

Retek Extract Transform and Load (RETL) 10.2.1



Release Notes



The software described in this documentation is furnished under a license agreement, is the confidential information of Retek Inc., and may be used only in accordance with the terms of the agreement.

No part of this documentation may be reproduced or transmitted in any form or by any means without the express written permission of Retek Inc., Retek on the Mall, 950 Nicollet Mall, Minneapolis, MN 55403, and the copyright notice may not be removed without the consent of Retek Inc.

Information in this documentation is subject to change without notice.

Retek provides product documentation in a read-only-format to ensure content integrity. Retek Customer Support cannot support documentation that has been changed without Retek authorization.

Corporate Headquarters:

Retek Inc.
Retek on the Mall
950 Nicollet Mall
Minneapolis, MN 55403

888.61.RETEK (toll free US)
+1 612 587 5000

Retek[®] Extract Transform and Load[™] is a trademark of Retek Inc.

Retek and the Retek logo are registered trademarks of Retek Inc.

European Headquarters:

Retek
110 Wigmore Street
London
W1U 3RW
United Kingdom

Switchboard:
+44 (0)20 7563 4600

Sales Enquiries:
+44 (0)20 7563 46 46
Fax: +44 (0)20 7563 46 10

This unpublished work is protected by confidentiality agreement, and by trade secret, copyright, and other laws. In the event of publication, the following notice shall apply:

©2002 Retek Inc. All rights reserved.

All other product names mentioned are trademarks or registered trademarks of their respective owners and should be treated as such.

Printed in the United States of America.

Customer Support

Customer Support hours:

Customer Support is available 7x24x365 via e-mail, phone, and Web access.

Depending on the Support option chosen by a particular client (Standard, Plus, or Premium), the times that certain services are delivered may be restricted. Severity 1 (Critical) issues are addressed on a 7x24 basis and receive continuous attention until resolved, for all clients on active maintenance.

Contact Method Contact Information

Internet (ROCS) www.retek.com/support
Retek's secure client Web site to update and view issues

E-mail support@retек.com

Phone US & Canada: 1-800-61-RETEK (1-800-617-3835)
World: +1 612-587-5800
EMEA: 011 44 1223 703 444
Asia Pacific: 61 425 792 927

Mail Retek Customer Support
Retek on the Mall
950 Nicollet Mall
Minneapolis, MN 55403

When contacting Customer Support, please provide:

- Product version and program/module name.
- Functional and technical description of the problem (include business impact).
- Detailed step by step instructions to recreate.
- Exact error message received.
- Screen shots of each step you take.

Retek Extract Transform and Load (RETL)

The Retek Extract Transform and Load (RETL) is a high-performance runtime tool used in parallel processing systems where high volumes of data must be processed quickly. By incorporating RETL into an application, the amount of time required to process data from databases and flat files may be reduced. Increasing the number of processors on Unix servers can scale RETL to handle larger volumes of data.

Functional enhancements

The RETL 10.2 product release and the subsequent 10.2.1 patch release include several important product improvements that enhance the product's performance, provide greater independence from environmental variables and dynamic libraries as well as allow the RETL to operate as a stand-alone binary.

We strongly recommend that clients who took delivery of the RETL 10.2 (or prior versions) upgrade to the RETL 10.2.1 to take advantage of the product improvements and issue resolutions.

The RETL releases include a tool to verify and diagnose issues, and the product's operating system and database compatibilities have been expanded to include AIX 5.1L and Oracle 9.2.

The RETL releases are backward compatible with previous releases; however they strictly enforce input schemas and data fields. In some instances, the RETL releases will find data and/or flow errors that previously went unreported. The result may be more discarded, rejected records and / or error messages than previously identified.

Clients that took delivery of the RETL 1.7.1 (or prior releases) are strongly encouraged to upgrade to the RETL 10.2 product release and apply the RETL 10.2.1 patch release in order to take advantage of the product improvements:

RETL 10.2.1

This is a patch release intended to resolve a memory related issue. The issue results in abnormally high memory consumption that impacts the server operations. The resolutions significantly reduce the amount of memory that RETL requires and allocates. In addition, locking mechanisms that protect the dataset connection operators were made more efficient.

