

pyspark-plaso

Tool for Distributed Extraction of Timestamped Events from Files

Reference Documentation

Marek Rychlý and Radek Burget



Acknowledgements

This work was supported by the Ministry of the Interior of the Czech Republic as a part of the project Integrated platform for analysis of digital data from security incidents VI20172020062.

PySpark Plaso

Release 2019

Generated by Doxygen 1.8.13

Contents

- 1 Namespace Index** **1**
 - 1.1 Packages 1

- 2 Hierarchical Index** **3**
 - 2.1 Class Hierarchy 3

- 3 Class Index** **5**
 - 3.1 Class List 5

- 4 File Index** **7**
 - 4.1 File List 7

- 5 Namespace Documentation** **9**
 - 5.1 plaso Namespace Reference 9
 - 5.2 plaso.tarzan Namespace Reference 9
 - 5.3 plaso.tarzan.app Namespace Reference 9
 - 5.4 plaso.tarzan.app.controllers Namespace Reference 9
 - 5.5 plaso.tarzan.app.controllers.controller Namespace Reference 10
 - 5.6 plaso.tarzan.app.controllers.filemancontroller Namespace Reference 10
 - 5.7 plaso.tarzan.app.controllers.plasocontroller Namespace Reference 10
 - 5.8 plaso.tarzan.app.extractors Namespace Reference 10
 - 5.9 plaso.tarzan.app.extractors.extractor Namespace Reference 10
 - 5.10 plaso.tarzan.app.extractors.manager Namespace Reference 10
 - 5.11 plaso.tarzan.app.extractors.pe Namespace Reference 11
 - 5.12 plaso.tarzan.app.extractors.sqlite Namespace Reference 11

5.13	plaso.tarzan.app.extractors.winreg Namespace Reference	11
5.14	plaso.tarzan.app.pyspark_plaso Namespace Reference	11
5.15	plaso.tarzan.app.pyspark_plaso_webapp Namespace Reference	11
5.15.1	Function Documentation	11
5.15.1.1	app_list_routes()	12
5.15.1.2	configure_app()	12
5.15.2	Variable Documentation	12
5.15.2.1	app	12
5.15.2.2	methods	12
5.16	plaso.tarzan.app.views Namespace Reference	12
5.17	plaso.tarzan.app.views.extractview Namespace Reference	12
5.18	plaso.tarzan.app.views.hdfsfileformview Namespace Reference	13
5.19	plaso.tarzan.app.views.hdfsfileview Namespace Reference	13
5.20	plaso.tarzan.app.views.hdfslsview Namespace Reference	13
5.21	plaso.tarzan.app.views.hdfsrmview Namespace Reference	13
5.22	plaso.tarzan.app.views.hdfszipformview Namespace Reference	13
5.23	plaso.tarzan.app.views.hdfszipview Namespace Reference	13
5.24	plaso.tarzan.dfvfs2 Namespace Reference	14
5.25	plaso.tarzan.dfvfs2.definitions Namespace Reference	14
5.25.1	Variable Documentation	14
5.25.1.1	TYPE_INDICATOR_HDFS	14
5.26	plaso.tarzan.dfvfs2.hdfs_file_entry Namespace Reference	14
5.27	plaso.tarzan.dfvfs2.hdfs_file_io Namespace Reference	14
5.28	plaso.tarzan.dfvfs2.hdfs_file_system Namespace Reference	14
5.29	plaso.tarzan.dfvfs2.hdfs_path_spec Namespace Reference	15
5.30	plaso.tarzan.dfvfs2.hdfs_resolver_helper Namespace Reference	15
5.31	plaso.tarzan.file Namespace Reference	15
5.32	plaso.tarzan.file.file Namespace Reference	15
5.33	plaso.tarzan.file.generic_file Namespace Reference	15
5.34	plaso.tarzan.file.hdfs_file Namespace Reference	15

5.35	plaso.tarzan.lib Namespace Reference	16
5.36	plaso.tarzan.lib.custom_pip_package_dir Namespace Reference	16
5.37	plaso.tarzan.lib.hdfs Namespace Reference	16
5.38	plaso.tarzan.lib.pyarrow_hdfs Namespace Reference	16
5.39	plaso.tarzan.lib.pyspark_hdfs Namespace Reference	16
5.40	plaso.tarzan.mediator Namespace Reference	16
5.41	plaso.tarzan.mediator.buffered_mediator Namespace Reference	17
5.41.1	Function Documentation	17
5.41.1.1	generate()	17
5.41.2	Variable Documentation	17
5.41.2.1	buffered_mediator	17
5.41.2.2	result	17
5.42	plaso.tarzan.mediator.mediator Namespace Reference	17
5.43	plaso.tarzan.mediator.print_mediator Namespace Reference	18
5.44	plaso.tarzan.run Namespace Reference	18
5.44.1	Function Documentation	18
5.44.1.1	run_PE()	18
5.44.1.2	run_SQLite()	18
5.44.1.3	run_WinReg()	18
5.44.2	Variable Documentation	18
5.44.2.1	m	18

6	Class Documentation	19
6.1	plaso.tarzan.mediator.buffered_mediator.BufferedMediator Class Reference	19
6.1.1	Detailed Description	20
6.1.2	Constructor & Destructor Documentation	20
6.1.2.1	__init__()	20
6.1.3	Member Function Documentation	21
6.1.3.1	flush_buffer()	21
6.1.3.2	get_buffer_queue()	21
6.1.3.3	ProduceEventWithEventData()	21
6.1.3.4	ProduceExtractionError()	21
6.1.4	Member Data Documentation	22
6.1.4.1	buffer_queue	22
6.1.4.2	Error	22
6.1.4.3	Event	22
6.2	plaso.tarzan.app.controllers.controller.Controller Class Reference	22
6.2.1	Detailed Description	23
6.2.2	Constructor & Destructor Documentation	23
6.2.2.1	__init__()	23
6.2.3	Member Function Documentation	23
6.2.3.1	make_hdfs_uri()	23
6.2.3.2	strip_hdfs_uri()	24
6.2.4	Member Data Documentation	24
6.2.4.1	hdfs_base_uri	24
6.3	plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir Class Reference	24
6.3.1	Detailed Description	24
6.3.2	Constructor & Destructor Documentation	25
6.3.2.1	__init__()	25
6.3.3	Member Function Documentation	25
6.3.3.1	check_installed_package()	25
6.3.3.2	install()	25

6.3.3.3	make_package_dir()	25
6.3.3.4	pip_install()	26
6.3.3.5	register_package_dir()	26
6.3.4	Member Data Documentation	26
6.3.4.1	package_dir	26
6.4	plaso.tarzan.app.extractors.extractor.Extractor Class Reference	26
6.4.1	Detailed Description	27
6.4.2	Member Function Documentation	27
6.4.2.1	extract()	27
6.4.2.2	filter()	28
6.4.2.3	get_description()	28
6.4.2.4	get_name()	28
6.5	plaso.tarzan.app.views.extractview.ExtractView Class Reference	29
6.5.1	Constructor & Destructor Documentation	29
6.5.1.1	__init__()	30
6.5.2	Member Function Documentation	30
6.5.2.1	get()	30
6.5.3	Member Data Documentation	30
6.5.3.1	controller	30
6.6	plaso.tarzan.app.controllers.filemancontroller.FileManController Class Reference	30
6.6.1	Detailed Description	31
6.6.2	Constructor & Destructor Documentation	31
6.6.2.1	__init__()	32
6.6.3	Member Function Documentation	32
6.6.3.1	get()	32
6.6.3.2	get_zip()	32
6.6.3.3	ls()	32
6.6.3.4	put()	33
6.6.3.5	put_form()	33
6.6.3.6	rm()	33

6.6.3.7	zip_put()	33
6.6.3.8	zip_put_form()	34
6.6.4	Member Data Documentation	34
6.6.4.1	hdfs	34
6.7	plaso.tarzan.file.generic_file.FileObject Class Reference	34
6.7.1	Detailed Description	35
6.7.2	Member Function Documentation	35
6.7.2.1	close()	35
6.7.2.2	read()	35
6.7.2.3	seek()	36
6.8	plaso.tarzan.lib.hdfs.Hdfs Class Reference	36
6.8.1	Detailed Description	37
6.8.2	Member Function Documentation	37
6.8.2.1	append_to_path()	37
6.8.2.2	basename()	37
6.8.2.3	close_filesystem()	38
6.8.2.4	close_stream()	38
6.8.2.5	dirname()	38
6.8.2.6	exists()	38
6.8.2.7	get_filesystem()	39
6.8.2.8	get_path_size()	39
6.8.2.9	get_stream_offset()	39
6.8.2.10	info()	39
6.8.2.11	list_files()	40
6.8.2.12	make_path()	40
6.8.2.13	make_qualified_path()	40
6.8.2.14	make_simple_path()	41
6.8.2.15	make_uri()	41
6.8.2.16	mkdir()	41
6.8.2.17	open_filesystem()	41

6.8.2.18	open_inputstream()	42
6.8.2.19	open_outputstream()	42
6.8.2.20	parse_uri()	42
6.8.2.21	read_inputstream()	42
6.8.2.22	remove()	43
6.8.2.23	seek_stream()	43
6.8.2.24	write_outputstream()	43
6.8.3	Member Data Documentation	43
6.8.3.1	PATH_SEPARATOR	43
6.8.3.2	SCHEME	44
6.9	plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory Class Reference	44
6.9.1	Detailed Description	44
6.10	plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile Class Reference	45
6.10.1	Detailed Description	45
6.10.2	Constructor & Destructor Documentation	45
6.10.2.1	__init__()	46
6.10.3	Member Function Documentation	46
6.10.3.1	get_offset()	46
6.10.3.2	get_size()	46
6.10.3.3	read()	46
6.10.3.4	seek()	47
6.11	plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry Class Reference	47
6.11.1	Detailed Description	48
6.11.2	Constructor & Destructor Documentation	48
6.11.2.1	__init__()	48
6.11.3	Member Function Documentation	48
6.11.3.1	access_time()	49
6.11.3.2	change_time()	49
6.11.3.3	GetLinkedFileEntry()	49
6.11.3.4	GetParentFileEntry()	49

6.11.3.5	link()	49
6.11.3.6	modification_time()	50
6.11.3.7	name()	50
6.11.4	Member Data Documentation	50
6.11.4.1	entry_type	50
6.11.4.2	TYPE_INDICATOR	50
6.12	plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView Class Reference	51
6.12.1	Constructor & Destructor Documentation	51
6.12.1.1	__init__()	52
6.12.2	Member Function Documentation	52
6.12.2.1	delete()	52
6.12.2.2	get()	52
6.12.2.3	post()	52
6.12.2.4	put()	52
6.12.3	Member Data Documentation	52
6.12.3.1	controller	53
6.13	plaso.tarzan.file.hdfs_file.HdfsFileObject Class Reference	53
6.13.1	Detailed Description	54
6.13.2	Constructor & Destructor Documentation	54
6.13.2.1	__init__()	54
6.13.3	Member Function Documentation	54
6.13.3.1	close()	54
6.13.3.2	open()	55
6.13.3.3	read()	55
6.13.3.4	seek()	55
6.13.3.5	tell()	55
6.13.4	Member Data Documentation	56
6.13.4.1	hdfs	56
6.13.4.2	input_stream	56
6.13.4.3	path	56

6.14	plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem Class Reference	56
6.14.1	Detailed Description	57
6.14.2	Constructor & Destructor Documentation	57
6.14.2.1	__init__()	58
6.14.3	Member Function Documentation	58
6.14.3.1	BasenamePath()	58
6.14.3.2	DirnamePath()	58
6.14.3.3	FileEntryExistsByPathSpec()	58
6.14.3.4	GetFileEntryByPathSpec()	59
6.14.3.5	GetRootFileEntry()	59
6.14.3.6	JoinPath()	59
6.14.3.7	SplitPath()	59
6.14.4	Member Data Documentation	60
6.14.4.1	hdfs	60
6.14.4.2	LOCATION_ROOT	60
6.14.4.3	PATH_SEPARATOR	60
6.14.4.4	TYPE_INDICATOR	60
6.15	plaso.tarzan.app.views.hdfsfileview.HdfsFileView Class Reference	60
6.15.1	Constructor & Destructor Documentation	61
6.15.1.1	__init__()	61
6.15.2	Member Function Documentation	61
6.15.2.1	delete()	61
6.15.2.2	get()	62
6.15.2.3	post()	62
6.15.2.4	put()	62
6.15.3	Member Data Documentation	62
6.15.3.1	controller	62
6.16	plaso.tarzan.app.views.hdfslsview.HdfsLsView Class Reference	63
6.16.1	Constructor & Destructor Documentation	63
6.16.1.1	__init__()	64

6.16.2	Member Function Documentation	64
6.16.2.1	get()	64
6.16.3	Member Data Documentation	64
6.16.3.1	controller	64
6.17	plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec Class Reference	64
6.17.1	Detailed Description	65
6.17.2	Constructor & Destructor Documentation	65
6.17.2.1	__init__()	65
6.17.3	Member Data Documentation	65
6.17.3.1	TYPE_INDICATOR	66
6.18	plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper Class Reference	66
6.18.1	Detailed Description	67
6.18.2	Member Function Documentation	67
6.18.2.1	NewFileObject()	67
6.18.2.2	NewFileSystem()	67
6.18.3	Member Data Documentation	67
6.18.3.1	TYPE_INDICATOR	67
6.19	plaso.tarzan.app.views.hdfsrmview.HdfsRmView Class Reference	68
6.19.1	Constructor & Destructor Documentation	68
6.19.1.1	__init__()	69
6.19.2	Member Function Documentation	69
6.19.2.1	delete()	69
6.19.2.2	get()	69
6.19.3	Member Data Documentation	69
6.19.3.1	controller	69
6.20	plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView Class Reference	70
6.20.1	Constructor & Destructor Documentation	70
6.20.1.1	__init__()	71
6.20.2	Member Function Documentation	71
6.20.2.1	delete()	71

6.20.2.2	get()	71
6.20.2.3	post()	71
6.20.2.4	put()	71
6.20.3	Member Data Documentation	71
6.20.3.1	controller	72
6.21	plaso.tarzan.app.views.hdfszipview.HdfsZipView Class Reference	72
6.21.1	Constructor & Destructor Documentation	73
6.21.1.1	__init__()	73
6.21.2	Member Function Documentation	73
6.21.2.1	delete()	73
6.21.2.2	get()	73
6.21.2.3	post()	73
6.21.2.4	put()	73
6.21.3	Member Data Documentation	74
6.21.3.1	controller	74
6.22	plaso.tarzan.file.file.LocalFileObject Class Reference	74
6.22.1	Detailed Description	75
6.22.2	Constructor & Destructor Documentation	75
6.22.2.1	__init__()	75
6.22.3	Member Function Documentation	76
6.22.3.1	close()	76
6.22.3.2	open()	76
6.22.3.3	read()	76
6.22.3.4	seek()	76
6.22.3.5	tell()	77
6.22.4	Member Data Documentation	77
6.22.4.1	file_object	77
6.22.4.2	path	77
6.23	plaso.tarzan.app.extractors.manager.Manager Class Reference	77
6.23.1	Detailed Description	78

6.23.2	Member Function Documentation	78
6.23.2.1	deregister_extractor()	78
6.23.2.2	get_extractors()	78
6.23.2.3	register_extractor()	79
6.24	plaso.tarzan.mediator.mediator.Mediator Class Reference	79
6.24.1	Detailed Description	80
6.24.2	Constructor & Destructor Documentation	80
6.24.2.1	__init__()	80
6.24.3	Member Function Documentation	80
6.24.3.1	abort()	80
6.24.3.2	AddEventAttribute()	81
6.24.3.3	AppendToParserChain()	81
6.24.3.4	GetFileEntry()	81
6.24.3.5	GetFilename()	81
6.24.3.6	PopFromParserChain()	82
6.24.3.7	ProduceEventWithEventData()	82
6.24.3.8	ProduceExtractionError()	82
6.24.3.9	RemoveEventAttribute()	82
6.24.3.10	SetFileEntry()	83
6.24.3.11	temporary_directory()	83
6.24.4	Member Data Documentation	83
6.24.4.1	codepage	83
6.24.4.2	collection_filters_helper	83
6.24.4.3	file_entry	83
6.24.4.4	parser_chain_components	83
6.25	plaso.tarzan.app.extractors.pe.PE Class Reference	84
6.25.1	Detailed Description	84
6.25.2	Member Function Documentation	85
6.25.2.1	extract()	85
6.25.2.2	filter()	85

6.26	plaso.tarzan.app.controllers.plasocontroller.PlasoController Class Reference	85
6.26.1	Detailed Description	86
6.26.2	Constructor & Destructor Documentation	86
6.26.2.1	__init__()	86
6.26.3	Member Function Documentation	87
6.26.3.1	extract()	87
6.26.4	Member Data Documentation	87
6.26.4.1	spark_context	87
6.27	plaso.tarzan.mediator.print_mediator.PrintMediator Class Reference	87
6.27.1	Detailed Description	88
6.27.2	Constructor & Destructor Documentation	88
6.27.2.1	__init__()	88
6.27.3	Member Function Documentation	89
6.27.3.1	ProduceEventWithEventData()	89
6.27.3.2	ProduceExtractionError()	89
6.27.3.3	ProduceExtractionWarning()	89
6.28	plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs Class Reference	90
6.28.1	Detailed Description	91
6.28.2	Member Function Documentation	91
6.28.2.1	close_filesystem()	91
6.28.2.2	close_stream()	91
6.28.2.3	exists()	91
6.28.2.4	get_filesystem()	92
6.28.2.5	get_path_size()	92
6.28.2.6	get_stream_offset()	92
6.28.2.7	info()	92
6.28.2.8	list_files()	93
6.28.2.9	mkdir()	93
6.28.2.10	open_filesystem()	93
6.28.2.11	open_inputstream()	93

6.28.2.12	<code>open_outputstream()</code>	94
6.28.2.13	<code>pass_to_outputstream()</code>	94
6.28.2.14	<code>read_inputstream()</code>	94
6.28.2.15	<code>remove()</code>	94
6.28.2.16	<code>seek_stream()</code>	95
6.28.2.17	<code>write_outputstream()</code>	95
6.28.3	Member Data Documentation	95
6.28.3.1	<code>fs</code>	95
6.29	<code>plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs</code> Class Reference	96
6.29.1	Detailed Description	97
6.29.2	Constructor & Destructor Documentation	97
6.29.2.1	<code>__init__()</code>	97
6.29.3	Member Function Documentation	97
6.29.3.1	<code>append_to_path()</code>	98
6.29.3.2	<code>close_filesystem()</code>	98
6.29.3.3	<code>close_stream()</code>	98
6.29.3.4	<code>exists()</code>	98
6.29.3.5	<code>get_filesystem()</code>	99
6.29.3.6	<code>get_path_size()</code>	99
6.29.3.7	<code>get_stream_offset()</code>	99
6.29.3.8	<code>info()</code>	99
6.29.3.9	<code>list_files()</code>	100
6.29.3.10	<code>make_path()</code>	100
6.29.3.11	<code>mkdir()</code>	100
6.29.3.12	<code>open_filesystem()</code>	101
6.29.3.13	<code>open_inputstream()</code>	101
6.29.3.14	<code>open_outputstream()</code>	101
6.29.3.15	<code>read_inputstream()</code>	101
6.29.3.16	<code>read_inputstream_with_chunk()</code>	102
6.29.3.17	<code>remove()</code>	102

6.29.3.18	seek_stream()	102
6.29.3.19	seek_stream_with_path()	103
6.29.3.20	write_outputstream()	103
6.29.4	Member Data Documentation	103
6.29.4.1	fs	103
6.29.4.2	fs_class	103
6.29.4.3	isr_class	103
6.29.4.4	path_class	104
6.29.4.5	spark_context	104
6.29.4.6	uri_class	104
6.30	plaso.tarzan.app.pyspark_plaso.PySparkPlaso Class Reference	104
6.30.1	Detailed Description	105
6.30.2	Member Function Documentation	105
6.30.2.1	action_events_rdd_by_collecting_into_json()	105
6.30.2.2	action_events_rdd_by_saving_into_halyard()	105
6.30.2.3	create_files_rdd()	106
6.30.2.4	get_java_rdd_helpers_package()	106
6.30.2.5	list_files()	106
6.30.2.6	transform_files_rdd_to_extracted_events_rdd()	106
6.31	plaso.tarzan.app.extractors.sqlite.SQLite Class Reference	107
6.31.1	Detailed Description	108
6.31.2	Member Function Documentation	108
6.31.2.1	extract()	108
6.31.2.2	filter()	108
6.32	plaso.tarzan.app.extractors.winreg.WinReg Class Reference	109
6.32.1	Detailed Description	110
6.32.2	Member Function Documentation	110
6.32.2.1	extract()	110
6.32.2.2	filter()	110

7	File Documentation	111
7.1	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/__init__.py File Reference	111
7.2	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/__init__.py File Reference	111
7.3	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/__init__.py File Reference	111
7.4	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/__init__.py File Reference	111
7.5	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/__init__.py File Reference	112
7.6	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/__init__.py File Reference	112
7.7	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/__init__.py File Reference	112
7.8	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/__init__.py File Reference	112
7.9	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/__init__.py File Reference	112
7.10	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controller.py File Reference	112
7.11	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/filemancontroller.py File Reference	113
7.12	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroller.py File Reference	113
7.13	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/extractor.py File Reference	113
7.14	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/manager.py File Reference	113
7.15	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.py File Reference	114
7.16	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite.py File Reference	114
7.17	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.py File Reference	114
7.18	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso.py File Reference	114
7.19	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso_webapp.py File Reference	115
7.20	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractview.py File Reference	115
7.21	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformview.py File Reference	115
7.22	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.py File Reference	116
7.23	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfslsview.py File Reference	116
7.24	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsrview.py File Reference	116
7.25	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipformview.py File Reference	116
7.26	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipview.py File Reference	117
7.27	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/definitions.py File Reference	117
7.28	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs_file_entry.py File Reference	117
7.29	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs_file_io.py File Reference	117

