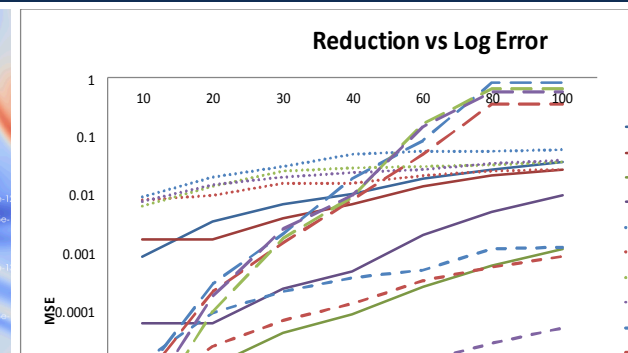
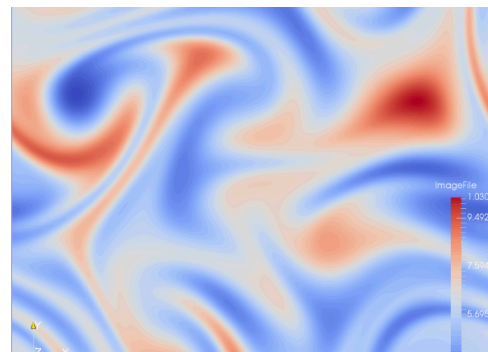
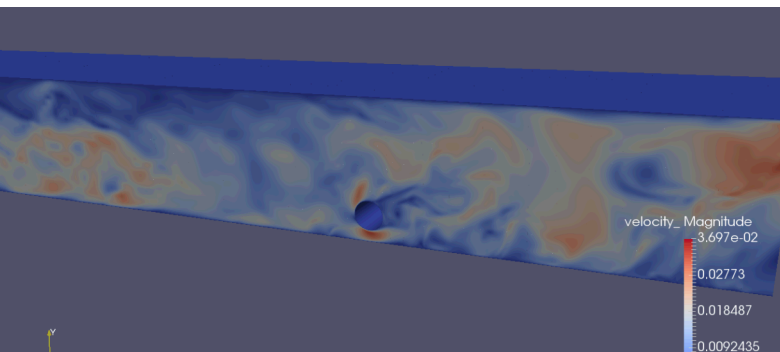


Exceptional service in the national interest



Performance Characteristics of Wavelet Compressive Sensing

Nathan Fabian, Maher Salloum, David Hensinger,
Jina Lee



Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

Introduction (motivation)

Problem: Post-Moore and exascale machines will generate datasets that are too large to be stored and transferred to remote sites, thus **impeding exploratory research** and visualization.

Ongoing research: **in-situ data compression** currently focuses on datasets defined on logically regular grids.

Hypothesis: scientific data has low information density

Our contribution: We are developing a compression/reconstruction method based on **compressive sensing (CS)** to address fields/PDE solutions defined on a cloud of points ("**point-clouds**") e.g. unstructured mesh

- *Fast, in-situ*
- *Non-intrusive*
- *Large compression ratios ($R > 50$)*

Standard Wavelet Compression

- Find wavelet transform, w , of field data, x .

$$w = \Psi x$$

- Ψ is Alpert's second-generation wavelet matrix
- Reduce w to w' by zeroing out low magnitude values
- Then recover x' from reduced w'

$$x' = \Psi^* w'$$

Wavelet Compressed Sensing

- It's similar, but instead we sample x

$$s = \Phi x$$

- s is smaller than x by our reduction factor
- Recover w_s using StOMP,
Stage-wise Orthogonal Matching Pursuit

$$s = \Theta w_s, \Theta = \Phi \Psi^*$$

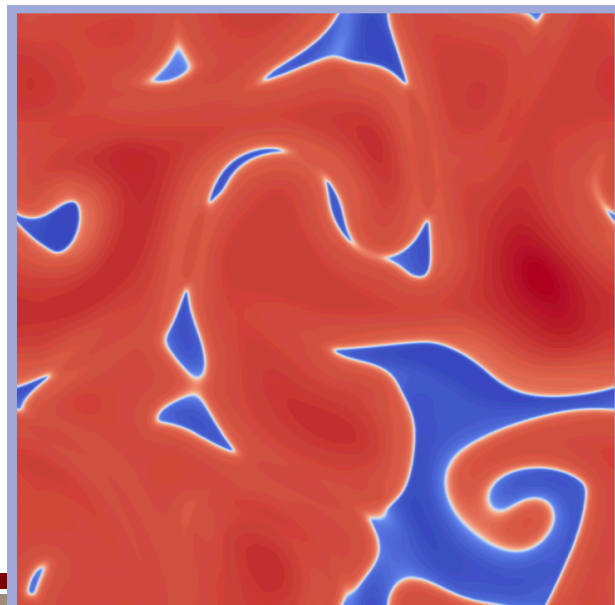
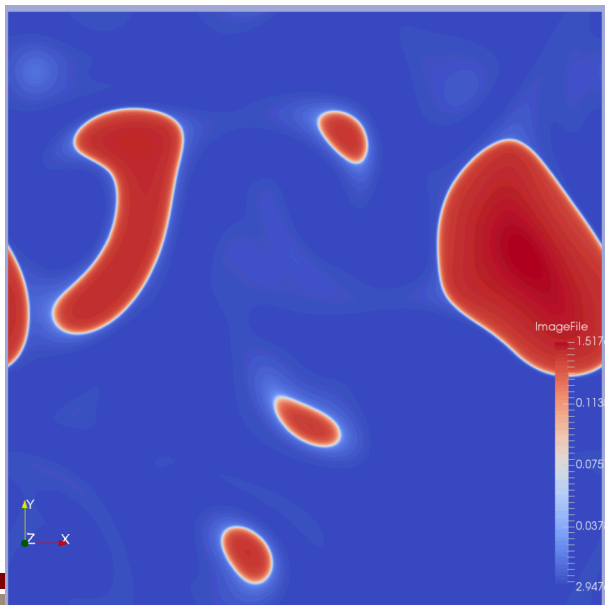
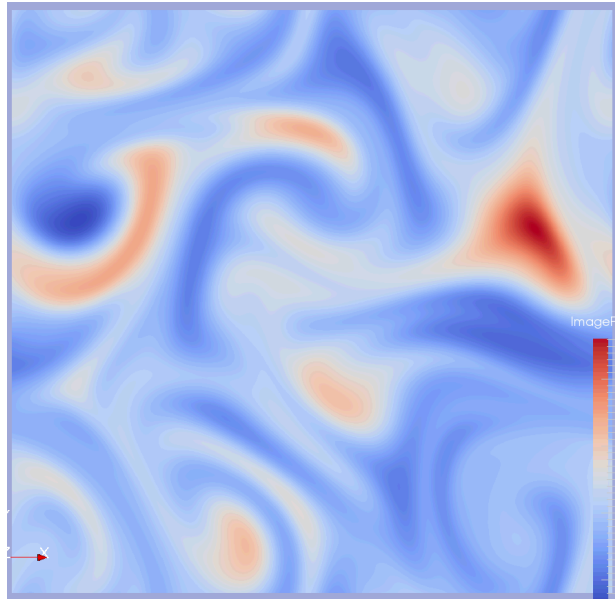
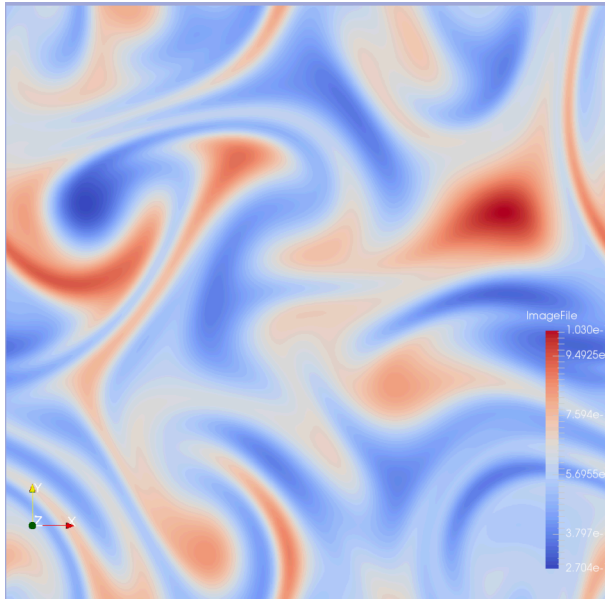
- Then recover x' from recovered w_s

$$x' = \Psi^* w_s$$

Comparison Algorithms

- ZFP
 - Divides a 3d array into 4x4x4 blocks
 - Convert to fixed point
 - Block transform similar to JPEG, but 3d and better
 - Encoded bitstream, ordered by impact on error, truncated to reduction factor
- Vector Quantization
 - Divides a 2d array into 8x8 blocks
 - K-means cluster on each block as 64-value vector
 - Number of centroids approximately reduction factor
 - Replace each block with its centroid index.

