

Compression Research: Involvement



- Development of algorithms for lossless compression
 - ▶ MAFISC: suite of preconditioners for HDF5, aims to pack data optimally
Reduced climate/weather data by additional 10-20%, simple filters are sufficient
- Cost-benefit analysis: e.g., for long-term storage MAFISC pays off
- Analysis of compression characteristics for earth-science related data sets
 - ▶ Lossless LZMA yields best ratio but is very slow, LZ4fast outperforms BLOSC
 - ▶ Lossy: GRIB+JPEG2000 vs. MAFISC and proprietary software
- Development of the Scientific Compression Library (SCIL)
 - ▶ Separates concern of data accuracy and choice of algorithms
 - ▶ Users specify necessary accuracy and performance parameters
 - ▶ Metacompression library makes the choice of algorithms
 - ▶ Supports also new algorithms
 - ▶ Ongoing: standardization of useful compression quantities
- A **method for system-wide determination** of data characteristics
 - ▶ Method has been integrated into a script suite to scan data centers