ACTUATE
The BIRT Company™

Serving 40 Million Financial Customers in Minutes with BIRT iHub

Actuate powers a scalable, high-performance application that creates and delivers personalized reports

A prospective customer challenged Actuate to prove that an application powered by BIRT iHub could prepare and present monthly transaction statements for millions of users on a tight deadline. Because high performance and linear scalability are hallmarks of solutions powered by BIRT iHub, Actuate accepted the challenge.

Table of Contents

Executive Summary

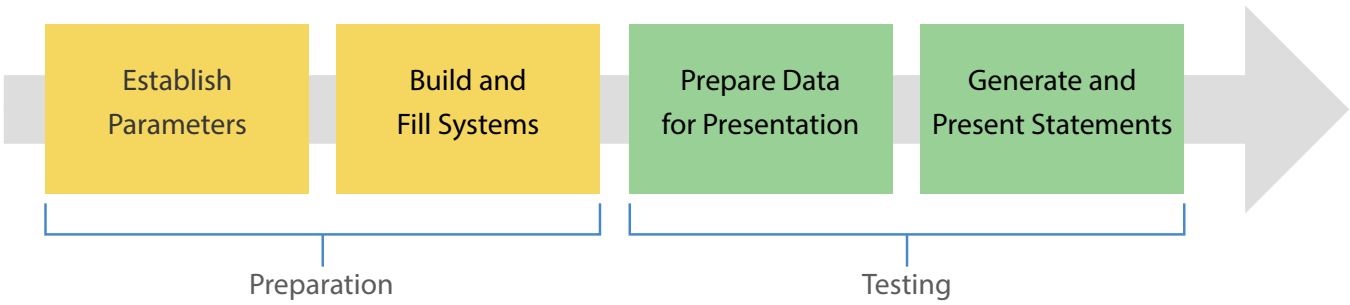
Actuate was challenged by a global financial services company to create a test application that would generate monthly transaction statements for 40 million users in five days, and then present those statements in near-real time. Because many users access their statements in the first two days after data is available, the statement presentation portion of the application needed to respond to a high peak workload.

User statements could range in length from a few entries to hundreds of pages. The company's customers reside in many countries, so the application needed to support multiple currencies. All statements needed to be available to users for seven years after creation. The application would potentially handle more than 200 million users, so the solution needed to scale.

Actuate met this challenge with an application built using BIRT iHub. Actuate proved that a simple, clustered four-server BIRT iHub system could meet all of the customer's data-preparation and statement-presentation goals – including handling the short-term spike in statement delivery – in less than two days, and easily scale to many more users.

A Project in Four Parts

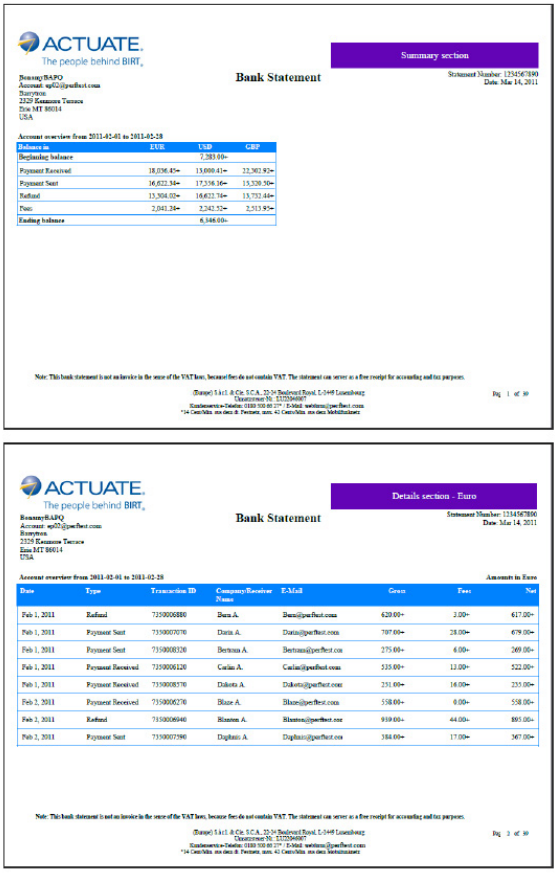
After establishing parameters for the test, Actuate engineers assembled an infrastructure and created data sources for the test. The testing itself was done in two phases: data preparation, followed by statement generation and presentation.



