Dremio Cloner is a python-based utility for Dremio Enterprise. It supports the following commands: get, put, cascade-acl, report-acl.

Command "get"

Command "get" selectively saves definitions for objects such as Source, Space, Folder, PDS, VDS, ACLs, Reflections, Queues, Rules, Tags, Wikis, and Votes from a Dremio environment into a JSON file.

The command is configured with a JSON file with configuration attributes listed below. For detailed description of the configuration JSON attributes, see Reference section below in Appendix 1.

- "command":"get"
- "source": defines Dremio Environment with
  - "endpoint"
  - "username"
  - "password"
  - "verify_ssl"
  - "is_community_edition"
  - "graph_api_support"
- "target": defines an output filename with
  - "filename"
- "options":
  - defines logging options with
    - "logging.level"
    - "logging.format"
    - "logging.filename"
    - "logging.verbose"
  - miscellaneous options with
    - "max_errors"
    - "http_timeout"
  - defines scope of Space processing with
    - "space.filter"
    - "space.exclude.filter"
    - "space.folder.filter"
    - "space.folder.exclude.filter"
  - defines scope of Source processing with
    - "source.filter"
    - "source.exclude.filter"
    - "source.folder.filter"
    - "source.folder.exclude.filter"

- defines scope of PDS processing with
    - "pds.filter"
    - "pds.exclude.filter
    - "pds.list.useapi"
- defines scope of VDS processing with
    - "vds.filter"
    - "vds.exclude.filter"
    - " vds.dependencies.process_mode"

Please see a sample JSON configuration file in the Appendix 2.

Command "put"

Command "put" selectively updates an existing Dremio Environment based on a JSON file previously generated by "get" command.

The command is configured with a JSON file with configuration attributes listed below. For detailed description of the configuration JSON attributes, see Reference section below in Appendix 1.

- "command":"put"
- "source": defines an output filename with
    - "filename"
- "target": defines Dremio Environment with
    - "endpoint"
    - "username"
    - "password"
    - "verify_ssl"
    - "is_community_edition"
- "options":
    - defines logging options with
        - "logging.level"
        - "logging.format"
        - "logging.filename
        - "logging.verbose"
    - miscellaneous options with
        - "max_errors"
        - "http_timeout"
        - "source.retry_timedout"
        - "dry_run"
    - Defines scope of User processing with
        - "user.process_mode"

- "space.ignore_missing_acl_user"
- "space.ignore_missing_acl_group"
- "folder.ignore_missing_acl_user"
- "folder.ignore_missing_acl_group"
- "source.ignore_missing_acl_user"
- "source.ignore_missing_acl_group"
- "pds.ignore_missing_acl_user"
- "pds.ignore_missing_acl_group"
- "vds.ignore_missing_acl_user"
- "vds.ignore_missing_acl_group"
  - defines scope of Space processing with
    - "space.process_mode"
    - "space.filter"
    - "space.exclude.filter"
    - "folder.process_mode"
    - "space.folder.filter"
    - "space.folder.exclude.filter"
  - defines scope of Source processing with
    - "source.process_mode"
    - "source.filter"
    - "source.exclude.filter"
    - "source.folder.filter"
    - "source.folder.exclude.filter"
  - defines scope of PDS processing with
    - "pds.process_mode"
    - "pds.filter"
    - "pds.exclude.filter
    - "pds.list.useapi"
  - defines scope of VDS processing with
    - "vds.process_mode"
    - "vds.filter"
    - "vds.exclude.filter"
    - "vds._max_hierarchy_depth"

Please see a sample JSON configuration file in the Appendix 3.

Command "cascade-acl"

Command "cascade-acl" selectively propagates ACLs in an object hierarchy.

The command is configured with a JSON file with configuration attributes listed below. For detailed description of the configuration JSON attributes, see Reference section below in Appendix 1.

- "command":"cascade-acl"
- "target": defines Dremio Environment with
  - "endpoint"
  - "username"
  - "password"
  - "verify_ssl"
- "options":
  - defines logging options with
    - "logging.level"
    - "logging.format"
    - "logging.filename
    - "logging.verbose"
  - miscellaneous options with
    - "max_errors"
    - "http_timeout"
    - "source.retry_timedout"
    - "dry_run"
  - defines scope of Space processing with
    - "space.filter"
    - "space.exclude.filter"
    - "space.cascade-acl-origin.override-object"
    - "space.folder.filter"
    - "space.folder.exclude.filter"
    - "space.folder.cascade-acl-origin.filter"
  - defines scope of Source processing with
    - "source.filter"
    - "source.exclude.filter"
    - "source.cascade-acl-origin.override-object"
    - "source.folder.filter"
    - "source.folder.exclude.filter"
  - defines scope of PDS processing with
    - "pds.filter"
    - "pds.exclude.filter
    - "pds.list.useapi"
  - defines scope of VDS processing with
    - "vds.filter"
    - "vds.exclude.filter"

Note, if none of *space.cascade-acl-origin.override-object, space.folder.cascade-acl-origin.filter, source.cascade-acl-origin.override-object* specified:
- each Space ACL will be propagated through its hierarchy and applied to Folders and VDSs as per filter configuration

- each Source ACL will be propagated through its hierarchy and applied to PDSs as per filter configuration

Please see a sample JSON configuration file in the Appendix 4.

