression

Wavelet compression is a class of transform-based techniques with its core operation being wavelet transforms. Such a transform decomposes data into coefficients in the wavelet domain, representing information in various time and frequency scales. In the use case of compression, wavelet kernels and transformation methods are chosen such

that the transform is 1) *nonexpansive*, meaning that the number of output coefficients equals the number of input data points; 2) *invertible*, meaning that input data could be reconstructed in its full fidelity; and 3) excellent in *information concentration*, meaning that the vast majority of information in the data is disproportionately kept in a small amount of coefficients, namely the ones with the largest magnitudes. Compression using wavelets happens when only those large-magnitude coefficients are used to reconstruct the data, and the rest coefficients are discarded.

In the scientific visualization community, wavelet compression is traditionally used to improve interactivity and accelerate renderings on large data sets [KS99, IP99, GWGS02, TBR*12]. In recent years, with the rise of big data and the growing I/O gap, wavelets are increasingly viewed and applied as a traditional compression operator. Such applications include compressing simulation checkpoint files [SSEM15], climate model output [WMB*11, BXH*17], turbulent flow simulations [LGP*15], and more generic volumetric scientific data [VED96, GGRE13, LSO*17]. Finally, wavelet compression has been proven to fit into the in situ analysis paradigm, as it is capable of achieving portable performance across modern massively parallel architectures (i.e., multi-core CPUs and GPUs) [LMC*17] and reducing overall I/O time for simulation runs with hundreds to thousands of compute nodes [LLCC17].

3. Our Method

The goal of this research is to divert I/O resources to where they are needed most. Our technique consists of two phases. In the first phase, we perform a lightweight analysis step to determine the most pertinent data among all the compute nodes. In the second phase, we perform wavelet compression, adapting the I/O budget per node according to results from the first phase.

Within the first phase of our technique, we explore two different strategies for assessing data content. Our first strategy is inherent to wavelet compression — utilizing the coefficient magnitudes that are calculated during the transform process. Our second strategy uses Shannon entropy to discern the most salient data; these results are then used to derive the budgets for each compute node's wavelet compression. We also compare these two strategies against the standard strategy, i.e., equally allocating the I/O budget to all compute nodes.

3.1. Wavelet Magnitude

The first strategy, involving the calculated wavelet magnitudes, is innate to wavelet compression. We note that in the setting of multi-domain simulations (domain meaning the data set owned by one rank), compression is applied on individual domains to generate coefficients. Magnitude difference across domains no longer necessarily warrant difference in information content; nevertheless, magnitude still remains as a strong indicator of the global importance of each

domain. During the wavelet compression process, the data values are prioritized with each pass of the filter along each axis. First, each rank will calculate its total local magnitude. Second, an MPI_AllReduce summation will calculate the total global magnitude. And then lastly, each rank will determine their respective ratio. This ratio is then used to determine that rank's I/O budget.

3.2. Entropy

The second strategy utilizes Shannon Entropy, also called Information Entropy, a widely used strategy to determine the importance of data. From a high level, Shannon Entropy calculated the number of bits required to save the given data. The more bits that are required, then the more information that is present.

Each rank will calculate its local entropy value from the input data. Then, an MPI_AllReduce summation is used to calculate the total global entropy, and from there each rank will calculate their respective ratio. As with the wavelet strategy, the calculated ratio will determine each rank's I/O budget and will save data accordingly.

3.3. Standard

The two reallocation strategies will be compared against the standard wavelet compression where each rank has the same I/O budget.

4. Experiment Overview

Our experiments studied a cross-product of options over three factors, for a total of 18 tests. The three factors and their options were:

- **Allocation Strategy.** This factor had three options (all described in Section 3): Magnitude, Entropy, and Standard.
- **I/O Budget.** This factor also had three options: data reductions of 32:1, 64:1, and 128:1.
- **Simulation Code.** This factor had two options:
 - Lulesh [lul] is a hydrodynamics simulation that models the Sedov blast problem. This simulation had size 1000^3 and ran for 10,200 cycles with 512 ranks on 32 nodes. The experiments for Lulesh were performed on Cori.
 - CloverLeaf3D [MRG*15] is a simulation that solves the compressible Euler equations. This simulation had size 1000^3 and ran for 500 cycles with 64 ranks on 32 nodes. The experiments for CloverLeaf3D were performed on Cheyenne.

This research was implemented within the Ascent framework [LAA*17] and the experiments were run on NERSC's Cori supercomputer, as well as NCAR's Cheyenne supercomputer. Ascent provides a lightweight in situ infrastructure that includes a vtk-m wavelet implementation [MSU*16].

The effects of data reduction were measured by considering the Normalized Root Mean Square Error (NRMSE)

for the compressed output compared to the original data. We also measured the overhead to do reallocation, i.e., the amount of time spent coordinating between the nodes.

