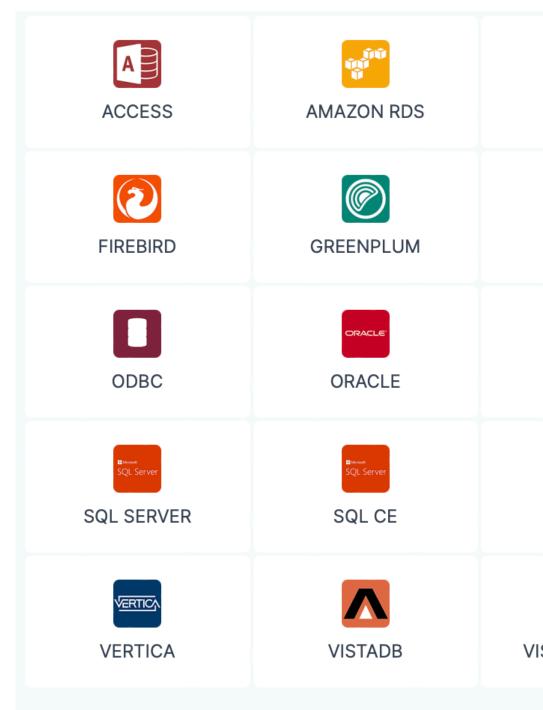
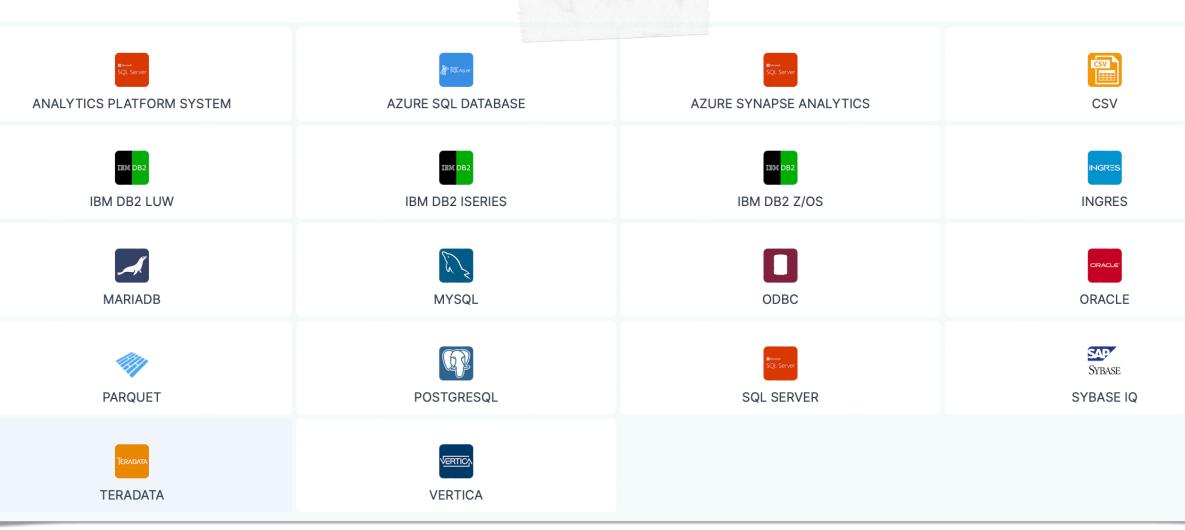
Archiving via DB conversion

Spectral Core Damir Bulic · dbulic@spectralcore.com · @BulicDamir

Databases supported Over 40 formats

- File-based
- Relational
- Cloud
- Data lakes





| CSV | IBM DB2 LUW | IBM DB2 Z/OS | IBM DB2 AS/400 | dBASE DBASE | EXCEL |
|-------------------------------|-------------|--------------|------------------|------------------------|----------|
| INGRES | INTERBASE | SQL Server | MARIADB | MYSQL | NEXUSE |
| PARADOX | PERVASIVE | POSTGRESQL | OPENEDGE | SIA RD SIARD | SQL AZU |
| opentext ⊽GUPTA SQLBASE | SQLITE | SYBASE ASE | SYBASE ADVANTAGE | SYBASE SQL ANYWHERE | TERADATA |
| | | | | | |

VISUAL FOXPRO

dBASE



Full Convert

spectralcore.com/fullconvert

- Single-click migration
- Drivers built-in, no external dependencies
- Scheduler built-in for recurring migrations
- Console included for easy scripting
- Customers in more than 90 countries
- Many success stories
- Full SIARD support!

