

PostgreSQL Partition Manager Extension (pg_partman)

Contents

PG Partition Manager	2
INSTALLATION	2
UPGRADE	3
EXAMPLE	3
TESTING	3
LICENSE AND COPYRIGHT	3
About	4
Child Table Property Inheritance	4
Sub-partitioning	4
Retention	4
Constraint Exclusion	5
Custom Time Interval Considerations	5
Naming Length Limits	5
Unique Constraints & Upsert	5
Logging/Monitoring	6
Background Worker	6
Extension Objects	6
Creation Functions	6
Maintenance Functions	9
Destruction Functions	11
Tables	13
Scripts	15
Simple Time Based: 1 Partition Per Day	19
Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table	20
Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table and using upsert to drop conflicting rows	22
Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table and using upsert to update conflicting rows	24
Sub-partition Time->Time->Time: Yearly -> Monthly -> Daily	25
Sub-partition ID->ID->ID: 10,000 -> 1,000 -> 100	27
Set run_maintenance() to run often enough	34
Use Retention Policy	34
Undo Partitioning: Simple Time Based	34
Undo Partitioning: Simple Serial ID	35
Undo Partitioning: Sub-partition ID->ID->ID	35
Undo Partitioning: Sub-partition Time->Time->Time	36

PG Partition Manager

`pg_partman` is an extension to create and manage both time-based and serial-based table partition sets. Sub-partitioning is also supported. Child table & trigger function creation is all managed by the extension itself. Tables with existing data can also have their data partitioned in easily managed smaller batches. Optional retention policy can automatically drop partitions no longer needed. A background worker (BGW) process is included to automatically run partition maintenance without the need of an external scheduler (cron, etc) in most cases.

All bug reports, feature requests and general questions can be directed to the Issues section on Github. Please feel free to post here no matter how minor you may feel your issue or question may be. - https://github.com/keithf4/pg_partman/issues

If you're looking for a partitioning system that handles any range type beyond just time & serial, check out https://github.com/moat/range_partitioning. Note that if you are doing time/serial, the methods used in `pg_partman` are much more efficient and will provide better performance. But if you need greater flexibility, the `range_partitioning` extension should work very well for you.

WARNING: The `pg_partman` background worker is currently incompatible with PostgreSQL 9.6. It runs fine in 9.4 and 9.5. I'm still working on trying to figure out what the issue is. If you must upgrade to 9.6 before this issue is fixed, you can disable the background worker by removing the entry for it in `shared_preload_libraries` and restarting postgres. Instead schedule the `run_maintenance()` `plpgsql` function to run via another scheduler method (cron). Note that the background worker technically runs in 9.6 and won't cause errors, but it essentially ignores the interval configuration option, runs without pause and will very quickly fill up your postgres logs or cause excessive system load.

INSTALLATION

Requirement: PostgreSQL 9.4 & 9.5 (currently incompatible with 9.6)

Recommended: `pg_jobmon` (>=v1.3.2). PG Job Monitor will automatically be used if it is installed and setup properly. https://github.com/omniti-labs/pg_jobmon

In the directory where you downloaded `pg_partman`, run

```
make install
```

If you do not want the background worker compiled and just want the plain PL/PGSQL functions, you can run this instead:

```
make NO_BGW=1 install
```

The background worker must be loaded on database start by adding the library to `shared_preload_libraries` in `postgresql.conf`

```
shared_preload_libraries = 'pg_partman_bgw'      # (change requires restart)
```

You can also set other control variables for the BGW in `postgresql.conf`. "dbname" is required at a minimum for maintenance to run on the given database(s). These can be added/changed at anytime with a simple reload. See the documentation for more details. An example with some of them:

```
pg_partman_bgw.interval = 3600
pg_partman_bgw.role = 'keith'
pg_partman_bgw.dbname = 'keith'
```

Log into PostgreSQL and run the following commands. Schema is optional (but recommended) and can be whatever you wish, but it cannot be changed after installation. If you're using the BGW, the database cluster can be safely started without having the extension first created in the configured database(s). You can create the extension at any time and the BGW will automatically pick up that it exists without restarting the cluster (as long as `shared_preload_libraries` was set) and begin running maintenance as configured.

```
CREATE SCHEMA partman;
CREATE EXTENSION pg_partman SCHEMA partman;
```

Functions must either be run as a superuser or you can set the ownership of the extension functions to a superuser role and they will also work (SECURITY DEFINER is set).

I've received many requests for being able to install this extension on Amazon RDS. RDS does not support third-party extension management outside of the ones it has approved and provides itself. Therefore, I cannot provide support for running this extension in RDS if the limitations are RDS related. If you'd like to see this extension available there, please send an email to rds-postgres-extensions-request@amazon.com requesting that they include it. The more people that do so, the more likely it will happen!

Version 1.8.8 of `pg_partman` is still available on github if you're running a version of PostgreSQL older than 9.4. Note however that no further updates (bug fixes, features, etc) are being released for the 1.x series. If you encounter any issues, please plan for upgrading your database to 9.4+ so that you can use the 2.x series of `pg_partman`.

UPGRADE

Run “make install” same as above to put the script files and libraries in place. Then run the following in PostgreSQL itself:

```
ALTER EXTENSION pg_partman UPDATE TO '<latest version>';
```

If you are doing a `pg_dump/restore` and you’ve upgraded `pg_partman` in place from previous versions, it is recommended you use the `-column-inserts` option when dumping and/or restoring `pg_partman`’s configuration tables. This is due to ordering of the configuration columns possibly being different (upgrades just add the columns onto the end, whereas the default of a new install may be different).

If upgrading from 1.x to 2.x, please see the `CHANGELOG` or the notes in the update script itself for additional instructions for updating your trigger functions to the newer version and other important considerations for the update.

EXAMPLE

First create a parent table with an appropriate column type for the partitioning type you will do. Apply all defaults, indexes, constraints, privileges & ownership to the parent table and they will be inherited to newly created child tables automatically (not already existing partitions, see docs for how to fix that). Here’s one with columns that can be used for either

```
CREATE schema test;  
CREATE TABLE test.part_test (col1 serial, col2 text, col3 timestamptz NOT NULL DEFAULT now());
```

Then just run the `create_parent()` function with the appropriate parameters

```
SELECT partman.create_parent('test.part_test', 'col3', 'time', 'daily');  
or  
SELECT partman.create_parent('test.part_test', 'col1', 'id', '100000');
```

This will turn your table into a parent table and premake 4 future partitions and also make 4 past partitions. To make new partitions for time-based partitioning, schedule the `run_maintenance()` function to run periodically or use the background worker settings in `postgresql.conf` (the latter is recommended). Serial based partitioning does not always require `run_maintenance()` (see doc file below).

This should be enough to get you started. Please see the [pg_partman.md](#) file in the doc folder for more information on the types of partitioning supported and what the parameters in the `create_parent()` function mean.

TESTING

This extension can use the `pgTAP` unit testing suite to evaluate if it is working properly (<http://www.pgtap.org>). **WARNING:** You **MUST** increase `max_locks_per_transaction` above the default value of 64. For me, 128 has worked well so far. This is due to the sub-partitioning tests that create/destroy several hundred tables in a single transaction. If you don’t do this, you risk a cluster crash when running subpartitioning tests.

LICENSE AND COPYRIGHT

PG Partition Manager (`pg_partman`) is released under the PostgreSQL License, a liberal Open Source license, similar to the BSD or MIT licenses.

Copyright (c) 2016 OmniTI, Inc.

Permission to use, copy, modify, and distribute this software and its documentation for any purpose, without fee, and without a written agreement is hereby granted, provided that the above copyright notice and this paragraph and the following two paragraphs appear in all copies.

IN NO EVENT SHALL THE AUTHOR BE LIABLE TO ANY PARTY FOR DIRECT, INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES, INCLUDING LOST PROFITS, ARISING OUT OF THE USE OF THIS SOFTWARE AND ITS DOCUMENTATION, EVEN IF THE AUTHOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

THE AUTHOR SPECIFICALLY DISCLAIMS ANY WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE SOFTWARE PROVIDED HEREUNDER IS ON AN “AS IS” BASIS, AND THE AUTHOR HAS NO OBLIGATIONS TO PROVIDE MAINTENANCE, SUPPORT, UPDATES, ENHANCEMENTS, OR MODIFICATIONS.

About

PostgreSQL Partition Manager is an extension to help make managing time or serial id based table partitioning easier. It has many options, but usually only a few are needed, so it's much easier to use than it may seem (and definitely easier than implementing it yourself). Currently the trigger functions only handle inserts to the parent table. Updates that would move a value from one partition to another are not yet supported. Some features of this extension have been expanded upon in the author's blog - http://www.keithf4.com/tag/pg_partman

If you attempt to insert data into a partition set that contains data for a partition that does not exist, that data will be placed into the set's parent table. This is preferred over automatically creating new partitions to match that data since a mistake that is causing non-partitioned data to be inserted could cause a lot of unwanted child tables to be made. The `check_parent()` function provides monitoring for any data getting inserted into parents and the `partition_data_*` set of functions can easily partition that data for you if it is valid data. That is much easier than having to clean up potentially hundreds or thousands of unwanted partitions. And also better than throwing an error and losing the data!

Note that future child table creation is based on the data currently in the partition set for both time & serial partitioning. This means that if you put "future" data into a partition set, newly created tables will be based off that value. This may cause intervening data to go to the parent as stated above if no child table exists. It is recommended that you set the `premake` value high enough to encompass your expected data range being inserted and the `optimize_trigger` value to efficiently handle your most frequent data range. See below for further explanations on these configuration values.

If you have an existing partition set and you'd like to migrate it to `pg_partman`, please see the `migration.md` file in the `doc` folder.

Child Table Property Inheritance

For this extension, most of the attributes of the child partitions are all obtained from the original parent. This includes defaults, indexes (primary keys, unique, clustering, etc), foreign keys (optional), tablespace, constraints, privileges & ownership. This also includes the OID and UNLOGGED table properties. For managing privileges, whenever a new partition is created it will obtain its privilege & ownership information from what the parent has at that time. Previous partition privileges are not changed. If previous partitions require that their privileges be updated, a separate function is available. This is kept as a separate process due to being an expensive operation when the partition set grows larger. The defaults, indexes, tablespace & constraints on the parent are only applied to newly created partitions and are not retroactively set on ones that already existed. While you would not normally create indexes on the parent of a partition set, doing so makes it much easier to manage in this case. There will be no data in the parent table (if everything is working right), so they will not take up any space or have any impact on system performance. Using the parent table as a control to the details of the child tables like this gives a better place to manage things that's a little more natural than a configuration table or using setup functions.

Sub-partitioning

Sub-partitioning with multiple levels is supported. You can do time->time, id->id, time->id and id->time. There is no set limit on the level of subpartitioning you can do, but be sensible and keep in mind performance considerations on managing many tables in a single inheritance set. Also, if the number of tables in a single partition set gets very high, you may have to adjust the `max_locks_per_transaction` `postgresql.conf` setting above the default of 64. Otherwise you may run into shared memory issues or even crash the cluster. By default all subpartition sets require `run_maintenance()` for the creation of new partitions. Single level time-based partition sets already do this, but single level serial sets do not. If you have contention issues when `run_maintenance()` is called for general maintenance of all partition sets, you can set the `use_run_maintenance` column in the `part_config` table to false if you do not want that general call to manage your subpartition set. But you must then call `run_maintenance(parent_table)` directly, and often enough, to have to future partitions made. See the `create_parent_sub()` & `run_maintenance()` functions below for more information.

Retention

If you don't need to keep data in older partitions, a retention system is available to automatically drop unneeded child partitions. By default, they are only uninherited not actually dropped, but that can be configured if desired. If the old partitions are kept, dropping their indexes can also be configured to recover disk space. Note that this will also remove any primary key or unique constraints in order to allow the indexes to be dropped. There is also a method available to dump the tables out if they don't need to be in the database anymore but still need to be kept. To set the retention policy, enter either an interval or integer value into the `retention` column of the `part_config` table. For time-based partitioning, the interval value will set that any partitions containing only data older than that will be dropped. For id-based partitioning, the integer value will set that any partitions with an id value less than the current maximum id value minus the retention value will be dropped. For example, if the current max id is 100 and the retention value is 30, any partitions with id values less than 70 will be dropped. The current maximum id value at the time the drop function is run is always used. Keep in mind that for subpartition sets, when a parent table has a child dropped, if that child table is in turn partitioned, the drop is a CASCADE and ALL child tables down the entire inheritance tree will be dropped.

Constraint Exclusion

One of the big advantages of partitioning is a feature called **constraint exclusion** (see docs for explanation of functionality and examples <http://www.postgresql.org/docs/current/static/ddl-partitioning.html#DDL-PARTITIONING-CONSTRAINT-EXCLUSION>). The problem with most partitioning setups however, is that this will only be used on the partitioning control column. If you use a WHERE condition on any other column in the partition set, a scan across all child tables will occur unless there are also constraints on those columns. And predicting what a column's values will be to precreate constraints can be very hard or impossible. `pg_partman` has a feature to apply constraints on older tables in a partition set that may no longer have any edits done to them ("old" being defined as older than the `optimize_constraint` config value). It checks the current min/max values in the given columns and then applies a constraint to that child table. This can allow the constraint exclusion feature to potentially eliminate scanning older child tables when other columns are used in WHERE conditions. Be aware that this limits being able to edit those columns, but for the situations where it is applicable it can have a tremendous affect on query performance for very large partition sets. So if you are only inserting new data this can be very useful, but if data is regularly being inserted throughout the entire partition set, this is of limited use. Functions for easily recreating constraints are also available if data does end up having to be edited in those older partitions. Note that constraints managed by PG Partman SHOULD NOT be renamed in order to allow the extension to manage them properly for you. For a better example of how this works, please see this blog post: <http://www.keithf4.com/managing-constraint-exclusion-in-table-partitioning>

NOTE: This may not work with sub-partitioning. It will work on the first level of partitioning, but is not guaranteed to work properly on further sub-partition sets depending on the interval combinations and the `optimize_constraint` value. Ex: Weekly -> Daily with a daily `optimize_constraint` of 7 won't work as expected. Weekly constraints will get created but daily sub-partition ones likely will not.