5. Results

We tested our strategies on two different simulations, Lulesh and CloverLeaf3D. For each simulation we measured the average NRMSE compared to the original data and compared our results with standard compression. We also measured the overhead of our strategies.

5.1. Lulesh

Lulesh takes time to advance to an interesting state, so for the majority of the cycles only rank 0 has pertinent data. Thus, the reallocation strategies were able to devote their entire I/O budget to saving rank 0's data, whereas the standard compression would only save a fixed portion of rank 0's data.

Moreover, by dedicating the entire I/O budget to rank 0, the reallocation strategies saved out less data overall than the standard compression. This is due to the fact that rank 0's total data size is smaller than the total budget (i.e. $120^3 < \frac{1000^3}{128} < \frac{1000^3}{64} < \frac{1000^3}{32}$).

The results in Table 1 demonstrates that the disproportionate makeup of the data heavily favored our reallocation strategies, which improved on standard compression by several orders of magnitude. Figure 1 shows the visual artifacts present in the standard compression. Here, our reallocation strategies prioritized this data in order to save a more accurate representation of the original data compared to the standard compression.

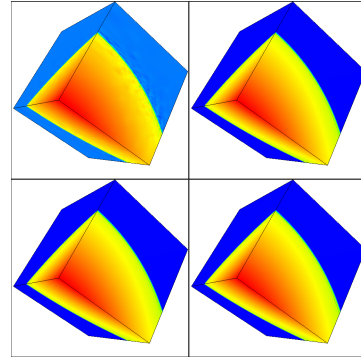


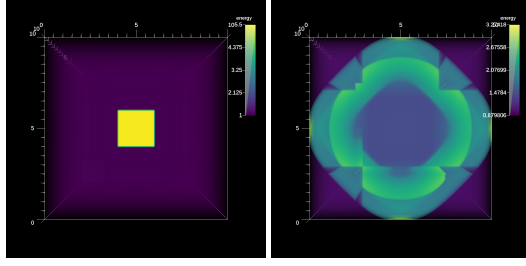
Figure 1: **Top Left:** The standard compression using 128:1 compression ratio. **Top Right:** Reallocation strategy based on magnitude. This strategy determined the data should be saved using a 7:1 compression ratio. **Bottom Left:** The original data. **Bottom Right:** Reallocation strategy based on entropy. This strategy determined the data should be saved using a 3:1 compression ratio.

5.2. CloverLeaf3D

CloverLeaf3D advances to an interesting state more quickly than Lulesh as shown in Figure 3, hence the reduced cycles for the experiment.

Compression Ratio	Standard Avg. NRMSE	Magnitude Avg. NRMSE	Entropy Avg. NRMSE
32:1	0.00055	3.73E-13	3.97E-08
64:1	0.00184	4.58E-09	9.58E-08
128:1	0.00436	9.32E-06	4.98E-07

Table 1: The average NRMSE for the reallocation strategies for each I/O budget on Lulesh.



(a) Timestep 10

(b) Timestep 63,640

Figure 2: CloverLeaf3D at (a) the beginning of the simulation and (b) at the end of the simulation.

The results for CloverLeaf3D in Figure 3 shows that all allocation strategies performed similarly as the simulation progressed, but showed information gains during early cycles. With a more dispersed makeup of important data, the reallocation strategies were similar to the fixed compression ratio in terms of distributing the I/O budget.

Similar to Lulesh, the reallocation strategies attained increased I/O savings during the early cycles when the data was more condensed and only several ranks were saving out their entire slice of data. However, those savings decreased as the simulation progressed, eventually both reallocation strategies were using the full extent of their I/O budget.

Simulation	Magnitude MPI	Entropy MPI	Entropy Calc
Lulesh	0.2378385	0.1733445	0.0063935
CloverLeaf3D	0.023089	0.0188065	8.00E-06

Table 2: The average MPI timing for each strategy and the average time to calculate entropy.

5.3. Timings

With any in situ analysis it is important to know if the proposed algorithm creates a detrimental overhead to the simulation as a whole. We measured the time it takes for each rank to calculate entropy as well as the MPI coordination time. For both simulations we found the MPI coordination and entropy calculation time to be acceptable for in situ analysis as shown in Table 2.

