Enhancing geophysical data analysis with Open Source software

BG GROUP

June 2012

C. Jones, J.I. Selvage, J.A. Edgar and H. Macintyre

BG GROUP

Why use open source

- Can help to improve the rate of technology distribution from Academia to Industry
- Open source modular software model allows
  - Outsourcing of coding to any software developer
  - Easier integration with other software as other companies can access the source code.
- Open source software helps to build competitive market of service providers allowing outsourcing of work
- Open source software aids ‘Reproducible Research’

Requirements to allow geoscientists to experiment

BG GROUP

- Software with flexibility
  - To allow geoscientists to be able to experiment with their ideas without significant data replication
- Scalable software
  - To work on realistically sized datasets without additional development time
- Software with high level language access
  - To remove the barrier of complex programming languages, for many geoscientists application development is not part of their job
  - e.g. macros or graphical flow tools
- These requirements do not imply open source software but many open source projects adopt agile development schemes which then reflect in their capabilities

Example software that meets our requirements to allow geoscientists to experiment

BG GROUP

- Software with flexibility
  - Attribute engine in OpendTect, flexibility that starts to be a toolbox for seismic data analysis
- Scalable software
  - Needs scalable hardware and data access e.g. Lustre and SLURM
  - OpendTect, ability to immediately scale solutions to realistic datasets
- Software with high level language access
  - Can be direct with visual programming tools e.g. Scratch, App Inventor Edu, attribute editor in OpendTect and any seismic processing system
  - Indirect using a high level language e.g. a Perl script to manipulate another application

BG GROUP

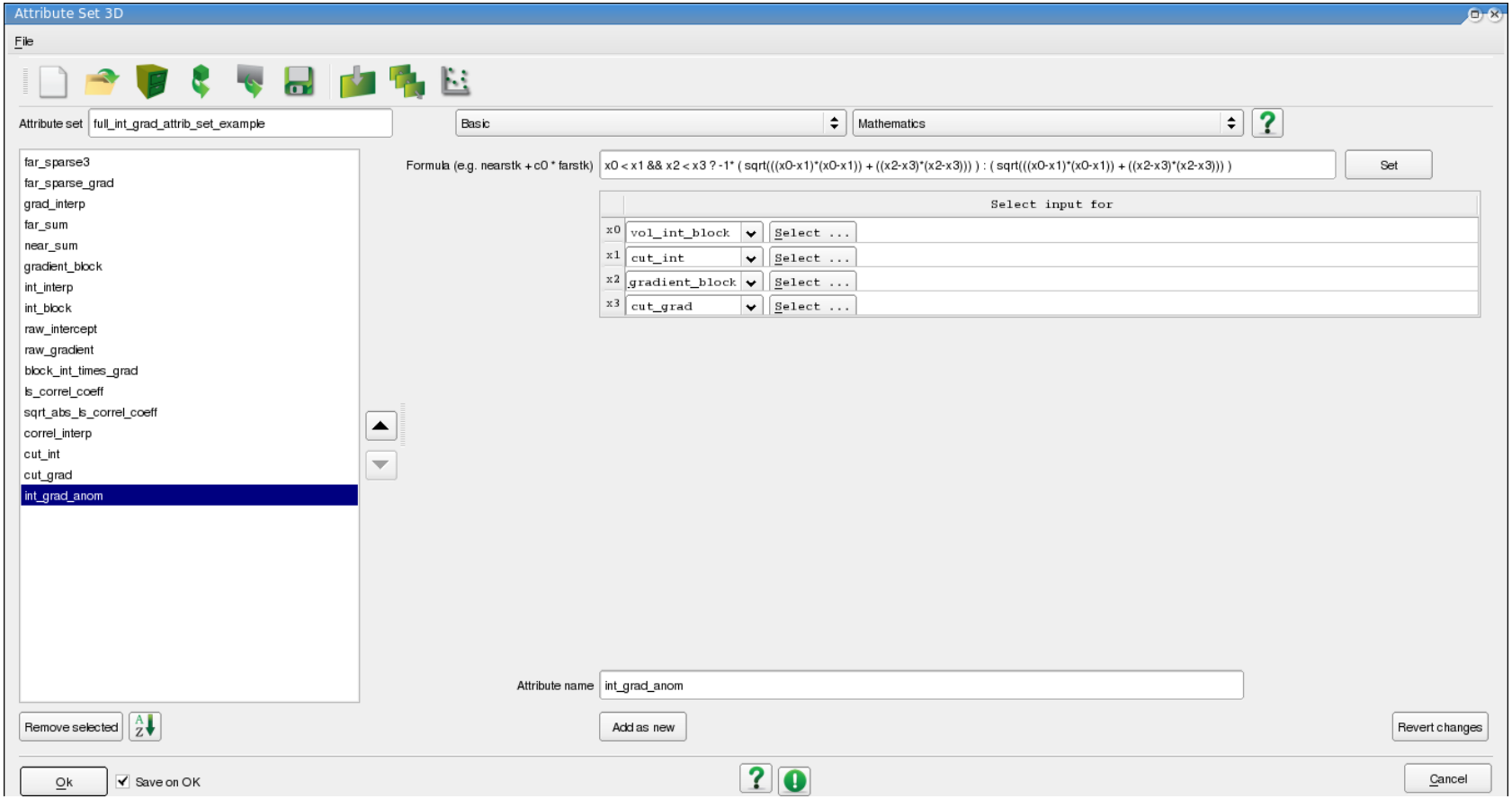
Examples of experimenting using OpendTect

- Use OpendTect as a volumetric analysis tool
- The attributes and the attribute engine form a 3D seismic analysis toolkit
- Built many complex attribute flows to perform a variety of analysis on both pre and post stack seismic datasets
- Also integrated academic code into OpendTect
  - very successful – Wang Multichannel matching pursuit
  - less successful – Madagascar GUI plug-in, focus on reproducible research

