



ChilIETL Fact Sheet



ChilIETL™ by Cloud Data Solutions is a master scheduler for Azure Data Factory. It is completely metadata driven.

ChilIETL eliminates the redundant grunt work associated with ETL so developers can spend more time on enhancing business value. It's complete, integrated, scalable, architecture elegantly addresses the common design-patterns associated with scheduling and logging of ETL processes. ChilIETL provides intelligent restarts on-failures, real time status reporting, documentation, notifications, and it encourages code re-use. The entire framework is built on an open Azure architecture: Azure Data Factory, SQL Azure, Azure Logic Apps and Power BI.

We offer clients an alternative to reinventing the wheel and building their own backend ETL architecture. Our clients realize a solid ETL architecture is crucial to a successful BI project but that the business simply can't appreciate their spending time on this. Implementing ChilIETL allows you to fast track your BI Project and focus on delivering data to the business.

Common ETL Challenges and the ChilIETL Solution

Cloud Data Solutions designed ChilIETL based on our years of experience developing ETL Solutions. The following are typical problems IT departments have encountered in developing ETL and how ChilIETL solves them.

ETL Backend Development Takes Much Too Long Before Businesses See Tangible Results

The lag between the start of the development process and when meaningful results can be demonstrated is typically unacceptably slow to the business. Developers must spend considerable time building out the ETL backend architecture and methodically staging each table.

ChilIETL Solution: Complete Out of the Box Architecture Speeds Up Development

ChilIETL is a clean, complete, well documented solution built on industry best practices. Long implementation times are eliminated because backend work is already done.

Flavor of the Day Tools GUI Interfaces Fail in the Long Run

GUI, no code interfaces demo well but fail to deliver in the real world because they can't handle complex business logic. Also, these tools are difficult to maintain since the logic is not as easy to see as viewing a SQL statement.

ChilIETL Solution: Relies on SQL and Standard Azure Resources

The solution utilizes Azure Data Factory, Azure SQL, Azure Logic Apps, Excel, and Power BI.

Maintenance Costs Increase Over Time Due to varying, inconsistent coding techniques

Many developers come up with their own coding patterns, which means that inconsistencies abound. There is considerable redundancy in ETL and many developers copy the code from a similar process and then modify it as needed. This may save time in the short term but it results in a large number of processes that can't possibly be kept up to date. This problem is compounded with developer turnover as the inconsistent approach relies on tribal knowledge.

ChillETL Solution: Meta-Data Driven Approach

Meta data replaces inefficient copy/paste coding, with a low code development process. It enforces a standard, repeatable development process. This approach facilitates onboarding new developers.

Patchwork Scheduling Adds Complexity

One typical ETL scheduling approach is embedding processing chains in other processes. This approach creates cumbersome schedules. Inefficiencies, such as spaghetti code, sequential processing, and added complexity arise which slows development and troubleshooting.

ChillETL Solution: Flexible Scheduling

Replace patch work schedules with unlimited, instant changes by simply updating Meta data. Processes can execute concurrently with some processes and sequentially with others based on dependencies.

Stagnant Documentation

In many cases, documentation is done once at the "hand off" time and never updated. Because data flow and source to target mapping spreadsheets are not updated, whatever is provided to the business and future developers comes with the carveout, "it isn't current".

ChillETL Solution: Create Updated, Accessible Documentation

Documentation is stored in the meta-data database, is always up to date, and easily accessible via Power BI reports.

