Graph Processing Toolkit (gpt)

SeaDAS Dev Team

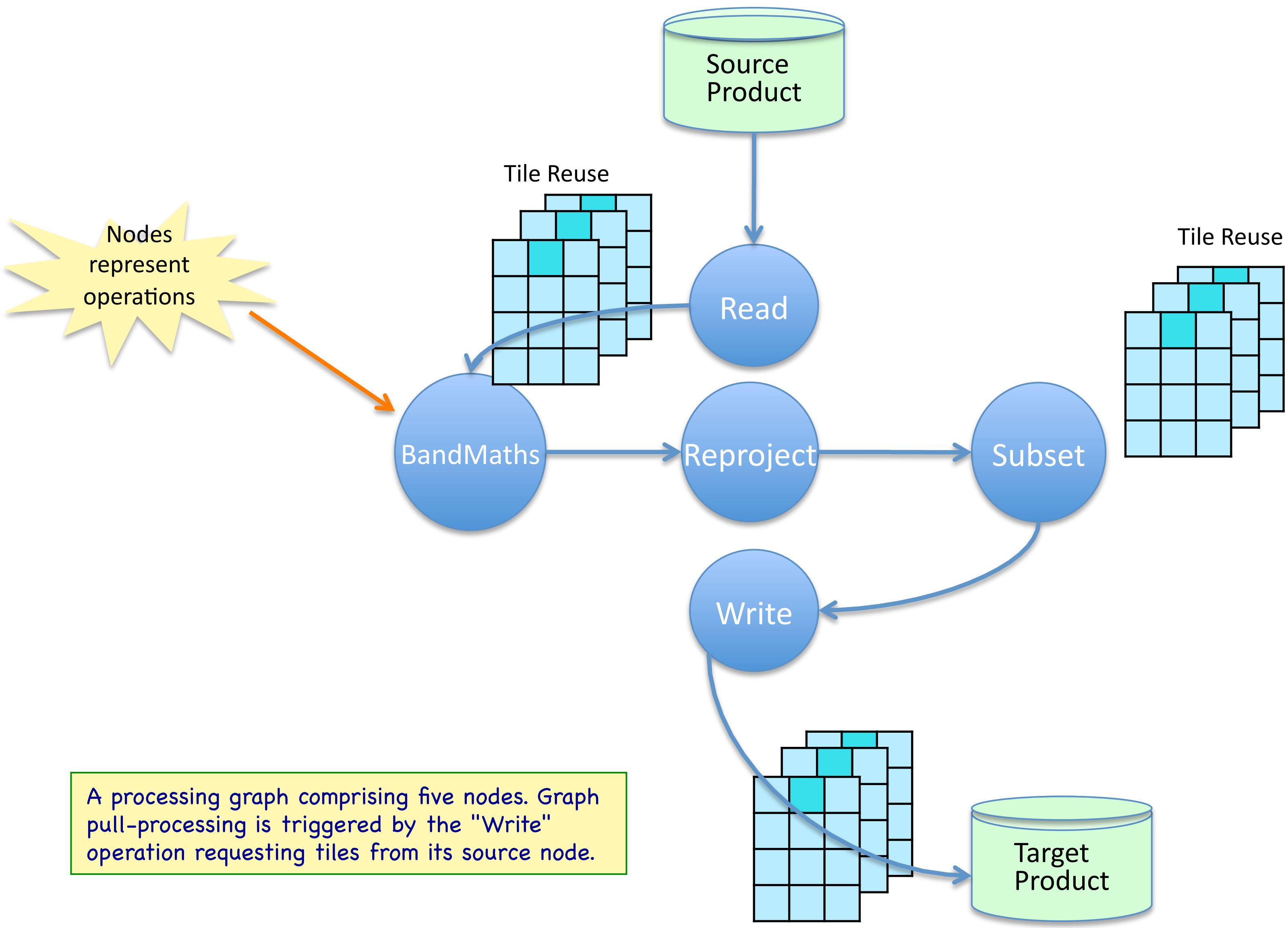
Agenda

- What is gpt?
- How does it work?
- Use case

Graph Processing Framework

- GPF – Graph Processing Framework
  - Allows to construct directed, acyclic graphs (DAG) of processing nodes
  - A node in the graph refers to a data processor or operator, such as Read, BandMaths, Collocate, etc
  - Pull processing, each node pulls at its source node first in order to perform the algorithm it implements
  - The actual processing of a graph is triggered by requesting samples from one of its nodes, usually the final node in the DAG.
- GPF Operators can be invoked in two ways:
  - command-line using the GPF Graph Processing Tool (gpt), located in SeaDAS bin directory
  - dedicated user interfaces in SeaDAS application.

Graph Processing Framework



A processing graph comprising five nodes. Graph pull-processing is triggered by the "Write" operation requesting tiles from its source node.

The BEAM Graph Processing Tool (gpt)

- *gpt* is used to execute BEAM raster data operators in batch-mode from command line.
  - operators can be used stand-alone or combined as a directed acyclic graph (DAG).
  - Processing graphs are represented using XML.
- Usage:
  - gpt <op>|<graph-file> [options] [<source-file-1> <source-file-2> ...]

gpt Operators

- BandMaths - Creates a product with one or more bands using mathematical expressions
- Bathymetry - Creates a bathymetry band, elevation band, topography band and bathymetry mask
- Binning - Performs spatial and temporal aggregation of pixel values into cells ('bins') of a planetary grid
- Collocate - Collocates two products based on their geo-coding
- EMClusterAnalysis - Performs an expectation-maximization (EM) cluster analysis
- KMeansClusterAnalysis - Performs a K-Means cluster analysis
- LandWaterMask - Creates a single band target product for a land/water-mask, using SRTM-shapefiles [60° N, 60° S] and the GlobCover world map (above 60° N)
- Merge - Copies raster data from a number of source products to a specified 'master' product
- Meris.N1Patcher - Copies an existing N1 file and replaces the data for the radiance bands

gpt Operators (contn'd)

- Mosaic - Creates a mosaic out of a set of source products.
- PCA - Performs a Principle Component Analysis.
- PixEx - Extracts pixels from given locations and source products.
- Read - Reads a product from disk.
