Apache Ignite binary client Python API Documentation

Apache Software Foundation (ASF)

Jun 18, 2021

Contents:

CHAPTER 1

Basic Information

1.1 What it is

This is an Apache Ignite thin (binary protocol) client library, written in Python 3, abbreviated as *pyignite*.

Apache Ignite is a memory-centric distributed database, caching, and processing platform for transactional, analytical, and streaming workloads delivering in-memory speeds at petabyte scale.

Ignite binary client protocol provides user applications the ability to communicate with an existing Ignite cluster without starting a full-fledged Ignite node. An application can connect to the cluster through a raw TCP socket.

1.2 Prerequisites

- *Python 3.6* or above (3.6, 3.7, 3.8 and 3.9 are tested),
- Access to *Apache Ignite* node, local or remote. The current thin client version was tested on *Apache Ignite 2.10.0* (binary client protocol 1.7.0).

1.3 Installation

1.3.1 for end user

If you want to use *pyignite* in your project, you may install it from PyPI:

```
$ pip install pyignite
```

1.3.2 for developer

If you want to run tests, examples or build documentation, clone the whole repository:

```
$ git clone git@github.com:apache/ignite-python-thin-client.git
$ pip install -e .
```

This will install the repository version of *pyignite* into your environment in so-called "develop" or "editable" mode. You may read more about editable installs in the *pip* manual.

Then run through the contents of *requirements* folder to install the the additional requirements into your working Python environment using

```
$ pip install -r requirements/<your task>.txt
```

For development, it is recommended to install *tests* requirements

```
$ pip install -r requirements/tests.txt
```

For checking codestyle run:

```
$ flake8
```

You may also want to consult the setuptools manual about using *setup.py*.

1.4 Examples

Some examples of using pyignite are provided in *examples* folder. They are extensively commented in the *Examples of usage* section of the documentation.

This code implies that it is run in the environment with *pyignite* package installed, and Apache Ignite node is running on localhost:10800, unless otherwise noted.

There is also a possibility to run examples alone with tests. For the explanation of testing, look up the *Testing* section.

1.5 Testing

Create and activate virtualenv environment.

Install a binary release of Ignite with *log4j2* enabled and set *IGNITE_HOME* environment variable.

```
$ cd <ignite_binary_release>
$ export IGNITE_HOME=$(pwd)
$ cp -r $IGNITE_HOME/libs/optional/ignite-log4j2 $IGNITE_HOME/libs/
```

Run

```
$ pip install -e .
$ pytest
```

Other *pytest* parameters:

`--examples` run the examples as one test. If you wish to run *only* the examples, supply also the name of the test function to *pytest* launcher:

```
$ pytest --examples ../tests/test_examples.py::test_examples
```

In this test assertion fails if any of the examples' processes ends with non-zero exit code.

Examples are not parameterized for the sake of simplicity. They always run with default parameters (host and port) regardless of any other *pytest* option.

Since failover, SSL and authentication examples are meant to be controlled by user or depend on special configuration of the Ignite cluster, they can not be automated.

1.5.1 Using tox

For automate running tests against different python version, it is recommended to use tox

```
$ pip install tox
$ tox
```

1.6 Documentation

To recompile this documentation, do this from your virtualenv environment:

```
$ pip install -r requirements/docs.txt
$ cd docs
$ make html
```

Then open docs/generated/html/index.html in your browser.

If you feel that old version is stuck, do

```
$ make clean
$ sphinx-apidoc -feM -o source/ ../ ../setup.py
$ make html
```

And that should be it.

1.7 Licensing

This is a free software, brought to you on terms of the Apache License v2.

CHAPTER 2

API Specification

The modules and subpackages listed here are the basis of a stable API of *pyignite*, intended for end users.

2.1 pyignite.client module

This module contains *Client* class, that lets you communicate with Apache Ignite cluster node by the means of Ignite binary client protocol.

To start the communication, you may connect to the node of their choice by instantiating the *Client* object and calling `connect()` method with proper parameters.

The whole storage room of Ignite cluster is split up into named structures, called caches. For accessing the particular cache in key-value style (a-la Redis or memcached) you should first create the *Cache* object by calling *create_cache()* or *get_or_create_cache()* methods, than call *Cache* methods. If you wish to create a cache object without communicating with server, there is also a *get_cache()* method that does just that.

For using Ignite SQL, call *sql()* method. It returns a generator with result rows.

register_binary_type() and *query_binary_type()* methods operates the local (class-wise) registry for Ignite Complex objects.

class pyignite.client.**Client**(*compact_footer: bool = None*, *partition_aware: bool = True*, ***kwargs*)

Synchronous Client implementation.

__init__(*compact_footer: bool = None*, *partition_aware: bool = True*, ***kwargs*)

Initialize client.

Parameters

- **compact_footer** – (optional) use compact (True, recommended) or full (False) schema approach when serializing Complex objects. Default is to use the same approach the server is using (None). Apache Ignite binary protocol documentation on this topic: https://ignite.apache.org/docs/latest/binary-client-protocol/data-format#schema

- **partition_aware** – (optional) try to calculate the exact data placement from the key before to issue the key operation to the server node, *True* by default.

close()

compact_footer

This property remembers Complex object schema encoding approach when decoding any Complex object, to use the same approach on Complex object encoding.

Returns True if compact schema was used by server or no Complex object decoding has yet taken place, False if full schema was used.

connect(**args*)

Connect to Ignite cluster node(s).

Parameters **args** – (optional) host(s) and port(s) to connect to.

create_cache(*settings: Union[str, dict]*) → pyignite.cache.Cache

Creates Ignite cache by name. Raises *CacheError* if such a cache is already exists.

Parameters **settings** – cache name or dict of cache properties' codes and values. All cache properties are documented here: *Cache Properties*. See also the *cache creation example*,

Returns `Cache` object.

get_best_node()

Returns the node from the list of the nodes, opened by client, that most probably contains the needed key-value pair. See IEP-23.

This method is not a part of the public API. Unless you wish to extend the *pyignite* capabilities (with additional testing, logging, examining connections, et c.) you probably should not use it.

Parameters

- **cache** – Ignite cache, cache name or cache id,
- **key** – (optional) pythonic key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns Ignite connection object.

get_binary_type(*binary_type: Union[str, int]*) → dict

Gets the binary type information from the Ignite server. This is quite a low-level implementation of Ignite thin client protocol's *OP_GET_BINARY_TYPE* operation. You would probably want to use `query_binary_type()` instead.

Parameters **binary_type** – binary type name or ID,

Returns

binary type description a dict with the following fields:

- *type_exists*: True if the type is registered, False otherwise. In the latter case all the following fields are omitted,
- *type_id*: Complex object type ID,
- *type_name*: Complex object type name,
- *affinity_key_field*: string value or None,
- *is_enum*: False in case of Complex object registration,
- *schemas*: a list, containing the Complex object schemas in format: OrderedDict[field name: field type hint]. A schema can be empty.

get_cache(*settings: Union[str, dict]*) → pyignite.cache.Cache
Creates Cache object with a given cache name without checking it up on server. If such a cache does not exist, some kind of exception (most probably *CacheError*) may be raised later.

Parameters **settings** – cache name or cache properties (but only *PROP_NAME* property is allowed),

Returns *Cache* object.

get_cache_names() → list
Gets existing cache names.

Returns list of cache names.

get_cluster() → pyignite.cluster.Cluster
Get client cluster facade.

Returns `Cluster` instance.

get_or_create_cache(*settings: Union[str, dict]*) → pyignite.cache.Cache
Creates Ignite cache, if not exist.

Parameters **settings** – cache name or dict of cache properties' codes and values. All cache properties are documented here: *Cache Properties*. See also the *cache creation example*,

Returns *Cache* object.

partition_aware

partition_awareness_supported_by_protocol

protocol_context
Returns protocol context, or None, if no connection to the Ignite cluster was not yet established.

This method is not a part of the public API. Unless you wish to extend the *pyignite* capabilities (with additional testing, logging, examining connections, et c.) you probably should not use it.

put_binary_type(*type_name: str*, *affinity_key_field: str = None*, *is_enum=False*, *schema: dict = None*)
Registers binary type information in cluster. Do not update binary registry. This is a literal implementation of Ignite thin client protocol's *OP_PUT_BINARY_TYPE* operation. You would probably want to use *register_binary_type()* instead.

Parameters

- **type_name** – name of the data type being registered,
- **affinity_key_field** – (optional) name of the affinity key field,
- **is_enum** – (optional) register enum if True, binary object otherwise. Defaults to False,
- **schema** – (optional) when register enum, pass a dict of enumerated parameter names as keys and an integers as values. When register binary type, pass a dict of field names: field types. Binary type with no fields is OK.

query_binary_type(*binary_type: Union[int, str], schema: Union[int, dict] = None*)
Queries the registry of Complex object classes.

Parameters

- **binary_type** – Complex object type name or ID,
- **schema** – (optional) Complex object schema or schema ID

Returns found dataclass or None, if *schema* parameter is provided, a dict of {schema ID: dataclass} format otherwise.

random_node
Returns random usable node.

This method is not a part of the public API. Unless you wish to extend the *pyignite* capabilities (with additional testing, logging, examining connections, et c.) you probably should not use it.

register_binary_type(*data_class: Type[CT_co], affinity_key_field: str = None*)
Register the given class as a representation of a certain Complex object type. Discards autogenerated or previously registered class.

Parameters

- **data_class** – Complex object class,
- **affinity_key_field** – (optional) affinity parameter.

register_cache(*cache_id: int*)

sql(*query_str: str*, *page_size: int = 1024*, *query_args: Iterable[T_co] = None*, *schema: str = 'PUBLIC'*, *statement_type: int = 0*, *distributed_joins: bool = False*, *local: bool = False*, *replicated_only: bool = False*, *enforce_join_order: bool = False*, *collocated: bool = False*, *lazy: bool = False*, *include_field_names: bool = False*, *max_rows: int = -1*, *timeout: int = 0*, *cache: Union[int, str, pyignite.cache.Cache] = None*) → pyignite.cursors.SqlFieldsCursor
Runs an SQL query and returns its result.

Parameters

- **query_str** – SQL query string,
- **page_size** – (optional) cursor page size. Default is 1024, which means that client makes one server call per 1024 rows,
- **query_args** – (optional) query arguments. List of values or (value, type hint) tuples,
- **schema** – (optional) schema for the query. Defaults to *PUBLIC*,
- **statement_type** – (optional) statement type. Can be:
  - StatementType.ALL any type (default),
  - StatementType.SELECT select,
  - StatementType.UPDATE update.
- **distributed_joins** – (optional) distributed joins. Defaults to False,
- **local** – (optional) pass True if this query should be executed on local node only. Defaults to False,
- **replicated_only** – (optional) whether query contains only replicated tables or not. Defaults to False,
- **enforce_join_order** – (optional) enforce join order. Defaults to False,
- **collocated** – (optional) whether your data is co-located or not. Defaults to False,
- **lazy** – (optional) lazy query execution. Defaults to False,
- **include_field_names** – (optional) include field names in result. Defaults to False,
- **max_rows** – (optional) query-wide maximum of rows. Defaults to -1 (all rows),
- **timeout** – (optional) non-negative timeout value in ms. Zero disables timeout (default),
- **cache** – (optional) Name or ID of the cache to use to infer schema. If set, 'schema' argument is ignored,

Returns sql fields cursor with result rows as a lists. If *include_field_names* was set, the first row will hold field names.

tx_start(*concurrency: pyignite.datatypes.transactions.TransactionConcurrency = <TransactionConcurrency.PESSIMISTIC: 1>*, *isolation: pyignite.datatypes.transactions.TransactionIsolation = <TransactionIsolation.REPEATABLE_READ: 1>*, *timeout: int = 0*, *label: Union[str, NoneType] = None*) → pyignite.transaction.Transaction

Start thin client transaction.