Custom Time Interval Considerations

The smallest interval supported is 1 second and the upper limit is bounded by the minimum and maximum timestamp values that PostgreSQL supports (<http://www.postgresql.org/docs/current/static/datatype-datetime.html>).

When first running `create_parent()` to create a partition set, intervals less than a day round down when determining what the first partition to create will be. Intervals less than 24 hours but greater than 1 minute use the nearest hour rounded down. Intervals less than 1 minute use the nearest minute rounded down. However, enough partitions will be made to support up to what the real current time is. This means that when `create_parent()` is run, more previous partitions may be made than expected and all future partitions may not be made. The first run of `run_maintenance()` will fix the missing future partitions. This happens due to the nature of being able to support custom time intervals. Any intervals greater than or equal to 24 hours should set things up as would be expected.

Keep in mind that for intervals equal to or greater than 100 years, the extension will use the real start of the century or millennium to determine the partition name & constraint rules. For example, the 21st century and 3rd millennium started January 1, 2001 (not 2000). This also means there is no year "0". It's much too difficult to try to work around this and make nice "even" partition names & rules to handle all possible time periods people may need. Blame the Gregorian creators.

Naming Length Limits

PostgreSQL has an object naming length limit of 63 characters. If you try and create an object with a longer name, it truncates off any characters at the end to fit that limit. This can cause obvious issues with partition names that rely on having a specifically named suffix. PG Partman automatically handles this for all child tables, trigger functions and triggers. It will truncate off the existing parent table name to fit the required suffix. Be aware that if you have tables with very long, similar names, you may run into naming conflicts if they are part of separate partition sets. With serial based partitioning, be aware that over time the table name will be truncated more and more to fit a longer partition suffix. So while the extension will try and handle this edge case for you, it is recommended to keep table names that will be partitioned as short as possible.

Unique Constraints & Upsert

Table inheritance in PostgreSQL does not allow a primary key or unique index/constraint on the parent to apply to all child tables. The constraint is applied to each individual table, but not on the entire partition set as a whole. For example, this means a careless application can cause a primary key value to be duplicated in a partition set. This is one of the "big issues" that causes performance issues with partitioning on other database systems and one of the reasons for the delay in getting partitioning built in to PostgreSQL. In the mean time, a python script is included with `pg_partman` that can provide monitoring to help ensure the lack of this feature doesn't cause long term harm. See `check_unique_constraint.py` in the **Scripts** section.

INSERT ... ON CONFLICT (upsert) is supported in the partitioning trigger, but is very limited. The major limitations are that the constraint violations that would trigger the ON CONFLICT clause only occur on individual child tables that actually contain data due to reasons explained above. Of a larger concern than data duplication is an ON CONFLICT DO UPDATE clause which may not fire and cause wildly inconsistent data if not accounted for. It is unclear whether this limitation will be able to be overcome while partitioning is based around inheritance and triggers. For situations where only new data is being inserted, upsert can provide significant performance improvements. However, if you're relying on data in older partitions to cause a constraint violation that upsert would normally handle, it likely will not work. Also, if the resulting UPDATE would end up violating the partitioning constraint of that child table, it will fail. `pg_partman` does not currently support UPDATES that would require moving a row from one child table to another.

Upsert is optional, turned off by default and is recommended you test it out extensively before implementing in production and monitor it carefully. See <https://www.postgresql.org/docs/9.5/static/sql-insert.html>

Logging/Monitoring

The PG Jobmon extension (https://github.com/omniti-labs/pg_jobmon) is optional and allows auditing and monitoring of partition maintenance. If jobmon is installed and configured properly, it will automatically be used by partman with no additional setup needed. Jobmon can also be turned on or off individually for each partition set by using the `jobmon` column in the `part_config` table or with the option to `create_parent()` during initial setup. Note that if you try to partition `pg_jobmon`'s tables you **MUST** set the option in `create_parent()` to false, otherwise it will be put into a permanent lockwait since `pg_jobmon` will be trying to write to the table it's trying to partition. By default, any function that fails to run successfully 3 consecutive times will cause jobmon to raise an alert. This is why the default pre-make value is set to 4 so that an alert will be raised in time for intervention with no additional configuration of jobmon needed. You can of course configure jobmon to alert before (or later) than 3 failures if needed. If you're running partman in a production environment it is **HIGHLY** recommended to have jobmon installed and some sort of 3rd-party monitoring configured with it to alert when partitioning fails (Nagios, Circonus, etc).

Background Worker

With PostgreSQL 9.4, the ability to create custom background workers and dynamically load them during runtime was introduced. `pg_partman`'s BGW is basically just a scheduler that runs the `run_maintenance()` function for you so that you don't have to use an external scheduler (cron, etc). Right now it doesn't do anything differently than calling `run_maintenance()` directly, but that may change in the future. See the README.md file for installation instructions. If you need to call `run_maintenance()` directly on any specific partition sets, you will still need to do so manually using an outside scheduler. This only maintains partition sets that have `use_run_maintenance` in `**part_config**` set to true. LOG messages are output to the normal PostgreSQL log file to indicate when the BGW runs. Additional logging messages are available if `log_min_messages` is set to "DEBUG1".

The following configuration options are available to add into `postgresql.conf` to control the BGW process:

- `pg_partman_bgw.dbname`
 - Required. The database(s) that `run_maintenance()` will run on. If more than one, use a comma separated list. If not set, BGW will do nothing.
- `pg_partman_bgw.interval`
 - Number of seconds between calls to `run_maintenance()`. Default is 3600 (1 hour).
 - See further documentation below on suggested values for this based on partition types & intervals used.
- `pg_partman_bgw.role`
 - The role that `run_maintenance()` will run as. Default is "postgres". Only a single role name is allowed.
- `pg_partman_bgw.analyze`
 - Same purpose as the `p_analyze` argument to `run_maintenance()`. See below for more detail. Set to 'on' for TRUE. Set to 'off' for FALSE. Default is 'on'.
- `pg_partman_bgw.jobmon`
 - Same purpose as the `p_jobmon` argument to `run_maintenance()`. See below for more detail. Set to 'on' for TRUE. Set to 'off' for FALSE. Default is 'on'.

Extension Objects

A superuser must be used to run all these functions in order to set privileges & ownership properly in all cases. All are set with SECURITY DEFINER, so if you cannot have a superuser running them just assign a superuser role as the owner.

As a note for people that were not aware, you can name arguments in function calls to make calling them easier and avoid confusion when there are many possible arguments. If a value has a default listed, it is not required to pass a value to that argument. As an example: `SELECT create_parent('schema.table', 'col1', 'time', 'daily', p_start_partition := '2015-10-20');`

Creation Functions

```
create_parent(p_parent_table text, p_control text, p_type text, p_interval text, p_constraint_cols text[] DEFAULT NULL, p_premake int DEFAULT 4, p_use_run_maintenance boolean DEFAULT NULL, p_start_partition text DEFAULT NULL, p_inherit_fk boolean DEFAULT true, p_epoch boolean DEFAULT 'none', p_upsert text DEFAULT '', p_trigger_return_null boolean DEFAULT true, p_jobmon boolean DEFAULT true, p_debug boolean DEFAULT false)
```

- Main function to create a partition set with one parent table and inherited children. Parent table must already exist. Please apply all defaults, indexes, constraints, privileges & ownership to parent table so they will propagate to children.
- An ACCESS EXCLUSIVE lock is taken on the parent table during the running of this function. No data is moved when running this function, so lock should be brief.
- `p_parent_table` - the existing parent table. MUST be schema qualified, even if in public schema.
- `p_control` - the column that the partitioning will be based on. Must be a time or integer based column.
- `p_type` - one of the following values to set the partitioning type that will be used:
 - **time**
 - * Create a time-based partition set using a predefined interval below.
 - * The number of partitions most efficiently managed behind and ahead of the current one is determined by the **optimize_trigger** config value in the `part_config` table (default of 4 means data for 4 previous and 4 future partitions are handled best).
 - * *Beware setting the `optimize_trigger` value too high as that will lessen the efficiency boost.*
 - * Inserts to the parent table outside the `optimize_trigger` window will go to the proper child table if it exists, but performance will be degraded due to the higher overhead of handling that condition.
 - * If the child table does not exist for that time value, the row will go to the parent.
 - * Child table creation & trigger function is kept up to date by `run_maintenance()` function.
 - **time-custom**
 - * Allows use of any time interval instead of the predefined ones below. Works the same as “time”.
 - * Note this method uses a lookup table as well as a dynamic insert statement, so performance will not be as good as the predefined intervals. So, while it is more flexible, it sacrifices speed.
 - * Child table creation is kept up to date by `run_maintenance()` function.
 - **id**
 - * Create a serial/id-based partition set. Same functionality & use of `optimize_trigger` value as the “time” method.
 - * By default, when the id value reaches 50% of the max value for that partition, it will automatically create the next partition in sequence if it doesn’t yet exist. This can be changed to use `run_maintenance()` instead. See the notes for this function below.
 - * Note that the 50% rule is NOT true if the id set is sub-partitioned. Then `run_maintenance()` must be used.
 - * Only supports id values greater than or equal to zero.
- `p_interval` - the time or numeric range interval for each partition. No matter the partitioning type, value must be given as text. The generic intervals of “yearly -> quarter-hour” are for the “time” type and allow better performance than using an arbitrary time interval (time-custom).
 - *yearly* - One partition per year
 - *quarterly* - One partition per yearly quarter. Partitions are named as YYYYqQ (ex: 2012q4)
 - *monthly* - One partition per month
 - *weekly* - One partition per week. Follows ISO week date format (http://en.wikipedia.org/wiki/ISO_week_date). Partitions are named as IYYYwIW (ex: 2012w36)
 - *daily* - One partition per day
 - *hourly* - One partition per hour
 - *half-hour* - One partition per 30 minute interval on the half-hour (1200, 1230)
 - *quarter-hour* - One partition per 15 minute interval on the quarter-hour (1200, 1215, 1230, 1245)
 - *<interval>* - For the time-custom partitioning type, this can be any interval value that is valid for the PostgreSQL interval data type. Do not type cast the parameter value, just leave as text.
 - *<integer>* - For ID based partitions, the integer value range of the ID that should be set per partition. Enter this as an integer in text format (‘100’ not 100). Must be greater than or equal to 10.
- `p_constraint_cols` - an optional array parameter to set the columns that will have additional constraints set. See the **About** section above for more information on how this works and the **apply_constraints()** function for how this is used.
- `p_premake` - is how many additional partitions to always stay ahead of the current partition. Default value is 4. This will keep at minimum 5 partitions made, including the current one. For example, if today was Sept 6th, and `premake` was set to 4 for a daily partition, then partitions would be made for the 6th as well as the 7th, 8th, 9th and 10th. Note some intervals may occasionally cause an extra partition to be premade or one to be missed due to leap years, differing month lengths, daylight savings (on non-UTC systems), etc. This won’t hurt anything and will self-correct. If partitioning ever falls behind the `premake` value, normal running of `run_maintenance()` and data insertion should automatically catch things up.
- `p_use_run_maintenance` - Used to tell partman whether you’d like to override the default way that child partitions are created. Set this value to TRUE to allow you to use the `run_maintenance()` function, without any table parameter, to create new child tables for serial partitioning instead of using 50% method mentioned above. Time based partitioning MUST use `run_maintenance()`, so either leave this value true or call the `run_maintenance()` function directly on a partition set by passing the parent table as a parameter. See **run_maintenance** in Maintenance Functions section below for more info.
- `p_start_partition` - allows the first partition of a set to be specified instead of it being automatically determined. Must be a valid timestamp (for time-based) or positive integer (for id-based) value. Be aware, though, the actual parameter data type is text. For

time-based partitioning, all partitions starting with the given timestamp up to `CURRENT_TIMESTAMP` (plus `premake`) will be created. For id-based partitioning, only the partition starting at the given value (plus `premake`) will be made.

- `p_inherit_fk` - allows `pg_partman` to automatically manage inheriting any foreign keys that exist on the parent table to all its children. Defaults to `TRUE`.
- `p_epoch` - tells `pg_partman` that the control column is an integer type, but actually represents an epoch time value. You can also specify whether the value is seconds or milliseconds. Valid values for this option are: 'seconds', 'milliseconds' & 'none'. The default is 'none'. All triggers, constraints & table names will be time-based. In addition to a normal index on the control column, be sure you create a functional, time-based index on the control column (`to_timestamp(controlcolumn)`) as well so this works efficiently.
- `p_upsert` - adds upsert to insert queries in the partition trigger to allow handling of conflicts Defaults to "" (empty string) which means it's inactive.
 - the value entered here is the entire `ON CONFLICT` clause which will then be appended to the `INSERT` statement(s) in the trigger
 - Ex: to ignore conflicting rows on a table with primary key "id" set `p_upsert` to '`ON CONFLICT (id)DO NOTHING`'
 - Ex: to update a conflicting row on a table with columns (id(pk), val) set `p_upsert` to '`ON CONFLICT (id)DO UPDATE SET val=EXCLUDED.val`'
 - Requires postgresql 9.5
 - See *About* section above for more info.
- `p_trigger_return_null` - Boolean value that allows controlling the behavior of the partition trigger `RETURN`. By default this is true and the trigger returns `NULL` to prevent data going into the parent table as well as the children. However, if you have multiple triggers and are relying on the return to be the `NEW` column value, this can cause a problem. Setting this config value to false will cause the partition trigger to `RETURN NEW`. You are then responsible for handling the return value in another trigger appropriately. Otherwise, this will cause new data to go into both the child and parent table of the partition set.
- `p_jobmon` - allow `pg_partman` to use the `pg_jobmon` extension to monitor that partitioning is working correctly. Defaults to `TRUE`.
- `p_debug` - turns on additional debugging information.

```
create_sub_parent(p_top_parent text, p_control text, p_type text, p_interval text, p_constraint_cols text[]
DEFAULT NULL, p_premake int DEFAULT 4, p_start_partition text DEFAULT NULL, p_inherit_fk boolean DEFAULT true,
p_epoch boolean DEFAULT 'none', p_jobmon boolean DEFAULT true, p_debug boolean DEFAULT false)RETURNS boolean;
```