Assumptions and Expectations

The following test parameters were established based on guidance from Actuate's prospective customer:

- 40 million user accounts, complete with logins and one month's worth of transaction details in multiple currencies, were created.
- Those 40 million users would generate a total of 4.2 billion transaction records (686 GB of data) per month.
- Transaction data would be organized and prepared for statements in a batch within five days of the end of each month.
- Statements would be created and presented when individual users requested them.
- Each statement would have a summary table showing the beginning balance, total transactions grouped by transaction type and currency type, and ending balance. The statement would also present transaction details grouped by currency and sorted chronologically. Each page of the statement would include the client's name and address, a header with a company logo, and footer with boilerplate text. Figure 1 shows sample statement pages.
- Statements would be accessible to users on demand for a minimum of seven years.


ACTUATE
The people behind BIRT

Summary section

Bank Statement

Statement Number: 1234567890
Date: Mar 14, 2011

BonanyBAPO
Account: sql2@perfect.com
Bonytron
2320 Kenmore Terrace
Erie MT 86014
USA

Account overview from 2011-02-01 to 2011-02-28

Balance in	EUR	USD	GBP
Beginning balance		7,383.00-	
Payment Received	18,036.45-	13,000.41-	22,302.92-
Payment Sent	16,622.34-	17,536.16-	15,320.50-
Refund	13,304.02-	16,621.74-	13,732.44-
Fees	2,041.34-	2,242.52-	2,513.97-
Ending balance		6,346.00-	

Note: This bank statement is not an invoice in the sense of the VAT laws, because they do not contain VAT. The statement can serve as a free receipt for accounting and tax purposes.

(Banque) S.à r.l. & Cie. S.C.A., 23-34 Boulevard Royal, L-2449 Luxembourg
Unternehmen-Telefon: 0180 500 00 27* | E-Mail: webteam@perfleet.com
*14 Cent/Min. aus dem dt. Festnetz, max. 42 Cent/Min. aus dem Mobilfunknetz

Pg. 1 of 10

ACTUATE
The people behind BIRT

Details section - Euro

Bank Statement

Statement Number: 1234567890
Date: Mar 14, 2011

BonanyBAPO
Account: sql2@perfleet.com
Bonytron
2320 Kenmore Terrace
Erie MT 86014
USA

Account overview from 2011-02-01 to 2011-02-28

Amounts in Euro

Date	Type	Transaction ID	Company/Receiver Name	E-Mail	Gross	Fees	Net
Feb 1, 2011	Refund	7350006880	Bern A.	Bern@perfleet.com	620.00-	3.00-	617.00-
Feb 1, 2011	Payment Sent	7350007070	Darin A.	Darin@perfleet.com	707.00-	28.00-	679.00-
Feb 1, 2011	Payment Sent	7350008320	Bertram A.	Bertram@perfleet.com	275.00-	6.00-	269.00-
Feb 1, 2011	Payment Received	7350006120	Carlin A.	Carlin@perfleet.com	535.00+	13.00-	522.00+
Feb 1, 2011	Payment Received	7350008570	Dakota A.	Dakota@perfleet.com	251.00+	16.00-	235.00+
Feb 2, 2011	Payment Received	7350006270	Blane A.	Blane@perfleet.com	558.00+	0.00-	558.00+
Feb 2, 2011	Refund	7350006940	Blanton A.	Blanton@perfleet.com	919.00+	44.00-	893.00+
Feb 2, 2011	Payment Sent	7350007590	Daphani A.	Daphani@perfleet.com	384.00+	17.00+	367.00+

Note: This bank statement is not an invoice in the sense of the VAT laws, because they do not contain VAT. The statement can serve as a free receipt for accounting and tax purposes.

(Banque) S.à r.l. & Cie. S.C.A., 23-34 Boulevard Royal, L-2449 Luxembourg
Unternehmen-Telefon: 0180 500 00 27* | E-Mail: webteam@perfleet.com
*14 Cent/Min. aus dem dt. Festnetz, max. 42 Cent/Min. aus dem Mobilfunknetz

Pg. 2 of 10

Figure 1: A sample statement generated by the application includes an account summary, transaction details, logos and boilerplate text.

- Statements would vary in size according to this distribution:

Population	Database size	Page Count	PDF Document Size
94% of users	75 rows	12 Pages	30 KB
5% of users	300 rows	40 Pages	80 KB
1% of users	2,000 rows	250 Pages	400 KB

- 20 percent of the user population – 8 million users in all – would access their statements within two days of statements becoming available.
- During those two days, 85 percent of users would access and view their statements online in HTML format, while 15 percent would request statements as PDFs. The application would therefore need to generate 1.2 million PDFs in the first two days.
- Two-thirds of users in the first two days – 5.3 million users total – would retrieve their statements on the first day of availability. The time distribution of first-day users would follow the curve in Figure 2. The peak number of simultaneous active users (i.e., the peak system load) would be 439,000.