Parameters

- **concurrency** – (optional) transaction concurrency, see `TransactionConcurrency`,
- **isolation** – (optional) transaction isolation level, see `TransactionIsolation`,
- **timeout** – (optional) transaction timeout in milliseconds,
- **label** – (optional) transaction label.

Returns `Transaction` instance.

unwrap_binary(*value: Any*) → Any

Detects and recursively unwraps Binary Object or collections of BinaryObject.

Parameters **value** – anything that could be a Binary Object or collection of BinaryObject,

Returns the result of the Binary Object unwrapping with all other data left intact.

2.2 pyignite.aio_client module

class pyignite.aio_client.**AioClient**(*compact_footer: bool = None*, *partition_aware: bool = True*, ***kwargs*)

Bases: `pyignite.client.BaseClient`

Asynchronous Client implementation.

__init__(*compact_footer: bool = None*, *partition_aware: bool = True*, ***kwargs*)

Initialize client.

Parameters

- **compact_footer** – (optional) use compact (True, recommended) or full (False) schema approach when serializing Complex objects. Default is to use the same approach the server is using (None). Apache Ignite binary protocol documentation on this topic: https://ignite.apache.org/docs/latest/binary-client-protocol/data-format#schema
- **partition_aware** – (optional) try to calculate the exact data placement from the key before to issue the key operation to the server node, *True* by default.

close()

connect(**args*)

Connect to Ignite cluster node(s).

Parameters **args** – (optional) host(s) and port(s) to connect to.

create_cache(*settings: Union[str, dict]*) → pyignite.aio_cache.AioCache

Creates Ignite cache by name. Raises *CacheError* if such a cache is already exists.

Parameters **settings** – cache name or dict of cache properties' codes and values. All cache properties are documented here: *Cache Properties*. See also the *cache creation example*,

Returns *Cache* object.

get_best_node()
Returns the node from the list of the nodes, opened by client, that most probably contains the needed key-value pair. See IEP-23.

This method is not a part of the public API. Unless you wish to extend the *pyignite* capabilities (with additional testing, logging, examining connections, et c.) you probably should not use it.

Parameters

- **cache** – Ignite cache, cache name or cache id,
- **key** – (optional) pythonic key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns Ignite connection object.

get_binary_type(*binary_type: Union[str, int]*) → dict
Gets the binary type information from the Ignite server. This is quite a low-level implementation of Ignite thin client protocol's *OP_GET_BINARY_TYPE* operation. You would probably want to use *query_binary_type()* instead.

Parameters **binary_type** – binary type name or ID,

Returns

binary type description a dict with the following fields:

- *type_exists*: True if the type is registered, False otherwise. In the latter case all the following fields are omitted,
- *type_id*: Complex object type ID,
- *type_name*: Complex object type name,
- *affinity_key_field*: string value or None,
- *is_enum*: False in case of Complex object registration,
- *schemas*: a list, containing the Complex object schemas in format: OrderedDict[field name: field type hint]. A schema can be empty.

get_cache(*settings: Union[str, dict]*) → pyignite.aio_cache.AioCache
Creates Cache object with a given cache name without checking it up on server. If such a cache does not exist, some kind of exception (most probably *CacheError*) may be raised later.

Parameters **settings** – cache name or cache properties (but only *PROP_NAME* property is allowed),

Returns *Cache* object.

get_cache_names() → list
Gets existing cache names.

Returns list of cache names.

get_cluster() → pyignite.aio_cluster.AioCluster
Get client cluster facade.

Returns `AioCluster` instance.

get_or_create_cache(*settings: Union[str, dict]*) → pyignite.aio_cache.AioCache
Creates Ignite cache, if not exist.

Parameters **settings** – cache name or dict of cache properties' codes and values. All cache properties are documented here: *Cache Properties*. See also the *cache creation example*,

Returns `Cache` object.

put_binary_type(*type_name: str*, *affinity_key_field: str = None*, *is_enum=False*, *schema: dict = None*)
Registers binary type information in cluster. Do not update binary registry. This is a literal implementation of Ignite thin client protocol's *OP_PUT_BINARY_TYPE* operation. You would probably want to use `register_binary_type()` instead.

Parameters

- **type_name** – name of the data type being registered,
- **affinity_key_field** – (optional) name of the affinity key field,
- **is_enum** – (optional) register enum if True, binary object otherwise. Defaults to False,
- **schema** – (optional) when register enum, pass a dict of enumerated parameter names as keys and an integers as values. When register binary type, pass a dict of field names: field types. Binary type with no fields is OK.

query_binary_type(*binary_type: Union[int, str]*, *schema: Union[int, dict] = None*)
Queries the registry of Complex object classes.

Parameters

- **binary_type** – Complex object type name or ID,
- **schema** – (optional) Complex object schema or schema ID,

Returns found dataclass or None, if *schema* parameter is provided, a dict of {schema ID: dataclass} format otherwise.

random_node() → pyignite.connection.aio_connection.AioConnection
Returns random usable node.

This method is not a part of the public API. Unless you wish to extend the *pyignite* capabilities (with additional testing, logging, examining connections, et c.) you probably should not use it.

register_binary_type(*data_class: Type[CT_co]*, *affinity_key_field: str = None*)
Register the given class as a representation of a certain Complex object type. Discards autogenerated or previously registered class.

Parameters

- **data_class** – Complex object class,
- **affinity_key_field** – (optional) affinity parameter.

sql()
Runs an SQL query and returns its result.

Parameters

- **query_str** – SQL query string,
- **page_size** – (optional) cursor page size. Default is 1024, which means that client makes one server call per 1024 rows,
- **query_args** – (optional) query arguments. List of values or (value, type hint) tuples,
- **schema** – (optional) schema for the query. Defaults to *PUBLIC*,
- **statement_type** – (optional) statement type. Can be:

  - StatementType.ALL any type (default),
  - StatementType.SELECT select,
  - StatementType.UPDATE update.
- **distributed_joins** – (optional) distributed joins. Defaults to False,
- **local** – (optional) pass True if this query should be executed on local node only. Defaults to False,
- **replicated_only** – (optional) whether query contains only replicated tables or not. Defaults to False,
- **enforce_join_order** – (optional) enforce join order. Defaults to False,
- **collocated** – (optional) whether your data is co-located or not. Defaults to False,
- **lazy** – (optional) lazy query execution. Defaults to False,
- **include_field_names** – (optional) include field names in result. Defaults to False,
- **max_rows** – (optional) query-wide maximum of rows. Defaults to -1 (all rows),
- **timeout** – (optional) non-negative timeout value in ms. Zero disables timeout (default),
- **cache** – (optional) Name or ID of the cache to use to infer schema. If set, 'schema' argument is ignored,

Returns async sql fields cursor with result rows as a lists. If *include_field_names* was set, the first row will hold field names.

tx_start(*concurrency: pyignite.datatypes.transactions.TransactionConcurrency = <TransactionConcurrency.PESSIMISTIC: 1>*, *isolation: pyignite.datatypes.transactions.TransactionIsolation = <TransactionIsolation.REPEATABLE_READ: 1>*, *timeout: int = 0*, *label: Union[str, NoneType] = None*) → pyignite.transaction.AioTransaction

Start async thin client transaction. **Supported only python 3.7+**

Parameters

- **concurrency** – (optional) transaction concurrency, see `TransactionConcurrency`,
- **isolation** – (optional) transaction isolation level, see `TransactionIsolation`,
- **timeout** – (optional) transaction timeout in milliseconds,
- **label** – (optional) transaction label.

Returns `AioTransaction` instance.

unwrap_binary(*value: Any*) → Any

Detects and recursively unwraps Binary Object.

Parameters **value** – anything that could be a Binary Object,

Returns the result of the Binary Object unwrapping with all other data left intact.

2.3 pyignite.cache module

class pyignite.cache.**BaseCache**(*client: BaseClient*, *name: str*, *expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)

__init__(*client: BaseClient*, *name: str*, *expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)
Initialize self. See help(type(self)) for accurate signature.

cache_id
Cache ID.

Returns integer value of the cache ID.

cache_info
Cache meta info.

client

Returns Client object, through which the cache is accessed.

name

Returns cache name string.

with_expire_policy(*expiry_policy: Union[pyignite.datatypes.expiry_policy.ExpiryPolicy, NoneType] = None*, *create: Union[int, datetime.timedelta] = -2*, *update: Union[int, datetime.timedelta] = -2*, *access: Union[int, datetime.timedelta] = -2*)

Parameters

- **expiry_policy** – optional `ExpiryPolicy` object. If it is set, other params will be ignored,
- **create** – TTL for create in milliseconds or `timedelta`,
- **update** – TTL for update in milliseconds or `timedelta`,
- **access** – TTL for access in milliseconds or `timedelta`,

Returns cache decorator with expiry policy set.

class pyignite.cache.**Cache**(*client: Client*, *name: str*, *expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)
Ignite cache abstraction. Users should never use this class directly, but construct its instances with *create_cache()*, *get_or_create_cache()* or *get_cache()* methods instead. See *this example* on how to do it.

__init__(*client: Client*, *name: str*, *expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)
Initialize cache object. For internal use.

Parameters

- **client** – Ignite client,
- **name** – Cache name.

cache_id
Cache ID.

Returns integer value of the cache ID.

cache_info
Cache meta info.

clear(*keys: Union[list, NoneType] = None*)
Clears the cache without notifying listeners or cache writers.

Parameters **keys** – (optional) list of cache keys or (key, key type hint) tuples to clear (default: clear all).

clear_key(*key*, *key_hint: object = None*)
Clears the cache key without notifying listeners or cache writers.

Parameters

- **key** – key for the cache entry,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

clear_keys(*keys: Iterable[T_co]*)
Clears the cache key without notifying listeners or cache writers.

Parameters **keys** – a list of keys or (key, type hint) tuples

client

Returns Client object, through which the cache is accessed.

contains_key(*key*, *key_hint=None*) → bool
Returns a value indicating whether given key is present in cache.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns boolean *True* when key is present, *False* otherwise.

contains_keys(*keys: Iterable[T_co]*) → bool
Returns a value indicating whether all given keys are present in cache.

Parameters **keys** – a list of keys or (key, type hint) tuples,

Returns boolean *True* when all keys are present, *False* otherwise.

destroy()
Destroys cache with a given name.

get(*key*, *key_hint: object = None*) → Any
Retrieves a value from cache by key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns value retrieved.

get_all(*keys: list*) → dict
Retrieves multiple key-value pairs from cache.

Parameters **keys** – list of keys or tuples of (key, key_hint),

Returns a dict of key-value pairs.

get_and_put(*key*, *value*, *key_hint=None*, *value_hint=None*) → Any
Puts a value with a given key to cache, and returns the previous value for that key, or null value if there was not such key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

Returns old value or None.

get_and_put_if_absent(*key*, *value*, *key_hint=None*, *value_hint=None*)
Puts a value with a given key to cache only if the key does not already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted,

Returns old value or None.

get_and_remove(*key*, *key_hint=None*) → Any
Removes the cache entry with specified key, returning the value.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns old value or None.

get_and_replace(*key*, *value*, *key_hint=None*, *value_hint=None*) → Any
Puts a value with a given key to cache, returning previous value for that key, if and only if there is a value currently mapped for that key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

Returns old value or None.

get_size(*peek_modes=None*)
Gets the number of entries in cache.

Parameters **peek_modes** – (optional) limit count to near cache partition (PeekModes.NEAR), primary cache (PeekModes.PRIMARY), or backup cache (PeekModes.BACKUP). Defaults to primary cache partitions (PeekModes.PRIMARY),

Returns integer number of cache entries.

name

Returns cache name string.

put(*key*, *value*, *key_hint: object = None*, *value_hint: object = None*)
Puts a value with a given key to cache (overwriting existing value if any).

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

put_all(*pairs: dict*)
Puts multiple key-value pairs to cache (overwriting existing associations if any).