7.30	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs_file_system.py File Reference	118
7.31	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs_path_spec.py File Reference	118
7.32	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs_resolver_helper.py File Reference	118
7.33	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/file.py File Reference	118
7.34	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/generic_file.py File Reference	119
7.35	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/hdfs_file.py File Reference	119
7.36	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom_pip_package_dir.py File Reference	119
7.37	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/hdfs.py File Reference	119
7.38	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow_hdfs.py File Reference	120
7.39	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyspark_hdfs.py File Reference	120
7.40	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered_mediator.py File Reference	120
7.41	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.py File Reference	121
7.42	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/print_mediator.py File Reference	121
7.43	/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/run.py File Reference	121
Index		123

Chapter 1

Namespace Index

1.1 Packages

Here are the packages with brief descriptions (if available):

- plaso 9
- plaso.tarzan 9
- plaso.tarzan.app 9
- plaso.tarzan.app.controllers 9
- plaso.tarzan.app.controllers.controller 10
- plaso.tarzan.app.controllers.filemancontroller 10
- plaso.tarzan.app.controllers.plasocontroller 10
- plaso.tarzan.app.extractors 10
- plaso.tarzan.app.extractors.extractor 10
- plaso.tarzan.app.extractors.manager 10
- plaso.tarzan.app.extractors.pe 11
- plaso.tarzan.app.extractors.sqlite 11
- plaso.tarzan.app.extractors.winreg 11
- plaso.tarzan.app.pyspark_plaso 11
- plaso.tarzan.app.pyspark_plaso_webapp 11
- plaso.tarzan.app.views 12
- plaso.tarzan.app.views.extractview 12
- plaso.tarzan.app.views.hdfsfileformview 13
- plaso.tarzan.app.views.hdfsfileview 13
- plaso.tarzan.app.views.hdfslsview 13
- plaso.tarzan.app.views.hdfsrview 13
- plaso.tarzan.app.views.hdfszipformview 13
- plaso.tarzan.app.views.hdfszipview 13
- plaso.tarzan.dfvfs2 14
- plaso.tarzan.dfvfs2.definitions 14
- plaso.tarzan.dfvfs2.hdfs_file_entry 14
- plaso.tarzan.dfvfs2.hdfs_file_io 14
- plaso.tarzan.dfvfs2.hdfs_file_system 14
- plaso.tarzan.dfvfs2.hdfs_path_spec 15
- plaso.tarzan.dfvfs2.hdfs_resolver_helper 15
- plaso.tarzan.file 15
- plaso.tarzan.file.file 15
- plaso.tarzan.file.generic_file 15
- plaso.tarzan.file.hdfs_file 15
- plaso.tarzan.lib 16

plaso.tarzan.lib.custom_pip_package_dir	16
plaso.tarzan.lib.hdfs	16
plaso.tarzan.lib.pyarrow_hdfs	16
plaso.tarzan.lib.pyspark_hdfs	16
plaso.tarzan.mediator	16
plaso.tarzan.mediator.buffered_mediator	17
plaso.tarzan.mediator.mediator	17
plaso.tarzan.mediator.print_mediator	18
plaso.tarzan.run	18

Chapter 2

Hierarchical Index

2.1 Class Hierarchy

This inheritance list is sorted roughly, but not completely, alphabetically:

plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir	24
plaso.tarzan.lib.hdfs.Hdfs	36
plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs	90
plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs	96
object	
plaso.tarzan.app.controllers.controller.Controller	22
plaso.tarzan.app.controllers.filemancontroller.FileManController	30
plaso.tarzan.app.controllers.plasocontroller.PlasoController	85
plaso.tarzan.app.extractors.extractor.Extractor	26
plaso.tarzan.app.extractors.pe.PE	84
plaso.tarzan.app.extractors.sqlite.SQLite	107
plaso.tarzan.app.extractors.winreg.WinReg	109
plaso.tarzan.app.extractors.manager.Manager	77
plaso.tarzan.app.pyspark_plaso.PySparkPlaso	104
plaso.tarzan.file.generic_file.FileObject	34
plaso.tarzan.file.file.LocalFileObject	74
plaso.tarzan.file.hdfs_file.HdfsFileObject	53
plaso.tarzan.mediator.mediator.Mediator	79
plaso.tarzan.mediator.buffered_mediator.BufferedMediator	19
plaso.tarzan.mediator.print_mediator.PrintMediator	87
Directory	
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory	44
FileEntry	
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry	47
FileIO	
plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile	45
FileSystem	
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem	56
LocationPathSpec	
plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec	64
MethodView	
plaso.tarzan.app.views.extractview.ExtractView	29
plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView	51
plaso.tarzan.app.views.hdfsfileview.HdfsFileView	60

plaso.tarzan.app.views.hdfsview.HdfsLsView	63
plaso.tarzan.app.views.hdfsrmview.HdfsRmView	68
plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView	70
plaso.tarzan.app.views.hdfszipview.HdfsZipView	72
ResolverHelper	
plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper	66

Chapter 3

Class Index

3.1 Class List

Here are the classes, structs, unions and interfaces with brief descriptions:

plaso.tarzan.mediator.buffered_mediator.BufferedMediator	19
plaso.tarzan.app.controllers.controller.Controller	22
plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir	24
plaso.tarzan.app.extractors.extractor.Extractor	26
plaso.tarzan.app.views.extractview.ExtractView	29
plaso.tarzan.app.controllers.filemancontroller.FileManController	30
plaso.tarzan.file.generic_file.FileObject	34
plaso.tarzan.lib.hdfs.Hdfs	36
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory	44
plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile	45
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry	47
plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView	51
plaso.tarzan.app.hdfs_file.HdfsFileObject	53
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem	56
plaso.tarzan.app.views.hdfsfileview.HdfsFileView	60
plaso.tarzan.app.views.hdfsfileview.HdfsLsView	63
plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec	64
plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper	66
plaso.tarzan.app.views.hdfsrmview.HdfsRmView	68
plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView	70
plaso.tarzan.app.views.hdfszipview.HdfsZipView	72
plaso.tarzan.file.file.LocalFileObject	74
plaso.tarzan.app.extractors.manager.Manager	77
plaso.tarzan.mediator.mediator.Mediator	79
plaso.tarzan.app.extractors.pe.PE	84
plaso.tarzan.app.controllers.plasocontroller.PlasoController	85
plaso.tarzan.mediator.print_mediator.PrintMediator	87
plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs	90
plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs	96
plaso.tarzan.app.pyspark_plaso.PySparkPlaso	104
plaso.tarzan.app.extractors.sqlite.SQLite	107
plaso.tarzan.app.extractors.winreg.WinReg	109

Chapter 4

File Index

4.1 File List

Here is a list of all files with brief descriptions:

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/__init__.py	111
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/run.py	121
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/__init__.py	111
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso.py	114
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso_webapp.py	115
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/__init__.py	111
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controller.py	112
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/filemancontroller.py	113
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroller.py	113
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/__init__.py	111
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/extractor.py	113
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/manager.py	113
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.py	114
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite.py	114
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.py	114
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/__init__.py	112
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractview.py	115
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformview.py	115
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.py	116
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfslsview.py	116
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsrview.py	116
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipformview.py	116
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipview.py	117
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/__init__.py	112
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/definitions.py	117
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/hdfs_file_entry.py	117
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/hdfs_file_io.py	117
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/hdfs_file_system.py	118
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/hdfs_path_spec.py	118
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvfs2/hdfs_resolver_helper.py	118
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/__init__.py	112
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/file.py	118
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/generic_file.py	119
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/hdfs_file.py	119
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/__init__.py	112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom_pip_package_dir.py	119
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/hdfs.py	119
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow_hdfs.py	120
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyspark_hdfs.py	120
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/__init__.py	112
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered_mediator.py	120
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.py	121
/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/print_mediator.py	121

Chapter 5

Namespace Documentation

5.1 plaso Namespace Reference

Namespaces

- [tarzan](#)

5.2 plaso.tarzan Namespace Reference

Namespaces

- [app](#)
- [dfvfs2](#)
- [file](#)
- [lib](#)
- [mediator](#)
- [run](#)

5.3 plaso.tarzan.app Namespace Reference

Namespaces

- [controllers](#)
- [extractors](#)
- [pyspark_plaso](#)
- [pyspark_plaso_webapp](#)
- [views](#)

5.4 plaso.tarzan.app.controllers Namespace Reference

Namespaces

- [controller](#)
- [filemancontroller](#)
- [plasocontroller](#)

5.5 plaso.tarzan.app.controllers.controller Namespace Reference

Classes

- class [Controller](#)

5.6 plaso.tarzan.app.controllers.filemancontroller Namespace Reference

Classes

- class [FileManController](#)

5.7 plaso.tarzan.app.controllers.plasocontroller Namespace Reference

Classes

- class [PlasoController](#)

5.8 plaso.tarzan.app.extractors Namespace Reference

Namespaces

- [extractor](#)
- [manager](#)
- [pe](#)
- [sqlite](#)
- [winreg](#)

5.9 plaso.tarzan.app.extractors.extractor Namespace Reference

Classes

- class [Extractor](#)

5.10 plaso.tarzan.app.extractors.manager Namespace Reference

Classes

- class [Manager](#)

5.11 plaso.tarzan.app.extractors.pe Namespace Reference

Classes

- class [PE](#)

5.12 plaso.tarzan.app.extractors.sqlite Namespace Reference

Classes

- class [SQLite](#)

5.13 plaso.tarzan.app.extractors.winreg Namespace Reference

Classes

- class [WinReg](#)

5.14 plaso.tarzan.app.pyspark_plaso Namespace Reference

Classes

- class [PySparkPlaso](#)

5.15 plaso.tarzan.app.pyspark_plaso_webapp Namespace Reference

Functions

- def [app_list_routes](#) ()
- def [configure_app](#) (spark_context, hdfs_uri)

Variables

- [app](#) = Flask(__name__)
- [methods](#)

5.15.1 Function Documentation

5.15.1.1 `app_list_routes()`

```
def plaso.tarzan.app.pyspark_plaso_webapp.app_list_routes ( )
```

5.15.1.2 `configure_app()`

```
def plaso.tarzan.app.pyspark_plaso_webapp.configure_app (
    spark_context,
    hdfs_uri )
```

5.15.2 Variable Documentation

5.15.2.1 `app`

```
plaso.tarzan.app.pyspark_plaso_webapp.app = Flask(__name__)
```

5.15.2.2 `methods`

```
plaso.tarzan.app.pyspark_plaso_webapp.methods
```

5.16 `plaso.tarzan.app.views` Namespace Reference

Namespaces

- [extractview](#)
- [hdfsfileformview](#)
- [hdfsfileview](#)
- [hdfslsview](#)
- [hdfsrmview](#)
- [hdfszipformview](#)
- [hdfszipview](#)

5.17 `plaso.tarzan.app.views.extractview` Namespace Reference

Classes

- class [ExtractView](#)

5.18 plaso.tarzan.app.views.hdfsfileformview Namespace Reference

Classes

- class [HdfsFileFormView](#)

5.19 plaso.tarzan.app.views.hdfsfileview Namespace Reference

Classes

- class [HdfsFileView](#)

5.20 plaso.tarzan.app.views.hdfsListView Namespace Reference

Classes

- class [HdfsListView](#)

5.21 plaso.tarzan.app.views.hdfsrmview Namespace Reference

Classes

- class [HdfsRmView](#)

5.22 plaso.tarzan.app.views.hdfszipformview Namespace Reference

Classes

- class [HdfsZipFormView](#)

5.23 plaso.tarzan.app.views.hdfszipview Namespace Reference

Classes

- class [HdfsZipView](#)

5.24 plaso.tarzan.dfvfs2 Namespace Reference

Namespaces

- [definitions](#)
- [hdfs_file_entry](#)
- [hdfs_file_io](#)
- [hdfs_file_system](#)
- [hdfs_path_spec](#)
- [hdfs_resolver_helper](#)

5.25 plaso.tarzan.dfvfs2.definitions Namespace Reference

Variables

- string [TYPE_INDICATOR_HDFS](#) = 'HDFS'

5.25.1 Variable Documentation

5.25.1.1 TYPE_INDICATOR_HDFS

```
string plaso.tarzan.dfvfs2.definitions.TYPE_INDICATOR_HDFS = 'HDFS'
```

5.26 plaso.tarzan.dfvfs2.hdfs_file_entry Namespace Reference

Classes

- class [HDFSDirectory](#)
- class [HDFSFileEntry](#)

5.27 plaso.tarzan.dfvfs2.hdfs_file_io Namespace Reference

Classes

- class [HDFSFile](#)

5.28 plaso.tarzan.dfvfs2.hdfs_file_system Namespace Reference

Classes

- class [HDFSFileSystem](#)

5.29 plaso.tarzan.dfvfs2.hdfs_path_spec Namespace Reference

Classes

- class [HDFSPathSpec](#)

5.30 plaso.tarzan.dfvfs2.hdfs_resolver_helper Namespace Reference

Classes

- class [HDFSResolverHelper](#)

5.31 plaso.tarzan.file Namespace Reference

Namespaces

- [file](#)
- [generic_file](#)
- [hdfs_file](#)

5.32 plaso.tarzan.file.file Namespace Reference

Classes

- class [LocalFileObject](#)

5.33 plaso.tarzan.file.generic_file Namespace Reference

Classes

- class [FileObject](#)

5.34 plaso.tarzan.file.hdfs_file Namespace Reference

Classes

- class [HdfsFileObject](#)

5.35 plaso.tarzan.lib Namespace Reference

Namespaces

- [custom_pip_package_dir](#)
- [hdfs](#)
- [pyarrow_hdfs](#)
- [pyspark_hdfs](#)

5.36 plaso.tarzan.lib.custom_pip_package_dir Namespace Reference

Classes

- class [CustomPipPackageDir](#)

5.37 plaso.tarzan.lib.hdfs Namespace Reference

Classes

- class [Hdfs](#)

5.38 plaso.tarzan.lib.pyarrow_hdfs Namespace Reference

Classes

- class [PyArrowHdfs](#)

5.39 plaso.tarzan.lib.pyspark_hdfs Namespace Reference

Classes

- class [PySparkHdfs](#)

5.40 plaso.tarzan.mediator Namespace Reference

Namespaces

- [buffered_mediator](#)
- [mediator](#)
- [print_mediator](#)

5.41 plaso.tarzan.mediator.buffered_mediator Namespace Reference

Classes

- class [BufferedMediator](#)

Functions

- def [generate](#) (bm)

Variables

- [buffered_mediator](#) = [BufferedMediator](#)()
- [result](#) = buffered_mediator.flush_buffer()

5.41.1 Function Documentation

5.41.1.1 generate()

```
def plaso.tarzan.mediator.buffered_mediator.generate (  
    bm )
```

5.41.2 Variable Documentation

5.41.2.1 buffered_mediator

```
plaso.tarzan.mediator.buffered_mediator.buffered_mediator = BufferedMediator()
```

5.41.2.2 result

```
plaso.tarzan.mediator.buffered_mediator.result = buffered_mediator.flush_buffer()
```

5.42 plaso.tarzan.mediator.mediator Namespace Reference

Classes

- class [Mediator](#)

5.43 plaso.tarzan.mediator.print_mediator Namespace Reference

Classes

- class [PrintMediator](#)

5.44 plaso.tarzan.run Namespace Reference

Functions

- def [run_PE](#) (mediator)
- def [run_SQLite](#) (mediator)
- def [run_WinReg](#) (mediator)

Variables

- `m = plaso.tarzan.mediator.print_mediator.PrintMediator()`

5.44.1 Function Documentation

5.44.1.1 [run_PE\(\)](#)

```
def plaso.tarzan.run.run_PE (  
    mediator )
```

5.44.1.2 [run_SQLite\(\)](#)

```
def plaso.tarzan.run.run_SQLite (  
    mediator )
```

5.44.1.3 [run_WinReg\(\)](#)

```
def plaso.tarzan.run.run_WinReg (  
    mediator )
```

5.44.2 Variable Documentation

5.44.2.1 [m](#)

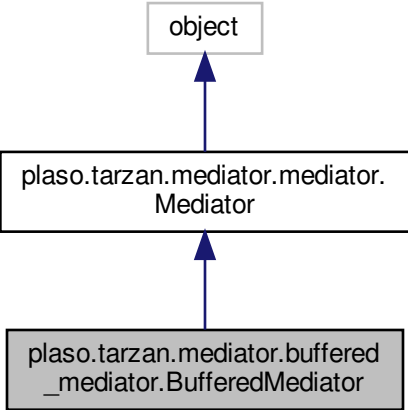
```
plaso.tarzan.run.m = plaso.tarzan.mediator.print\_mediator.PrintMediator\(\)
```


Chapter 6

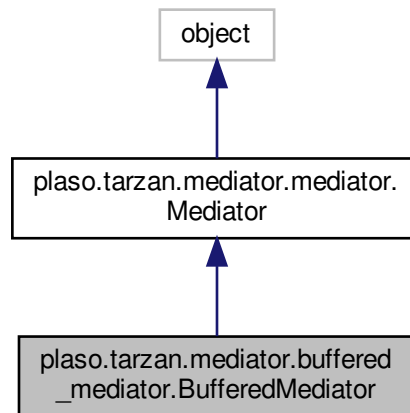
Class Documentation

6.1 plaso.tarzan.mediator.buffered_mediator.BufferedMediator Class Reference

Inheritance diagram for plaso.tarzan.mediator.buffered_mediator.BufferedMediator:



Collaboration diagram for `plaso.tarzan.mediator.buffered_mediator.BufferedMediator`:



Public Member Functions

- `def __init__ (self, buffer_queue=None)`
- `def ProduceEventWithEventData (self, event, event_data)`
- `def ProduceExtractionError (self, message, path_spec=None)`
- `def get_buffer_queue (self)`
- `def flush_buffer (self, exception_on_error=False)`

Public Attributes

- [buffer_queue](#)
- [Event](#)
- [Error](#)

6.1.1 Detailed Description

Tarzan mediator putting events into a buffer.

6.1.2 Constructor & Destructor Documentation

6.1.2.1 `__init__()`

```
def plaso.tarzan.mediator.buffered_mediator.BufferedMediator.__init__ (
    self,
    buffer_queue = None )
```

Initialize the mediator including its queue if not given.

6.1.3 Member Function Documentation

6.1.3.1 flush_buffer()

```
def plaso.tarzan.mediator.buffered_mediator.BufferedMediator.flush_buffer (
    self,
    exception_on_error = False )
```

Flush the buffer into a list of events and return the list.
:param exception_on_error: True to raise an exception of an error event.
:return: the list

6.1.3.2 get_buffer_queue()

```
def plaso.tarzan.mediator.buffered_mediator.BufferedMediator.get_buffer_queue (
    self )
```

Get the queue of buffered events.
:return: the queue

6.1.3.3 ProduceEventWithEventData()

```
def plaso.tarzan.mediator.buffered_mediator.BufferedMediator.ProduceEventWithEventData (
    self,
    event,
    event_data )
```

Produce a particular event with its data.
:param event: the event
:param event_data: the event's data

6.1.3.4 ProduceExtractionError()

```
def plaso.tarzan.mediator.buffered_mediator.BufferedMediator.ProduceExtractionError (
    self,
    message,
    path_spec = None )
```

Produce an extraction error with a particular message and a path.
:param message: the error message
:param path_spec: the erroneous path

6.1.4 Member Data Documentation

6.1.4.1 buffer_queue

`plaso.tarzan.mediator.buffered_mediator.BufferedMediator.buffer_queue`

6.1.4.2 Error

`plaso.tarzan.mediator.buffered_mediator.BufferedMediator.Error`

6.1.4.3 Event

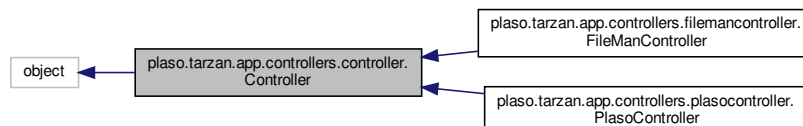
`plaso.tarzan.mediator.buffered_mediator.BufferedMediator.Event`

The documentation for this class was generated from the following file:

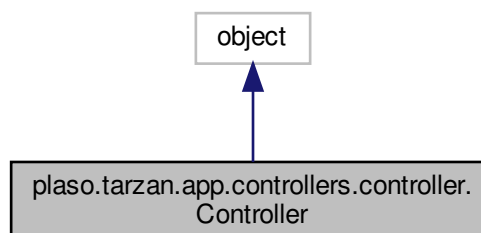
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered_mediator.py

6.2 plaso.tarzan.app.controllers.controller.Controller Class Reference

Inheritance diagram for `plaso.tarzan.app.controllers.controller.Controller`:



Collaboration diagram for `plaso.tarzan.app.controllers.controller.Controller`:



Public Member Functions

- def `__init__` (self, `hdfs_base_uri`)
- def `make_hdfs_uri` (self, `hdfs_path`)
- def `strip_hdfs_uri` (self, `hdfs_path`)

Public Attributes

- `hdfs_base_uri`

6.2.1 Detailed Description

Base-class for controllers.

6.2.2 Constructor & Destructor Documentation

6.2.2.1 `__init__()`

```
def plaso.tarzan.app.controllers.controller.Controller.__init__ (
    self,
    hdfs_base_uri )
```

Create a new controller that is able to store and utilize HDFS URI.
:param `hdfs_base_uri`: the base HDFS URI to store

6.2.3 Member Function Documentation

6.2.3.1 `make_hdfs_uri()`

```
def plaso.tarzan.app.controllers.controller.Controller.make_hdfs_uri (
    self,
    hdfs_path )
```

Get a full HDFS URI by adding a given path to the base HDFS URI.
:param `hdfs_path`: HDFS path
:return: HDFS URI

6.2.3.2 strip_hdfs_uri()

```
def plaso.tarzan.app.controllers.controller.Controller.strip_hdfs_uri (
    self,
    hdfs_path )
```

Get a given HDFS path without its HDFS URI prefix.
:param hdfs_path: the HDFS path
:return: the HDFS path without the HDFS URI prefix

6.2.4 Member Data Documentation

6.2.4.1 hdfs_base_uri

```
plaso.tarzan.app.controllers.controller.Controller.hdfs_base_uri
```

The documentation for this class was generated from the following file:

- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controller.py>

6.3 plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir Class Reference

Public Member Functions

- def [__init__](#) (self, [package_dir](#)=None)
- def [make_package_dir](#) (self)
- def [register_package_dir](#) (self)
- def [pip_install](#) (self, packages)
- def [install](#) (self, packages)

Static Public Member Functions

- def [check_installed_package](#) (package)

Public Attributes

- [package_dir](#)

6.3.1 Detailed Description

Python packages install class (not in a shell) to be executable by PySpark workers by
pyspark-driver:
 sparkContext.addPyFile("custom_pip_package_dir.py")
pyspark-worker:
 from custom_pip_package_dir import CustomPipPackageDir
 cppd = CustomPipPackageDir()
 cppd.install("future")

6.3.2 Constructor & Destructor Documentation

6.3.2.1 `__init__()`

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.__init__ (
    self,
    package_dir = None )
```

Initializes the Python package install class.

:param package_dir: a directory where to load packages from on the installation

6.3.3 Member Function Documentation

6.3.3.1 `check_installed_package()`

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.check_installed_package (
    package ) [static]
```

Check if the package is available.

