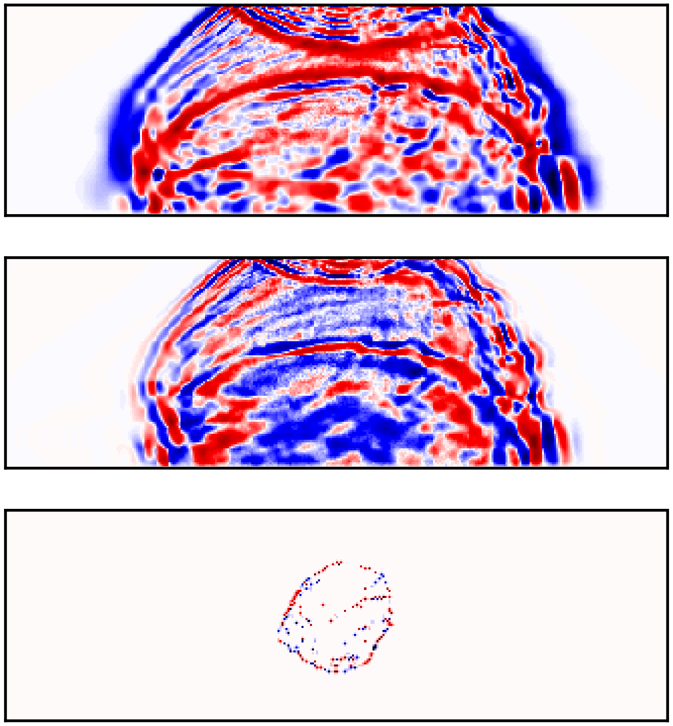
COMPRESSION FOR EFFECTIVE MEMORY BANDWIDTH USE IN FORWARD MODELING

EILEEN MARTIN

SEP ANNUAL MEETING
REPORT 152, PAGE 267
JUNE 2014

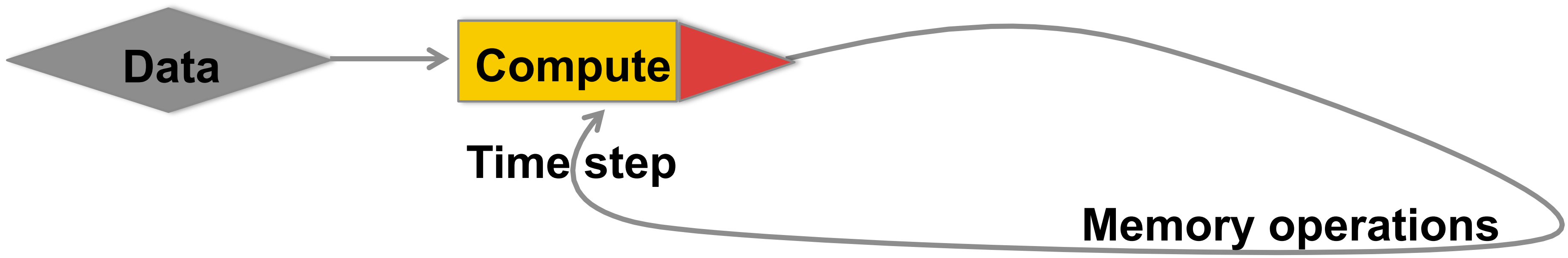


OUTLINE

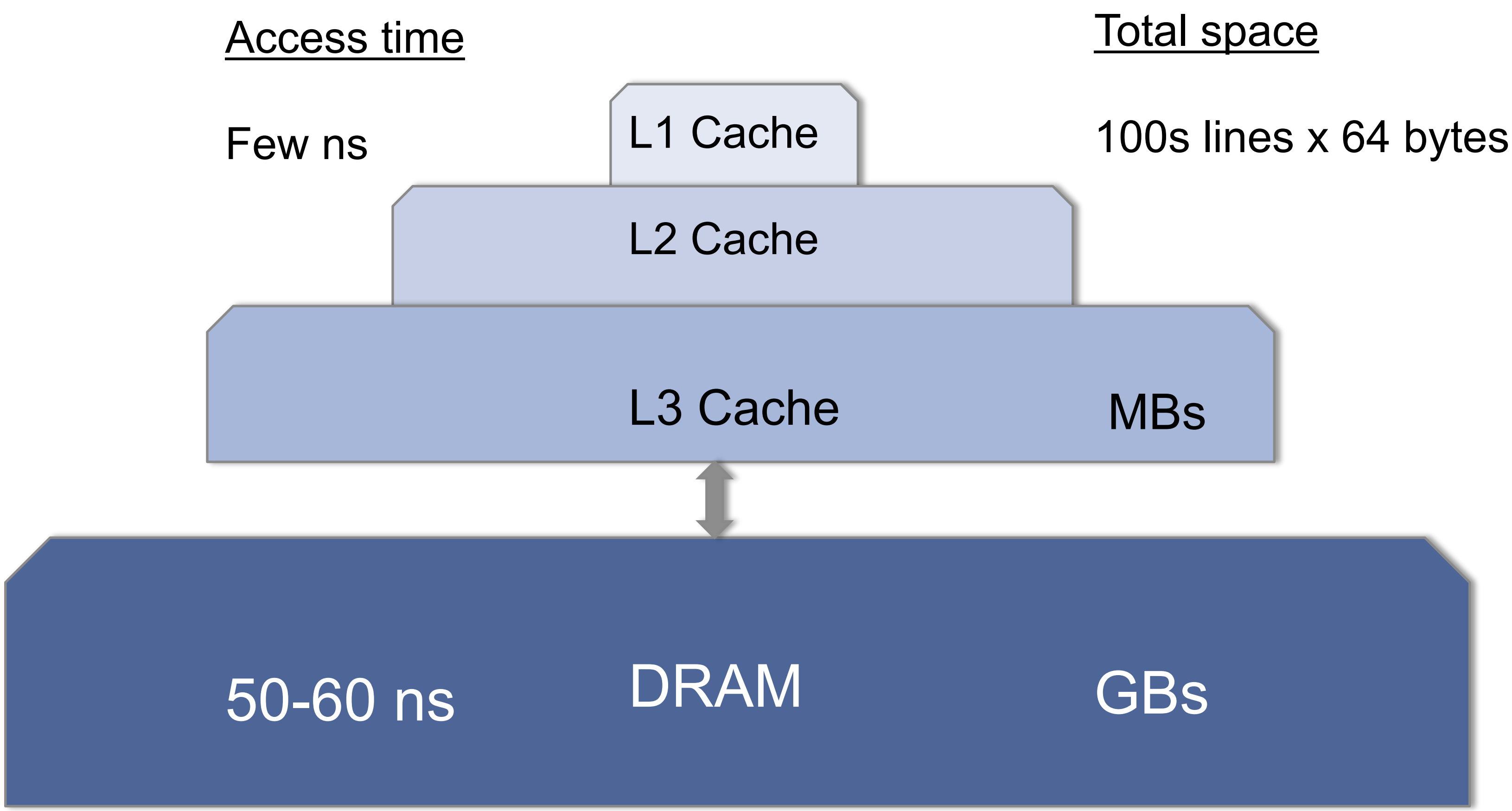
- **Introduction**
- **Compression in cache**
  - Theory, previous work
  - Numerical results
- **Compression when writing to disk**
  - Theory
  - Numerical results
- **Future work & conclusions**

WHY COMPRESS?