Parameters **pairs** – dictionary type parameters, contains key-value pairs to save. Each key or value can be an item of representable Python type or a tuple of (item, hint),

put_if_absent(*key*, *value*, *key_hint=None*, *value_hint=None*)
Puts a value with a given key to cache only if the key does not already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

remove_all()
Removes all cache entries, notifying listeners and cache writers.

remove_if_equals(*key*, *sample*, *key_hint=None*, *sample_hint=None*)
Removes an entry with a given key if provided value is equal to actual value, notifying listeners and cache writers.

Parameters

- **key** – key for the cache entry,
- **sample** – a sample to compare the stored value with,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **sample_hint** – (optional) Ignite data type, for whic the given sample should be converted.

remove_key(*key*, *key_hint=None*)
Clears the cache key without notifying listeners or cache writers.

Parameters

- **key** – key for the cache entry,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

remove_keys(*keys: list*)
Removes cache entries by given list of keys, notifying listeners and cache writers.

Parameters **keys** – list of keys or tuples of (key, key_hint) to remove.

replace(*key*, *value*, *key_hint: object = None*, *value_hint: object = None*)
Puts a value with a given key to cache only if the key already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

replace_if_equals(*key*, *sample*, *value*, *key_hint=None*, *sample_hint=None*, *value_hint=None*) → Any

Puts a value with a given key to cache only if the key already exists and value equals provided sample.

Parameters

- **key** – key for the cache entry,
- **sample** – a sample to compare the stored value with,
- **value** – new value for the given key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **sample_hint** – (optional) Ignite data type, for whic the given sample should be converted
- **value_hint** – (optional) Ignite data type, for which the given value should be converted,

Returns boolean *True* when key is present, *False* otherwise.

scan(*page_size: int = 1*, *partitions: int = -1*, *local: bool = False*) → pyignite.cursors.ScanCursor

Returns all key-value pairs from the cache, similar to *get_all*, but with internal pagination, which is slower, but safer.

Parameters

- **page_size** – (optional) page size. Default size is 1 (slowest and safest),
- **partitions** – (optional) number of partitions to query (negative to query entire cache),
- **local** – (optional) pass True if this query should be executed on local node only. Defaults to False,

Returns Scan query cursor.

select_row(*query_str: str*, *page_size: int = 1*, *query_args: Union[list, NoneType] = None*, *distributed_joins: bool = False*, *replicated_only: bool = False*, *local: bool = False*, *timeout: int = 0*) → pyignite.cursors.SqlCursor

Executes a simplified SQL SELECT query over data stored in the cache. The query returns the whole record (key and value).

Parameters

- **query_str** – SQL query string,
- **page_size** – (optional) cursor page size. Default is 1, which means that client makes one server call per row,
- **query_args** – (optional) query arguments,
- **distributed_joins** – (optional) distributed joins. Defaults to False,
- **replicated_only** – (optional) whether query contains only replicated tables or not. Defaults to False,
- **local** – (optional) pass True if this query should be executed on local node only. Defaults to False,
- **timeout** – (optional) non-negative timeout value in ms. Zero disables timeout (default),

Returns Sql cursor.

settings

Lazy Cache settings. See the *example* of reading this property.

All cache properties are documented here: *Cache Properties*.

Returns dict of cache properties and their values.

with_expire_policy(*expiry_policy: Union[pyignite.datatypes.expiry_policy.ExpiryPolicy, NoneType] = None, create: Union[int, datetime.timedelta] = -2, update: Union[int, datetime.timedelta] = -2, access: Union[int, datetime.timedelta] = -2*)

Parameters

- **expiry_policy** – optional `ExpiryPolicy` object. If it is set, other params will be ignored,
- **create** – TTL for create in milliseconds or `timedelta`,
- **update** – TTL for update in milliseconds or `timedelta`,
- **access** – TTL for access in milliseconds or `timedelta`,

Returns cache decorator with expiry policy set.

`pyignite.cache.`**create_cache**(*client: Client, settings: Union[str, dict]*) → Cache

`pyignite.cache.`**get_cache**(*client: Client, settings: Union[str, dict]*) → Cache

`pyignite.cache.`**get_or_create_cache**(*client: Client, settings: Union[str, dict]*) → Cache

2.4 pyignite.aio_cache module

class `pyignite.aio_cache.`**AioCache**(*client: AioClient, name: str, expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)

Bases: *pyignite.cache.BaseCache*

Ignite cache abstraction. Users should never use this class directly, but construct its instances with *create_cache()*, *get_or_create_cache()* or *get_cache()* methods instead. See *this example* on how to do it.

__init__(*client: AioClient, name: str, expiry_policy: pyignite.datatypes.expiry_policy.ExpiryPolicy = None*)

Initialize async cache object. For internal use.

Parameters

- **client** – Async Ignite client,
- **name** – Cache name.

clear(*keys: Union[list, NoneType] = None*)

Clears the cache without notifying listeners or cache writers.

Parameters **keys** – (optional) list of cache keys or (key, key type hint) tuples to clear (default: clear all).

clear_key(*key, key_hint: object = None*)

Clears the cache key without notifying listeners or cache writers.

Parameters

- **key** – key for the cache entry,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

clear_keys(*keys: Iterable[T_co]*)

Clears the cache key without notifying listeners or cache writers.

Parameters **keys** – a list of keys or (key, type hint) tuples

contains_key(*key*, *key_hint=None*) → bool
Returns a value indicating whether given key is present in cache.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns boolean *True* when key is present, *False* otherwise.

contains_keys(*keys: Iterable[T_co]*) → bool
Returns a value indicating whether all given keys are present in cache.

Parameters **keys** – a list of keys or (key, type hint) tuples,

Returns boolean *True* when all keys are present, *False* otherwise.

destroy()
Destroys cache with a given name.

get(*key*, *key_hint: object = None*) → Any
Retrieves a value from cache by key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns value retrieved.

get_all(*keys: list*) → dict
Retrieves multiple key-value pairs from cache.

Parameters **keys** – list of keys or tuples of (key, key_hint),

Returns a dict of key-value pairs.

get_and_put(*key*, *value*, *key_hint=None*, *value_hint=None*) → Any
Puts a value with a given key to cache, and returns the previous value for that key, or null value if there was not such key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

Returns old value or None.

get_and_put_if_absent(*key*, *value*, *key_hint=None*, *value_hint=None*)
Puts a value with a given key to cache only if the key does not already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted,

Returns old value or None.

get_and_remove(*key*, *key_hint=None*) → Any
Removes the cache entry with specified key, returning the value.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

Returns old value or None.

get_and_replace(*key*, *value*, *key_hint=None*, *value_hint=None*) → Any
Puts a value with a given key to cache, returning previous value for that key, if and only if there is a value currently mapped for that key.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

Returns old value or None.

get_size(*peek_modes=None*)
Gets the number of entries in cache.

Parameters **peek_modes** – (optional) limit count to near cache partition (PeekModes.NEAR), primary cache (PeekModes.PRIMARY), or backup cache (PeekModes.BACKUP). Defaults to primary cache partitions (PeekModes.PRIMARY),

Returns integer number of cache entries.

put(*key*, *value*, *key_hint: object = None*, *value_hint: object = None*)
Puts a value with a given key to cache (overwriting existing value if any).

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

put_all(*pairs: dict*)
Puts multiple key-value pairs to cache (overwriting existing associations if any).

Parameters **pairs** – dictionary type parameters, contains key-value pairs to save. Each key or value can be an item of representable Python type or a tuple of (item, hint),

put_if_absent(*key*, *value*, *key_hint=None*, *value_hint=None*)
Puts a value with a given key to cache only if the key does not already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

remove_all()
Removes all cache entries, notifying listeners and cache writers.

remove_if_equals(*key*, *sample*, *key_hint=None*, *sample_hint=None*)
Removes an entry with a given key if provided value is equal to actual value, notifying listeners and cache writers.

Parameters

- **key** – key for the cache entry,
- **sample** – a sample to compare the stored value with,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **sample_hint** – (optional) Ignite data type, for whic the given sample should be converted.

remove_key(*key*, *key_hint=None*)
Clears the cache key without notifying listeners or cache writers.

Parameters

- **key** – key for the cache entry,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,

remove_keys(*keys: list*)
Removes cache entries by given list of keys, notifying listeners and cache writers.

Parameters **keys** – list of keys or tuples of (key, key_hint) to remove.

replace(*key*, *value*, *key_hint: object = None*, *value_hint: object = None*)
Puts a value with a given key to cache only if the key already exist.

Parameters

- **key** – key for the cache entry. Can be of any supported type,
- **value** – value for the key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **value_hint** – (optional) Ignite data type, for which the given value should be converted.

replace_if_equals(*key*, *sample*, *value*, *key_hint=None*, *sample_hint=None*, *value_hint=None*) → Any
Puts a value with a given key to cache only if the key already exists and value equals provided sample.

Parameters

- **key** – key for the cache entry,
- **sample** – a sample to compare the stored value with,
- **value** – new value for the given key,
- **key_hint** – (optional) Ignite data type, for which the given key should be converted,
- **sample_hint** – (optional) Ignite data type, for whic the given sample should be converted
- **value_hint** – (optional) Ignite data type, for which the given value should be converted,

Returns boolean *True* when key is present, *False* otherwise.

scan(*page_size: int = 1*, *partitions: int = -1*, *local: bool = False*) → pyignite.cursors.AioScanCursor
Returns all key-value pairs from the cache, similar to *get_all*, but with internal pagination, which is slower, but safer.

Parameters

- **page_size** – (optional) page size. Default size is 1 (slowest and safest),
- **partitions** – (optional) number of partitions to query (negative to query entire cache),
- **local** – (optional) pass True if this query should be executed on local node only. Defaults to False,

Returns async scan query cursor

settings() → Union[dict, NoneType]
Lazy Cache settings. See the *example* of reading this property.

All cache properties are documented here: *Cache Properties*.

Returns dict of cache properties and their values.

pyignite.aio_cache.**create_cache**(*client: AioClient, settings: Union[str, dict]*) → AioCache

pyignite.aio_cache.**get_cache**(*client: AioClient, settings: Union[str, dict]*) → AioCache

pyignite.aio_cache.**get_or_create_cache**(*client: AioClient, settings: Union[str, dict]*) → AioCache

2.5 Data Types

Apache Ignite uses a sophisticated system of serializable data types to store and retrieve user data, as well as to manage the configuration of its caches through the Ignite binary protocol.

The complexity of data types varies from simple integer or character types to arrays, maps, collections and structures.

Each data type is defined by its code. *Type code* is byte-sized. Thus, every data object can be represented as a payload of fixed or variable size, logically divided into one or more fields, prepended by the *type_code* field.

Most of Ignite data types can be represented by some of the standard Python data type or class. Some of them, however, are conceptually alien, overly complex, or ambiguous to Python dynamic type system.

The following table summarizes the notion of Apache Ignite data types, as well as their representation and handling in Python. For the nice description, as well as gory implementation details, you may follow the link to the parser/constructor class definition. Note that parser/constructor classes are not instantiatable. The *class* here is used mostly as a sort of tupperware for organizing methods together.