COPYING THE DATABASE...



8 OF 8 WORKERS ON THE JOB

STOP

COPYING DATA

| 537 of 720 in tot copying dat | |
|-------------------------------------|---------|
| вутеs 16.0 GB of 19.7 GB | 51.9MB/ |
| RECORDS 87.9M of 176.6M | 1.3M/s |

| | | 500 | | | | 6 M 10 / |
|-------|----------------------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 418k | ot | 500k records | O. | 162.5k/s | 0 | 6.4MB/s |
| 408k | of | 500k records | 0 | 163.3k/s | 0 | 6.5MB/s |
| 399k | of | 500k records | o | 163.5k/s | ٥ | 6.5MB/s |
| 2,916 | of | 500k records | 0 | 159.76/0 | 0 | 6 3MR/c |
| | | | | | | |
| 361k | of | 500k records | O | 164.3k/s | 0 | 6.5MB/s |
| 350k | of | 500k records | 0 | 160.6k/s | 0 | 6.4MB/s |
| | | | | | | |
| 284k | of | 500k records | 0 | 168.2k/s | 0 | 6.7MB/s |
| 244k | of | 500k records | 0 | 166.0k/s | 0 | 6.6MB/s |
| | 408k 399k 381k 361k 350k 284k | 408k of 399k of 381k of 361k of 350k of 284k of | 408kof500k records399kof500k records381kof500k records361kof500k records350kof500k records284kof500k records | 408k of 500k records • 399k of 500k records • 381k of 500k records • 361k of 500k records • 350k of 500k records • 284k of 500k records • | 408k of 500k records 163.3k/s 399k of 500k records 163.5k/s 381k of 500k records 163.5k/s 381k of 500k records 159.7k/s 361k of 500k records 164.3k/s 350k of 500k records 160.6k/s 284k of 500k records 168.2k/s | 418k of 500k records • 162.5k/s • 408k of 500k records • 163.3k/s • 399k of 500k records • 163.5k/s • 381k of 500k records • 159.7k/s • 361k of 500k records • 164.3k/s • 350k of 500k records • 160.6k/s • 284k of 500k records • 168.2k/s • 244k of 500k records • 166.0k/s • |



Omni Loader Distributed migration cluster

- Ideal solution for migrating databases on premises to the cloud
- Handles hundreds of terabytes easily
- Same ease of use as Full Convert, for the basic use-case
- Works on Windows and Linux
- Fetches data from several source databases via hundreds of connections from several agents, compresses and encrypts in-flights, ingests to the target
- Repartitions data on the fly

SQL SERVE

3B31023.DATABASE.WINDOWS.NET.DEMO



AZURE SYNAPSE ANALYTICS

CSYNAPSE.DATABASE.WINDOWS.NET.SCSYNAPSE

| bc94864 jobs | ∆ wk-caas-2500e34 job | ∆ wk-caas-7b9ff5 3 jobs | ∆ wk-caas-edb11c3 jobs | ∆ wk-caas-dc55314 |
|------------------|-------------------------|--------------------------|--------------------------|----------------------|
| 3.6 thousand/s | 10.0 MB/s 5.7 thousand/ | 5.7 MB/s 5.8 thousand/s | 3.8 MB/s 5.3 thousand/s | 14.5 MB/s 5.7 thousa |
| 246c58 2 jobs | ∆ wk-caas-3a58153 job | ☆ wk-caas-6a616e 3 jobs | ∆ wk-caas-149c22 2 jobs | & wk-caas-756abc4 |
| 3.9 thousand/s | 11.2 MB/s 4.4 thousand/ | 11.5 MB/s 4.3 thousand/s | 7.9 MB/s 2.9 thousand/s | 11.6 MB/s 5.3 thous |
| 6dcda34 jobs | ∆ wk-caas-13c5c1 4 job | & wk-caas-16b1eb4 jobs | ∆ wk-caas-3b28482 jobs | ∆ wk-caas-b6591f 4 |
| 170.9 thousand/s | 4.1 MB/s 2.4 thousand/ | 12.2 MB/s 5.9 thousand/s | 5.6 MB/s 23.6 thousand/s | 6.5 MB/s 8.7 thousa |

CANCEL

TA 70.3% 5 OBJECTS IN PROGRESS (58 WORKERS) 14 DONE, 0 IN QU DB0.COMMENTS UPLOADING... 58% • 75.0 million records • 27.1 thousand/s • 25.3 GB • 2 7 million CHUNK 1 OF 6 75% • 12.8 million records • 6.6 thousand/s • 4.3 GB • 2 7 million CHUNK 2 OF 6 35% • 12.7 million records • 3.1 thousand/s • 4.3 GB • 1

