PigErrorHandlingFunctionalSpecification

This document describes the functional specification for the Error Handling feature in Pig.

Error types and mechanism to handle errors

The #cookbook discusses the classification of errors within Pig and proposes a guideline for exceptions that are to be used by developers. A reclassification of the errors is presented below.

Frontend errors

. The front-end consists of multiple components - parser, type checker, optimizer, translators, etc. These errors usually occur at the client side before the execution begins in Hadoop. All the errors from these components can be categorized as front-end errors. Components that are part of the front end will throw specific exceptions that capture the context. For example, the parser throws a `ParseException`, the type checker will throw a `TypeCheckerException`, the optimizer will throw a `LogicalOptimizerException`, etc. A list of the exceptions thrown in the front-end are as follows.

- `FrontendException` Generic front-end exception
- `JobCreationException` Used for indicating errors during Map Reduce job creation
- `LogicalToPhysicalTranslatorException` Used for indicating errors during logical plan to physical plan translation
- `MRCompilerException` Used for indicating errors during map reduce plan compilation from physical plan
- `OptimizerException` Used for indicating errors during logical plan optimization
- `PigException` Generic exception in Pig
- `PlanException` Used for indicating errors during plan/graph operations
- `PlanValidationException` Used for indicating errors during plan validation
- `SchemaMergeException` Used for indicating errors during schema merges
- `TypeCheckerException` Used for indicating errors due to type checking
- `VisitorException` Used for indicating errors when visiting a plan

Backend errors

. The execution pipeline, the operators that form the pipeline and the map reduce classes fall into the back-end. The errors that occur in the back-end are generally at run-time. Exceptions such as `ExecException` and `RunTimeException` fall into this category. These errors will be propagated to the user facing system and an appropriate error message indicating the source of the error will be displayed.

Internal errors

. Any error that is not reported via an explicit exception is indicative of a bug in the system. Such errors are flagged as internal errors and will be reported as possible bugs. Any non Hadoop backend error is a bug in Pig.

While the aforementioned errors describe a developer's viewpoint of errors, the user is interested in the source of the errors. A classification of the source of errors is given below.

1. User Input - Sources of user input error are syntax error, semantic error, etc.
2. Bug - An internal error in the Pig code and not related to the user's input
3. User Environment - The client side environment
4. Remote Environment - The Hadoop execution environment

Error codes

Error codes are categorized into ranges depending on the nature of the error. The following table indicates the ranges for the error types in Pig. Normally, errors due to user input and bugs in the software are not retriable. Errors due to the user environment and remote environment may be retriable based on the context. Error codes will be used for documentation to help users address common errors.

Error type	Range
User Input	1000 - 1999

Bug	2000 - 2999
User Environment (retriable)	3000 - 3999
User Environment	4000 - 4999
Remote Environment (retriable)	5000 - 5999
Remote Environment	6000 - 6999

Programmatic access via Java APIs can query if exceptions are retriable or not. For external processes that rely on the return code of the process, the table given below will indicate the status of the process execution. Front-end exceptions will result in failures as far the user is concerned. Hadoop's errors are not retriable and return fatal error code (2) or partial failure code (3).

Status	Return Code
Successful Execution	0
Retriable error	1
Fatal error	2
Partial failure	3

Error message

The format of the error message shown to the user will be as follows:

ERROR <Error code>: <Error message>

E.g.: ERROR 1005: The arity of the group by columns do not match.

Requirement on UDF authors

In order to enable warning message aggregation, UDF authors should use Pig's abstraction to handle warning message aggregation. For more details refer to the [#design].

```
attachment:PigLogger.jpg
```

Warning Codes

Each warning message will be associated with a warning code. The warning code will correspond to the enum type used for the warning message aggregation.

Note

1. The exact warning codes along with the warning messages are to be decided.

Additional command line switches

In order to support the ability to turn on/off warning message aggregation, log error messages to client side logs and specify the location of the client side log, the following switches will be added to and/or extended in Pig.

1. -wagg to turn on warning aggregation; by default warning aggregation is turned off.
2. -v to include printing error messages on screen; by default error messages will be written to client side log. Using -v will also print the messages on the screen
3. <<Anchor(clientSideLog)>>-l directory where the client side log is stored; by default, logs will be stored in the current working directory and named pig_<datestamp>.log. When used in batch mode (i.e., when a script to be executed is provided), the log file name will also contain the name of the script - pig_<datestam>_<pig_file>.log.