Note: you are not obliged to actually use those parser/constructor classes. Pythonic types will suffice to interact with Apache Ignite binary API. However, in some rare cases of type ambiguity, as well as for the needs of interoperability, you may have to sneak one or the other class, along with your data, in to some API function as a *type conversion hint*.

type_code	Apache Ignite docs reference	Python type or class	Parser/constructor class
Primitive data types			
0x01	Byte	int	`ByteObject`
0x02	Short	int	`ShortObject`
0x03	Int	int	`IntObject`
0x04	Long	int	`LongObject`
0x05	Float	float	`FloatObject`
0x06	Double	float	`DoubleObject`

Continued on next page

Table 1 – continued from previous page

type_code	Apache Ignite docs reference	Python type or class	Parser/constructor class
0x07	Char	str	`CharObject`
0x08	Bool	bool	`BoolObject`
0x65	Null	NoneType	`Null`
Standard objects			
0x09	String	Str	`String`
0x0a	UUID	uuid.UUID	`UUIDObject`
0x21	Timestamp	tuple	`TimestampObject`
0x0b	Date	datetime.datetime	`DateObject`
0x24	Time	datetime.timedelta	`TimeObject`
0x1e	Decimal	decimal.Decimal	`DecimalObject`
0x1c	Enum	tuple	`EnumObject`
0x67	Binary enum	tuple	`BinaryEnumObject`
Arrays of primitives			
0x0c	Byte array	iterable/bytearray	`ByteArrayObject`
0x0d	Short array	iterable/list	`ShortArrayObject`
0x0e	Int array	iterable/list	`IntArrayObject`
0x0f	Long array	iterable/list	`LongArrayObject`
0x10	Float array	iterable/list	`FloatArrayObject`
0x11	Double array	iterable/list	`DoubleArrayObject`
0x12	Char array	iterable/list	`CharArrayObject`
0x13	Bool array	iterable/list	`BoolArrayObject`
Arrays of standard objects			
0x14	String array	iterable/list	`StringArrayObject`
0x15	UUID array	iterable/list	`UUIDArrayObject`
0x22	Timestamp array	iterable/list	`TimestampArrayObject`
0x16	Date array	iterable/list	`DateArrayObject`
0x23	Time array	iterable/list	`TimeArrayObject`
0x1f	Decimal array	iterable/list	`DecimalArrayObject`
Object collections, special types, and complex object			
0x17	Object array	tuple[int, iterable/list]	`ObjectArrayObject`
0x18	Collection	tuple[int, iterable/list]	`CollectionObject`
0x19	Map	tuple[int, dict/OrderedDict]	`MapObject`
0x1d	Enum array	iterable/list	`EnumArrayObject`
0x67	Complex object	object	`BinaryObject`
0x1b	Wrapped data	tuple[int, bytes]	`WrappedDataObject`

2.6 Cache Properties

The `prop_codes` module contains a list of ordinal values, that represent various cache settings.

Please refer to the Configuring Caches documentation on cache synchronization, rebalance, affinity and other cache configuration-related matters.

Property name	Ordinal value	Property type	Description
Read/write cache properties, used to configure cache via *create_cache()* or *get_or_create_* of *Client* (*create_cache()* or *get_or_create_cache()* of *AioClient*).			
PROP_NAME	0	str	Cache name.
PROP_CACHE_MODE	1	int	Cache mode:

Table 2 – continued from

Property name	Ordinal value	Property type	Description
PROP_CACHE_ATOMICITY_MODE	2	int	Cache atomici
PROP_BACKUPS_NUMBER	3	int	Number of ba
PROP_WRITE_SYNCHRONIZATION_MODE	4	int	Write synchro
PROP_COPY_ON_READ	5	bool	Copy-on-read
PROP_READ_FROM_BACKUP	6	bool	Read from bac
PROP_DATA_REGION_NAME	100	str	Data region na
PROP_IS_ONHEAP_CACHE_ENABLED	101	bool	Is OnHeap cac
PROP_QUERY_ENTITIES	200	list	A list of query
PROP_QUERY_PARALLELISM	201	int	Query parallel
PROP_QUERY_DETAIL_METRIC_SIZE	202	int	Query detail m
PROP_SQL_SCHEMA	203	str	SQL schema
PROP_SQL_INDEX_INLINE_MAX_SIZE	204	int	SQL index inl
PROP_SQL_ESCAPE_ALL	205	bool	Turns on SQL
PROP_MAX_QUERY_ITERATORS	206	int	Maximum nur
PROP_REBALANCE_MODE	300	int	Rebalance mo
PROP_REBALANCE_DELAY	301	int	Rebalance del
PROP_REBALANCE_TIMEOUT	302	int	Rebalance tim
PROP_REBALANCE_BATCH_SIZE	303	int	Rebalance bat
PROP_REBALANCE_BATCHES_PREFETCH_COUNT	304	int	Rebalance bat
PROP_REBALANCE_ORDER	305	int	Rebalance ord
PROP_REBALANCE_THROTTLE	306	int	Rebalance thr
PROP_GROUP_NAME	400	str	Group name
PROP_CACHE_KEY_CONFIGURATION	401	list	Cache key con
PROP_DEFAULT_LOCK_TIMEOUT	402	int	Default lock t
PROP_MAX_CONCURRENT_ASYNC_OPERATIONS	403	int	Maximum nur
PROP_PARTITION_LOSS_POLICY	404	int	Partition loss p
PROP_EAGER_TTL	405	bool	Eager TTL
PROP_STATISTICS_ENABLED	406	bool	Statistics enab
PROP_EXPIRY_POLICY	407	`ExpiryPolicy`	Set expiry pol

2.6.1 Query entity

A dict with all ot the following keys:

- *table_name*: SQL table name,
- *key_field_name*: name of the key field,
- *key_type_name*: name of the key type (Java type or complex object),
- *value_field_name*: name of the value field,
- *value_type_name*: name of the value type,
- *field_name_aliases*: a list of 0 or more dicts of aliases (see *Field name alias*),
- *query_fields*: a list of 0 or more query field names (see *Query field*),
- *query_indexes*: a list of 0 or more query indexes (see *Query index*).

Field name alias

- *field_name*: field name,

- *alias*: alias (str).

Query field

- *name*: field name,
- *type_name*: name of Java type or complex object,
- *is_key_field*: (optional) boolean value, *False* by default,
- *is_notnull_constraint_field*: boolean value,
- *default_value*: (optional) anything that can be converted to *type_name* type. *None* (`Null`) by default,
- *precision* (optional) decimal precision: total number of digits in decimal value. Defaults to -1 (use cluster default). Ignored for non-decimal SQL types (other than *java.math.BigDecimal*),
- *scale* (optional) decimal precision: number of digits after the decimal point. Defaults to -1 (use cluster default). Ignored for non-decimal SQL types.

Query index

- *index_name*: index name,
- *index_type*: index type code as an integer value in unsigned byte range,
- *inline_size*: integer value,
- *fields*: a list of 0 or more indexed fields (see *Fields*).

Fields

- *name*: field name,
- *is_descending*: (optional) boolean value, *False* by default.

2.6.2 Cache key

A dict of the following format:

- *type_name*: name of the complex object,
- *affinity_key_field_name*: name of the affinity key field.

2.6.3 Expiry policy

Set expiry policy to cache (see `ExpiryPolicy`). If set to *None*, expiry policy will not be set.

2.7 pyignite.exceptions module

exception pyignite.exceptions.**AuthenticationError**(*message: str*)

Bases: `Exception`

This exception is raised on authentication failure.

__init__(*message: str*)
Initialize self. See help(type(self)) for accurate signature.

exception pyignite.exceptions.**BinaryTypeError**
Bases: *pyignite.exceptions.CacheError*

A remote error in operation with Complex Object registry.

exception pyignite.exceptions.**CacheCreationError**
Bases: *pyignite.exceptions.CacheError*

This exception is raised, when any complex operation failed on cache creation phase.

exception pyignite.exceptions.**CacheError**
Bases: Exception

This exception is raised, whenever any remote Thin client cache operation returns an error.

exception pyignite.exceptions.**ClusterError**
Bases: Exception

This exception is raised, whenever any remote Thin client cluster operation returns an error.

exception pyignite.exceptions.**HandshakeError**(*expected_version: Tuple[int, int, int], message: str*)
Bases: OSError

This exception is raised on Ignite binary protocol handshake failure, as defined in https://ignite.apache.org/docs/latest/binary-client-protocol/binary-client-protocol#connection-handshake

__init__(*expected_version: Tuple[int, int, int], message: str*)
Initialize self. See help(type(self)) for accurate signature.

exception pyignite.exceptions.**NotSupportedByClusterError**
Bases: Exception

This exception is raised, whenever cluster does not supported specific operation probably because it is outdated.

exception pyignite.exceptions.**NotSupportedError**
Bases: Exception

This exception is raised, whenever client does not support specific operation.

exception pyignite.exceptions.**ParameterError**
Bases: Exception

This exception represents the parameter validation error in any *pyignite* method.

exception pyignite.exceptions.**ParseError**
Bases: Exception

This exception is raised, when *pyignite* is unable to build a query to, or parse a response from, Ignite node.

exception pyignite.exceptions.**ReconnectError**
Bases: Exception

This exception is raised by *Client.reconnect* method, when no more nodes are left to connect to. It is not meant to be an error, but rather a flow control tool, similar to *StopIteration*.

exception pyignite.exceptions.**SQLError**
Bases: *pyignite.exceptions.CacheError*

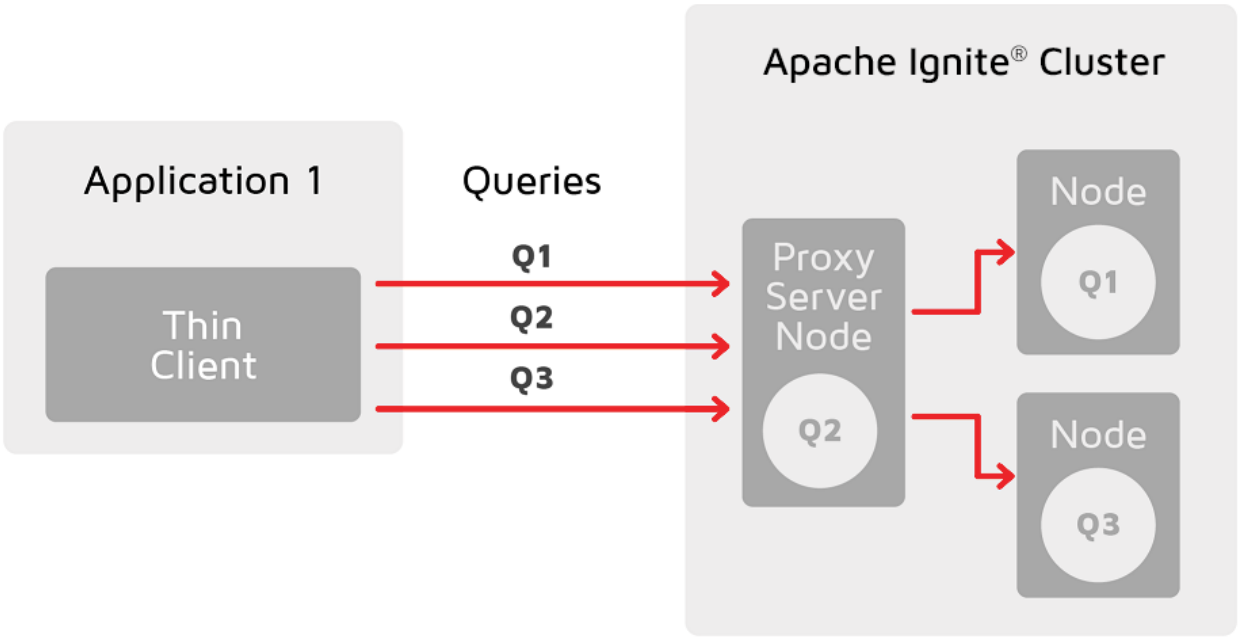
An error in SQL query.