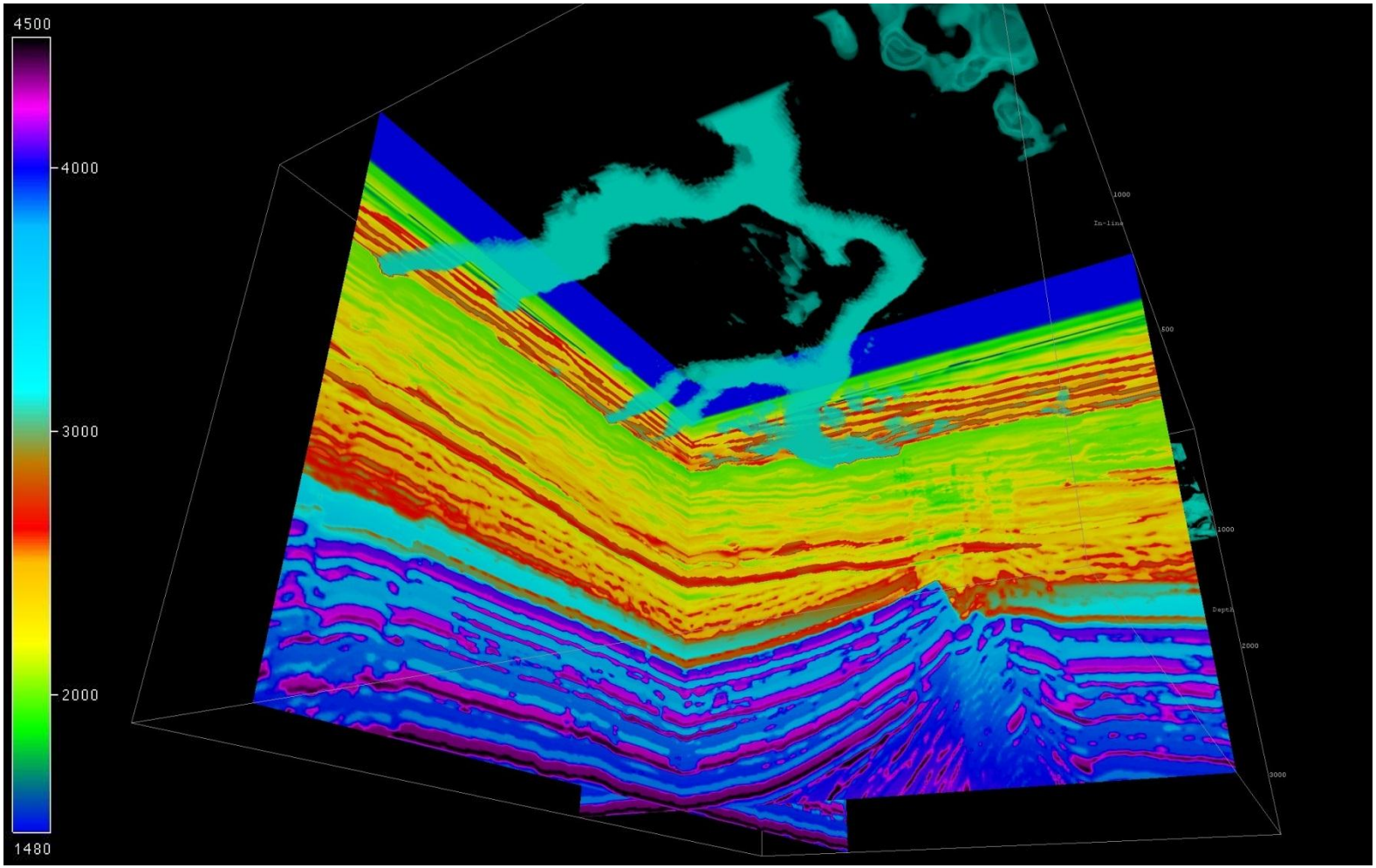
Wang, Y., 2010, Multichannel matching pursuit for seismic trace decomposition, Geophysics 75, V61, DOI:10.1190/1.3462015

Flow editor in OpendTect (high level language access)

- Easy to build complex inter-dependent data flows from multiple volumes and dimensions