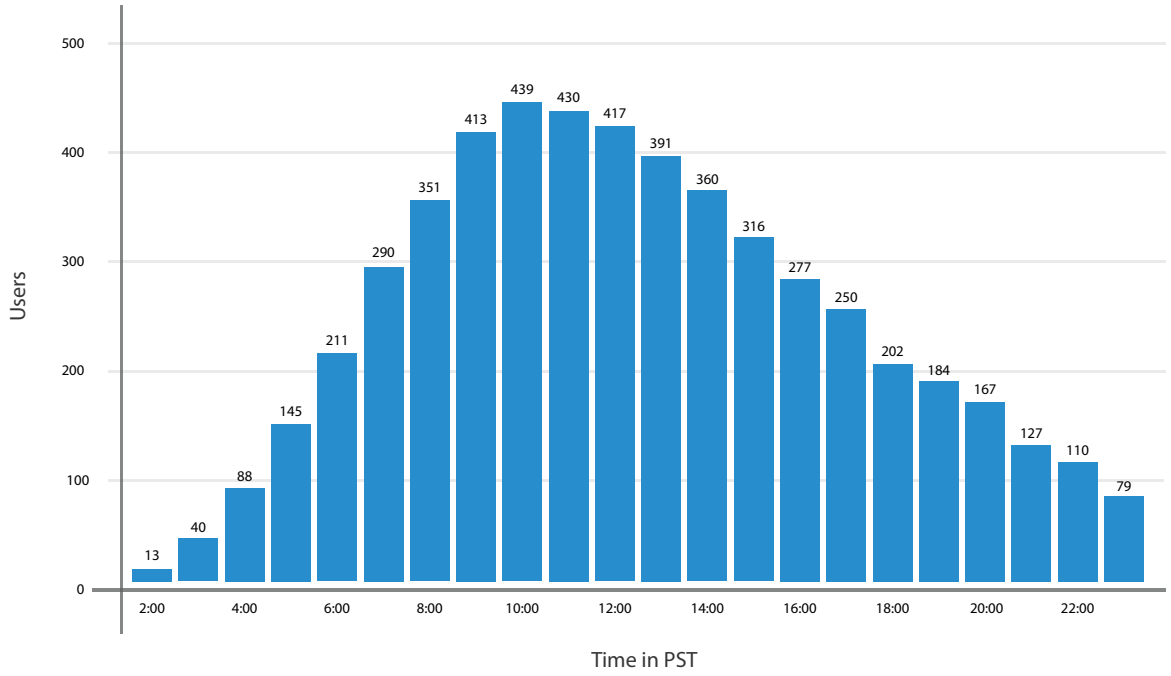


Figure 2: Anticipated hour-by-hour distribution of users accessing statements on the first day of availability.

Test Preparation: Building a 40 Million-User Architecture

Actuate created a scalable architecture based on industry-standard hardware and BIRT iHub software. Certain components of the software (such as a database for metadata) were assigned their own hardware to reduce bottlenecks and resource contention.

The following systems comprised the test infrastructure:

- One four-node cluster of BIRT iHub (*iHub Cluster* in Figure 3).
- One ETL database (*ETL Data* in Figure 3) to host transaction data in comma separated value (CSV) format for 10 million customers. The same data was used four times to create simulated records for 40 million users. The ETL database would hold more than a billion rows of transaction records.
- One iHub Encyclopedia database (*iHub DB* in Figure 3) to host metadata such as file access rights and file locations. Postgres was used for the database.

What is BIRT iHub?

BIRT iHub is the commercial platform for all BIRT content. The out-of-the-box deployment server removes the burden on developers to manually build infrastructure components, and automatically enriches content for end users with interactivity and personalization. Learn more at actuate.com/products/birt-ihub-visualization-platform/

What is an iHub Encyclopedia?

An iHub Encyclopedia is a partitioned set of content, users, and permissions that encapsulate an application. In a production or OEM environment, the Encyclopedia can have multiple volumes, each with unique content, users and permissions. Actuate supports DB2, Microsoft SQL Server 2008, Oracle and Postgres for the Encyclopedia. The iHub Encyclopedia database can be replicated for fault tolerance.