Compendium of error messages

A list of possible error messages is listed below. This list is not comprehensive and will be modified to reflect the true error message along with the error code.

Error Code	Error Message
1000	Error during parsing
1001	Unable to descirbe schema for alias <alias>
1002	Unable to store alias <id>
1003	Unable to find an operator for alias <alias>
1004	No alias <alias> to <operation>
1005	No plan for <alias> to <operation>
1006	Could not find operator in plan
1007	Found duplicates in schema. <list of duplicate column names> . Please alias the columns with unique names.
1008	Expected a bag with a single element of type tuple but got a bag schema with multiple elements
1009	Expected a bag with a single element of type tuple but got an element of type <type>
1010	getAtomicGroupByType is used only when dealing with atomic <group/join> col
1011	getTupleGroupBySchema is used only when dealing with <tuple/join> group col
1012	Each <COGroup/join> input has to have the same number of inner plans
1013	attributes can either be star ⭐ or a list of expressions, but not both.
1014	Problem with input <operator> of User-defined function: <function>
1015	Error determining fieldschema of constant: <constant>
1016	Problems in merging user defined schema
1017	Schema mismatch. A basic type on flattening cannot have more than one column. User defined schema: <schema>
1018	Problem determining schema during load
1019	Unable to merge schemas
1020	Only a BAG or TUPLE can have schemas. Got <type>
1021	Type mismatch. No useful type for merging. Field Schema: <field schema>. Other Fileld Schema: + otherFs
1022	Type mismatch. Field Schema: <field schema>. Other Fileld Schema: + otherFs
1023	Unable to create field schema
1024	Found duplicate aliases: <alias>
1025	Found more than one match: <list of aliases>
1026	Attempt to fetch field: <field> from schema of size <size>
1027	Cannot reconcile schemas with different sizes. This schema has size <size> other has size of <size>
1028	Access to the tuple <alias> of the bag is disallowed. Only access to the elements of the tuple in the bag is allowed.
1029	One of the schemas is null for merging schemas. Schema: <schema> Other schema: <schema>
1030	Different schema sizes for merging schemas. Schema size: <size> Other schema size: <size>
1031	Incompatible types for merging schemas. Field schema type: <type> Other field schema type: <type>
1032	Incompatible inner schemas for merging schemas. Field schema: <schema> Other field schema: <schema>
1033	Schema size mismatch for merging schemas. Other schema size greater than schema size. Schema: <schema>. Other schema: <schema>
1034	TypeCastInserter invoked with an invalid operator class name: <operator class name>
1035	Error getting LOProject's input schema
1036	Map key should be a basic type
1037	Operand of Regex can be CharArray only
1038	Operands of AND/OR can be boolean only
1039	Incompatible types in <Addition/Subtraction/Division/Multiplication/Mod/GreaterThan/LesserThan/> operator. left hand side: <type> right hand size: type
1040	Could not set <Add/Subtract/Multiply/Divide/Mod/UserFunc/BinCond> field schema
1041	NEG can be used with numbers or Bytearray only
1042	NOT can be used with boolean only
1043	Unable to retrieve field schema of operator.
1044	Unable to get list of overloaded methods.