- Create a subpartition set of an already existing partitioned set.
- `p_top_parent` - This parameter is the parent table of an already existing partition set. It tells `pg_partman` to turn all child tables of the given partition set into their own parent tables of their own partition sets using the rest of the parameters for this function.
- All other parameters to this function have the same exact purpose as those of `create_parent()`, but instead are used to tell `pg_partman` how each child table shall itself be partitioned.
- For example if you have an existing partition set done by year and you then want to partition each of the year partitions by day, you would use this function.
- It is advised that you keep table names short for subpartition sets if you plan on relying on the table names for organization. The suffix added on to the end of a table name is always guaranteed to be there for whatever partition type is active for that set, but if the total length is longer than 63 chars, the original name will get truncated. Longer table names may cause the original parent table names to be truncated and possibly cut off the top level partitioning suffix. I cannot control this and made the requirement that the lowest level partitioning suffix survives.
- Note that for the first level of subpartitions, the `p_parent_table` argument you originally gave to `create_parent()` would be the exact same value you give to `create_sub_parent()`. If you need further subpartitioning, you would then start giving `create_sub_parent()` a different value (the child tables of the top level partition set).
- Note that for ID sub-partitioning, future partition maintenance must be done with `run_maintenance()` and does not use the 50% rule mentioned above.

```
partition_data_time(p_parent_table text, p_batch_count int DEFAULT 1, p_batch_interval interval DEFAULT NULL,
p_lock_wait numeric DEFAULT 0, p_order text DEFAULT 'ASC', p_analyze boolean DEFAULT true)RETURNS bigint
```

- This function is used to partition data that may have existed prior to setting up the parent table as a time-based partition set, or to fix data that accidentally gets inserted into the parent.
- If the needed partition does not exist, it will automatically be created. If the needed partition already exists, the data will be moved there.
- If you are trying to partition a large amount of data automatically, it is recommended to run this function with an external script and appropriate batch settings. This will help avoid transactional locks and prevent a failure from causing an extensive rollback. See **Scripts** section for an included python script that will do this for you.
- For sub-partitioned sets, you must start partitioning data at the highest level and work your way down each level. All data will not automatically go to the lowest level when run from the top for a sub-partitioned set.
- `p_parent_table` - the existing parent table. This is assumed to be where the unpartitioned data is located. **MUST** be schema qualified, even if in public schema.

- `p_batch_interval` - optional argument, a time interval of how much of the data to move. This can be smaller than the partition interval, allowing for very large sized partitions to be broken up into smaller commit batches. Defaults to the configured partition interval if not given or if you give an interval larger than the partition interval.
- `p_batch_count` - optional argument, how many times to run the `batch_interval` in a single call of this function. Default value is 1.
- `p_lock_wait` - optional argument, sets how long in seconds to wait for a row to be unlocked before timing out. Default is to wait forever.
- `p_order` - optional argument, by default data is migrated out of the parent in ascending order (ASC). Allows you to change to descending order (DESC).
- `p_analyze` - optional argument, by default whenever a new child table is created, an analyze is run on the parent table of the partition set to ensure constraint exclusion works. This analyze can be skipped by setting this to false and help increase the speed of moving large amounts of data. If this is set to false, it is highly recommended that a manual analyze of the partition set be done upon completion to ensure statistics are updated properly.
- Returns the number of rows that were moved from the parent table to partitions. Returns zero when parent table is empty and partitioning is complete.

partition_data_id(p_parent_table text, p_batch_count int DEFAULT 1, p_batch_interval int DEFAULT NULL, p_lock_wait numeric DEFAULT 0, p_order text DEFAULT 'ASC')RETURNS bigint

- This function is used to partition data that may have existed prior to setting up the parent table as a serial id partition set, or to fix data that accidentally gets inserted into the parent.
- If the needed partition does not exist, it will automatically be created. If the needed partition already exists, the data will be moved there.
- If you are trying to partition a large amount of data automatically, it is recommended to run this function with an external script and appropriate batch settings. This will help avoid transactional locks and prevent a failure from causing an extensive rollback. See **Scripts** section for an included python script that will do this for you.
- For sub-partitioned sets, you must start partitioning data at the highest level and work your way down each level. All data will not automatically go to the lowest level when run from the top for a sub-partitioned set.
- `p_parent_table` - the existing parent table. This is assumed to be where the unpartitioned data is located. MUST be schema qualified, even if in public schema.
- `p_batch_interval` - optional argument, an integer amount representing an interval of how much of the data to move. This can be smaller than the partition interval, allowing for very large sized partitions to be broken up into smaller commit batches. Defaults to the configured partition interval if not given or if you give an interval larger than the partition interval.
- `p_batch_count` - optional argument, how many times to run the `batch_interval` in a single call of this function. Default value is 1.
- `p_lock_wait` - optional argument, sets how long in seconds to wait for a row to be unlocked before timing out. Default is to wait forever.
- `p_order` - optional argument, by default data is migrated out of the parent in ascending order (ASC). Allows you to change to descending order (DESC).
- `p_analyze` - optional argument, by default whenever a new child table is created, an analyze is run on the parent table of the partition set to ensure constraint exclusion works. This analyze can be skipped by setting this to false and help increase the speed of moving large amounts of data. If this is set to false, it is highly recommended that a manual analyze of the partition set be done upon completion to ensure statistics are updated properly.
- Returns the number of rows that were moved from the parent table to partitions. Returns zero when parent table is empty and partitioning is complete.

Maintenance Functions

run_maintenance(p_parent_table text DEFAULT NULL, p_analyze boolean DEFAULT true, p_jobmon boolean DEFAULT true, p_debug boolean DEFAULT false)RETURNS void

- Run this function as a scheduled job (cron, etc) to automatically create child tables for partition sets configured to use it.
- You can also use the included background worker (BGW) to have this automatically run for you by PostgreSQL itself. Note that the `p_parent_table` parameter is not available with this method, so if you need to run it for a specific partition set, you must do that manually or scheduled as noted above. The other parameters have postgresql.conf values that can be set. See BGW section above.
- This function also maintains the partition retention system for any partitions sets that have it turned on.
- Every run checks for all tables listed in the `part_config` table with `use_run_maintenance` set to true and either creates new partitions for them or runs their retention policy.
- By default, time-based partition sets and all sub-partition sets have `use_run_maintenance` set to true. This function is required to be run to maintain time-based partitioning & sub-partition sets.

- By default, serial-based partition sets have `use_run_maintenance` set to false (except if they are sub-partitioned) and don't require `run_maintenance()` for partition maintenance, but can be overridden to do so. By default serial partitioning creates new partitions when the current one reaches 50% of its max capacity. This can cause contention on very high transaction tables. If configured to use `run_maintenance()` for serial partitioning, you must call it often enough to keep partition creation ahead of your insertion rate, otherwise data will go into the parent.
- New partitions are only created if the number of child tables ahead of the current one is less than the `premake` value, so you can run this more often than needed without fear of needlessly creating more partitions.
- Every run checks all tables of all types listed in the `part_config` table with a value in the `retention` column and drops tables as needed (see **About** and config table below).
- Will automatically update the function for **time** partitioning (and **id** if configured) to keep the parent table pointing at the correct partitions. When using time, run this function more often than the partitioning interval to keep the trigger function running its most efficient. For example, if using quarter-hour, run every 5 minutes; if using daily, run at least twice a day, etc.
- `p_parent_table` - an optional parameter that if passed will cause `run_maintenance()` to be run for ONLY that given table. This occurs no matter what `use_run_maintenance` in `part_config` is set to. High transaction rate tables can cause contention when maintenance is being run for many tables at the same time, so this allows finer control of when partition maintenance is run for specific tables. Note that this will also cause the retention system to only be run for the given table as well.
- `p_analyze` - By default when a new child table is created, an analyze is run on the parent to ensure statistics are updated for constraint exclusion. However, for large partition sets, this analyze can take a while and if `run_maintenance()` is managing several partitions in a single run, this can cause contention while the analyze finishes. Set this to false to disable the analyze run and avoid this contention. Please note that you must then schedule an analyze of the parent table at some point for constraint exclusion to work properly on all child tables.
- `p_jobmon` - an optional parameter to control whether `run_maintenance()` itself uses the `pg_jobmon` extension to log what it does. Whether the maintenance of a particular table uses `pg_jobmon` is controlled by the setting in the `part_config` table and this setting will have no effect on that. Defaults to true if not set.
- `p_debug` - Output additional notices to help with debugging problems or to more closely examine what is being done during the run.

show_partitions (p_parent_table text, p_order text DEFAULT 'ASC')

- List all child tables of a given partition set. Each child table returned as a single row.
- Tables are returned in the order that makes sense for the partition interval, not by the locale ordering of their names.
- `p_order` - optional parameter to set the order the child tables are returned in. Defaults to ASCending. Set to 'DESC' to return in descending order.

show_partition_name(p_parent_table text, p_value text, OUT partition_table text, OUT suffix_timestamp timestamp, OUT suffix_id bigint, OUT table_exists boolean)

- Given a parent table managed by `pg_partman` (`p_parent_table`) and an appropriate value (time or id but given in text form for `p_value`), return the name of the child partition that that value would exist in.
- If using epoch time partitioning, give the timestamp value, NOT the integer epoch value (use `to_timestamp()` to convert an epoch value).
- Returns a child table name whether the child table actually exists or not
- Also returns a raw value (`suffix_timestamp` or `suffix_id`) for the partition suffix for the given child table
- Also returns a boolean value (`table_exists`) to say whether that child table actually exists

check_parent(p_exact_count boolean DEFAULT true)

- Run this function to monitor that the parent tables of the partition sets that `pg_partman` manages do not get rows inserted to them.
- Returns a row for each parent table along with the number of rows it contains. Returns zero rows if none found.
- `partition_data_time()` & `partition_data_id()` can be used to move data from these parent tables into the proper children.
- `p_exact_count` will tell the function to give back an exact count of how many rows are in each parent if any is found. This is the default if the parameter is left out. If you don't care about an exact count, you can set this to false and it will return if it finds even just a single row in any parent. This can significantly speed up the check if a lot of data ends up in a parent or there are many partitions being managed.

apply_constraints(p_parent_table text, p_child_table text DEFAULT NULL, p_job_id bigint DEFAULT NULL, p_debug BOOLEAN DEFAULT FALSE)

- Apply constraints to child tables in a given partition set for the columns that are configured (constraint names are all prefixed with "partmanconstr_").
- Note that this does not need to be called manually to maintain custom constraints. The creation of new partitions automatically manages adding constraints to old child tables.

- Columns that are to have constraints are set in the **part_config** table **constraint_cols** array column or during creation with the parameter to **create_parent()**.
- If the **pg_partman** constraints already exists on the child table, the function will cleanly skip over the ones that exist and not create duplicates.
- If the column(s) given contain all NULL values, no constraint will be made.
- If the child table parameter is given, only that child table will have constraints applied.
- If the **p_child_table** parameter is not given, constraints are placed on the last child table older than the **optimize_constraint** value. For example, if the **optimize_constraint** value is 30, then constraints will be placed on the child table that is 31 back from the current partition (as long as partition pre-creation has been kept up to date).
- If you need to apply constraints to all older child tables, use the included python script (**reapply_constraint.py**). This script has options to make constraint application easier with as little impact on performance as possible.
- The **p_job_id** parameter is optional. It's for internal use and allows job logging to be consolidated into the original job that called this function if applicable.
- The **p_debug** parameter will show you the constraint creation statement that was used.

drop_constraints(p_parent_table text, p_child_table text, p_debug boolean DEFAULT false)

- Drop constraints that have been created by **pg_partman** for the columns that are configured in *part_config*. This makes it easy to clean up constraints if old data needs to be edited and the constraints aren't allowing it.
- Will only drop constraints that begin with **partmanconstr_*** for the given child table and configured columns.
- If you need to drop constraints on all child tables, use the included python script (**reapply_constraint.py**). This script has options to make constraint removal easier with as little impact on performance as possible.
- The debug parameter will show you the constraint drop statement that was used.

reapply_privileges(p_parent_table text)

- This function is used to reapply ownership & grants on all child tables based on what the parent table has set.
- Privileges that the parent table has will be granted to all child tables and privileges that the parent does not have will be revoked (with CASCADE).
- Privileges that are checked for are SELECT, INSERT, UPDATE, DELETE, TRUNCATE, REFERENCES, & TRIGGER.
- Be aware that for large partition sets, this can be a very long running operation and is why it was made into a separate function to run independently. Only privileges that are different between the parent & child are applied, but it still has to do system catalog lookups and comparisons for every single child partition and all individual privileges on each.
- **p_parent_table** - parent table of the partition set. Must be schema qualified and match a parent table name already configured in **pg_partman**.

apply_foreign_keys(p_parent_table text, p_child_table text DEFAULT NULL, p_job_id bigint DEFAULT NULL, p_debug boolean DEFAULT false) * Applies any foreign keys that exist on a parent table in a partition set to all the child tables. * This function is automatically called whenever a new child table is created, so there is no need to manually run it unless you need to fix an existing child table. * If you need to apply this to an entire partition set, see the **reapply_foreign_keys.py** python script. This will commit after every FK creation to avoid contention. * This function can be used on any table inheritance set, not just ones managed by **pg_partman**. * The **p_job_id** parameter is optional. It's for internal use and allows job logging to be consolidated into the original job that called this function if applicable. * The **p_debug** parameter will show you the constraint creation statement that was used.