Command "report-acl"

Command "report-acl" command produces a security report on all objects in a Dremio environment.

The command is configured with a JSON file with configuration attributes listed below. For detailed description of the configuration JSON attributes, see Reference section below in Appendix 1.

- "command":"report-acl"
- "source": defines Dremio Environment with
  - "endpoint"
  - "username"
  - "password"
  - "verify_ssl"
- "target": defines an output filename with
  - "filename"
- "options":
  - defines logging options with
    - "logging.level"
    - "logging.format"
    - "logging.filename
    - "logging.verbose"
  - miscellaneous options with
    - "max_errors"
    - "http_timeout"
    - "source.retry_timedout"
  - defines scope of Space processing with
    - "report.csv.delimiter"
    - "report.csv.newline"
  - defines scope of Space processing with
    - "space.filter"
    - "space.exclude.filter"
    - "space.folder.filter"
    - "space.folder.exclude.filter"

- defines scope of Source processing with
  - "source.filter"
  - "source.exclude.filter"
  - "source.folder.filter"
  - "source.folder.exclude.filter"
- defines scope of PDS processing with
  - "pds.filter"
  - "pds.exclude.filter
  - "pds.list.useapi"
- defines scope of VDS processing with
  - "vds.filter"
  - "vds.exclude.filter"

Note, that this command does not provide any option for Scope definition. Please see a sample JSON configuration file in the Appendix 5.

Appendix 1. Reference - Configuration Options

Configuration Option	Description
target or *source* section, when Dremio Environment definition is required.	
endpoint	Defines Dremio API endpoint. For example, http://localhost:9047/. Mandatory attribute.
username	Dremio user name. Must be an Admin. Mandatory attribute.
password	Dremio user password. Optional field. If not provided, CLI will request password at runtime.
verify_ssl	If set to False, Dremio Cloner will not validate SSL certificate of the Dremio Environment. Default is True.
is_community_edition	Set it to True if reading Dremio CE. Writing to Dremio CE is not supported.
graph_api_support	Dremio Graph API is only available in EE starting version 4.0.0. Default value is False.
target or *source* section, when defined with a file name.	
filename	Defines a filename for input or output.
logging options	
logging.level	Defines logging level: DEBUG, INFO, WARN, ERROR
logging.format	Logging format. For example: "%(levelname)s:%(asctime)s:%(message)s"
logging.filename	Filename for logging. File will be appended if exists. If this option is omitted, standard output will be used for logging.
logging.verbose	Default False. Produce verbose logging: - Log entire entity definitions.
Miscellaneous options	
max_errors	Defines a number of errors at which processing will be terminated.
http_timeout	Timeout for each API call. This parameter might become important in certain situations when Sources defined in Dremio are not available.
dry_run	Defines a Dremio Cloner execution that will not update a target Dremio environment. In conjunction with logging.level set to WARN allows to execute Dremio Cloner without an impact on the target environment and check the log file for all activities that would have been submitted to the target Dremio Environment otherwise. Respective log entries will include *dry_run* keyword.
vds._max_hierarchy_depth	Defines maximum level of VDS hierarchy supported by Dremio Cloner. It's a guard rail with default value of 100.
Scope of Dremio Space processing	

space.filter	A filter that defines what Spaces will be **included** into processing. "*" will include all Spaces. Empty field will exclude all Spaces. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *space.exclude.filter*.
space.exclude.filter	A filter that defines what Spaces will be **excluded** into processing. "*" will exclude all Spaces. Empty field will include all Spaces. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *space.filter*.
space.folder.filter	A filter that defines what Space Folders will be **included** into processing. "*" will include all Folders. Empty field will exclude all Folders. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *space.folder.exclude.filter*.
space.folder.exclude.filter	A filter that defines what Space Folders will be **excluded** into processing. "*" will exclude all Folders. Empty field will include all Spaces. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *space.folder.filter*.
Scope of Dremio Space processing	
source.filter	A filter that defines what Sources will be **included** into processing. "*" will include all Sources. Empty field will exclude all Sources. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *source.exclude.filter*.
source.exclude.filter	A filter that defines what Spaces will be **excluded** into processing. "*" will exclude all Spaces. Empty field will include all Sources. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *source.filter*.
source.folder.filter	A filter that defines what Source Folders will be **included** into processing. "*" will include all Folders. Empty field will exclude all Folders. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *source.exclude.filter*.
source.folder.exclude.filter	A filter that defines what Source Folders will be **excluded** into processing. "*" will exclude all Folders. Empty field will include all Spaces. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *source.filter*.
Scope of Dremio PDS processing	