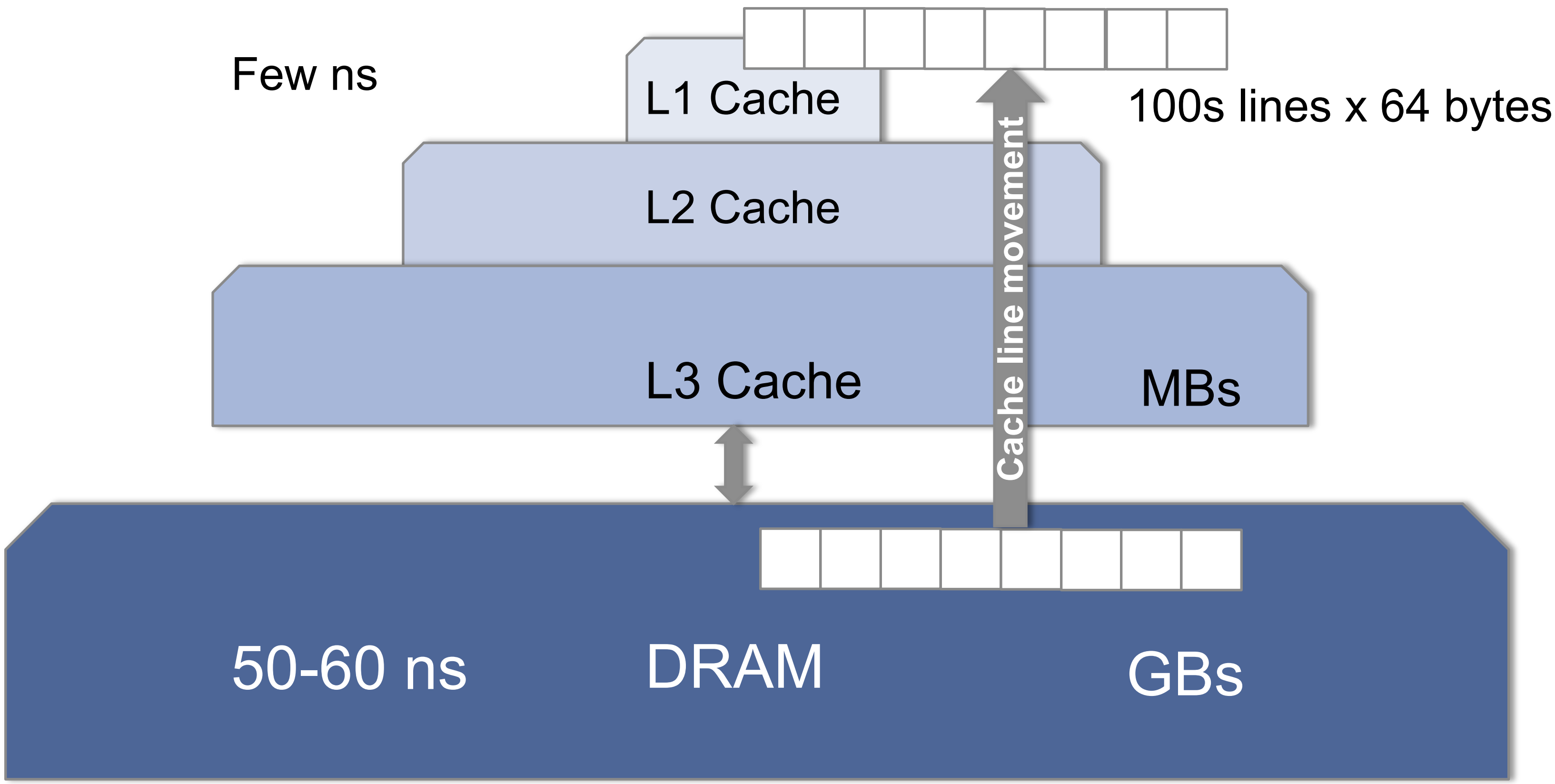
- Algorithms are limited by data movement
- Lossy compression gives higher compression rates
- Could use resources in other ways (larger data, UQ)
- Some imaging algorithms robust to incoherent errors



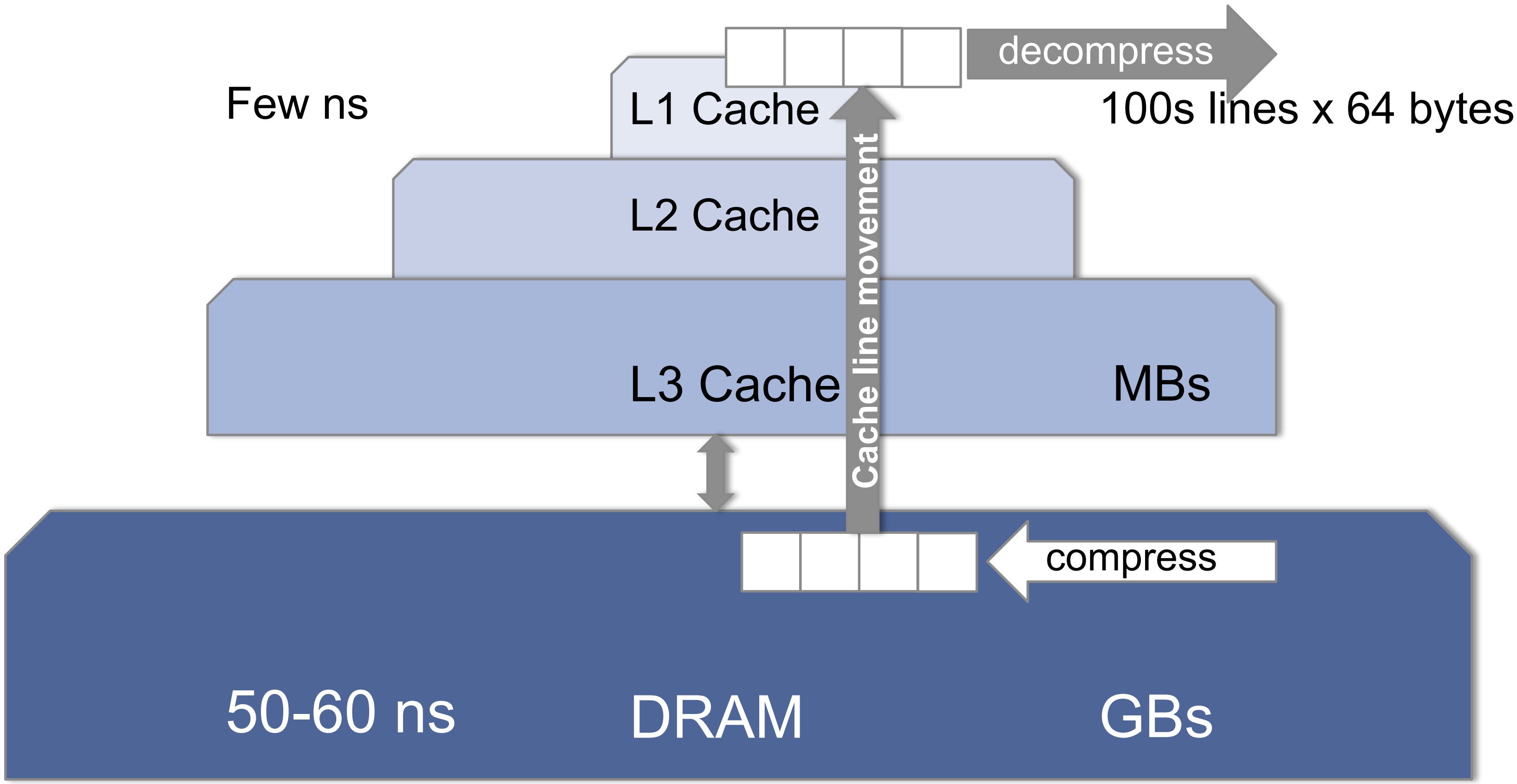
CACHE OVERVIEW



CACHE OVERVIEW



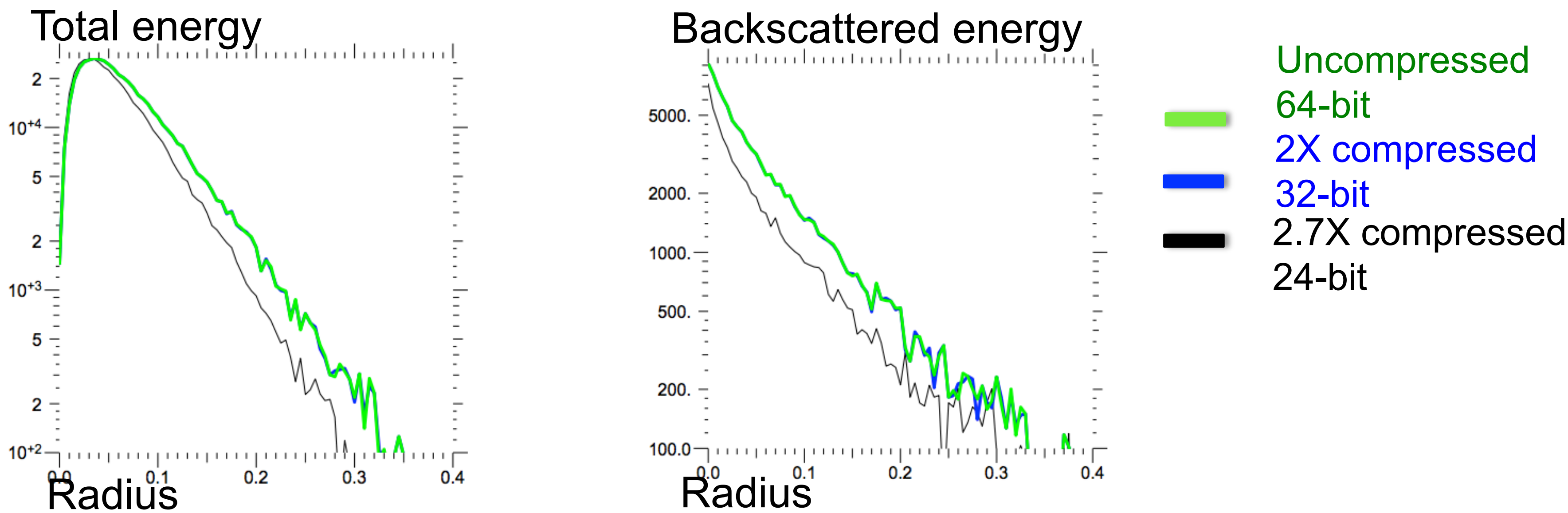
CACHE OVERVIEW



PREVIOUS WORK

Compression by fpzip ~2x, APAX up to 5x (Laney et al, 2013)