Destruction Functions

undo_partition_time(p_parent_table text, p_batch_count int DEFAULT 1, p_batch_interval interval DEFAULT NULL, p_keep_table boolean DEFAULT true, p_lock_wait numeric DEFAULT 0) RETURNS *bigint* * Undo a time-based partition set created by **pg_partman**. This function MOVES the data from existing child partitions to the parent table. * When this function is run, the trigger on the parent table & the trigger function are immediately dropped (if they still exist). This means any further writes are done to the parent. * When this function is run, the **undo_in_progress** column in the configuration table is set to true. This causes all partition creation and retention management to stop. * If you are trying to un-partition a large amount of data automatically, it is recommended to run this function with an external script and appropriate batch settings. This will help avoid transactional locks and prevent a failure from causing an extensive rollback. See **Scripts** section for an included python script that will do this for you. * By default, partitions are not DROPPED, they are UNINHERITED. This leave previous child tables as empty, independent tables. * Without setting either batch argument manually, each run of the function will move all the data from a single partition into the parent. * Once all child tables have been uninherited/dropped, the configuration data is removed from **pg_partman** automatically. * For subpartitioned tables, you must start at the lowest level parent table and undo from there then work your way up. If you attempt to undo partitioning on a subpartition set, the function will stop with a warning to let you know. * **p_parent_table** - parent table of the partition set. Must be schema qualified and match a parent table name already configured in **pg_partman**. * **p_batch_count** - an optional argument, this sets how many times to move the amount of data equal to the **p_batch_interval** argument (or default partition interval if not set) in a single run of the function. Defaults to 1. * **p_batch_interval** - an optional argument, a time interval of how much of the data to move. This

can be smaller than the partition interval, allowing for very large sized partitions to be broken up into smaller commit batches. Defaults to the configured partition interval if not given or if you give an interval larger than the partition interval. * **p_keep_table** - an optional argument, setting this to false will cause the old child table to be dropped instead of uninherited after all of its data has been moved. Note that it takes at least two batches to actually uninherit/drop a table from the set. * **p_lock_wait** - optional argument, sets how long in seconds to wait for either the table or a row to be unlocked before timing out. Default is to wait forever. * Returns the number of rows moved to the parent table. Returns zero when all child tables are empty.

undo_partition_id(p_parent_table text, p_batch_count int DEFAULT 1, p_batch_interval bigint DEFAULT NULL, p_keep_table boolean DEFAULT true, p_lock_wait numeric DEFAULT 0)RETURNS bigint

- Undo an id-based partition set created by **pg_partman**. This function MOVES the data from existing child partitions to the parent table.
- When this function is run, the trigger on the parent table & the trigger function are immediately dropped (if they still exist). This means any further writes are done to the parent.
- When this function is run, the **undo_in_progress** column in the configuration table is set to true. This causes all partition creation and retention management to stop.
- If you are trying to un-partition a large amount of data automatically, it is recommended to run this function with an external script and appropriate batch settings. This will help avoid transactional locks and prevent a failure from causing an extensive rollback. See **Scripts** section for an included python script that will do this for you.
- By default, partitions are not DROPPED, they are UNINHERITED. This leave previous child tables as empty, independent tables.
- Without setting either batch argument manually, each run of the function will move all the data from a single partition into the parent.
- Once all child tables have been uninherited/dropped, the configuration data is removed from **pg_partman** automatically.
- For subpartitioned tables, you must start at the lowest level parent table and undo from there then work your way up. If you attempt to undo partitioning on a subpartition set, the function will stop with a warning to let you know.
- **p_parent_table** - parent table of the partition set. Must be schema qualified and match a parent table name already configured in **pg_partman**.
- **p_batch_count** - an optional argument, this sets how many times to move the amount of data equal to the **p_batch_interval** argument (or default partition interval if not set) in a single run of the function. Defaults to 1.
- **p_batch_interval** - an optional argument, an integer amount representing an interval of how much of the data to move. This can be smaller than the partition interval, allowing for very large sized partitions to be broken up into smaller commit batches. Defaults to the configured partition interval if not given or if you give an interval larger than the partition interval.
- **p_keep_table** - an optional argument, setting this to false will cause the old child table to be dropped instead of uninherited after all of it's data has been moved. Note that it takes at least two batches to actually uninherit/drop a table from the set (second batch sees it has no more data and drops it).
- **p_lock_wait** - optional argument, sets how long in seconds to wait for either the table or a row to be unlocked before timing out. Default is to wait forever.
- Returns the number of rows moved to the parent table. Returns zero when all child tables are empty.

undo_partition(p_parent_table text, p_batch_count int DEFAULT 1, p_keep_table boolean DEFAULT true, p_jobmon boolean DEFAULT true, p_lock_wait numeric DEFAULT 0)RETURNS bigint

- Undo the parent/child table inheritance of any partition set, not just ones managed by **pg_partman**. This function COPIES the data from existing child partitions to the parent table.
 - WARNING: If used on a sub-partitioned set not managed by **pg_partman**, results could be unpredictable. It is not recommended to do so.
- If you need to keep the data in your child tables after it is put into the parent, use this function.
- Unlike the other undo functions, data cannot be copied in batches smaller than the partition interval. Every run of the function copies an entire partition to the parent.
- When this function is run, the **undo_in_progress** column in the configuration table is set to true if it was managed by **pg_partman**. This causes all partition creation and retention management to stop ONLY if it was managed by **pg_partman**.
- If you are trying to un-partition a large amount of data automatically, it is recommended to run this function with an external script and appropriate batch settings. This will help avoid transactional locks and prevent a failure from causing an extensive rollback. See **Scripts** section for an included python script that will do this for you.
- By default, partitions are not DROPPED, they are UNINHERITED. This leave previous child tables exactly as they were but no longer inherited from the parent. Does not work on multiple levels of inheritance (subpartitions) if dropping tables.
- **p_parent_table** - parent table of the partition set. Must be schema qualified but does NOT have to be managed by **pg_partman**.
- **p_batch_count** - an optional argument, this sets how many partitions to copy data from in a single run. Defaults to 1.
- **p_keep_table** - an optional argument, setting this to false will cause the old child table to be dropped instead of uninherited.
- **p_jobmon** - an optional paramter to stop **undo_partition()** from using the **pg_jobmon** extension to log what it does. Defaults to true if not set.

- `p_lock_wait` - optional argument, sets how long in seconds to wait for either the table or a row to be unlocked before timing out. Default is to wait forever.
- Returns the number of rows moved to the parent table. Returns zero when child tables are all empty.

drop_partition_time(p_parent_table text, p_retention interval DEFAULT NULL, p_keep_table boolean DEFAULT NULL, p_keep_index boolean DEFAULT NULL, p_retention_schema text DEFAULT NULL)RETURNS int

- This function is used to drop child tables from a time-based partition set. By default, the table is just uninherited and not actually dropped. For automatically dropping old tables, it is recommended to use the `run_maintenance()` function with retention configured instead of calling this directly.
- `p_parent_table` - the existing parent table of a time-based partition set. MUST be schema qualified, even if in public schema.
- `p_retention` - optional parameter to give a retention time interval and immediately drop tables containing only data older than the given interval. If you have a retention value set in the config table already, the function will use that, otherwise this will override it. If not, this parameter is required. See the **About** section above for more information on retention settings.
- `p_keep_table` - optional parameter to tell partman whether to keep or drop the table in addition to uninheriting it. TRUE means the table will not actually be dropped; FALSE means the table will be dropped. This function will just use the value configured in `part_config` if not explicitly set. This option is ignored if `retention_schema` is set.
- `p_keep_index` - optional parameter to tell partman whether to keep or drop the indexes of the child table when it is uninherited. TRUE means the indexes will be kept; FALSE means all indexes will be dropped. This function will just use the value configured in `part_config` if not explicitly set. This option is ignored if `p_keep_table` is set to FALSE or if `retention_schema` is set.
- `p_retention_schema` - optional parameter to tell partman to move a table to another schema instead of dropping it. Set this to the schema you want the table moved to. This function will just use the value configured in `part_config` if not explicitly set. If this option is set, the retention `p_keep_table` & `p_keep_index` parameters are ignored.
- Returns the number of partitions affected.

drop_partition_id(p_parent_table text, p_retention bigint DEFAULT NULL, p_keep_table boolean DEFAULT NULL, p_keep_index boolean DEFAULT NULL, p_retention_schema text DEFAULT NULL)RETURNS int

- This function is used to drop child tables from an id-based partition set. By default, the table just uninherited and not actually dropped. For automatically dropping old tables, it is recommended to use the `run_maintenance()` function with retention configured instead of calling this directly.
- `p_parent_table` - the existing parent table of a time-based partition set. MUST be schema qualified, even if in public schema.
- `p_retention` - optional parameter to give a retention integer interval and immediately drop tables containing only data less than the current maximum id value minus the given retention value. If you have a retention value set in the config table already, the function will use that, otherwise this will override it. If not, this parameter is required. See the **About** section above for more information on retention settings.
- `p_keep_table` - optional parameter to tell partman whether to keep or drop the table in addition to uninheriting it. TRUE means the table will not actually be dropped; FALSE means the table will be dropped. This function will just use the value configured in `part_config` if not explicitly set. This option is ignored if `retention_schema` is set.
- `p_keep_index` - optional parameter to tell partman whether to keep or drop the indexes of the child table when it is uninherited. TRUE means the indexes will be kept; FALSE means all indexes will be dropped. This function will just use the value configured in `part_config` if not explicitly set. This option is ignored if `p_keep_table` is set to FALSE or if `retention_schema` is set.
- `p_retention_schema` - optional parameter to tell partman to move a table to another schema instead of dropping it. Set this to the schema you want the table moved to. This function will just use the value configured in `part_config` if not explicitly set. If this option is set, the retention `p_keep_table` & `p_keep_index` parameters are ignored.
- Returns the number of partitions affected.

drop_partition_column(p_parent_table text, p_column text)RETURNS void

- Depending on when a column was added (before or after partitioning was set up), dropping it on the parent may or may not drop it from all children. This function is used to ensure a column is always dropped from the parent and all children in a partition set.
- Uses the IF EXISTS clause in all drop statements, so it may spit out notices/warnings that a column was not found. You can safely ignore these warnings. It should not spit out any errors.

Tables

part_config

Stores all configuration data for partition sets managed by the extension. The only columns in this table that should ever need to be manually changed are:

1. `retention`, `retention_schema`, `retention_keep_table` & `retention_keep_index` to configure the partition set's retention policy

2. `constraint_cols` to have partman manage additional constraints and `optimize_constraint` to control when they're added
3. `premake`, `optimize_trigger`, `inherit_fk`, `use_run_maintenance` & `jobmon` to change the default behavior.

The rest are managed by the extension itself and should not be changed unless absolutely necessary.

- `parent_table`
 - Parent table of the partition set
- `control`
 - Column used as the control for partition constraints. Must be a time or integer based column.
- `partition_type`
 - Type of partitioning. Must be one of the types mentioned above in the `create_parent()` info.
- `partition_interval`
 - Text type value that determines the interval for each partition.
 - Must be a value that can either be cast to the interval or bigint data types.
- `constraint_cols`
 - Array column that lists columns to have additional constraints applied. See **About** section for more information on how this feature works.
- `premake`
 - How many partitions to keep pre-made ahead of the current partition. Default is 4.
- `optimize_trigger`
 - Manages number of partitions which are handled most efficiently by trigger. See `create_parent()` function for more info. Default 4.
- `optimize_constraint`
 - Manages which old tables get additional constraints set if configured to do so. See **About** section for more info. Default 30.
- `epoch`
 - Flag the table to be partitioned by time by an integer epoch value instead of a timestamp. See `create_parent()` function for more info. Default 'none'.
- `inherit_fk`
 - Set whether `pg_partman` manages inheriting foreign keys from the parent table to all children.
 - Defaults to TRUE. Can be set with the `create_parent()` function at creation time as well.
- `retention`
 - Text type value that determines how old the data in a child partition can be before it is dropped.
 - Must be a value that can either be cast to the interval (for time-based partitioning) or bigint (for serial partitioning) data types.
 - Leave this column NULL (the default) to always keep all child partitions. See **About** section for more info.
- `retention_schema`
 - Schema to move tables to as part of the retentions system instead of dropping them. Overrides `retention_keep_*` options.
- `retention_keep_table`
 - Boolean value to determine whether dropped child tables are kept or actually dropped.
 - Default is TRUE to keep the table and only uninherit it. Set to FALSE to have the child tables removed from the database completely.
- `retention_keep_index`
 - Boolean value to determine whether indexes are dropped for child tables that are uninherited.
 - Default is TRUE. Set to FALSE to have the child table's indexes dropped when it is uninherited.
- `infinite_time_partitions`
 - By default, new partitions in a time-based set will not be created if new data is not inserted to keep an infinite amount of empty tables from being created.
 - If you'd still like new partitions to be made despite there being no new data, set this to TRUE.
 - Defaults to FALSE.
- `datetime_string`

- For time-based partitioning, this is the datetime format string used when naming child partitions.
- **use_run_maintenance**
 - Boolean value that tells `run_maintenance()` function whether it should manage new child table creation automatically when `run_maintenance()` is called without a table parameter.
 - If `run_maintenance()` is given a table parameter, this option is ignored and maintenance will always run.
 - Defaults to TRUE for time-based partitioning.
 - Defaults to FALSE for single-level serial-based partitioning and can be changed to TRUE if desired.
 - If changing an existing serial partitioned set from FALSE to TRUE, you must run `create_function_id('parent_schema.parent_table')` to change the trigger function so it no longer creates new partitions.
 - Defaults to TRUE for all sub-partition tables
- **jobmon**
 - Boolean value to determine whether the `pg_jobmon` extension is used to log/monitor partition maintenance. Defaults to true.
- **sub_partition_set_full**
 - Boolean value to denote that the final partition for a sub-partition set has been created. Allows `run_maintenance()` to run more efficiently when there are large numbers of subpartition sets.
- **undo_in_progress**
 - Set by the `undo_partition` functions whenever they are run. If true, this causes all partition creation and retention management by the `run_maintenance()` function to stop. Default is false.
- **trigger_exception_handling**
 - Boolean value that can be set to allow the partitioning trigger function to handle any exceptions encountered while writing to this table. Handling it in this case means putting the data into the parent table to try and ensure no data loss in case of errors. Be aware that catching the exception here will override any other exception handling that may be done when writing to this partitioned set (Ex. handling a unique constraint violation to ignore it). Just the existence of this exception block will also increase xid consumption since every row inserted will increment the global xid value. If this is table has a high insert rate, you can quickly reach xid wraparound, so use this carefully. This option is set to false by default to avoid causing unexpected behavior in other exception handling situations.
- **p_upsert**
 - text value of the ON CONFLICT clause to include in the partition trigger Defaults to '' (empty string) which means it's inactive. See `create_parent()` function definition & *About* section for more info.
- **trigger_return_null**
 - Boolean value that allows controlling the behavior of the partition trigger RETURN. By default this is true and the trigger returns NULL to prevent data going into the parent table as well as the children. However, if you have multiple triggers and are relying on the return to be the NEW column value, this can cause a problem. Setting this config value to false will cause the partition trigger to RETURN NEW. You are then responsible for handling the return value in another trigger appropriately. Otherwise, this will cause new data to go into both the child and parent table of the partition set.

part_config_sub

- Stores all configuration data for sub-partitioned sets managed by `pg_partman`.
- The **sub_parent** column is the parent table of the subpartition set and all other columns govern how that parent's children are subpartitioned.
- All columns except **sub_parent** work the same exact way as their counterparts in the **part_config** table.