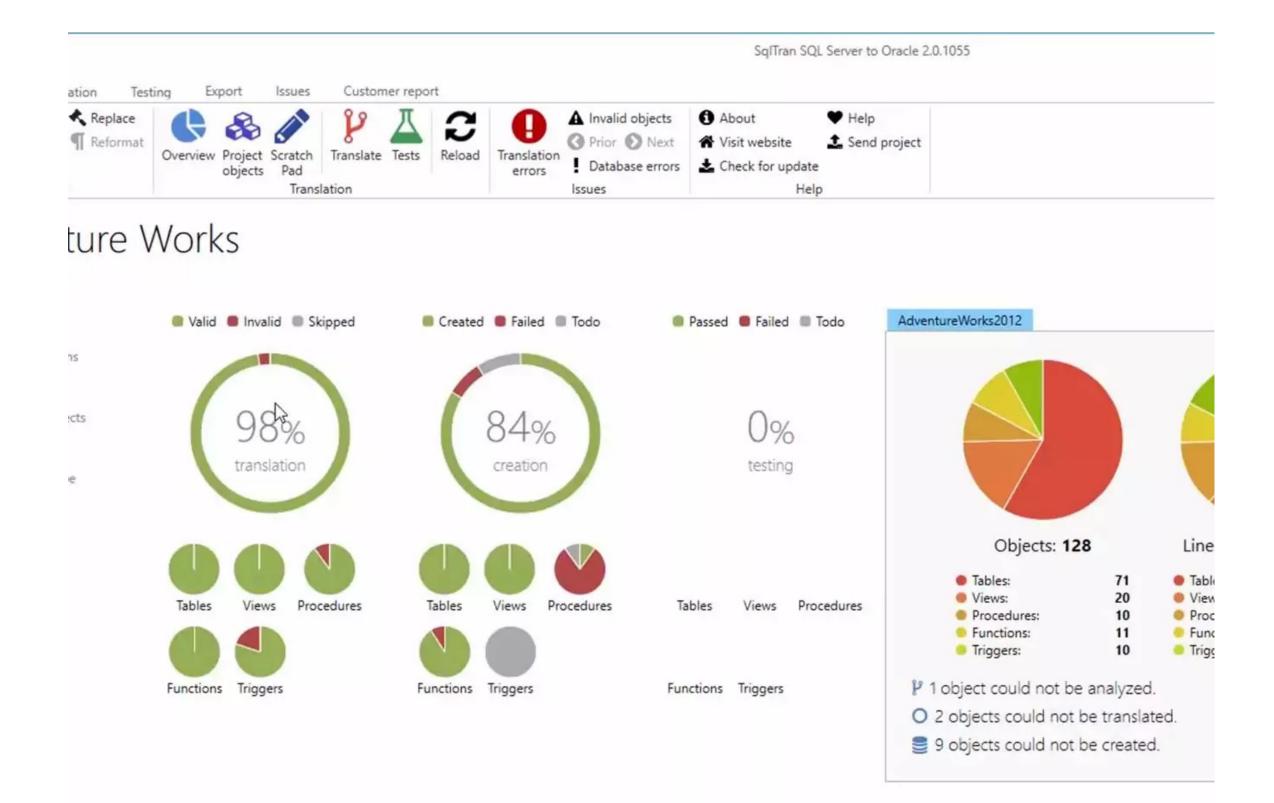
| | CHUNK 3 OF 6 | 44% • 12.5 million records • 3.8 thousand/s • 4.2 GB • |
|------|--------------|-----------------------------------------------------------|
| 2 GB | CHUNK 4 OF 6 | 19% • 12.4 million records • 1.6 thousand/s • 4.2 GB • 57 |
| MB/s | CHUNK 5 OF 6 | 77% • 12.1 million records • 6.5 thousand/s • 4.1 GB • |



SQL Tran

DDL translation of very complex schema

- Custom parser generator engine
- 150x faster translation than closest competitor
- Goal of >98% fully automated translation
- Manual translation available for code that needs complete rewrite
- Static analysis used to generate tests
- Uses OmniLoader engine for data migration



Documenter

Database schema documentation

- Reads full database metadata (same engine as SQL Tran)
- Builds a HTML website with the navigable schema
- Discontinued! To be included into upcoming database manager.

