

# Retek<sup>®</sup> Extract Transform and Load<sup>™</sup> 11.1

## 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<sup>®</sup> Extract Transform and Load<sup>™</sup> is a trademark of Retek Inc.

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

This unpublished work is protected by confidentiality agreement, and by trade secret, copyright, and other laws.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:

©2003 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.

**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



## ***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](http://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.



## RETL 11.1 Overview

The Retek Extract Transform and Load (RETL) is a high-performance runtime tool useful 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.

## RETL 11.1 Functional and technical enhancements

The RETL 11.1 generally available release includes significant improvements that enhance the product's configuration management, operations, performance and translation support.

**We strongly recommend that before any installation or development work is performed, that the RETL 11.1 Programmer's Guide, provided with the release, is read cover to cover. In particular, read and understand Chapter 2 and the "Upgrade From A Previous Versions Of RETL" and "Backwards Compatibility Notes" sections as well as the "Known Issues" section of these Release Notes. All certified Retek products using RETL must be regression and volume tested with the RETL 11.1 tool prior to its production deployment at customer sites.**

Issues identified in the 11.0 release and prior releases have been resolved with enhancements and fixes being delivered through RETL 11.1 and subsequent patch releases. No additional releases of RETL 10.x are planned. This upgrade is recommended to take advantage of several improvements.

- Retek Product Certification** – The primary focus of the RETL 11.1 effort was certification of all generally available Retek products using RETL. Known backward compatibility issues to 10.x release are identified in the Known Issues section of these release notes. The generally available Retek product release versions that have been certified and required with RETL 11.1 include:

Retek Product	Certified Release	Planned Release Date
Retek Data Warehouse (RDW)	10.2.2	30-Sep-2003
Retek Merchandising System (RMS)	10.1.5	30-Sep-2003
Retek Inventory Optimization (RIO)	10.3.1	14-Nov-2003
Retek Demand Forecasting (RDF)	10.1 – Obtain patch to RDF interfaces through ROCS	26-Nov-2003

- **True Platform Independence** - The new architecture of the 11.x releases has been written entirely in Java. As a result, RETL may now run on platforms that previously would have required significant efforts to port to. See the F.A.Q question in the Programmer's Guide "Since RETL has been re-written in Java, can I run RETL on other platforms that have a JRE?" and the Compatibility Matrix for supported platforms. As a direct result of the platform independence, significant RETL resources in platform maintenance will be freed up to develop new features going forward.
- **True Database Independence** - With the 11.x releases, generic 'dbread'/'dbwrite' operators have been introduced. These operators use JDBC-compliant technology to connect to databases in a truly database-independent manner. Additionally, previous 10.x releases were unable to connect to different database types in the same flow. 11.x releases allow full interoperability among different databases in the same flow (e.g. among db2, oracle, teradata all in the same flow). See the F.A.Q question in the Programmer's Guide "With the introduction of RETL 11.x, JDBC technology has been introduced as a mechanism for connecting and reading/writing to databases. Does this mean I can connect to any database (e.g. Sybase, MySQL, Acumate, etc)?"
- **Performance Improvements** - RETL 11.x releases include a new architecture that optimizes pipeline parallelism by consolidating connected operators into the same pipeline where possible. This reduces the number of threads required, minimizing context switching and thrashing among threads, and thereby allowing a framework that supports improved performance and scalability. Additionally, optimizations have been made in certain operators, namely oraread and export, which have been shown to positively affect performance in comparison to 10.x releases. Internal performance benchmarking between 10.x and 11.x releases has shown 11.x releases to be anywhere from 15%-150% faster.
- **Orawrite Improvements** - Significant configuration and performance improvements in the orawrite operator have been implemented. Users may now run orawrite using sql\*loader in parallel mode. Internal testing has shown improvements up to 100% faster than when running in non-parallel mode. Users may also specify the following properties: ROWS, SORTEDINDEXES, SINGLEROW, PRELOAD, and POSTLOAD. See the [RETL 11.1 Programmer's Guide](#) for more information on how to use these properties.
- **Native Load Utilities** - RETL 11.1 includes support for the native database load utilities in orawrite (sql\*loader) and terawrite (mload/fastload). Performance increases up to 40-50% have been realized for flows with these operators in them when compared to 10.x versions.
- **Error Management** - 11.x releases introduce better flow debugging by giving line and column specific error messages when RETL encounters problems reading an XML flow definition. Additionally, 11.x's exception handling allows for better error handling and debugging.