pds.filter	A filter that defines what PDSs will be **included** into processing. "*" will include all PDSs. Empty field will exclude all PDSs. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *pds.exclude.filter*.
pds.exclude.filter	A filter that defines what PDSs will be **excluded** into processing. "*" will exclude all PDSs. Empty field will include all PDSs. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *pds.filter*.
"pds.list.useapi"	Forces to use API for collecting list of PDSs if set to True. Default value is False which means that INFOMRATION_SCHEMA will be utilized instead of API. False is a recommended value.
Scope of Dremio VDS processing	
vds.filter	A filter that defines what VDSs will be **included** into processing. "*" will include all VDSs. Empty field will exclude all VDSs. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *vds.exclude.filter*.
vds.exclude.filter	A filter that defines what VDSs will be **excluded** into processing. "*" will exclude all VDSs. Empty field will include all VDSs. Star may be used multiple times in the filter to define a pattern. Works in logical AND with *vds.filter*.
Scope of user and group processing	
user.process_mode	Determines if users will be created in the target Dremio Environment if they are referenced in the source JSON file but not in the target environment. Applicable for "put" command only. However, user creation is not possible with the current Dremio API. This parameter can only have a single value *skip*.
space.ignore_missing_acl_user *space.ignore_missing_acl_group* *folder.ignore_missing_acl_user* *folder.ignore_missing_acl_group* *source.ignore_missing_acl_user* *source.ignore_missing_acl_group* *pds.ignore_missing_acl_user* *pds.ignore_missing_acl_group* *vds.ignore_missing_acl_user* *vds.ignore_missing_acl_group*	These configuration parameters define if Dremio Cloner ignores a situation when a user or a group is defined in an ACL in the source JSON file but is not present in the target Dremio Environment. This situation is a potential security risk as an ACL may be created with no limitations in the target environment when all referenced users and groups cannot be found. Default value is False.

space.process_mode folder.process_mode source.process_mode pds.process_mode vds.process_mode	Defines whether Dremio Cloner will - insert new objects only or - update existing objects only or - do an upsert. These parameters can be set to: *skip*, *create_only*, *update_only*, *create_overwrite*. *skip* will prevent any changes to the target Dremio Environment for the specified object type. Note, *pds.process_mode* can only take *skip* and *promote.* An example of usage would be the following settings: *space.process_mode: skip* *folder.process_mode: skip* *vds.process_mode: create_overwrite* This configuration will not update any Space or Folder. However, it will update VDSs as per VDS related filters.
vds.dependencies.process_mode	Possible values: *ignore*, *get*. Default *ignore*. If set to *get*, Dremio Cloner will collect information on all decencies throughout the object hierarchy (VDS and PDS) required for each VDS that satisfies VDS filter criteria.
Cascade-acl specific parameters	
space.cascade-acl-origin.override-object	If specified, overrides default behavior for Space hierarchy and an ACL of the object specified in this parameter will be used through **all Spaces all hierarchies** instead of the respective Spaces' ACLs.
source.cascade-acl-origin.override-object	If specified, overrides default behavior for Source hierarchy and an ACL of the object specified in this parameter will be used through **all Source all hierarchies** instead of the respective Sources' ACLs.
space.folder.cascade-acl-origin.filter	If specified, overrides default behavior for Space hierarchy and an ACLs of the Folders selected by this will be used through **its Folder hierarchy** instead of the respective Source's ACL.
Report-acl specific parameters	
report.csv.delimiter	A field delimiter used to generate a report.
report.csv.newline	A new line delimiter used to generate a report.

Appendix 2. Sample GET command JSON configuration file

```json
{"dremio_cloner": [
    {"command":"get"},
    {"source": [
        {"endpoint": "http://localhost:9047/"},
        {"username": "mikhail"},
        {"password": ""},
        {"verify_ssl": "False"}]
    },
    {"target": [
        {"filename":"dremio_backup.json"}]
    },
    {"options": [
        {"logging.level":"logging.INFO"},
        {"logging.format":"%(levelname)s:%(asctime)s:%(message)s"},
        {"logging.filename":""},

        {"max_errors":"9999"},
        {"http_timeout":"10"},

        {"space.filter": "*"},
        {"space.exclude.filter": ""},
        {"space.folder.filter":"*"},
        {"space.folder.exclude.filter":""},

        {"source.filter":"*"},
        {"source.exclude.filter":""},
        {"source.folder.filter":"*"},
        {"source.folder.exclude.filter":""},

        {"pds.filter":"*"},
        {"pds.exclude.filter":""},
        {"pds.list.useapi":"False"},

        {"vds.filter":"*"},
        {"vds.exclude.filter":""} ]
    }]
}
```