- Reproject - Reprojects a source product to a target Coordinate Reference System.
- StatisticsOp - Computes statistics for an arbitrary number of source products.
- Subset - Creates a spatial and/or spectral subset of a data product.
- TemporalPercentile - Computes percentiles over a given time period.
- Unmix - Performs a linear spectral unmixing.
- Write - Writes a data product to a file.

gpt command line execution

- The gpt is located in $SEADAS_INSTALL_DIR$/bin
- Help - *gpt.command –h (on mac)*

  - *gpt.sh –h (on linux)*
- Help for a particular operator - *gpt $OperatorName$ -h.*

Sample gpt Operator

```
gs616-seadas1:Sergio aabduraz$ gpt -h Subset
Usage:
  gpt Subset [options]

Description:
  Create a spatial and/or spectral subset of a data product.


Source Options:
  -Ssource=<file>    The source product to create the subset from.
                     This is a mandatory source.

Parameter Options:
  -PbandNames=<string,string,string,...>          The comma-separated list of names of bands to be copied.
                                                  If not given, all bands are copied.
  -PcopyMetadata=<boolean>                        Whether to copy the metadata of the source product.
                                                  Default value is 'false'.
  -PfullSwath=<boolean>                           Forces the operator to extend the subset region to the full swath.
                                                  Default value is 'false'.
  -PgeoRegion=<geometry>                          The subset region in geographical coordinates using WKT-format,
                                                  e.g. POLYGON((<lon1> <lat1>, <lon2> <lat2>, ..., <lon1> <lat1>))
                                                  (make sure to quote the option due to spaces in <geometry>).
                                                  If not given, the entire scene is used.
  -PsubSamplingX=<int>                            The pixel sub-sampling step in X (horizontal image direction)
                                                  Default value is '1'.
  -PsubSamplingY=<int>                            The pixel sub-sampling step in Y (vertical image direction)
                                                  Default value is '1'.
  -PtiePointGridNames=<string,string,string,...>  The comma-separated list of names of tie-point grids to be copied.
                                                  If not given, all bands are copied.
```