CHAPTER 3

Partition Awareness

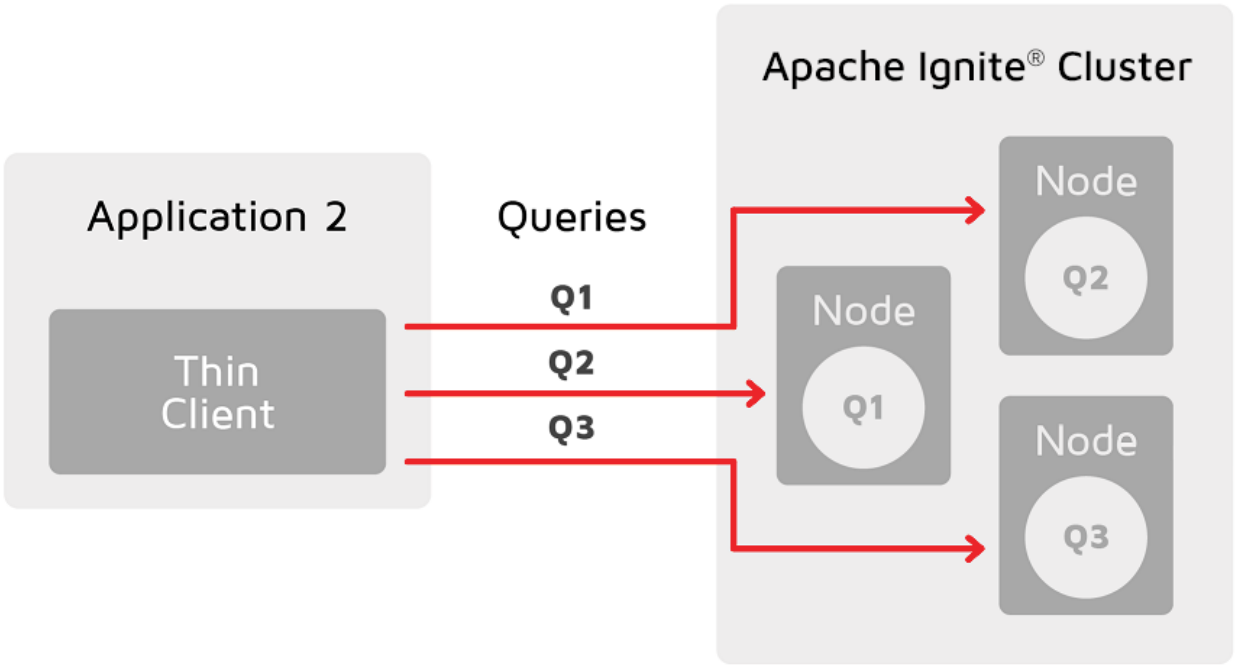
Partition awareness allows the thin client to send query requests directly to the node that owns the queried data.

Without partition awareness, an application that is connected to the cluster via a thin client executes all queries and operations via a single server node that acts as a proxy for the incoming requests. These operations are then re-routed to the node that stores the data that is being requested. This results in a bottleneck that could prevent the application from scaling linearly.



Notice how queries must pass through the proxy server node, where they are routed to the correct node.

With partition awareness in place, the thin client can directly route queries and operations to the primary nodes that own the data required for the queries. This eliminates the bottleneck, allowing the application to scale more easily.



Partition awareness can be enabled or disabled by setting *partition_aware* parameter in *pyignite.client.Client.__init__()* or *pyignite.aio_client.AioClient.__init__()* to *True* (by default) or *False*.

Also, it is recommended to pass list of address and port pairs of all server nodes to *pyignite.client.Client.connect()* or to *pyignite.aio_client.AioClient.connect()*.

For example:

```
from pyignite import Client

client = Client(
    partition_awareness=True
)
nodes = [('10.128.0.1', 10800), ('10.128.0.2', 10800),...]
with client.connect(nodes):
    ....
```

```
from pyignite import AioClient

client = AioClient(
    partition_awareness=True
)
nodes = [('10.128.0.1', 10800), ('10.128.0.2', 10800),...]
async with client.connect(nodes):
    ....
```

CHAPTER 4

Examples of usage

File: get_and_put.py.

4.1 Key-value

4.1.1 Open connection

```
from pyignite import Client

client = Client()
with client.connect('127.0.0.1', 10800):
```

4.1.2 Create cache

```
my_cache = client.create_cache('my cache')
```

4.1.3 Put value in cache

```
my_cache.put('my key', 42)
```

4.1.4 Get value from cache

```
result = my_cache.get('my key')
print(result)  # 42

result = my_cache.get('non-existent key')
print(result)  # None
```

4.1.5 Get multiple values from cache

```
result = my_cache.get_all([
    'my key',
    'non-existent key',
    'other-key',
])
print(result)  # {'my key': 42}
```

4.1.6 Type hints usage

File: type_hints.py

```
my_cache.put('my key', 42)
# value '42' takes 9 bytes of memory as a LongObject

my_cache.put('my key', 42, value_hint=ShortObject)
# value '42' takes only 3 bytes as a ShortObject

my_cache.put('a', 1)
# 'a' is a key of type String

my_cache.put('a', 2, key_hint=CharObject)
# another key 'a' of type CharObject was created

value = my_cache.get('a')
print(value)
# 1

value = my_cache.get('a', key_hint=CharObject)
print(value)
# 2

# now let us delete both keys at once
my_cache.remove_keys([
    'a',                # a default type key
    ('a', CharObject),  # a key of type CharObject
])
```

As a rule of thumb:

- when a *pyignite* method or function deals with a single value or key, it has an additional parameter, like *value_hint* or *key_hint*, which accepts a parser/constructor class,
- nearly any structure element (inside dict or list) can be replaced with a two-tuple of (said element, type hint).

Refer the *Data Types* section for the full list of parser/constructor classes you can use as type hints.

4.1.7 ExpiryPolicy

File: expiry_policy.py.

You can enable expiry policy (TTL) by two approaches.

Firstly, expiry policy can be set for entire cache by setting `PROP_EXPIRY_POLICY` in cache settings dictionary on creation.

```
ttl_cache = client.create_cache({
    PROP_NAME: 'test',
    PROP_EXPIRY_POLICY: ExpiryPolicy(create=timedelta(seconds=1.0))
})
```

```
ttl_cache.put(1, 1)
time.sleep(0.5)
print(f"key = {1}, value = {ttl_cache.get(1)}")
# key = 1, value = 1
time.sleep(1.2)
print(f"key = {1}, value = {ttl_cache.get(1)}")
# key = 1, value = None
```

Secondly, expiry policy can be set for all cache operations, which are done under decorator. To create it use *with_expire_policy()*

```
ttl_cache = simple_cache.with_expire_policy(access=timedelta(seconds=1.0))
ttl_cache.put(1, 1)
time.sleep(0.5)
print(f"key = {1}, value = {ttl_cache.get(1)}")
# key = 1, value = 1
time.sleep(1.7)
print(f"key = {1}, value = {ttl_cache.get(1)}")
# key = 1, value = None
```

4.1.8 Scan

File: scans.py.

Cache's *scan()* method queries allows you to get the whole contents of the cache, element by element.

Let us put some data in cache.

```
my_cache = client.create_cache('my cache')
my_cache.put_all({'key_{}'.format(v): v for v in range(20)})
# {
#     'key_0': 0,
#     'key_1': 1,
#     'key_2': 2,
#     ... 20 elements in total...
#     'key_18': 18,
#     'key_19': 19
# }
```

scan() returns a cursor, that yields two-tuples of key and value. You can iterate through the generated pairs in a safe manner:

```
with my_cache.scan() as cursor:
    for k, v in cursor:
        print(k, v)
# 'key_17' 17
# 'key_10' 10
# 'key_6' 6,
# ... 20 elements in total...
# 'key_16' 16
# 'key_12' 12
```

Or, alternatively, you can convert the cursor to dictionary in one go:

```
with my_cache.scan() as cursor:
    print(dict(cursor))
# {
#     'key_17': 17,
#     'key_10': 10,
#     'key_6': 6,
#     ... 20 elements in total...
#     'key_16': 16,
#     'key_12': 12
# }
```

But be cautious: if the cache contains a large set of data, the dictionary may consume too much memory!

4.2 Object collections

File: get_and_put_complex.py.

Ignite collection types are represented in *pyignite* as two-tuples. First comes collection type ID or deserialization hint, which is specific for each of the collection type. Second comes the data value.

```
from pyignite.datatypes import CollectionObject, MapObject, ObjectArrayObject
```

4.2.1 Map

For Python prior to 3.6, it might be important to distinguish between ordered (*collections.OrderedDict*) and unordered (*dict*) dictionary types, so you could use `LINKED_HASH_MAP` for the former and `HASH_MAP` for the latter.

Since CPython 3.6 all dictionaries became de facto ordered. You can always use *LINKED_HASH_MAP* as a safe default.

```
value = OrderedDict([(1, 'test'), ('key', 2.0)])

# saving ordered dictionary
type_id = MapObject.LINKED_HASH_MAP
my_cache.put('my dict', (type_id, value))
result = my_cache.get('my dict')
print(result)  # (2, OrderedDict([(1, 'test'), ('key', 2.0)]))

# saving unordered dictionary
type_id = MapObject.HASH_MAP
my_cache.put('my dict', (type_id, value))
result = my_cache.get('my dict')
print(result)  # (1, {'key': 2.0, 1: 'test'})
```

4.2.2 Collection

See `CollectionObject` and Ignite documentation on Collection type for the description of various Java collection types. Note that not all of them have a direct Python representative. For example, Python do not have ordered sets (it is indeed recommended to use *OrderedDict*'s keys and disregard its values).

As for the *pyignite*, the rules are simple: pass any iterable as a data, and you always get *list* back.

```
type_id = CollectionObject.LINKED_LIST
value = [1, '2', 3.0]

my_cache.put('my list', (type_id, value))

result = my_cache.get('my list')
print(result)  # (2, [1, '2', 3.0])

type_id = CollectionObject.HASH_SET
value = [4, 4, 'test', 5.6]

my_cache.put('my set', (type_id, value))

result = my_cache.get('my set')
print(result)  # (3, [5.6, 4, 'test'])
```

4.2.3 Object array

`ObjectArrayObject` has a very limited functionality in *pyignite*, since no type checks can be enforced on its contents. But it still can be used for interoperability with Java.

```
type_id = ObjectArrayObject.OBJECT
value = [7, '8', 9.0]

my_cache.put(
    'my array of objects',
    (type_id, value),
    value_hint=ObjectArrayObject  # this hint is mandatory!
)
result = my_cache.get('my array of objects')
print(result)  # (-1, [7, '8', 9.0])
```

4.3 Transactions

File: transactions.py.

Client transactions are supported for caches with `TRANSACTIONAL` mode.

Let's create transactional cache:

```
cache = client.get_or_create_cache({
    PROP_NAME: 'tx_cache',
    PROP_CACHE_ATOMICITY_MODE: CacheAtomicityMode.TRANSACTIONAL
})
```

Let's start a transaction and commit it:

```
key = 1
with client.tx_start(
        isolation=TransactionIsolation.REPEATABLE_READ,
        concurrency=TransactionConcurrency.PESSIMISTIC
) as tx:
    cache.put(key, 'success')
    tx.commit()
```

Let's check that the transaction was committed successfully:

```
# key=1 value=success
print(f"key={key} value={cache.get(key)}")
```

Let's check that raising exception inside *with* block leads to transaction's rollback

```
try:
    with client.tx_start(
            isolation=TransactionIsolation.REPEATABLE_READ,
            concurrency=TransactionConcurrency.PESSIMISTIC
    ):
        cache.put(key, 'fail')
        raise RuntimeError('test')
except RuntimeError:
    pass

# key=1 value=success
print(f"key={key} value={cache.get(key)}")
```

Let's check that timed out transaction is successfully rolled back

```
try:
    with client.tx_start(timeout=1000, label='long-tx') as tx:
        cache.put(key, 'fail')
        time.sleep(2.0)
        tx.commit()
except CacheError as e:
    # Cache transaction timed out: GridNearTxLocal[...timeout=1000, ... label=long-tx]
    print(e)

# key=1 value=success
print(f"key={key} value={cache.get(key)}")
```

See more info about transaction's parameters in a documentation of *tx_start()*

4.4 SQL

File: sql.py.

These examples are similar to the ones given in the Apache Ignite SQL Documentation: Getting Started.