- **Online Help** - A command-line option has been added to 11.x releases that allow flow developers to view operator syntax and usage via the command-line without having to refer to this document.
- **Installation and Configuration Simplification** - 11.x releases consolidate code into a single binary rather than the 26 binaries of 10.x release. This considerably eases installation. In addition, there is less environment setup that needs to be done in order to complete an installation.
- **Backwards Compatibility** - A fundamental requirement of 11.x versions is that they be backwardly-compatible with the 10.x versions. However, 11.x versions more strictly enforce valid XML flow interfaces, input schemas, and data fields. In some instances, the 11.x product will find data and/or flow errors that may have previously been unreported. The result may be more discarded or rejected records and/or error messages than previously identified.

OS	Version	Arch	DB	Version	RETL Executable
IBM AIX	64 bit	5.1	Oracle	9.2.0.2	IBM AIX
HP-UX	64 bit	11i	Oracle	9.2.0.2	HP-UX
HP-UX	64 bit	11i	Oracle	9.0.1.4	HP-UX
HP-UX	32 bit	11i	Oracle	8.1.7.4	HP-UX
HP-UX	32 bit	11i	NCR Teradata	2r4.2	HP-UX
HP-UX	64 bit	11i	None	None	HP-UX
HP-UX	32 bit	11i	None	None	HP-UX
Sun Solaris	64 bit	5.8	Oracle	9.2.0.2	Sun Solaris
Sun Solaris	64 bit	5.8	Oracle	9.0.1.4	Sun Solaris
Sun Solaris	32 bit	5.8	Oracle	9.0.1.4	Sun Solaris
Sun Solaris	32 bit	5.8	Oracle	8.1.7.4	Sun Solaris
Sun Solaris	32 bit	5.8	NCR Teradata	2r4.2	Sun Solaris
Sun Solaris	64 bit	5.8	None	None	Sun Solaris
Sun Solaris	32 bit	5.8	None	None	Sun Solaris

**Notes:**

- Certification with DB2 v7.2.3 was not performed as part of the RETL 11.1 certification effort. Currently, no Retek clients are using RETL with DB2 so the determination was made to re-schedule the DB2 certification effort of DB2 v8 for the next generally available release of RETL.
- 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. The RETL is able to process datasets either directly from RDBMS database tables or from Unix flat file based systems. The RETL stores information about each dataset that describes the data structures (the metadata). The means by which the metadata is supplied is dependent on the data interface operators provided by the RETL and may require use of a schema to ensure that datasets are consistent from the source data to the target data.

## IBM De-Support of AIX 4.3.3

Effective December of 2003, IBM will no longer support the AIX 4.3.3 operating system. As a result, RETL will discontinue support of the operating system once existing AIX 4.3.3 customers have migrated to the supported AIX operating system version, AIX 5.1 (and higher), or another supported operating system. AIX 4.3.3 does not support JRE v1.4. Retek is working with each impacted client regarding this migration.

OS	Version	Arch	DB	Version	RETL Executable
IBM AIX	64 bit	4.3.3	Oracle	9.2	IBM AIX
IBM AIX	64 bit	4.3.3	Oracle	9.0.1	IBM AIX
IBM AIX	32 bit	4.3.3	IBM DB2	7.2.3	IBM AIX
IBM AIX	32 bit	4.3.3	Oracle	8.1.7	IBM AIX
IBM AIX	32 bit	4.3.3	NCR Teradata	2r4.1	IBM AIX
IBM AIX	64 bit	4.3.3	None	None	IBM AIX
IBM AIX	64 bit	4.3.3	None	None	IBM AIX

# RETL 11.1 known issues with resolution

## DB2 Certification

Certification with DB2 was not performed as part of the RETL 11.1 certification effort. Currently, no Retek clients are using RETL with DB2 so the determination was made to re-schedule the DB2 certification effort of DB2 v8 for the next generally available release of RETL.