Scripts

If the extension was installed using `make`, the below script files should have been installed to the PostgreSQL binary directory.

partition_data.py

- A python script to make partitioning in committed batches easier.
- Calls either `partition_data_time()` or `partition_data_id()` depending on the value given for `-type`.
- A commit is done at the end of each `-interval` and/or fully created partition.
- Returns the total number of rows moved to partitions. Automatically stops when parent is empty.

- To help avoid heavy load and contention during partitioning, autovacuum is turned off for the parent table and all child tables when this script is run. When partitioning is complete, autovacuum is set back to its default value and the parent table is vacuumed when it is emptied.
- `--parent (-p)`: Parent table of an already created partition set. Required.
- `--type (-t)`: Type of partitioning. Valid values are “time” and “id”. Required.
- `--connection (-c)`: Connection string for use by psycopg. Defaults to “host=” (local socket).
- `--interval (-i)`: Value that is passed on to the partitioning function as `p_batch_interval` argument. Use this to set an interval smaller than the partition interval to commit data in smaller batches. Defaults to the partition interval if not given.
- `--batch (-b)`: How many times to loop through the value given for `-interval`. If `-interval` not set, will use default partition interval and make at most `-b` partition(s). Script commits at the end of each individual batch. (NOT passed as `p_batch_count` to partitioning function). If not set, all data in the parent table will be partitioned in a single run of the script.
- `--wait (-w)`: Cause the script to pause for a given number of seconds between commits (batches).
- `--order (-o)`: Allows you to specify the order that data is migrated from the parent to the children, either ascending (ASC) or descending (DESC). Default is ASC.
- `--lockwait (-l)`: Have a lock timeout of this many seconds on the data move. If a lock is not obtained, that batch will be tried again.
- `--lockwait_tries`: Number of times to allow a lockwait to time out before giving up on the partitioning. Defaults to 10.
- `--autovacuum_on`: Turning autovacuum off requires a brief lock to ALTER the table property. Set this option to leave autovacuum on and avoid the lock attempt.
- `--quiet (-q)`: Switch setting to stop all output during and after partitioning.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.
- `--debug` Show additional debugging output
- Examples:

```
Partition all data in a parent table. Commit after each partition is made.
    python partition_data.py -c "host=localhost dbname=mydb" -p schema.parent_table -t time
Partition by id in smaller intervals and pause between them for 5 seconds (assume >100 partition
interval)
    python partition_data.py -p schema.parent_table -t id -i 100 -w 5
Partition by time in smaller intervals for at most 10 partitions in a single run (assume monthly
partition interval)
    python partition_data.py -p schema.parent_table -t time -i "1 week" -b 10
```

undo_partition.py

- A python script to make undoing partitions in committed batches easier.
- Can also work on any parent/child partition set not managed by `pg_partman` if `-type` option is not set.
- This script calls either `undo_partition()`, `undo_partition_time()` or `undo_partition_id` depending on the value given for `-type`.
- A commit is done at the end of each `-interval` and/or emptied partition.
- Returns the total number of rows put into the to parent. Automatically stops when last child table is empty.
- `--parent (-p)`: Parent table of the partition set. Required.
- `--type (-t)`: Type of partitioning. Valid values are “time” and “id”. Not setting this argument will use `undo_partition()` and work on any parent/child table set.
- `--connection (-c)`: Connection string for use by psycopg. Defaults to “host=” (local socket).
- `--interval (-i)`: Value that is passed on to the partitioning function as `p_batch_interval`. Use this to set an interval smaller than the partition interval to commit data in smaller batches. Defaults to the partition interval if not given.
- `--batch (-b)`: How many times to loop through the value given for `-interval`. If `-interval` not set, will use default partition interval and undo at most `-b` partition(s). Script commits at the end of each individual batch. (NOT passed as `p_batch_count` to undo function). If not set, all data will be moved to the parent table in a single run of the script.
- `--wait (-w)`: Cause the script to pause for a given number of seconds between commits (batches).
- `--droptable (-d)`: Switch setting for whether to drop child tables when they are empty. Leave off option to just uninherit.
- `--quiet (-q)`: Switch setting to stop all output during and after partitioning undo.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.

- `--debug`: Show additional debugging output

dump_partition.py

- A python script to dump out tables contained in the given schema. Uses `pg_dump`, creates a SHA-512 hash file of the dump file, and then drops the table.
- When combined with the `retention_schema` configuration option, provides a way to reliably dump out tables that would normally just be dropped by the retention system.
- Tables are not dropped if `pg_dump` does not return successfully.
- The connection options for `psycopg` and `pg_dump` were separated out due to distinct differences in their requirements depending on your database connection configuration.
- All `dump_*` option defaults are the same as they would be for `pg_dump` if they are not given.
- Will work on any given schema, not just the one used to manage `pg_partman` retention.
- `--schema (-n)`: The schema that contains the tables that will be dumped. (Required).
- `--connection (-c)`: Connection string for use by `psycopg`. Role used must be able to select from `pg_catalog.pg_tables` in the relevant database and drop all tables in the given schema. Defaults to “host=” (local socket). Note this is distinct from the parameters sent to `pg_dump`.
- `--output (-o)`: Path to dump file output location. Default is where the script is run from.
- `--dump_database (-d)`: Used for `pg_dump`, same as its `-dbname` option or final database name parameter.
- `--dump_host`: Used for `pg_dump`, same as its `-host` option.
- `--dump_username`: Used for `pg_dump`, same as its `-username` option.
- `--dump_port`: Used for `pg_dump`, same as its `-port` option.
- `--pg_dump_path`: Path to `pg_dump` binary location. Must set if not in current `PATH`.
- `--Fp`: Dump using `pg_dump` plain text format. Default is binary custom (`-Fc`).
- `--nohashfile`: Do NOT create a separate file with the SHA-512 hash of the dump. If dump files are very large, hash generation can possibly take a long time.
- `--nodrop`: Do NOT drop the tables from the given schema after dumping/hashing.
- `--verbose (-v)`: Provide more verbose output.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.

vacuum_maintenance.py

- A python script to perform additional `VACUUM` maintenance on a given partition set. The main purpose of this is to provide an easier means of freezing tuples in older partitions that are no longer written to. This allows `autovacuum` to skip over them safely without causing transaction id wraparound issues. See the PostgreSQL documentation for more information on this maintenance issue: <http://www.postgresql.org/docs/current/static/routine-vacuuming.html#VACUUM-FOR-WRAPAROUND>.
- Vacuums all child tables in a given partition set who’s `age(relfrozenxid)` is greater than `vacuum_freeze_min_age`, including the parent table.
- Highly recommend scheduled runs of this script with the `-freeze` option if you have child tables that never have writes after a certain period of time.
- `-parent (-p)`: Parent table of an already created partition set. (Required)
- `-connection (-c)`: Connection string for use by `psycopg`. Defaults to “host=” (local socket).
- `-freeze (-z)`: Sets the `FREEZE` option to the `VACUUM` command.
- `-full (-f)`: Sets the `FULL` option to the `VACUUM` command. Note that `-freeze` is not necessary if you set this. Recommend reviewing `-dryrun` before running this since it will lock all tables it runs against, possibly including the parent.
- `-vacuum_freeze_min_age (-a)`: By default the script obtains this value from the system catalogs. By setting this, you can override the value obtained from the database. Note this does not change the value in the database, only the value this script uses.
- `-noparent`: Normally the parent table is included in the list of tables to vacuum if its `age(relfrozenxid)` is higher than `vacuum_freeze_min_age`. Set this to force exclusion of the parent table, even if it meets that criteria.
- `-dryrun`: Show what the script will do without actually running it against the database. Highly recommend reviewing this before running for the first time.
- `-quiet (-q)`: Turn off all output.
- `-debug`: Show additional debugging output.

reapply_indexes.py

- A python script for reapplying indexes on child tables in a partition set after they are changed on the parent table.
- Any indexes that currently exist on the children and match the definition on the parent will be left as is. There is an option to recreate matching as well indexes if desired, as well as the primary key.
- Indexes that do not exist on the parent will be dropped from all children.

- Commits are done after each index is dropped/created to help prevent long running transactions & locks.
- NOTE: New index names are made based off the child table name & columns used, so their naming may differ from the name given on the parent. This is done to allow the tool to account for long or duplicate index names. If an index name would be duplicated, an incremental counter is added on to the end of the index name to allow it to be created. Use the `--dryrun` option first to see what it will do and which names may cause dupes to be handled like this.
- `--parent (-p)`: Parent table of an already created partition set. Required.
- `--connection (-c)`: Connection string for use by pycpg. Defaults to "host=" (local socket).
- `--concurrent`: Create indexes with the CONCURRENTLY option. Note this does not work on primary keys when `--primary` is given.
- `--drop_concurrent`: Drop indexes concurrently when recreating them (PostgreSQL >= v9.2). Note this does not work on primary keys when `--primary` is given.
- `--recreate_all (-R)`: By default, if an index exists on a child and matches the parent, it will not be touched. Setting this option will force all child indexes to be dropped & recreated. Will obey the `--concurrent` & `--drop_concurrent` options if given. Will not recreate primary keys unless `--primary` option is also given.
- `--primary`: By default the primary key is not recreated. Set this option if that is needed. Note this will cause an exclusive lock on the child table for the duration of the recreation.
- `--jobs (-j)`: Use the python multiprocessing library to recreate indexes in parallel. Note that this is per table, not per index. Be very careful setting this option if load is a concern on your systems.
- `--wait (-w)`: Wait the given number of seconds after indexes have finished being created on a table before moving on to the next. When used with `-j`, this will set the pause between the batches of parallel jobs instead.
- `--dryrun`: Show what the script will do without actually running it against the database. Highly recommend reviewing this before running. Note that if multiple indexes would get the same default name, the duplicated names will show in the dryrun (because the index doesn't exist in the catalog to check for it). When the real thing is run, the duplicated names will be handled as stated in the NOTE above.
- `--quiet`: Turn off all output.
- `--nonpartman` If the partition set you are running this on is not managed by `pg_partman`, set this flag otherwise this script may not work. Note that the `pg_partman` extension is still required to be installed for this to work since it uses certain internal functions. When this is set the order that the tables are reindexed is alphabetical instead of logical.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.

reapply_constraints.py * A python script for redoing constraints on child tables in a given partition set for the columns that are configured in `part_config` table. * Typical usage would be `-d` mode to drop constraints, edit the data as needed, then `-a` mode to reapply constraints. * `--parent (-p)`: Parent table of an already created partition set. (Required) * `--connection (-c)`: Connection string for use by pycpg. Defaults to "host=" (local socket). * `--drop_constraints (-d)`: Drop all constraints managed by `pg_partman`. Drops constraints on ALL child tables in the partition set. * `--add_constraints (-a)`: Apply constraints on configured columns to all child tables older than the premake value. * `--jobs (-j)`: Use the python multiprocessing library to recreate indexes in parallel. Value for `-j` is number of simultaneous jobs to run. Note that this is per table, not per index. Be very careful setting this option if load is a concern on your systems. * `--wait (-w)`: Wait the given number of seconds after a table has had its constraints dropped or applied before moving on to the next. When used with `-j`, this will set the pause between the batches of parallel jobs instead. * `--dryrun`: Show what the script will do without actually running it against the database. Highly recommend reviewing this before running. * `--quiet (-q)`: Turn off all output. * `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.

reapply_foreign_keys.py

- A python script for redoing the inherited foreign keys for an entire partition set.
- All existing foreign keys on all child tables are dropped and the foreign keys that exist on the parent at the time this is run will be applied to all children.
- Commits after each foreign key is created to avoid long periods of contention.
- `--parent (-p)`: Parent table of an already created partition set. (Required)
- `--connection (-c)`: Connection string for use by pycpg. Defaults to "host=" (local socket).
- `--quiet (-q)`: Switch setting to stop all output during and after partitioning undo.
- `--dryrun`: Show what the script will do without actually running it against the database. Highly recommend reviewing this before running.
- `--nonpartman` If the partition set you are running this on is not managed by `pg_partman`, set this flag. Otherwise internal `pg_partman` functions are used and this script may not work. When this is set the order that the tables are rekeyed is alphabetical instead of logical.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.
- `--debug`: Show additional debugging output

check_unique_constraints.py

- Partitioning using inheritance has the shortcoming of not allowing a unique constraint to apply to all tables in the entire partition set without causing large performance issues once the partition set begins to grow very large. This script is used to check that all rows in a partition set are unique for the given columns.
- Note that on very large partition sets this can be an expensive operation to run that can consume a large chunk of storage space. The amount of storage space required is enough to dump out the entire index's column data as a plaintext file.
- If there is a column value that violates the unique constraint, this script will return those column values along with a count of how many of each value there are. Output can also be simplified to a single, total integer value to make it easier to use with monitoring applications.
- `--parent (-p)`: Parent table of the partition set to be checked. (Required)
- `--column_list (-l)`: Comma separated list of columns that make up the unique constraint to be checked. (Required)
- `--connection (-c)`: Connection string for use by pycpg. Defaults to "host=" (local socket).
- `--temp (-t)`: Path to a writable folder that can be used for temp working files. Defaults system temp folder.
- `--psql`: Full path to psql binary if not in current PATH.
- `--simple`: Output a single integer value with the total duplicate count. Use this for monitoring software that requires a simple value to be checked for.
- `--quiet (-q)`: Suppress all output unless there is a constraint violation found.
- `--version`: Print out the minimum version of `pg_partman` this script is meant to work with. The version of `pg_partman` installed may be greater than this.