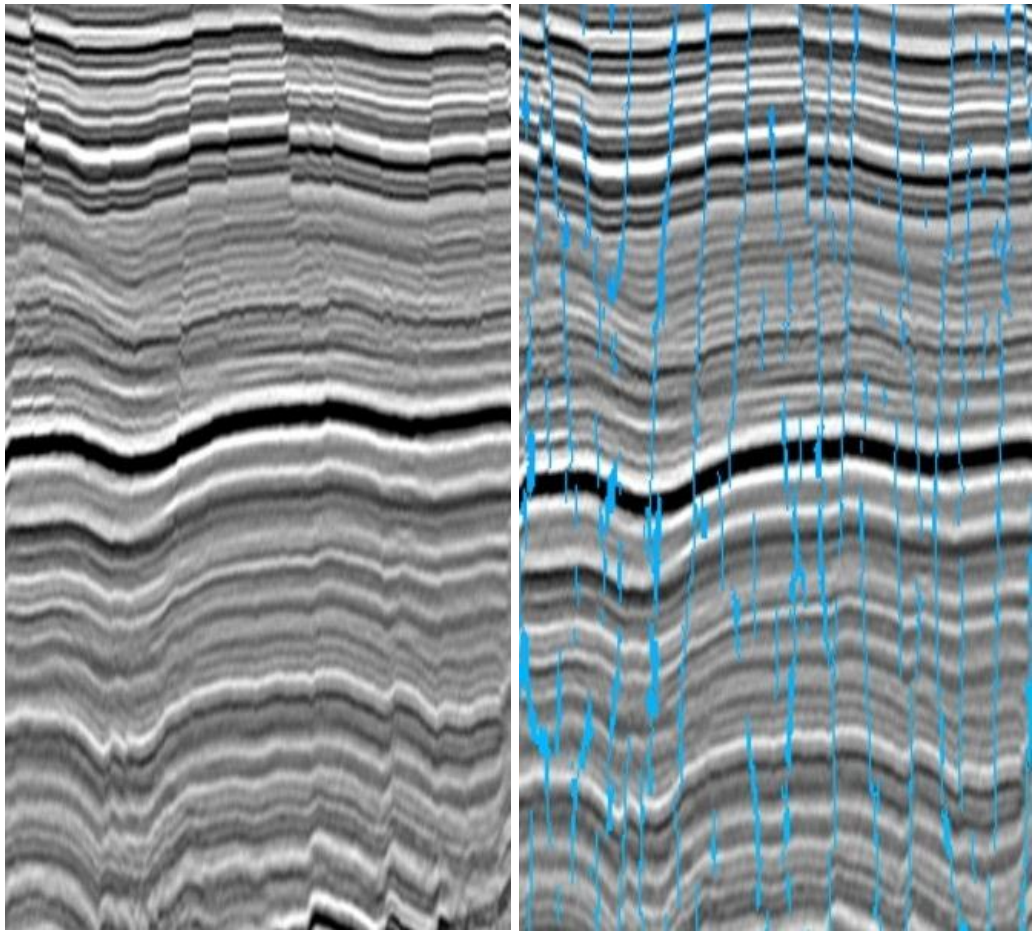


Synthetic velocity model for testing FWI



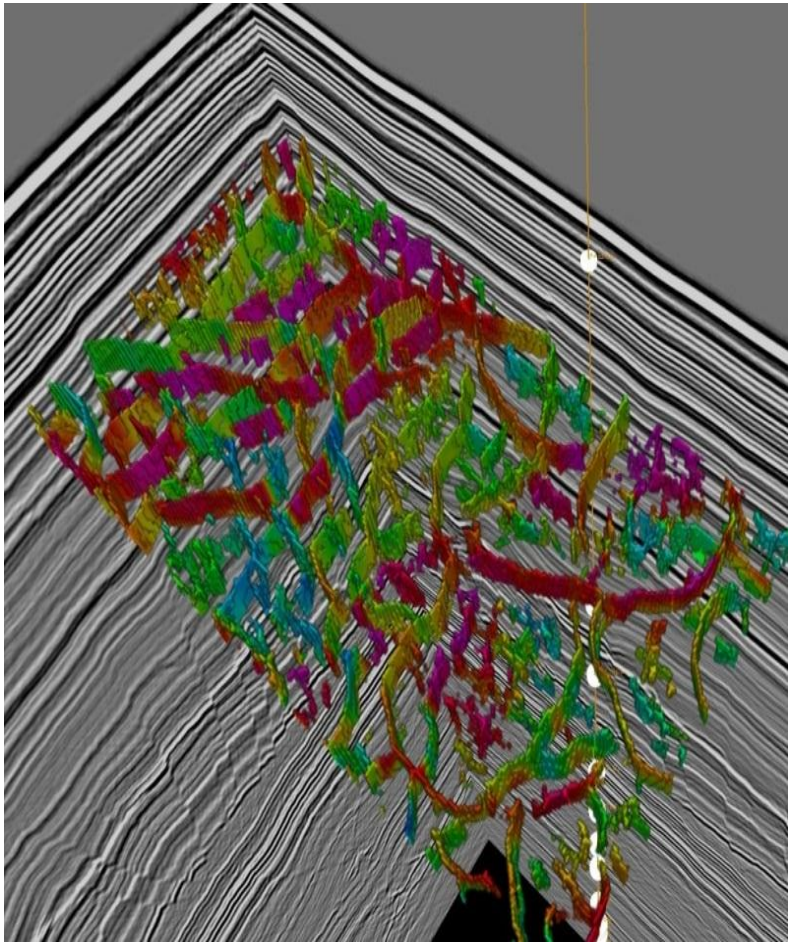
Attributes used to extract faults volume for geohazard analysis