## General Known Issues of RETL 11.1 and prior releases

- The RETL does not handle arbitrary precision math.
- In the RETL 11.0 or prior releases, carefully check all property names to ensure they exactly match the property name and value spellings noted in the Programmer's Guide. RETL may error on misspelled or unknown properties.
- The export operator will default to use '|' delimited fields if there isn't a 'schemafilename' property specified. This can cause problems if the incoming data contains the '|' character. The recommendation is to use an export schemafilename if it is possible for data to contain a '|' character.
- The orawrite operator will cause RETL to abend with a return code != 0 only in the case that sql\*loader fails. Should sql\*loader exit prematurely after having reached its maximum number of allowed rejected records, RETL will not abort. Adding this functionality is an enhancement and logged as #579.
- Join operators may use a large amount of memory when the distribution of key-equivalent records is low and the number of records is high. It is important to test join operations with appropriate data volumes on each input side to verify the maximum amount of memory your application may require.
- DB2 users need to have the db2 database located on the same server as RETL because the port is not configurable. Bug #510

## Backwards-Compatibility to 10.x versions

- RETL 11 versions are stricter when enforcing valid XML flow interfaces, input schemas, and data fields. In some instances, the RETL 11 product will find data and/or flow errors that may have previously been unreported. The result may be more discarded or rejected records and/or error messages than previously identified.
- An unintended feature of 10.x allowed input datafiles to contain more fields than are specified in the schemafilename. RETL 11 does not allow more fields to exist in the datafile and will throw an exception should this happen. This is logged as Bug #632 and will be fixed in a future version of RETL 11.
- Output field order between 10.x and 11.x may be different for those flows that don't specify a 'schemafilename' property in the export operator.

- RETL 10.x versions always format float fields with 6 digits of precision. In 11.x versions, the precision is the same as the input data to RETL. This hasn't been shown to cause any issues in certification efforts, but it should be noted as a behavioral difference. This is documented as Bug #441.
- RETL 11.x versions require additional properties for dbread/dbwrite operators. 'hostname' and 'port' are now required properties and may be specified in rfx.conf for convenience. See the RETL 11.1 Programmer's Guide for more information on syntax and usage.
- When upgrading an oracle database, a separate instance of RETL should be used in order to test the new database. This is because only one port can be specified in rfx.conf at a time. The other option is to use RETL with the '-c' option to use a different config file. See the RETL 11.1 Programmer's Guide for more information.
- RETL 11.x versions may require more physical memory for certain flows than 10.x releases.
- Certification with DB2 was not performed as part of the RETL 11.1 certification effort. Currently, no Retek clients are using RETL with DB2 so the determination was made to re-schedule the DB2 certification effort of DB2 v8 for the next generally available release of RETL.

### Known issues specific to RETL 11.1

- Bug #407 - RETL 11.1 does not reject numeric fields based on maximum values for each particular RETL datatype. For example, the value '100' will not be rejected for an 'int8' datatype.
- Bug #461 – RETL 11.1 should provide a more specific error messages with more record and field level details.
- Bug #541 - Incorrect Handling of Too Few Fields in Import. RETL 10.x will reject incoming records with too few fields. 11.x versions will treat the missing fields as null values. This will be fixed to reject records in future versions of RETL.
- Bug #632 - An unintended feature of 10.x allowed input datafiles to contain more fields than are specified in the schemafile. RETL 11.1 does not allow more fields to exist in the datafile and will throw an exception should this happen.
- Bug #646 - Terawrite performs a 'delete' on a table when terawrite's 'mode' property == 'truncate'. This results in slow truncation on large tables.
- Bug #652 - NOT Can Only be Used Once in Filter Expression. The workaround is to reverse all other operations to obtain the same effect. For example, "not STATUS\_CDE eq 'A' and not STATUS\_CDE eq 'C' and not STATUS\_CDE eq 'X' becomes "STATUS\_CDE ne 'A' and STATUS\_CDE ne 'C' and STATUS\_CDE ne 'X'".
- Bug #678 - Multibyte fixed-width exports are slower than single-byte fixed-width exports.