MVAPICH
MPI, PGAS and Hybrid MPI+PGAS Library

Follow us on

https://twitter.com/mvapich

HiDL
High-Performance Deep Learning

OMB-Compr: An Extension to OSU Micro Benchmarks for Collective Compression Error Measurement

Jacob Queiser, **Nick Contini,** Hari Subramoni and Dhabaleswar K Panda

Introduction

Compression uses compute hardware to reduce data volume for transactions over system and network interconnects

- Lossy Compression
  - Introduces noise. Fidelity is either acceptable, or noise is restricted
  - Achieve higher compression ratios and throughput
  - Acceptable for images (JPEG), but may cause unease with high-precision, low tolerance data
- MVAPICH-Plus can leverage lossy compression in point-to-point and collective operations
  - Increases effective bandwidth
  - Reduces transmission time for large messages

Motivation

No benchmarking utility to target compression-enhanced lossy communication in the Message Passing Interface (MPI).

We want to...

- Provide MPI users a mechanism to quickly examine how lossy communication will impact their application data.
- Allow users to examine if the loss is tolerable or justifies the performance gains.
- Provide a standard tool for MPI designers to assess their performance trade-offs.

Contributions

We introduce OMB-Compr

1. We develop a method for measuring the error rate introduced by an MPI collective algorithm and integrate it into OSU Micro Benchmarks (OMB).
2. We detail a method for generating random communication data using different distributions. We provide a JSON format allowing users to specify the properties of the random distribution to model application data.
3. We present osu_alltoall results showing how the content of the send buffer affects error rate, demonstrating the utility of our extension.

Design: Metrics

Perform element wise calculations of the decompressed data against the expected

- Absolute Mean Error (ME)
- Mean Squared Error (MSE)
- Mean Relative Error (MRE)
- Peak Signal to Noise Ratio (PSNR)
    - Quality of the compression operation. The ratio of the maximum data signal amplitude to the MSE, in decibel (dBs) domain
- Also provide min/max error

$$MSE = \frac{\sum_{i=1}^{N}(decompressed_i - expected_i)^2}{N}$$

$$PSNR = 10\log_{10}\frac{\left(\frac{1}{2}(Expected\ Max - Expected\ Min)\right)^2}{MSE}$$

Design: Datasets

OMB originally uses empty send buffer for benchmarking.

OMB-Compr extends OMB with the ability to use Generated or Custom datasets to fill the send buffers

- Generative Data: Gaussian, Uniform, Gamma, Exponential.
  - Configured from a JSON file. Obtained through the GNU Scientific Library
- Custom Data: User provides dataset from a binary file

```
{
    "distribution": "uniform",
    "a": 0,
    "b": 1,
    "shift": 0,
    "scale": 7,
    "sparsity": 0.3
}
```

(a) Uniform Distribution

```
{
    "distribution": "gaussian",
    "sigma": 1.1,
    "mu": 0.3,
    "shift": 0,
    "scale": 1,
    "sparsity": 0
}
```

(b) Gaussian Distribution

Design: Challenges

Approach: Use the dataset to fill send buffers. Compare lossy communicated received buffers against the expected to obtain the loss metrics.

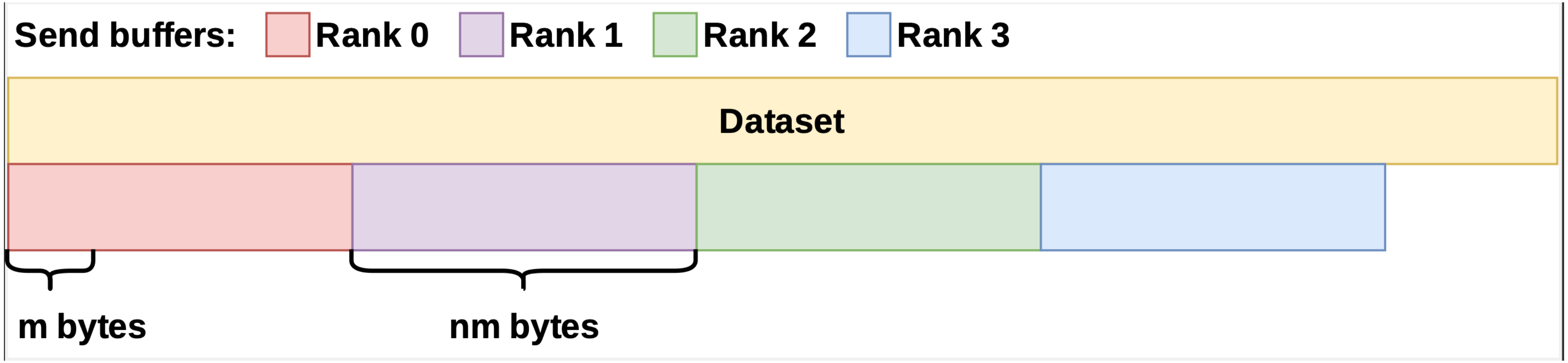
- Solution 1: Perform both compressed and normal communication versions of the collective.
  - **MPI standard doesn't specify how a compressed collective is enabled**
  - **OMB cannot assume if a compressed collective exists or how to turn it off**
- Solution 2: Use point-to-point primitives
  - Follows from the previous issues.
- Solution 3: Have each rank generate its send data in a deterministic manner
  - **Amicable to lossy compressors**
  - **Defeats use case of the benchmark**