- LULESH- shock hydrodynamics
- Miranda- Navier-Stokes
- pF3D- laser-plasma interaction



COMPRESSED FORWARD MODELING

Initial conditions: $t = 0,\ u(0,x) = u_0(x),\ d_t u(0,x) = 0$

$u(0) \xrightarrow{\text{compress}} u_c(0)$

For i=1:nt

$u_c(t_{old}) \xrightarrow{\text{decompress}} u_{dc}(t_{old})$

$u_{dc}(t_{old}) \xrightarrow{\text{update}} u_u(t_i) = L(u(t_{old}))$

$u_u(t_i) \xrightarrow{\text{compress}} u_c(t_i)$

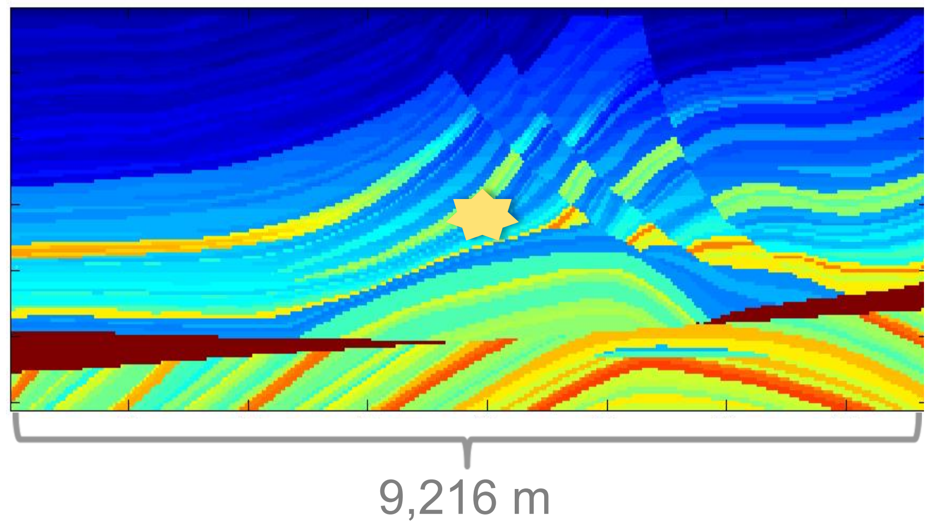
End

***Note:**

L is FDTD 2nd order time, 2nd order space wave operator

(De)compression with *fpzip_file_read/write*

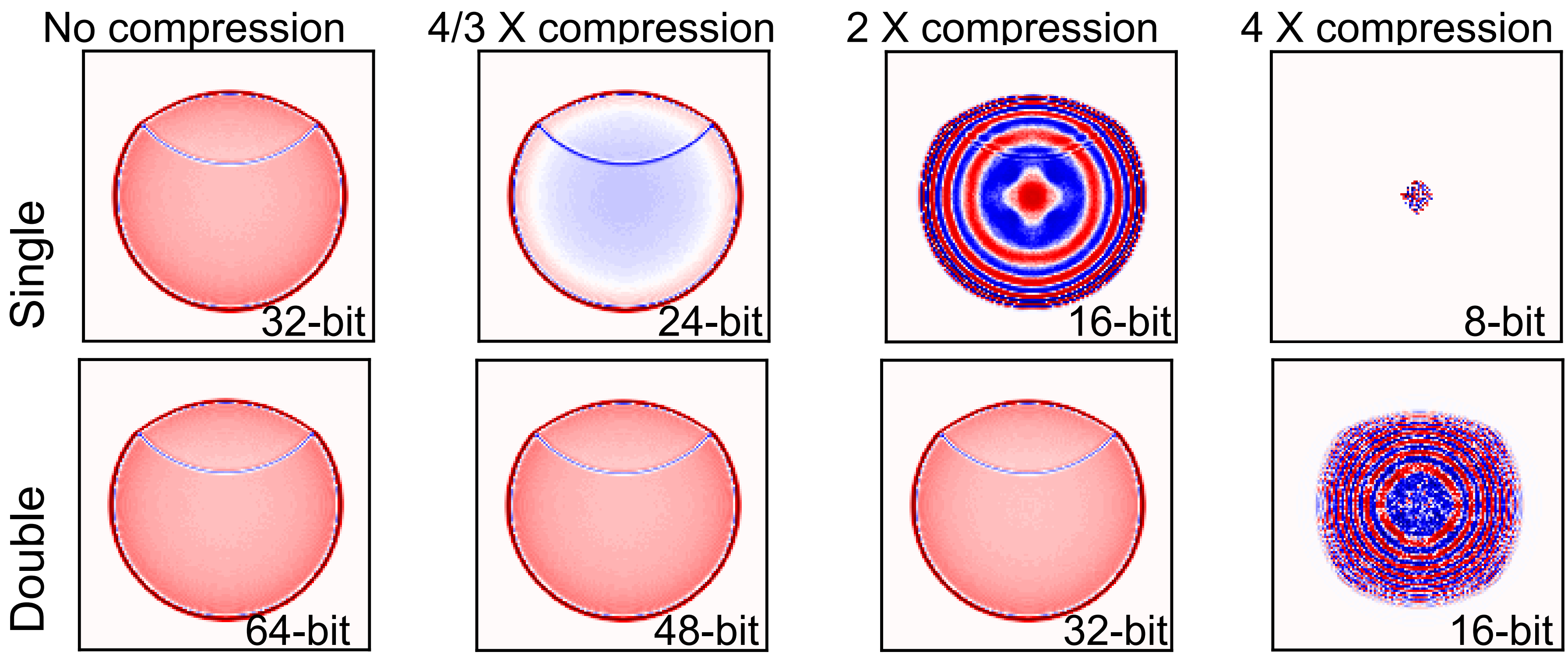
NUMERICAL EXAMPLES



Source: 10 Hz peak frequency Ricker wavelet
FDTD: 2nd order space, 2nd order time
Reflective boundaries on all sides

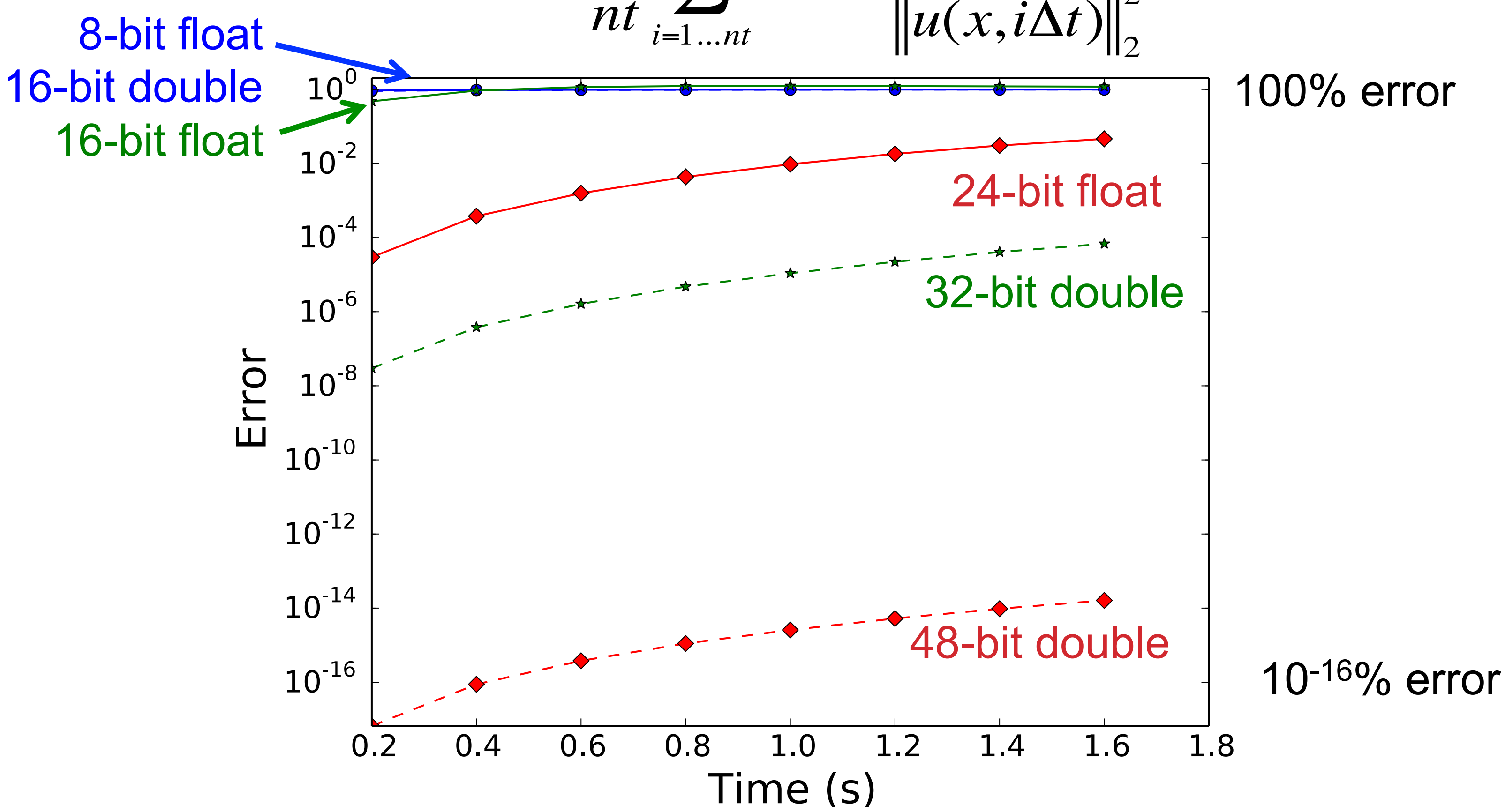
Marmousi model downloaded from Benamou, 1996