Structured Data



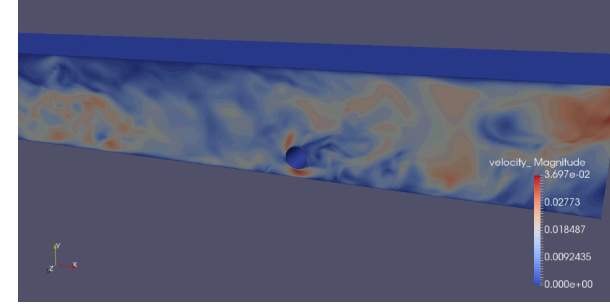
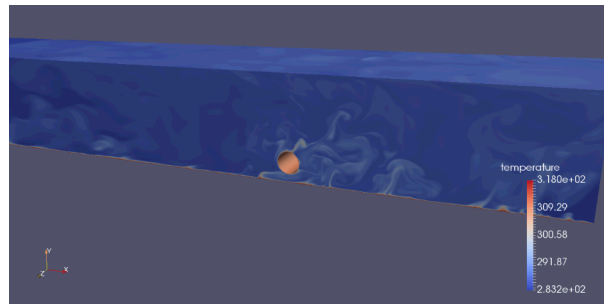
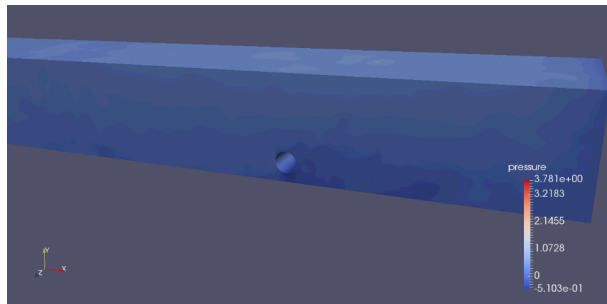
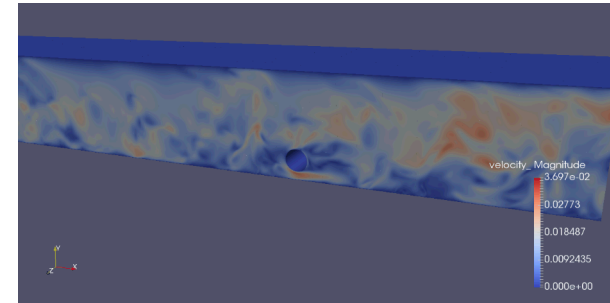
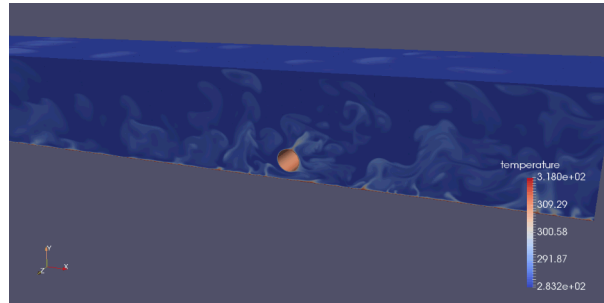
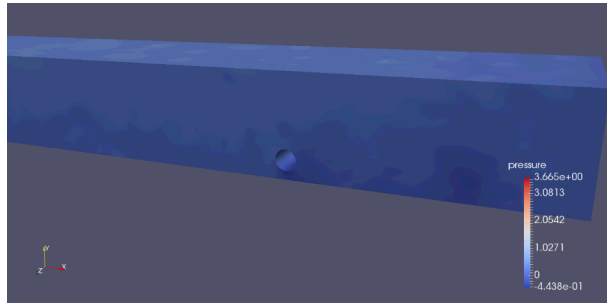
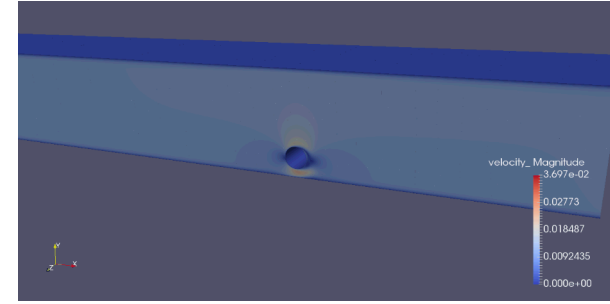
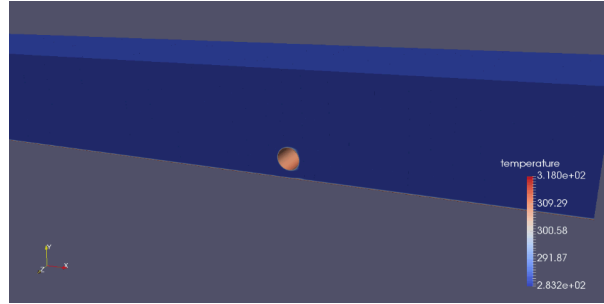
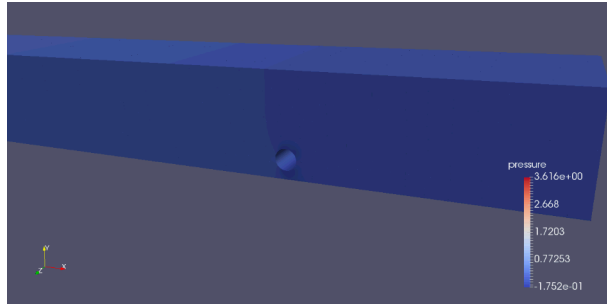
Unstructured Data

Time

Pressure

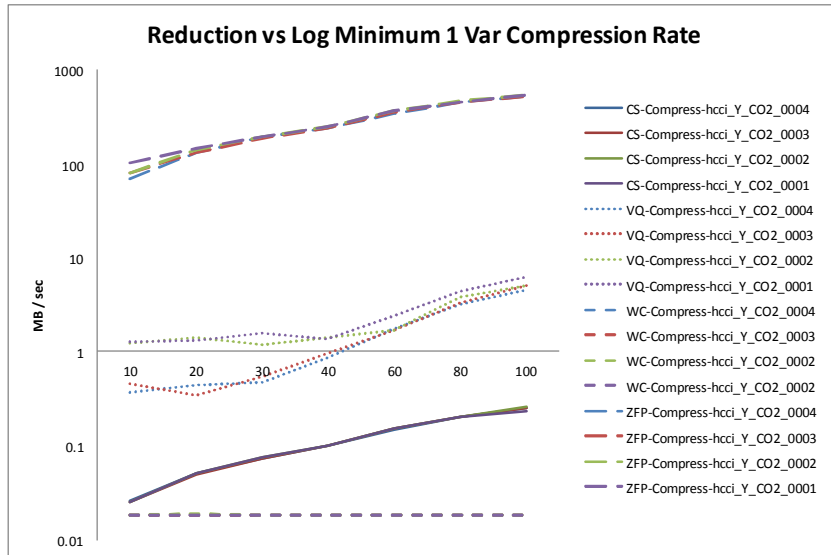
Temperature

Velocity magnitude

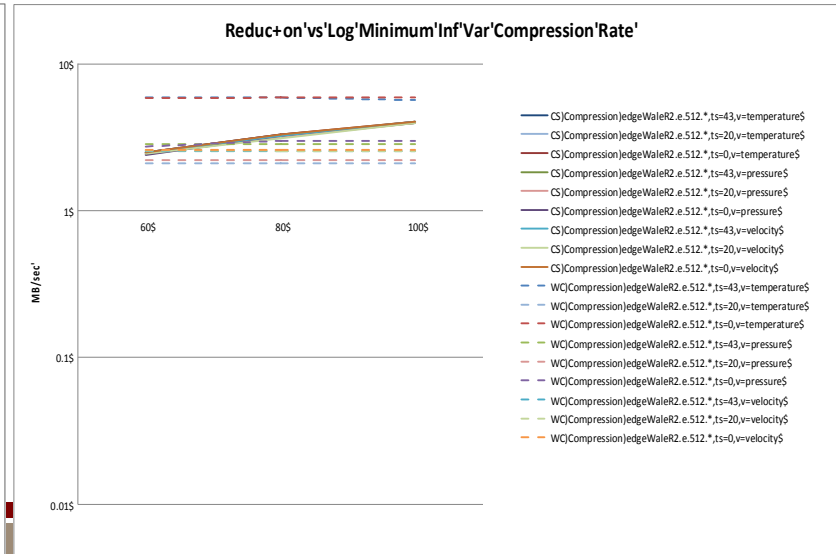
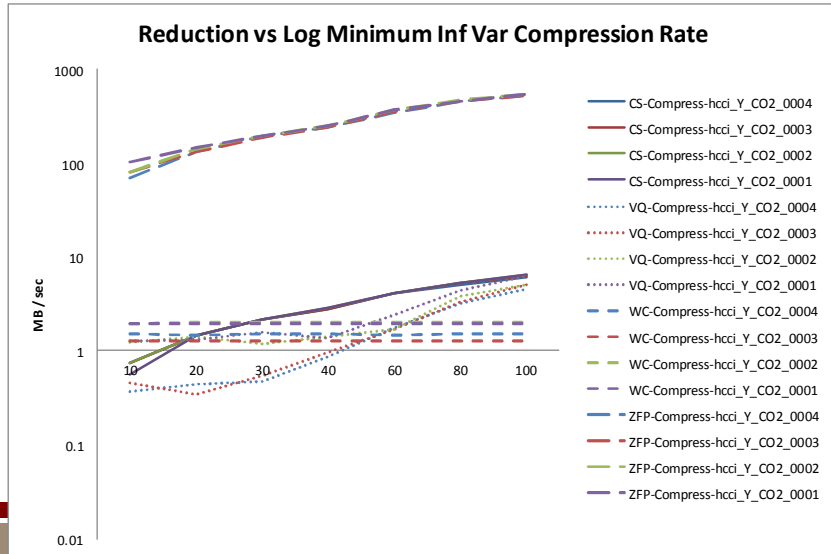
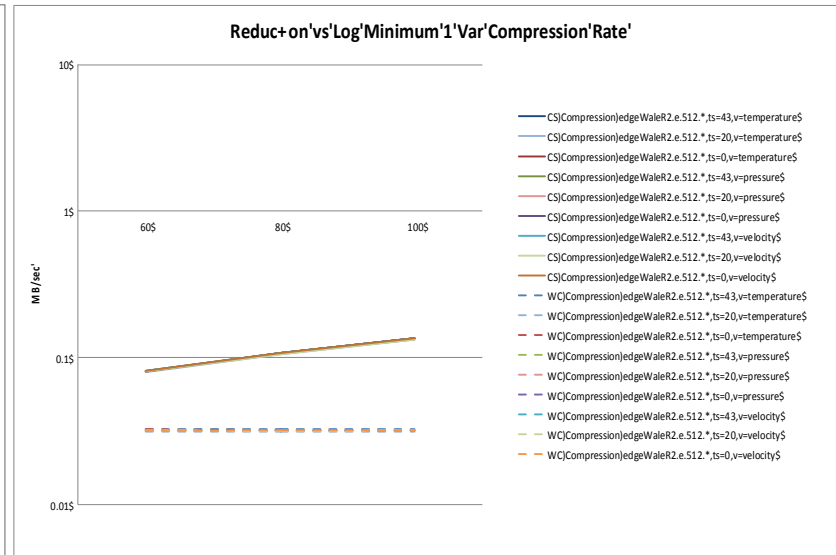