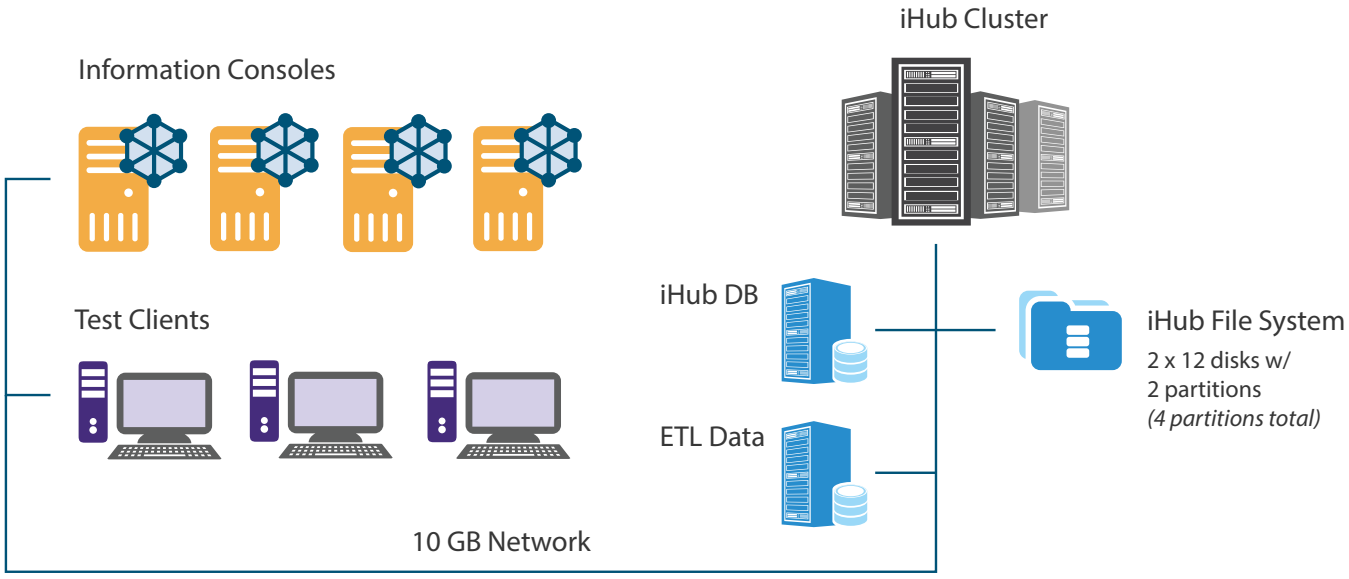


Figure 3: The infrastructure built by Actuate for the test.

- Hardware for the iHub cluster, ETL database and iHub Encyclopedia database was:
  - Dual-processor server, 6 cores per processor, Intel Xeon X5650, 2.66 GHz
  - 48 GB RAM
  - SPECint 2006 rating: 353
  - Red Hat Linux operating system
- Two storage systems (*iHub File System* in Figure 3) to hold of BIRT Data Objects and generated reports.
  - 600 GB SAS (Serial Attached SCSI) disk drives, 6 gigabit per second I/O
  - 12 disk drives per file system
  - Disks in each file system were arranged into two partitions (four partitions total)
- Four load-balanced Information Consoles (*Information Consoles* in Figure 3) providing portal access to the iHub Cluster
- Three load-generating clients (*Test Clients* in Figure 3) to generate simulated user actions
- One 10-gigabit network

Note that BIRT iHub is responsible for both preparing transaction data to be statement-ready, and for generating and presenting the statements themselves. These two processes exploit different capabilities of BIRT:

- *BIRT Factories* build BIRT Data Objects from CSV files in the data preparation phase.
- *BIRT Designer Professional* and *BIRT Viewing Services* structure data and create PDF and HTML documents during the statement generation phase.

What is the Information Console?

The Information Console provides an interface that enables users to browse, run, view and organize reports. It seamlessly integrates BIRT content within any webpage, following the site's look and feel.

What is a BIRT Data Object?

The BIRT Data Object (BDO) is an efficient, compressed data format. BDOs are normalized and indexed; they do not require a database to read, and are built for analytics. For these reasons, BDOs provide an ideal format for archiving monthly transaction data – a key element of the customer's requirements.

Test Phase 1: Data Preparation

The first test phase was designed to prove that Actuate could prepare a month's worth of account data for presentation in an acceptable timeframe. Five days (7,200 minutes) was chosen as a goal.

The test followed a procedure common to financial institutions: An extract, transform, load (ETL) process would pull a month's data from a production database in comma-separated value (CSV) format. At many institutions this large CSV file is loaded into a separate data warehouse for statement generation. However, because BIRT supports CSV files directly as data sources, the data warehousing step was unnecessary.

The test CSV file was 686 GB in size and held 4.2 billion rows of raw transactional data. From this CSV file, BIRT iHub produced 56 BIRT Data Objects (BDOs), each containing approximately one million user statements and ranging in size from 11 to 13 GB. The total size of all BDOs was 718 GB.

Actuate employed BIRT's BDO caching utility to prepare and cache statement data in memory for faster access.