1045	Could not infer the matching function for <func spec> as multiple or none of them fit. Please use an explicit cast.
1046	Multiple matching functions for <funcspec> with input schemas: (<schema> , <schema>). Please use an explicit cast.
1047	Condition in BinCond must be boolean
1048	Two inputs of BinCond must have compatible schemas
1049	Problem during evaluaton of BinCond output type
1050	Unsupported input type for BinCond: lhs = <type>; rhs = <type>
1051	Cannot cast to bytearray
1052	Cannot cast <type> with schema <schema> to <type> with schema <schema>
1053	Cannot resolve load function to use for casting from <type> to <type>
1054	Cannot merge schemas from inputs of UNION
1055	Problem while reading schemas from inputs of <Union/SplitOutput/Distinct/Limit/Cross>
1056	Problem while casting inputs of Union
1057	's inner plan can only have one output (leaf)
1058	Split's condition must evaluate to boolean. Found: <type>
1059	Problem while reconciling output schema of <Sort/Filter/Split>
1060	Cannot resolve <COGroup/Foreach/Fragment Replicate Join> output schema
1061	Sorry, group by complex types will be supported soon
1062	COGroup by incompatible types
1063	Problem while reading field schema from input while inserting cast
1064	Problem reading column <col> from schema: <schema>
1065	Found more than one load function to use: <list of load functions>
1066	Unable to open iterator for alias <alias>
1067	Unable to explain alias <alias>
1068	Using <Map/Bag> as key not supported.
1069	Problem resolving class version numbers for class <class>
1070	Could not resolve <class> using imports: <package import list>
1071	Cannot convert a <type> to <a/an> <type>
1072	Out of bounds access: Request for field number <number> exceeds tuple size of <size>
1073	Cannot determine field schema for <object>
1074	Problem with formatting. Could not convert <object> to <Integer/Long/Float/Double>.
1075	Received a bytearray from the UDF. Cannot determine how to convert the bytearray to <int/float/long/double/string/tuple/bag/map>
1076	Problem while reading field schema of cast operator.
1077	Two operators that require a cast in between are not adjacent.
1078	Schema size mismatch for casting. Input schema size: <size>. Target schema size: <size>
1079	Undefined type checking logic for unary operator: " <operator>
1080	Did not find inputs for operator: " <operator>
1081	Cannot cast to <int/float/long/double/string/tuple/bag/map>. Expected bytearray but received: <type>
1082	Cogroups with more than 127 inputs not supported.
1083	setBatchOn() must be called first.
1084	Invalid Query: Query is null or of size 0.
1085	operator in <pushBefore/pushAfter> is null. Cannot <pushBefore/pushAfter> null operators.
1086	First operator in <pushBefore/pushAfter> should have multiple <inputs/outputs>. Found first operator with <size> <inputs/outputs>.
1087	The <inputNum/outputNum> <num> should be lesser than the number of <inputs/outputs> of the first operator. Found first operator with <size> <inputs/outputs>.
1088	operator in <pushBefore/pushAfter> should have <at least> one <output/input>. Found <first/second> operator with <no/<size> > <outputs/inputs>.
1089	Second operator in <pushBefore/pushAfter> should be the <successor/predecessor> of the First operator.
1090	Second operator can have at most one <incoming/outgoing> edge from First operator. Found <num> edges.

1091	First operator does not support multiple <outputs/inputs>. On completing the <pushBefore/pushAfter> operation First operator will end up with <num> edges
1092	operator in swap is null. Cannot swap null operators.
1093	Swap supports swap of operators with at most one <input/output>. Found <first/second> operator with <size> <inputs/outputs>
1094	Attempt to insert between two nodes that were not connected.
1095	Attempt to remove and reconnect for node with multiple <predecessors/successors>.
1096	Attempt to remove and reconnect for node with <<size>/no> <predecessors/successors>.
1097	Containing node cannot be null.
1098	Node index cannot be negative.
1099	Node to be replaced cannot be null.
1100	Replacement node cannot be null.
1101	Merge Join must have exactly two inputs. Found : + <size> + inputs
1102	Data is not sorted on <left/right> side. Last two keys encountered were: <previous key>, <current key>
1103	Merge join only supports Filter, Foreach and Load as its predecessor. Found : <operator>
1104	Right input of merge-join must implement SamplableLoader interface. This loader doesn't implement it.
1105	Heap percentage / Conversion factor cannot be set to 0
1106	Merge join is possible only for simple column or '*' join keys when using <funcspec> as the loader
1107	Try to merge incompatible types (eg. numerical type vs non-numeircal type)
1108	Duplicated schema
1109	Input (<input alias>) on which outer join is desired should have a valid schema
1110	Unsupported query: You have an partition column (<colname>) inside a <regexp operator/function/cast/null check operator/bincond operator> in the filter condition.
1111	Use of partition column/condition with non partition column/condition in filter expression is not supported.
1112	Unsupported query: You have an partition column (<column name>) in a construction like: (pcond and ...) or (pcond and ...) where pcond is a condition on a partition column.
1113	Unable to describe schema for nested expression <alias>
1114	Unable to find schema for nested alias <nested alias>
1115	Place holder for Howl related errors
1116	Duplicate udf script (in scripting language)
1117	Cannot merge schema
1118	Cannot convert bytes load from BinStorage
1119	Cannot find LoadCaster class
1120	Cannot cast complex data
1121	Python error
1122	The arity of cogroup/group by columns do not match
1123	Cogroup/Group by * is only allowed if the input has a schema
1124	Mismatch merging expression field schema .. with user specified schema ..
1125	Error determining field schema from object in constant expression"
1126	Schema having field with null alias cannot be merged using alias.
1127	Dereference index out of range in schema.
1128	Cannot find field dereference field in schema.
1129	Referring to column(s) within a column of type .. is not allowed
1130	Datatype of i'th group/join column in j'th relation of statement is incompatible with corresponding column in other relations in the statement