Sample gpt Operator Configuration in XML

```
<node id="subsetNode">
    <operator>Subset</operator>
    <sources>
        <source>${source}</source>
    </sources>
    <parameters>
        <geoRegion>POLYGON((-77.5 40, -77.5 35, -72.5 35, -72.5 40,   -77.5 40))</geoRegion>
        <bandNames>chlor_a, Rrs_443 </bandNames>
        <copyMetadata>true</copyMetadata>
    </parameters>
</node>
```

Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
   1. Expression: !mask ? product: NaN
   2. Name new products product_mask
4. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
5. Reproject the two cropped files and all its products using Mercator 1 SP projection
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file --> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

Steps in SeaDAS Application

Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152

Open File

Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3

BandMaths

Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)

- Expression: !mask ? product: NaN
- Name new products product_mask

BandMaths

Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)

Subset

Reproject the two cropped files and all its products using Mercator 1 SP projection

Reproject

Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7

Collocate

Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim

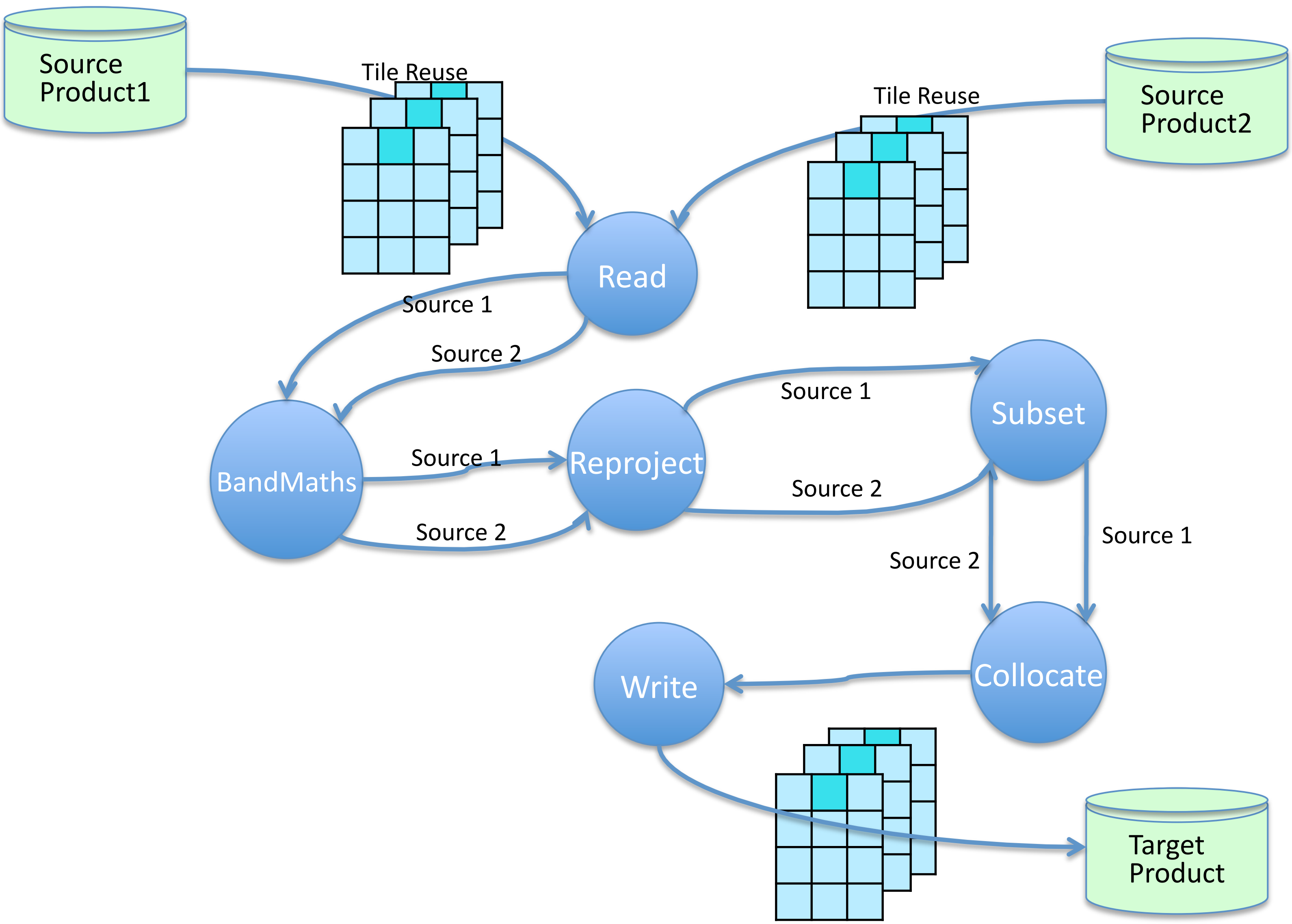
Save

Use raster export to create a netCDF file ---> MODIS_SeaWiFS_Collocated_Mask.nc

Raster Export

Delete all intermediate files

Graph Processing in Batch



Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
    1. Expression: !mask ? product: NaN
    2. Name new products product_mask
4. Reproject the two cropped files and all its products using Mercator 1 SP projection
5. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file –> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

BandMaths Operator

```
<node id="bandMathsNode">
    <operator>BandMaths</operator>
    <sources>
        <sourceProducts>${source}</sourceProducts>
    </sources>
    <parameters>
        <targetBands>
            <targetBand>
                <name>chlor_a</name>
                <expression>!MY_MASK ? chlor_a : NaN</expression>
                <description>chlor_a</description>