TWO-LAYER COMPRESSION

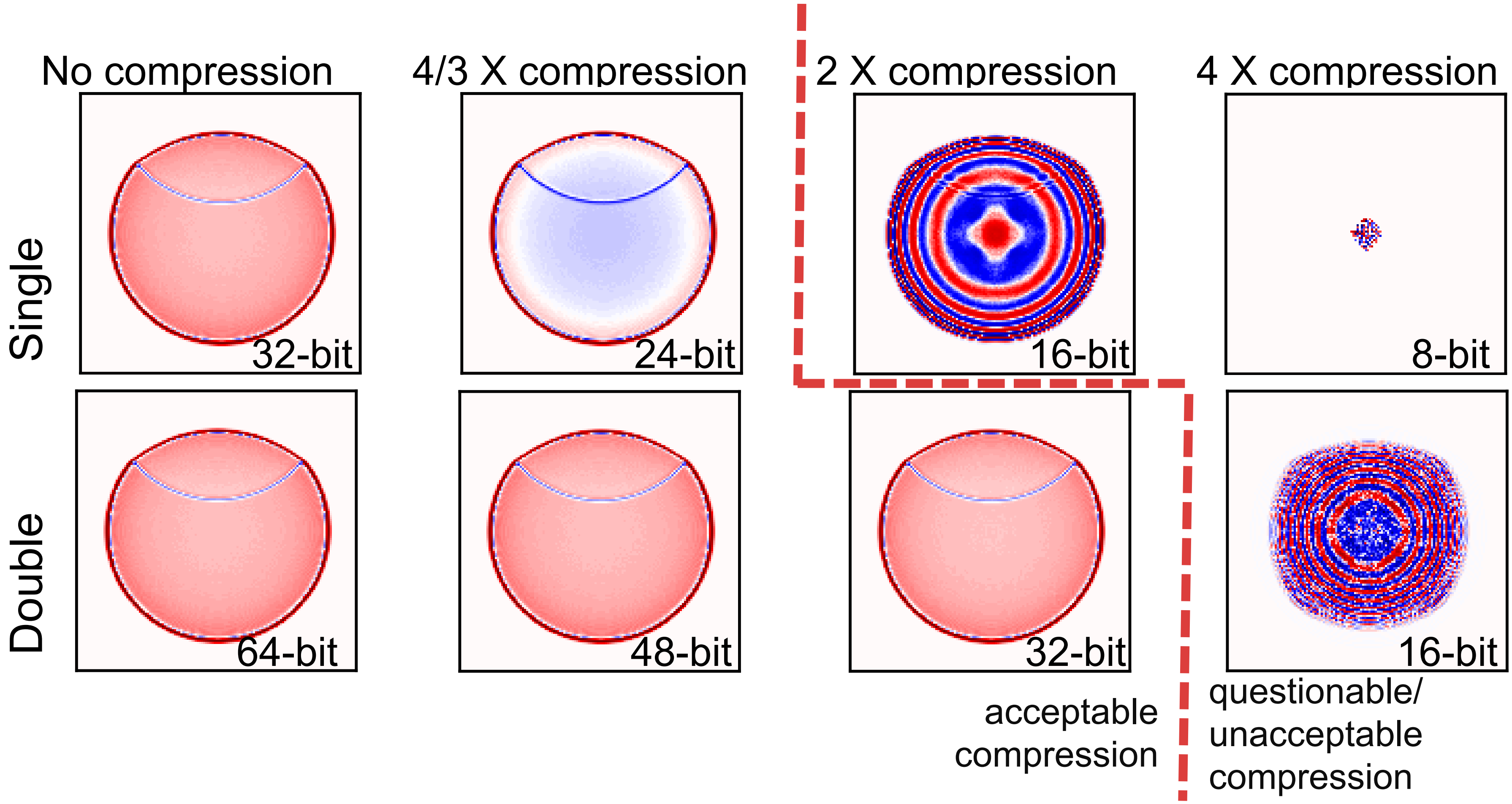


COMPARISON OF ERRORS

$$\frac{1}{nt}\sum_{i=1\ldots nt}\frac{\left\|u(x,i\Delta t)-u_c(x,i\Delta t)\right\|_2^2}{\left\|u(x,i\Delta t)\right\|_2^2}$$