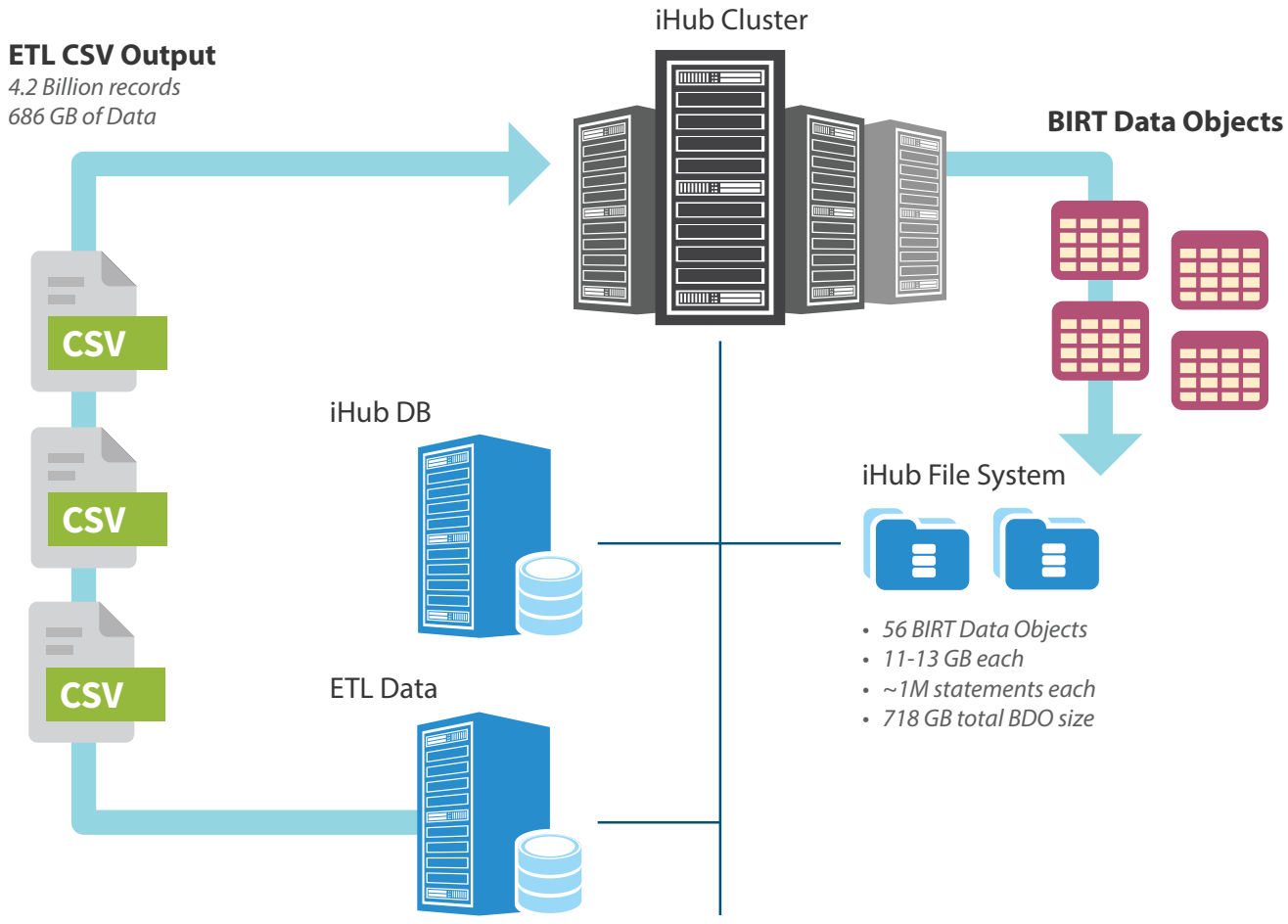


Figure 4. Data flow for Phase 1 of Actuate's test application.

Two rounds of tests were conducted in Phase 1, with each round run first on a two-node iHub cluster and then on a four-node iHub cluster. Each iHub node was configured to run eight BIRT Factories. The goal of running both two-node and four-node tests was to prove that the task could be completed successfully with two nodes, and also to show that BIRT performance scaled (with additional hardware) to support the customer's eventual goal of supporting 200 million users.

Round one of the test involved organizing the data by account but not sorting the data within the BIRT Data Objects. This table shows the time and throughput for the task, and confirms BIRT's linear scalability.