This HowTo guide will show you some examples of how to set up both simple, single level partitioning as well as multi-level sub-partitioning. It will also show you how to partition data out of a table that has existing data (see **Sub-partition ID->ID->ID**) and undo the partitioning of an existing partition set. For more details on what each function does and the additional features in this extension, please see the `pg_partman.md` documentation file.

Simple Time Based: 1 Partition Per Day

```
keith=# \d partman_test.time_taptest_table
          Table "partman_test.time_taptest_table"
  Column |          Type          |          Modifiers
-----+-----+-----
 col1   | integer                | not null
 col2   | text                    |
 col3   | timestamp with time zone | not null default now()
Indexes:
    "time_taptest_table_pkey" PRIMARY KEY, btree (col1)

keith=# SELECT create_parent('partman_test.time_taptest_table', 'col3', 'time', 'daily');
 create_parent
-----
 t
(1 row)

keith=# \d+ partman_test.time_taptest_table
          Table "partman_test.time_taptest_table"
  Column |          Type          |          Modifiers          | Storage | Stats target |
-----+-----+-----+-----+-----+
  Description
 col1   | integer                | not null                    | plain   |               |
 col2   | text                    |                              | extended |               |
 col3   | timestamp with time zone | not null default now()     | plain   |               |
Indexes:
    "time_taptest_table_pkey" PRIMARY KEY, btree (col1)
Triggers:
    time_taptest_table_part_trig BEFORE INSERT ON partman_test.time_taptest_table FOR EACH ROW
    EXECUTE PROCEDURE partman_test.time_taptest_table_part_trig_func()
Child tables: partman_test.time_taptest_table_p2015_05_09,
               partman_test.time_taptest_table_p2015_05_10,
               partman_test.time_taptest_table_p2015_05_11,
               partman_test.time_taptest_table_p2015_05_12,
               partman_test.time_taptest_table_p2015_05_13,
               partman_test.time_taptest_table_p2015_05_14,
               partman_test.time_taptest_table_p2015_05_15,
               partman_test.time_taptest_table_p2015_05_16,
               partman_test.time_taptest_table_p2015_05_17
```

The trigger function most efficiently covers a specific period of time for 4 days before and 4 days after today. This can be adjusted with the `optimize_trigger` config option in the `part_config` table. Outside of that, a dynamic statement tries to find the appropriate child table to put the data into. Note this dynamic statement is far less efficient since a catalog lookup is required and the statement plan cannot be cached as well as looking up the that the child table exists. If the child table does not exist at all for the time value given, the data goes to the parent:

```
keith=# \sf partman_test.time_taptest_table_part_trig_func
CREATE OR REPLACE FUNCTION partman_test.time_taptest_table_part_trig_func()
  RETURNS trigger
  LANGUAGE plpgsql
AS $function$
    DECLARE
        v_count          int;
        v_partition_name text;
        v_partition_timestamp timestampz;
    BEGIN
    IF TG_OP = 'INSERT' THEN
        v_partition_timestamp := date_trunc('day', NEW.col3);
        IF NEW.col3 >= '2015-05-13 00:00:00-04' AND NEW.col3 < '2015-05-14 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_13 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-12 00:00:00-04' AND NEW.col3 < '2015-05-13 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_12 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-14 00:00:00-04' AND NEW.col3 < '2015-05-15 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_14 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-11 00:00:00-04' AND NEW.col3 < '2015-05-12 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_11 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-15 00:00:00-04' AND NEW.col3 < '2015-05-16 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_15 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-10 00:00:00-04' AND NEW.col3 < '2015-05-11 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_10 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-16 00:00:00-04' AND NEW.col3 < '2015-05-17 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_16 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-09 00:00:00-04' AND NEW.col3 < '2015-05-10 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_09 VALUES (NEW.*);
        ELSIF NEW.col3 >= '2015-05-17 00:00:00-04' AND NEW.col3 < '2015-05-18 00:00:00-04' THEN
            INSERT INTO partman_test.time_taptest_table_p2015_05_17 VALUES (NEW.*);
        ELSE
            v_partition_name := partman.check_name_length('time_taptest_table',
                'partman_test', to_char(v_partition_timestamp, 'YYYY-MM-DD'), TRUE);
            SELECT count(*) INTO v_count FROM pg_tables WHERE schemaname || '.' || tablename =
                v_partition_name;
            IF v_count > 0 THEN
                EXECUTE 'INSERT INTO ' || v_partition_name || ' VALUES($1.*)' USING NEW;
            ELSE
                RETURN NEW;
            END IF;
        END IF;
    END IF;
    RETURN NULL;
END $function$
```

Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table

```
keith=# \d partman_test.id_taptest_table
Table "partman_test.id_taptest_table"
Column | Type | Modifiers
-----+-----+-----
col1 | integer | not null
col2 | text | not null default 'stuff'::text
col3 | timestamp with time zone | default now()
Indexes:
    "id_taptest_table_pkey" PRIMARY KEY, btree (col1)

keith=# SELECT create_parent('partman_test.id_taptest_table', 'col1', 'id', '10');
create_parent
```



```

v_current_partition_id := NEW.col1 - (NEW.col1 % 10);
IF (NEW.col1 % 10) > (10 / 2) THEN
    v_id_position := (length(v_last_partition) - position('p_' in
        reverse(v_last_partition))) + 2;
    v_next_partition_id := (substring(v_last_partition from v_id_position)::bigint) + 10;
    WHILE ((v_next_partition_id - v_current_partition_id) / 10) <= 4 LOOP
        v_partition_created :=
            partman.create_partition_id('partman_test.id_taptest_table',
                ARRAY[v_next_partition_id]);
        IF v_partition_created THEN
            PERFORM partman.create_function_id('partman_test.id_taptest_table');
            PERFORM partman.apply_constraints('partman_test.id_taptest_table');
        END IF;
        v_next_partition_id := v_next_partition_id + 10;
    END LOOP;
END IF;
END IF;
RETURN NULL;
END $function$

```

Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table and using upsert to drop conflicting rows

Uses same example table as above

```

keith=# SELECT create_parent('partman_test.id_taptest_table', 'col1', 'id', '10', p_upsert :=
'ON CONFLICT (col1) DO NOTHING');
 create_parent
-----

```

```

t
(1 row)

```

```

keith=# \d+ partman_test.id_taptest_table

```

```

Table "partman_test.id_taptest_table"
Column |          Type          |          Modifiers          | Storage | Stats target
-----+-----+-----+-----+-----+
col1   | integer                | not null                    | plain  |
col2   | text                    | not null default 'stuff'::text | extended |
col3   | timestamp with time zone | default now()                | plain  |

```

Indexes:

```

" id_taptest_table_pkey" PRIMARY KEY, btree (col1)

```

Triggers:

```

 id_taptest_table_part_trig BEFORE INSERT ON partman_test.id_taptest_table FOR EACH ROW
 EXECUTE PROCEDURE partman_test.id_taptest_table_part_trig_func()

```

```

Child tables: partman_test.id_taptest_table_p0,
               partman_test.id_taptest_table_p10,
               partman_test.id_taptest_table_p20,
               partman_test.id_taptest_table_p30,
               partman_test.id_taptest_table_p40

```

Other than the new ON CONFLICT clause, this trigger function works exactly the same as the previous ID example.

```

CREATE OR REPLACE FUNCTION partman_test.id_taptest_table_part_trig_func()
RETURNS trigger
LANGUAGE plpgsql
AS $function$
DECLARE
    v_count          int;
    v_current_partition_id  bigint;
    v_current_partition_name text;
    v_id_position    int;
    v_last_partition text := 'partman_test.id_taptest_table_p40';
    v_next_partition_id  bigint;
    v_next_partition_name text;
    v_partition_created boolean;

```

```

BEGIN
IF TG_OP = 'INSERT' THEN
  IF NEW.col1 >= 0 AND NEW.col1 < 10 THEN
    INSERT INTO partman_test.id_taptest_table_p0 VALUES (NEW.*) ON CONFLICT (col1) DO
      NOTHING;
  ELSIF NEW.col1 >= 10 AND NEW.col1 < 20 THEN
    INSERT INTO partman_test.id_taptest_table_p10 VALUES (NEW.*) ON CONFLICT (col1) DO
      NOTHING;
  ELSIF NEW.col1 >= 20 AND NEW.col1 < 30 THEN
    INSERT INTO partman_test.id_taptest_table_p20 VALUES (NEW.*) ON CONFLICT (col1) DO
      NOTHING;
  ELSIF NEW.col1 >= 30 AND NEW.col1 < 40 THEN
    INSERT INTO partman_test.id_taptest_table_p30 VALUES (NEW.*) ON CONFLICT (col1) DO
      NOTHING;
  ELSIF NEW.col1 >= 40 AND NEW.col1 < 50 THEN
    INSERT INTO partman_test.id_taptest_table_p40 VALUES (NEW.*) ON CONFLICT (col1) DO
      NOTHING;
  ELSE
    v_current_partition_id := NEW.col1 - (NEW.col1 % 10);
    v_current_partition_name := partman.check_name_length('id_taptest_table',
      'partman_test', v_current_partition_id::text, TRUE);
    SELECT count(*) INTO v_count FROM pg_tables WHERE schemaname || '.' || tablename =
      v_current_partition_name;
    IF v_count > 0 THEN
      EXECUTE 'INSERT INTO ' || v_current_partition_name || ' VALUES($1.*) ON CONFLICT
        (col1) DO NOTHING' USING NEW;
    ELSE
      RETURN NEW;
    END IF;
  END IF;
END IF;
v_current_partition_id := NEW.col1 - (NEW.col1 % 10);
IF (NEW.col1 % 10) > (10 / 2) THEN
  v_id_position := (length(v_last_partition) - position('p_' in
    reverse(v_last_partition))) + 2;
  v_next_partition_id := (substring(v_last_partition from v_id_position)::bigint) + 10;
  WHILE ((v_next_partition_id - v_current_partition_id) / 10) <= 4 LOOP
    v_partition_created :=
      partman.create_partition_id('partman_test.id_taptest_table',
        ARRAY[v_next_partition_id]);
    IF v_partition_created THEN
      PERFORM partman.create_function_id('partman_test.id_taptest_table');
      PERFORM partman.apply_constraints('partman_test.id_taptest_table');
    END IF;
    v_next_partition_id := v_next_partition_id + 10;
  END LOOP;
END IF;
END IF;
RETURN NULL;
END $function$

```

Running the following query will insert a row in the table

```

keith@keith=# INSERT INTO partman_test.id_taptest_table(col1,col2) VALUES(1,'insert1');
INSERT 0 0
Time: 1.668 ms
keith@keith=# select * from partman_test.id_taptest_table;
 col1 | col2 | col3
-----+-----+-----
    1 | insert1 | 2016-07-12 15:03:47.396159-04
(1 row)
Time: 1.708 ms

```

Running the following query will not fail but the row in the table will not change and col2 will still be 'insert1'

```

keith@keith=# INSERT INTO partman_test.id_taptest_table(col1,col2) VALUES(1,'insert2');

```

```

INSERT 0 0
Time: 1.995 ms
keith@keith=# select * from partman_test.id_taptest_table;
 col1 |  col2  |          col3
-----+-----+-----
    1 | insert1 | 2016-07-12 15:03:47.396159-04
(1 row)

Time: 1.721 ms

```

Simple Serial ID: 1 Partition Per 10 ID Values Starting With Empty Table and using upsert to update conflicting rows

Uses same example table as above

```

keith@keith=# SELECT create_parent('partman_test.id_taptest_table', 'col1', 'id', '10',
    p_upsert := 'ON CONFLICT (col1) DO UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3');
 create_parent
-----

```

```

t
(1 row)

```

```

keith=# \d+ partman_test.id_taptest_table

```

```

Table "partman_test.id_taptest_table"
Column |          Type          | Modifiers | Storage | Stats target
-----+-----+-----+-----+-----+
 col1  | integer                | not null  | plain   |
 col2  | text                   | not null default 'stuff'::text | extended |
 col3  | timestamp with time zone | default now() | plain   |

```

```

Indexes:
    "id_taptest_table_pkey" PRIMARY KEY, btree (col1)

```

```

Triggers:
    id_taptest_table_part_trig BEFORE INSERT ON partman_test.id_taptest_table FOR EACH ROW
    EXECUTE PROCEDURE partman_test.id_taptest_table_part_trig_func()

```

```

Child tables: partman_test.id_taptest_table_p0,
               partman_test.id_taptest_table_p10,
               partman_test.id_taptest_table_p20,
               partman_test.id_taptest_table_p30,
               partman_test.id_taptest_table_p40

```

Other than the new ON CONFLICT clause, this trigger function works exactly the same as the previous ID example.

```

keith@keith=# \sf partman_test.id_taptest_table_part_trig_func
CREATE OR REPLACE FUNCTION partman_test.id_taptest_table_part_trig_func()
RETURNS trigger
LANGUAGE plpgsql
AS $function$
DECLARE
    v_count          int;
    v_current_partition_id    bigint;
    v_current_partition_name  text;
    v_id_position      int;
    v_last_partition   text := 'id_taptest_table_p40';
    v_next_partition_id    bigint;
    v_next_partition_name  text;
    v_partition_created  boolean;
BEGIN
IF TG_OP = 'INSERT' THEN
IF NEW.col1 >= 0 AND NEW.col1 < 10 THEN
INSERT INTO partman_test.id_taptest_table_p0 VALUES (NEW.*) ON CONFLICT (col1) DO
UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3;
ELSIF NEW.col1 >= 10 AND NEW.col1 < 20 THEN
INSERT INTO partman_test.id_taptest_table_p10 VALUES (NEW.*) ON CONFLICT (col1) DO
UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3;
ELSIF NEW.col1 >= 20 AND NEW.col1 < 30 THEN