:param package: the package to check

:return: True if the package is available, False otherwise

6.3.3.2 `install()`

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.install (
    self,
    packages )
```

Install the packages by external call of pip if they are not already installed.

:param packages: a list of packages to install

:return: the output buffer of the pip

6.3.3.3 `make_package_dir()`

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.make_package_dir (
    self )
```

Create a custom module storage.

6.3.3.4 pip_install()

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.pip_install (
    self,
    packages )
```

Install the packages by external call of pip.
 :param packages: a list of packages to install
 :return: the output buffer of the pip

6.3.3.5 register_package_dir()

```
def plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.register_package_dir (
    self )
```

Register the package directory as a custom module storage.

6.3.4 Member Data Documentation

6.3.4.1 package_dir

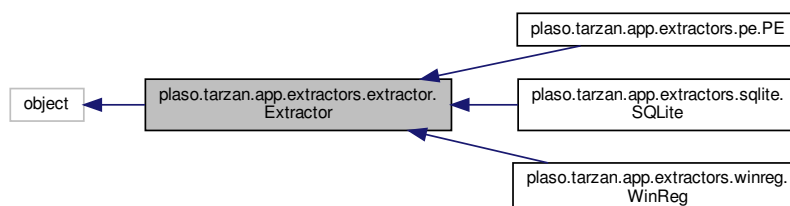
```
plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir.package_dir
```

The documentation for this class was generated from the following file:

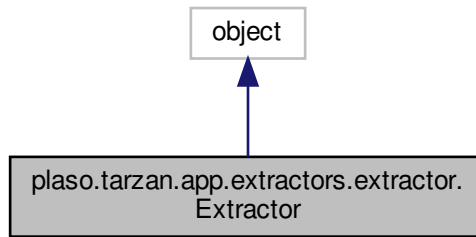
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom_pip_package_dir.py

6.4 plaso.tarzan.app.extractors.extractor.Extractor Class Reference

Inheritance diagram for plaso.tarzan.app.extractors.extractor.Extractor:



Collaboration diagram for plaso.tarzan.app.extractors.extractor.Extractor:



Public Member Functions

- def [get_name](#) (cls)
- def [get_description](#) (cls)
- def [filter](#) (cls, path)
- def [extract](#) (cls, path)

6.4.1 Detailed Description

Serve as a dummy extractor for its sub-classes.

6.4.2 Member Function Documentation

6.4.2.1 `extract()`

```
def plaso.tarzan.app.extractors.extractor.Extractor.extract (
    cls,
    path )
```

Extract events from a file given by its HDFS path.
:param path: the HDFS path of the file
:return: the list of events

6.4.2.2 filter()

```
def plaso.tarzan.app.extractors.extractor.Extractor.filter (
    cls,
    path )
```

Check if a file given by its HDFS path can be processed by this extractor.
:param path: the HDFS path of the file
:return: true iff the file can be processed

6.4.2.3 get_description()

```
def plaso.tarzan.app.extractors.extractor.Extractor.get_description (
    cls )
```

Get a description of the extractor.
:return: the description of the extractor

6.4.2.4 get_name()

```
def plaso.tarzan.app.extractors.extractor.Extractor.get_name (
    cls )
```

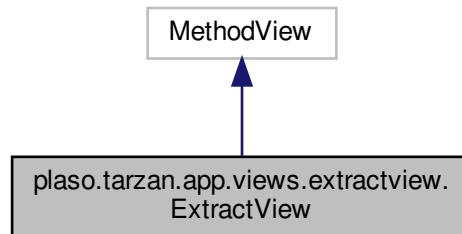
Get a name of the extractor.
:return: the name of the extractor

The documentation for this class was generated from the following file:

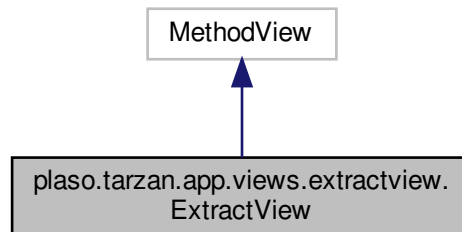
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/extractor.py>

6.5 plaso.tarzan.app.views.extractview.ExtractView Class Reference

Inheritance diagram for plaso.tarzan.app.views.extractview.ExtractView:



Collaboration diagram for plaso.tarzan.app.views.extractview.ExtractView:



Public Member Functions

- `def __init__(self, hdfs_base_uri, spark_context)`
- `def get(self, hdfs_path="")`

Public Attributes

- `controller`

6.5.1 Constructor & Destructor Documentation

6.5.1.1 `__init__()`

```
def plaso.tarzan.app.views.extractview.ExtractView.__init__ (
    self,
    hdfs_base_uri,
    spark_context )
```

6.5.2 Member Function Documentation

6.5.2.1 `get()`

```
def plaso.tarzan.app.views.extractview.ExtractView.get (
    self,
    hdfs_path = "" )
```

6.5.3 Member Data Documentation

6.5.3.1 `controller`

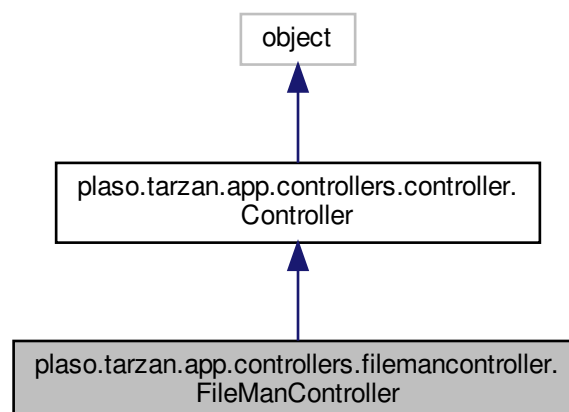
```
plaso.tarzan.app.views.extractview.ExtractView.controller
```

The documentation for this class was generated from the following file:

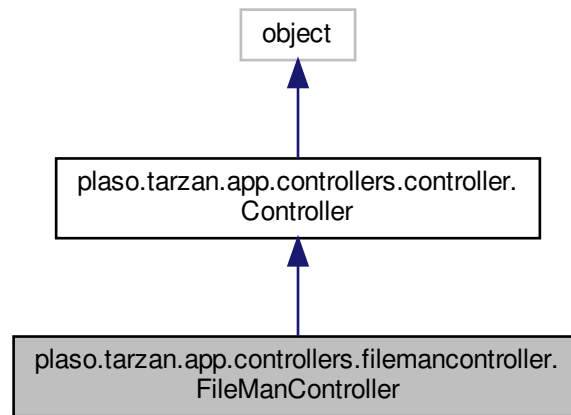
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractview.py>

6.6 `plaso.tarzan.app.controllers.filemancontroller.FileManController` Class Reference

Inheritance diagram for `plaso.tarzan.app.controllers.filemancontroller.FileManController`:



Collaboration diagram for plaso.tarzan.app.controllers.filemancontroller.FileManController:



Public Member Functions

- `def __init__ (self, hdfs_base_uri)`
- `def ls (self, hdfs_path= "")`
- `def rm (self, hdfs_path= "")`
- `def get (self, hdfs_path= "")`
- `def get_zip (self, hdfs_path= "")`
- `def put (self, request, hdfs_path)`
- `def put_form (self, request, hdfs_path)`
- `def zip_put (self, request, hdfs_path= "")`
- `def zip_put_form (self, request, hdfs_path= "")`

Public Attributes

- `hdfs`

6.6.1 Detailed Description

Controller for HDFS-related management operations.

6.6.2 Constructor & Destructor Documentation

6.6.2.1 `__init__()`

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.__init__ (
    self,
    hdfs_base_uri )
```

Create a new controller that is able to store and utilize HDFS URI in HDFS queries and operations.
:param hdfs_base_uri: the base HDFS URI to store

6.6.3 Member Function Documentation

6.6.3.1 `get()`

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.get (
    self,
    hdfs_path = "" )
```

Get a Flask Response to download a file in a given HDFS path.
:param hdfs_path: the HDFS path of the file to download
:return: the Flask Response of the download

6.6.3.2 `get_zip()`

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.get_zip (
    self,
    hdfs_path = "" )
```

Get a Flask Response to download a ZIP archive of files and directories in a given HDFS path.
:param hdfs_path: the HDFS path which include into the ZIP archive
:return: the Flask Response to download the ZIP file

6.6.3.3 `ls()`

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.ls (
    self,
    hdfs_path = "" )
```

Get a Flask Response listing all files and directories (the dirs are suffixed by "/") that are in a given HDFS path.
:param hdfs_path: the HDFS path to search for the files and directories
:return: the Flask Response of the list of files and directories in the path

6.6.3.4 put()

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.put (
    self,
    request,
    hdfs_path )
```

Copy the content uploaded by a PUT request in the REST API to a file in a given HDFS path.
:param request: the REST API PUT request
:param hdfs_path: the HDFS path where to upload
:return: the Flask Response confirming the upload

6.6.3.5 put_form()

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.put_form (
    self,
    request,
    hdfs_path )
```

Copy the content uploaded by a HTML FORM to a file in a given HDFS path.
:param request: the request which contains the FORM upload
:param hdfs_path: the HDFS path where to upload
:return: the Flask Response confirming the upload

6.6.3.6 rm()

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.rm (
    self,
    hdfs_path = "" )
```

Get a Flask Response confirming a recursive removal of files and directories in a given HDFS path.
:param hdfs_path: the HDFS path which should be removed including to remove its files and directories
:return: the Flask Response of the confirmation of the remove

6.6.3.7 zip_put()

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.zip_put (
    self,
    request,
    hdfs_path = "" )
```

Extract the content of a ZIP archive uploaded by a PUT request in the REST API into a given HDFS path.
:param request: the REST API PUT request
:param hdfs_path: the HDFS path where to extract the ZIP archive
:return: the Flask Response confirming the upload and extraction

6.6.3.8 zip_put_form()

```
def plaso.tarzan.app.controllers.filemancontroller.FileManController.zip_put_form (
    self,
    request,
    hdfs_path = "" )
```

Extract the content of a ZIP archive uploaded by a HTML FORM into a given HDFS path.
 :param request: the request which contains the FORM upload
 :param hdfs_path: the HDFS path where to extract the ZIP archive
 :return: the Flask Response confirming the upload and extraction

6.6.4 Member Data Documentation

6.6.4.1 hdfs

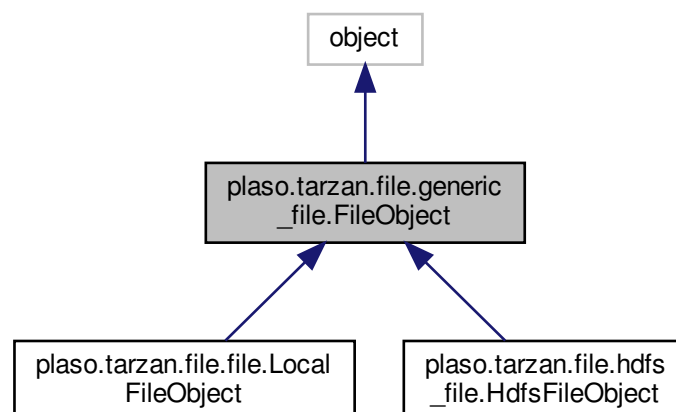
```
plaso.tarzan.app.controllers.filemancontroller.FileManController.hdfs
```

The documentation for this class was generated from the following file:

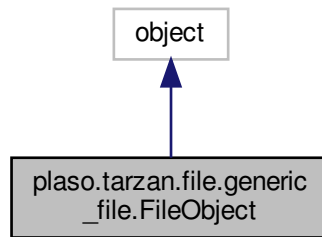
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/filemancontroller.py>

6.7 plaso.tarzan.file.generic_file.FileObject Class Reference

Inheritance diagram for plaso.tarzan.file.generic_file.FileObject:



Collaboration diagram for plaso.tarzan.file.generic_file.FileObject:



Public Member Functions

- def [close](#) (self)
- def [read](#) (self, size)
- def [seek](#) (self, offset, whence)

6.7.1 Detailed Description

Abstract class for a generic file.

6.7.2 Member Function Documentation

6.7.2.1 close()

```
def plaso.tarzan.file.generic_file.FileObject.close (
    self )
```

Close the file.

6.7.2.2 read()

```
def plaso.tarzan.file.generic_file.FileObject.read (
    self,
    size )
```

Reads a byte string from the file-like object at the current offset.
The function will read a byte string of the specified size or all of the remaining data if no size was specified.
:param size: number of bytes to read, where None is all remaining data
:return: data read

6.7.2.3 seek()

```
def plaso.tarzan.file.generic_file.FileObject.seek (
    self,
    offset,
    whence )
```

Set a position in the file for future reading or writing.

:param offset: the position as an offset

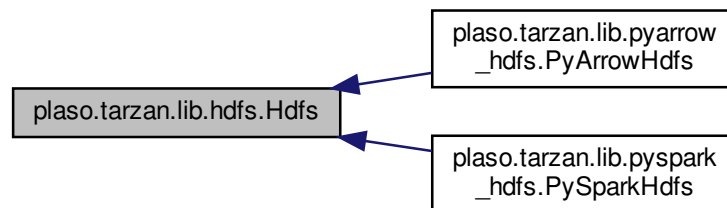
:param whence: from where the position should be reached (the beginning, the end, etc.)

The documentation for this class was generated from the following file:

- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/generic_file.py

6.8 plaso.tarzan.lib.hdfs.Hdfs Class Reference

Inheritance diagram for plaso.tarzan.lib.hdfs.Hdfs:



Public Member Functions

- def [make_uri](#) (self, filesystem=None)
- def [open_filesystem](#) (self, hdfs_uri, user)
- def [close_filesystem](#) (self, filesystem)
- def [get_filesystem](#) (self, force_filesystem)
- def [exists](#) (self, path_string, filesystem)
- def [make_path](#) (self, path_string, qualified=True, filesystem=None)
- def [make_simple_path](#) (self, path_string, filesystem=None)
- def [make_qualified_path](#) (self, path_string, filesystem=None)
- def [append_to_path](#) (self, original_path, new_child_string)
- def [basename](#) (self, path_string)
- def [dirname](#) (self, path_string)
- def [info](#) (self, path_string, filesystem=None)
- def [remove](#) (self, path_string, recursive=True, filesystem=None)
- def [mkdir](#) (self, path_string, filesystem=None)
- def [list_files](#) (self, path_string, recursion_level=0, include_dir_names=False, filesystem=None)

- def `open_inputstream` (self, path, filesystem)
- def `open_outputstream` (self, path, filesystem=None)
- def `close_stream` (self, input_stream)
- def `get_stream_offset` (self, input_stream)
- def `get_path_size` (self, path, filesystem)
- def `read_inputstream` (self, input_stream, size)
- def `write_outputstream` (self, output_stream, data)
- def `seek_stream` (self, stream, offset, whence=os.SEEK_SET)

Static Public Member Functions

- def `parse_uri` (hdfs_uri)

Static Public Attributes

- string `PATH_SEPARATOR` = '/'
- string `SCHEME` = 'hdfs'

6.8.1 Detailed Description

An abstract class for drivers to access HDFS.

6.8.2 Member Function Documentation

6.8.2.1 `append_to_path()`

```
def plaso.tarzan.lib.hdfs.Hdfs.append_to_path (
    self,
    original_path,
    new_child_string )
```

Append a path/directory/file into another HDFS path.
:param original_path: the another path
:param new_child_string: the path to append
:return: the resulting path

6.8.2.2 `basename()`

```
def plaso.tarzan.lib.hdfs.Hdfs.basename (
    self,
    path_string )
```

Get a basename from a given HDFS path.
:param path_string: the path
:return: the basename

6.8.2.3 close_filesystem()

```
def plaso.tarzan.lib.hdfs.Hdfs.close_filesystem (
    self,
    filesystem )
```

Close a given HDFS filesystem.
:param filesystem: the filesystem

6.8.2.4 close_stream()

```
def plaso.tarzan.lib.hdfs.Hdfs.close_stream (
    self,
    input_stream )
```

Close a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream

6.8.2.5 dirname()

```
def plaso.tarzan.lib.hdfs.Hdfs.dirname (
    self,
    path_string )
```

Get a dirname from a given HDFS path.
:param path_string: the path
:return: the dirname

6.8.2.6 exists()

```
def plaso.tarzan.lib.hdfs.Hdfs.exists (
    self,
    path_string,
    filesystem )
```

Check if a given HDFS path exists in the given HDFS filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: True if the path exists in the filesystem, False otherwise

6.8.2.7 get_filesystem()

```
def plaso.tarzan.lib.hdfs.Hdfs.get_filesystem (
    self,
    force_filesystem )
```

Get a given or a default HDFS filesystem.
:param force_filesystem: the given filesystem
:return: the filesystem

6.8.2.8 get_path_size()

```
def plaso.tarzan.lib.hdfs.Hdfs.get_path_size (
    self,
    path,
    filesystem )
```

Get the size of a given HDFS path (a file) in a given filesystem.
:param path: the path
:param filesystem: the filesystem
:return: the size

6.8.2.9 get_stream_offset()

```
def plaso.tarzan.lib.hdfs.Hdfs.get_stream_offset (
    self,
    input_stream )
```

Get the current position (an offset) in a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream
:return: the position/offset

6.8.2.10 info()

```
def plaso.tarzan.lib.hdfs.Hdfs.info (
    self,
    path_string,
    filesystem = None )
```

Get metadata of a given path in a given or a default filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: the metadata dictionary

6.8.2.11 list_files()

```
def plaso.tarzan.lib.hdfs.Hdfs.list_files (
    self,
    path_string,
    recursion_level = 0,
    include_dir_names = False,
    filesystem = None )
```

Get a list of files (optionally recursively to the specified level) in a given HDFS path of a given filesystem.

:param path_string: the path
:param recursion_level: the level for the recursion (0 is just a current directory without any recursion)
:param include_dir_names: True to include also names of directories (suffixed by /)
:param filesystem: the filesystem
:return: the list of files in the path

6.8.2.12 make_path()

```
def plaso.tarzan.lib.hdfs.Hdfs.make_path (
    self,
    path_string,
    qualified = True,
    filesystem = None )
```

Get a (qualified) HDFS URI from a given path and a given or a default filesystem.

:param path_string: the path
:param qualified: True to get the qualified path
:param filesystem: the filesystem
:return: the resulting path

6.8.2.13 make_qualified_path()

```
def plaso.tarzan.lib.hdfs.Hdfs.make_qualified_path (
    self,
    path_string,
    filesystem = None )
```

Get a qualified HDFS URI from a given path and a given or a default filesystem.

:param path_string: the path
:param filesystem: the filesystem
:return: the resulting path

6.8.2.14 make_simple_path()

```
def plaso.tarzan.lib.hdfs.Hdfs.make_simple_path (
    self,
    path_string,
    filesystem = None )
```

Get an unqualified HDFS URI from a given path and a given or a default filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: the resulting path

6.8.2.15 make_uri()

```
def plaso.tarzan.lib.hdfs.Hdfs.make_uri (
    self,
    filesystem = None )
```

Get HDFS URI of a given or a default filesystem.
:param filesystem: the filesystem
:return: the HDFS URI

6.8.2.16 mkdir()

```
def plaso.tarzan.lib.hdfs.Hdfs.mkdir (
    self,
    path_string,
    filesystem = None )
```

Make a new directory (if does not exist) of a given path in a given or a default filesystem.
:param path_string: the path
:param filesystem: the filesystem

6.8.2.17 open_filesystem()

```
def plaso.tarzan.lib.hdfs.Hdfs.open_filesystem (
    self,
    hdfs_uri,
    user )
```

Open HDFS filesystem of a given URI and as a given HDFS user.
:param hdfs_uri: HDFS URI to open
:param user: HDFS user to act as when opening
:return: the opened filesystem

6.8.2.18 open_inputstream()

```
def plaso.tarzan.lib.hdfs.Hdfs.open_inputstream (
    self,
    path,
    filesystem )
```

Open and get an input-stream for a HDFS file given by its path in a given filesystem.
:param path: the path
:param filesystem: the filesystem
:return: the opened input-stream

6.8.2.19 open_outputstream()

```
def plaso.tarzan.lib.hdfs.Hdfs.open_outputstream (
    self,
    path,
    filesystem = None )
```

Open and get an output-stream for a HDFS file given by its path in a given filesystem.
:param path: the path
:param filesystem: the filesystem
:return: the opened output-stream

6.8.2.20 parse_uri()

```
def plaso.tarzan.lib.hdfs.Hdfs.parse_uri (
    hdfs_uri ) [static]
```

Parse HDFS URI into individual components.
:param hdfs_uri: HDFS URI of the file to parse
:return: an object that provide individual component of the URI

6.8.2.21 read_inputstream()

```
def plaso.tarzan.lib.hdfs.Hdfs.read_inputstream (
    self,
    input_stream,
    size )
```

Read data form a given opened input stream for a HDFS file.
:param input_stream: the input-stream
:param size: the size of data to read
:return: the data

6.8.2.22 remove()

```
def plaso.tarzan.lib.hdfs.Hdfs.remove (
    self,
    path_string,
    recursive = True,
    filesystem = None )
```

Remove (optionally recursively) a HDFS file/directory given by its path in a given or a default filesystem.
:param path_string: the path
:param recursive: True to remove recursively
:param filesystem: the filesystem

6.8.2.23 seek_stream()

```
def plaso.tarzan.lib.hdfs.Hdfs.seek_stream (
    self,
    stream,
    offset,
    whence = os.SEEK_SET )
```

Set a given position (an offset) in a given (previously opened) input-stream for a HDFS file.
:param stream: the opened stream
:param offset: the position/offset
:param whence: the direction

6.8.2.24 write_outputstream()