4.4.1 Setup

First let us establish a connection.

```
client = Client()
with client.connect('127.0.0.1', 10800):
```

Then create tables. Begin with *Country* table, than proceed with related tables *City* and *CountryLanguage*.

```
COUNTRY_CREATE_TABLE_QUERY = '''CREATE TABLE Country (
    Code CHAR(3) PRIMARY KEY,
    Name CHAR(52),
    Continent CHAR(50),
```

(continues on next page)

(continued from previous page)

```
    Region CHAR(26),
    SurfaceArea DECIMAL(10,2),
    IndepYear SMALLINT(6),
    Population INT(11),
    LifeExpectancy DECIMAL(3,1),
    GNP DECIMAL(10,2),
    GNPOld DECIMAL(10,2),
    LocalName CHAR(45),
    GovernmentForm CHAR(45),
    HeadOfState CHAR(60),
    Capital INT(11),
    Code2 CHAR(2)
)'''

CITY_CREATE_TABLE_QUERY = '''CREATE TABLE City (
    ID INT(11),
    Name CHAR(35),
    CountryCode CHAR(3),
    District CHAR(20),
    Population INT(11),
    PRIMARY KEY (ID, CountryCode)
) WITH "affinityKey=CountryCode"'''

LANGUAGE_CREATE_TABLE_QUERY = '''CREATE TABLE CountryLanguage (
    CountryCode CHAR(3),
    Language CHAR(30),
    IsOfficial BOOLEAN,
    Percentage DECIMAL(4,1),
    PRIMARY KEY (CountryCode, Language)
) WITH "affinityKey=CountryCode"'''
```

```
for query in [
    COUNTRY_CREATE_TABLE_QUERY,
    CITY_CREATE_TABLE_QUERY,
    LANGUAGE_CREATE_TABLE_QUERY,
]:
    client.sql(query)
```

Create indexes.

```
CITY_CREATE_INDEX = '''
CREATE INDEX idx_country_code ON city (CountryCode)'''

LANGUAGE_CREATE_INDEX = '''
CREATE INDEX idx_lang_country_code ON CountryLanguage (CountryCode)'''
```

```
for query in [CITY_CREATE_INDEX, LANGUAGE_CREATE_INDEX]:
    client.sql(query)
```

Fill tables with data.

```
COUNTRY_INSERT_QUERY = '''INSERT INTO Country(
    Code, Name, Continent, Region,
    SurfaceArea, IndepYear, Population,
```

(continues on next page)

(continued from previous page)

```
    LifeExpectancy, GNP, GNPOld,
    LocalName, GovernmentForm, HeadOfState,
    Capital, Code2
) VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)'''

CITY_INSERT_QUERY = '''INSERT INTO City(
    ID, Name, CountryCode, District, Population
) VALUES (?, ?, ?, ?, ?)'''

LANGUAGE_INSERT_QUERY = '''INSERT INTO CountryLanguage(
    CountryCode, Language, IsOfficial, Percentage
) VALUES (?, ?, ?, ?)'''
```

```
for row in COUNTRY_DATA:
    client.sql(COUNTRY_INSERT_QUERY, query_args=row)

for row in CITY_DATA:
    client.sql(CITY_INSERT_QUERY, query_args=row)

for row in LANGUAGE_DATA:
    client.sql(LANGUAGE_INSERT_QUERY, query_args=row)
```

Data samples are taken from PyIgnite GitHub repository.

That concludes the preparation of data. Now let us answer some questions.

4.4.2 What are the 10 largest cities in our data sample (population-wise)?

```
MOST_POPULATED_QUERY = '''
SELECT name, population FROM City ORDER BY population DESC LIMIT 10'''

with client.sql(MOST_POPULATED_QUERY) as cursor:
    print('Most 10 populated cities:')
    for row in cursor:
        print(row)
# Most 10 populated cities:
# ['Mumbai (Bombay)', 10500000]
# ['Shanghai', 9696300]
# ['New York', 8008278]
# ['Peking', 7472000]
# ['Delhi', 7206704]
# ['Chongqing', 6351600]
# ['Tianjin', 5286800]
# ['Calcutta [Kolkata]', 4399819]
# ['Wuhan', 4344600]
# ['Harbin', 4289800]
```

The `sql()` method returns a generator, that yields the resulting rows.

4.4.3 What are the 10 most populated cities throughout the 3 chosen countries?

If you set the *include_field_names* argument to *True*, the `sql()` method will generate a list of column names as a first yield. You can access field names with Python built-in *next* function.

```
MOST_POPULATED_IN_3_COUNTRIES_QUERY = '''
SELECT country.name as country_name, city.name as city_name, MAX(city.population) AS␣
↪max_pop FROM country
    JOIN city ON city.countrycode = country.code
    WHERE country.code IN ('USA','IND','CHN')
    GROUP BY country.name, city.name ORDER BY max_pop DESC LIMIT 10
'''

with client.sql(MOST_POPULATED_IN_3_COUNTRIES_QUERY, include_field_names=True) as␣
↪cursor:
    print('Most 10 populated cities in USA, India and China:')
    print(next(cursor))
    print('----------------------------------------')
    for row in cursor:
        print(row)
# Most 10 populated cities in USA, India and China:
# ['COUNTRY_NAME', 'CITY_NAME', 'MAX_POP']
# ----------------------------------------
# ['India', 'Mumbai (Bombay)', 10500000]
# ['China', 'Shanghai', 9696300]
# ['United States', 'New York', 8008278]
# ['China', 'Peking', 7472000]
# ['India', 'Delhi', 7206704]
# ['China', 'Chongqing', 6351600]
# ['China', 'Tianjin', 5286800]
# ['India', 'Calcutta [Kolkata]', 4399819]
# ['China', 'Wuhan', 4344600]
# ['China', 'Harbin', 4289800]
```

4.4.4 Display all the information about a given city

```
# show city info
CITY_INFO_QUERY = '''SELECT * FROM City WHERE id = ?'''

with client.sql(CITY_INFO_QUERY, query_args=[3802], include_field_names=True) as␣
↪cursor:
    field_names = next(cursor)
    field_data = list(*cursor)

    print('City info:')
    for field_name, field_value in zip(field_names * len(field_data), field_data):
        print('{}: {}'.format(field_name, field_value))
# City info:
# ID: 3802
# NAME: Detroit
# COUNTRYCODE: USA
# DISTRICT: Michigan
# POPULATION: 951270
```

Finally, delete the tables used in this example with the following queries:

```
DROP_TABLE_QUERY = '''DROP TABLE {} IF EXISTS'''
```

```
# clean up
for table_name in [
    CITY_TABLE_NAME,
    LANGUAGE_TABLE_NAME,
    COUNTRY_TABLE_NAME,
]:
    result = client.sql(DROP_TABLE_QUERY.format(table_name))
```

4.5 Complex objects

File: binary_basics.py.

Complex object (that is often called 'Binary object') is an Ignite data type, that is designed to represent a Java class. It have the following features:

- have a unique ID (type id), which is derives from a class name (type name),
- have one or more associated schemas, that describes its inner structure (the order, names and types of its fields). Each schema have its own ID,
- have an optional version number, that is aimed towards the end users to help them distinguish between objects of the same type, serialized with different schemas.

Unfortunately, these distinctive features of the Complex object have few to no meaning outside of Java language. Python class can not be defined by its name (it is not unique), ID (object ID in Python is volatile; in CPython it is just a pointer in the interpreter's memory heap), or complex of its fields (they do not have an associated data types, moreover, they can be added or deleted in run-time). For the *pyignite* user it means that for all purposes of storing native Python data it is better to use Ignite `CollectionObject` or `MapObject` data types.

However, for interoperability purposes, *pyignite* has a mechanism of creating special Python classes to read or write Complex objects. These classes have an interface, that simulates all the features of the Complex object: type name, type ID, schema, schema ID, and version number.

Assuming that one concrete class for representing one Complex object can severely limit the user's data manipulation capabilities, all the functionality said above is implemented through the metaclass: `GenericObjectMeta`. This metaclass is used automatically when reading Complex objects.

```
person_cache = client.get_or_create_cache('person')

person_cache.put(
print(person.__class__.__name__)
# Person

print(person.__class__ is Person)
# Person(first_name='Ivan', last_name='Ivanov', age=33, version=1)
```

Here you can see how `GenericObjectMeta` uses attrs package internally for creating nice *__init__()* and *__repr__()* methods.

In this case the autogenerated dataclass's name *Person* is exactly matches the type name of the Complex object it represents (the content of the `type_name` property). But when Complex object's class name contains characters, that can not be used in a Python identifier, for example:

- ., when fully qualified Java class names are used,
- *$*, a common case for Scala classes,
- +, internal class name separator in C#,

then *pyignite* can not maintain this match. In such cases *pyignite* tries to sanitize a type name to derive a "good" dataclass name from it.

If your code needs consistent naming between the server and the client, make sure that your Ignite cluster is configured to use simple class names.

Anyway, you can reuse the autogenerated dataclass for subsequent writes:

```
Person = person.__class__

person_cache.put(
    1, Person(first_name='Ivan', last_name='Ivanov', age=33)
)
```

`GenericObjectMeta` can also be used directly for creating custom classes:

```
from pyignite import Client, GenericObjectMeta
from pyignite.datatypes import String, IntObject


class Person(metaclass=GenericObjectMeta, schema=OrderedDict([
    ('first_name', String),
    ('last_name', String),
    ('age', IntObject),
])):
    pass
```

Note how the *Person* class is defined. *schema* is a `GenericObjectMeta` metaclass parameter. Another important `GenericObjectMeta` parameter is a *type_name*, but it is optional and defaults to the class name ('Person' in our example).

Note also, that *Person* do not have to define its own attributes, methods and properties (*pass*), although it is completely possible.

Now, when your custom *Person* class is created, you are ready to send data to Ignite server using its objects. The client will implicitly register your class as soon as the first Complex object is sent. If you intend to use your custom class for reading existing Complex objects' values before all, you must register said class explicitly with your client:

```
client.register_binary_type(Person)
```

Now, when we dealt with the basics of *pyignite* implementation of Complex Objects, let us move on to more elaborate examples.

4.5.1 Read

File: read_binary.py.

Ignite SQL uses Complex objects internally to represent keys and rows in SQL tables. Normally SQL data is accessed via queries (see *SQL*), so we will consider the following example solely for the demonstration of how Binary objects (not Ignite SQL) work.

In the *previous examples* we have created some SQL tables. Let us do it again and examine the Ignite storage afterwards.

```
result = client.get_cache_names()
print(result)
# [
#     'SQL_PUBLIC_CITY',
#     'SQL_PUBLIC_COUNTRY',
#     'PUBLIC',
#     'SQL_PUBLIC_COUNTRYLANGUAGE'
# ]
```

We can see that Ignite created a cache for each of our tables. The caches are conveniently named using '*SQL_<schema name>_<table name>*' pattern.

Now let us examine a configuration of a cache that contains SQL data using a `settings` property.

```
city_cache = client.get_or_create_cache('SQL_PUBLIC_CITY')
print(city_cache.settings[PROP_NAME])
# 'SQL_PUBLIC_CITY'

print(city_cache.settings[PROP_QUERY_ENTITIES])
# {
#     'key_type_name': (
#         'SQL_PUBLIC_CITY_9ac8e17a_2f99_45b7_958e_06da32882e9d_KEY'
#     ),
#     'value_type_name': (
#         'SQL_PUBLIC_CITY_9ac8e17a_2f99_45b7_958e_06da32882e9d'
#     ),
#     'table_name': 'CITY',
#     'query_fields': [
#         ...
#     ],
#     'field_name_aliases': [
#         ...
#     ],
#     'query_indexes': []
# }
```

The values of *value_type_name* and *key_type_name* are names of the binary types. The *City* table's key fields are stored using *key_type_name* type, and the other fields *value_type_name* type.

