

OmniSci 5.8 の新機能

OmniSci バージョン 5.8 を発表できることを大変うれしく思います。このリリースでは、OmniSci のレンダリングエンジンに基本的な変更を加え、新しいデータベース機能、地理空間演算子、システム管理ツールを導入しています。

長年のユーザーの皆様の中には、"OmniSci はバージョン 5.7 をリリースしたばかりではないか？"と疑問に思われる方もいらっしゃるかもしれません。その答えは、"Yes!"です。

実際、Pey Silvester 氏のリーダーシップのもと、当社の有能なエンジニアリングチームは、より短期間で一貫性のあるリリーススケジュールに取り組んでおり、お客様、パートナーそして社内スタッフはすでにその恩恵を受けています。

この記事では、OmniSci バージョン 5.8 リリースの主な機能を紹介し、それらの機能をどのように活用できるかについていくつかの例を示します。

◆レンダリング

OmniSci Version 5.7 では、現行の OpenGL レンダラーに代わる新しい Vulkan を搭載したレンダラーの公式ベータ版を発表しました。Vulkan は、従来のものに比べて以下のような利点があります。

- 低オーバーヘッド
- より直接的な GPU 制御
- 低 CPU 使用率
- クロスプラットフォームサポート
- マルチスレッド処理対応

バージョン 5.8 では、OpenGL に代わって、正式に新しい Vulkan バックエンドレンダラーがデフォルトで有効となりました。ただし、OpenGL は今後のリリースで非推奨となり、削除される予定です。

また、Vulkan だけでなく、マルチ GPU コンポジターのメモリーフットプリントとパフォーマンスも改善されています。簡単に言えば、ノード上の GPU の数が多ければ多いほど、潜在的な利益は大きくなります。

◆データベース

クエリエンジン

バージョン 5.7 では、パブリックベータ機能として並列エグゼキュータが導入されましたこれにより、クエリを非同期的に並行して実行できるようになり、全体的な並行性が向上しました。並列エグゼキュータは、設定ファイルに `--num-executors=2` フラグを指定することで、デフォルトでオンになります。

```
port = 6274
http-port = 6278
data = "/var/lib/omnisci/data"
null-div-by-zero = true
num-executors = 3

[web]
port = 6273
frontend = "/opt/omnisci/frontend"
servers-json = "/var/lib/omnisci/servers.json"
enable-https = true
```

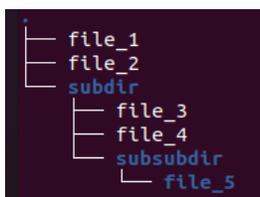
経験豊富な SQL 開発者は、以前のバージョンでは、当社のウィンドウ関数は特定のテーブルにしか適用できず、パーティション句が必要であることに気付きました。現在のバージョンでは、ウィンドウ関数は複数のフラグメントやシャードを持つテーブル上で動作し、パーティション句を指定する必要はありません。

最後に、WIDTH_BUCKET 演算子のサポートを有効にして、数値のビンングを容易にし、Immerse でのヒストグラムのパフォーマンスを向上させました。

ストレージ I/O

OmniSci バージョン 5.8 では、COPY FROM コマンドを実行する際に、ファイルパスの正規表現によるグローブ、フィルタリング、ソートが可能になりました。

例えば、以下のようなフォルダとファイル構成になっているとします。



フォルダとファイル構成

ワイルドカード付きのパス名、またはフォルダ名を指定して、ローカルにある Parquet ファイルや CSV ファイルをグロブすることができます。フォルダを指定してグロブすると、指定されたフォルダ以下のすべてのファイルが再帰的に返され、ワイルドカードを指定してグロブすると、展開されたファイルパスに一致するすべてのファイルパスが返されます。

```
#Recursive Globbing

COPY table_1 FROM ".../subdir";

returns file_3, file_4, file_5

#Wildcard Globbing

COPY table_1 FROM ".../subdir/file*";

returns file_3, file_4
```

ファイルフィルタリングは、グロブされた不要なファイルを削除する機能を拡張します。フィルタリングを使用するには、REGEX_PATH_FILTER オプションを指定し、このパターンに一致しないファイルは省略されます。

```
#Filtering

COPY table_1 from ".../" WITH (REGEX_PATH_FILTER=".*file_[4-5]");

returns file_4, file_5
```

ソート機能は、ファイルパス正規表現の新しい機能を補完するもので、次のようなオプションでファイルのインポート順を指定することができます。

- Pathname (デフォルト)
- date_modified
- regex *
- regex_date *
- regex_number *

```
#FILE_SORT_ORDER_BY  
  
COPY table_1 from ".../" WITH (FILE_SORT_ORDER_BY="date_modified");
```