The combination of improved memory efficiency and locking mechanisms have resulted in noticeable performance improvement and greatly decreased memory consumption.

RETL 10.2

The RETL 10.2 product release includes the following technical improvements.

- Several performance enhancements have significantly improved processing speed. The processing improvements will be most noticeable with complex flows that process larger numbers of records.
- Dynamic library dependencies are minimized; therefore dependencies on environment variables have been reduced.
- Support for Oracle Version 9, Release 2 (9.2).
- Support for AIX 5.1L.

Technical enhancements

RETL 10.2.1

The RETL 10.2.1 patch release includes the following technical improvements:

- RETL was consuming large amounts of memory and as a result, negatively impacting server operation. This was a result of a circular dependency that caused dataset expansion when not needed.
- Allocates memory chunks more conservatively. The flows may now run with significantly less memory than in release 10.2 and 1.7.1.
- Adds the ability to clean up memory as operators finish. This reduces overall memory consumption and allows memory to be released prior to completion of the entire flow.
- Improves efficiency of locking within Datasets. This improves overall efficiency of data processing and results in a performance increase. This change should also allow better scale because operators no longer need to synchronize on a per record basis. In short, the more complex the flow, the more noticeable the improvement should be.
- An improvement to the AIX 4.3.3 platform allowing the use of an enhanced DataSet. In the RETL 1.7.1 and 10.2 releases, a timed mutex wait was required within the more advanced datasets. This operation impacted performance on AIX 4.3.3. As a result, RETL used a less efficient DataSet for this platform. This has been improved by removing the need for the timed wait and may result in significant performance improvements for that platform.

RETL 10.2

- Several performance enhancements have significantly improved processing speed. The processing improvements will be most noticeable with complex flows that process larger numbers of records.
- Dynamic library dependencies are minimized; therefore dependencies on environment variables have been reduced.
- A new tool to verify and diagnose problems with a RETL installation is now being packaged. The `verify_retl` tool is run as part of the RETL installation. (See the `README.verify_retl` in the base of the package for more information regarding this tool.)
- The OraRead operator may run stored procedures if "sp_prequery" and /or "sp_postquery" properties are specified. One of usages is Flashback for Oracle 9i. See the RETL 10.2 Programmer's Guide for more details.
- The new property "primaryindex" has been added to the TeraWrite operator allowing users to specify primary indexes for created tables. This property improves performance when tables are created or recreated. If the "primaryindex" property is not specified, no primary indexes will be created.
- TeraWrite uses FastLoad rather than MLoad when writing to empty tables. This should result in better performance in that specific case.
- Improved messages within TeraWrite when `Rfx_Show_sql` is exported - shows when MLoad and FastTLoad are being called.
- If `Rfx_Debug` and `Rfx_Show_sql` are specified TeraWrite will write the data that is being written to the DB also to a flat file for debugging.
- Modified Filter operator to alert user when multiple outputs have been specified but 'rejects' property is not specified or set to false. Previously the user would have seen the message: 'Num outputs does not match attached datasets'.
- Differences of int64 on 32bit and 64bit platforms have been resolved to check underflow and overflow conditions. The differences resulted in erroneous data being returned when performing conversions of large numbers.
- DB2Write now defaults to conventional mode loading.
- Improved error messaging with sql loader failures. The error messages include more information (LogFile and control file, rather than just ParFile) when `RFX_SHOW_SQL=1`.
- RETL no longer accepts blank lines as valid input.