Compression Timing Results (in situ)

Structured Data



Unstructured Data

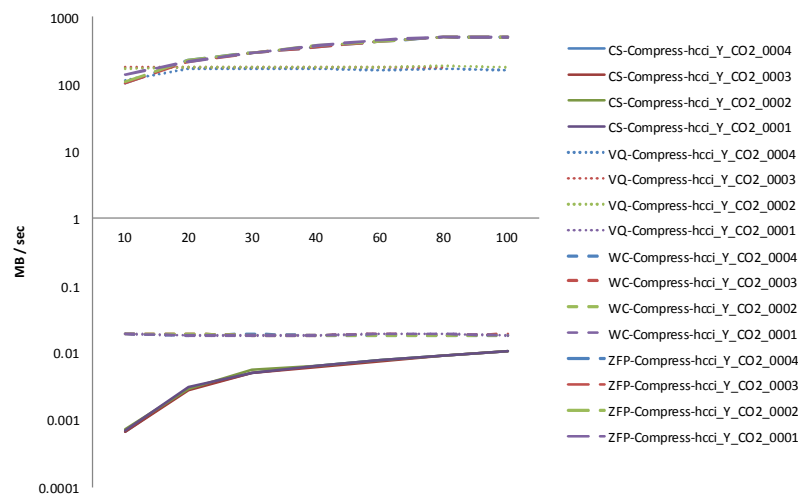


Decompression Timing Results (post)

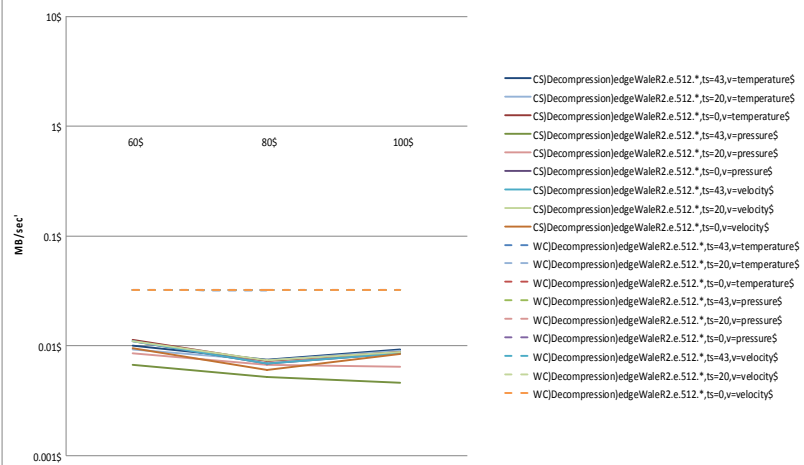
Structured Data

Unstructured Data

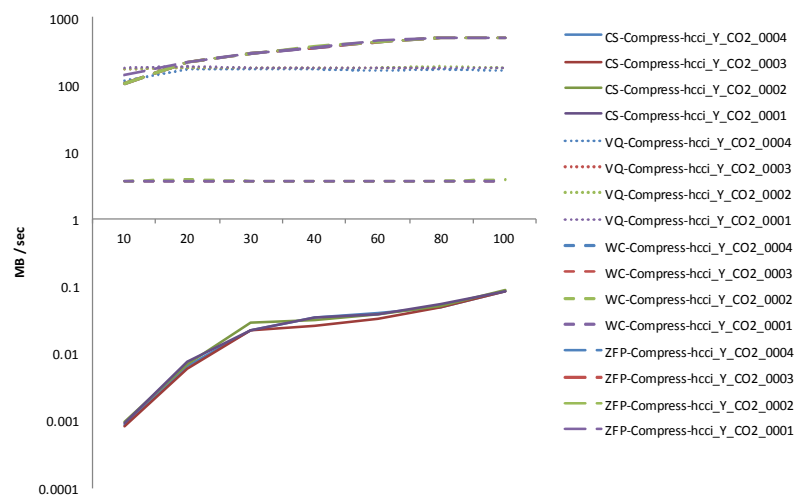
Reduction vs Log Minimum 1 Var Decompression Rate



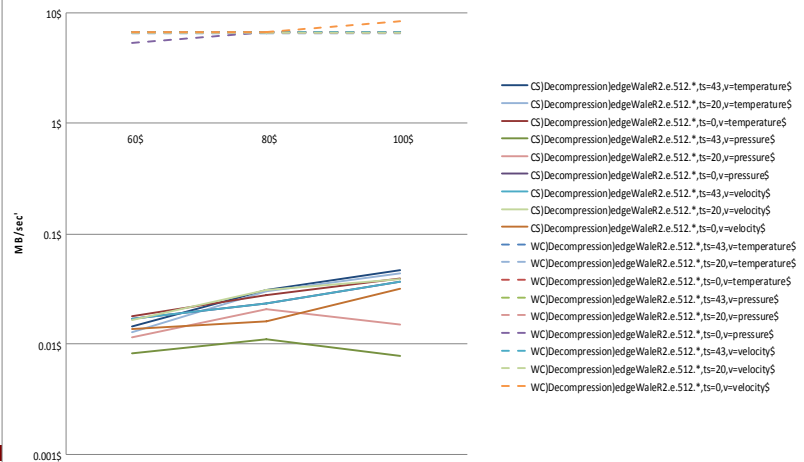
Reduction vs Log Minimum 1 Var Decompression Rate



Reduction vs Log Minimum Inf Var Decompression Rate



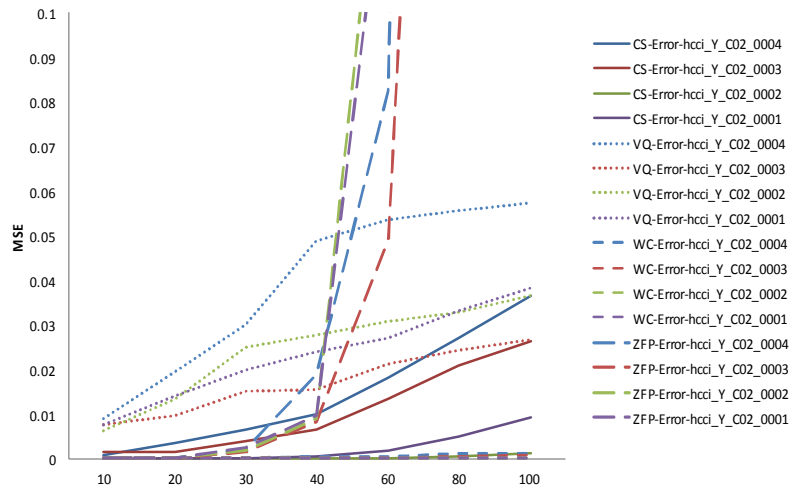
Reduction vs Log Minimum Inf Var Decompression Rate