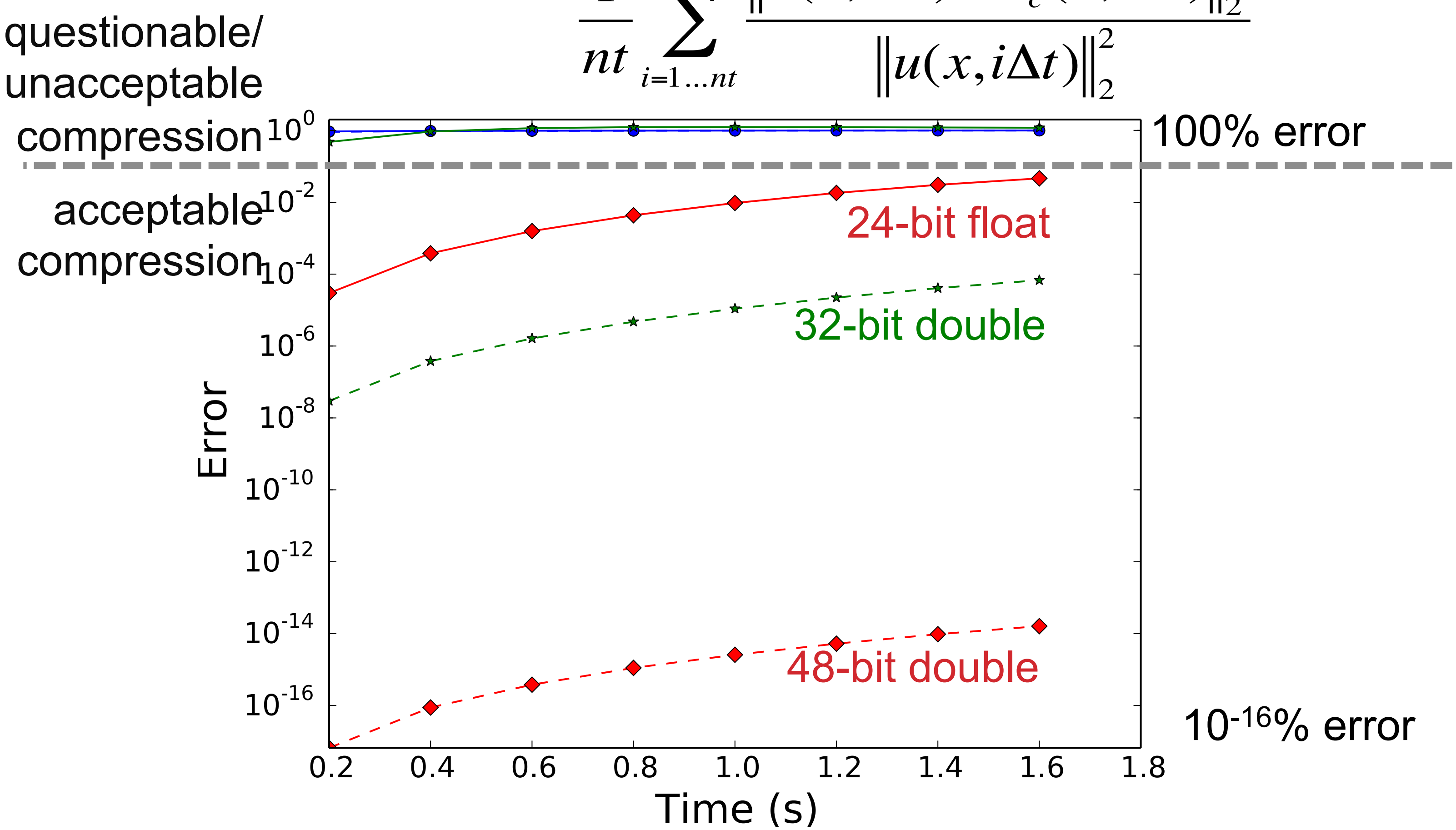


TWO-LAYER COMPRESSION

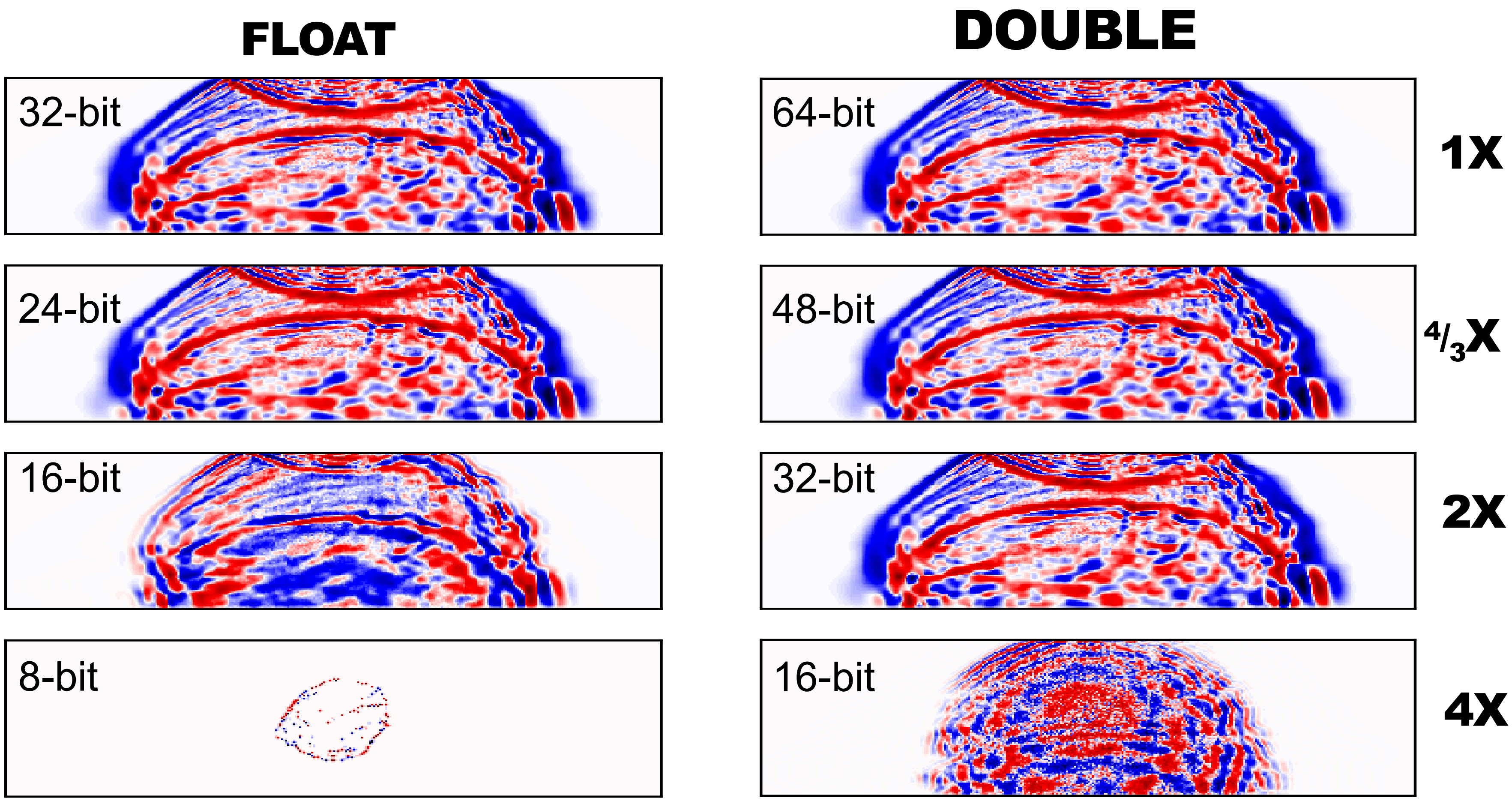


COMPARISON OF ERRORS

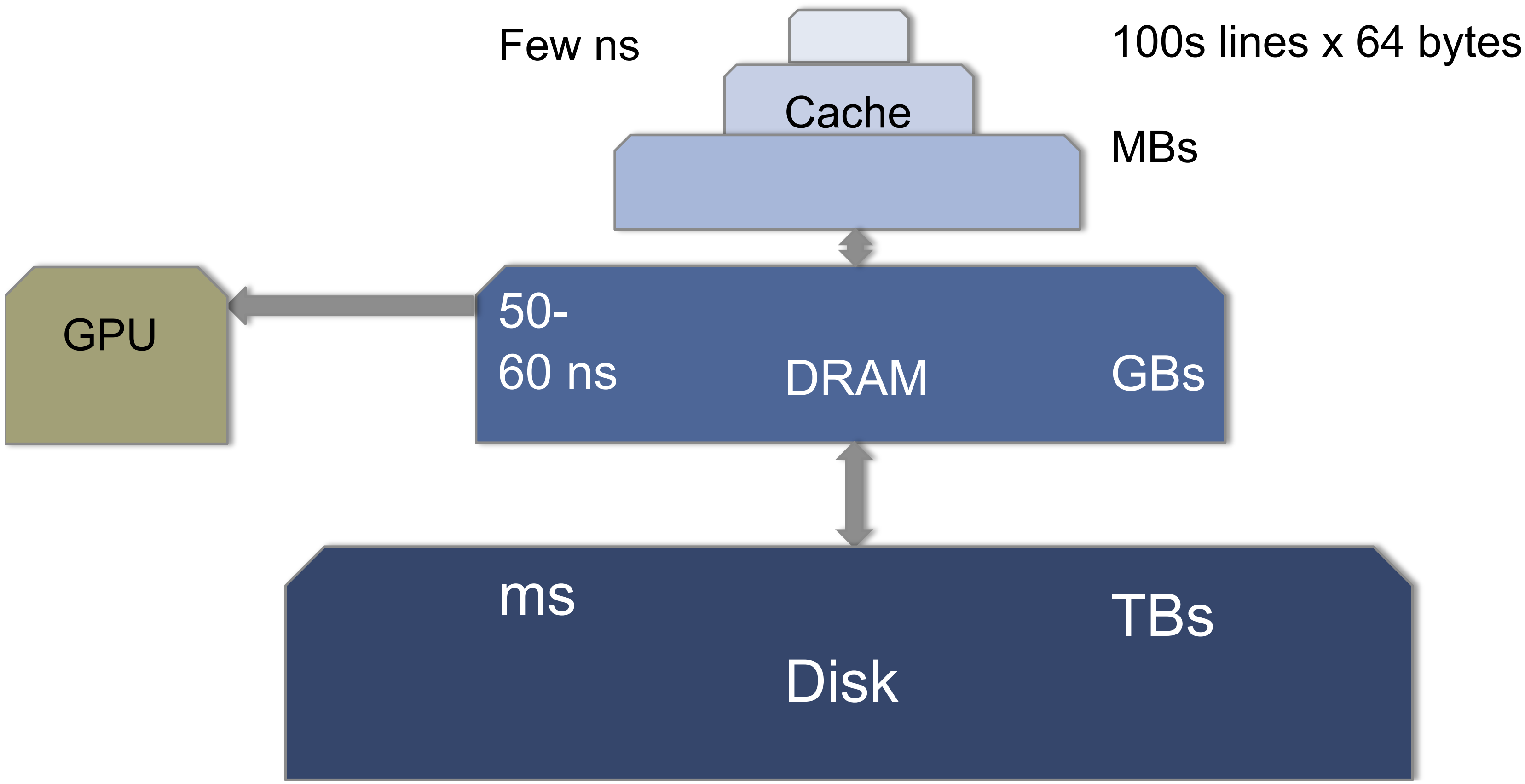
$$\frac{1}{nt}\sum_{i=1\ldots nt}\frac{\|u(x,i\Delta t)-u_c(x,i\Delta t)\|_2^2}{\|u(x,i\Delta t)\|_2^2}$$