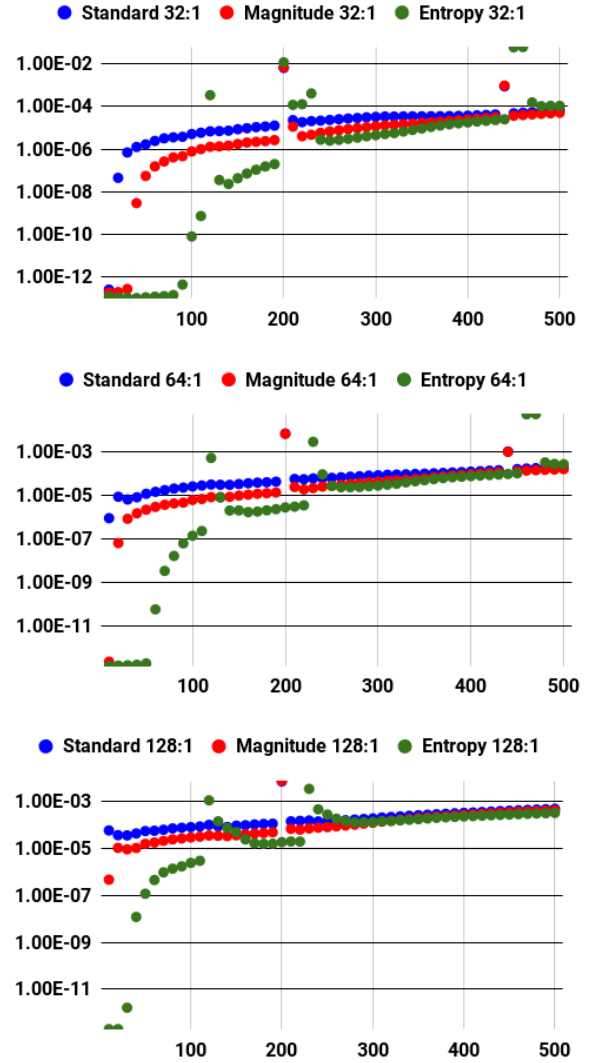


Figure 3: The average RMSE over time for each strategy on CloverLeaf3D with each compression ratio.

6. Conclusion

This work introduces an approach for automatically adapting I/O resources based on data complexity. The results show benefit for a simulation where data complexity is imbalanced and also for a simulation in early phases of simulation. Further, among our two strategies, Shannon entropy better captured information content than an inherently wavelet based approach. In terms of future work, we would like to integrate this into in situ frameworks, and also evaluate on more simulation codes.