A Single Process Failure Often Leads to a Missed SLA

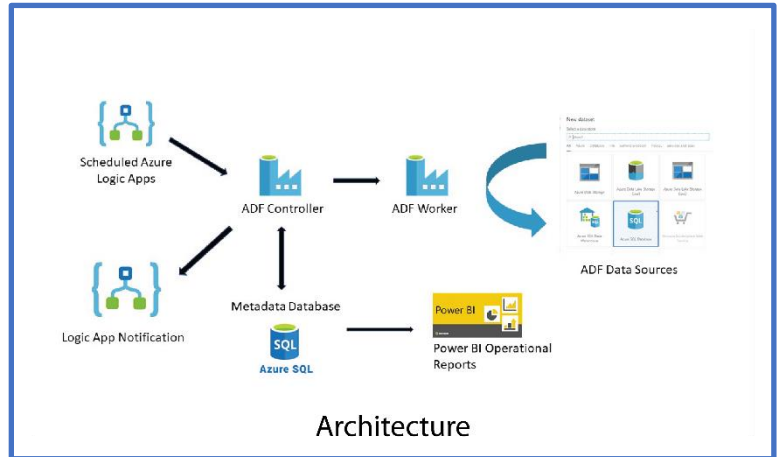
Troubleshooting production ETL is a time critical activity. Inconsistent coding, patchwork schedules, and a reliance on tribal knowledge make it difficult to quickly identify what process failed and how to recover. To compensate, developers rerun the entire job, which frequently means the data is not available on time.

ChillETL Solution: Robust Error Handling

Failure notifications identify the process that failed and the error message. A failed job can be retried so that only the process that failed and their dependencies are executed.

ChilIETL Components

- **Controller ADF:** a complete meta data driven process scheduler. Deploy the ARM template and you are ready to go.
- **Worker ADF Pipeline Templates:** prebuilt, pipeline templates for common ETL processes are included, which can be customized as needed using standard ADF features.
- **Logic App Templates:** Azure Logic App templates are provided for email notification and master job scheduling.
- **Power BI Operational Reports:** Power BI reports provide a real time view of ETL processing. Historical visualizations identify what processes are becoming bottlenecks.
- **Metadata SQL Database:** a SQL database that controls the entire ETL process. It stores multiple environments, schedules, process dependencies and pipeline parameter values.
- **Excel Metadata Editor:** an Excel Frontend that populates the SQL Metadata database. This editor can also generate the SQL objects typically required in an ETL solution and update a Power BI Model.



ScheduleCode	Process Type	Start	End	Min	Status	Error Message
01 - Setup	ScaleSQLPipeline	01:00 AM	01:01 AM	1	Completed	
02 - Staging	CopyOnPremisetoSQLAzure	01:02 AM	01:03 AM	1	Completed	
03 - Dimensions	ExecuteProcPipeline	01:03 AM	01:04 AM	1	Completed	
04 - Facts	ExecuteProcPipeline	01:05 AM	01:08 AM	3	Completed	
05 - Power BI	RefreshPBIDataSetPipeline	01:11 AM	01:15 AM	4	Running	
06 - Cleanup	ScaleSQLPipeline				Pending	

ChilIETL Status Report

The screenshot shows the Excel Metadata Editor interface. The main data table is as follows:

ProcessName	ProcessDescription	SchedulePhaseName	procname	ProcessTag	TimeoutSec	PollingInterval	IsDisable
Update customers	Execute customers_merge stored procedure to update customer	Dimensions	customers_merge	customers	180	30	FALSE
Update products	Execute products_merge stored procedure to update products	Dimensions	products_merge	products	180	30	FALSE
Update distcenters	Execute distcenters_merge stored procedure to update distcent	Dimensions	distcenters_merge	distcenters	180	30	FALSE
Update shippers	Execute shippers_merge stored procedure to update shippers	Dimensions	shippers_merge	shippers	180	30	FALSE
Update salesreps	Execute salesreps_merge stored procedure to update salesreps	Dimensions	salesreps_merge	salesreps	180	30	FALSE
Update contacts	Execute contacts_merge stored procedure to update contacts	Dimensions	contacts_merge	contacts	180	30	FALSE

ChilIETL Metadata

Try Before You Buy

Organizations want to be comfortable that a solution fits their environment before making a significant investment. To meet this need we offer a ChilIETL trial. It allows you to test all the functionality of ChilIETL in your environment with your data before licensing ChilIETL.

Consulting Services Available as Needed

We offer ChilIETL installation assistance, ADF pipeline development, and Training. We can also provide consulting services in designing and building out your BI solution. Our goal is to provide the services you need for a successful Microsoft Azure Cloud Based BI solution.

For More Information



<http://www.Clouddata.solutions>

Info@Clouddata.Solutions

(480) 389-6905