S HumanResources.Employee ×

🔶 🔶 C 🕧 File 🛛 C:/Users/Damir/Documents/Spectral%20Core/Documenter/Database%20documentation/R12.AdventureWorks2017/_Table_HumanResource... 🟠 🗧

R12.AdventureWorks2017

Assemblies | Data Types | Roles | Schemas | Tables | Views | Procedures | Functions | Triggers

R12.AdventureWorks2017 Database

HumanResources.Employee Table

Information

| Name | Employee |
|----------------------|--------------------|
| Schema | HumanResources |
| Row Count | 0 |
| Data Size | |
| Index Size | |
| Reserved Size | |
| Unused Size | |
| Created | 27-Oct-17 14:33:01 |
| Modified | 27-Oct-17 14:33:14 |

Columns

| PK | Key | Identity | Name | Data Type | Allow Nulls | Collation | References | Default | Computed | Compute Expressio |
|----|-----|----------|-------------------|------------------|-------------|------------------------------|------------|----------------------|----------|---------------------------|
| 9 | | | BusinessEntityID | int | | | | | | |
| | | | NationalIDNumber | nvarchar(15) | | SQL_Latin1_General_CP1_CI_AS | | | | |
| | | | LoginID | nvarchar(256) | | SQL_Latin1_General_CP1_CI_AS | | | | |
| | | | OrganizationNode | hierarchyid | V | | | | | |
| | | | OrganizationLevel | smallint | V | | | | M | ([OrganizationNode].[GetL |
| | | | JobTitle | nvarchar(50) | | SQL_Latin1_General_CP1_CI_AS | | | | |
| | | | BirthDate | date | | | | | | |
| | | | MaritalStatus | nchar(1) | | SQL_Latin1_General_CP1_CI_AS | | | | |
| | | | Gender | nchar(1) | | SQL_Latin1_General_CP1_CI_AS | | | | |
| | | | HireDate | date | | | | | | |
| | | | SalariedFlag | bit | | | | 1 | | |
| | | | VacationHours | smallint | | | | 0 | | |
| | | | SickLeaveHours | smallint | | | | 0 | | |
| | | | CurrentFlag | bit | | | | 1 | | |
| | | | rowguid | uniqueidentifier | | | | <pre>newid()</pre> | | |
| | | | ModifiedDate | datetime | | | | <pre>getdate()</pre> | | |



Archival needs today We got you covered

- Full Convert supports SIARD for reading and writing
- Wide range of databases can be easily converted to SIARD with minimal work
- SIARD can be extracted to wide range of databases with minimal work
- Full customization available (renaming, mapping, BLOB placement)
- High performance reader and writer (limited by inherent limitations of SIARD) format)



SIARD challenges It's XML. Zipped XML.

- We can handle up to 16 connections to speed things up, but...
- cores. In the end, just one core is used to compress data.
- XML is verbose and clunky
- solution.
- Reasonably, SIARD is useful only for small datasets
- extreme pace and I expect SIARD to be less applicable in the future.

• ZIP writing can't be parallelized. I have a 12-core CPU. We generate as much as possible using all

• XML is not well suited to handle a lot of data. We can spill BLOB data outside, but it's a clunky

• SIARD already can't handle much of what we see in real world today. Datasets are growing at

That's why SIARD is not supported in Omni Loader, our most powerful database migration software.

Looking into the future **Horizontal scaling**

- Vertical scaling is dead. Moore's law is dead.
- cores.
- Looking at cloud, it is all about horizontal, on-demand scaling. This trend will continue.
- Workflows are moving away from personal computers. Developers are often remoting from laptops to remote development environments with less limitations (network, CPU).
- (data lakes, OLTP, OLAP, NoSQL)
- spatial data etc.

• Vertical scaling of a single server today is in reality a horizontal scaling - more cores, not faster

Data is streamed from many sources (IoT, heterogeneous databases) into specialized storage

• While we can expect relational engines to stay with us for a very long time, we can also expect new engines to appear. We have graph data, hierarchical data, custom data types, CLR types,

Future proofing archives A new archival data format

- As described in an email I sent
- In short:
 - Separate structure for data and schema
 - Compressed chunks of columnar data (we use that in Omni Loader already, similar to Parquet)
 - Extensible data types
 - Schema can be written in a single SQLite file (even though JSON would be good enough)
 - Data should be written in one file, or many files on a local network, or anywhere in the world. lacksquareInstead of referencing data separately for each value, we should be referencing a chunk of data.
 - Data chunk can contain from a few records all the way to billions of records.





Discussion time.