2000	Internal error. Mismatch in group by arities. Expected: <schema>. Found: <schema>
2001	Unable to clone plan before compiling
2002	The output file(s): <filename> already exists

2003	Cannot read from the storage where the output <filename> will be stored
2004	Internal error while trying to check if type casts are needed
2005	Expected <class>, got <class>
2006	TypeCastInserter invoked with an invalid operator class name: <class>
2007	Unable to insert type casts into plan
2008	cannot have more than one input. Found <n> inputs.
2009	Can not move LOLimit up
2010	LOFilter should have one input
2011	Can not insert LOLimit clone
2012	Can not remove LOLimit after <class>
2013	Moving LOLimit in front of <class> is not implemented
2014	Unable to optimize load-stream-store optimization
2015	Invalid physical operators in the physical plan
2016	Unable to obtain a temporary path.
2017	Internal error creating job configuration.
2018	Internal error. Unable to introduce the combiner for optimization.
2019	Expected to find plan with single leaf. Found <n> leaves.
2020	Expected to find plan with UDF leaf. Found <class>
2021	Internal error. Unexpected operator project ⭐ in local rearrange inner plan.
2022	Both map and reduce phases have been done. This is unexpected while compiling.
2023	Received a multi input plan when expecting only a single input one.
2024	Expected reduce to have single leaf. Found <n> leaves.
2025	Expected leaf of reduce plan to always be POStore. Found <class>
2026	No expression plan found in POSort.
2027	Both map and reduce phases have been done. This is unexpected for a merge.
2028	ForEach can only have one successor. Found <n> successors.
2029	Error rewriting POJoinPackage.
2030	Expected reduce plan leaf to have a single predecessor. Found <n> predecessors.
2031	Found map reduce operator with POLocalRearrange as last oper but with no succesor.
2032	Expected map reduce operator to have a single successor. Found <n> successors.

2033	Problems in rearranging map reduce operators in plan.
2034	Error compiling operator <class>
2035	Internal error. Could not compute key type of sort operator.
2036	Unhandled key type <type>
2037	Invalid ship specification. File doesn't exist: <file>
2038	Unable to rename <oldName> to <newName>
2039	Unable to copy <src> to <dst>
2040	Unknown exec type: <type>
2041	No Plan to compile
2042	Internal error. Unable to translate logical plan to physical plan.
2043	Unexpected error during execution.
2044	The type <type> cannot be collected as a Key type
2045	Internal error. Not able to check if the leaf node is a store operator.
2046	Unable to create FileInputHandler.
2047	Internal error. Unable to introduce split operators.
2048	Error while performing checks to introduce split operators.
2049	Error while performing checks to optimize limit operator.
2050	Internal error. Unable to optimize limit operator.
2051	Did not find a predecessor for <Distinct/Filter/Limit/Negative/Null/Sort/Split/Split Output/Store/Stream>.
2052	Internal error. Cannot retrieve operator from null or empty list.
2053	Internal error. Did not find roots in the physical plan.
2054	Internal error. Could not convert <object> to <Integer/Long/Float/Double/Tuple/Bag/Map>
2055	Did not find exception name to create exception from string: <string>
2056	Cannot create exception from empty string.
2057	Did not find fully qualified method name to reconstruct stack trace: <line>
2058	Unable to set index on the newly created POLocalRearrange.
2059	Problem with inserting cast operator for <regular expression/binary conditional/unary operator/user defined function/fragment replicate join/cogroup/project/<operator>> in plan.
2060	Expected one leaf. Found <n> leaves.
2061	Expected single group by element but found multiple elements.
2062	Each COGroup input has to have the same number of inner plans."