```



```

INSERT INTO partman_test.id_taptest_table_p20 VALUES (NEW.*) ON CONFLICT (col1) DO
UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3;
ELSIF NEW.col1 >= 30 AND NEW.col1 < 40 THEN
INSERT INTO partman_test.id_taptest_table_p30 VALUES (NEW.*) ON CONFLICT (col1) DO
UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3;
ELSIF NEW.col1 >= 40 AND NEW.col1 < 50 THEN
INSERT INTO partman_test.id_taptest_table_p40 VALUES (NEW.*) ON CONFLICT (col1) DO
UPDATE SET col2=EXCLUDED.col2, col3=EXCLUDED.col3;
ELSE
v_current_partition_id := NEW.col1 - (NEW.col1 % 10);
v_current_partition_name := partman.check_name_length('id_taptest_table',
v_current_partition_id::text, TRUE);
SELECT count(*) INTO v_count FROM pg_catalog.pg_tables WHERE schemaname =
'partman_test'::name AND tablename = v_current_partition_name::name;
IF v_count > 0 THEN
EXECUTE format('INSERT INTO %I.%I VALUES($1.*) ON CONFLICT (col1) DO UPDATE SET
col2=EXCLUDED.col2, col3=EXCLUDED.col3', 'partman_test',
v_current_partition_name) USING NEW;
ELSE
RETURN NEW;
END IF;
END IF;
v_current_partition_id := NEW.col1 - (NEW.col1 % 10);
IF (NEW.col1 % 10) > (10 / 2) THEN
v_id_position := (length(v_last_partition) - position('p_' in
reverse(v_last_partition))) + 2;
v_next_partition_id := (substring(v_last_partition from v_id_position)::bigint) + 10;
WHILE ((v_next_partition_id - v_current_partition_id) / 10) <= 4 LOOP
v_partition_created :=
partman.create_partition_id('partman_test.id_taptest_table',
ARRAY[v_next_partition_id]);
IF v_partition_created THEN
PERFORM partman.create_function_id('partman_test.id_taptest_table');
PERFORM partman.apply_constraints('partman_test.id_taptest_table');
END IF;
v_next_partition_id := v_next_partition_id + 10;
END LOOP;
END IF;
END IF;
RETURN NULL;
END $function$

```

Runnign the following query will insert a row in the table

```

keith@keith=# INSERT INTO partman_test.id_taptest_table(col1,col2) VALUES(1,'insert1');
INSERT 0 0
Time: 5.440 ms
keith@keith=# select * from partman_test.id_taptest_table;
 col1 | col2 | col3
-----+-----+-----
    1 | insert1 | 2016-07-12 15:09:25.138428-04
(1 row)

```

Running the following query will not fail and the row in the table will change and col2 will now be 'insert2' and the timestamp in col3 will update to the default value now()

```

keith@keith=# INSERT INTO partman_test.id_taptest_table(col1,col2) VALUES(1,'insert2');
INSERT 0 0
Time: 6.334 ms
keith@keith=# select * from partman_test.id_taptest_table;
 col1 | col2 | col3
-----+-----+-----
    1 | insert2 | 2016-07-12 15:10:04.50675-04
(1 row)

```

Sub-partition Time->Time->Time: Yearly -> Monthly -> Daily

```

keith=# \d partman_test.time_taptest_table
          Table "partman_test.time_taptest_table"
  Column |          Type          |          Modifiers
-----+-----+-----
 col1   | integer                | not null
 col2   | text                    |
 col3   | timestamp with time zone | not null default now()
Indexes:
    "time_taptest_table_pkey" PRIMARY KEY, btree (col1)

```

Create top yearly partition set that only covers 2 years forward/back

```

keith=# SELECT create_parent('partman_test.time_taptest_table', 'col3', 'time', 'yearly',
    p_premake := 2);
 create_parent
-----
 t
(1 row)

```

Now tell pg_partman to partition all yearly child tables by month. Do this by giving it the parent table of the yearly partition set (happens to be the same as above)

```

keith=# SELECT create_sub_parent('partman_test.time_taptest_table', 'col3', 'time', 'monthly',
    p_premake := 2);
 create_sub_parent
-----
 t
(1 row)

```

```

keith=# select tablename from pg_tables where schemaname = 'partman_test' order by tablename;
          tablename
-----
 time_taptest_table
 time_taptest_table_p2013
 time_taptest_table_p2013_p2013_01
 time_taptest_table_p2014
 time_taptest_table_p2014_p2014_01
 time_taptest_table_p2015
 time_taptest_table_p2015_p2015_03
 time_taptest_table_p2015_p2015_04
 time_taptest_table_p2015_p2015_05
 time_taptest_table_p2015_p2015_06
 time_taptest_table_p2015_p2015_07
 time_taptest_table_p2016
 time_taptest_table_p2016_p2016_01
 time_taptest_table_p2017
 time_taptest_table_p2017_p2017_01
(15 rows)

```

The day this tutorial was written is 2015-05-13. You now see that this causes only 2 new future partitions to be created. And for the monthly partitions, they have been created to cover 2 months ahead as well. Note that the trigger will still cover 4 ahead and 4 behind for both partition levels unless you change the `optimize_trigger` option in the config table. A parent table ALWAYS has at least one child, so for the time period that is outside of what the premake covers, just a single table has been made for the lowest possible month in that yearly time period (January). Now tell pg_partman to partition every monthly table that currently exists by day. Do this by giving it the parent table of each monthly partition set (the parent with the just the year suffix since its children are the monthly partitions).

```

SELECT create_sub_parent('partman_test.time_taptest_table_p2013', 'col3', 'time', 'daily',
    p_premake := 2);
SELECT create_sub_parent('partman_test.time_taptest_table_p2014', 'col3', 'time', 'daily',
    p_premake := 2);
SELECT create_sub_parent('partman_test.time_taptest_table_p2015', 'col3', 'time', 'daily',
    p_premake := 2);
SELECT create_sub_parent('partman_test.time_taptest_table_p2016', 'col3', 'time', 'daily',
    p_premake := 2);
SELECT create_sub_parent('partman_test.time_taptest_table_p2017', 'col3', 'time', 'daily',
    p_premake := 2);

```

```

keith=# select tablename from pg_tables where schemaname = 'partman_test' order by tablename;
          tablename
-----
time_taptest_table
time_taptest_table_p2013
time_taptest_table_p2013_p2013_01
time_taptest_table_p2013_p2013_01_p2013_01_01
time_taptest_table_p2014
time_taptest_table_p2014_p2014_01
time_taptest_table_p2014_p2014_01_p2014_01_01
time_taptest_table_p2015
time_taptest_table_p2015_p2015_03
time_taptest_table_p2015_p2015_03_p2015_03_01
time_taptest_table_p2015_p2015_04
time_taptest_table_p2015_p2015_04_p2015_04_01
time_taptest_table_p2015_p2015_05
time_taptest_table_p2015_p2015_05_p2015_05_11
time_taptest_table_p2015_p2015_05_p2015_05_12
time_taptest_table_p2015_p2015_05_p2015_05_13
time_taptest_table_p2015_p2015_05_p2015_05_14
time_taptest_table_p2015_p2015_05_p2015_05_15
time_taptest_table_p2015_p2015_06
time_taptest_table_p2015_p2015_06_p2015_06_01
time_taptest_table_p2015_p2015_07
time_taptest_table_p2015_p2015_07_p2015_07_01
time_taptest_table_p2016
time_taptest_table_p2016_p2016_01
time_taptest_table_p2016_p2016_01_p2016_01_01
time_taptest_table_p2017
time_taptest_table_p2017_p2017_01
time_taptest_table_p2017_p2017_01_p2017_01_01
(28 rows)

```

Again, assuming today's date is 2015-05-13, it has created the sub-partitions to cover 2 days in the future. All other parent tables outside of the current time period have the lowest possible day created for them.

Sub-partition ID->ID->ID: 10,000 -> 1,000 -> 100

This partition set has existing data already in it. We will partition it out using the python script found in the "bin" directory of the repo. It's possible to use the `partition_data_id()` function in postgres as well, but that would partition all the data out in a single transaction which, for a live table, could cause serious contention issues. The python script allows commits to be done in batches and avoid that contention.

```

keith=# \d partman_test.id_taptest_table
          Table "partman_test.id_taptest_table"
  Column |          Type          |          Modifiers
-----+-----+-----
 col1   | integer                | not null
 col2   | text                   | not null default 'stuff'::text
 col3   | timestamp with time zone | default now()
Indexes:

keith=# SELECT count(*) FROM partman_test.id_taptest_table ;
 count
-----
100000
(1 row)

keith=# SELECT min(col1), max(col1) FROM partman_test.id_taptest_table ;
 min | max
-----+-----
   1 | 100000
(1 row)

```

Since there is already data in the table, the child tables initially created will be based around the max value, two before it and two after it. As stated above for time, the trigger still covers for 4 partitions before & after most efficiently, so if you need to adjust that as well, see the `part_config` table.

```
keith=# SELECT create_parent('partman_test.id_taptest_table', 'col1', 'id', '10000',
    p_use_run_maintenance := true, p_jobmon := false, p_premake := 2);
 create_parent
-----
 t
(1 row)

keith=# select tablename from pg_tables where schemaname = 'partman_test' order by tablename;
      tablename
-----
 id_taptest_table
 id_taptest_table_p100000
 id_taptest_table_p110000
 id_taptest_table_p120000
 id_taptest_table_p80000
 id_taptest_table_p90000
```

However, the data still resides in the parent table at this time. To partition it out, use the python script as mentioned above. The options below will cause it to commit every 100 rows. If the interval option was not given, it would commit them at the configured interval of 10,000. Allowing a lower interval decreases the possible contention and allows the data to be more readily available in the newly created partitions:

```
$ python partition_data.py -c host=localhost -p partman_test.id_taptest_table -t id -i 100
Attempting to turn off autovacuum for partition set...
... Success!
Rows moved: 100
Rows moved: 100
...
Rows moved: 100
Rows moved: 100
Rows moved: 1
Total rows moved: 100000
Running vacuum analyze on parent table...
Attempting to reset autovacuum for old parent table and all child tables...
... Success!
```

Partitioning the data like this has also made the partitions that were needed to store the data

```
keith=# select tablename from pg_tables where schemaname = 'partman_test' order by tablename;
      tablename
-----
 id_taptest_table
 id_taptest_table_p0
 id_taptest_table_p10000
 id_taptest_table_p100000
 id_taptest_table_p110000
 id_taptest_table_p120000
 id_taptest_table_p20000
 id_taptest_table_p30000
 id_taptest_table_p40000
 id_taptest_table_p50000
 id_taptest_table_p60000
 id_taptest_table_p70000
 id_taptest_table_p80000
 id_taptest_table_p90000
(14 rows)
```

Now create the sub-partitions for 1000. As was noted above for time, we give the parent table who's children we want partitioned along with the properties to give those children:

```
keith=# SELECT create_sub_parent('partman_test.id_taptest_table', 'col1', 'id', '1000',
    p_jobmon := false, p_premake := 2);
 create_sub_parent
```

```
t
(1 row)
```

All children tables get at least their minimum sub-partition made and the sub-partitions based around the current max value are also created.

```
keith=# select tablename from pg_tables where schemaname = 'partman_test' order by tablename;
          tablename
-----
id_taptest_table
id_taptest_table_p0
id_taptest_table_p0_p0
id_taptest_table_p10000
id_taptest_table_p100000
id_taptest_table_p100000_p100000
id_taptest_table_p100000_p101000
id_taptest_table_p100000_p102000
id_taptest_table_p10000_p10000
id_taptest_table_p110000
id_taptest_table_p110000_p110000
id_taptest_table_p120000
id_taptest_table_p120000_p120000
id_taptest_table_p20000
id_taptest_table_p20000_p20000
id_taptest_table_p30000
id_taptest_table_p30000_p30000
id_taptest_table_p40000
id_taptest_table_p40000_p40000
id_taptest_table_p50000
id_taptest_table_p50000_p50000
id_taptest_table_p60000
id_taptest_table_p60000_p60000
id_taptest_table_p70000
id_taptest_table_p70000_p70000
id_taptest_table_p80000
id_taptest_table_p80000_p80000
id_taptest_table_p90000
id_taptest_table_p90000_p98000
id_taptest_table_p90000_p99000
(30 rows)
```

If you're wondering why, even with data in them, the children didn't get all their sub-partitions created, it's for the same reason that the top partition only initially had the 2 previous and 2 after created: the data still exists in the sub-partition parents. You can see this by running the monitoring function built into `pg_partman` here:

```
keith=# select * from check_parent() order by 1;
          parent_table          | count
-----+-----
partman_test.id_taptest_table_p0 | 9999
partman_test.id_taptest_table_p10000 | 10000
partman_test.id_taptest_table_p100000 | 1
partman_test.id_taptest_table_p20000 | 10000
partman_test.id_taptest_table_p30000 | 10000
partman_test.id_taptest_table_p40000 | 10000
partman_test.id_taptest_table_p50000 | 10000
partman_test.id_taptest_table_p60000 | 10000
partman_test.id_taptest_table_p70000 | 10000
partman_test.id_taptest_table_p80000 | 10000
partman_test.id_taptest_table_p90000 | 10000
(11 rows)
```

So, lets fix that:

```
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p0 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p10000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p20000 -t id -i 100
```

```
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p30000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p40000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p50000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p60000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p70000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p80000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p90000 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p100000 -t id -i
100
```

Now the monitoring function returns nothing (as should be the norm):

```
keith=# select * from check_parent() order by 1;
parent_table | count
-----+-----
(0 rows)
```