Error Results

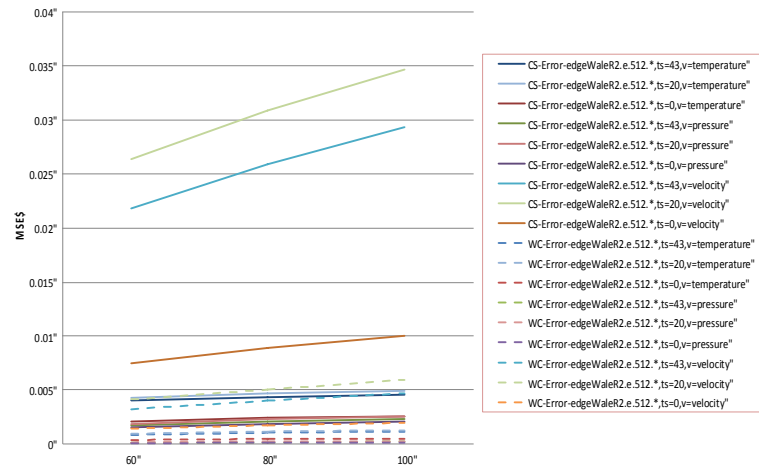
Structured Data

Reduction vs Error

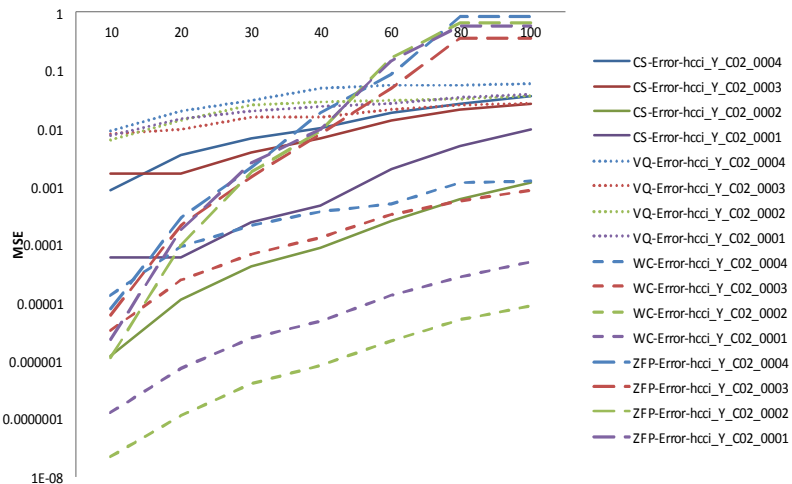


Unstructured Data

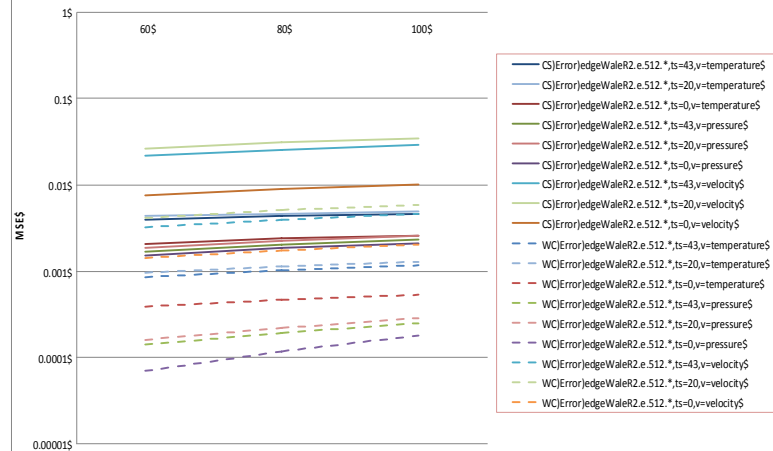
Reduction vs Error



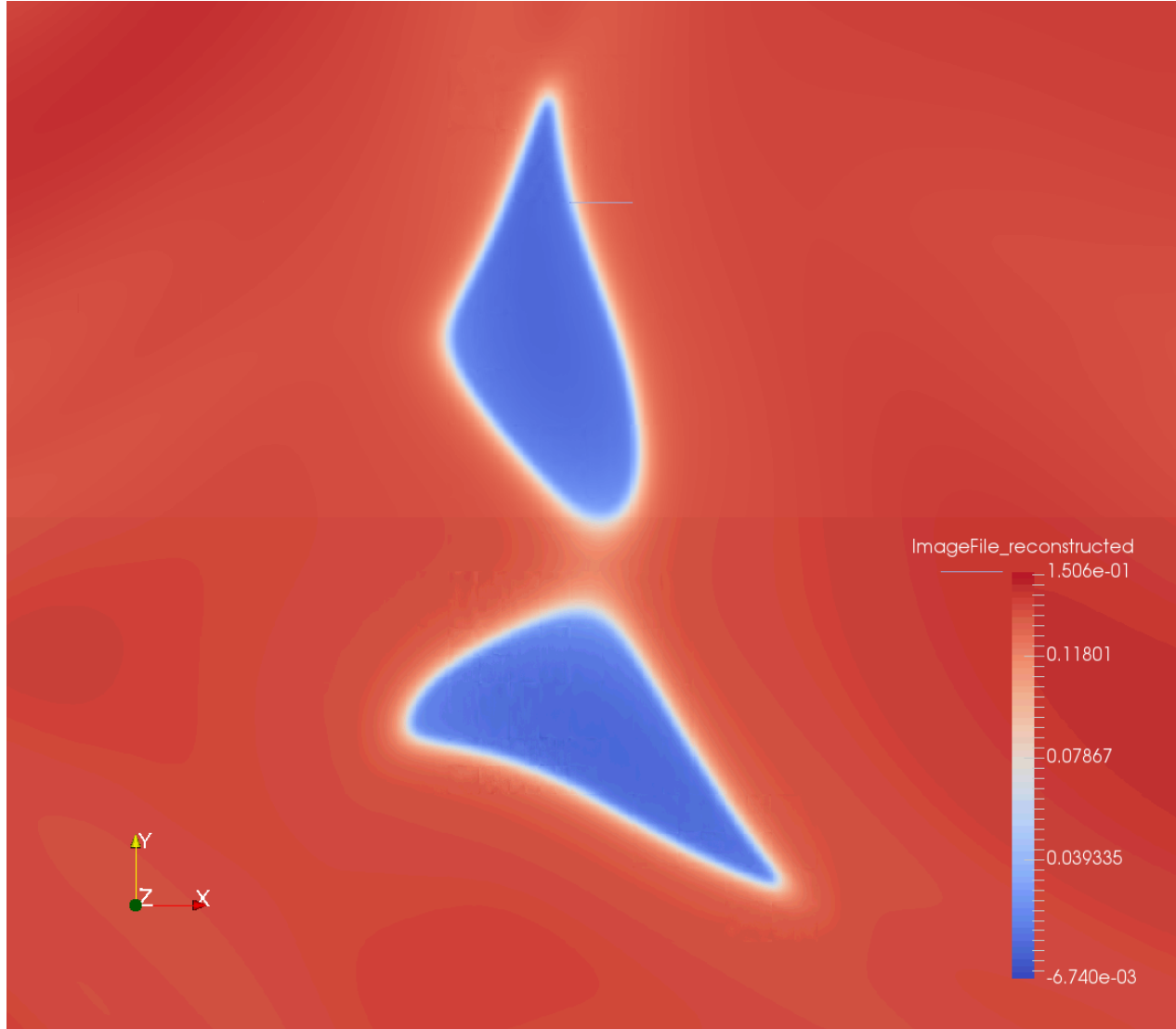
Reduction vs Log Error



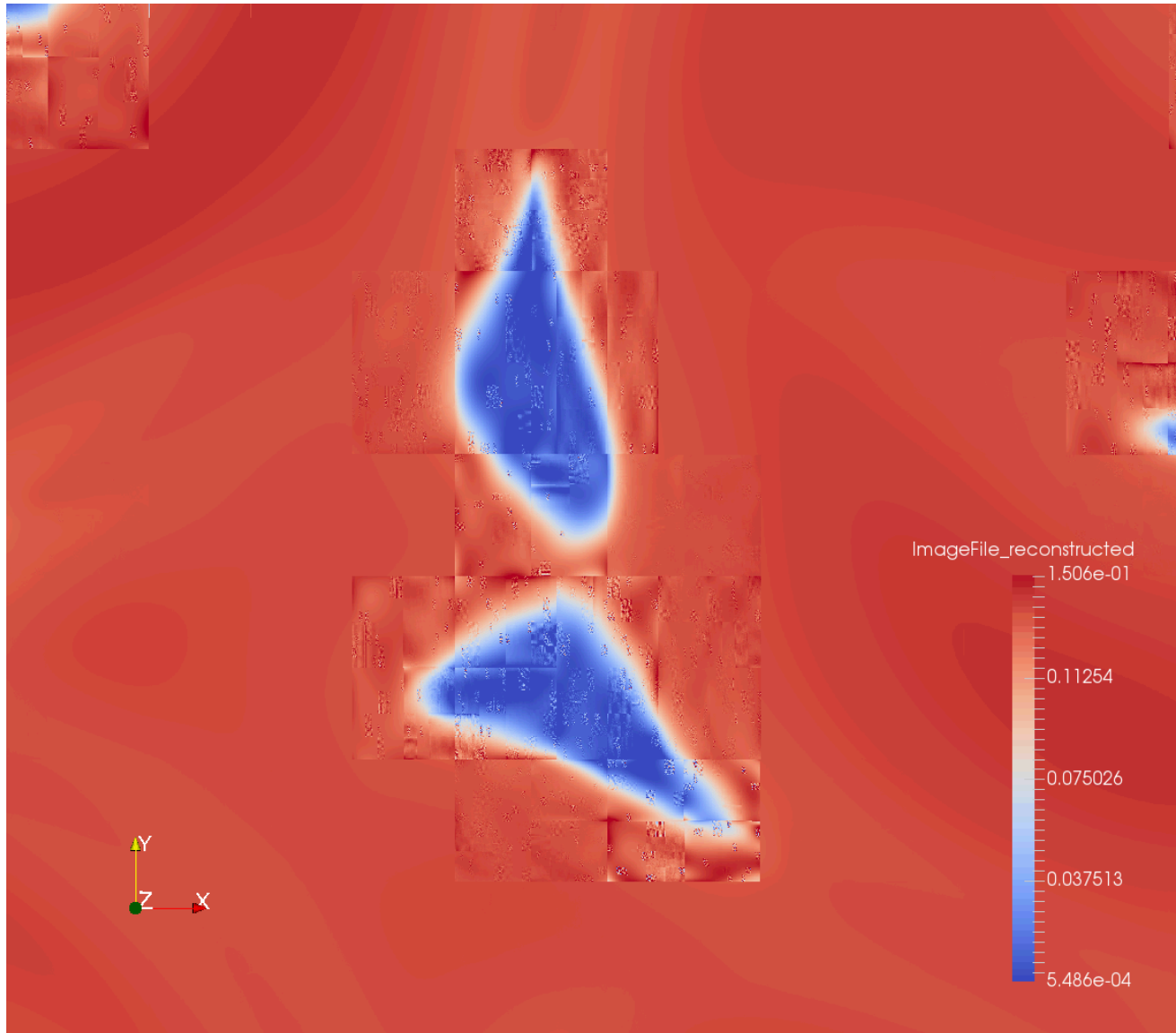
Reduction vs Log Error



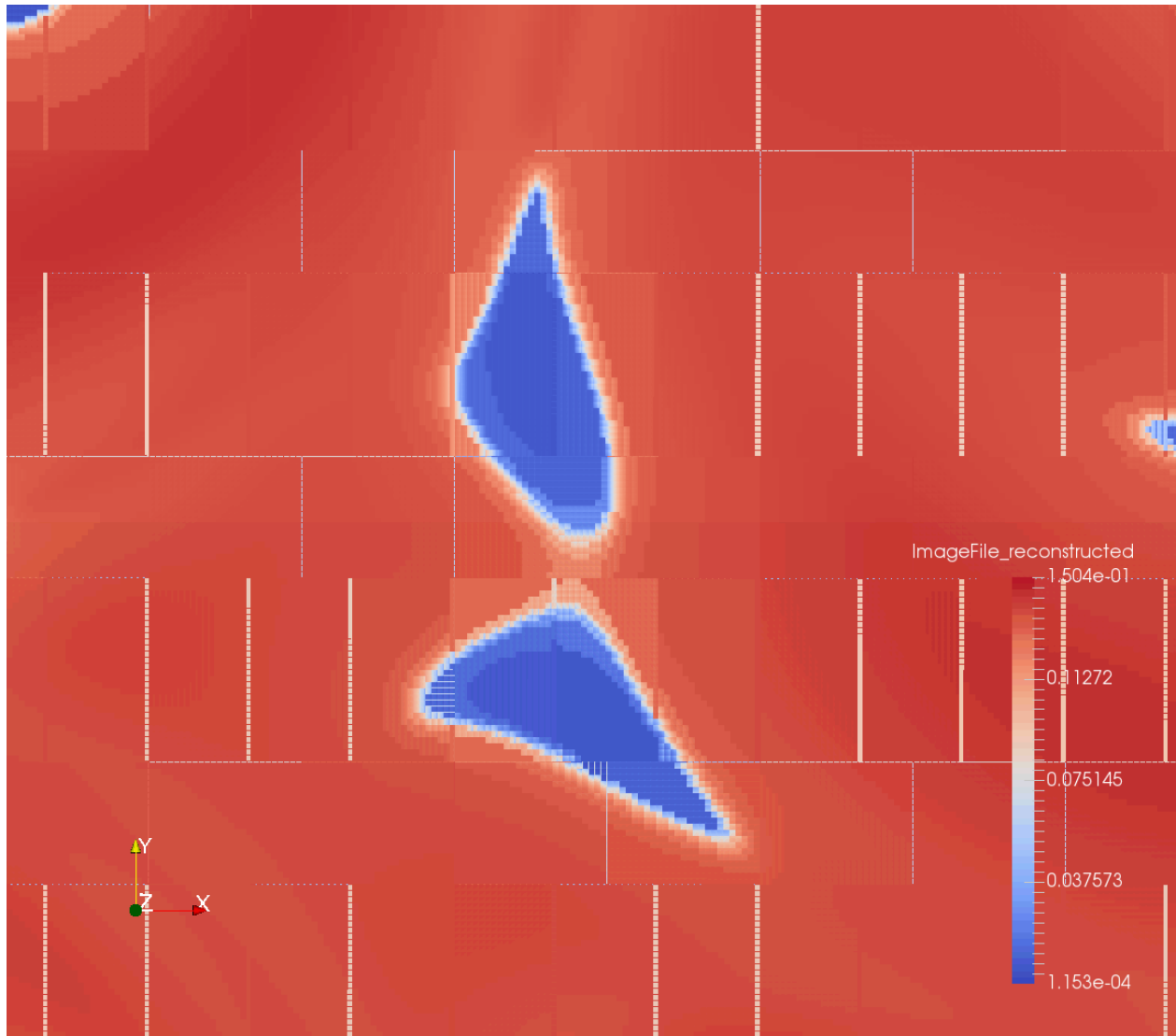