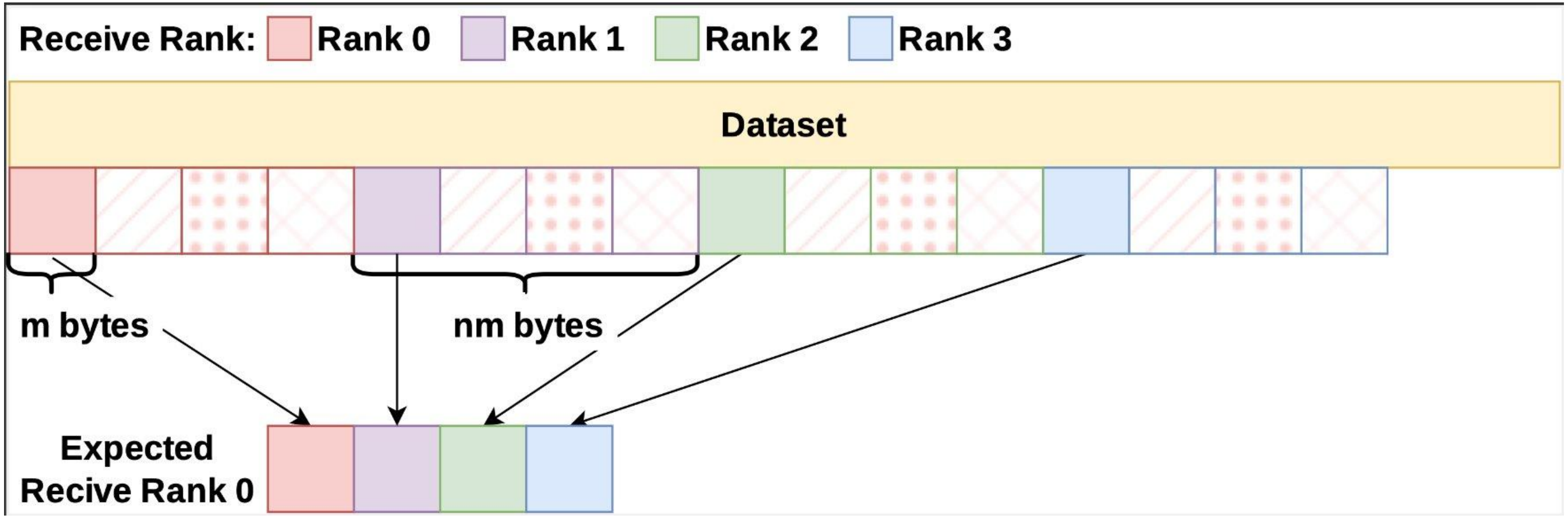
Design: Our Approach

Use the filesystem to communicate send buffers without compression

- A single rank can construct the dataset and store to disk
- Every rank can now see the contents of the others' send buffers to construct their expectation through this file
- Ranks read the file in the same pattern as the collective operation