Now we also see all child partitons were created for the data that exists:

```
keith=# SELECT tablename FROM pg_tables WHERE schemaname = 'partman_test' order by tablename;
-----
id_taptest_table
id_taptest_table_p0
id_taptest_table_p0_p0
id_taptest_table_p0_p1000
id_taptest_table_p0_p2000
id_taptest_table_p0_p3000
id_taptest_table_p0_p4000
id_taptest_table_p0_p5000
id_taptest_table_p0_p6000
id_taptest_table_p0_p7000
id_taptest_table_p0_p8000
id_taptest_table_p0_p9000
id_taptest_table_p10000
id_taptest_table_p100000
id_taptest_table_p100000_p100000
id_taptest_table_p100000_p101000
id_taptest_table_p100000_p102000
id_taptest_table_p10000_p10000
id_taptest_table_p10000_p11000
id_taptest_table_p10000_p12000
id_taptest_table_p10000_p13000
id_taptest_table_p10000_p14000
id_taptest_table_p10000_p15000
id_taptest_table_p10000_p16000
id_taptest_table_p10000_p17000
id_taptest_table_p10000_p18000
id_taptest_table_p10000_p19000
id_taptest_table_p110000
id_taptest_table_p110000_p110000
id_taptest_table_p120000
id_taptest_table_p120000_p120000
id_taptest_table_p20000
id_taptest_table_p20000_p20000
id_taptest_table_p20000_p21000
id_taptest_table_p20000_p22000
id_taptest_table_p20000_p23000
id_taptest_table_p20000_p24000
id_taptest_table_p20000_p25000
id_taptest_table_p20000_p26000
id_taptest_table_p20000_p27000
id_taptest_table_p20000_p28000
id_taptest_table_p20000_p29000
id_taptest_table_p30000
id_taptest_table_p30000_p30000
```

id_taptest_table_p30000_p31000
id_taptest_table_p30000_p32000
id_taptest_table_p30000_p33000
id_taptest_table_p30000_p34000
id_taptest_table_p30000_p35000
id_taptest_table_p30000_p36000
id_taptest_table_p30000_p37000
id_taptest_table_p30000_p38000
id_taptest_table_p30000_p39000
id_taptest_table_p40000
id_taptest_table_p40000_p40000
id_taptest_table_p40000_p41000
id_taptest_table_p40000_p42000
id_taptest_table_p40000_p43000
id_taptest_table_p40000_p44000
id_taptest_table_p40000_p45000
id_taptest_table_p40000_p46000
id_taptest_table_p40000_p47000
id_taptest_table_p40000_p48000
id_taptest_table_p40000_p49000
id_taptest_table_p50000
id_taptest_table_p50000_p50000
id_taptest_table_p50000_p51000
id_taptest_table_p50000_p52000
id_taptest_table_p50000_p53000
id_taptest_table_p50000_p54000
id_taptest_table_p50000_p55000
id_taptest_table_p50000_p56000
id_taptest_table_p50000_p57000
id_taptest_table_p50000_p58000
id_taptest_table_p50000_p59000
id_taptest_table_p60000
id_taptest_table_p60000_p60000
id_taptest_table_p60000_p61000
id_taptest_table_p60000_p62000
id_taptest_table_p60000_p63000
id_taptest_table_p60000_p64000
id_taptest_table_p60000_p65000
id_taptest_table_p60000_p66000
id_taptest_table_p60000_p67000
id_taptest_table_p60000_p68000
id_taptest_table_p60000_p69000
id_taptest_table_p70000
id_taptest_table_p70000_p70000
id_taptest_table_p70000_p71000
id_taptest_table_p70000_p72000
id_taptest_table_p70000_p73000
id_taptest_table_p70000_p74000
id_taptest_table_p70000_p75000
id_taptest_table_p70000_p76000
id_taptest_table_p70000_p77000
id_taptest_table_p70000_p78000
id_taptest_table_p70000_p79000
id_taptest_table_p80000
id_taptest_table_p80000_p80000
id_taptest_table_p80000_p81000
id_taptest_table_p80000_p82000
id_taptest_table_p80000_p83000
id_taptest_table_p80000_p84000
id_taptest_table_p80000_p85000
id_taptest_table_p80000_p86000
id_taptest_table_p80000_p87000
id_taptest_table_p80000_p88000
id_taptest_table_p80000_p89000
id_taptest_table_p90000

```

id_taptest_table_p90000_p90000
id_taptest_table_p90000_p91000
id_taptest_table_p90000_p92000
id_taptest_table_p90000_p93000
id_taptest_table_p90000_p94000
id_taptest_table_p90000_p95000
id_taptest_table_p90000_p96000
id_taptest_table_p90000_p97000
id_taptest_table_p90000_p98000
id_taptest_table_p90000_p99000
(119 rows)

```

We can still take this another level deeper as well. Normally with a large amount of data, it's not recommended to partition down to an interval this low since the benefit gained is minimal compared the management of such a large number of tables. But it's being done here as an example. Just as with the time example above, we now have to sub-partition each one of the sub-parent tables to say how we want their children sub-partitioned:

```

SELECT create_sub_parent('partman_test.id_taptest_table_p0', 'col1', 'id', '100', p_jobmon :=
    false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p10000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p20000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p30000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p40000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p50000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p60000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p70000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p80000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p90000', 'col1', 'id', '100', p_jobmon
    := false, p_premake := 2);
SELECT create_sub_parent('partman_test.id_taptest_table_p100000', 'col1', 'id', '100',
    p_jobmon := false, p_premake := 2);

```

I won't show the full list here, but you can see how every child table of the above parents is now a parent table itself with the appropriate minimal child table created where needed as well as the child tables around the current max:

```

keith=# SELECT tablename FROM pg_tables WHERE schemaname = 'partman_test' order by tablename;
          tablename
-----
id_taptest_table
id_taptest_table_p0
id_taptest_table_p0_p0
id_taptest_table_p0_p0_p0
id_taptest_table_p0_p1000
id_taptest_table_p0_p1000_p1000
id_taptest_table_p0_p2000
id_taptest_table_p0_p2000_p2000
...
id_taptest_table_p10000
id_taptest_table_p100000
id_taptest_table_p100000_p100000
id_taptest_table_p100000_p100000_p100000
id_taptest_table_p100000_p100000_p100100
id_taptest_table_p100000_p100000_p100200
id_taptest_table_p100000_p101000
id_taptest_table_p100000_p101000_p101000
id_taptest_table_p100000_p102000
id_taptest_table_p100000_p102000_p102000
id_taptest_table_p10000_p10000

```



```

id_taptest_table_p10000_p10000_p10000
id_taptest_table_p10000_p11000
id_taptest_table_p10000_p11000_p11000
...
id_taptest_table_p90000_p98000
id_taptest_table_p90000_p98000_p98000
id_taptest_table_p90000_p99000
id_taptest_table_p90000_p99000_p99800
id_taptest_table_p90000_p99000_p99900
(225 rows)

```

If you ran the `check_parent()` function, you'd see that now each one of these new parent tables now needs to have its data moved. Now's a good time show a trick for generating many individual statements based on values returned from a query:

```

SELECT 'python partition_data.py -c host=localhost -p '||parent_table||' -t id -i 100' FROM
    part_config order by parent_table;

                                ?column?
-----
python partition_data.py -c host=localhost -p partman_test.id_taptest_table -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p0 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p0_p0 -t id -i 100
python partition_data.py -c host=localhost -p partman_test.id_taptest_table_p0_p1000 -t id -i
    100
...

```

This will generate the commands to partition out the data found in any parent table managed by `pg_partman`. Yes some are already empty, but that won't matter since they'll just do nothing and it makes the query to generate these commands easier. Recommend putting the output from this into an executable shell file vs just pasting it all into the shell directly. Now if you get a list of all the tables, you can see there's quite a lot now (the row count returned is the number of tables).

```

keith=# SELECT tablename FROM pg_tables WHERE schemaname = 'partman_test' order by tablename;
          tablename
-----
id_taptest_table
id_taptest_table_p0
id_taptest_table_p0_p0
id_taptest_table_p0_p0_p0
id_taptest_table_p0_p0_p100
id_taptest_table_p0_p0_p200
id_taptest_table_p0_p0_p300
id_taptest_table_p0_p0_p400
id_taptest_table_p0_p0_p500
id_taptest_table_p0_p0_p600
id_taptest_table_p0_p0_p700
id_taptest_table_p0_p0_p800
id_taptest_table_p0_p0_p900
id_taptest_table_p0_p1000
id_taptest_table_p0_p1000_p1000
id_taptest_table_p0_p1000_p1100
id_taptest_table_p0_p1000_p1200
id_taptest_table_p0_p1000_p1300
id_taptest_table_p0_p1000_p1400
id_taptest_table_p0_p1000_p1500
id_taptest_table_p0_p1000_p1600
id_taptest_table_p0_p1000_p1700
id_taptest_table_p0_p1000_p1800
id_taptest_table_p0_p1000_p1900
id_taptest_table_p0_p2000
id_taptest_table_p0_p2000_p2000
id_taptest_table_p0_p2000_p2100
...
id_taptest_table_p90000_p98000_p98800
id_taptest_table_p90000_p98000_p98900
id_taptest_table_p90000_p99000
id_taptest_table_p90000_p99000_p99000

```

```
id_taptest_table_p90000_p99000_p99100
id_taptest_table_p90000_p99000_p99200
id_taptest_table_p90000_p99000_p99300
id_taptest_table_p90000_p99000_p99400
id_taptest_table_p90000_p99000_p99500
id_taptest_table_p90000_p99000_p99600
id_taptest_table_p90000_p99000_p99700
id_taptest_table_p90000_p99000_p99800
id_taptest_table_p90000_p99000_p99900
(1124 rows)
```

Now all 100,000 rows are properly partitioned where they should be and any new rows should go where they're supposed to.

Set `run_maintenance()` to run often enough

Using the above time-based partitions, `run_maintenance()` should be called at least twice a day to ensure it keeps up with the requirements of the smallest time partition interval (daily).

For serial based partitioning that uses `run_maintenance()` (the sub-partitioning above does so), you must know your data ingestion rate and call it often enough to keep new partitions created ahead of that rate.

If you're using the Background Worker (BGW), set the `pg_partman_bgw.interval` value in `postgresql.conf`. This example sets it to run every 12 hrs (43200 seconds). See the `doc/pg_partman.md` file for more information on the BGW settings.

```
pg_partman_bgw.interval = 43200
pg_partman_bgw.role = 'keith'
pg_partman_bgw.dbname = 'keith'
```

If you're not using the BGW, you must use a third-party scheduling tool like cron to schedule the calls to `run_maintenance()`

```
03 01,13 * * * psql -c "SELECT run_maintenance()"
```

Use Retention Policy

To drop partitions on the first example above that are older than 30 days, set the following:

```
UPDATE part_config SET retention = '30 days', retention_keep_table = false WHERE parent_table
= 'partman_test.time_taptest_table';
```

To drop partitions on the second example above that contain a value 100 less than the current max (`max(col1) - 100`), set the following:

```
UPDATE part_config SET retention = '100', retention_keep_table = false WHERE parent_table =
'partman_test.id_taptest_table';
```

For example, once the current id value of `col1` reaches 1000, all partitions with values less than 900 will be dropped.

If you'd like to keep the old data stored offline in dump files, set the `retention_schema` column as well (the `keep*` config options will be overridden if this is set):

```
UPDATE part_config SET retention = '30 days', retention_schema = 'archive' WHERE parent_table
= 'partman_test.time_taptest_table';
```

Then use the included python script `dump_partition.py` to dump out all tables contained in the archive schema:

```
$ python dump_partition.py -c "host=localhost username=postgres" -d mydatabase -n archive -o
/path/to/dump/location
```

To implement any retention policy, just ensure `run_maintenance()` is called often enough for your needs. That function handles both partition creation and the retention policies.

Undo Partitioning: Simple Time Based

As with partitioning data out, it's best to use the python script to undo partitioning as well to avoid contention and moving large amounts of data in a single transaction. Except for the final example, there's no data in these partition sets, but the example would work either way. This also shows how you can give time-based partition sets a lower interval than what they are partitioned at. This set was daily above, but the batches are committed at the hourly marks (if there was data).

```
$ python undo_partition.py -p partman_test.time_taptest_table -c host=localhost -t time -i "1
hour"
Attempting to turn off autovacuum for partition set...
... Success!
Total rows moved: 0
Running vacuum analyze on parent table...
Attempting to reset autovacuum for old parent table...
... Success!
```

Undo Partitioning: Simple Serial ID

This just undoes the id partitions committing at the default partition interval of 10 given above.

```
$ python undo_partition.py -p partman_test.id_taptest_table -c host=localhost -t id
Attempting to turn off autovacuum for partition set...
... Success!
Total rows moved: 0
Running vacuum analyze on parent table...
Attempting to reset autovacuum for old parent table...
... Success!
```

Undo Partitioning: Sub-partition ID->ID->ID

Undoing sub-partitioning involves a little more work (or possibly a lot if it's a large set). You have to start from the bottom up. Just as I did above for generating statements for partitioning the data out, I can do the same for the `undo_partition.py` script. Keep in mind this gets the undo statement for ALL the parents at once. You do have to go through and parse out the top level calls as well as the mid-level partition, but this at least saves you a lot of potential typing (and typos). The bottom partitions must all be done first and the top last. Also, in this case I have no intention of keeping the old, empty tables anymore, so the `-droptable` option is given. `pg_partman` tries to be as safe as possible, so it only uninherits tables by default when undoing partitioning. If you want something dropped, you have to be sure and tell it.

```
SELECT 'python undo_partition.py -c host=localhost -p '||parent_table||' -t id -i 100 --droptable'
FROM part_config order by parent_table;
```

First do the lowest level sub-partitions:

```
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p0_p0 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p0_p1000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p0_p2000 -t id -i 100
--droptable
...
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p100000_p100000 -t id
-i 100 --droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p100000_p101000 -t id
-i 100 --droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p100000_p102000 -t id
-i 100 --droptable
```

Next do what were the mid level sub-partitions:

```
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p0 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p10000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p100000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p110000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p120000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p20000 -t id -i 100
--droptable
```

```
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p30000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p40000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p50000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p60000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p70000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p80000 -t id -i 100
--droptable
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table_p90000 -t id -i 100
--droptable
```

And finally do the last, top level partition:

```
python undo_partition.py -c host=localhost -p partman_test.id_taptest_table -t id -i 100
--droptable
```

Now there is only one table left with all the data

```
keith=# SELECT tablename FROM pg_tables WHERE schemaname = 'partman_test' order by tablename;
      tablename
-----
 id_taptest_table

keith=# SELECT count(*) FROM partman_test.id_taptest_table ;
      count
-----
 100000
(1 row)
```

Undo Partitioning: Sub-partition Time->Time->Time

This is done in the same exact way as for ID->ID->ID except the undo_partition.py script would use the -t time setting and -i would use a time interval value.

Hopefully these working examples can help you get started. Again, please see the `pg_partman.md` doc for the full details on all the functions and features of this extension. If you have any issues or questions, feel free to open an issue on the github page: https://github.com/keithf4/pg_partman