BG GROUP

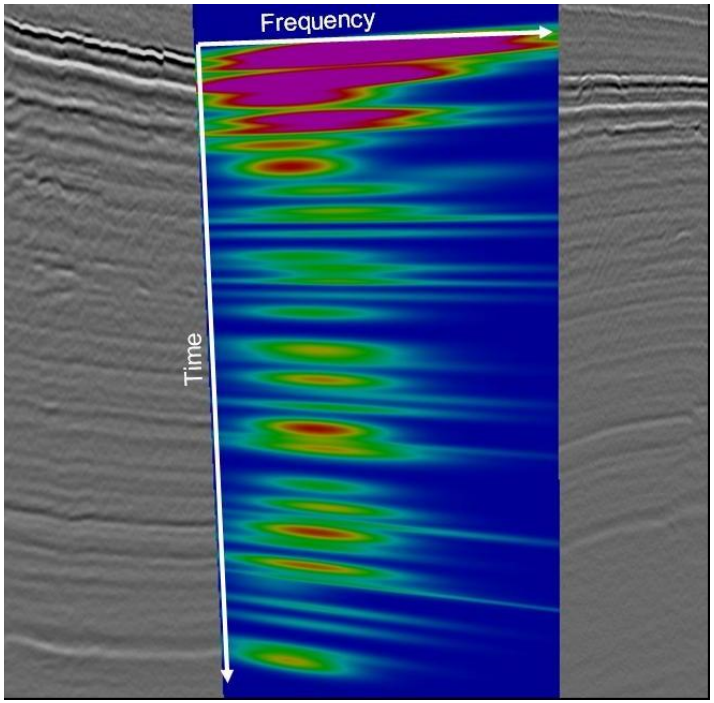


With and without automatically extracted faults

Automatic tracked faults coloured by azimuth

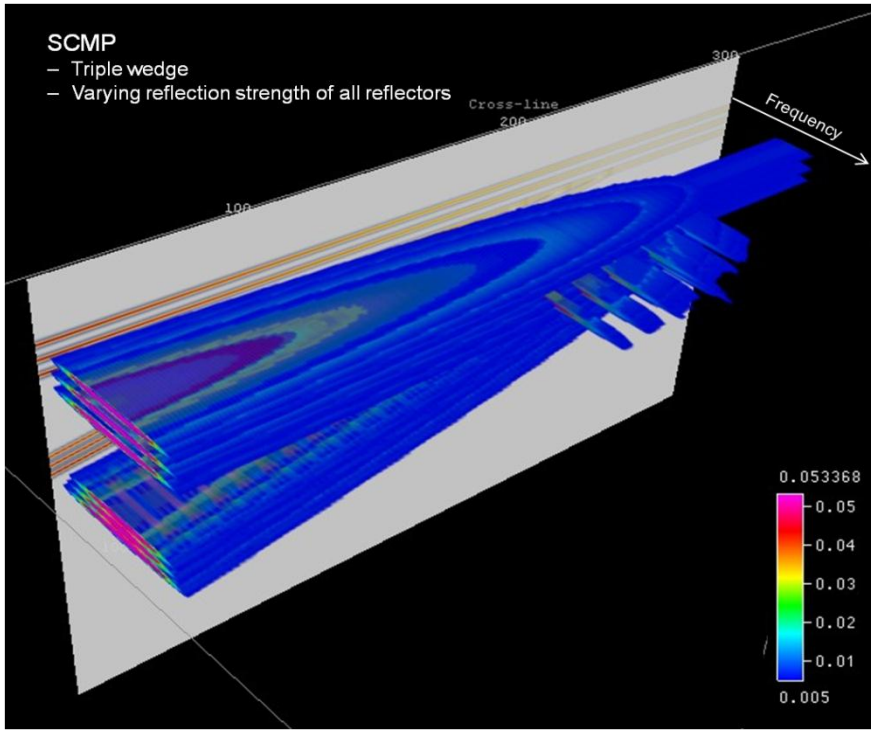


Matching pursuit spectral decomposition



Viewing spectral decomposition data using a pre-stack seismic view

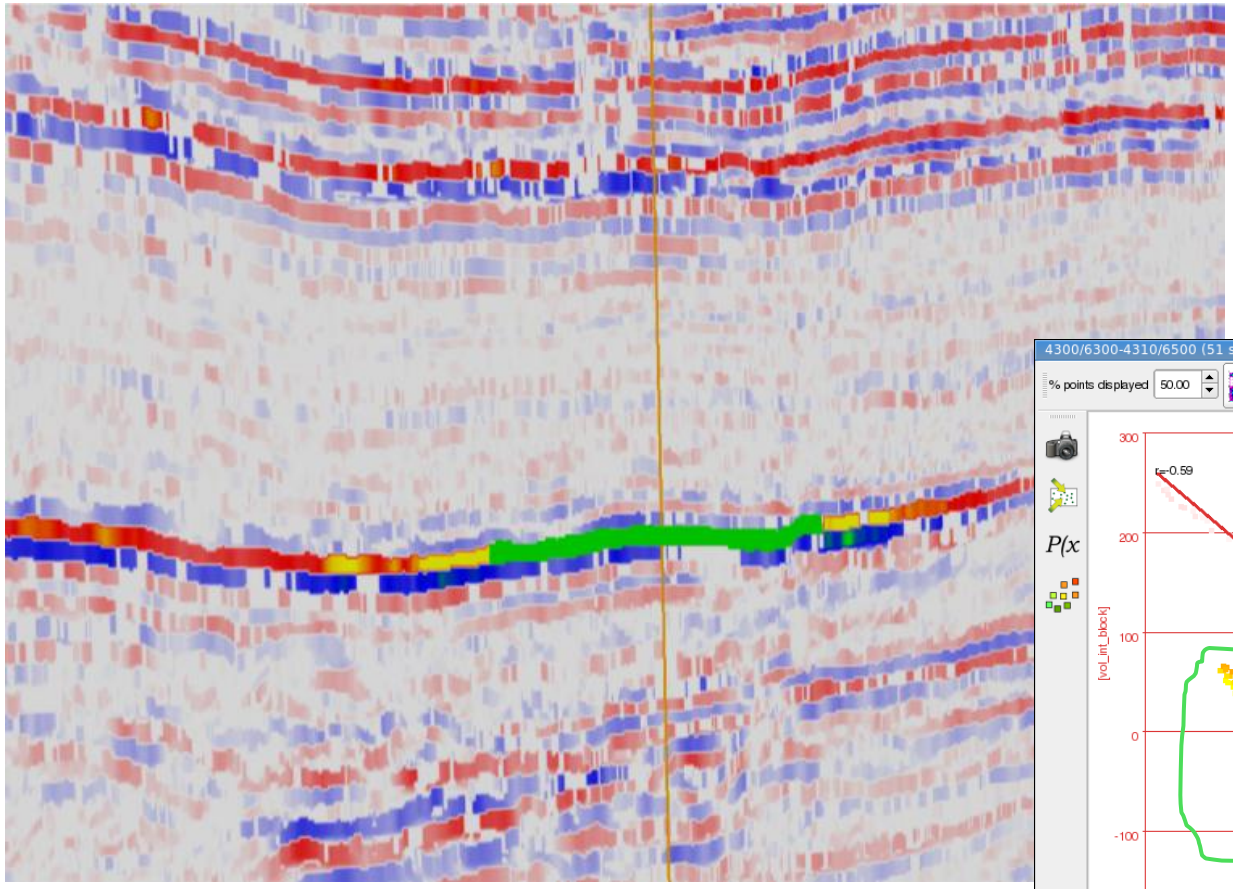
3D X, Y and Frequency of matching pursuit frequency decomposition of a wedge model



Code developed by Yanghua Wang (2010) was ported to OpendTect plug-in by Helene Huck of dGB. Plugin is freely available to Centre for Reservoir Geophysics sponsors