```
def plaso.tarzan.lib.hdfs.Hdfs.write_outputstream (
    self,
    output_stream,
    data )
```

Write data buffer to a given opened output stream for a HDFS file.
:param output_stream: the output-stream
:param data: the data buffer to write
:return: the number of bytes written

6.8.3 Member Data Documentation

6.8.3.1 PATH_SEPARATOR

```
string plaso.tarzan.lib.hdfs.Hdfs.PATH_SEPARATOR = '/' [static]
```

6.8.3.2 SCHEME

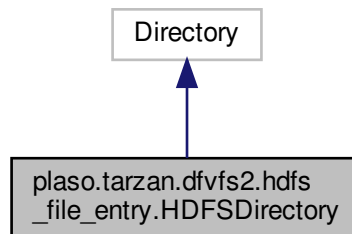
```
string plaso.tarzan.lib.hdfs.Hdfs.SCHEME = 'hdfs' [static]
```

The documentation for this class was generated from the following file:

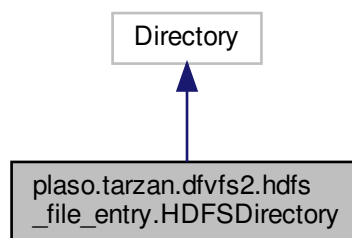
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/hdfs.py](#)

6.9 plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory Class Reference

Inheritance diagram for plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory:



6.9.1 Detailed Description

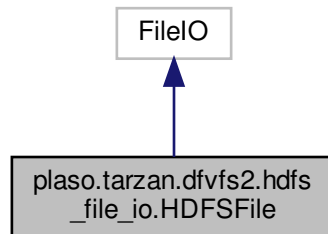
HDFS directory.

The documentation for this class was generated from the following file:

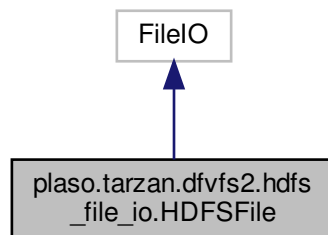
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_entry.py](#)

6.10 plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile Class Reference

Inheritance diagram for plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile:



Public Member Functions

- def `__init__` (self, resolver_context)
- def `read` (self, size=None)
- def `seek` (self, offset, whence=os.SEEK_SET)
- def `get_offset` (self)
- def `get_size` (self)

6.10.1 Detailed Description

File-like object for HDFS.

6.10.2 Constructor & Destructor Documentation

6.10.2.1 `__init__()`

```
def plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile.__init__ (
    self,
    resolver_context )
```

Initializes a file-like object.
:param resolver_context: resolver context

6.10.3 Member Function Documentation

6.10.3.1 `get_offset()`

```
def plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile.get_offset (
    self )
```

Retrieves the current offset into the file-like object.
:return: current offset into the file-like object
:raises: IOError: if the file-like object has not been opened

6.10.3.2 `get_size()`

```
def plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile.get_size (
    self )
```

Retrieves the size of the file-like object.
:return: size of the file-like object data
:raises IOError: if the file-like object has not been opened

6.10.3.3 `read()`

```
def plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile.read (
    self,
    size = None )
```

Reads a byte string from the file-like object at the current offset.
The function will read a byte string of the specified size or all of the remaining data if no size was specified.
:param size: number of bytes to read, where None is all remaining data.
:return: data read
:raises IOError: if the read failed

6.10.3.4 seek()

```
def plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile.seek (
    self,
    offset,
    whence = os.SEEK_SET )
```

Seeks to an offset within the file-like object.

:param offset: offset to seek to

:param whence: value that indicates whether offset is an absolute or relative position within the file

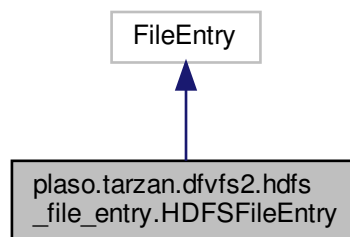
:raises IOError: if the seek failed

The documentation for this class was generated from the following file:

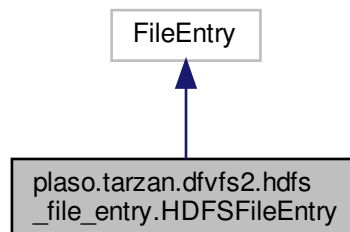
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_io.py

6.11 plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry Class Reference

Inheritance diagram for plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry:



Public Member Functions

- def `__init__` (self, resolver_context, file_system, path_spec, is_root=False)
- def `access_time` (self)
- def `change_time` (self)
- def `link` (self)
- def `name` (self)
- def `modification_time` (self)
- def `GetLinkedFileEntry` (self)
- def `GetParentFileEntry` (self)

Public Attributes

- `entry_type`

Static Public Attributes

- `TYPE_INDICATOR` = `definitions.TYPE_INDICATOR_HDFS`

6.11.1 Detailed Description

HDFS file entry.

6.11.2 Constructor & Destructor Documentation

6.11.2.1 `__init__`()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.__init__ (
    self,
    resolver_context,
    file_system,
    path_spec,
    is_root = False )
```

Initializes a file entry.

:param resolver_context: resolver context

:param file_system: file system

:param path_spec: path specification

:param is_root: True if the file entry is the root file entry of the corresponding file system

:raises BackEndError: If there is a HDFS backend error

6.11.3 Member Function Documentation

6.11.3.1 access_time()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.access_time (
    self )
```

Access time or None if not available.

6.11.3.2 change_time()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.change_time (
    self )
```

Change time or None if not available.

6.11.3.3 GetLinkedFileEntry()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.GetLinkedFileEntry (
    self )
```

Retrieves the linked file entry, for example for a symbolic link.
:return: linked file entry or None if not available

6.11.3.4 GetParentFileEntry()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.GetParentFileEntry (
    self )
```

Retrieves the parent file entry.
:return: parent file entry or None if not available

6.11.3.5 link()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.link (
    self )
```

Full path of the linked file entry.

6.11.3.6 modification_time()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.modification_time (
    self )
```

Modification time or None if not available.

6.11.3.7 name()

```
def plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.name (
    self )
```

Name of the file entry, without the full path.

6.11.4 Member Data Documentation

6.11.4.1 entry_type

```
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.entry_type
```

6.11.4.2 TYPE_INDICATOR

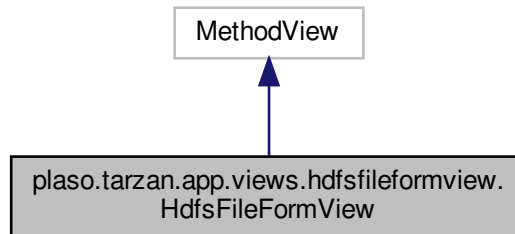
```
plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry.TYPE_INDICATOR = definitions.TYPE_INDICATOR_←  
R_HDFS [static]
```

The documentation for this class was generated from the following file:

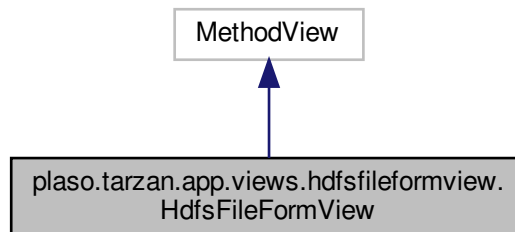
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_entry.py](#)

6.12 plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView Class Reference

Inheritance diagram for plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView:



Collaboration diagram for plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView:



Public Member Functions

- def `__init__` (self, hdfs_base_uri)
- def `get` (self, hdfs_path="")
- def `post` (self, hdfs_path="")
- def `delete` (self, hdfs_path="")
- def `put` (self, hdfs_path="")

Public Attributes

- `controller`

6.12.1 Constructor & Destructor Documentation

6.12.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.__init__ (
    self,
    hdfs_base_uri )
```

6.12.2 Member Function Documentation

6.12.2.1 `delete()`

```
def plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.delete (
    self,
    hdfs_path = "" )
```

6.12.2.2 `get()`

```
def plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.get (
    self,
    hdfs_path = "" )
```

6.12.2.3 `post()`

```
def plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.post (
    self,
    hdfs_path = "" )
```

6.12.2.4 `put()`

```
def plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.put (
    self,
    hdfs_path = "" )
```

6.12.3 Member Data Documentation

6.12.3.1 controller

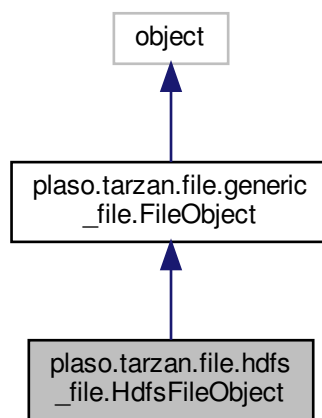
```
plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView.controller
```

The documentation for this class was generated from the following file:

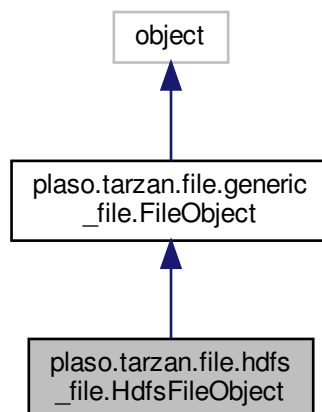
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformview.py>

6.13 plaso.tarzan.file.hdfs_file.HdfsFileObject Class Reference

Inheritance diagram for plaso.tarzan.file.hdfs_file.HdfsFileObject:



Collaboration diagram for plaso.tarzan.file.hdfs_file.HdfsFileObject:



Public Member Functions

- def `__init__` (self)
- def `open` (self, hdfs_uri, path_spec=None)
- def `close` (self)
- def `read` (self, size=None)
- def `seek` (self, offset, whence=os.SEEK_SET)
- def `tell` (self)

Public Attributes

- `hdfs`
- `path`
- `input_stream`

6.13.1 Detailed Description

HDFS file access.

6.13.2 Constructor & Destructor Documentation

6.13.2.1 `__init__()`

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.__init__ (  
    self )
```

Initializes a HDFS file object.

6.13.3 Member Function Documentation

6.13.3.1 `close()`

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.close (  
    self )
```

Close the file.

6.13.3.2 open()

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.open (
    self,
    hdfs_uri,
    path_spec = None )
```

Open a HDFS file identified by the HDFS URI and optionally also by its path.
:param hdfs_uri: the HDFS URI of the file to open
:param path_spec: the path of the file to open

6.13.3.3 read()

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.read (
    self,
    size = None )
```

Reads a byte string from the file-like object at the current offset.
The function will read a byte string of the specified size or all of the remaining data if no size was specified.
:param size: number of bytes to read, where None is all remaining data
:return: data read

6.13.3.4 seek()

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.seek (
    self,
    offset,
    whence = os.SEEK_SET )
```

Set a position in the file for future reading or writing.
:param offset: the position as an offset
:param whence: from where the position should be reached (the beginning, the end, etc.)

6.13.3.5 tell()

```
def plaso.tarzan.file.hdfs_file.HdfsFileObject.tell (
    self )
```

Get a current position in the file.
:return: the current position (an offset)

6.13.4 Member Data Documentation

6.13.4.1 hdfs

```
plaso.tarzan.file.hdfs_file.HdfsFileObject.hdfs
```

6.13.4.2 input_stream

```
plaso.tarzan.file.hdfs_file.HdfsFileObject.input_stream
```

6.13.4.3 path

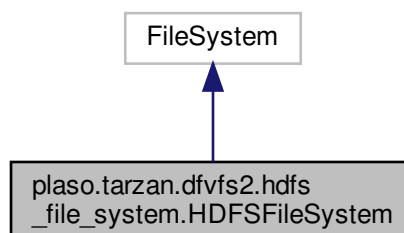
```
plaso.tarzan.file.hdfs_file.HdfsFileObject.path
```

The documentation for this class was generated from the following file:

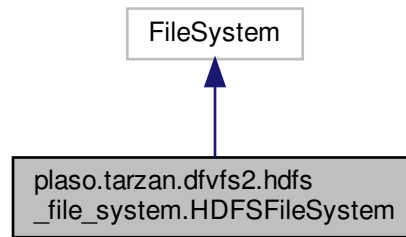
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/hdfs_file.py

6.14 plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem Class Reference

Inheritance diagram for plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem:



Public Member Functions

- def `__init__` (self, resolver_context)
- def `FileEntryExistsByPathSpec` (self, path_spec)
- def `GetFileEntryByPathSpec` (self, path_spec)
- def `GetRootFileEntry` (self)
- def `BasenamePath` (self, path)
- def `DirnamePath` (self, path)
- def `JoinPath` (self, path_segments, qualified=False)
- def `SplitPath` (self, path, qualified=False)

Public Attributes

- `hdfs`

Static Public Attributes

- `LOCATION_ROOT` = `Hdfs.PATH_SEPARATOR`
- `PATH_SEPARATOR` = `Hdfs.PATH_SEPARATOR`
- `TYPE_INDICATOR` = `definitions.TYPE_INDICATOR_HDFS`

6.14.1 Detailed Description

HDFS file system.

6.14.2 Constructor & Destructor Documentation

6.14.2.1 `__init__()`

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.__init__ (
    self,
    resolver_context )
```

Initializes a file system.

:param resolver_context: resolver context

:raises ValueError: if a derived file system class does not define a type indicator

6.14.3 Member Function Documentation

6.14.3.1 `BasenamePath()`

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.BasenamePath (
    self,
    path )
```

Determines the basename of the path.

:param path: path

:return: basename of the path

6.14.3.2 `DirnamePath()`

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.DirnamePath (
    self,
    path )
```

Determines the directory name of the path. The file system root is represented by an empty string.

:param path: path

:return: directory name of the path or None

6.14.3.3 `FileEntryExistsByPathSpec()`

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.FileEntryExistsByPathSpec (
    self,
    path_spec )
```

Determines if a file entry for a path specification exists.

:param path_spec: a path specification

:return: True if the file entry exists, false otherwise

6.14.3.4 GetFileEntryByPathSpec()

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.GetFileEntryByPathSpec (
    self,
    path_spec )
```

Retrieves a file entry for a path specification.
:param path_spec: a path specification
:return: a file entry or None if not available

6.14.3.5 GetRootFileEntry()

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.GetRootFileEntry (
    self )
```

Retrieves the root file entry.
:return: a file entry or None if not available

6.14.3.6 JoinPath()

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.JoinPath (
    self,
    path_segments,
    qualified = False )
```

Joins the path segments into a path.
:param path_segments: path segments where the first can be a fully qualified HDFS URI
:param qualified: if the resulting path should be a fully qualified HDFS URI
:return: joined path segments prefixed with the path separator

6.14.3.7 SplitPath()

```
def plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.SplitPath (
    self,
    path,
    qualified = False )
```

Splits the path into path segments.
:param path: path
:param qualified: if the first segment should be an HDFS URI
:return: path segments without the root path segment, which is an empty string.

6.14.4 Member Data Documentation

6.14.4.1 hdfs

```
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.hdfs
```

6.14.4.2 LOCATION_ROOT

```
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.LOCATION_ROOT = Hdfs.PATH_SEPARATOR [static]
```

6.14.4.3 PATH_SEPARATOR

```
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.PATH_SEPARATOR = Hdfs.PATH_SEPARATOR [static]
```

6.14.4.4 TYPE_INDICATOR

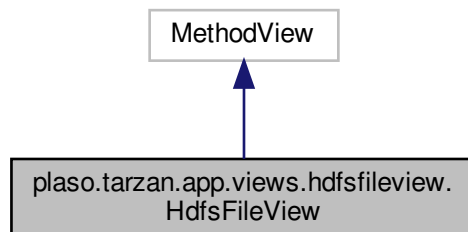
```
plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem.TYPE_INDICATOR = definitions.TYPE_INDICATOR_HDFS [static]
```

The documentation for this class was generated from the following file:

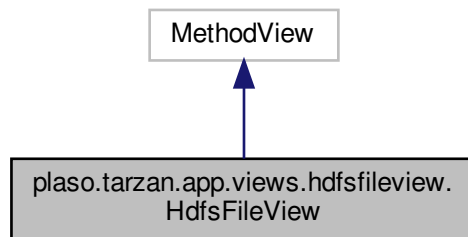
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_system.py](#)

6.15 plaso.tarzan.app.views.hdfsfileview.HdfsFileView Class Reference

Inheritance diagram for plaso.tarzan.app.views.hdfsfileview.HdfsFileView:



Collaboration diagram for plaso.tarzan.app.views.hdfsfileview.HdfsFileView:



Public Member Functions

- def `__init__` (self, hdfs_base_uri)
- def `get` (self, hdfs_path="")
- def `post` (self, hdfs_path="")
- def `delete` (self, hdfs_path="")
- def `put` (self, hdfs_path="")

Public Attributes

- `controller`

6.15.1 Constructor & Destructor Documentation

6.15.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfsfileview.HdfsFileView.__init__ (
    self,
    hdfs_base_uri )
```

6.15.2 Member Function Documentation

6.15.2.1 `delete()`

```
def plaso.tarzan.app.views.hdfsfileview.HdfsFileView.delete (
    self,
    hdfs_path = "" )
```

6.15.2.2 get()

```
def plaso.tarzan.app.views.hdfsfileview.HdfsFileView.get (
    self,
    hdfs_path = "" )
```

6.15.2.3 post()

```
def plaso.tarzan.app.views.hdfsfileview.HdfsFileView.post (
    self,
    hdfs_path = "" )
```

6.15.2.4 put()

```
def plaso.tarzan.app.views.hdfsfileview.HdfsFileView.put (
    self,
    hdfs_path = "" )
```

6.15.3 Member Data Documentation

6.15.3.1 controller

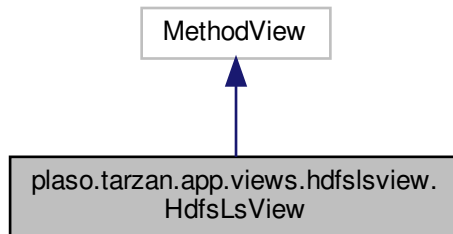
```
plaso.tarzan.app.views.hdfsfileview.HdfsFileView.controller
```

The documentation for this class was generated from the following file:

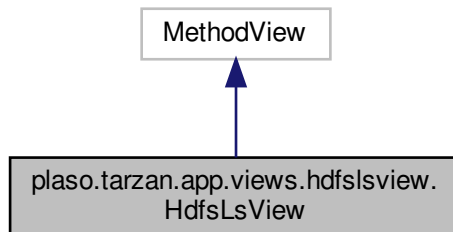
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.py](#)

6.16 plaso.tarzan.app.views.hdfsview.HdfsLsView Class Reference

Inheritance diagram for plaso.tarzan.app.views.hdfsview.HdfsLsView:



Collaboration diagram for plaso.tarzan.app.views.hdfsview.HdfsLsView:



Public Member Functions

- `def __init__(self, hdfs_base_uri)`
- `def get(self, hdfs_path="")`

Public Attributes

- `controller`

6.16.1 Constructor & Destructor Documentation

6.16.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfsview.HdfsLsView.__init__ (
    self,
    hdfs_base_uri )
```

6.16.2 Member Function Documentation

6.16.2.1 `get()`

```
def plaso.tarzan.app.views.hdfsview.HdfsLsView.get (
    self,
    hdfs_path = "" )
```

6.16.3 Member Data Documentation

6.16.3.1 `controller`

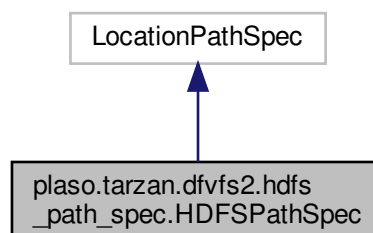
```
plaso.tarzan.app.views.hdfsview.HdfsLsView.controller
```

The documentation for this class was generated from the following file:

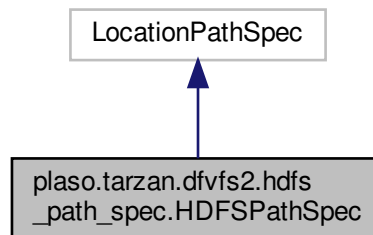
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsview.py>

6.17 `plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec` Class Reference

Inheritance diagram for `plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec`:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec:



Public Member Functions

- `def __init__(self, location=None, kwargs)`

Static Public Attributes

- `TYPE_INDICATOR = definitions.TYPE_INDICATOR_HDFS`

6.17.1 Detailed Description

HDFS path specification.

6.17.2 Constructor & Destructor Documentation

6.17.2.1 __init__()

```
def plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec.__init__(
    self,
    location = None,
    kwargs )
```

Initializes a path specification. Note that the HDFS path specification cannot have a parent.
:param location: HDFS specific location URI string, e.g., "hdfs://hadoop@namenode:8020/test_data"
:raises ValueError: when location is not set or parent is set

6.17.3 Member Data Documentation

6.17.3.1 TYPE_INDICATOR

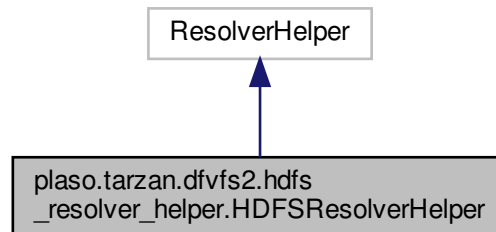
```
plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec.TYPE_INDICATOR = definitions.TYPE_INDICATOR_←
HDFS [static]
```

The documentation for this class was generated from the following file:

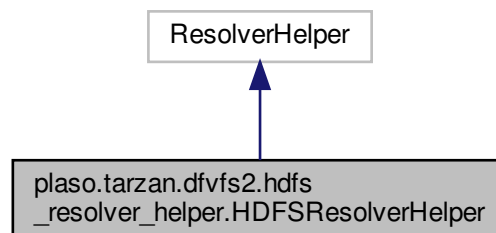
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_path_spec.py

6.18 plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper Class Reference

Inheritance diagram for plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper:



Collaboration diagram for plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper:



Public Member Functions

- def [NewFileObject](#) (self, resolver_context)
- def [NewFileSystem](#) (self, resolver_context)

Static Public Attributes

- `TYPE_INDICATOR` = `definitions.TYPE_INDICATOR_HDFS`

6.18.1 Detailed Description

HDFS resolver helper.

6.18.2 Member Function Documentation

6.18.2.1 NewFileObject()