2063	Expected multiple group by element but found single element.
2064	Unsupported root type in LOForEach: <operator>
2065	Did not find roots of the inner plan.
2066	Unsupported (root) operator in inner plan: <operator>
2067	does not know how to handle type: <type>
2068	Internal error. Improper use of method getColumn() in POProject
2069	Error during map reduce compilation. Problem in accessing column from project operator.
2070	Problem in accessing column from project operator.
2071	Problem with setting up local rearrange's plans.
2072	Attempt to run a non-algebraic function as an algebraic function
2073	Problem with replacing distinct operator with distinct built-in function.
2074	Could not configure distinct's algebraic functions in map reduce plan.
2075	Could not set algebraic function type.
2076	Unexpected Project-Distinct pair while trying to set up plans for use with combiner.
2077	Problem with reconfiguring plan to add distinct built-in function.
2078	Caught error from UDF: <class> <message from UDF>
2079	Unexpected error while printing physical plan.
2080	Foreach currently does not handle type <type>
2081	Unable to setup the <load/store> function.
2082	Did not expect result of type: <type>
2083	Error while trying to get next result in POStream.
2084	Error while running streaming binary.
2085	Unexpected problem during optimization. Could not find LocalRearrange in combine plan.
2086	Unexpected problem during optimization. Could not find all LocalRearrange operators.
2087	Unexpected problem during optimization. Found index: <index> in multiple LocalRearrange operators.
2088	Unable to get results for: <file specification>
2089	Unable to flag project operator to use single tuple bag.
2090	Received Error while processing the <combine/reduce> plan.
2091	Packaging error while processing group.
2092	No input paths specified in job.

2093	Encountered error in package operator while processing group.
2094	Unable to deserialize object
2095	Did not get reduce key type from job configuration.
2096	Unexpected class in SortPartitioner: <class name>
2097	Failed to copy from: <src> to: <dst>
2098	Invalid seek option: <options>
2099	Problem in constructing slices.
2100	does not exist.
2101	should not be used for storing.
2102	"Cannot test a <type> for emptiness.
2103	Problem while computing <max/min/sum> of <doubles/floats/ints/longs/strings>.
2104	Error while determining schema of <BinStorage data/input>.
2105	Error while converting <int/long/float/double/chararray/tuple/bag/map> to bytes
2106	Error while computing <arity/count/concat/min/max/sum/size> in <class name>
2107	DIFF expected two inputs but received <n> inputs.
2108	Could not determine data type of field: <object>
2109	TextLoader does not support conversion <from/to> <Bag/Tuple/Map/Integer/Long/Float/Double>.
2110	Unable to deserialize optimizer rules.
2111	Unable to create temporary directory: <path>
2112	Unexpected data while reading tuple from binary file.
2113	SingleTupleBag should never be serialized or serialized.
2114	Expected input to be chararray, but got <class name>
2115	Internal error. Expected to throw exception from the backend. Did not find any exception to throw.
2116	Unexpected error. Could not check for the existence of the file(s): <filename>
2117	Unexpected error when launching map reduce job.
2118	Unable to create input slice for: <filename>
2119	Internal Error: Found multiple data types for map key
2120	Internal Error: Unable to determine data type for map key
2121	Error while calling finish method on UDFs.
2122	Sum of probabilities should be one

212 3	Internal Error: Unable to discover required fields from the loads
212 4	Internal Error: Unexpected error creating field schema
212 5	Expected at most one predecessor of load
212 6	Predecessor of load should be store
212 7	Cloning of plan failed.
212 8	Failed to connect store with dependent load.
212 9	Internal Error. Unable to add store to the split plan for optimization.
213 0	Internal Error. Unable to merge split plans for optimization.
213 1	Internal Error. Unable to connect split plan for optimization.
213 2	Internal Error. Unable to replace store with split operator for optimization.
213 3	Internal Error. Unable to connect map plan with successors for optimization.
213 4	Internal Error. Unable to connect map plan with predecessors for optimization.
213 5	Received error from store function.
213 6	Internal Error. Unable to set multi-query index for optimization.
213 7	Internal Error. Unable to add demux to the plan as leaf for optimization.
213 8	Internal Error. Unable to connect package to local rearrange operator in pass-through combiner for optimization.
213 9	Invalid value type: <type>. Expected value type is DataBag.
214 0	Invalid package index: <index>. Should be in the range between 0 and <package array size>.
214 1	Internal Error. Cannot merge non-combiner with combiners for optimization.
214 2	ReadOnceBag should never be serialized.
214 3	Expected index value within POPackageLite is 0, but found 'index'.
214 4	Problem while fixing project inputs during rewiring.
214 5	Problem while rebuilding schemas after transformation.
214 6	Internal Error. Inconsistency in key index found during optimization.
214 7	Error cloning POLocalRearrange for limit after sort.
214 8	Error cloning POPackageLite for limit after sort
214 9	Internal error while trying to check if filters can be pushed up.
215 0	Internal error. The push before input is not set.
215 1	Internal error while pushing filters up.
215 2	Internal error while trying to check if foreach with flatten can be pushed down.