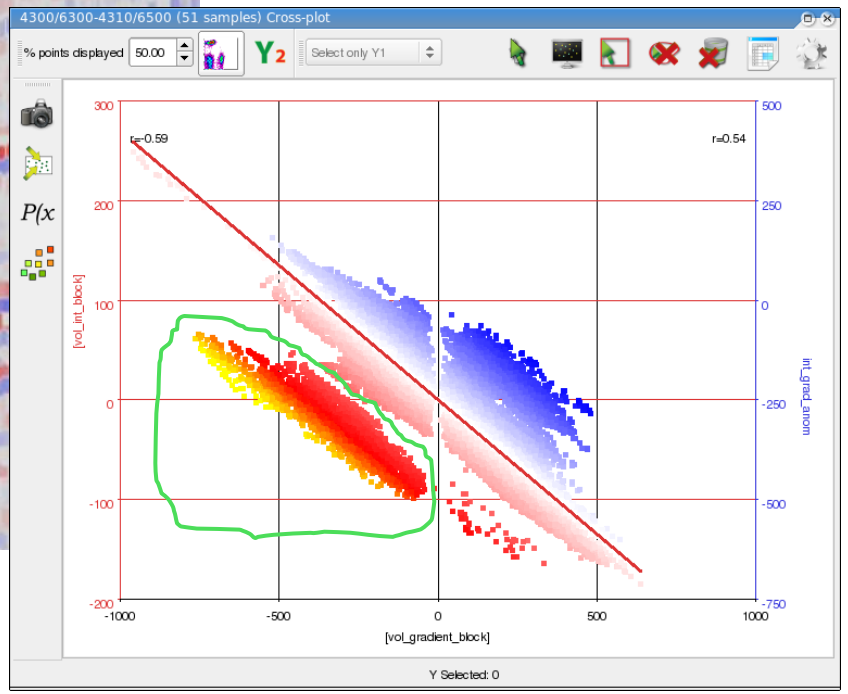
Cross plotting pre-stack attributes

BG GROUP



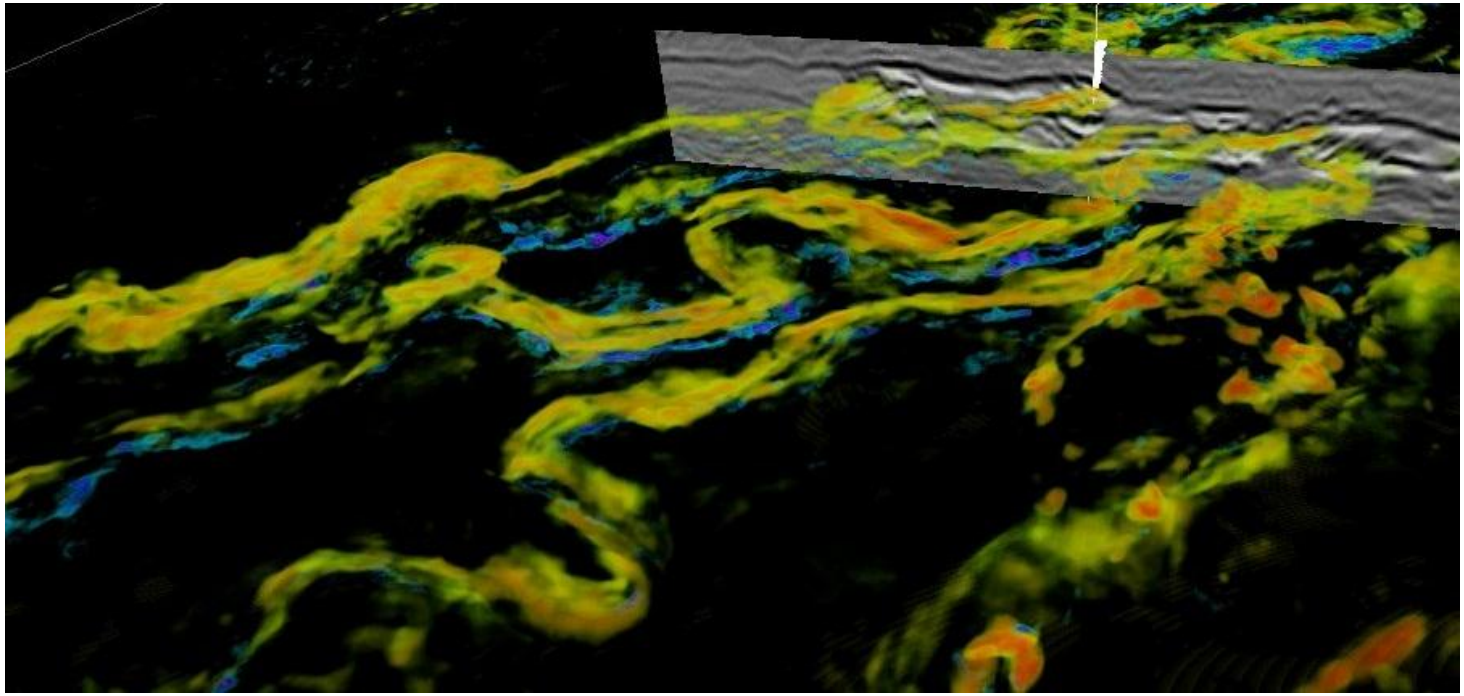
Blocked intercept and gradient cross-plotted