```
def plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper.NewFileObject (
    self,
    resolver_context )
```

Creates a new file-like object.
:param resolver_context: resolver context
:return: file-like object

6.18.2.2 NewFileSystem()

```
def plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper.NewFileSystem (
    self,
    resolver_context )
```

Creates a new file system object.
:param resolver_context: resolver context
:return:: file system

6.18.3 Member Data Documentation

6.18.3.1 TYPE_INDICATOR

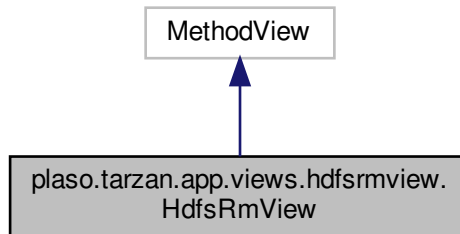
```
plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper.TYPE_INDICATOR = definitions.TYP↔  
E\_INDICATOR\_HDFS [static]
```

The documentation for this class was generated from the following file:

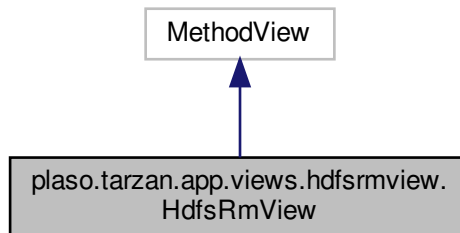
- `/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_resolver_helper.py`

6.19 plaso.tarzan.app.views.hdfsrmview.HdfsRmView Class Reference

Inheritance diagram for plaso.tarzan.app.views.hdfsrmview.HdfsRmView:



Collaboration diagram for plaso.tarzan.app.views.hdfsrmview.HdfsRmView:



Public Member Functions

- def `__init__` (self, hdfs_base_uri)
- def `get` (self, hdfs_path="")
- def `delete` (self, hdfs_path="")

Public Attributes

- `controller`

6.19.1 Constructor & Destructor Documentation

6.19.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfsrmview.HdfsRmView.__init__ (
    self,
    hdfs_base_uri )
```

6.19.2 Member Function Documentation

6.19.2.1 `delete()`

```
def plaso.tarzan.app.views.hdfsrmview.HdfsRmView.delete (
    self,
    hdfs_path = "" )
```

6.19.2.2 `get()`

```
def plaso.tarzan.app.views.hdfsrmview.HdfsRmView.get (
    self,
    hdfs_path = "" )
```

6.19.3 Member Data Documentation

6.19.3.1 `controller`

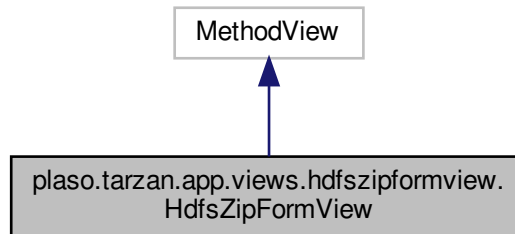
```
plaso.tarzan.app.views.hdfsrmview.HdfsRmView.controller
```

The documentation for this class was generated from the following file:

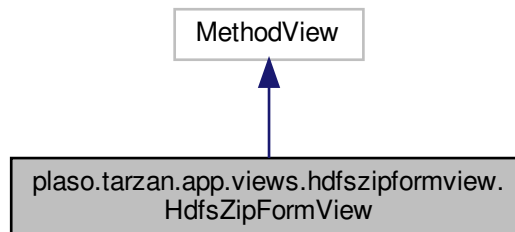
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsrmview.py](#)

6.20 plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView Class Reference

Inheritance diagram for plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView:



Collaboration diagram for plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView:



Public Member Functions

- def `__init__` (self, hdfs_base_uri)
- def `get` (self, hdfs_path="")
- def `post` (self, hdfs_path="")
- def `delete` (self, hdfs_path="")
- def `put` (self, hdfs_path="")

Public Attributes

- `controller`

6.20.1 Constructor & Destructor Documentation

6.20.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.__init__ (
    self,
    hdfs_base_uri )
```

6.20.2 Member Function Documentation

6.20.2.1 `delete()`

```
def plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.delete (
    self,
    hdfs_path = "" )
```

6.20.2.2 `get()`

```
def plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.get (
    self,
    hdfs_path = "" )
```

6.20.2.3 `post()`

```
def plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.post (
    self,
    hdfs_path = "" )
```

6.20.2.4 `put()`

```
def plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.put (
    self,
    hdfs_path = "" )
```

6.20.3 Member Data Documentation

6.20.3.1 controller

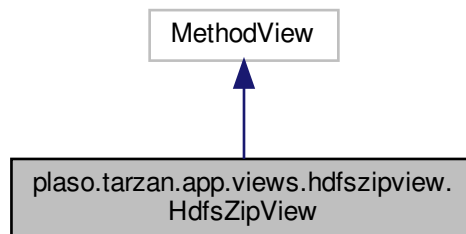
`plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView.controller`

The documentation for this class was generated from the following file:

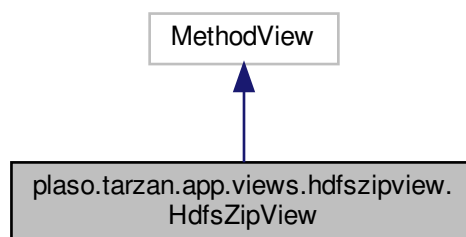
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipformview.py>

6.21 plaso.tarzan.app.views.hdfszipview.HdfsZipView Class Reference

Inheritance diagram for `plaso.tarzan.app.views.hdfszipview.HdfsZipView`:



Collaboration diagram for `plaso.tarzan.app.views.hdfszipview.HdfsZipView`:



Public Member Functions

- `def __init__(self, hdfs_base_uri)`
- `def get(self, hdfs_path='')`
- `def post(self, hdfs_path='')`
- `def delete(self, hdfs_path='')`
- `def put(self, hdfs_path='')`

Public Attributes

- [controller](#)

6.21.1 Constructor & Destructor Documentation

6.21.1.1 `__init__()`

```
def plaso.tarzan.app.views.hdfszipview.HdfsZipView.__init__ (
    self,
    hdfs_base_uri )
```

6.21.2 Member Function Documentation

6.21.2.1 `delete()`

```
def plaso.tarzan.app.views.hdfszipview.HdfsZipView.delete (
    self,
    hdfs_path = "" )
```

6.21.2.2 `get()`

```
def plaso.tarzan.app.views.hdfszipview.HdfsZipView.get (
    self,
    hdfs_path = "" )
```

6.21.2.3 `post()`

```
def plaso.tarzan.app.views.hdfszipview.HdfsZipView.post (
    self,
    hdfs_path = "" )
```

6.21.2.4 `put()`

```
def plaso.tarzan.app.views.hdfszipview.HdfsZipView.put (
    self,
    hdfs_path = "" )
```

6.21.3 Member Data Documentation

6.21.3.1 controller

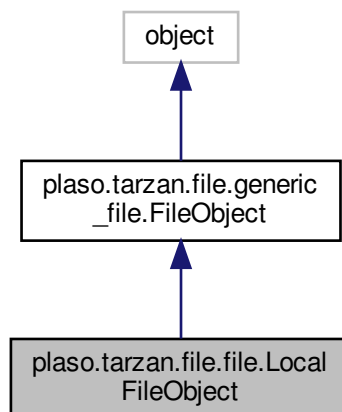
```
plaso.tarzan.app.views.hdfszipview.HdfsZipView.controller
```

The documentation for this class was generated from the following file:

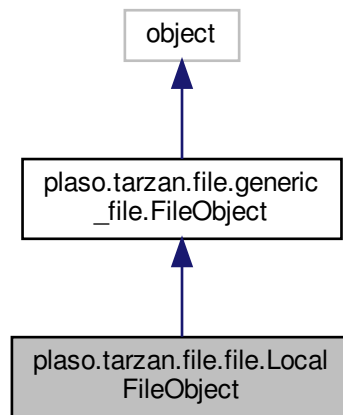
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipview.py>

6.22 plaso.tarzan.file.file.LocalFileObject Class Reference

Inheritance diagram for plaso.tarzan.file.file.LocalFileObject:



Collaboration diagram for plaso.tarzan.file.file.LocalFileObject:



Public Member Functions

- def `__init__` (self)
- def `open` (self, path_spec=None, mode='rb')
- def `close` (self)
- def `read` (self, size=None)
- def `seek` (self, offset, whence=os.SEEK_SET)
- def `tell` (self)

Public Attributes

- `path`
- `file_object`

6.22.1 Detailed Description

Local file access.

6.22.2 Constructor & Destructor Documentation

6.22.2.1 `__init__()`

```
def plaso.tarzan.file.file.LocalFileObject.__init__(  
    self )
```

Initializes a local file object.

6.22.3 Member Function Documentation

6.22.3.1 close()

```
def plaso.tarzan.file.file.LocalFileObject.close (
    self )
```

Close the file.

6.22.3.2 open()

```
def plaso.tarzan.file.file.LocalFileObject.open (
    self,
    path_spec = None,
    mode = 'rb' )
```

Open a local file identified by its path and in a particular mode.

:param path_spec: the path of the file to open

:param mode: the mode of the file opening (and its future usage, i.e., reading, writing, etc.)

6.22.3.3 read()

```
def plaso.tarzan.file.file.LocalFileObject.read (
    self,
    size = None )
```

Reads a byte string from the file-like object at the current offset.

The function will read a byte string of the specified size or all of the remaining data if no size was specified.

:param size: number of bytes to read, where None is all remaining data

:return: data read

:raises IOError: if the read failed

:raises OSError: if the read failed

6.22.3.4 seek()

```
def plaso.tarzan.file.file.LocalFileObject.seek (
    self,
    offset,
    whence = os.SEEK_SET )
```

Set a position in the file for future reading or writing.

:param offset: the position as an offset

:param whence: from where the position should be reached (the beginning, the end, etc.)

6.22.3.5 tell()

```
def plaso.tarzan.file.file.LocalFileObject.tell (
    self )
```

Get a current position in the file.
:return: the current position (an offset)

6.22.4 Member Data Documentation

6.22.4.1 file_object

```
plaso.tarzan.file.file.LocalFileObject.file_object
```

6.22.4.2 path

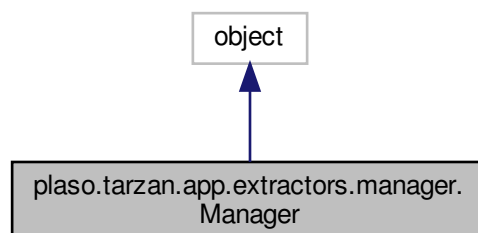
```
plaso.tarzan.file.file.LocalFileObject.path
```

The documentation for this class was generated from the following file:

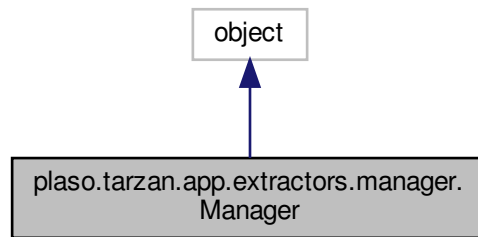
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/file.py>

6.23 plaso.tarzan.app.extractors.manager.Manager Class Reference

Inheritance diagram for plaso.tarzan.app.extractors.manager.Manager:



Collaboration diagram for `plaso.tarzan.app.extractors.manager.Manager`:



Public Member Functions

- def [get_extractors](#) (cls)
- def [deregister_extractor](#) (cls, extractor_class)
- def [register_extractor](#) (cls, extractor_class)

6.23.1 Detailed Description

Manager of extractors.

6.23.2 Member Function Documentation

6.23.2.1 `deregister_extractor()`

```
def plaso.tarzan.app.extractors.manager.Manager.deregister_extractor (
    cls,
    extractor_class )
```

De-registers an extractor class.

:param extractor_class: extractor class (subclass of `Extractor`)

:raise `KeyError`: if extractor class is not set for the corresponding name

6.23.2.2 `get_extractors()`

```
def plaso.tarzan.app.extractors.manager.Manager.get_extractors (
    cls )
```

Get a list of registered extractor classes.

:return: the list of extractor classes

6.23.2.3 register_extractor()

```
def plaso.tarzan.app.extractors.manager.Manager.register_extractor (
    cls,
    extractor_class )
```

Registers an extractor class.

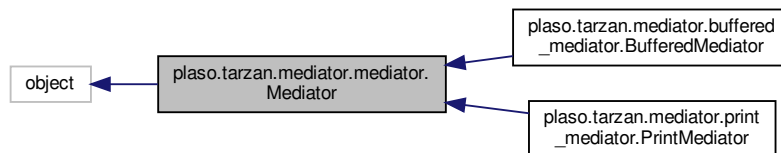
```
:param extractor_class: extractor class (subclass of Extractor)
:raise KeyError: if extractor class is already set for the corresponding name
```

The documentation for this class was generated from the following file:

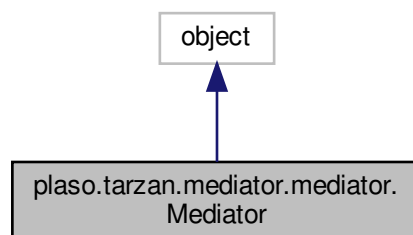
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/manager.py](#)

6.24 plaso.tarzan.mediator.mediator.Mediator Class Reference

Inheritance diagram for plaso.tarzan.mediator.mediator.Mediator:



Collaboration diagram for plaso.tarzan.mediator.mediator.Mediator:



Public Member Functions

- def `__init__` (self)
- def `ProduceEventWithEventData` (self, event, event_data)
- def `ProduceExtractionError` (self, message, path_spec)
- def `SetFileEntry` (self, file_entry)
- def `GetFileEntry` (self)
- def `GetFilename` (self)
- def `AppendToParserChain` (self, plugin_or_parser)
- def `PopFromParserChain` (self)
- def `AddEventAttribute` (self, attribute_name, attribute_value)
- def `RemoveEventAttribute` (self, attribute_name)
- def `abort` (self)
- def `temporary_directory` (self)

Public Attributes

- `file_entry`
- `parser_chain_components`
- `collection_filters_helper`
- `codepage`

6.24.1 Detailed Description

An abstract base-class for Tarzan mediators.

6.24.2 Constructor & Destructor Documentation

6.24.2.1 `__init__()`

```
def plaso.tarzan.mediator.mediator.Mediator.__init__ (
    self )
```

Initialize the mediator.

6.24.3 Member Function Documentation

6.24.3.1 `abort()`

```
def plaso.tarzan.mediator.mediator.Mediator.abort (
    self )
```

Check if the parsing should be aborted.
:return: True if parsing should be aborted

6.24.3.2 AddEventAttribute()

```
def plaso.tarzan.mediator.mediator.Mediator.AddEventAttribute (
    self,
    attribute_name,
    attribute_value )
```

Adds an attribute that will be set on all events produced. Setting attributes using this method will cause events produced via this mediator to have an attribute with the provided name set with the provided value.

:param attribute_name: name of the attribute to add
:param attribute_value: value of the attribute to add
:raises KeyError: if the event attribute is already set

6.24.3.3 AppendToParserChain()

```
def plaso.tarzan.mediator.mediator.Mediator.AppendToParserChain (
    self,
    plugin_or_parser )
```

Adds a parser or parser plugin to the parser chain.

:param plugin_or_parser: parser or parser plugin

6.24.3.4 GetFileEntry()

```
def plaso.tarzan.mediator.mediator.Mediator.GetFileEntry (
    self )
```

Retrieves the active file entry.

:return: file entry

6.24.3.5 GetFilename()

```
def plaso.tarzan.mediator.mediator.Mediator.GetFilename (
    self )
```

Retrieves the name of the active file entry.

:return: name of the active file entry or None

6.24.3.6 PopFromParserChain()

```
def plaso.tarzan.mediator.mediator.Mediator.PopFromParserChain (
    self )
```

Removes the last added parser or parser plugin from the parser chain.

6.24.3.7 ProduceEventWithEventData()

```
def plaso.tarzan.mediator.mediator.Mediator.ProduceEventWithEventData (
    self,
    event,
    event_data )
```

Produce a particular event with its data.

:param event: the event
:param event_data: the event's data

6.24.3.8 ProduceExtractionError()

```
def plaso.tarzan.mediator.mediator.Mediator.ProduceExtractionError (
    self,
    message,
    path_spec )
```

Produce an extraction error with a particular message and a path.

:param message: the error message
:param path_spec: the erroneous path

6.24.3.9 RemoveEventAttribute()

```
def plaso.tarzan.mediator.mediator.Mediator.RemoveEventAttribute (
    self,
    attribute_name )
```

Removes an attribute from being set on all events produced.

:param attribute_name: name of the attribute to remove
:raises KeyError: if the event attribute is not set

6.24.3.10 SetFileEntry()

```
def plaso.tarzan.mediator.mediator.Mediator.SetFileEntry (
    self,
    file_entry )
```

Sets the active file entry.
:param file_entry: file entry

6.24.3.11 temporary_directory()

```
def plaso.tarzan.mediator.mediator.Mediator.temporary_directory (
    self )
```

A path of the directory for temporary files.
:return: path of the directory for temporary files.

6.24.4 Member Data Documentation

6.24.4.1 codepage

plaso.tarzan.mediator.mediator.Mediator.codepage

6.24.4.2 collection_filters_helper

plaso.tarzan.mediator.mediator.Mediator.collection_filters_helper

6.24.4.3 file_entry

plaso.tarzan.mediator.mediator.Mediator.file_entry

6.24.4.4 parser_chain_components

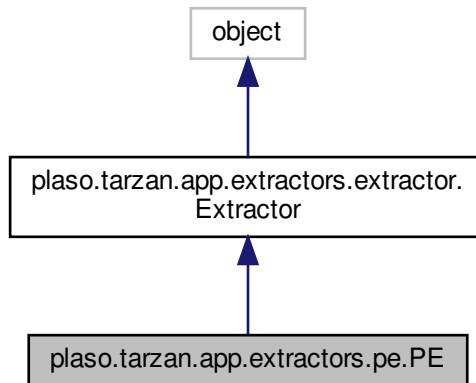
plaso.tarzan.mediator.mediator.Mediator.parser_chain_components

The documentation for this class was generated from the following file:

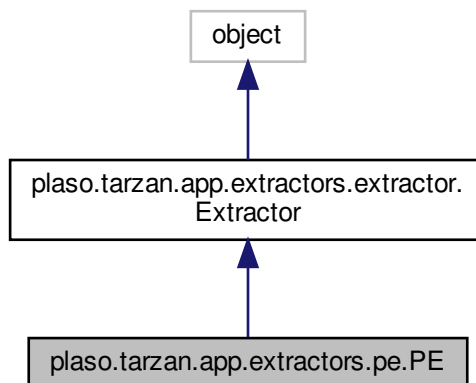
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.py>

6.25 plaso.tarzan.app.extractors.pe.PE Class Reference

Inheritance diagram for plaso.tarzan.app.extractors.pe.PE:



Collaboration diagram for plaso.tarzan.app.extractors.pe.PE:



Public Member Functions

- def [filter](#) (cls, path)
- def [extract](#) (cls, path)

6.25.1 Detailed Description

Extract events from Portable Executable (PE) files.

6.25.2 Member Function Documentation

6.25.2.1 extract()

```
def plaso.tarzan.app.extractors.pe.PE.extract (
    cls,
    path )
```

Extract events from a PE file in the given path.
:param path: the path of the file to extract
:return: a list of events

6.25.2.2 filter()

```
def plaso.tarzan.app.extractors.pe.PE.filter (
    cls,
    path )
```

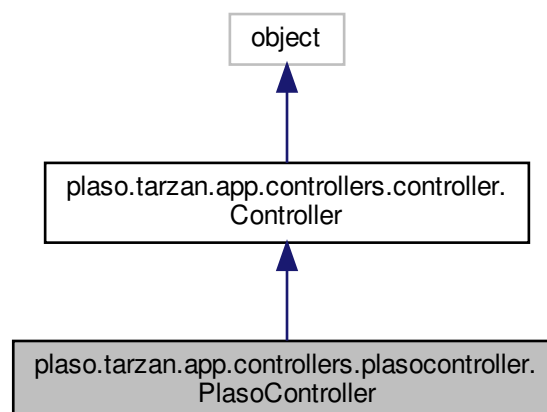
Check if the file-path is a file-path of a PE file.
For the filename extensions see https://en.wikipedia.org/wiki/Portable_Executable
:param path: the file-path to check
:return: True iff it is a PE file

The documentation for this class was generated from the following file:

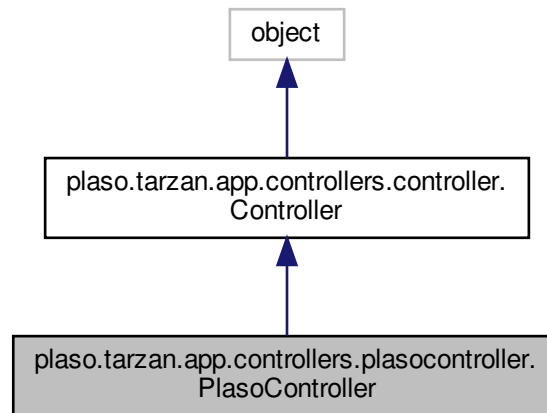
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.py>

6.26 plaso.tarzan.app.controllers.plasocontroller.PlasoController Class Reference

Inheritance diagram for plaso.tarzan.app.controllers.plasocontroller.PlasoController:



Collaboration diagram for `plaso.tarzan.app.controllers.plasocontroller.PlasoController`:



Public Member Functions

- `def __init__ (self, hdfs_base_uri, spark_context)`
- `def extract (self, hdfs_path="")`

Public Attributes

- `spark_context`

6.26.1 Detailed Description

Controller for extraction of events by the Palso.

6.26.2 Constructor & Destructor Documentation

6.26.2.1 `__init__()`

```

def plaso.tarzan.app.controllers.plasocontroller.PlasoController.__init__ (
    self,
    hdfs_base_uri,
    spark_context )
  
```

Create a new controller that will be utilizing HDFS URI and SparkContext.
 :param `hdfs_base_uri`: the base HDFS URI to store
 :param `spark_context`: the Spark context

6.26.3 Member Function Documentation

6.26.3.1 extract()