最後に、インポート時に長い 10 進数の文字列をより強力に処理し、ArrowCSVForeignStorage の日付型をサポートしました。

◆地理空間 (Geospatial)

OmniSci は、すべての UTM ゾーンと EPSG:4326 (緯度/経度) および EPSG:900913 (ウェブメルカトル) との間の ST_Transform をネイティブにサポートしました。これらの改良により、ニッチな座標系や地域の座標系で保存された地理空間データのインポート、操作、分析が容易になりました。

バージョン 5.8 では、サポート対象の地理空間事業者を強化するための取り組みを続けています。

- ST_BUFFER
- ST_DIFFERENCE
- ST_UNION
- ST_Equals

最後に、ST_Distance を使用したポイントタイプ間の Spatial 結合が最大 100 倍高速化され、ST_DWithin は測地線距離の計算に GEOGRAPHY 入力を使用して実行されます。

◆システム管理

どのようなクエリが実行され、不要なジョブが中断されているかを理解することは、OmniSci で大規模なクエリをテストし、最適化する際に便利です。

これまでのバージョンでは、クエリを表示したり中断したりする機能は、スーパーユーザのみに許可されていました。今回、スーパーユーザ以外も同じことができるようになりました。

```
show queries;
```

query_session_id	current_status	executor_id	submitted	query_str	login_name	client_address	db_name	exec_device_type
713-t1ax	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	John	http:::1	omnisci	GPU
491-xpfb	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	Patrick	http:::1	omnisci	GPU
451-gp2c	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	John	http:::1	omnisci	GPU
190-5pax	PENDING_EXECUTOR	1	2021-08-03 ...	SELECT ...	Cavin	http:::1	omnisci	GPU
720-nQtV	RUNNING_QUERY_KERNEL	2	2021-08-03 ...	SELECT ...	Cavin	tcp:::ffff:127.0.0.1:50142	omnisci	GPU
946-ooNP	RUNNING_IMPORTER	0	2021-08-03 ...	IMPORT_GEO_TABLE	Rio	tcp:::ffff:127.0.0.1:47314	omnisci	CPU

```
kill query '946-ooNP'
```

```
show queries;
```

query_session_id	current_status	executor_id	submitted	query_str	login_name	client_address	db_name	exec_device_type
713-t1ax	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	John	http:::1	omnisci	GPU
491-xpfb	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	Patrick	http:::1	omnisci	GPU
451-gp2c	PENDING_QUEUE	0	2021-08-03 ...	SELECT ...	John	http:::1	omnisci	GPU
190-5pax	PENDING_EXECUTOR	1	2021-08-03 ...	SELECT ...	Cavin	http:::1	omnisci	GPU
720-nQtV	RUNNING_QUERY_KERNEL	2	2021-08-03 ...	SELECT ...				

バージョン 5.8 では、スーパーユーザは、新しい ALTER SYSTEM CLEAR コマンドを使用して、CPU、GPU、または Render のメモリをクリアすることができます。メモリ管理はサーバで行いますので、通常はこのコマンドを使用する必要はありません。

しかし、スーパーユーザが予期せぬメモリ関連の問題に遭遇した場合、メモリをクリアしてパフォーマンスが改善されるかどうかを試すことができます。

```
ALTER SYSTEM CLEAR CPU MEMORY
```

```
ALTER SYSTEM CLEAR GPU MEMORY
```

```
ALTER SYSTEM CLEAR RENDER MEMORY
```

OmniSci の最新版をダウンロードすることができます！

いつものように、素晴らしいリリースを生み出してくれた驚異的なエンジニアリングチームに感謝したいと思います。リリースノートの内容をご覧になりたい方は <https://docs.omnisci.com/overview/release-notes> をご参照ください。

OmniSci バージョン 5.8 へのアップグレードを予定されている方は、データディレクトリをバックアップし、OmniSci のアップグレードドキュメントに従って作業を行うことをお勧めします。企業のお客様は、必要に応じて OmniSci サポートに連絡してサポートを受けることができます。

ご不明な点がございましたら、LinkedIn や Twitter、コミュニティフォーラムなどでご質問ください。

Antonio Cotroneo は、OmniSci の技術コンテンツ戦略担当ディレクターで、これまでのキャリアを通じて、世界中の人々が地理空間データ、マッピング技術、空間分析を最大限に活用して、顧客やコミュニティにとって重要な意思決定を行うことができるよう支援してきました。南カリフォルニア大学で地理情報科学技術の修士号を取得し、Esri に 5 年以上勤務しました。現在は、妻と 2 人の子供と一緒にノースカロライナ州シャーロットに住んでいます。

※本文章は <https://www.omnisci.com/blog/whats-new-in-omnisci-5-8> を弊社で翻訳したものです。