Attribute made to be distance away from the background trend line

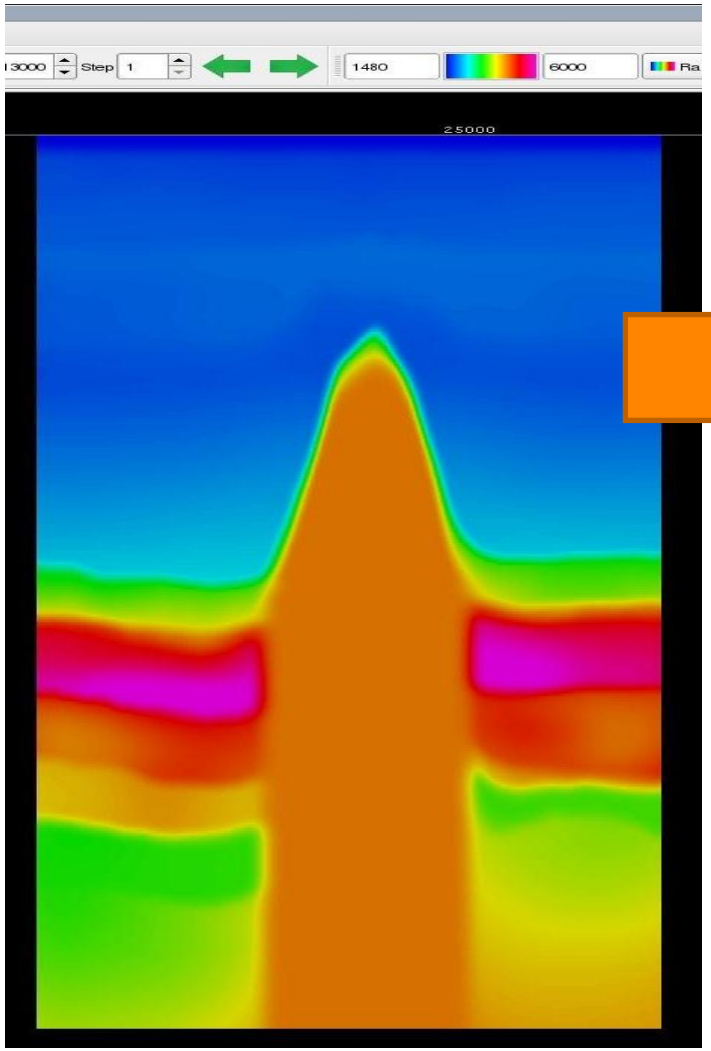


Bayesian classification

- Generation of 3D facies probability volumes from seismic elastic properties using facies probability density functions created from seismic elastic properties at well locations
- Multiple a priori models built from volume attributes with horizon defined attribute manipulations



Velocity model manipulation

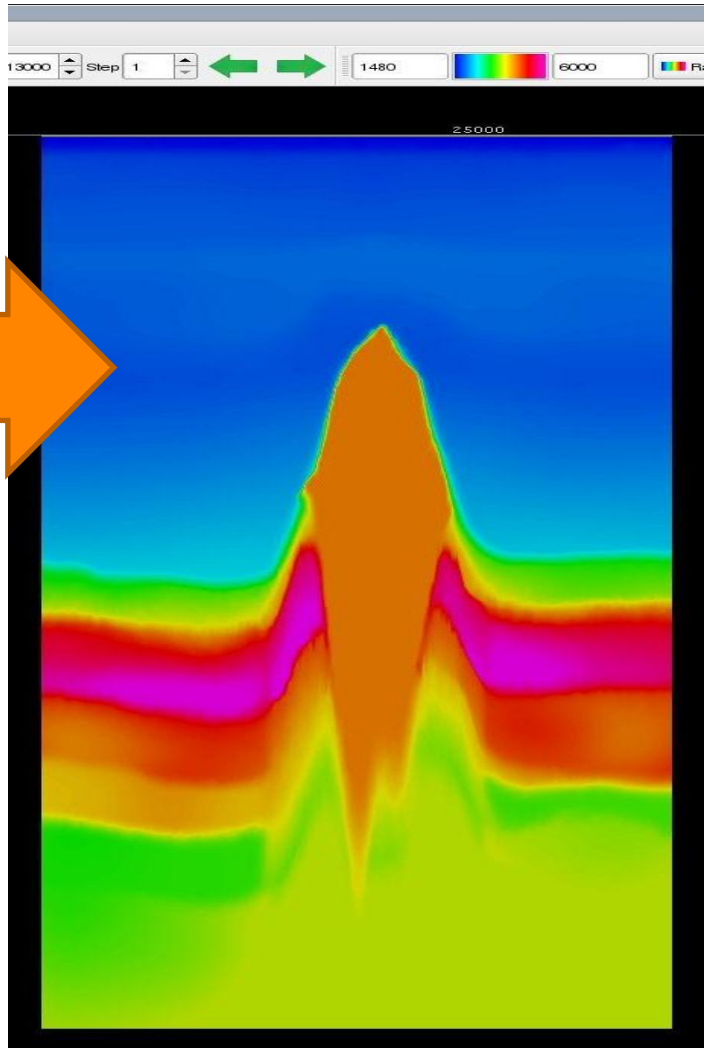


Input:

- Velocity model
- Seismic horizons

Attributes:

- Volume Builder
- Created Maths attribute for interpolating velocities up to new salt side



Integration with other software

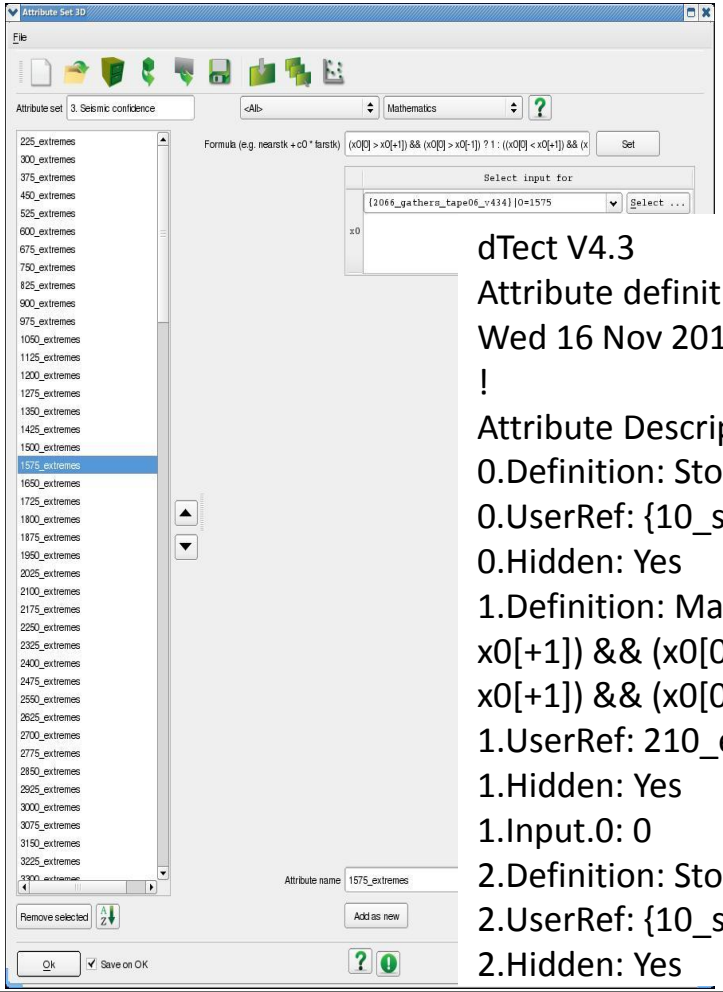
- Use several pieces of open source software together to improve our analysis of the geophysical data
  - GMT for many mapping needs
  - SLURM which is a resource manager for clusters
  - OpendTect for seismic data analysis, it is used in conjunction with SLURM
  - Lustre which is a parallel file system to provide high speed data transfer to the cluster.
  - Perl to script large repetitive OpendTect attribute files and to write command driver scripts for OpendTect