2153	Internal error. The mapping for the flattened columns is empty
2154	Internal error. Schema of successor cannot be null for pushing down foreach with flatten.
2155	Internal error while pushing foreach with flatten down.
2156	Error while fixing projections. Projection map of node to be replaced is null.
2157	Error while fixing projections. No mapping available in old predecessor to replace column.
2158	Error during fixing projections. No mapping available in old predecessor for column to be replaced.
2159	Error during fixing projections. Could not locate replacement column from the old predecessor.
2160	Error during fixing projections. Projection map of new predecessor is null.
2161	Error during fixing projections. No mapping available in new predecessor to replace column.
2162	Error during fixing projections. Could not locate mapping for column <column> in new predecessor.
2163	Error during fixing projections. Could not locate replacement column for column: <column> in the new predecessor.
2164	Expected EOP as return status. Found: <returnStatus>
2165	Problem in index construction.
2166	Key type mismatch. Found key of type <type> on left side. But, found key of type <type> in index built for right side.
2167	LocalRearrange used to extract keys from tuple isn't configured correctly.
2168	Expected physical plan with exactly one root and one leaf.
2169	Physical operator preceding <right/left> predicate not found in compiled MR jobs.
2170	Physical operator preceding both left and right predicate found to be same. This is not expected.
2171	Expected one but found more then one root physical operator in physical plan.
2172	Expected physical operator at root to be POLoad. Found : <PhysicalOperator>
2173	One of the preceding compiled MR operator is null. This is not expected.
2174	Internal exception. Could not create the sampler job.
2175	Internal error. Could not retrieve file size for the sampler.
2176	Error processing right input during merge join
2177	Prune column optimization: Cannot retrieve operator from null or empty list
2178	Prune column optimization: The matching node from the optimizor framework is null
2179	Prune column optimization: Error while performing checks to prune columns.
2180	Prune column optimization: Only LOForEach and LOSplit are expected
2181	Prune column optimization: Unable to prune columns.
2182	Prune column optimization: Only relational operator can be used in column prune optimization.

2183	Prune column optimization: LOLoad must be the root logical operator.
2184	Prune column optimization: Fields list inside RequiredFields is null.
2185	Prune column optimization: Unable to prune columns.
2186	Prune column optimization: Cannot locate node from successor
2187	Column pruner: Cannot get predessors
2188	Column pruner: Cannot prune columns
2189	Column pruner: Expect schema
2190	PruneColumns: Cannot find predecessors for logical operator
2191	PruneColumns: No input to prune
2192	PruneColumns: Column to prune does not exist
2193	PruneColumns: Foreach can only have 1 predecessor
2194	PruneColumns: Expect schema
2195	PruneColumns: Fail to visit foreach inner plan
2196	RelationalOperator: Exception when traversing inner plan
2197	RelationalOperator: Cannot drop column which require *
2198	LOLoad: load only take 1 input
2199	LOLoad: schema mismatch
2200	PruneColumns: Error getting top level project
2201	Could not validate schema alias
2202	Error change distinct/sort to use secondary key optimizer
2203	Sort on columns from different inputs
2204	Error setting secondary key plan
2205	Error visiting POForEach inner plan
2206	Error visiting POSort inner plan
2207	POForEach inner plan has more than 1 root
2208	Exception visiting foreach inner plan
2209	Internal error while processing any partition filter conditions in the filter after the load
2210	Internal Error in logical optimizer.
2211	Column pruner: Unable to prune columns.
2212	Unable to prune plan.