```
def plaso.tarzan.app.controllers.plasocontroller.PlasoController.extract (
    self,
    hdfs_path = "" )
```

Run Plaso Extractors on a given HDFS path to generate events.
:param hdfs_path: the path where to extract events from
:return: the Flask Response with a JSON document of extracted events

6.26.4 Member Data Documentation

6.26.4.1 spark_context

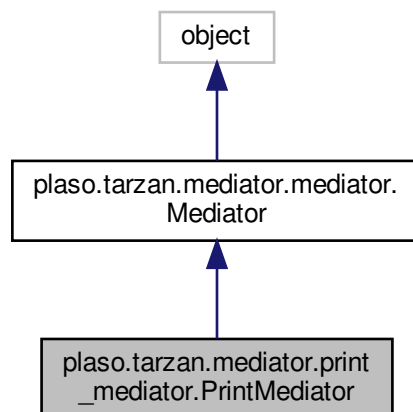
```
plaso.tarzan.app.controllers.plasocontroller.PlasoController.spark_context
```

The documentation for this class was generated from the following file:

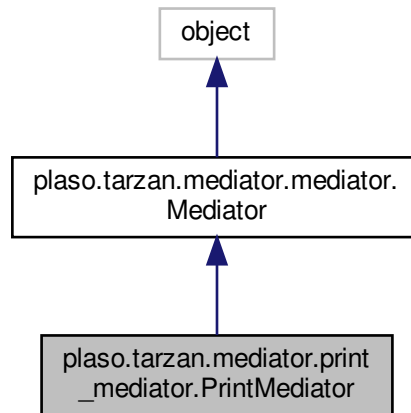
- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroller.py>

6.27 plaso.tarzan.mediator.print_mediator.PrintMediator Class Reference

Inheritance diagram for plaso.tarzan.mediator.print_mediator.PrintMediator:



Collaboration diagram for `plaso.tarzan.mediator.print_mediator.PrintMediator`:



Public Member Functions

- `def __init__ (self)`
- `def ProduceEventWithEventData (self, event, event_data)`
- `def ProduceExtractionError (self, message, path_spec=None)`
- `def ProduceExtractionWarning (self, message, path_spec=None)`

Additional Inherited Members

6.27.1 Detailed Description

Tarzan mediator printing events to stdout.

6.27.2 Constructor & Destructor Documentation

6.27.2.1 `__init__()`

```
def plaso.tarzan.mediator.print_mediator.PrintMediator.__init__ (
    self )
```

Initialize the mediator.

6.27.3 Member Function Documentation

6.27.3.1 ProduceEventWithEventData()

```
def plaso.tarzan.mediator.print_mediator.PrintMediator.ProduceEventWithEventData (
    self,
    event,
    event_data )
```

Produce a particular event with its data.
:param event: the event
:param event_data: the event's data

6.27.3.2 ProduceExtractionError()

```
def plaso.tarzan.mediator.print_mediator.PrintMediator.ProduceExtractionError (
    self,
    message,
    path_spec = None )
```

Produce an extraction error with a particular message and a path.
:param message: the error message
:param path_spec: the erroneous path

6.27.3.3 ProduceExtractionWarning()

```
def plaso.tarzan.mediator.print_mediator.PrintMediator.ProduceExtractionWarning (
    self,
    message,
    path_spec = None )
```

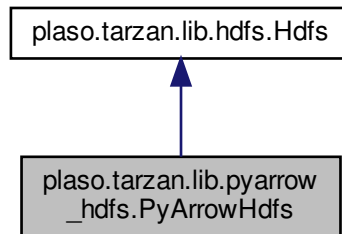
Produce an extraction warning with a particular message and a path.
:param message: the warning message
:param path_spec: the path that caused the warning

The documentation for this class was generated from the following file:

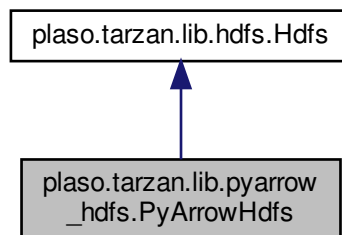
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/print_mediator.py](#)

6.28 plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs Class Reference

Inheritance diagram for plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs:



Collaboration diagram for plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs:



Public Member Functions

- def [open_filesystem](#) (self, hdfs_uri, user="hadoop")
- def [close_filesystem](#) (self, filesystem=None)
- def [get_filesystem](#) (self, force_filesystem=None)
- def [exists](#) (self, path_string, filesystem=None)
- def [info](#) (self, path_string, filesystem=None)
- def [remove](#) (self, path_string, recursive=True, filesystem=None)
- def [mkdir](#) (self, path_string, filesystem=None)
- def [list_files](#) (self, path_string, recursion_level=0, include_dir_names=False, filesystem=None)
- def [open_inputstream](#) (self, path, filesystem=None)
- def [open_outputstream](#) (self, path, filesystem=None)
- def [close_stream](#) (self, input_stream)
- def [get_stream_offset](#) (self, input_stream)
- def [get_path_size](#) (self, path, filesystem=None)
- def [read_inputstream](#) (self, input_stream, size=None)
- def [write_outputstream](#) (self, output_stream, data)
- def [pass_to_outputstream](#) (self, output_stream, input_stream)
- def [seek_stream](#) (self, stream, offset, whence=os.SEEK_SET)

Static Public Attributes

- `fs` = None

Additional Inherited Members

6.28.1 Detailed Description

HDFS driver utilizing PyArrow library.

6.28.2 Member Function Documentation

6.28.2.1 `close_filesystem()`

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.close_filesystem (
    self,
    filesystem = None )
```

Close a given HDFS filesystem.
:param filesystem: the filesystem

6.28.2.2 `close_stream()`

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.close_stream (
    self,
    input_stream )
```

Close a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream

6.28.2.3 `exists()`

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.exists (
    self,
    path_string,
    filesystem = None )
```

Check if a given HDFS path exists in the given HDFS filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: True if the path exists in the filesystem, False otherwise

6.28.2.4 get_filesystem()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.get_filesystem (
    self,
    force_filesystem = None )
```

Get a given or a default HDFS filesystem.
:param force_filesystem: the given filesystem
:return: the filesystem

6.28.2.5 get_path_size()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.get_path_size (
    self,
    path,
    filesystem = None )
```

Get the size of a given HDFS path (a file) in a given filesystem.
:param path: the path
:param filesystem: the filesystem
:return: the size

6.28.2.6 get_stream_offset()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.get_stream_offset (
    self,
    input_stream )
```

Get the current position (an offset) in a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream
:return: the position/offset

6.28.2.7 info()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.info (
    self,
    path_string,
    filesystem = None )
```

Get metadata of a given path in a given or a default filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: the metadata dictionary

6.28.2.8 list_files()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.list_files (
    self,
    path_string,
    recursion_level = 0,
    include_dir_names = False,
    filesystem = None )
```

Get a list of files (optionally recursively to the specified level) in a given HDFS path of a given filesystem.

- :param path_string: the path
- :param recursion_level: the level for the recursion (0 is just a current directory without any recursion)
- :param include_dir_names: True to include also names of directories (suffixed by /)
- :param filesystem: the filesystem
- :return: the list of files in the path

6.28.2.9 mkdir()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.mkdir (
    self,
    path_string,
    filesystem = None )
```

Make a new directory (if does not exist) of a given path in a given or a default filesystem.

- :param path_string: the path
- :param filesystem: the filesystem

6.28.2.10 open_filesystem()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.open_filesystem (
    self,
    hdfs_uri,
    user = "hadoop" )
```

Open HDFS filesystem of a given URI and as a given HDFS user.

- :param hdfs_uri: HDFS URI to open
- :param user: HDFS user to act as when opening
- :return: the opened filesystem as an object of pyarrow.HadoopFileSystem

6.28.2.11 open_inputstream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.open_inputstream (
    self,
    path,
    filesystem = None )
```

Open and get an input-stream for a HDFS file given by its path in a given filesystem.

- :param path: the path
- :param filesystem: the filesystem
- :return: the opened input-stream

6.28.2.12 open_outputstream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.open_outputstream (
    self,
    path,
    filesystem = None )
```

Open and get an output-stream for a HDFS file given by its path in a given filesystem.

:param path: the path
:param filesystem: the filesystem
:return: the opened output-stream

6.28.2.13 pass_to_outputstream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.pass_to_outputstream (
    self,
    output_stream,
    input_stream )
```

Pass data from the input stream to the output stream to a HDFS file.

:param output_stream: the output-stream of the HDFS file
:param input_stream: the input-stream

6.28.2.14 read_inputstream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.read_inputstream (
    self,
    input_stream,
    size = None )
```

Read data form a given opened input stream for a HDFS file.

:param input_stream: the input-stream
:param size: the size of data to read
:return: the data

6.28.2.15 remove()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.remove (
    self,
    path_string,
    recursive = True,
    filesystem = None )
```

Remove (optionally recursively) a HDFS file/directory given by its path in a given or a default filesystem.

:param path_string: the path
:param recursive: True to remove recursively
:param filesystem: the filesystem

6.28.2.16 seek_stream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.seek_stream (
    self,
    stream,
    offset,
    whence = os.SEEK_SET )
```

Set a given position (an offset) in a given (previously opened) input-stream for a HDFS file.

:param stream: the opened stream
:param offset: the position/offset
:param whence: the direction

6.28.2.17 write_outputstream()

```
def plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.write_outputstream (
    self,
    output_stream,
    data )
```

Write data buffer to a given opened output stream for a HDFS file.

:param output_stream: the output-stream
:param data: the data buffer to write
:return: the number of bytes written

6.28.3 Member Data Documentation

6.28.3.1 fs

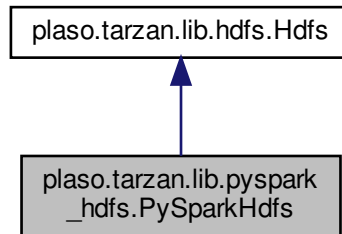
```
plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs.fs = None [static]
```

The documentation for this class was generated from the following file:

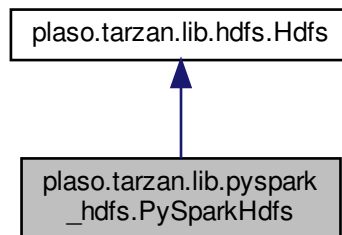
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow_hdfs.py](#)

6.29 plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs Class Reference

Inheritance diagram for plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs:



Collaboration diagram for plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs:



Public Member Functions

- def `__init__` (self, `spark_context`)
- def `open_filesystem` (self, `hdfs_uri`, `user`="hadoop")
- def `close_filesystem` (self, `filesystem`=None)
- def `make_path` (self, `path_string`, `qualified`=True, `filesystem`=None)
- def `append_to_path` (self, `original_path`, `new_child_string`)
- def `list_files` (self, `path_string`, `recursion_level`=0, `include_dir_names`=False, `filesystem`=None)
- def `open_inputstream` (self, `path`, `filesystem`=None)
- def `close_stream` (self, `input_stream`)
- def `get_stream_offset` (self, `input_stream`)
- def `get_path_size` (self, `path`, `filesystem`=None)
- def `read_inputstream` (self, `input_stream`, `size`=sys.maxsize)
- def `read_inputstream_with_chunk` (self, `input_stream`, `size`=sys.maxsize, `chunk_size`=2048)
- def `seek_stream` (self, `stream`, `offset`, `whence`=os.SEEK_SET)
- def `seek_stream_with_path` (self, `stream`, `offset`, `whence`=os.SEEK_SET, `path`=None)

- def [get_filesystem](#) (self, force_filesystem)
- def [exists](#) (self, path_string, filesystem)
- def [info](#) (self, path_string, filesystem=None)
- def [remove](#) (self, path_string, recursive=True, filesystem=None)
- def [mkdir](#) (self, path_string, filesystem=None)
- def [open_outputstream](#) (self, path, filesystem=None)
- def [write_outputstream](#) (self, output_stream, data)

Public Attributes

- [spark_context](#)
- [uri_class](#)
- [path_class](#)
- [fs_class](#)
- [isr_class](#)
- [fs](#)

Additional Inherited Members

6.29.1 Detailed Description

HDFS driver utilizing JVM gateway of the Spark Context.

6.29.2 Constructor & Destructor Documentation

6.29.2.1 `__init__()`

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.__init__ (
    self,
    spark_context )
```

Initialize the driver.
:param spark_context: the Spark Context

6.29.3 Member Function Documentation

6.29.3.1 `append_to_path()`

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.append_to_path (
    self,
    original_path,
    new_child_string )
```

Append a path/directory/file into another HDFS path.
:param original_path: the another path
:param new_child_string: the path to append
:return: the resulting path

6.29.3.2 `close_filesystem()`

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.close_filesystem (
    self,
    filesystem = None )
```

Close a given HDFS filesystem.
:param filesystem: the filesystem

6.29.3.3 `close_stream()`

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.close_stream (
    self,
    input_stream )
```

Close a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream

6.29.3.4 `exists()`

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.exists (
    self,
    path_string,
    filesystem )
```

Check if a given HDFS path exists in the given HDFS filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: True if the path exists in the filesystem, False otherwise

6.29.3.5 get_filesystem()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.get_filesystem (
    self,
    force_filesystem )
```

Get a given or a default HDFS filesystem.
:param force_filesystem: the given filesystem
:return: the filesystem

6.29.3.6 get_path_size()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.get_path_size (
    self,
    path,
    filesystem = None )
```

Get the size of a given HDFS path (a file) in a given filesystem.
:param path: the path
:param filesystem: the filesystem
:return: the size

6.29.3.7 get_stream_offset()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.get_stream_offset (
    self,
    input_stream )
```

Get the current position (an offset) in a given (previously opened) input-stream for a HDFS file.
:param input_stream: the opened input-stream
:return: the position/offset

6.29.3.8 info()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.info (
    self,
    path_string,
    filesystem = None )
```

Get metadata of a given path in a given or a default filesystem.
:param path_string: the path
:param filesystem: the filesystem
:return: the metadata dictionary

6.29.3.9 list_files()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.list_files (
    self,
    path_string,
    recursion_level = 0,
    include_dir_names = False,
    filesystem = None )
```

Get a list of files (optionally recursively to the specified level) in a given HDFS path of a given filesystem.

- :param path_string: the path
- :param recursion_level: the level for the recursion (0 is just a current directory without any recursion)
- :param include_dir_names: True to include also names of directories (suffixed by /)
- :param filesystem: the filesystem
- :return: the list of files in the path

6.29.3.10 make_path()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.make_path (
    self,
    path_string,
    qualified = True,
    filesystem = None )
```

Get a (qualified) HDFS URI from a given path and a given or a default filesystem.

- :param path_string: the path
- :param qualified: True to get the qualified path
- :param filesystem: the filesystem
- :return: the resulting path

6.29.3.11 mkdir()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.mkdir (
    self,
    path_string,
    filesystem = None )
```

Make a new directory (if does not exist) of a given path in a given or a default filesystem.

- :param path_string: the path
- :param filesystem: the filesystem

6.29.3.12 open_filesystem()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.open_filesystem (
    self,
    hdfs_uri,
    user = "hadoop" )
```

Open HDFS filesystem of a given URI and as a given HDFS user.

:param hdfs_uri: HDFS URI to open

:param user: HDFS user to act as when opening

:return: the opened filesystem as an object of org.apache.hadoop.fs.FileSystem class

6.29.3.13 open_inputstream()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.open_inputstream (
    self,
    path,
    filesystem = None )
```

Open and get an input-stream for a HDFS file given by its path in a given filesystem.

:param path: the path

:param filesystem: the filesystem

:return: the opened input-stream

6.29.3.14 open_outputstream()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.open_outputstream (
    self,
    path,
    filesystem = None )
```

Open and get an output-stream for a HDFS file given by its path in a given filesystem.

:param path: the path

:param filesystem: the filesystem

:return: the opened output-stream

6.29.3.15 read_inputstream()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.read_inputstream (
    self,
    input_stream,
    size = sys.maxsize )
```

Read data form a given opened input stream for a HDFS file.

:param input_stream: the input-stream

:param size: the size of data to read

:return: the data

6.29.3.16 read_inputstream_with_chunk()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.read_inputstream_with_chunk (
    self,
    input_stream,
    size = sys.maxsize,
    chunk_size = 2048 )
```

Read data form a given opened input stream for a HDFS file.

:param input_stream: the input-stream
:param size: the size of data to read
:param chunk_size: the size of a chunk
:return: the data

6.29.3.17 remove()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.remove (
    self,
    path_string,
    recursive = True,
    filesystem = None )
```

Remove (optionally recursively) a HDFS file/directory given by its path in a given or a default filesystem.

:param path_string: the path
:param recursive: True to remove recursively
:param filesystem: the filesystem

6.29.3.18 seek_stream()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.seek_stream (
    self,
    stream,
    offset,
    whence = os.SEEK_SET )
```

Set a given position (an offset) in a given (previously opened) input-stream for a HDFS file.

:param stream: the opened stream
:param offset: the position/offset
:param whence: the direction

6.29.3.19 seek_stream_with_path()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.seek_stream_with_path (
    self,
    stream,
    offset,
    whence = os.SEEK_SET,
    path = None )
```

Set a given position (an offset) in a given (previously opened) input-stream for a HDFS file.

- :param stream: the opened stream
- :param offset: the position/offset
- :param whence: the direction
- :param path: the path of the file required to be able to seek from the end of the file

6.29.3.20 write_outputstream()

```
def plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.write_outputstream (
    self,
    output_stream,
    data )
```

Write data buffer to a given opened output stream for a HDFS file.

- :param output_stream: the output-stream
- :param data: the data buffer to write
- :return: the number of bytes written

6.29.4 Member Data Documentation

6.29.4.1 fs

plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.fs

6.29.4.2 fs_class

plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.fs_class

6.29.4.3 isr_class

plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.isr_class

6.29.4.4 path_class

`plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.path_class`

6.29.4.5 spark_context

`plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.spark_context`

6.29.4.6 uri_class

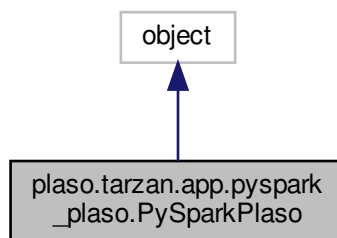
`plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs.uri_class`

The documentation for this class was generated from the following file:

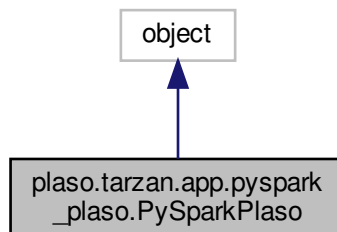
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyspark_hdfs.py](#)

6.30 plaso.tarzan.app.pyspark_plaso.PySparkPlaso Class Reference

Inheritance diagram for `plaso.tarzan.app.pyspark_plaso.PySparkPlaso`:



Collaboration diagram for `plaso.tarzan.app.pyspark_plaso.PySparkPlaso`:



Public Member Functions

- def [create_files_rdd](#) (cls, sc, hdfs_uri)

Static Public Member Functions

- def [get_java_rdd_helpers_package](#) (sc)
- def [list_files](#) (hdfs_uri)
- def [transform_files_rdd_to_extracted_events_rdd](#) (sc, files_rdd)
- def [action_events_rdd_by_saving_into_halyard](#) (sc, events_rdd, table_name, hbase_zk_quorum, hbase_zk_port)
- def [action_events_rdd_by_collecting_into_json](#) (sc, events_rdd)

6.30.1 Detailed Description

Provides methods to create, modify, and evaluate RDD for the Plaso extraction.

6.30.2 Member Function Documentation

6.30.2.1 [action_events_rdd_by_collecting_into_json\(\)](#)

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.action_events_rdd_by_collecting_into_json (
    sc,
    events_rdd ) [static]
```

Collects the content of a given RDD with (HDFS URI, event) pairs into a JSON documents collection.
:param sc: Spark Context of the RDD
:param events_rdd: the RDD of (HDFS URI, event) paris to save
:return the JSON documents collection as a string

6.30.2.2 [action_events_rdd_by_saving_into_halyard\(\)](#)

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.action_events_rdd_by_saving_into_halyard (
    sc,
    events_rdd,
    table_name,
    hbase_zk_quorum,
    hbase_zk_port ) [static]
```

Save the content of a given RDD with (HDFS URI, event) pairs into Halyard into a given table.
:param sc: Spark Context of the RDD
:param events_rdd: the RDD of (HDFS URI, event) paris to save
:param table_name: Halyard repository HBase table name
:param hbase_zk_quorum: HBase Zookeeper quorum of HBase config path
:param hbase_zk_port: the Zookeeper client port

6.30.2.3 create_files_rdd()

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.create_files_rdd (
    cls,
    sc,
    hdfs_uri )
```

Create a new RDD with HDFS URIs to all files (recursively) from the HDFS base URI.
:param sc: Spark Context of the RDD
:param hdfs_uri: the HDFS base URI
:return the RDD of HDFS URIs

6.30.2.4 get_java_rdd_helpers_package()

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.get_java_rdd_helpers_package (
    sc ) [static]
```

Access a JVM gateway of the Spark Context and get Java package `jvm.tarzan.helpers.rdd`.
:param sc: Spark Context
:return: the Java package

6.30.2.5 list_files()

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.list_files (
    hdfs_uri ) [static]
```

Get all files (recursively) from a given HDFS URI.
:param hdfs_uri: the HDFS URI to get files from
:return a list of HDFS URI in the given HDFS URI base-dir

6.30.2.6 transform_files_rdd_to_extracted_events_rdd()

```
def plaso.tarzan.app.pyspark_plaso.PySparkPlaso.transform_files_rdd_to_extracted_events_rdd (
    sc,
    files_rdd ) [static]
```

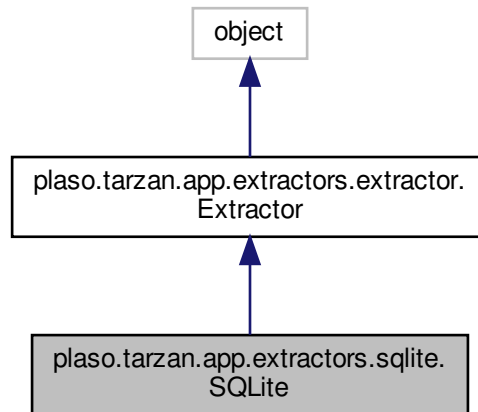
Transform RDD of HDFS URIs into a new RDDs of pairs produced by extractors where each pair consists of the HDFS URI of a file and one of event extracted from the file.
:param sc: Spark Context of the RDD
:param files_rdd: the RDD of HDFS URIs
:return: the RDD of (HDFS URI, event) pairs

The documentation for this class was generated from the following file:

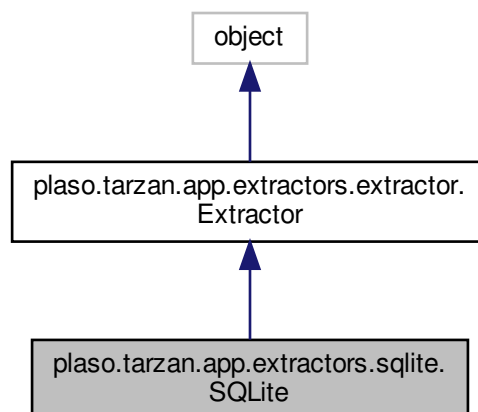
- /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso.py

6.31 plaso.tarzan.app.extractors.sqlite.SQLite Class Reference

Inheritance diagram for plaso.tarzan.app.extractors.sqlite.SQLite:



Collaboration diagram for plaso.tarzan.app.extractors.sqlite.SQLite:



Public Member Functions

- def [filter](#) (cls, path)
- def [extract](#) (cls, path)

6.31.1 Detailed Description

Extract events from SQLite files such as web-browser history files.