Perl and OpendTect

- Open format definitions (e.g. OpendTect .attr files) provide flexibility

Example piece of perl code to generate tens of seismic confidence attributes and merge them together

```
#!/usr/bin/perl
 # User needs to enter values here
# $attrDir =
"/Attribs/gather_confidence_perl.attr";
$namePreStkGath = "10_statics_3dcmp";
$objectIDPreStkGath = 3; #need to check this
$startOffset = 210;
$endOffset = 5310;
$incOffset = 100;
 open (MYFILE,
'>>/Attribs/gather_confidence_perl_final.attr');
 # DO NOT EDIT BELOW HERE
$ngathers = ($endOffset-
$startOffset)/$incOffset + 1;
$objectIDattrgath = $objectIDPreStkGath+1; #
CHECK THIS OpendTect seems to assign an
object id to the gathers that is 1 +
$objectIDPreStkGath;
```

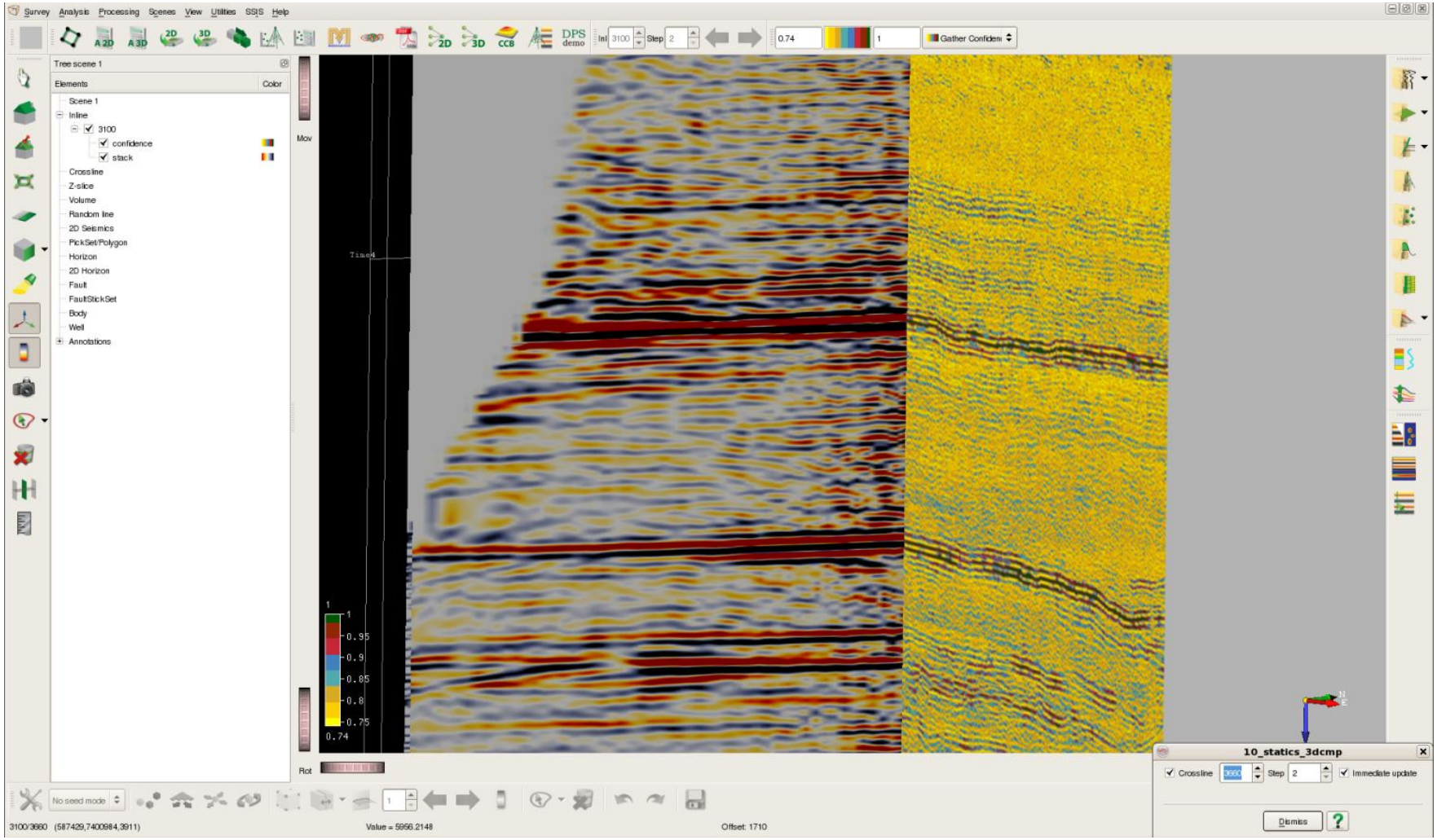


Example produced single attribute file

```
dTect V4.3
Attribute definitions
Wed 16 Nov 2011, 15:29:07
!
Attribute Descriptions
0.Definition: Storage id=100010.4 output=0
0.UserRef: {10_statics_3dcmp}|O=210
0.Hidden: Yes
1.Definition: Math expression="(x0[0] >
x0[+1]) && (x0[0] > x0[-1]) ? 1 : ((x0[0] <
x0[+1]) && (x0[0] < x0[-1]) ? 1 : 0)" output=0
1.UserRef: 210_extremes
1.Hidden: Yes
1.Input.0: 0
2.Definition: Storage id=100010.4 output=1
2.UserRef: {10_statics_3dcmp}|O=310
2.Hidden: Yes
```