Visual Errors – Wavelet – 100x



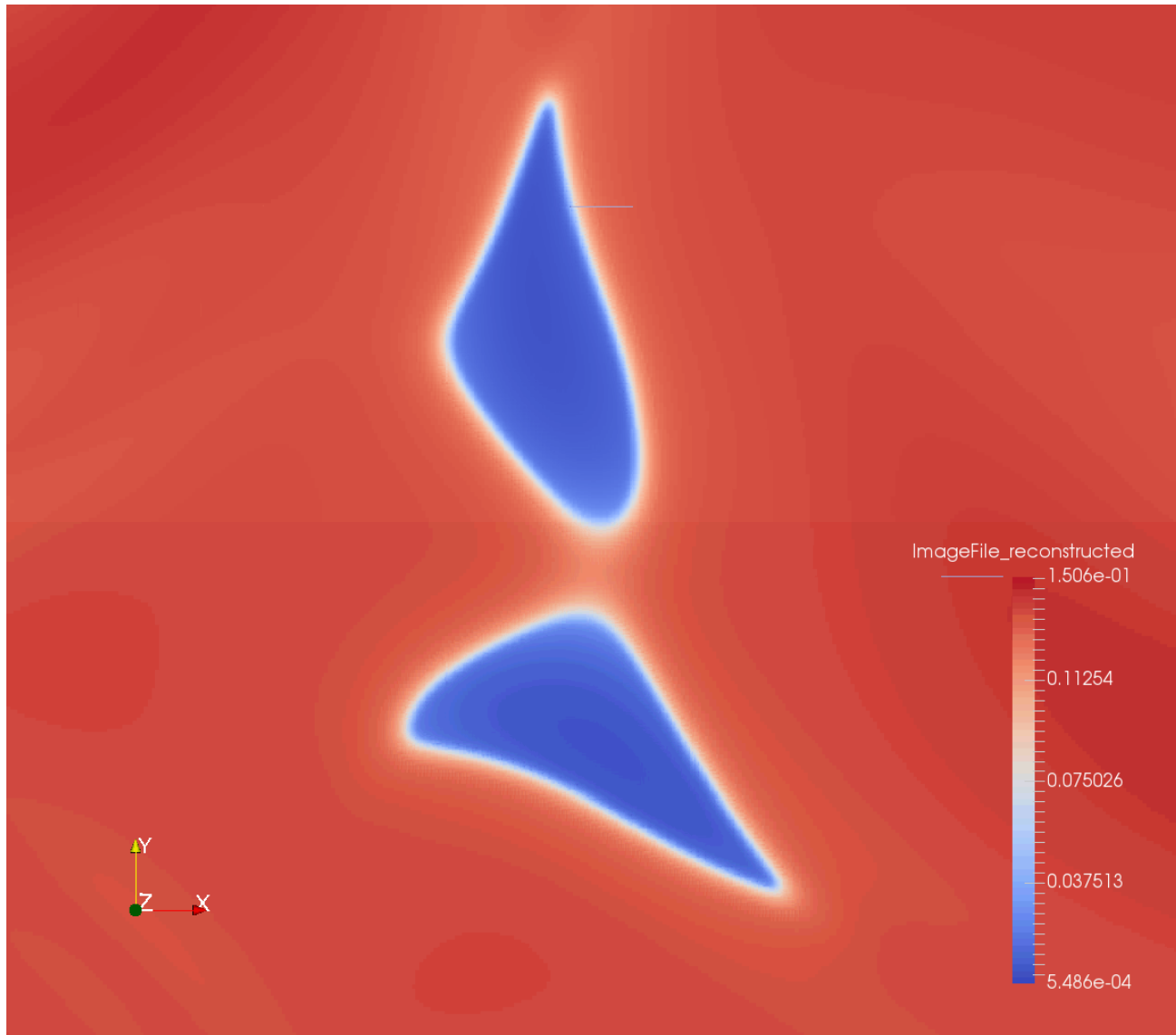
Visual Errors – CS – 100x



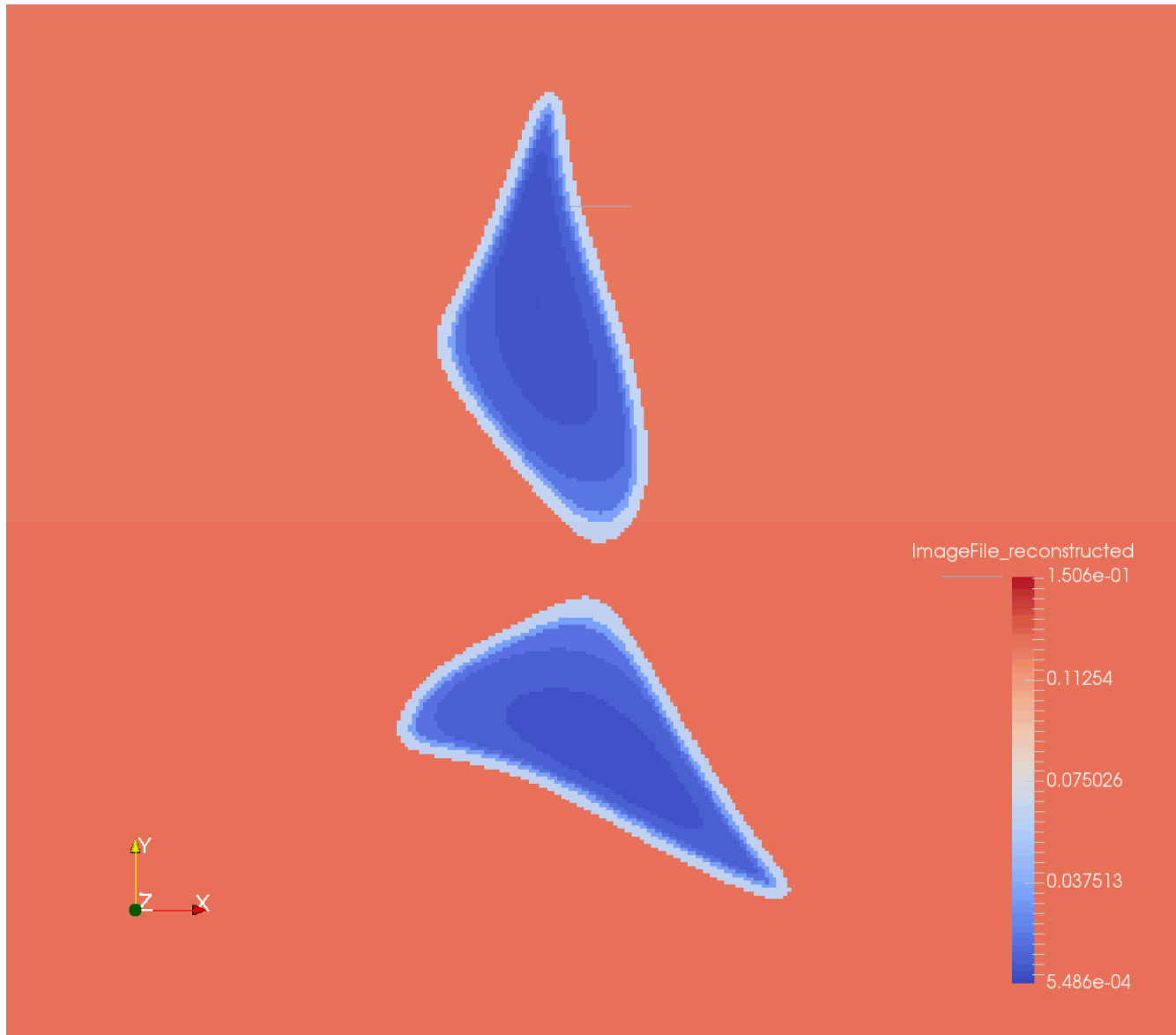
Visual Errors – VQ – 60x



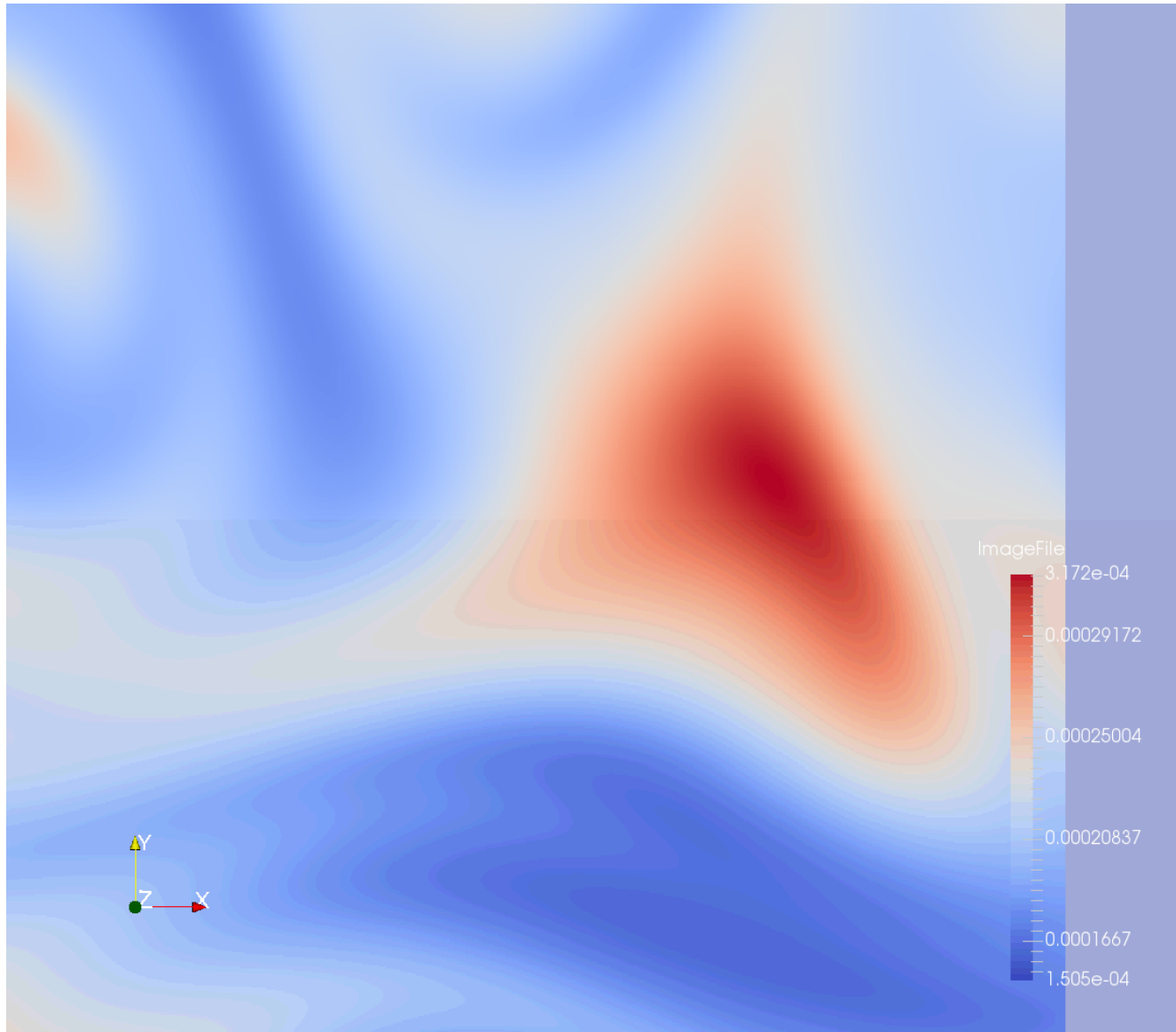
Visual Errors – ZFP – 32x



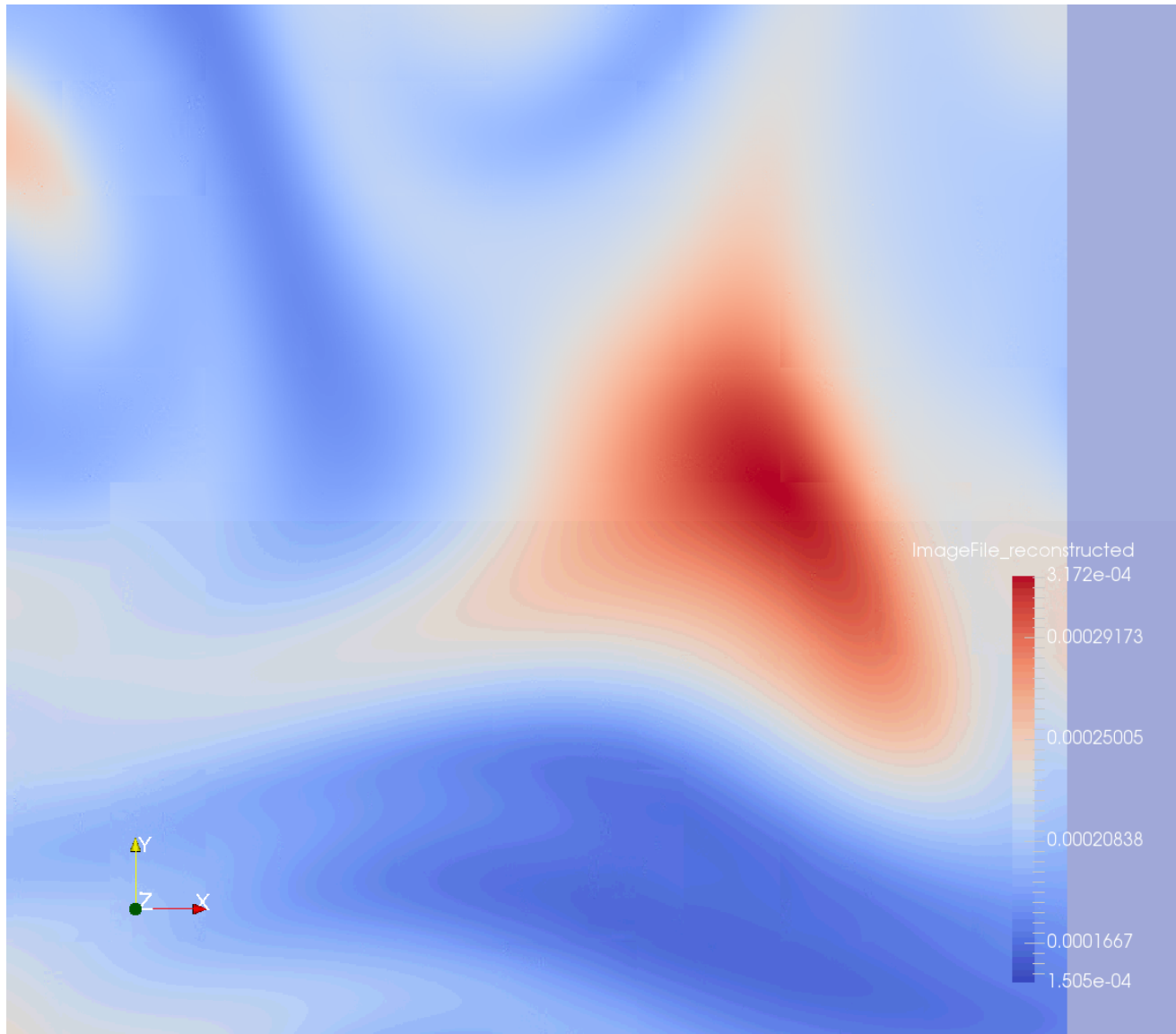
Visual Errors – ZFP – 60x



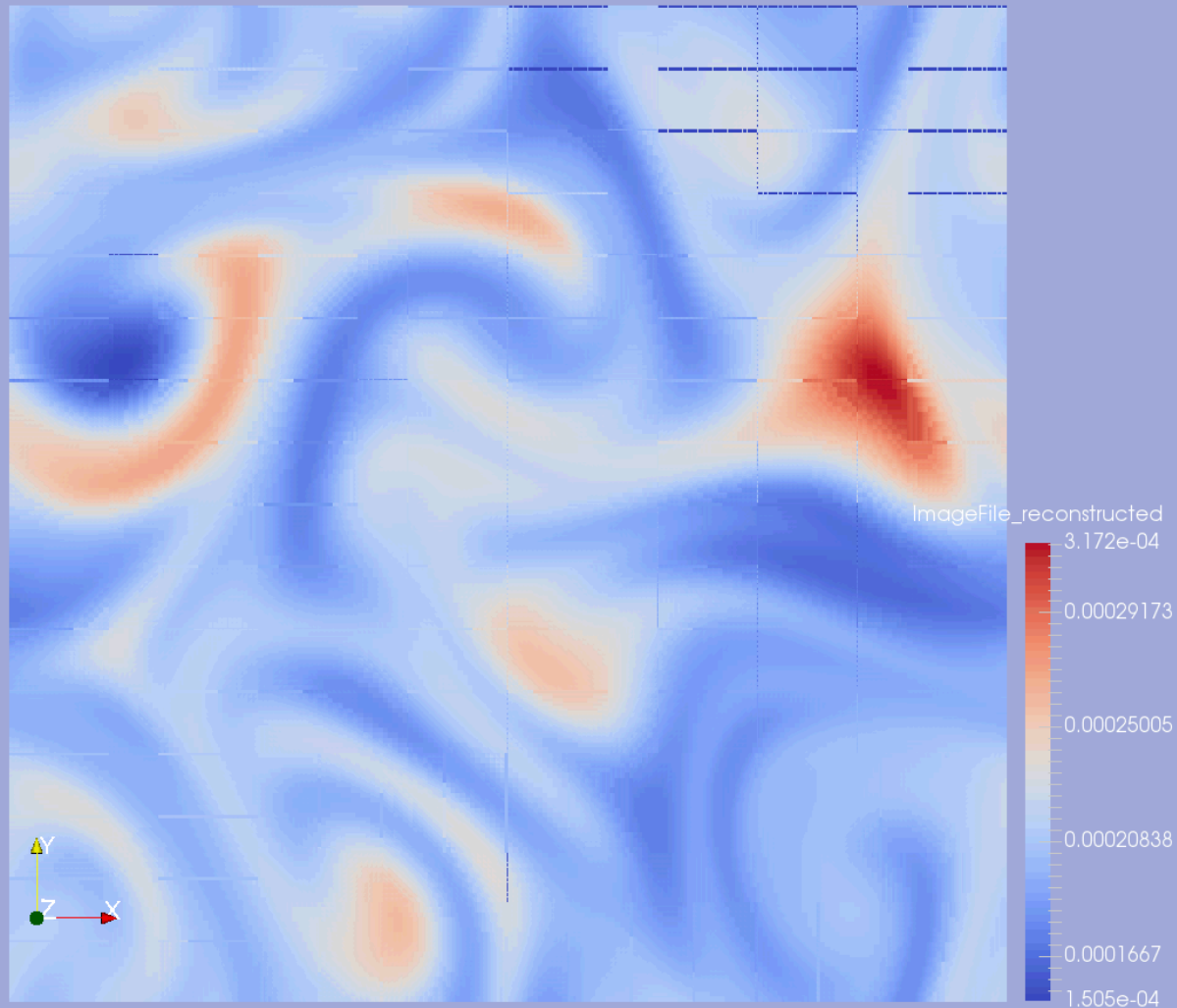