Now when we have the cache, in which the SQL data resides, and the names of the key and value data types, we can read the data without using SQL functions and verify the correctness of the result.

```
with city_cache.scan() as cursor:
    print(next(cursor))
# (
#     SQL_PUBLIC_CITY_6fe650e1_700f_4e74_867d_58f52f433c43_KEY(
#         ID=1890,
#         COUNTRYCODE='CHN',
#         version=1
#     ),
#     SQL_PUBLIC_CITY_6fe650e1_700f_4e74_867d_58f52f433c43(
#         NAME='Shanghai',
#         DISTRICT='Shanghai',
#         POPULATION=9696300,
#         version=1
#     )
# )
```

What we see is a tuple of key and value, extracted from the cache. Both key and value are represent Complex objects.

The dataclass names are the same as the *value_type_name* and *key_type_name* cache settings. The objects' fields correspond to the SQL query.

4.5.2 Create

File: create_binary.py.

Now, that we aware of the internal structure of the Ignite SQL storage, we can create a table and put data in it using only key-value functions.

For example, let us create a table to register High School students: a rough equivalent of the following SQL DDL statement:

```
CREATE TABLE Student (
    sid CHAR(9),
    name VARCHAR(20),
    login CHAR(8),
    age INTEGER(11),
    gpa REAL
)
```

These are the necessary steps to perform the task.

1. Create table cache.

```
student_cache = client.create_cache({
    PROP_NAME: 'SQL_PUBLIC_STUDENT',
    PROP_SQL_SCHEMA: 'PUBLIC',
    PROP_QUERY_ENTITIES: [
        {
            'table_name': 'Student'.upper(),
            'key_field_name': 'SID',
            'key_type_name': 'java.lang.Integer',
            'field_name_aliases': [],
            'query_fields': [
                {
                    'name': 'SID',
                    'type_name': 'java.lang.Integer',
                    'is_key_field': True,
                    'is_notnull_constraint_field': True,
                },
                {
                    'name': 'NAME',
                    'type_name': 'java.lang.String',
                },
                {
                    'name': 'LOGIN',
                    'type_name': 'java.lang.String',
                },
                {
                    'name': 'AGE',
                    'type_name': 'java.lang.Integer',
                },
                {
                    'name': 'GPA',
                    'type_name': 'java.math.Double',
                },
```

(continues on next page)

(continued from previous page)

```
            ],
            'query_indexes': [],
            'value_type_name': 'SQL_PUBLIC_STUDENT_TYPE',
            'value_field_name': None,
        },
    ],
})
```

2. Define Complex object data class.

```
class Student(
    metaclass=GenericObjectMeta,
    type_name='SQL_PUBLIC_STUDENT_TYPE',
    schema=OrderedDict([
        ('NAME', String),
        ('LOGIN', String),
        ('AGE', IntObject),
        ('GPA', DoubleObject),
    ])
):
    pass
```

3. Insert row.

```
student_cache.put(
    1,
    Student(LOGIN='jdoe', NAME='John Doe', AGE=17, GPA=4.25),
    key_hint=IntObject
)
```

Now let us make sure that our cache really can be used with SQL functions.

```
with client.sql(r'SELECT * FROM Student', include_field_names=True) as cursor:
    print(next(cursor))
    # ['SID', 'NAME', 'LOGIN', 'AGE', 'GPA']

    print(*cursor)
    # [1, 'John Doe', 'jdoe', 17, 4.25]
```

Note, however, that the cache we create can not be dropped with DDL command. It should be deleted as any other key-value cache.

```
# DROP_QUERY = 'DROP TABLE Student'
# client.sql(DROP_QUERY)
#
# pyignite.exceptions.SQLError: class org.apache.ignite.IgniteCheckedException:
# Only cache created with CREATE TABLE may be removed with DROP TABLE
# [cacheName=SQL_PUBLIC_STUDENT]

student_cache.destroy()
```

4.5.3 Migrate

File: migrate_binary.py.

Suppose we have an accounting app that stores its data in key-value format. Our task would be to introduce the following changes to the original expense voucher's format and data:

- rename *date* to *expense_date*,
- add *report_date*,
- set *report_date* to the current date if *reported* is True, None if False,
- delete *reported*.

First get the vouchers' cache.

```
accounting = client.get_or_create_cache('accounting')
```

If you do not store the schema of the Complex object in code, you can obtain it as a dataclass property with *query_binary_type()* method.

```
data_classes = client.query_binary_type('ExpenseVoucher')
print(data_classes)
# {
#     -231598180: <class '__main__.ExpenseVoucher'>
# }
```

Let us modify the schema and create a new Complex object class with an updated schema.

```
s_id, data_class = data_classes.popitem()
schema = data_class.schema

schema['expense_date'] = schema['date']
del schema['date']
schema['report_date'] = DateObject
del schema['reported']
schema['sum'] = DecimalObject


# define new data class
class ExpenseVoucherV2(
    metaclass=GenericObjectMeta,
    type_name='ExpenseVoucher',
    schema=schema,
):
    pass
```

Now migrate the data from the old schema to the new one.

```
def migrate(cache, data, new_class):
    """ Migrate given data pages. """
    for key, old_value in data:
        # read data
        print(old_value)
        # ExpenseVoucher(
        #     date=datetime(2017, 9, 21, 0, 0),
        #     reported=True,
        #     purpose='Praesent eget fermentum massa',
        #     sum=Decimal('666.67'),
        #     recipient='John Doe',
        #     cashier_id=8,
        #     version=1
```

(continues on next page)

(continued from previous page)

```
        # )

        # create new binary object
        new_value = new_class()

        # process data
        new_value.sum = old_value.sum
        new_value.purpose = old_value.purpose
        new_value.recipient = old_value.recipient
        new_value.cashier_id = old_value.cashier_id
        new_value.expense_date = old_value.date
        new_value.report_date = date.today() if old_value.reported else None

        # replace data
        cache.put(key, new_value)

        # verify data
        verify = cache.get(key)
        print(verify)
        # ExpenseVoucherV2(
        #     purpose='Praesent eget fermentum massa',
        #     sum=Decimal('666.67'),
        #     recipient='John Doe',
        #     cashier_id=8,
        #     expense_date=datetime(2017, 9, 21, 0, 0),
        #     report_date=datetime(2018, 8, 29, 0, 0),
        #     version=1,
        # )


# migrate data
with client.connect('127.0.0.1', 10800):
    accounting = client.get_or_create_cache('accounting')

    with accounting.scan() as cursor:
        migrate(accounting, cursor, ExpenseVoucherV2)
```

At this moment all the fields, defined in both of our schemas, can be available in the resulting binary object, depending on which schema was used when writing it using `put()` or similar methods. Ignite Binary API do not have the method to delete Complex object schema; all the schemas ever defined will stay in cluster until its shutdown.

This versioning mechanism is quite simple and robust, but it have its limitations. The main thing is: you can not change the type of the existing field. If you try, you will be greeted with the following message:

```
`org.apache.ignite.binary.BinaryObjectException: Wrong value has been set
[typeName=SomeType, fieldName=f1, fieldType=String, assignedValueType=int]`
```

As an alternative, you can rename the field or create a new Complex object.

4.6 Failover

File: failover.py.

When connection to the server is broken or timed out, `Client` object propagates an original exception (*OSError* or *SocketError*), but keeps its constructor's parameters intact and tries to reconnect transparently.

When *Client* detects that all nodes in the list are failed without the possibility of restoring connection, it raises a special *ReconnectError* exception.

Gather 3 Ignite nodes on *localhost* into one cluster and run:

```
from pyignite import Client
from pyignite.datatypes.cache_config import CacheMode
from pyignite.datatypes.prop_codes import PROP_NAME, PROP_CACHE_MODE, PROP_BACKUPS_
↪NUMBER
from pyignite.exceptions import SocketError


nodes = [
    ('127.0.0.1', 10800),
    ('127.0.0.1', 10801),
    ('127.0.0.1', 10802),
]


client = Client(timeout=4.0)
with client.connect(nodes):
    print('Connected')

    my_cache = client.get_or_create_cache({
        PROP_NAME: 'my_cache',
        PROP_CACHE_MODE: CacheMode.PARTITIONED,
        PROP_BACKUPS_NUMBER: 2,
    })
    my_cache.put('test_key', 0)
    test_value = 0

    # abstract main loop
    while True:
        try:
            # do the work
            test_value = my_cache.get('test_key') or 0
            my_cache.put('test_key', test_value + 1)
        except (OSError, SocketError) as e:
            # recover from error (repeat last command, check data
            # consistency or just continue  depends on the task)
            print(f'Error: {e}')
            print(f'Last value: {test_value}')
            print('Reconnecting')
```

Then try shutting down and restarting nodes, and see what happens.

```
# Connected
# Error: Connection broken.
# Last value: 2650
# Reconnecting
# Error: Connection broken.
# Last value: 10204
# Reconnecting
# Error: Connection broken.
# Last value: 18932
# Reconnecting
# Traceback (most recent call last):
#   ...
# pyignite.exceptions.ReconnectError: Can not reconnect: out of nodes.
```

Client reconnection do not require an explicit user action, like calling a special method or resetting a parameter. It means that instead of checking the connection status it is better for *pyignite* user to just try the supposed data operations and catch the resulting exception.

4.7 SSL/TLS

There are some special requirements for testing SSL connectivity.

The Ignite server must be configured for securing the binary protocol port. The server configuration process can be split up into these basic steps:

1. Create a key store and a trust store using Java keytool. When creating the trust store, you will probably need a client X.509 certificate. You will also need to export the server X.509 certificate to include in the client chain of trust.
2. Turn on the *SslContextFactory* for your Ignite cluster according to this document: Securing Connection Between Nodes.
3. Tell Ignite to encrypt data on its thin client port, using the settings for ClientConnectorConfiguration. If you only want to encrypt connection, not to validate client's certificate, set *sslClientAuth* property to *false*. You'll still have to set up the trust store on step 1 though.

Client SSL settings is summarized here: *Client*.

To use the SSL encryption without certificate validation just *use_ssl*.

```
from pyignite import Client

client = Client(use_ssl=True)
client.connect('127.0.0.1', 10800)
```

To identify the client, create an SSL keypair and a certificate with openssl command and use them in this manner:

```
from pyignite import Client

client = Client(
    use_ssl=True,
    ssl_keyfile='etc/.ssl/keyfile.key',
    ssl_certfile='etc/.ssl/certfile.crt',
)
client.connect('ignite-example.com', 10800)
```

To check the authenticity of the server, get the server certificate or certificate chain and provide its path in the *ssl_ca_certfile* parameter.

```
import ssl

from pyignite import Client

client = Client(
    use_ssl=True,
    ssl_ca_certfile='etc/.ssl/ca_certs',
    ssl_cert_reqs=ssl.CERT_REQUIRED,
)
client.connect('ignite-example.com', 10800)
```

You can also provide such parameters as the set of ciphers (*ssl_ciphers*) and the SSL version (*ssl_version*), if the defaults (`ssl._DEFAULT_CIPHERS` and TLS 1.1) do not suit you.

4.8 Password authentication

To authenticate you must set *authenticationEnabled* property to *true* and enable persistance in Ignite XML configuration file, as described in Authentication section of Ignite documentation.

Be advised that sending credentials over the open channel is greatly discouraged, since they can be easily intercepted. Supplying credentials automatically turns SSL on from the client side. It is highly recommended to secure the connection to the Ignite server, as described in *SSL/TLS* example, in order to use password authentication.

Then just supply *username* and *password* parameters to `Client` constructor.

```
from pyignite import Client

client = Client(username='ignite', password='ignite')
client.connect('ignite-example.com', 10800)
```

If you still do not wish to secure the connection is spite of the warning, then disable SSL explicitly on creating the client object:

```
client = Client(username='ignite', password='ignite', use_ssl=False)
```

Note, that it is not possible for Ignite thin client to obtain the cluster's authentication settings through the binary protocol. Unexpected credentials are simply ignored by the server. In the opposite case, the user is greeted with the following message:

```
# pyignite.exceptions.HandshakeError: Handshake error: Unauthenticated sessions are␣
↪prohibited.
```

CHAPTER 5

Asynchronous client examples

File: async_key_value.py.