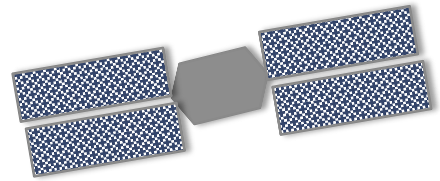
MARMOUSI COMPRESSION



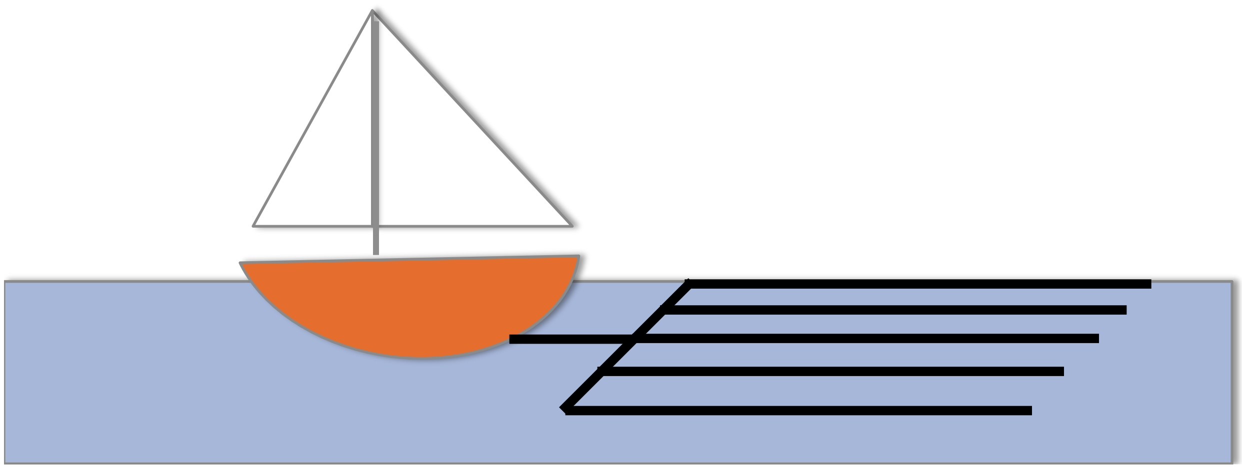
COMPRESSION WHEN WRITING TO DISK



WAVELETS

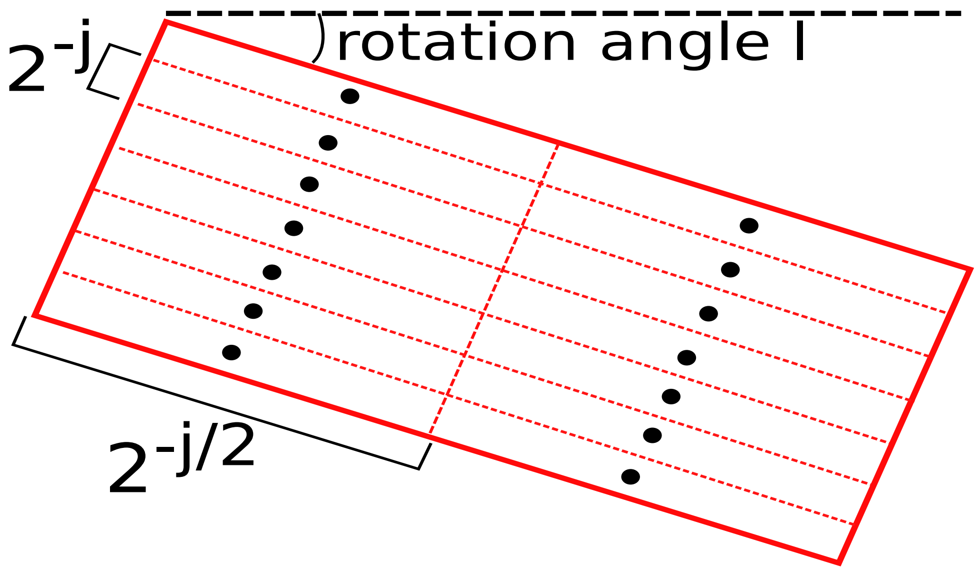


- **Redundant function space basis**
- **2D seismic data, ~20x compression (Bosman & Reiter, 1993)**
- **3D seismic data, ~100x compression (Villasenor et al, 1996)**
  - Using CDF 9/7 wavelet
- **Efficient 2D wavelet compression on GPU (Wong et al, 2007)**
- **Wavelet frame induces sparsity for wave propagation (Wu et al, 2008)**



CURVELETS

- **Idea: long thin wavelets at an angle designed to approximate curves well (Candès et al, 2006; Mallat, 2008)**



COMPRESSED FORWARD MODELING (CACHE LINES)

Initial conditions: $t = 0,\ u(0,x) = u_0(x),\ d_t u(0,x) = 0$

$u(0) \xrightarrow{\text{compress}} u_c(0)$

For i=1:nt

$u_c(t_{old}) \xrightarrow{\text{decompress}} u_{dc}(t_{old})$

$u_{dc}(t_{old}) \xrightarrow{\text{update}} u_u(t_i) = L(u(t_{old}))$

$u_u(t_i) \xrightarrow{\text{compress}} u_c(t_i)$

End

***Note:**

L **is FDTD 2nd order time, 2nd order space wave operator**

(De)compression with *fpzip_file_read/write*

COMPRESSED FORWARD MODELING RESTARTING

Initial conditions: $t = 0, u(0,x) = u_0(x), d_t u(0,x) = 0$

$u(0) \xrightarrow{\text{compress}} u_c(0)$

For i=1:nt

if(i mod f == 1) $u_c(t_{old}) \xrightarrow{\text{decompress}} u_{dc}(t_{old})$

$u_{dc}(t_{old}) \xrightarrow{\text{update}} u_u(t_i) = L(u(t_{old}))$

if(i mod f == 0) $u_u(t_i) \xrightarrow{\text{compress}} u_c(t_i)$

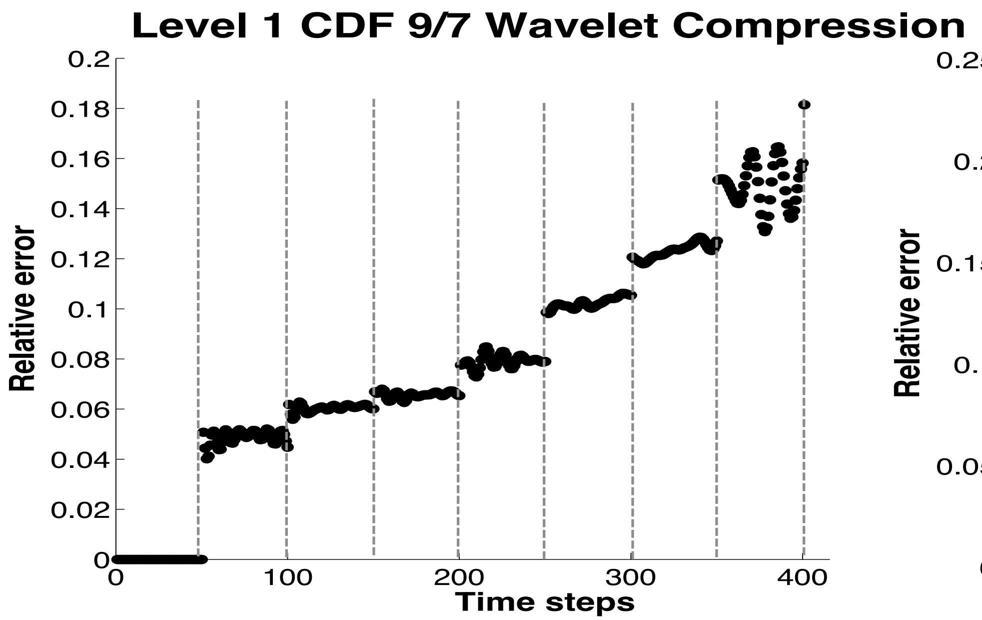
End

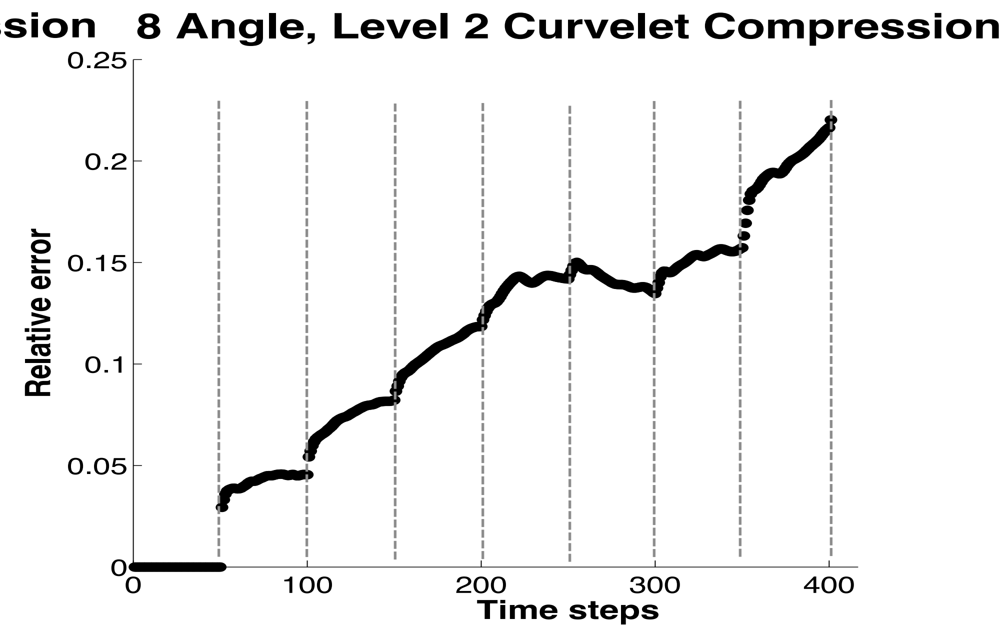
***Note:** f is the frequency of restarting

