

*Ardent™ DataStage® Suite, the leading data warehouse creation and management solution, is now the premier e-business information infrastructure. The DataStage Suite provides data movement, meta data management, and data quality assurance to more than 1,200 customers worldwide. It's a complete, modular solution offering:*

- *native access to data from mainframe, Unix, and Windows NT environments;*
- *data quality assurance capabilities; and*
- *full life-cycle meta data management.*



Web server logs reveal who visits a site, what they look at, where they travel from click to click, and so on. When coupled with back-end transactional data, these files are an invaluable source of information for e-commerce purposes.

Specialized "clickstream" tools for extracting and analyzing this wealth of data help e-businesses manage their Web sites by enabling them to understand site usage patterns. A significant drawback of these tools, however, is that the information they provide can't be integrated with other types of data to present a complete picture of the e-business. As a result, the exciting potential for true e-commerce customer intimacy remains largely unfulfilled.

**ClickPack**, a software option for DataStage Suite, unlocks the power to realize this potential. ClickPack automatically extracts and transforms data from Web server log files and e-mail systems, and offloads it to a data warehouse. By integrating online data with existing customer, product, and company data, ClickPack lets your sales and marketing professionals compare online customer demographic information against traditional demographics, or use online information to augment their existing customer knowledge base. In short, ClickPack collects the information that helps you turn visitors into customers.

#### **Clickstream Analytics**

Supplemented with ClickPack, DataStage can perform five essential Web-analytic functions:

##### **Reading, extracting, and parsing data**

DataStage supplies specific stages written to extract data from the most common Web servers, including Netscape, Microsoft IIS, and Apache. These stages parse data into IP address, URL, referral URL (RURL), time, event type, and other components. In addition, the stages immediately filter out unwanted data such as image requests or Web Crawler-generated records, greatly improving performance by reducing data volumes prior to movement and transformation.

##### **Aggregating and calculating**

Once the data has been parsed, the Web server log files are sorted and specific functions calculated—such as how many unique users visited a specific page, and what a user did during a session or visit to a Web site.

##### **Categorizing and grouping**

ClickPack can also create derived columns of information to provide additional grouping and categorizing of the raw log data. These might include Content area groups, Registration categorization, and so on.

**ClickPack collects the information that helps you turn visitors into customers. ClickPack:**

- adds clickstream analysis capability to market-leading data warehouse creation and management solution;
- extracts data from Microsoft IIS, Netscape, and Apache log files;
- extracts e-mail data from POP3 and IMAP protocols; and
- transforms data using new built-in transformations or with existing Perl functions.

## ClickPack

### ClickPack Functionality

- Log Reader Stage for Netscape
- Log Reader Stage for Microsoft IIS
- Log Reader Stage for Apache
- Log Parsing Functions and Transforms
- Perl Function Stage
- E-Mail Reader Stage for POP3
- E-Mail Reader Stage for IMAP
- E-Mail Reader Stage for SMTP

### Converting and validating IP addresses

Converting IP addresses to domain names lets analysts more easily group users within the same overall domain and see where their users come from. Domain names are validated for accuracy before becoming part of the data warehouse.

### Integrating Web and other enterprise data

Once data has been fully extracted and transformed into usable information, it can be integrated with any other supported DataStage source, including legacy data, Enterprise Resource Planning (ERP) data, and customer databases.

### E-Mail Transformation

ClickPack makes DataStage the first data integration solution to offer e-mail extraction and parsing.

#### E-mail extraction

ClickPack lets DataStage attach to, and extract e-mails from, e-mail servers. Standard Post Office Protocol 3 (POP3) and Internet Mail Access Protocol (IMAP) protocols are supported.

#### E-Mail parsing

Users can parse e-mails by To, From, Subject, Date, and Text. The body of the text can also be parsed through a keyword search function.

#### Embedded e-mail parsing

Embedded messages, such as replies or forwards, can be parsed into either additional columns or separate rows per embedded e-mail.

### Perl Support

ClickPack provides enhanced native support for the popular Perl programming language via a graphical interface. Data from Web server logs and e-mail systems can be parsed either with new, built-in transformations or with user-written Perl functions. Existing or new Perl transformations can be seamlessly integrated into any DataStage activity.

### About Informix

Based in Menlo Park, CA, Informix Software is the technology leader in software infrastructure solutions for the Internet. Informix is the first and only company to integrate e-commerce and business intelligence on a true Internet infrastructure. We provide a complete, fast and simple way to bring businesses to the Web, personalize content management, and analyze information real-time. This, with our highly scalable Web engines and media asset management capabilities, gives our customers a unique competitive advantage. For more information, contact the sales office nearest you or visit our Web site at [www.informix.com](http://www.informix.com).

**Informix**  
SOFTWARE  
way to web

4100 Bohannon Drive  
Menlo Park, CA 94025  
Tel: 650 926-6300  
[www.informix.com](http://www.informix.com)

#### INFORMIX REGIONAL SALES OFFICES

Asia Pacific	65 259 1715	Japan	01 3 3262 4500
Canada (Toronto)	416 730 9009	Latin America	300 591 8982
Europe/Middle East/Africa	44 208 818 3000	North America	800 966 9815
Federal	703 847 2900		650 926 8300

© 2000 Informix Corporation. All rights reserved. The following are trademarks of Informix Corporation or its affiliates, some or none of which may be registered in the U.S. or other jurisdictions: Informix®, way to web®, Arlink®, and DataStage®.

Printed in U.S.A. 400  
800-2207878