6.31.2 Member Function Documentation

6.31.2.1 extract()

```
def plaso.tarzan.app.extractors.sqlite.SQLite.extract (
    cls,
    path )
```

Extract events from an SQLite file in the given path.
:param path: the path of the file to extract
:return: a list of events

6.31.2.2 filter()

```
def plaso.tarzan.app.extractors.sqlite.SQLite.filter (
    cls,
    path )
```

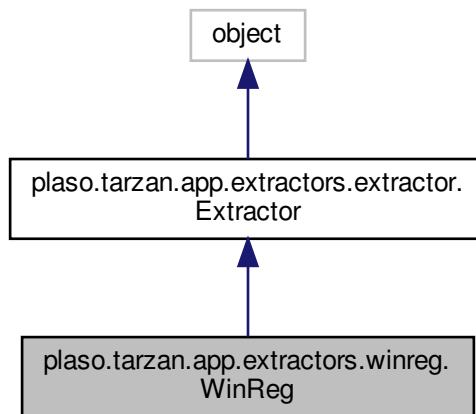
Check if the file-path is a file-path of an SQLite file.
:param path: the file-path to check
:return: True iff it is an SQLite file

The documentation for this class was generated from the following file:

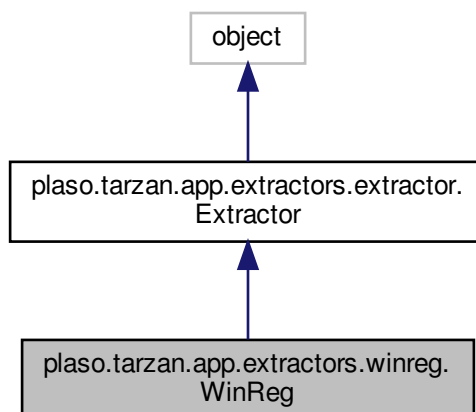
- [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite.py](#)

6.32 plaso.tarzan.app.extractors.winreg.WinReg Class Reference

Inheritance diagram for plaso.tarzan.app.extractors.winreg.WinReg:



Collaboration diagram for plaso.tarzan.app.extractors.winreg.WinReg:



Public Member Functions

- def [filter](#) (cls, path)
- def [extract](#) (cls, path)

6.32.1 Detailed Description

Extract events from Windows Registry files.

6.32.2 Member Function Documentation

6.32.2.1 extract()

```
def plaso.tarzan.app.extractors.winreg.WinReg.extract (
    cls,
    path )
```

Extract events from a Windows Registry file in the given path.
:param path: the path of the file to extract
:return: a list of events

6.32.2.2 filter()

```
def plaso.tarzan.app.extractors.winreg.WinReg.filter (
    cls,
    path )
```

Check if the file-path is a file-path of a Windows Registry file.
For the filename extensions see https://en.wikipedia.org/wiki/Windows_Registry#File_locations
:param path: the file-path to check
:return: True iff it is a PE file

The documentation for this class was generated from the following file:

- </home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.py>

Chapter 7

File Documentation

7.1 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/__init__.py File Reference

Namespaces

- [plaso.tarzan](#)

7.2 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/__init__.py File Reference

Namespaces

- [plaso.tarzan.app](#)

7.3 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/__init__.py File Reference

Namespaces

- [plaso.tarzan.app.controllers](#)

7.4 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/__init__.py File Reference

Namespaces

- [plaso.tarzan.app.extractors](#)

7.5 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/__init__.py File Reference

Namespaces

- [plaso.tarzan.app.views](#)

7.6 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/__init__.py File Reference

Namespaces

- [plaso.tarzan.dfvfs2](#)

7.7 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/__init__.py File Reference

Namespaces

- [plaso.tarzan.file](#)

7.8 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/__init__.py File Reference

Namespaces

- [plaso.tarzan.lib](#)

7.9 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/__init__.py File Reference

Namespaces

- [plaso.tarzan.mediator](#)

7.10 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controller.py File Reference

Classes

- class [plaso.tarzan.app.controllers.controller.Controller](#)

Namespaces

- [plaso.tarzan.app.controllers.controller](#)

7.11 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/filemancontroller.py File Reference

Classes

- class [plaso.tarzan.app.controllers.filemancontroller.FileManController](#)

Namespaces

- [plaso.tarzan.app.controllers.filemancontroller](#)

7.12 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroller.py File Reference

Classes

- class [plaso.tarzan.app.controllers.plasocontroller.PlasoController](#)

Namespaces

- [plaso.tarzan.app.controllers.plasocontroller](#)

7.13 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/extractor.py File Reference

Classes

- class [plaso.tarzan.app.extractors.extractor.Extractor](#)

Namespaces

- [plaso.tarzan.app.extractors.extractor](#)

7.14 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/manager.py File Reference

Classes

- class [plaso.tarzan.app.extractors.manager.Manager](#)

Namespaces

- [plaso.tarzan.app.extractors.manager](#)

7.15 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.py](#) File Reference

Classes

- class [plaso.tarzan.app.extractors.pe.PE](#)

Namespaces

- [plaso.tarzan.app.extractors.pe](#)

7.16 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite.py](#) File Reference

Classes

- class [plaso.tarzan.app.extractors.sqlite.SQLite](#)

Namespaces

- [plaso.tarzan.app.extractors.sqlite](#)

7.17 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.py](#) File Reference

Classes

- class [plaso.tarzan.app.extractors.winreg.WinReg](#)

Namespaces

- [plaso.tarzan.app.extractors.winreg](#)

7.18 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso.py](#) File Reference

Classes

- class [plaso.tarzan.app.pyspark_plaso.PySparkPlaso](#)

Namespaces

- [plaso.tarzan.app.pyspark_plaso](#)

7.19 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark_plaso_webapp.py File Reference

Namespaces

- [plaso.tarzan.app.pyspark_plaso_webapp](#)

Functions

- [def plaso.tarzan.app.pyspark_plaso_webapp.app_list_routes \(\)](#)
- [def plaso.tarzan.app.pyspark_plaso_webapp.configure_app \(spark_context, hdfs_uri\)](#)

Variables

- [plaso.tarzan.app.pyspark_plaso_webapp.app = Flask\(__name__\)](#)
- [plaso.tarzan.app.pyspark_plaso_webapp.methods](#)

7.20 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractview.py File Reference

Classes

- [class plaso.tarzan.app.views.extractview.ExtractView](#)

Namespaces

- [plaso.tarzan.app.views.extractview](#)

7.21 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformview.py File Reference

Classes

- [class plaso.tarzan.app.views.hdfsfileformview.HdfsFileFormView](#)

Namespaces

- [plaso.tarzan.app.views.hdfsfileformview](#)

7.22 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.py](#) File Reference

Classes

- class [plaso.tarzan.app.views.hdfsfileview.HdfsFileView](#)

Namespaces

- [plaso.tarzan.app.views.hdfsfileview](#)

7.23 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfslsview.py](#) File Reference

Classes

- class [plaso.tarzan.app.views.hdfslsview.HdfsLsView](#)

Namespaces

- [plaso.tarzan.app.views.hdfslsview](#)

7.24 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsrview.py](#) File Reference

Classes

- class [plaso.tarzan.app.views.hdfsrview.HdfsRmView](#)

Namespaces

- [plaso.tarzan.app.views.hdfsrview](#)

7.25 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipformview.py](#) File Reference

Classes

- class [plaso.tarzan.app.views.hdfszipformview.HdfsZipFormView](#)

Namespaces

- [plaso.tarzan.app.views.hdfszipformview](#)

7.26 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfszipview.py File Reference

Classes

- class [plaso.tarzan.app.views.hdfszipview.HdfsZipView](#)

Namespaces

- [plaso.tarzan.app.views.hdfszipview](#)

7.27 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/definitions.py File Reference

Namespaces

- [plaso.tarzan.dfvfs2.definitions](#)

Variables

- string [plaso.tarzan.dfvfs2.definitions.TYPE_INDICATOR_HDFS](#) = 'HDFS'

7.28 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_entry.py File Reference

Classes

- class [plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory](#)
- class [plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry](#)

Namespaces

- [plaso.tarzan.dfvfs2.hdfs_file_entry](#)

7.29 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_io.py File Reference

Classes

- class [plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile](#)

Namespaces

- [plaso.tarzan.dfvfs2.hdfs_file_io](#)

7.30 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_file_system.py](#) File Reference

Classes

- class [plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem](#)

Namespaces

- [plaso.tarzan.dfvfs2.hdfs_file_system](#)

7.31 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_path_spec.py](#) File Reference

Classes

- class [plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec](#)

Namespaces

- [plaso.tarzan.dfvfs2.hdfs_path_spec](#)

7.32 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dfvfs2/hdfs_resolver_helper.py](#) File Reference

Classes

- class [plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFSResolverHelper](#)

Namespaces

- [plaso.tarzan.dfvfs2.hdfs_resolver_helper](#)

7.33 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/file.py](#) File Reference

Classes

- class [plaso.tarzan.file.file.LocalFileObject](#)

Namespaces

- [plaso.tarzan.file.file](#)

7.34 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/generic_file.py File Reference

Classes

- class [plaso.tarzan.file.generic_file.FileObject](#)

Namespaces

- [plaso.tarzan.file.generic_file](#)

7.35 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/hdfs_file.py File Reference

Classes

- class [plaso.tarzan.file.hdfs_file.HdfsFileObject](#)

Namespaces

- [plaso.tarzan.file.hdfs_file](#)

7.36 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom_pip_package_dir.py File Reference

Classes

- class [plaso.tarzan.lib.custom_pip_package_dir.CustomPipPackageDir](#)

Namespaces

- [plaso.tarzan.lib.custom_pip_package_dir](#)

7.37 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/hdfs.py File Reference

Classes

- class [plaso.tarzan.lib.hdfs.Hdfs](#)

Namespaces

- [plaso.tarzan.lib.hdfs](#)

7.38 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow_hdfs.py](#) File Reference

Classes

- class [plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs](#)

Namespaces

- [plaso.tarzan.lib.pyarrow_hdfs](#)

7.39 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyspark_hdfs.py](#) File Reference

Classes

- class [plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs](#)

Namespaces

- [plaso.tarzan.lib.pyspark_hdfs](#)

7.40 [/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered_mediator.py](#) File Reference

Classes

- class [plaso.tarzan.mediator.buffered_mediator.BufferedMediator](#)

Namespaces

- [plaso.tarzan.mediator.buffered_mediator](#)

Functions

- def [plaso.tarzan.mediator.buffered_mediator.generate](#) (bm)

Variables

- [plaso.tarzan.mediator.buffered_mediator.buffered_mediator](#) = BufferedMediator()
- [plaso.tarzan.mediator.buffered_mediator.result](#) = buffered_mediator.flush_buffer()

7.41 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.py File Reference

Classes

- class [plaso.tarzan.mediator.mediator.Mediator](#)

Namespaces

- [plaso.tarzan.mediator.mediator](#)

7.42 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/print_mediator.py File Reference

Classes

- class [plaso.tarzan.mediator.print_mediator.PrintMediator](#)

Namespaces

- [plaso.tarzan.mediator.print_mediator](#)

7.43 /home/burgetr/git/pyspark-plaso/src/plaso/tarzan/run.py File Reference

Namespaces

- [plaso.tarzan.run](#)

Functions

- def [plaso.tarzan.run.run_PE](#) (mediator)
- def [plaso.tarzan.run.run_SQLite](#) (mediator)
- def [plaso.tarzan.run.run_WinReg](#) (mediator)

Variables

- [plaso.tarzan.run.m](#) = [plaso.tarzan.mediator.print_mediator.PrintMediator](#)()

Index

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/___
init__.py, 111

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/___
__init__.py, 111

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/___
__init__.py, 111

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controll~~er~~
_path_spec.py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/controll~~er~~
_resolver_helper.py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroll~~er~~
_py, 113

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/controllers/plasocontroll~~er~~
_py, 113

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/___
__init__.py, 111

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/extract~~or~~
_file.py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/mana~~ger~~
_file.py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/mana~~ger~~
_py, 113

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.~~py~~
_init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/pe.~~py~~
_py, 114

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite_pip_package_dir.py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/sqlite_pip_package_dir.py, 114

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.~~py~~
_py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/extractors/winreg.~~py~~
_py, 114

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark___
_plaso.py, 114

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/pyspark___
_plaso_webapp.py, 115

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/___
__init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractviewmediator.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/extractviewmediator.py, 115

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformview.~~py~~
_py, 115

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformviewmediator.py, 121

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileformviewmediator.py, 116

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.~~py~~
_py, 116

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsfileview.~~py~~
_py, 116

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsplasocontroller.~~py~~
_py, 116

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsplasocontroller.~~py~~
_py, 116

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/app/views/hdfsplasocontroller.~~py~~
_py, 117

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/___
__init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/definitionplaso.
py, 117

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs___
_file_entry.py, 117

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs___
_file_io.py, 117

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs___
_file_system.py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs___
_path_spec.py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/dvdfs2/hdfs___
_resolver_helper.py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/___
__init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/file.___
_py, 118

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/generic___
_file.py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/file/hdfs___
_file.py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/___
__init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom___
_hdfs.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/custom___
_hdfs.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/hdfs.___
_py, 119

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow___
_hdfs.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/lib/pyarrow___
_hdfs.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/___
__init__.py, 112

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered___
_mediator.py, 120

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/buffered___
_mediator.py, 115

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.___
_py, 115

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/mediator/mediator.___
_py, 121

/home/burgetr/git/pyspark-plaso/src/plaso/tarzan/run.py, ___
__init__

plaso:tarzan::app::controllers::controller::___
Controller, 23

plaso:tarzan::app::controllers::filemancontroller___
::FileManController, 31

plaso:tarzan::app::controllers::plasocontroller::___
PlasoController, 86

plaso:tarzan::app::views::extractview::Extract___
View, 29

plaso:tarzan::app::views::hdfsfileformview::Hdfs___
FileFormView, 51

plaso:tarzan::app::views::hdfsfileview::HdfsFile___

- View, 61
- plaso::tarzan::app::views::hdfsview::HdfsLsView, 63
- plaso::tarzan::app::views::hdfsrmview::HdfsRm↔View, 68
- plaso::tarzan::app::views::hdfszipformview::Hdfs↔ZipFormView, 70
- plaso::tarzan::app::views::hdfszipview::HdfsZip↔View, 73
- plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔Entry, 48
- plaso::tarzan::dfvfs2::hdfs_file_io::HDFSFile, 45
- plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔FileSystem, 57
- plaso::tarzan::dfvfs2::hdfs_path_spec::HDFS↔PathSpec, 65
- plaso::tarzan::file::file::LocalFileObject, 75
- plaso::tarzan::file::hdfs_file::HdfsFileObject, 54
- plaso::tarzan::lib::custom_pip_package_dir::↔CustomPipPackageDir, 25
- plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 97
- plaso::tarzan::mediator::buffered_mediator::↔BufferedMediator, 20
- plaso::tarzan::mediator::mediator::Mediator, 80
- plaso::tarzan::mediator::print_mediator::Print↔Mediator, 88
- abort
 - plaso::tarzan::mediator::mediator::Mediator, 80
- access_time
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔Entry, 48
- action_events_rdd_by_collecting_into_json
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso, 105
- action_events_rdd_by_saving_into_halyard
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso, 105
- AddEventAttribute
 - plaso::tarzan::mediator::mediator::Mediator, 80
- app
 - plaso::tarzan::app::pyspark_plaso_webapp, 12
- app_list_routes
 - plaso::tarzan::app::pyspark_plaso_webapp, 11
- append_to_path
 - plaso::tarzan::lib::hdfs::Hdfs, 37
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 97
- AppendToParserChain
 - plaso::tarzan::mediator::mediator::Mediator, 81
- basename
 - plaso::tarzan::lib::hdfs::Hdfs, 37
- BasenamePath
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔FileSystem, 58
- buffer_queue
 - plaso::tarzan::mediator::buffered_mediator::↔BufferedMediator, 22
- buffered_mediator
 - plaso::tarzan::mediator::buffered_mediator, 17
- change_time
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔Entry, 49
- check_installed_package
 - plaso::tarzan::lib::custom_pip_package_dir::↔CustomPipPackageDir, 25
- close
 - plaso::tarzan::file::file::LocalFileObject, 76
 - plaso::tarzan::file::generic_file::FileObject, 35
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 54
- close_filesystem
 - plaso::tarzan::lib::hdfs::Hdfs, 37
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 91
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 98
- close_stream
 - plaso::tarzan::lib::hdfs::Hdfs, 38
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 91
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 98
- codepage
 - plaso::tarzan::mediator::mediator::Mediator, 83
- collection_filters_helper
 - plaso::tarzan::mediator::mediator::Mediator, 83
- configure_app
 - plaso::tarzan::app::pyspark_plaso_webapp, 12
- controller
 - plaso::tarzan::app::views::extractview::Extract↔View, 30
 - plaso::tarzan::app::views::hdfsfileformview::Hdfs↔FileFormView, 52
 - plaso::tarzan::app::views::hdfsfileview::HdfsFile↔View, 62
 - plaso::tarzan::app::views::hdfsview::HdfsLsView, 64
 - plaso::tarzan::app::views::hdfsrmview::HdfsRm↔View, 69
 - plaso::tarzan::app::views::hdfszipformview::Hdfs↔ZipFormView, 71
 - plaso::tarzan::app::views::hdfszipview::HdfsZip↔View, 74
- create_files_rdd
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso, 105
- delete
 - plaso::tarzan::app::views::hdfsfileformview::Hdfs↔FileFormView, 52
 - plaso::tarzan::app::views::hdfsfileview::HdfsFile↔View, 61
 - plaso::tarzan::app::views::hdfsrmview::HdfsRm↔View, 69
 - plaso::tarzan::app::views::hdfszipformview::Hdfs↔ZipFormView, 71
 - plaso::tarzan::app::views::hdfszipview::HdfsZip↔View, 73
- deregister_extractor
 - plaso::tarzan::app::extractors::manager::Manager, 78

- dirname
 - plaso::tarzan::lib::hdfs::Hdfs, 38
- DirnamePath
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔
FileSystem, 58
- entry_type
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔
Entry, 50
- Error
 - plaso::tarzan::mediator::buffered_mediator::↔
BufferedMediator, 22
- Event
 - plaso::tarzan::mediator::buffered_mediator::↔
BufferedMediator, 22
- exists
 - plaso::tarzan::lib::hdfs::Hdfs, 38
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 91
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 98
- extract
 - plaso::tarzan::app::controllers::plasocontroller::↔
PlasoController, 87
 - plaso::tarzan::app::extractors::extractor::Extractor,
27
 - plaso::tarzan::app::extractors::pe::PE, 85
 - plaso::tarzan::app::extractors::sqlite::SQLite, 108
 - plaso::tarzan::app::extractors::winreg::WinReg,
110
- file_entry
 - plaso::tarzan::mediator::mediator::Mediator, 83
- file_object
 - plaso::tarzan::file::file::LocalFileObject, 77
- FileEntryExistsByPathSpec
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔
FileSystem, 58
- filter
 - plaso::tarzan::app::extractors::extractor::Extractor,
27
 - plaso::tarzan::app::extractors::pe::PE, 85
 - plaso::tarzan::app::extractors::sqlite::SQLite, 108
 - plaso::tarzan::app::extractors::winreg::WinReg,
110
- flush_buffer
 - plaso::tarzan::mediator::buffered_mediator::↔
BufferedMediator, 21
- fs
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 95
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 103
- fs_class
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 103
- generate
 - plaso::tarzan::mediator::buffered_mediator, 17
- get
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 32
 - plaso::tarzan::app::views::extractview::Extract↔
View, 30
 - plaso::tarzan::app::views::hdfsfileformview::Hdfs↔
FileFormView, 52
 - plaso::tarzan::app::views::hdfsfileview::HdfsFile↔
View, 61
 - plaso::tarzan::app::views::hdfslsview::HdfsLsView,
64
 - plaso::tarzan::app::views::hdfsrmview::HdfsRm↔
View, 69
 - plaso::tarzan::app::views::hdfszipformview::Hdfs↔
ZipFormView, 71
 - plaso::tarzan::app::views::hdfszipview::HdfsZip↔
View, 73
- get_buffer_queue
 - plaso::tarzan::mediator::buffered_mediator::↔
BufferedMediator, 21
- get_description
 - plaso::tarzan::app::extractors::extractor::Extractor,
28
- get_extractors
 - plaso::tarzan::app::extractors::manager::Manager,
78
- get_filesystem
 - plaso::tarzan::lib::hdfs::Hdfs, 38
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 91
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 98
- get_java_rdd_helpers_package
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso,
106
- get_name
 - plaso::tarzan::app::extractors::extractor::Extractor,
28
- get_offset
 - plaso::tarzan::dfvfs2::hdfs_file_io::HDFSFile, 46
- get_path_size
 - plaso::tarzan::lib::hdfs::Hdfs, 39
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 92
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 99
- get_size
 - plaso::tarzan::dfvfs2::hdfs_file_io::HDFSFile, 46
- get_stream_offset
 - plaso::tarzan::lib::hdfs::Hdfs, 39
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 92
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 99
- get_zip
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 32
- GetFileEntry
 - plaso::tarzan::mediator::mediator::Mediator, 81
- GetFileEntryByPathSpec
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔
FileSystem, 58
- GetFilename
 - plaso::tarzan::mediator::mediator::Mediator, 81
- GetLinkedFileEntry
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔
Entry, 49
- GetParentFileEntry
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔

- Entry, 49
- GetRootFileEntry
 - plaso::tarzan::dfvs2::hdfs_file_system::HDFS↔
FileSystem, 59
- hdfs
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 34
 - plaso::tarzan::dfvs2::hdfs_file_system::HDFS↔
FileSystem, 60
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 56
- hdfs_base_uri
 - plaso::tarzan::app::controllers::controller:↔
Controller, 24
- info
 - plaso::tarzan::lib::hdfs::Hdfs, 39
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 92
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 99
- input_stream
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 56
- install
 - plaso::tarzan::lib::custom_pip_package_dir:↔
CustomPipPackageDir, 25
- isr_class
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 103
- JoinPath
 - plaso::tarzan::dfvs2::hdfs_file_system::HDFS↔
FileSystem, 59
- LOCATION_ROOT
 - plaso::tarzan::dfvs2::hdfs_file_system::HDFS↔
FileSystem, 60
- link
 - plaso::tarzan::dfvs2::hdfs_file_entry::HDFSFile↔
Entry, 49
- list_files
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso,
106
 - plaso::tarzan::lib::hdfs::Hdfs, 39
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 92
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 99
- ls
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 32
- m
 - plaso::tarzan::run, 18
- make_hdfs_uri
 - plaso::tarzan::app::controllers::controller:↔
Controller, 23
- make_package_dir
 - plaso::tarzan::lib::custom_pip_package_dir:↔
CustomPipPackageDir, 25
- make_path
 - plaso::tarzan::lib::hdfs::Hdfs, 40
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 100
- make_qualified_path
 - plaso::tarzan::lib::hdfs::Hdfs, 40
- make_simple_path
 - plaso::tarzan::lib::hdfs::Hdfs, 40
- make_uri
 - plaso::tarzan::lib::hdfs::Hdfs, 41
- methods
 - plaso::tarzan::app::pyspark_plaso_webapp, 12
- mkdir
 - plaso::tarzan::lib::hdfs::Hdfs, 41
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 93
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 100
- modification_time
 - plaso::tarzan::dfvs2::hdfs_file_entry::HDFSFile↔
Entry, 49
- name
 - plaso::tarzan::dfvs2::hdfs_file_entry::HDFSFile↔
Entry, 50
- NewFileObject
 - plaso::tarzan::dfvs2::hdfs_resolver_helper::HDF↔
SResolverHelper, 67
- NewFileSystem
 - plaso::tarzan::dfvs2::hdfs_resolver_helper::HDF↔
SResolverHelper, 67
- open
 - plaso::tarzan::file::file::LocalFileObject, 76
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 54
- open_filesystem
 - plaso::tarzan::lib::hdfs::Hdfs, 41
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 93
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 100
- open_inputstream
 - plaso::tarzan::lib::hdfs::Hdfs, 41
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 93
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 101
- open_outputstream
 - plaso::tarzan::lib::hdfs::Hdfs, 42
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 93
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 101
- PATH_SEPARATOR
 - plaso::tarzan::dfvs2::hdfs_file_system::HDFS↔
FileSystem, 60
 - plaso::tarzan::lib::hdfs::Hdfs, 43
- package_dir
 - plaso::tarzan::lib::custom_pip_package_dir:↔
CustomPipPackageDir, 26
- parse_uri
 - plaso::tarzan::lib::hdfs::Hdfs, 42
- parser_chain_components
 - plaso::tarzan::mediator::mediator::Mediator, 83
- pass_to_outputstream
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 94
- path
 - plaso::tarzan::file::file::LocalFileObject, 77
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 56
- path_class
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 103

- pip_install
 - plaso::tarzan::lib::custom_pip_package_dir::↔ CustomPipPackageDir, 25
- plaso, 9
- plaso.tarzan, 9
- plaso.tarzan.app, 9
- plaso.tarzan.app.controllers, 9
- plaso.tarzan.app.controllers.controller, 10
- plaso.tarzan.app.controllers.controller.Controller, 22
- plaso.tarzan.app.controllers.filemancontroller, 10
- plaso.tarzan.app.controllers.filemancontroller.FileMan↔ Controller, 30
- plaso.tarzan.app.controllers.plasocontroller, 10
- plaso.tarzan.app.controllers.plasocontroller.Plaso↔ Controller, 85
- plaso.tarzan.app.extractors, 10
- plaso.tarzan.app.extractors.extractor, 10
- plaso.tarzan.app.extractors.extractor.Extractor, 26
- plaso.tarzan.app.extractors.manager, 10
- plaso.tarzan.app.extractors.manager.Manager, 77
- plaso.tarzan.app.extractors.pe, 11
- plaso.tarzan.app.extractors.pe.PE, 84
- plaso.tarzan.app.extractors.sqlite, 11
- plaso.tarzan.app.extractors.sqlite.SQLite, 107
- plaso.tarzan.app.extractors.winreg, 11
- plaso.tarzan.app.extractors.winreg.WinReg, 109
- plaso.tarzan.app.pyspark_plaso, 11
- plaso.tarzan.app.pyspark_plaso.PySparkPlaso, 104
- plaso.tarzan.app.pyspark_plaso_webapp, 11
- plaso.tarzan.app.views, 12
- plaso.tarzan.app.views.extractview, 12
- plaso.tarzan.app.views.extractview.ExtractView, 29
- plaso.tarzan.app.views.hdfsfileformview, 13
- plaso.tarzan.app.views.hdfsfileformview.HdfsFileForm↔ View, 51
- plaso.tarzan.app.views.hdfsfileview, 13
- plaso.tarzan.app.views.hdfsfileview.HdfsFileView, 60
- plaso.tarzan.app.views.hdfslsview, 13
- plaso.tarzan.app.views.hdfslsview.HdfsLsView, 63
- plaso.tarzan.app.views.hdfsrview, 13
- plaso.tarzan.app.views.hdfsrview.HdfsRmView, 68
- plaso.tarzan.app.views.hdfszipformview, 13
- plaso.tarzan.app.views.hdfszipformview.HdfsZipForm↔ View, 70
- plaso.tarzan.app.views.hdfszipview, 13
- plaso.tarzan.app.views.hdfszipview.HdfsZipView, 72
- plaso.tarzan.dfvfs2, 14
- plaso.tarzan.dfvfs2.definitions, 14
- plaso.tarzan.dfvfs2.hdfs_file_entry, 14
- plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSDirectory, 44
- plaso.tarzan.dfvfs2.hdfs_file_entry.HDFSFileEntry, 47
- plaso.tarzan.dfvfs2.hdfs_file_io, 14
- plaso.tarzan.dfvfs2.hdfs_file_io.HDFSFile, 45
- plaso.tarzan.dfvfs2.hdfs_file_system, 14
- plaso.tarzan.dfvfs2.hdfs_file_system.HDFSFileSystem, 56
- plaso.tarzan.dfvfs2.hdfs_path_spec, 15
- plaso.tarzan.dfvfs2.hdfs_path_spec.HDFSPathSpec, 64
- plaso.tarzan.dfvfs2.hdfs_resolver_helper, 15
- plaso.tarzan.dfvfs2.hdfs_resolver_helper.HDFS↔ ResolverHelper, 66
- plaso.tarzan.file, 15
- plaso.tarzan.file.file, 15
- plaso.tarzan.file.file.LocalFileObject, 74
- plaso.tarzan.file.generic_file, 15
- plaso.tarzan.file.generic_file.FileObject, 34
- plaso.tarzan.file.hdfs_file, 15
- plaso.tarzan.file.hdfs_file.HdfsFileObject, 53
- plaso.tarzan.lib, 16
- plaso.tarzan.lib.custom_pip_package_dir, 16
- plaso.tarzan.lib.custom_pip_package_dir.CustomPip↔ PackageDir, 24
- plaso.tarzan.lib.hdfs, 16
- plaso.tarzan.lib.hdfs.Hdfs, 36
- plaso.tarzan.lib.pyarrow_hdfs, 16
- plaso.tarzan.lib.pyarrow_hdfs.PyArrowHdfs, 90
- plaso.tarzan.lib.pyspark_hdfs, 16
- plaso.tarzan.lib.pyspark_hdfs.PySparkHdfs, 96
- plaso.tarzan.mediator, 16
- plaso.tarzan.mediator.buffered_mediator, 17
- plaso.tarzan.mediator.buffered_mediator.Buffered↔ Mediator, 19
- plaso.tarzan.mediator.mediator, 17
- plaso.tarzan.mediator.mediator.Mediator, 79
- plaso.tarzan.mediator.print_mediator, 18
- plaso.tarzan.mediator.print_mediator.PrintMediator, 87
- plaso.tarzan.run, 18
- plaso::tarzan::app::controllers::controller::Controller
 - __init__, 23
 - hdfs_base_uri, 24
 - make_hdfs_uri, 23
 - strip_hdfs_uri, 23
- plaso::tarzan::app::controllers::filemancontroller::File↔ ManController
 - __init__, 31
 - get, 32
 - get_zip, 32
 - hdfs, 34
 - ls, 32
 - put, 32
 - put_form, 33
 - rm, 33
 - zip_put, 33
 - zip_put_form, 33
- plaso::tarzan::app::controllers::plasocontroller::Plaso↔ Controller
 - __init__, 86
 - extract, 87
 - spark_context, 87
- plaso::tarzan::app::extractors::extractor::Extractor
 - extract, 27
 - filter, 27
 - get_description, 28
 - get_name, 28
- plaso::tarzan::app::extractors::manager::Manager
 - deregister_extractor, 78

- get_extractors, 78
 - register_extractor, 78
- plaso::tarzan::app::extractors::pe::PE
 - extract, 85
 - filter, 85
- plaso::tarzan::app::extractors::sqlite::SQLite
 - extract, 108
 - filter, 108
- plaso::tarzan::app::extractors::winreg::WinReg
 - extract, 110
 - filter, 110
- plaso::tarzan::app::pyspark_plaso::PySparkPlaso
 - action_events_rdd_by_collecting_into_json, 105
 - action_events_rdd_by_saving_into_halyard, 105
 - create_files_rdd, 105
 - get_java_rdd_helpers_package, 106
 - list_files, 106
 - transform_files_rdd_to_extracted_events_rdd, 106
- plaso::tarzan::app::pyspark_plaso_webapp
 - app, 12
 - app_list_routes, 11
 - configure_app, 12
 - methods, 12
- plaso::tarzan::app::views::extractview::ExtractView
 - __init__, 29
 - controller, 30
 - get, 30
- plaso::tarzan::app::views::hdfsfileformview::HdfsFile↔
 - FormView
 - __init__, 51
 - controller, 52
 - delete, 52
 - get, 52
 - post, 52
 - put, 52
- plaso::tarzan::app::views::hdfsfileview::HdfsFileView
 - __init__, 61
 - controller, 62
 - delete, 61
 - get, 61
 - post, 62
 - put, 62
- plaso::tarzan::app::views::hdfslsview::HdfsLsView
 - __init__, 63
 - controller, 64
 - get, 64
- plaso::tarzan::app::views::hdfsrmview::HdfsRmView
 - __init__, 68
 - controller, 69
 - delete, 69
 - get, 69
- plaso::tarzan::app::views::hdfszipformview::HdfsZip↔
 - FormView
 - __init__, 70
 - controller, 71
 - delete, 71
 - get, 71
 - post, 71
 - put, 71
- plaso::tarzan::app::views::hdfszipview::HdfsZipView
 - __init__, 73
 - controller, 74
 - delete, 73
 - get, 73
 - post, 73
 - put, 73
- plaso::tarzan::dfvfs2::definitions
 - TYPE_INDICATOR_HDFS, 14
- plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFileEntry
 - __init__, 48
 - access_time, 48
 - change_time, 49
 - entry_type, 50
 - GetLinkedFileEntry, 49
 - GetParentFileEntry, 49
 - link, 49
 - modification_time, 49
 - name, 50
 - TYPE_INDICATOR, 50
- plaso::tarzan::dfvfs2::hdfs_file_io::HDFSFile
 - __init__, 45
 - get_offset, 46
 - get_size, 46
 - read, 46
 - seek, 46
- plaso::tarzan::dfvfs2::hdfs_file_system::HDFSFile↔
 - System
 - __init__, 57
 - BasenamePath, 58
 - DirnamePath, 58
 - FileEntryExistsByPathSpec, 58
 - GetFileEntryByPathSpec, 58
 - GetRootFileEntry, 59
 - hdfs, 60
 - JoinPath, 59
 - LOCATION_ROOT, 60
 - PATH_SEPARATOR, 60
 - SplitPath, 59
 - TYPE_INDICATOR, 60
- plaso::tarzan::dfvfs2::hdfs_path_spec::HDFSPathSpec
 - __init__, 65
 - TYPE_INDICATOR, 65
- plaso::tarzan::dfvfs2::hdfs_resolver_helper::HDFS↔
 - ResolverHelper
 - NewFileObject, 67
 - NewFileSystem, 67
 - TYPE_INDICATOR, 67
- plaso::tarzan::file::file::LocalFileObject
 - __init__, 75
 - close, 76
 - file_object, 77
 - open, 76
 - path, 77
 - read, 76
 - seek, 76
 - tell, 76

- plaso::tarzan::file::generic_file::FileObject
 - close, 35
 - read, 35
 - seek, 35
- plaso::tarzan::file::hdfs_file::HdfsFileObject
 - __init__, 54
 - close, 54
 - hdfs, 56
 - input_stream, 56
 - open, 54
 - path, 56
 - read, 55
 - seek, 55
 - tell, 55
- plaso::tarzan::lib::custom_pip_package_dir::Custom↔
 - PipPackageDir
 - __init__, 25
 - check_installed_package, 25
 - install, 25
 - make_package_dir, 25
 - package_dir, 26
 - pip_install, 25
 - register_package_dir, 26
- plaso::tarzan::lib::hdfs::Hdfs
 - append_to_path, 37
 - basename, 37
 - close_filesystem, 37
 - close_stream, 38
 - dirname, 38
 - exists, 38
 - get_filesystem, 38
 - get_path_size, 39
 - get_stream_offset, 39
 - info, 39
 - list_files, 39
 - make_path, 40
 - make_qualified_path, 40
 - make_simple_path, 40
 - make_uri, 41
 - mkdir, 41
 - open_filesystem, 41
 - open_inputstream, 41
 - open_outputstream, 42
 - PATH_SEPARATOR, 43
 - parse_uri, 42
 - read_inputstream, 42
 - remove, 42
 - SCHEME, 43
 - seek_stream, 43
 - write_outputstream, 43
- plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs
 - close_filesystem, 91
 - close_stream, 91
 - exists, 91
 - fs, 95
 - get_filesystem, 91
 - get_path_size, 92
 - get_stream_offset, 92
 - info, 92
 - list_files, 92
 - mkdir, 93
 - open_filesystem, 93
 - open_inputstream, 93
 - open_outputstream, 93
 - pass_to_outputstream, 94
 - read_inputstream, 94
 - remove, 94
 - seek_stream, 94
 - write_outputstream, 95
- plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs
 - __init__, 97
 - append_to_path, 97
 - close_filesystem, 98
 - close_stream, 98
 - exists, 98
 - fs, 103
 - fs_class, 103
 - get_filesystem, 98
 - get_path_size, 99
 - get_stream_offset, 99
 - info, 99
 - isr_class, 103
 - list_files, 99
 - make_path, 100
 - mkdir, 100
 - open_filesystem, 100
 - open_inputstream, 101
 - open_outputstream, 101
 - path_class, 103
 - read_inputstream, 101
 - read_inputstream_with_chunk, 101
 - remove, 102
 - seek_stream, 102
 - seek_stream_with_path, 102
 - spark_context, 104
 - uri_class, 104
 - write_outputstream, 103
- plaso::tarzan::mediator::buffered_mediator
 - buffered_mediator, 17
 - generate, 17
 - result, 17
- plaso::tarzan::mediator::buffered_mediator::Buffered↔
 - Mediator
 - __init__, 20
 - buffer_queue, 22
 - Error, 22
 - Event, 22
 - flush_buffer, 21
 - get_buffer_queue, 21
 - ProduceEventWithEventData, 21
 - ProduceExtractionError, 21
- plaso::tarzan::mediator::mediator::Mediator
 - __init__, 80
 - abort, 80
 - AddEventAttribute, 80
 - AppendToParserChain, 81

- codepage, 83
- collection_filters_helper, 83
- file_entry, 83
- GetFileEntry, 81
- GetFilename, 81
- parser_chain_components, 83
- PopFromParserChain, 81
- ProduceEventWithEventData, 82
- ProduceExtractionError, 82
- RemoveEventAttribute, 82
- SetFileEntry, 82
- temporary_directory, 83
- plaso::tarzan::mediator::print_mediator::PrintMediator
 - __init__, 88
 - ProduceEventWithEventData, 89
 - ProduceExtractionError, 89
 - ProduceExtractionWarning, 89
- plaso::tarzan::run
 - m, 18
 - run_PE, 18
 - run_SQLite, 18
 - run_WinReg, 18
- PopFromParserChain
 - plaso::tarzan::mediator::mediator::Mediator, 81
- post
 - plaso::tarzan::app::views::hdfsfileformview::HdfsFileFormView, 52
 - plaso::tarzan::app::views::hdfsfileview::HdfsFileView, 62
 - plaso::tarzan::app::views::hdfszipformview::HdfsZipFormView, 71
 - plaso::tarzan::app::views::hdfszipview::HdfsZipView, 73
- ProduceEventWithEventData
 - plaso::tarzan::mediator::buffered_mediator::BufferedMediator, 21
 - plaso::tarzan::mediator::mediator::Mediator, 82
 - plaso::tarzan::mediator::print_mediator::PrintMediator, 89
- ProduceExtractionError
 - plaso::tarzan::mediator::buffered_mediator::BufferedMediator, 21
 - plaso::tarzan::mediator::mediator::Mediator, 82
 - plaso::tarzan::mediator::print_mediator::PrintMediator, 89
- ProduceExtractionWarning
 - plaso::tarzan::mediator::print_mediator::PrintMediator, 89
- put
 - plaso::tarzan::app::controllers::filemancontroller::FileManController, 32
 - plaso::tarzan::app::views::hdfsfileformview::HdfsFileFormView, 52
 - plaso::tarzan::app::views::hdfsfileview::HdfsFileView, 62
 - plaso::tarzan::app::views::hdfszipformview::HdfsZipFormView, 71
 - plaso::tarzan::app::views::hdfszipview::HdfsZipView, 73
- plaso::tarzan::app::views::hdfszipview::HdfsZipView, 73
- put_form
 - plaso::tarzan::app::controllers::filemancontroller::FileManController, 33
- read
 - plaso::tarzan::dfvs2::hdfs_file_io::HDFSFile, 46
 - plaso::tarzan::file::file::LocalFileObject, 76
 - plaso::tarzan::file::generic_file::FileObject, 35
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 55
- read_inputstream
 - plaso::tarzan::lib::hdfs::Hdfs, 42
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 94
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 101
- read_inputstream_with_chunk
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 101
- register_extractor
 - plaso::tarzan::app::extractors::manager::Manager, 78
- register_package_dir
 - plaso::tarzan::lib::custom_pip_package_dir::CustomPipPackageDir, 26
- remove
 - plaso::tarzan::lib::hdfs::Hdfs, 42
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 94
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 102
- RemoveEventAttribute
 - plaso::tarzan::mediator::mediator::Mediator, 82
- result
 - plaso::tarzan::mediator::buffered_mediator, 17
- rm
 - plaso::tarzan::app::controllers::filemancontroller::FileManController, 33
- run_PE
 - plaso::tarzan::run, 18
- run_SQLite
 - plaso::tarzan::run, 18
- run_WinReg
 - plaso::tarzan::run, 18
- SCHEME
 - plaso::tarzan::lib::hdfs::Hdfs, 43
- seek
 - plaso::tarzan::dfvs2::hdfs_file_io::HDFSFile, 46
 - plaso::tarzan::file::file::LocalFileObject, 76
 - plaso::tarzan::file::generic_file::FileObject, 35
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 55
- seek_stream
 - plaso::tarzan::lib::hdfs::Hdfs, 43
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 94
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 102
- seek_stream_with_path
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 102
- SetFileEntry
 - plaso::tarzan::mediator::mediator::Mediator, 82
- spark_context
 - plaso::tarzan::app::controllers::plasocontroller::PlasoController, 87

- plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 104
- SplitPath
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔
FileSystem, 59
- strip_hdfs_uri
 - plaso::tarzan::app::controllers::controller::↔
Controller, 23
- TYPE_INDICATOR_HDFS
 - plaso::tarzan::dfvfs2::definitions, 14
- TYPE_INDICATOR
 - plaso::tarzan::dfvfs2::hdfs_file_entry::HDFSFile↔
Entry, 50
 - plaso::tarzan::dfvfs2::hdfs_file_system::HDFS↔
FileSystem, 60
 - plaso::tarzan::dfvfs2::hdfs_path_spec::HDFS↔
PathSpec, 65
 - plaso::tarzan::dfvfs2::hdfs_resolver_helper::HDF↔
SResolverHelper, 67
- tell
 - plaso::tarzan::file::file::LocalFileObject, 76
 - plaso::tarzan::file::hdfs_file::HdfsFileObject, 55
- temporary_directory
 - plaso::tarzan::mediator::mediator::Mediator, 83
- transform_files_rdd_to_extracted_events_rdd
 - plaso::tarzan::app::pyspark_plaso::PySparkPlaso,
106
- uri_class
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 104
- write_outputstream
 - plaso::tarzan::lib::hdfs::Hdfs, 43
 - plaso::tarzan::lib::pyarrow_hdfs::PyArrowHdfs, 95
 - plaso::tarzan::lib::pyspark_hdfs::PySparkHdfs, 103
- zip_put
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 33
- zip_put_form
 - plaso::tarzan::app::controllers::filemancontroller↔
::FileManController, 33