L is FDTD 2nd order time, 2nd order space wave operator

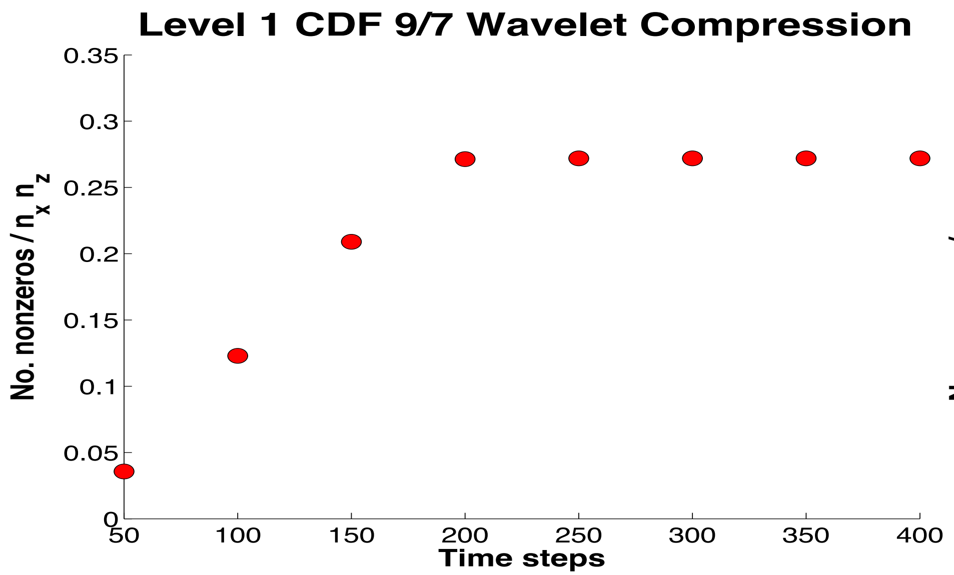
(De)compression with *CurveLab or Matlab Wavelet Toolbox*

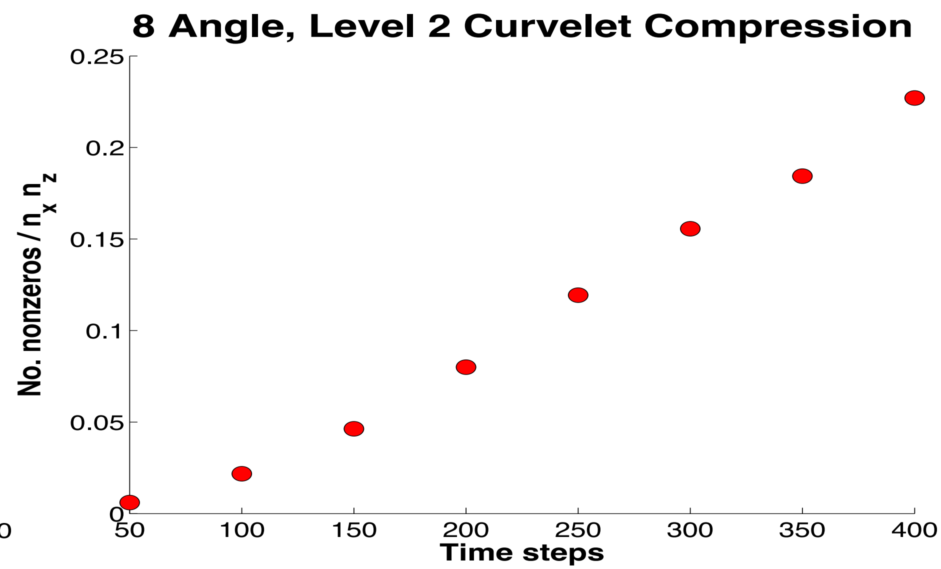
ERRORS: CHECKPOINT & RESTART





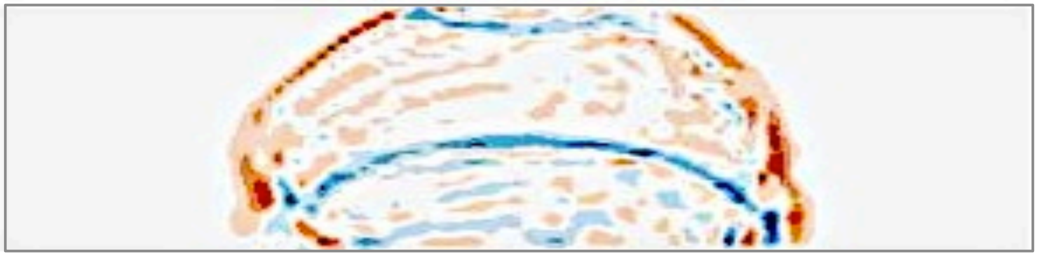
COMPRESSION RATIOS: CHECKPOINT & RESTART



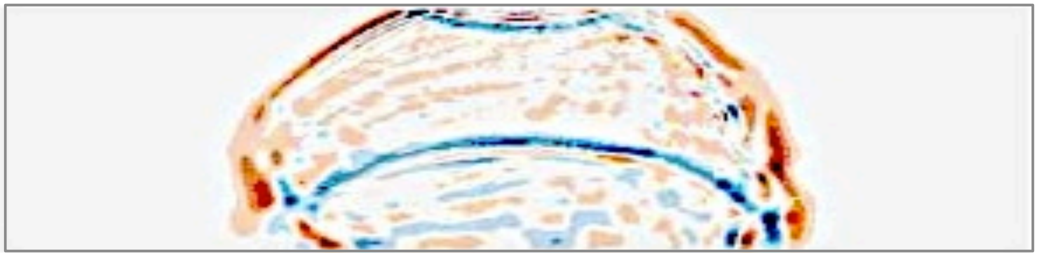


INDEPENDENT COMPRESSIONS

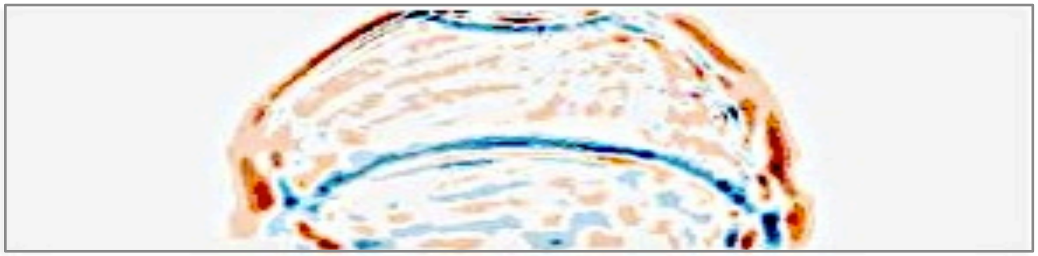
2 level CDF 9/7 wavelet
3800 nonzeros
35% relative error



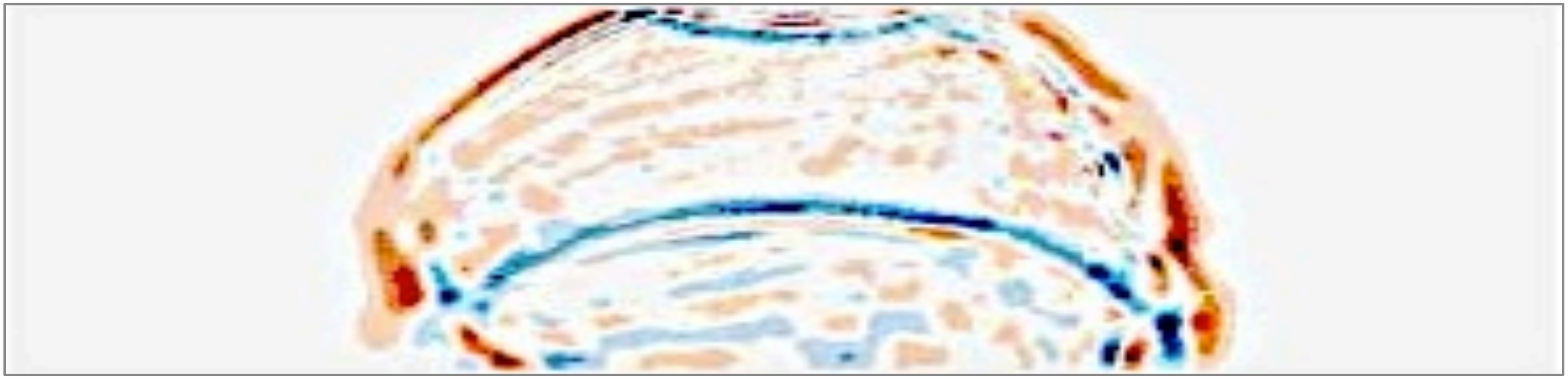
No compression
46848 grid
45840 nonzeros
~30,000 significant nonzeros