Appendix 3. Sample PUT command JSON configuration file

```json
{"dremio_cloner": [
    {"command":"put"},
    {"target": [
        {"endpoint": "http://localhost:9047/"},
        {"username": "mikhail"},
        {"password": ""},
        {"verify_ssl": "False"}]
    },
    {"source": [
        {"filename":"dremio_backup.json"}]
    },
    {"options": [

        {"dry_run":"True"},

        {"logging.level":"logging.INFO"},
        {"logging.format":"%(levelname)s:%(asctime)s:%(message)s"},
        {"logging.filename":""},

        {"max_errors":"9999"},
        {"http_timeout":"10"},
        {"source.retry_timedout":"false"},

        {"user.process_mode":"skip"},
        {"space.process_mode":"create_overwrite"},
        {"folder.process_mode":"create_overwrite"},
        {"source.process_mode":"create_overwrite"},
        {"pds.process_mode":"promote"},
        {"vds.process_mode":"create_overwrite"},

        {"space.ignore_missing_acl_user":"True"},
        {"space.ignore_missing_acl_group":"True"},
        {"folder.ignore_missing_acl_user":"True"},
        {"folder.ignore_missing_acl_group":"True"},
        {"source.ignore_missing_acl_user":"True"},
        {"source.ignore_missing_acl_group":"True"},
        {"pds.ignore_missing_acl_user":"True"},
        {"pds.ignore_missing_acl_group":"True"},
        {"vds.ignore_missing_acl_user":"True"},
        {"vds.ignore_missing_acl_group":"True"},

        {"space.filter": "*"},
        {"space.exclude.filter": ""},
        {"space.folder.filter":"*"},
        {"space.folder.exclude.filter":""},

        {"source.filter":"*"},
        {"source.exclude.filter":""},
        {"source.folder.filter":"*"},
        {"source.folder.exclude.filter":""},

        {"pds.filter":"*"},
        {"pds.exclude.filter":""},
        {"pds.list.useapi":"False"},

        {"vds.filter":"*"},
        {"vds.exclude.filter":""},
        {"vds._max_hierarchy_depth":"100"}]
```

```
    }]
}
```

Appendix 4. Sample CASCADE-ACL command JSON configuration file

```
{"dremio_cloner": [
    {"command": "cascade-acl"},
    {"target": [
        {"endpoint": "http://localhost:9047/"},
        {"username": "mikhail"},
        {"password": ""},
        {"verify_ssl": "False"}]
    },
    {"options": [

        {"dry_run": "True"},

        {"logging.level": "logging.INFO"},
        {"logging.format": "%(levelname)s:%(asctime)s:%(message)s"},
        {"logging.filename": ""},

        {"max_errors": "9999"},
        {"http_timeout": "10"},

        {"space.filter": "*"},
        {"space.exclude.filter": ""},
        {"space.cascade-acl-origin.override-object": "TestSpace/TestVDS"},
        {"space.folder.filter": "*"},
        {"space.folder.exclude.filter": ""},
        {"space.folder.cascade-acl-origin.filter": "TestFolder"},

        {"source.filter": "*"},
        {"source.exclude.filter": ""},
        {"source.cascade-acl-origin.override-object": "TestSource/TestFolder/TestPDS"},
        {"source.folder.filter": "*"},
        {"source.folder.exclude.filter": ""},

        {"pds.list.useapi": "False"},
        {"pds.filter": "*"},
        {"pds.exclude.filter": ""},

        {"vds.filter": "*"},
        {"vds.exclude.filter": ""} ]
    } ]
}
```

Appendix 5. Sample REPORT-ACL command JSON configuration file

```
{"dremio_cloner": [
    {"command":"report-acl"},
    {"source": [
        {"endpoint": "http://localhost:9047/"},
        {"username": "mikhail"},
        {"password": ""},
        {"verify_ssl": "False"}]
    },
    {"target": [
        {"filename":"dremio_audit_report.csv"}]
    },
    {"options": [

        {"logging.level":"logging.WARN"},
        {"logging.format":"%(levelname)s:%(asctime)s:%(message)s"},
        {"logging.filename":""},

        {"max_errors":"9999"},
        {"http_timeout":"10"},

        {"report.csv.delimiter":"\t"},
        {"report.csv.newline":"\n"}]
    }]
}
```