BG GROUP

Seismic confidence attribute (proxy for gather flatness)

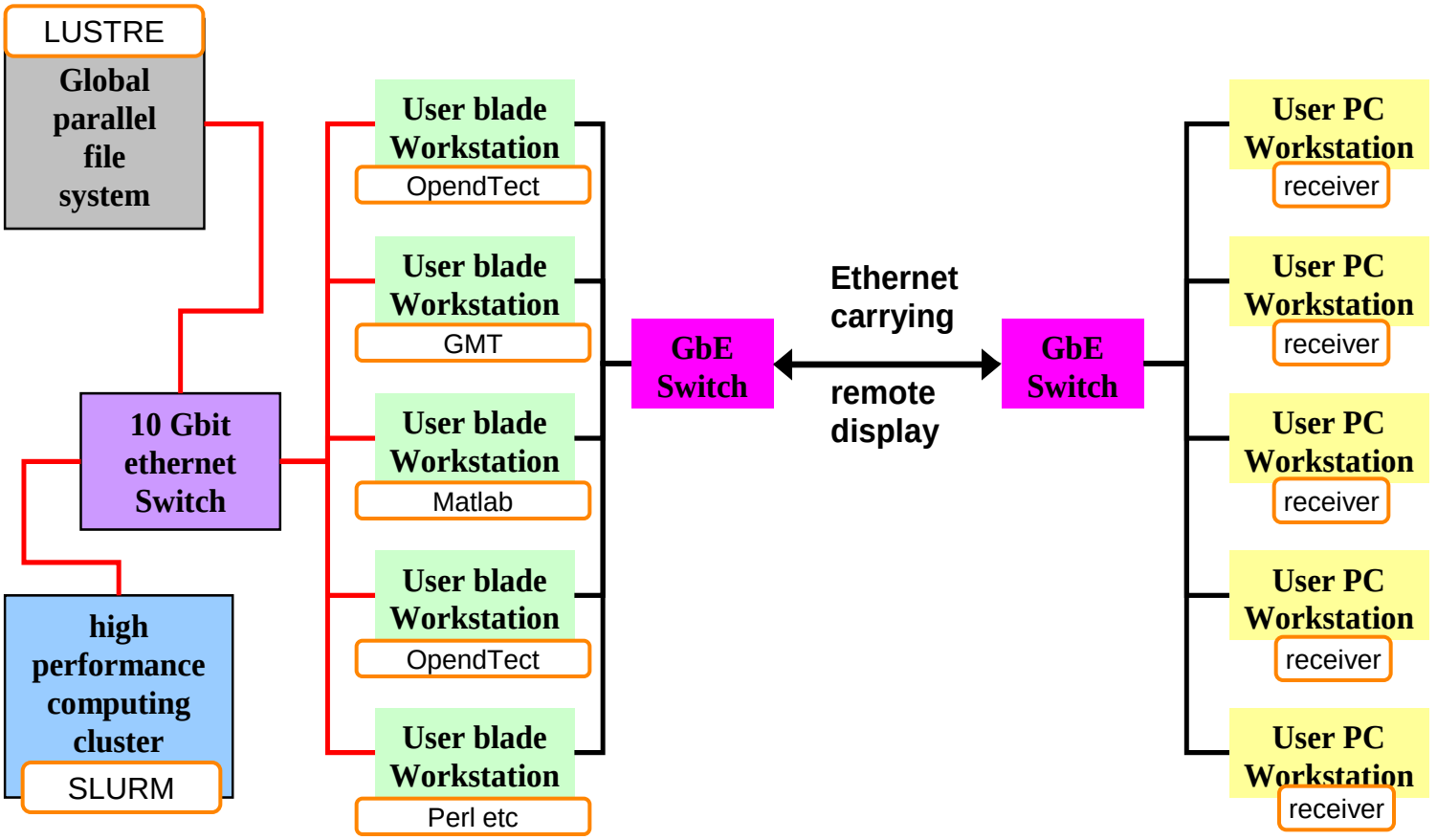


Matlab and SLURM

- Speed-up embarrassingly parallel problems which have been written in Matlab
  - Inversion codes written in Matlab
  - Segy IO library written in Matlab
  - Job control mechanism written in Matlab
  - Matlab compiler is used to produce executables
  - Executables are distributed by SLURM across a cluster
  - Results are then visualised in OpendTect

Integration of open source software to process large datasets

BG GROUP



Future directions

- Coalescing with ‘Big data’ technologies, Hadoop etc
  - Similar volumes of data
  - Need seismic shape decomposition algorithms
  - Seismic processing examples already e.g. SU and Madagascar
  - Offers ability to scale to work on realistic sized datasets without additional development time
- Need an open source seismic storage library to replace segy
  - e.g. Javaseis

Conclusions

- Not all items in all the examples are open source but open source software was required in all, so there is potential for service providers to be built on open source solutions
- The easier the toolkit to use, the more people use it
- A complex toolkit can have multiple interfaces to accommodate different users available time
- Flexible toolkits allow for more applications than originally anticipated
- Need more reusable seismic code libraries

BG GROUP

Acknowledgments

- OpendTect is developed by dGB Earth Sciences
- SLURM is developed at Lawrence Livermore National Laboratory
- GMT (Wessel and Smith, 1991) is supported by the National Science Foundation
- Seismic Unix (SU) is supported by the Center for Wave Phenomena at the Colorado School of Mines and maintained by John Stockwell
- Madagascar was started by Sergey Fomel
- Lustre™ is released by Oracle
- Matlab© 1984-2012- The MathWorks, Inc.
- Scratch and App Inventor Edu from MIT
- Perl was invented by Larry Wall

Wessel, P., and W. H. F. Smith, 1991, Free software helps map and display data, EOS Trans. Amer. Geophys. U., vol. 72 (41), pp. 441, 445-446.

BG GROUP