221 3	Error visiting inner plan for ForEach.
221 4	Cannot find POLocalRearrange to set secondary plan.
221 5	See more than 1 successors in the nested plan.
221 6	Cannot get field schema
221 7	Problem setFieldSchema
221 8	Invalid resource schema: bag schema must have tuple as its field
221 9	Attempt to disconnect operators which are not connected
222 0	Plan in inconssistent state, connected in fromEdges but not toEdges
222 1	No more walkers to pop
222 2	Expected LogicalExpressionVisitor to visit expression node
222 3	Expected LogicalPlanVisitor to visit relational node
222 4	Found LogicalExpressionPlan with more than one root
222 5	Projection with nothing to reference
222 6	Cannot fine reference for ProjectExpression
222 7	LogicalExpressionVisitor expects to visit expression plans
222 8	Could not find a related project Expression for Dereference
222 9	Couldn't find matching uid for project expression
223 0	Cannot get column from project
223 1	Unable to set index on newly create POLocalRearrange
223 2	Cannot get schema
223 3	Cannot get predecessor
223 4	Cannot get group key schema
223 5	Expected an ArrayList of Expression Plans
223 6	User defined load function should implement the LoadFunc interface
223 7	Unsupported operator in inner plan
223 8	Expected list of expression plans
223 9	Structure of schema change
224 0	LogicalPlanVisitor can only visit logical plan
224 1	UID is not found in the schema
224 2	TypeCastInserter invoked with an invalid operator

2243	Attempt to remove operator that is still connected to other operators
2244	Hadoop does not return any error message
2245	Cannot get schema from loadFunc
2246	Error merging schema
2247	Cannot determine skewed join schema
2248	twoLevelAccessRequired==true is not supported with" +"and isSubNameMatch==true.
2249	While using 'collected' on group; data must be loaded via loader implementing CollectableLoadFunc.
2250	Blocking operators are not allowed before Collected Group. Consider dropping using 'collected'.
2251	Merge Cogroup work on two or more relations. To use map-side group-by on single relation, use 'collected' qualifier.
2252	Base loader in Cogroup must implement CollectableLoadFunc.
2253	Side loaders in cogroup must implement IndexableLoadFunc.
2254	Currently merged cogroup is not supported after blocking operators.
2255	POSkewedJoin operator has " + compiledInputs.length + " inputs. It should have 2.
2256	Cannot remove and reconnect node with multiple inputs/outputs
2257	An unexpected exception caused the validation to stop
2258	Bug:Two different load functions mapped to an LOCast op
2259	Cannot instantiate class
2260	in split only one of input/output schema is null
2261	input and output schema size of split differ
2262	uid mapped to two different load functions
2263	expected only one predecessor
2264	more than one project star as leaf in plan
2265	Schema not expected for project-star
2266	Expected single LOGenerate output in innerplan of foreach
2267	reset on schema at pos greater than schema size
2268	More than one input found for scalar expression
2269	No input found for scalar expression

2997	Encountered IOException.
2998	Unexpected internal error.
2999	Unhandled internal error.
3000	IOException caught while compiling POMergeJoin
4000	The output file(s): <filename> already exists
4001	Cannot read from the storage where the output <filename> will be stored
4002	Can't read jar file: <name>

4003	Unable to obtain a temporary path.
4004	Invalid ship specification. File doesn't exist: <file>
4005	Unable to rename <oldName> to <newName>
4006	Unable to copy <src> to <dst>
4007	Missing <parameter> from hadoop configuration
4008	Failed to create local hadoop file <file>
4009	Failed to copy data to local hadoop file <file>
6000	The output file(s): <filename> already exists
6001	Cannot read from the storage where the output <filename> will be stored
6002	Unable to obtain a temporary path.
6003	Invalid cache specification. File doesn't exist: <file>
6004	Invalid ship specification. File doesn't exist: <file>
6005	Unable to rename <oldName> to <newName>
6006	Unable to copy <src> to <dst>
6007	Unable to check name <name>
6008	Failed to obtain glob for <pattern>
6009	Failed to create job client
6010	Could not connect to HOD
6011	Failed to run command <command> on server <server>; return code: <code>; error: <error message>
6012	Unable to run command: <command> on server <server>
6013	Unable to chmod <executable> . Thread interrupted.
6014	Failed to save secondary output '<fileName>' of task: <taskId>
6015	During execution, encountered a Hadoop error.
6016	Out of memory.
6017	Execution failed, while processing '<fileNames'>
6018	Error while reading input

References

1. "Pig Developer Cookbook" October 21, 2008, http://wiki.apache.org/pig/PigDeveloperCookbook
2. Santhosh Srinivasan, "Pig Error Handling Design", December 8, 2008, http://wiki.apache.org/pig/PigErrorHandlingDesign