5.1 Basic usage

Asynchronous client and cache (*AioClient* and *AioCache*) has mostly the same API as synchronous ones (*Client* and *Cache*). But there is some peculiarities.

5.1.1 Basic key-value

Firstly, import dependencies.

```
from pyignite import AioClient
```

Let's connect to cluster and perform key-value queries.

```
client = AioClient()
async with client.connect('127.0.0.1', 10800):
    # Create cache
    cache = await client.get_or_create_cache('test_async_cache')

    # Load data concurrently.
    await asyncio.gather(
        *[cache.put(f'key_{i}', f'value_{i}') for i in range(0, 20)]
    )

    # Key-value queries.
    print(await cache.get('key_10'))
    print(await cache.get_all([f'key_{i}' for i in range(0, 10)]))
    # value_10
    # {'key_3': 'value_3', 'key_2': 'value_2', 'key_1': 'value_1','....}
```

5.1.2 Scan

The `scan()` method returns `AioScanCursor`, that yields the resulting rows.

```
# Scan query.
async with cache.scan() as cursor:
    async for k, v in cursor:
        print(f'key = {k}, value = {v}')
# key = key_42, value = value_42
# key = key_43, value = value_43
# key = key_40, value = value_40
# key = key_41, value = value_41
# key = key_37, value = value_37
# key = key_51, value = value_51
# key = key_20, value = value_20
# ......
```

5.1.3 ExpiryPolicy

File: expiry_policy.py.

You can enable expiry policy (TTL) by two approaches.

Firstly, expiry policy can be set for entire cache by setting `PROP_EXPIRY_POLICY` in cache settings dictionary on creation.

```
ttl_cache = await client.create_cache({
    PROP_NAME: 'test',
    PROP_EXPIRY_POLICY: ExpiryPolicy(create=timedelta(seconds=1.0))
})
```

```
await ttl_cache.put(1, 1)
await asyncio.sleep(0.5)
value = await ttl_cache.get(1)
print(f"key = {1}, value = {value}")
# key = 1, value = 1
await asyncio.sleep(1.2)
value = await ttl_cache.get(1)
print(f"key = {1}, value = {value}")
# key = 1, value = None
```

Secondly, expiry policy can be set for all cache operations, which are done under decorator. To create it use *with_expire_policy()*

```
ttl_cache = simple_cache.with_expire_policy(access=timedelta(seconds=1.0))
await ttl_cache.put(1, 1)
await asyncio.sleep(0.5)
value = await ttl_cache.get(1)
print(f"key = {1}, value = {value}")
# key = 1, value = 1
await asyncio.sleep(1.7)
value = await ttl_cache.get(1)
print(f"key = {1}, value = {value}")
# key = 1, value = None
```

5.2 Transactions

File: transactions.py.

Client transactions are supported for caches with `TRANSACTIONAL` mode. **Supported only python 3.7+**

Let's create transactional cache:

```
cache = await client.get_or_create_cache({
    PROP_NAME: 'tx_cache',
    PROP_CACHE_ATOMICITY_MODE: CacheAtomicityMode.TRANSACTIONAL
})
```

Let's start a transaction and commit it:

```
key = 1
async with client.tx_start(
        isolation=TransactionIsolation.REPEATABLE_READ,
        concurrency=TransactionConcurrency.PESSIMISTIC
) as tx:
    await cache.put(key, 'success')
    await tx.commit()
```

Let's check that the transaction was committed successfully:

```
val = await cache.get(key)
print(f"key={key} value={val}")
```

Let's check that raising exception inside *async with* block leads to transaction's rollback

```
try:
    async with client.tx_start(
            isolation=TransactionIsolation.REPEATABLE_READ,
            concurrency=TransactionConcurrency.PESSIMISTIC
    ):
        await cache.put(key, 'fail')
        raise RuntimeError('test')
except RuntimeError:
    pass

# key=1 value=success
val = await cache.get(key)
print(f"key={key} value={val}")
```

Let's check that timed out transaction is successfully rolled back

```
try:
    async with client.tx_start(timeout=1000, label='long-tx') as tx:
        await cache.put(key, 'fail')
        await asyncio.sleep(2.0)
        await tx.commit()
except CacheError as e:
    # Cache transaction timed out: GridNearTxLocal[...timeout=1000, ... label=long-tx]
    print(e)

# key=1 value=success
val = await cache.get(key)
print(f"key={key} value={val}")
```

See more info about transaction's parameters in a documentation of *tx_start()*

5.3 SQL

File: async_sql.py.

First let us establish a connection.

```
client = AioClient()
async with client.connect('127.0.0.1', 10800):
```

Then create tables. Begin with *Country* table, than proceed with related tables *City* and *CountryLanguage*.

```

COUNTRY_CREATE_TABLE_QUERY = '''CREATE TABLE Country (
    Code CHAR(3) PRIMARY KEY,
    Name CHAR(52),
    Continent CHAR(50),
    Region CHAR(26),
    SurfaceArea DECIMAL(10,2),
    IndepYear SMALLINT(6),
    Population INT(11),
    LifeExpectancy DECIMAL(3,1),
    GNP DECIMAL(10,2),
    GNPOld DECIMAL(10,2),
    LocalName CHAR(45),
    GovernmentForm CHAR(45),
    HeadOfState CHAR(60),
    Capital INT(11),
    Code2 CHAR(2)
)'''

CITY_CREATE_TABLE_QUERY = '''CREATE TABLE City (
    ID INT(11),
    Name CHAR(35),
    CountryCode CHAR(3),
    District CHAR(20),
    Population INT(11),
    PRIMARY KEY (ID, CountryCode)
) WITH "affinityKey=CountryCode"'''

LANGUAGE_CREATE_TABLE_QUERY = '''CREATE TABLE CountryLanguage (
    CountryCode CHAR(3),
    Language CHAR(30),
    IsOfficial BOOLEAN,
    Percentage DECIMAL(4,1),
    PRIMARY KEY (CountryCode, Language)
) WITH "affinityKey=CountryCode"'''
```

```
    # create tables
    for query in [
        COUNTRY_CREATE_TABLE_QUERY,
        CITY_CREATE_TABLE_QUERY,
        LANGUAGE_CREATE_TABLE_QUERY,
    ]:
        await client.sql(query)
```

Create indexes.

```
CITY_CREATE_INDEX = '''
CREATE INDEX idx_country_code ON city (CountryCode)'''

LANGUAGE_CREATE_INDEX = '''
CREATE INDEX idx_lang_country_code ON CountryLanguage (CountryCode)'''
```

```
# create indices
for query in [CITY_CREATE_INDEX, LANGUAGE_CREATE_INDEX]:
    await client.sql(query)
```

Fill tables with data.

```
COUNTRY_INSERT_QUERY = '''INSERT INTO Country(
    Code, Name, Continent, Region,
    SurfaceArea, IndepYear, Population,
    LifeExpectancy, GNP, GNPOld,
    LocalName, GovernmentForm, HeadOfState,
    Capital, Code2
) VALUES (?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?)'''

CITY_INSERT_QUERY = '''INSERT INTO City(
    ID, Name, CountryCode, District, Population
) VALUES (?, ?, ?, ?, ?)'''

LANGUAGE_INSERT_QUERY = '''INSERT INTO CountryLanguage(
    CountryCode, Language, IsOfficial, Percentage
) VALUES (?, ?, ?, ?)'''
```

```
await asyncio.gather(*[
    client.sql(COUNTRY_INSERT_QUERY, query_args=row) for row in COUNTRY_DATA
])

await asyncio.gather(*[
    client.sql(CITY_INSERT_QUERY, query_args=row) for row in CITY_DATA
])

await asyncio.gather(*[
    client.sql(LANGUAGE_INSERT_QUERY, query_args=row) for row in LANGUAGE_DATA
])
```

Now let us answer some questions.

5.3.1 What are the 10 largest cities in our data sample (population-wise)?

```
MOST_POPULATED_QUERY = '''
SELECT name, population FROM City ORDER BY population DESC LIMIT 10'''

async with client.sql(MOST_POPULATED_QUERY) as cursor:
    print('Most 10 populated cities:')
```

(continues on next page)

(continued from previous page)

```
    async for row in cursor:
        print(row)
# Most 10 populated cities:
# ['Mumbai (Bombay)', 10500000]
# ['Shanghai', 9696300]
# ['New York', 8008278]
# ['Peking', 7472000]
# ['Delhi', 7206704]
# ['Chongqing', 6351600]
# ['Tianjin', 5286800]
# ['Calcutta [Kolkata]', 4399819]
# ['Wuhan', 4344600]
# ['Harbin', 4289800]
```

The *sql()* method returns `AioSqlFieldsCursor`, that yields the resulting rows.

5.3.2 What are the 10 most populated cities throughout the 3 chosen countries?

If you set the *include_field_names* argument to *True*, the *sql()* method will generate a list of column names as a first yield. Unfortunately, there is no async equivalent of *next* but you can await `__anext__()` of `AioSqlFieldsCursor`

```
MOST_POPULATED_IN_3_COUNTRIES_QUERY = '''
SELECT country.name as country_name, city.name as city_name, MAX(city.population) AS␣
↪max_pop FROM country
    JOIN city ON city.countrycode = country.code
    WHERE country.code IN ('USA','IND','CHN')
    GROUP BY country.name, city.name ORDER BY max_pop DESC LIMIT 10
'''

async with client.sql(MOST_POPULATED_IN_3_COUNTRIES_QUERY, include_field_names=True)␣
↪as cursor:
    print('Most 10 populated cities in USA, India and China:')
    print(await cursor.__anext__())
    print('----------------------------------------')
    async for row in cursor:
        print(row)
# Most 10 populated cities in USA, India and China:
# ['COUNTRY_NAME', 'CITY_NAME', 'MAX_POP']
# ----------------------------------------
# ['India', 'Mumbai (Bombay)', 10500000]
# ['China', 'Shanghai', 9696300]
# ['United States', 'New York', 8008278]
# ['China', 'Peking', 7472000]
# ['India', 'Delhi', 7206704]
# ['China', 'Chongqing', 6351600]
# ['China', 'Tianjin', 5286800]
# ['India', 'Calcutta [Kolkata]', 4399819]
# ['China', 'Wuhan', 4344600]
# ['China', 'Harbin', 4289800]
```

5.3.3 Display all the information about a given city

```
# show city info
CITY_INFO_QUERY = '''SELECT * FROM City WHERE id = ?'''

async with client.sql(CITY_INFO_QUERY, query_args=[3802], include_field_names=True)␣
↪as cursor:
    field_names = await cursor.__anext__()
    field_data = await cursor.__anext__()

    print('City info:')
    for field_name, field_value in zip(field_names * len(field_data), field_data):
        print('{}: {}'.format(field_name, field_value))
# City info:
# ID: 3802
# NAME: Detroit
# COUNTRYCODE: USA
# DISTRICT: Michigan
# POPULATION: 951270
```

Finally, delete the tables used in this example with the following queries:

```
DROP_TABLE_QUERY = '''DROP TABLE {} IF EXISTS'''
```

```
# clean up concurrently.
await asyncio.gather(*[
    client.sql(DROP_TABLE_QUERY.format(table_name)) for table_name in [
        CITY_TABLE_NAME,
        LANGUAGE_TABLE_NAME,
        COUNTRY_TABLE_NAME,
    ]
])
```

CHAPTER 6

Indices and tables

- genindex
- modindex
- search

Python Module Index

p

Index

Symbols

A

B

C

D

G

H

N

P

Q

R

S

T

U

W