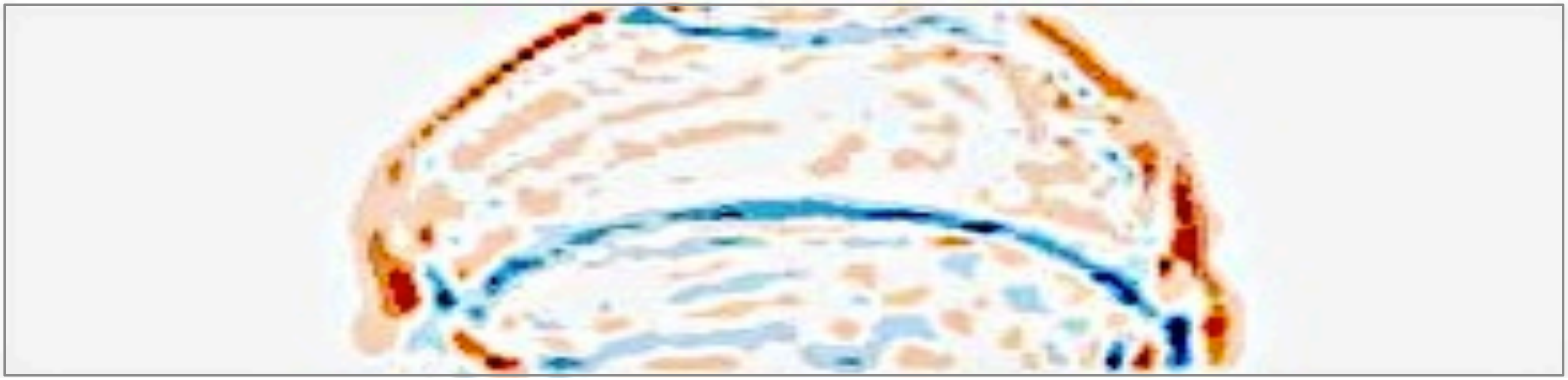
2 level 8 angle curvelet
6500 nonzeros
13% relative error



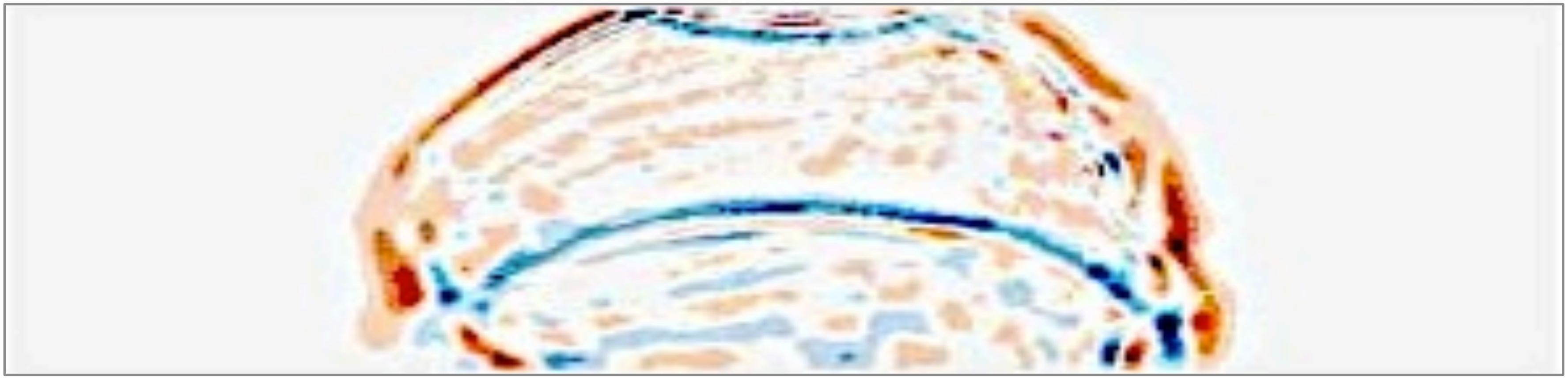
NO COMPRESSION



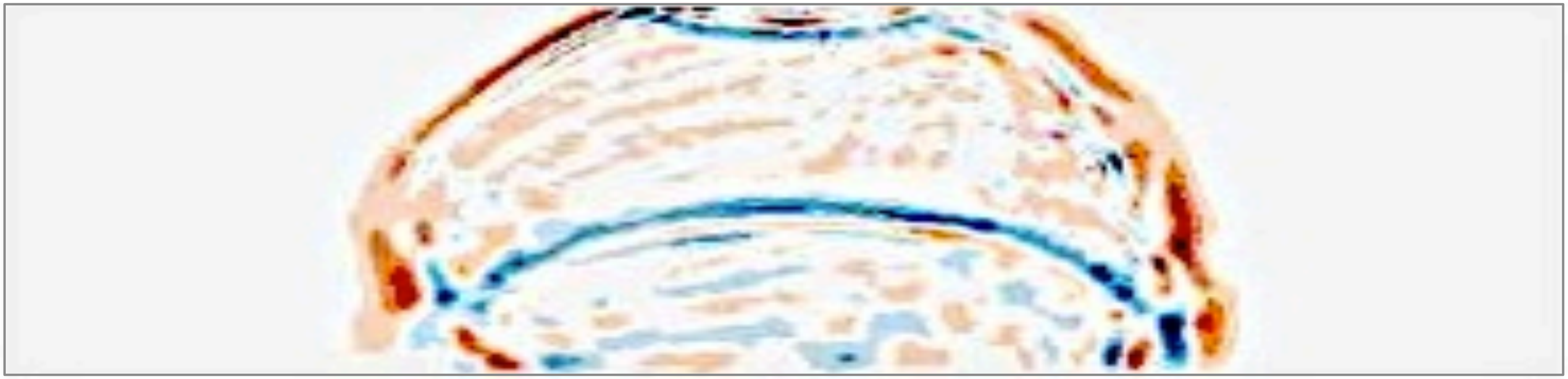
2 LEVEL CDF 9/7 WAVELET



NO COMPRESSION



2 LEVEL 8 ANGLE CURVELET



CONCLUSIONS

- **Modest compression possible with fpzip**
- **More compression possible with 2D wavelets & curvelets**
- **Checkpoint & restart errors build over time**
- **Compressing individual wave fields when writing to disk allows more constant control of error**

FUTURE WORK

- **Cache line scale compression:**
  - Try with APAX software at LLNL
- **Whole wave field compression:**
  - Compression of 3D scalar wave fields
  - Forward modeling on GPU, compress disk writes
  - Integration into imaging process
    - Use meaningful error metric (Clapp, 2008)
    - Focus on RTM initially

THANKS

- This work supported in part by DOE CSGF grant no. DE-FG02-97ER25308
- Thanks to Bob Clapp (SEP), Steve Langer (LLNL) and Peter Lindstrom (LLNL) for helpful conversations and suggestions.

REFERENCES

1. Benamou, J., 1996, Computation of multi-valued traveltimes in the marmousi model: Technical report.
2. Bosman, C. and E. Reiter, 1993, Seismic data compression using wavelet transforms: Expanded Abstracts of the 63rd Ann. Internat. Mtg., 1261–1264, Soc. Expl. Geo- phys.
3. Clapp, R., 2008, Prediction error filters to enhance differences: SEP Report, 131.
4. D. Laney, S. Langer, C. W. P. L. and A. Wegener, 2013, Assessing the effects of data compression in simulations using physically motivated metrics: Presented at the SC'13, Supercomputing.
5. E. Cand`es, L. Demanet, D. D. L. Y., 2007, Curvelab software.
6. J.D. Villasenor, R. E. and P. Donoho, 1996, Seismic data compression using high- dimensional wavelet transforms: , 396–405, IEEE.
7. Lindstrom, P. and M. Isenberg, 2006, Fast and efficient compression of floating-point data: IEEE Trans Vis Comput Graph, 12, 1245–1250.
8. ———, 2010, fpzip software.
9. Mallat, S., 2008, A wavelet tour of signal processing, 3 ed.: Academic Press.
10. R. Wu, B. W. and Y. Geng, 2008, Seismic wave propagation and imaging using time- space wavelets: Expanded Abstracts of the 78th Ann. Internat. Mtg., 2983–2987, Soc. Expl. Geophys.
11. T. Wong, C. Leung, P. H. and J. Wang, 2007, Discrete wavelet transform on consumer-level graphics hardware: IEEE Transactions on Multimedia, 9, 668–673.