- The RETL allows "" to represent a blank string for a null values of fixed length records within the schema files. Previously, RETL required the developer to specify a null value of exactly the same length as the fixed length field. Thus for a 500 character field RETL required that RETL developers specify "<with 500 blanks>". Developers may now simply enter "" to represent any number of blanks for the null value of the fixed length field.
- Better detection and reporting of field lengths and other properties.
- When converting from number to string, string now receives a Maxlength which is necessary for writing to a database. This enhancement allows developers to collapse flows into one flow that is capable of converting and writing to the database.
- Permissions of any files RETL uses are now checked. This includes making sure flows/schemas/configuration files/import files are readable, and any temp space RETL uses is writable. Appropriate error messages will be displayed if files don't have the correct permissions.
- During an import, if a null value existed in a date field and the field could not be null, a seg fault / core dump occurred. Now the record is rejected since it does not meet the schema specification.
- Previously, multiple string statements within Filter caused RETL to fail with the message "Invalid String passed to filter. String must be enclosed in single quotes." Filter now handles the problem without error.
- Import operator now generates an error message when a new line is the last character of the import data file. This empty line is equivalent to an empty record and is rejected as not meeting the file schema.
- Filter now treats all null values as being less than numeric values. Previously it would have treated, for example, '100' as being less than a 'null' value.
- If the "filter" property of the Filter operator is empty the following error will be displayed "RFX Unexpected Error: FilterOperator.cpp:119: [filter:0]: Empty value for 'filter' property.".
- Error message printing has been improved.
- A memory issues has been resolved. The issue occurred when delimited non-string fields were too long (invalid data). The result was a garbled error message.
- Environment checks to ensure that RFX_HOME, ORACLE_HOME and other needed variables are set in order. The user is now given better error messages when these variables aren't set.
- The MaxLength of string is set when a non-string is converting to string in the Convert operator.

- When BufSize is set to a very small number, there are no longer deadlocks.
- Time and Timestamp fields are now validated and will be dealt with as improperly formatted fields upon IMPORT.

Supported platforms

OS	Version	Arch	DB	Version	RETL Executable
AIX	4.3.3	32	none	None	rfx.a433.nodb.64
AIX	4.3.3	64	none	None	rfx.a433.nodb.32
AIX	4.3.3	64	Oracle	9.2	rfx.a433.o9x.64
AIX	4.3.3	64	Oracle	9.0.1	rfx.a433.o9x.64
AIX	4.3.3	32	DB2	7.2.3	rfx.a433.d723.32
AIX	4.3.3	32	Oracle	8.1.7	rfx.a433.o817.32
AIX	4.3.3	32	TeraData	2r4.2	rfx.a433.t241.32
AIX	5.1	64	Oracle	9.2	rfx.a51.o9x.64
HP-UX	11i	64	none	None	rfx.h11i.nodb.64
HP-UX	11i	32	none	None	rfx.h11i.nodb.32
HP-UX	11i	64	Oracle	9.2	rfx.h11i.o9x.64
HP-UX	11i	64	Oracle	9.0.1	rfx.h11i.o9x.64
HP-UX	11i	32	Oracle	8.1.7	rfx.h11i.o817.32
HP-UX	11i	32	TeraData	2r4.2	rfx.h11i.t241.32
Solaris	8	32	None	None	rfx.s58.nodb.32
Solaris	8	64	None	None	rfx.s58.nodb.64
Solaris	8	32	DB2	7.2.3	rfx.s58.d723.32
Solaris	8	64	Oracle	9.2	rfx.s58.o9x.64
Solaris	8	64	Oracle	9.0.1	rfx.s58.o9x.64
Solaris	8	32	Oracle	9.0.1	rfx.s58.o9x.32
Solaris	8	32	Oracle	8.1.7	rfx.s58.o817.32
Solaris	8	32	TeraData	2r4.2	rfx.s58.t241.32

Notes:

- Entries of 'none' (above) signify that the corresponding RETL executable may run standalone from a database. For example, if RETL is to be used standalone on Solaris 8 32bit, the 'rfx.s58.o9x.32' binary should be installed and used.
- Solaris 2.8, Solaris 8 and SunOS 5.8 are the same platform.
- If your current Operating System or Database is not specified above, please contact Retek Customer Support. Additional configuration will be supported on a case-by-case basis.

Known issues with resolution

The following issues have been identified in the RETL 10.2 product release:

- Translation from number to string for large arbitrary precision floating point numbers.
- When specifying fields in the Filter operator, they must all be uppercase otherwise they will not be recognized.
- The DropEdit of the Diff operator does not work properly.
- RFX leaves some temporary files in temp directories. These directories should be purged on a regular basis until this issues is resolved.