Round one: Statements Organized by Account		
Configuration	**Time**	**Throughput (MB/Sec)**
2 iHub cluster, 16 BIRT Factories	135 mins	88 MB/sec
4 iHub cluster, 32 BIRT Factories	69 mins	176 MB/sec

Round two of the test organized the data (from the CSV file) by account and then sorted it by date before creating BIRT Data Objects. Sorting, which makes data in the BDO easier to retrieve and use for calculation, was performed by BIRT iHub. Sorting requires considerable computing power, so adding sorting to the BDO creation process was intended to test iHub performance. This table shows the time and throughput for the task.

Round two: Statements Organized by Account and Sorted		
Configuration	**Time**	**Throughput (MB/Sec)**
2 iHub cluster, 16 BIRT Factories, 1 CSV file server	155 mins	74 MB/sec
4 iHub cluster, 32 BIRT Factories, 1 CSV file server	115 mins	99.2 MB/sec

***CSVs were on single NFS server, disk I/O saturated.*

Round two of the test identified a bottleneck unrelated to the sorting process: The single file system serving CSV data became saturated – that is, it was operating at peak I/O capacity. The BIRT engines could consume CSV data faster than the storage could serve it. Disk I/O is common bottleneck for systems working with large data volumes.

Actuate added a second CSV file system and reran the test. The bottleneck disappeared and the system returned to linear scalability. Solid-state drives would also eliminate the disk I/O problem, but at higher cost. This table shows the time and throughput for the task with the second file system in place.

Round two: Statements Organized by Account and Sorted		
Configuration	**Time**	**Throughput (MB/Sec)**
4 iHub cluster, 32 BIRT Factories, 2 CSV file servers	75 mins	152.4 MB/sec

The goal of Phase 1 was to prepare the data for presentation in 7,200 minutes. The efficiency of BIRT iHub enabled the process to be completed in 75 minutes or less using two clusters. Storage speed affected performance when iHub also sorted the CSV data before creating BDOs, but adding a second file server removed that bottleneck. Results are consolidated in Figure 5.

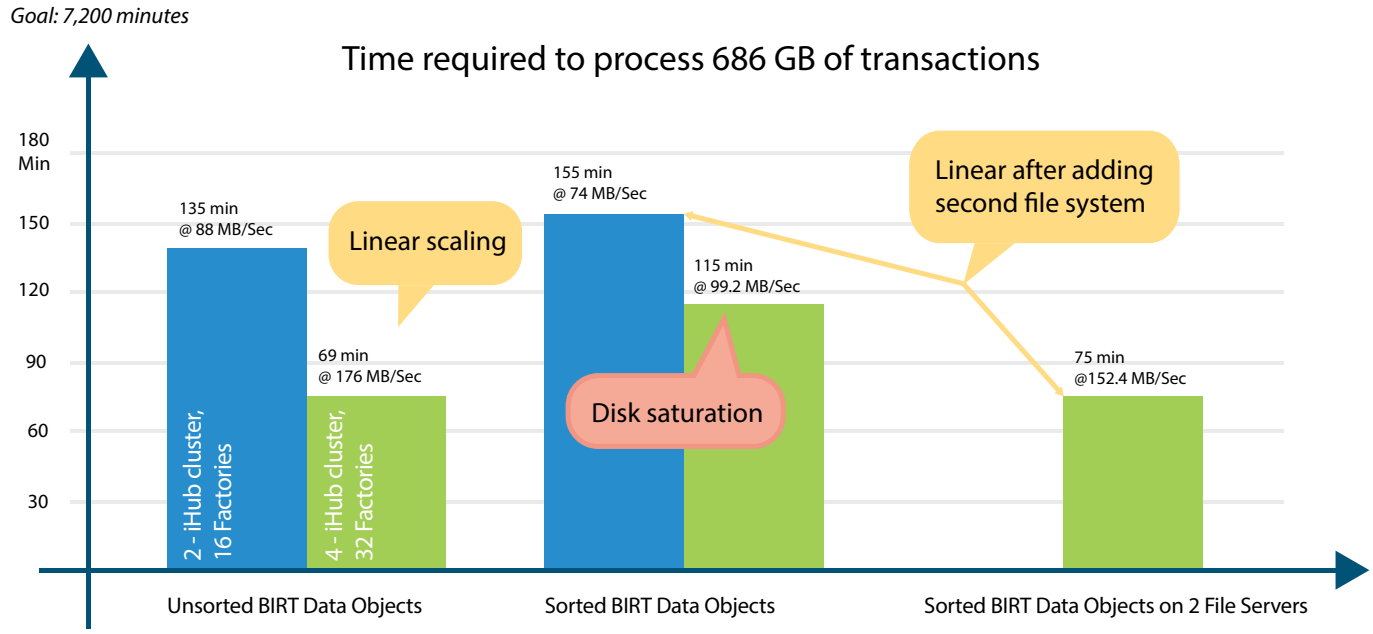


Figure 5: Phase 1 test results, showing iHub linear scaling in processing CSV data into BIRT Data Objects.