Visual Errors – Wavelet – 100x



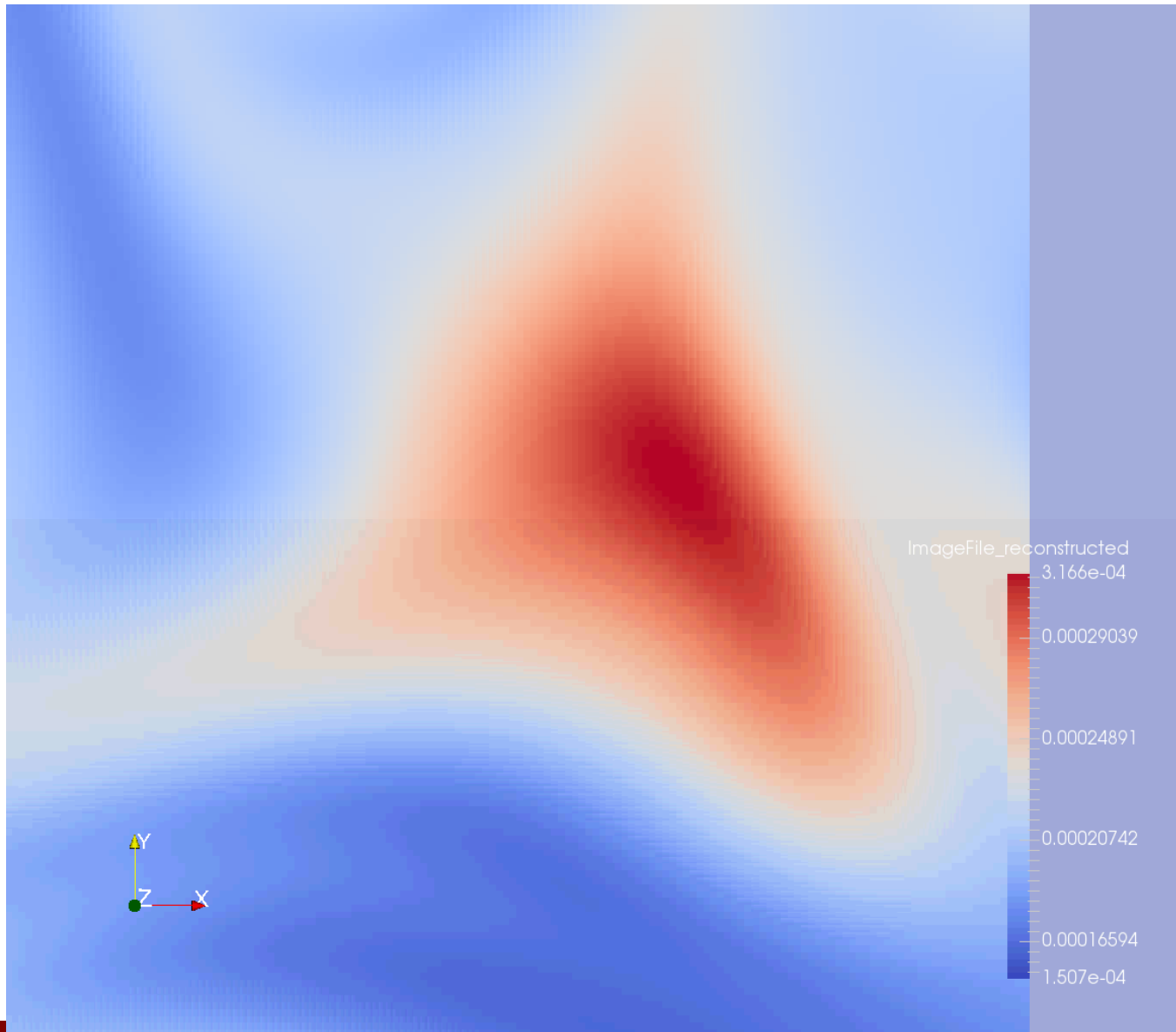
Visual Errors – CS – 100x



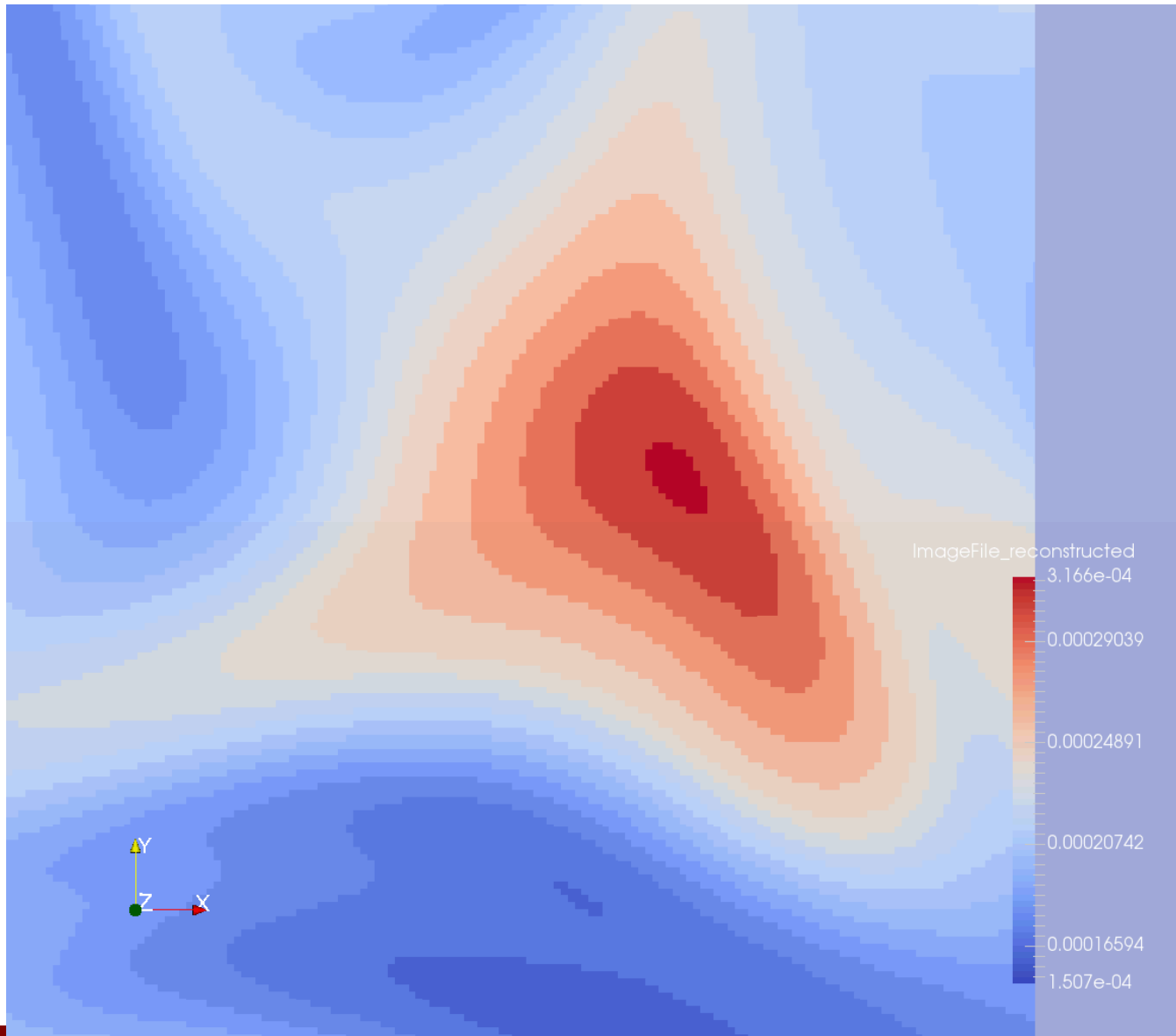
Visual Errors – VQ – 10x



Visual Errors – ZFP – 32x

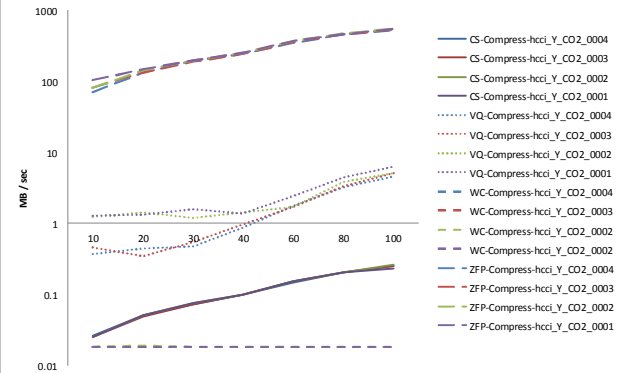


Visual Errors – ZFP – 40x

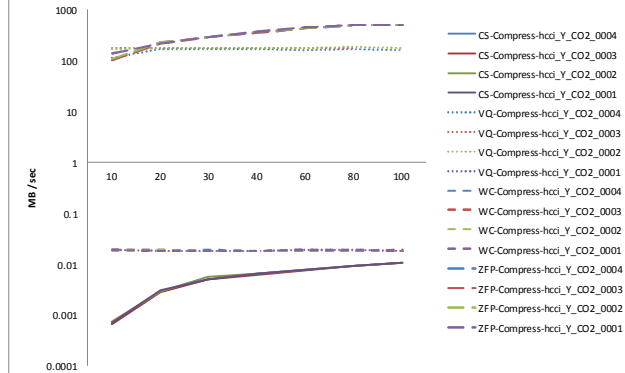


Structured Data Results

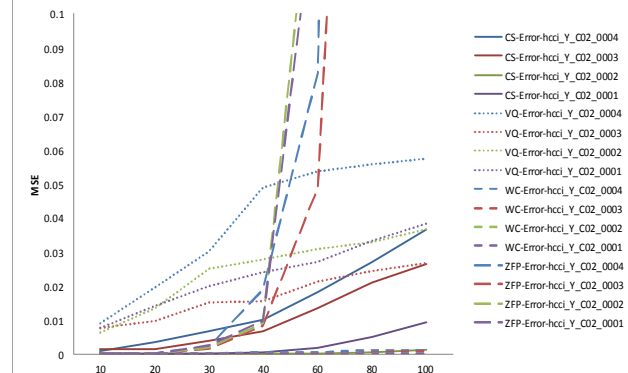