References

[BLZ*14] BOYUKA D. A., LAKSHMINARASIMHAM S., ZOU X., GONG Z., JENKINS J., SCHENDEL E. R., PODHORSZKI N.,

- LIU Q., KLASKY S., SAMATOVA N. F.: Transparent in situ data transformations in adios. In *2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing* (May 2014), pp. 256–266. [2](#)
- [BXH*17] BAKER A., XU H., HAMMERLING D., LI S., CLYNE J.: Toward a Multi-method Approach: Lossy Data Compression for Climate Simulation Data. In *Proceedings of ISC workshops on Data Reduction for Big Scientific Data (DRBSD-1)* (Frankfurt, Germany, Jun 2017). [1](#), [2](#)
- [CT06] COVER T., THOMAS J. A.: Elements of information theory. [2](#)
- [DSG*16] DORIER M., SISNEROS R., GOMEZ L. B., PETERKA T., ORF L., RAHMANI L., ANTONIU G., BOUGÂL L.: Adaptive performance-constrained in situ visualization of atmospheric simulations. In *2016 IEEE International Conference on Cluster Computing (CLUSTER)* (Sep. 2016), pp. 269–278. [2](#)
- [DSP*13] DORIER M., SISNEROS R., PETERKA T., ANTONIU G., SEMERARO D.: Damaris/viz: A nonintrusive, adaptable and user-friendly in situ visualization framework. pp. 67–75. [2](#)
- [GGRE13] GRALKA P., GROTTTEL S., REINA G., ERTL T.: Application-specific compression of large MD data preserving physical characteristics. In *IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)* (2013), IEEE, pp. 85–93. [1](#), [2](#)
- [GWGS02] GUTHE S., WAND M., GONSER J., STRASSER W.: Interactive rendering of large volume data sets. In *Proceedings of IEEE Visualization (VIS'02)* (2002), IEEE, pp. 53–60. [2](#)
- [IP99] IHM I., PARK S.: Wavelet-based 3D compression scheme for interactive visualization of very large volume data. In *Computer Graphics Forum* (1999), vol. 18, Wiley Online Library, pp. 3–15. [2](#)
- [KS99] KIM T.-Y., SHIN Y. G.: An efficient wavelet-based compression method for volume rendering. In *Proceedings of the Seventh Pacific Conference on Computer Graphics and Applications* (1999), IEEE, pp. 147–156. [2](#)
- [LAA*17] LARSEN M., AHRENS J., AYACHIT U., BRUGGER E., CHILDS H., GEVECI B., HARRISON C.: The alpine in situ infrastructure: Ascending from the ashes of strawman. In *Proceedings of the In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization* (New York, NY, USA, 2017), ISAV'17, ACM, pp. 42–46. [3](#)
- [LGP*15] LI S., GRUCHALLA K., POTTER K., CLYNE J., CHILDS H.: Evaluating the Efficacy of Wavelet Configurations on Turbulent-Flow Data. In *Proceedings of IEEE Symposium on Large Data Analysis and Visualization* (Chicago, IL, Oct 2015), pp. 81–89. [2](#)
- [LLCC17] LI S., LARSEN M., CLYNE J., CHILDS H.: Performance impacts of in situ wavelet compression on scientific simulations. In *Proceedings of the In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization Workshop* (New York, NY, USA, 2017), ISAV2017, ACM. [2](#)
- [LLCR17] LABASAN S., LARSEN M., CHILDS H., ROUNTREE B.: PaViz: A Power-Adaptive Framework for Optimizing Visualization Performance. In *Proceedings of EuroGraphics Symposium on Parallel Graphics and Visualization (EGPGV)* (Barcelona, Spain, 06 2017), pp. 1–10. [2](#)
- [LMC*17] LI S., MARSAGLIA N., CHEN V., SEWELL C., CLYNE J., CHILDS H.: Achieving Portable Performance For Wavelet Compression Using Data Parallel Primitives. In *Eurographics Symposium on Parallel Graphics and Visualization* (2017), Telea A., Bennett J., (Eds.), The Eurographics Association. [2](#)
- [LMG*18] LI S., MARSAGLIA N., GARTH C., WOODRING J., CLYNE J., CHILDS H.: Data reduction techniques for simulation, visualization and data analysis. *Computer Graphics Forum* 37, 6 (2018), 422–447. [2](#)
- [LSO*17] LI S., SANE S., ORF L., MININNI P., CLYNE J., CHILDS H.: Spatiotemporal wavelet compression for visualization of scientific simulation data. In *2017 IEEE International Conference on Cluster Computing (CLUSTER)* (Sept 2017), pp. 216–227. [2](#)
- [lul] *Hydrodynamics Challenge Problem, Lawrence Livermore National Laboratory*. Tech. Rep. LLNL-TR-490254. [3](#)
- [MRG*15] MUDALIGE G. R., REGULY I. Z., GILES M. B., GAUDIN W., HERDMAN J. A., MALLINSON A.: High-level abstractions for performance, portability and continuity of scientific software on future computing systems- cloverleaf 3d, 2015. [3](#)
- [MSU*16] MORELAND K., SEWELL C., USHER W., LO L., MEREDITH J., PUGMIRE D., KRESS J., SCHROOTS H., MA K., CHILDS H., LARSEN M., CHEN C., MAYNARD R., GEVECI B.: Vtk-m: Accelerating the visualization toolkit for massively threaded architectures. *IEEE Computer Graphics and Applications* 36, 3 (May 2016), 48–58. [3](#)
- [PAJKW] PURCHASE H., ANDRIENKO N., JANKUN-KELLY T., WARD M.: Theoretical foundations of information visualization: In information visualization: Human-centered issues and perspectives. *Springer-Verlag*. [2](#)
- [sem18] In situ visualization for computational science. Bennett J., Childs H., Garth C., Hentschel B., (Eds.), vol. 8 of *Report from Dagstuhl Seminar 18271*, Dagstuhl Publishing. [2](#)
- [Sha01] SHANNON C. E.: A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.* 5, 1 (01 2001), 3–55. [2](#)
- [SSEM15] SASAKI N., SATO K., ENDO T., MATSUOKA S.: Exploration of lossy compression for application-level checkpoint/restart. In *Parallel and Distributed Processing Symposium (IPDPS)* (2015), IEEE, pp. 914–922. [1](#), [2](#)
- [TBR*12] TREIB M., BURGER K., REICHL F., MENEVEAU C., SZALAY A., WESTERMANN R.: Turbulence visualization at the terascale on desktop pcs. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2169–2177. [2](#)
- [VED96] VILLASENOR J. D., ERGAS R., DONOHO P.: Seismic data compression using high-dimensional wavelet transforms. In *Proceedings of Data Compression Conference (DCC'96)* (1996), IEEE, pp. 396–405. [2](#)
- [WK08] WANG C., K.L. M.: A statistical approach to volume data quality assessment. *IEEE Trans Vis Comput Graph* 14 (2008), 590 – 602. [2](#)
- [WMB*11] WOODRING J., MNISZEWSKI S., BRISLAWN C., DEMARLE D., AHRENS J.: Revisiting wavelet compression for large-scale climate data using JPEG 2000 and ensuring data precision. In *IEEE Symposium on Large Data Analysis and Visualization (LDAV)* (Oct 2011), pp. 31–38. [2](#)
- [WS11] WANG C., SHEN H.-W.: Information theory in scientific visualization. *Entropy* 13, 1 (2011), 254–273. [2](#)
- [WYM08] WANG C., YU H., MA K.: Importance-driven time-varying data visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (Nov 2008), 1547–1554. [2](#)