Design: All-to-all

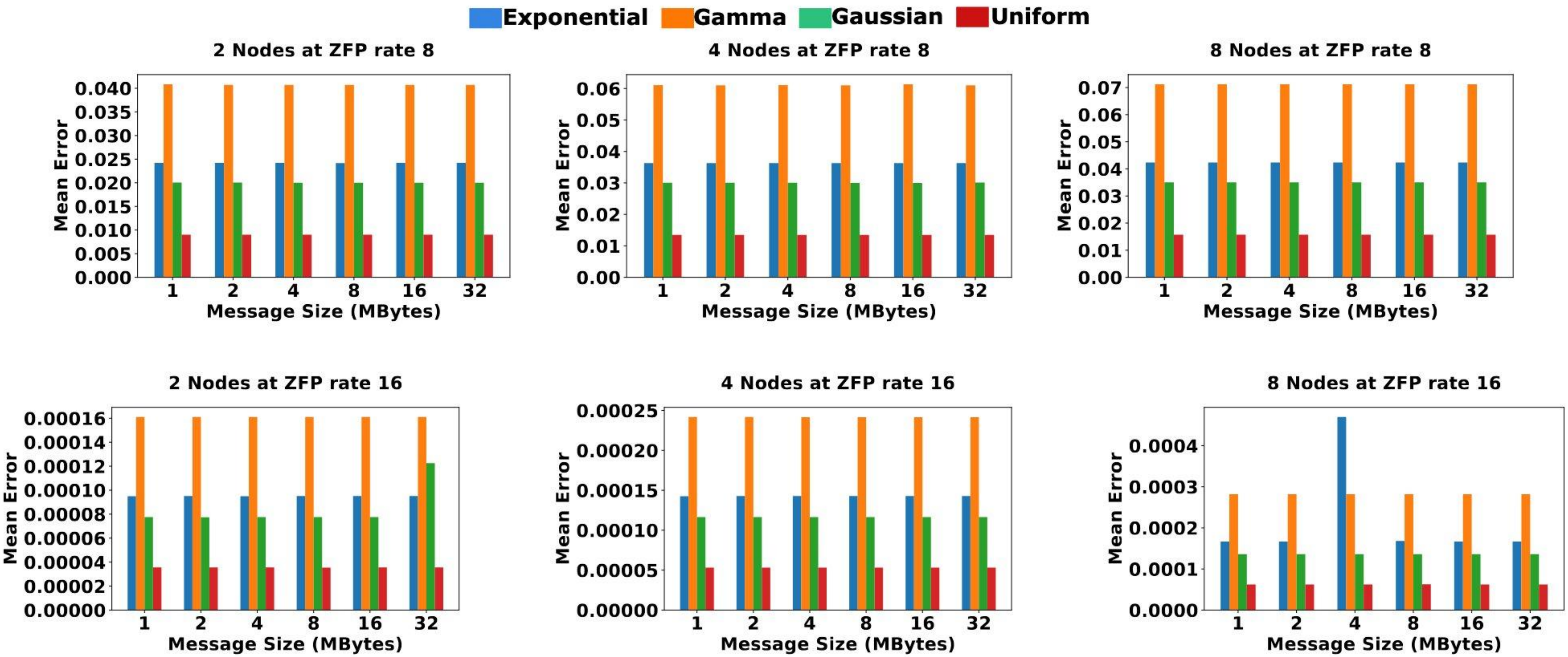




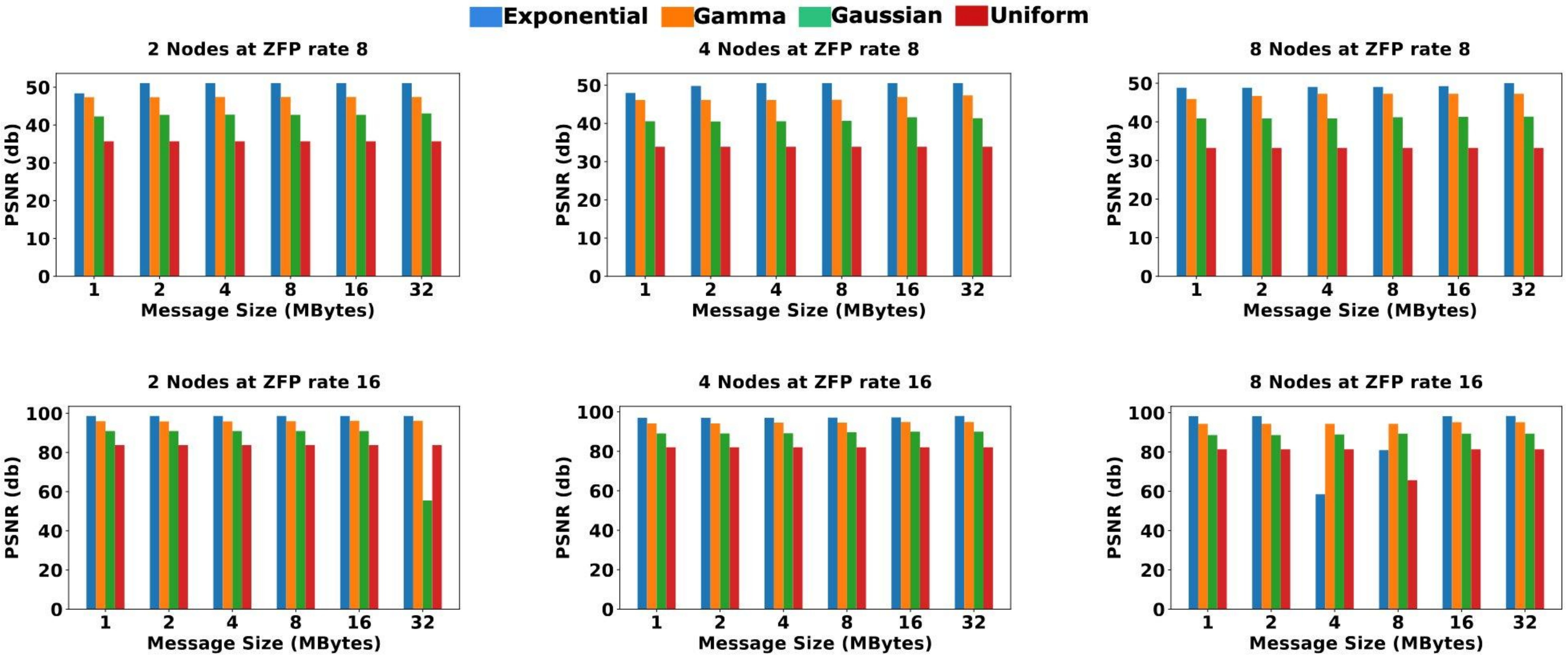
Evaluation

- MPI_Alltoall
- Enable ZFP-compression

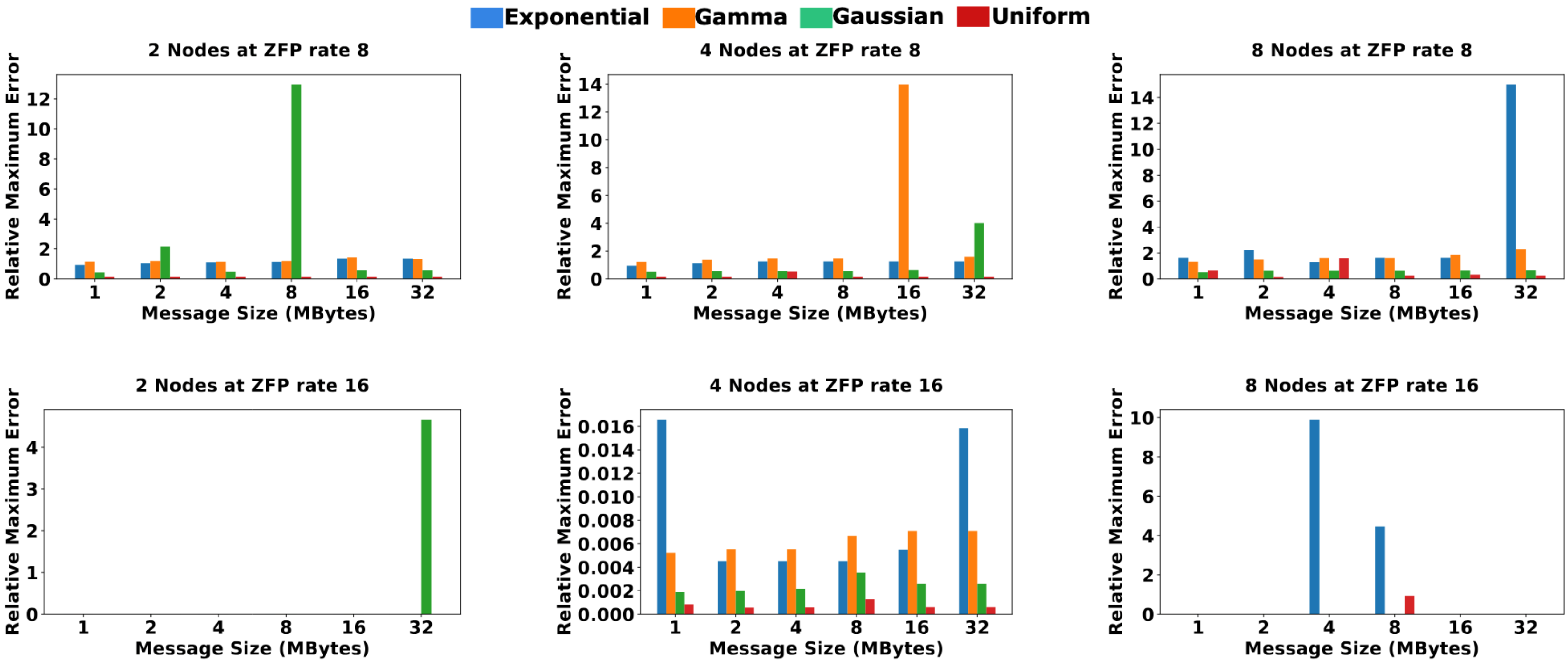
Results: Mean Error



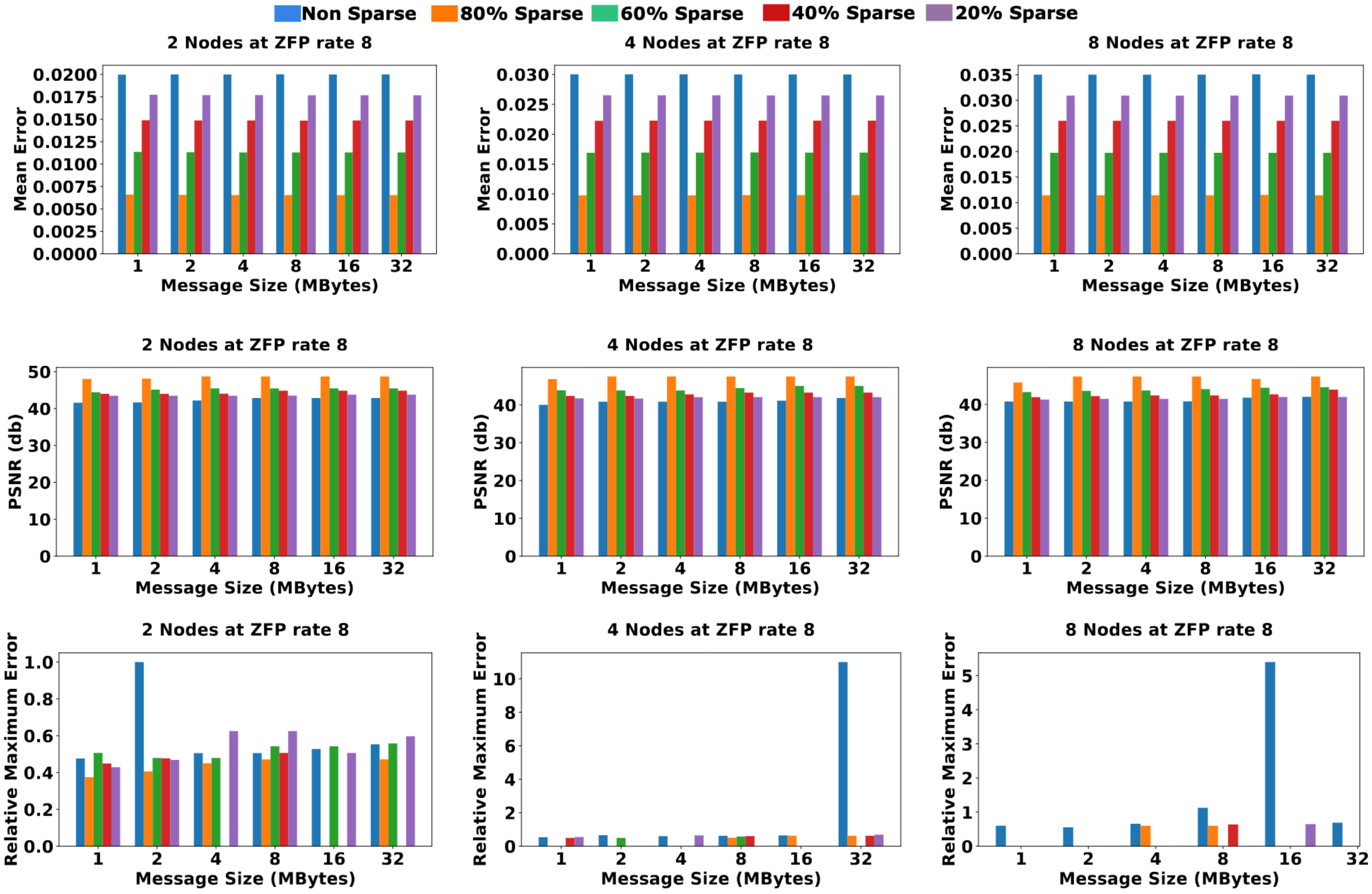
Results: PSNR



Results: Relative Max Error



Results: Sparse



Conclusion

Extended OMB for Error Detection

- Users can now gauge compression error
- Enabled testing with random distributions, sparsity, or custom user-provided data files.
- Showed how data content directly affects ZFP compression error.

Future Work

- Expand MPI Operations
- Compare Compressors (e.g., ZFP, SZ).
- Characterize real application data to create "best practice" recommendations for compression schemes.