Reduction vs Log Minimum 1 Var Compression Rate



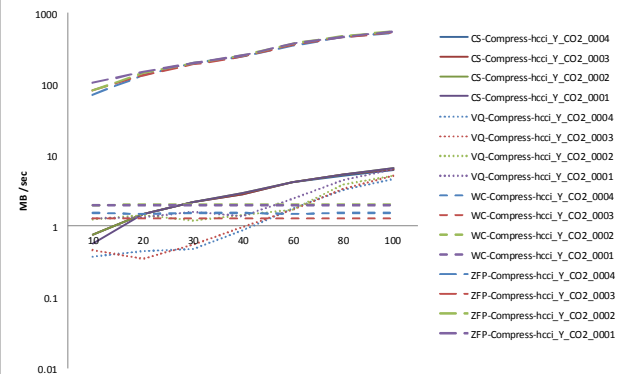
Reduction vs Log Minimum 1 Var Decompression Rate



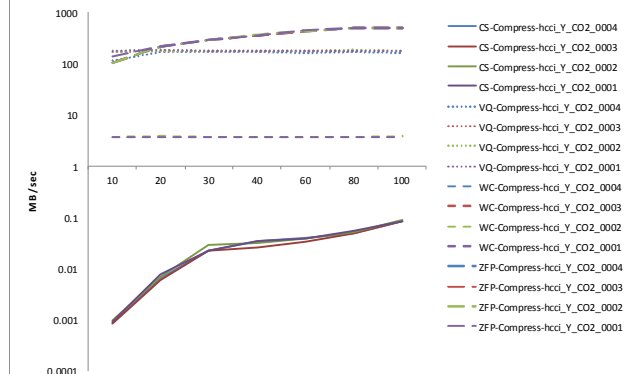
Reduction vs Error



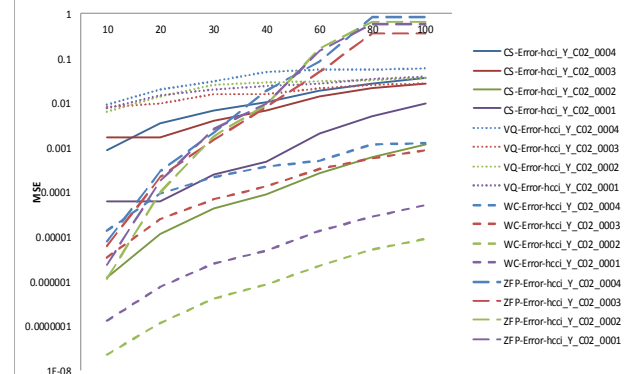
Reduction vs Log Minimum Inf Var Compression Rate



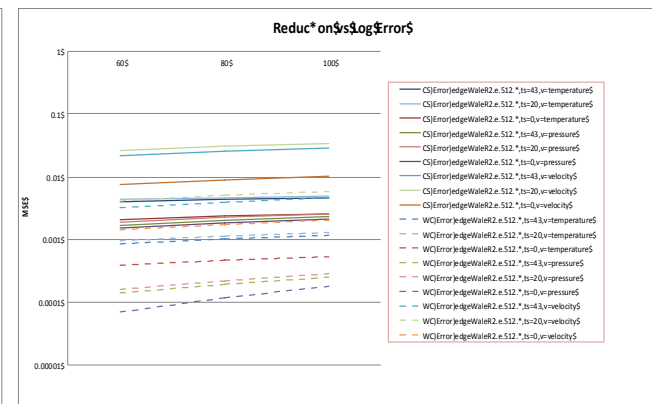
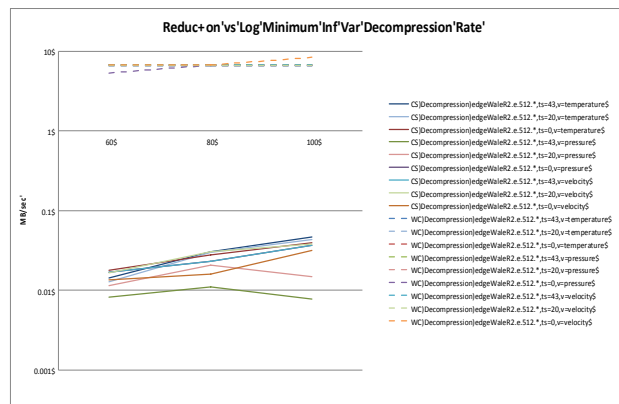