                <type>float32</type>
                <validExpression>chlor_a &gt; 0.001 AND chlor_a &lt; 100</validExpression>
                <noDataValue>-32767.0</noDataValue>
                <spectralBandIndex>0</spectralBandIndex>
            </targetBand>
        </targetBands>
        <variables>
            <variable>
                <name>MY_MASK</name>
                <type>boolean</type>
                <value>(l2_flags.HISATZEN ||l2_flags.STRAYLIGHT || l2_flags.CLDICE || l2_flags.MAXAERITER ||
                    l2_flags.MODGLINT)
                </value>
            </variable>
        </variables>
    </parameters>
</node>
```

Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
    1. Expression: !mask ? product: NaN
    2. Name new products product_mask
4. Reproject the two cropped files and all its products using Mercator 1 SP projection
5. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file —> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

Reproject Operator

```xml
<node id="reprojNode">
    <operator>Reproject</operator>
    <sources>
        <source>bandMathsNode</source>
    </sources>
    <parameters>
        <crs>PROJCS["Mercator_1SP / World Geodetic System 1984",
            GEOGCS["World Geodetic System 1984",
            DATUM["World Geodetic System 1984",
            SPHEROID["WGS 84", 6378137.0, 298.257223563, AUTHORITY["EPSG","7030"]],
            AUTHORITY["EPSG","6326"]],
            PRIMEM["Greenwich", 0.0, AUTHORITY["EPSG","8901"]],
            UNIT["degree", 0.017453292519943295],
            AXIS["Geodetic longitude", EAST],
            AXIS["Geodetic latitude", NORTH]],
            PROJECTION["Mercator_1SP"],
            PARAMETER["latitude_of_origin", 0.0],
            PARAMETER["central_meridian", 0.0],
            PARAMETER["scale_factor", 1.0],
            PARAMETER["false_easting", 0.0],
            PARAMETER["false_northing", 0.0],
            UNIT["m", 1.0],
            AXIS["Easting", EAST],
            AXIS["Northing", NORTH]]
        </crs>
        <resampling>Nearest</resampling>
        <orthorectify>false</orthorectify>
        <noDataValue>NaN</noDataValue>
        <includeTiePointGrids>true</includeTiePointGrids>
        <addDeltaBands>false</addDeltaBands>
    </parameters>
</node>
```

Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
   1. Expression: !mask ? product: NaN
   2. Name new products product_mask
4. Reproject the two cropped files and all its products using Mercator 1 SP projection
5. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file --> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

Subset Operator Sample Configuration

```xml
<node id="subsetNode">
    <operator>Subset</operator>
    <sources>
        <source>reprojNode</source>
    </sources>
    <parameters>
        <geoRegion>POLYGON((-77.5 40, -77.5 35, -72.5 35, -72.5 40, -77.5 40))</geoRegion>
        <bandNames>chlor_a,ag_412_mlrc,Rrs_443,Rrs_547,Rrs_555</bandNames>
        <copyMetadata>true</copyMetadata>
    </parameters>
</node>
```

Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
   1. Expression: !mask ? product: NaN
   2. Name new products product_mask
4. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
5. Reproject the two cropped files and all its products using Mercator 1 SP projection
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file —> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

Collocation Operator Sample Configuration

```
<node id="collocateNode">
    <operator>Collocate</operator>
    <sources>
      <master>${master}</master>
      <slave>${slave}</slave>
    </sources>
    <parameters>
      <targetProductName></targetProductName>
      <targetProductType>BEAM-DIMAP</targetProductType>
      <renameMasterComponents>true</renameMasterComponents>
      <renameSlaveComponents>true</renameSlaveComponents>
      <masterComponentPattern>${ORIGINAL_NAME}_R</masterComponentPattern>
      <slaveComponentPattern>${ORIGINAL_NAME}_D</slaveComponentPattern>
      <resamplingType>NEAREST_NEIGHBOUR</resamplingType>
    </parameters>
  </node>
```

Use Case

1. Load the two L2 files on SeaDAS: A2006132174000.L2_CHL_ag_LAC and S2006132174152
2. Create a mask: mask:l2_flags.HISATZEN or l2_flags.STRAYLIGHT or l2_flags.CLDICE or l2_flags.MAXAERITER or l2_flags.MODGLINT3
3. Apply the generated mask file to the following products: chlor_a, ag_412_mlrc, Rrs_443, Rrs_547 (MODIS only)
    1. Expression: !mask ? product: NaN
    2. Name new products product_mask
4. Crop (subset) each file with coordinates: 35N to 40N and -77.5W to -72.5W5)
5. Reproject the two cropped files and all its products using Mercator 1 SP projection
6. Collocate the two files into a single file. The products will be now named chlor_a_mask_R, ... for MODIS and chlor_a_mask_D, ... for SeaWiFS7
7. Save the collocated file with the masked and collocated products to: MODIS_SeaWiFS_Collocated_Mask.dim
8. Use raster export to create a netCDF file –> MODIS_SeaWiFS_Collocated_Mask.nc
9. Delete all intermediate files

Write Operator Sample Configuration

```xml
<node id="writeNode">
    <operator>Write</operator>
    <sources>
        <source>collocateNode</source>
    </sources>
    <parameters>
        <file>/Users/Shared/Tutorial/gpt/MODIS_SEAWIFS_Collocated_Mask.nc</file>
        <formatName>NetCDF-CF</formatName>
        <deleteOutputOnFailure>true</deleteOutputOnFailure>
        <writeEntireTileRows>true</writeEntireTileRows>
        <clearCacheAfterRowWrite>true</clearCacheAfterRowWrite>
    </parameters>
</node>
```

Putting It All Together

- gpt_batch.xml
- gpt_batch.properties
- files.tx
- process.bash

```
while read -r m s f; do
gpt gpt_batch.xml -p gpt_batch.properties -SsourceModis=$m -SsourceSeawifs=$s -PtargetFileName=$f
done < files.txt
```