Test Phase 2: On-Demand Generation and Viewing

The second phase of the test measured system performance when generating and presenting statements to users. During this phase, three test clients generated thousands of simulated user sessions at a rapid pace. Each simulated user performed the following actions:

1. Log in to the website. The test platform included integration between website login credentials and BIRT iHub to replicate a single sign-on user experience.
2. Request the statement report. This action invoked an Actuate on-demand request to generate and display the report.
3. View the first page of the statement in HTML via Actuate Viewer (85 percent of users); *or* Download the full statement as PDF (15 percent of users).
4. Log out.

Population	Database size	Page Count	PDF Document Size
94% of users	75 rows	12 Pages	30 KB
5% of users	300 rows	40 Pages	80 KB
1% of users	2,000 rows	250 Pages	400 KB

As previously noted, user statements varied in size according to this distribution:

The goal of the test was to deliver 439,000 statement views per hour (122 views per second) to meet the expected first day peak load (see Figure 2).

Four tests, each lasting one hour, were conducted using the entire complement of 56 BIRT Data Objects. The first test used one iHub and one Information Console. Each subsequent test added one iHub and one Information Console to the cluster. In all tests, each iHub drove four Viewing Services. Details on the test are shown in the following table.

Configuration IC x IH, VS	Throughput	Users per hour	Response time	CPU iHub	CPU IC	Errors	Error Rate
1x1, 4 VS	31 views/sec	111K	0.98 sec	79%	22%	42	0.0004%
2x2, 8 VS	66 views/sec	238K	0.92 sec	80%	24%	88	0.0004%
3x3, 12 VS	99 views/sec	358K	0.93 sec	79%	24%	131	0.0004%
4x4, 16 VS	131 views/sec	473K	0.94 sec	80%	25%	176	0.0004%

IC = Information Console, IH = iHub, VS = Viewing Services

A cluster of four BIRT iHub servers and four Information Consoles delivered 131 views per second, beating the goal of 122 views per second by more than 7 percent. iHub CPUs never exceeded 80 percent of capacity.

Additionally, tests showed that BIRT iHub scaled linearly as iHub servers and Information Consoles were added (see Figure 6).

Errors were documented and error rates were measured during the hour-long runs. Users who experienced errors re-executed their requests until the views were successfully completed. The very low error rate remained consistent as the iHub cluster grew.