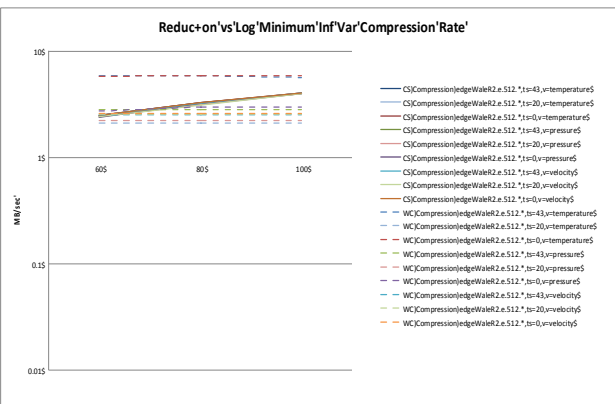
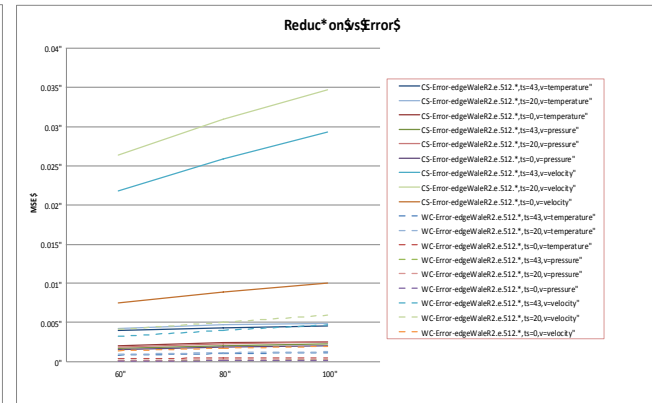
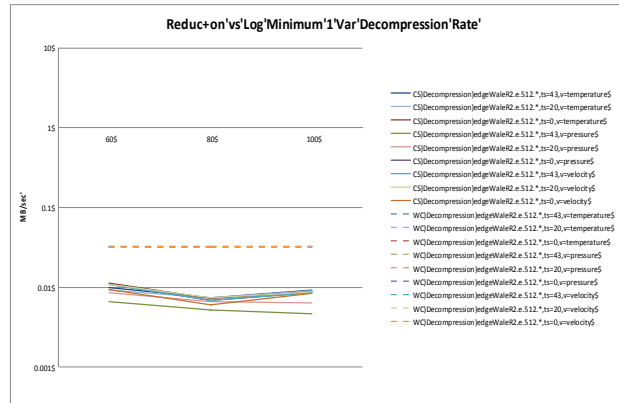
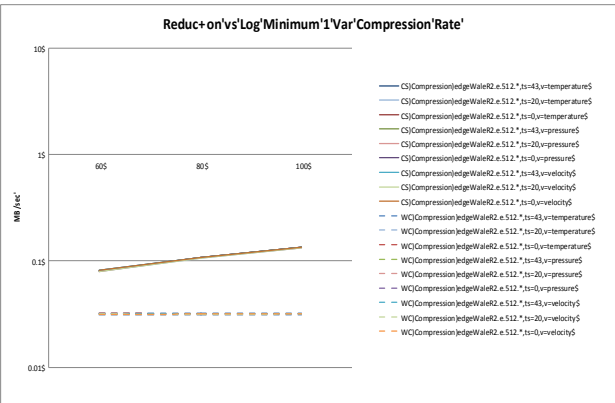
Reduction vs Log Minimum Inf Var Decompression Rate



Reduction vs Log Error



Unstructured Data Results



Future Work

- Understand the memory overheads
- Visual-system inspired error metric
- Higher levels of scale and compression