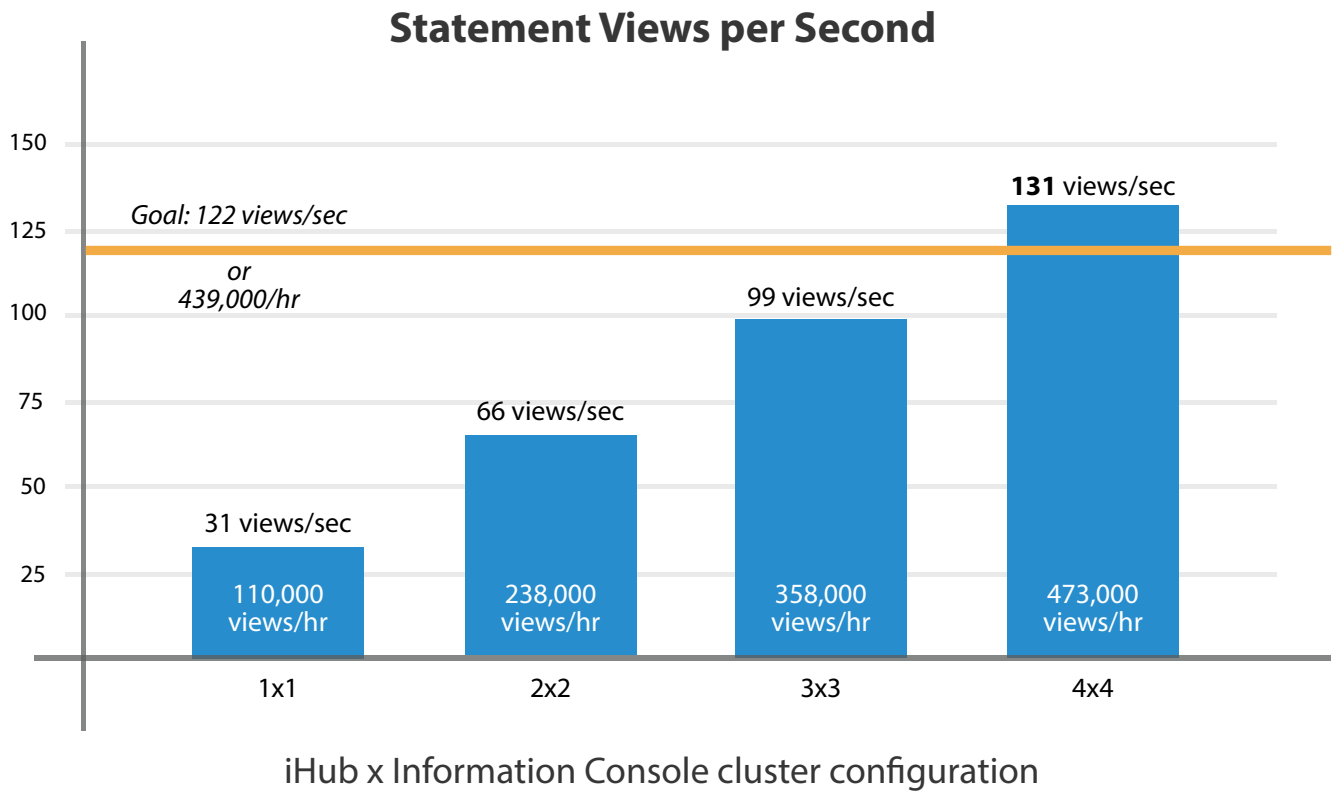


Figure 6. Phase 2 test results, showing iHub linear scaling in delivering statement views.

Conclusions

Actuate engineers were challenged to meet specific scalability and performance goals using BIRT. The test showed that BIRT iHub met all objectives:

- In a four-server cluster, BIRT iHub easily handled user and transaction data for 40 million accounts. All BIRT components proved to be linearly scalable and easily able to handle both more users and more transactions.
- BIRT iHub was very fast at preparing CSV data for presentation. Although 7,200 minutes were allotted to prepare 4.2 billion rows of data, iHub required less than 75 minutes for the task – just over one percent of the allotted time.
- BIRT iHub in a four-server cluster easily served expected workloads, including a first-day peak of 439,000 views per hour, with CPU capacity to spare. The peak workload included a variety of statement types and complexities.

BIRT iHub can meet demanding requirements for high-volume, high-velocity, rapidly scaling data preparation and presentation. This test was performed for a financial services company, but similar requirements exist in other markets and use cases, including healthcare, government, the Internet of Things, and customer-facing applications.

Download BIRT iHub

BIRT iHub is the commercial platform for all BIRT content. The out-of-the-box deployment server removes the burden on developers to manually build infrastructure components, and automatically enriches content for end users with interactivity and personalization.

actuate.com/resources/product-downloads/

Global Contacts

Actuate Corporation
Corporate Headquarters
951 Mariner's Island Blvd
San Mateo, CA 94404
Phone: (888) 422-8828
info@actuate.com
www.actuate.com

Actuate UK Ltd.
1st Floor Pickford House
46 Bow Lane
London, EC4M 9DL
United Kingdom
Phone:+44 (0)207 246 4700
ukinfo@actuate.com
www.actuate.com/uk/

Actuate (Deutschland) GmbH
Kaiserstrasse 44
60329 Frankfurt am Main
Germany
Phone:+49(0) 69 66 90 25-0
info@actuate.de
www.actuate.de

Actuate Spain
BIRT Analytics
Frederic Mompou 5
Edificio Euro 3
08960 Sant Just Desvern
Barcelona, Spain
Phone:(+ 34) 93 371 44 70
www.actuate.com/birtanalytics/

Actuate France
57 rue de Châteaudun
Paris 75009
France
Phone:+33 (0) 176 21 5650
fr_contact@actuate.com
www.actuate.fr

Actuate Australia Pty Ltd
Level 10, Suite 7
100 Walker Street
North Sydney
NSW 2060
Australia
Phone: +61 (2) 9900 6300
apac_marketing@actuate.com

Actuate Asia Pacific
8 Temasek Boulevard
#35-01, Suntec Tower Three
038988
Singapore
Phone:+65 6887 3777
apac_marketing@actuate.com
www.actuate.com/ap

Actuate Japan Co., Ltd.
アクチュエイトジャパン株式会社
〒102-0083
東京都千代田区麹町 5 丁目 1
NK真和ビル 9 階
Japan
Phone:(03) 5357-1631
jpinfo@actuate.com
www.actuate.com/jp

Actuate India
Level 8, Vibgyor Towers,
G Block C62,
Bandra Kurla Complex
Mumbai, 400051, India
Phone:+91-22-4090-7281
Fax:+91-22-4090-7283
india_sales@actuate.com

Notice

The information in this white paper is proprietary to Actuate Corporation ("Actuate") and may not be used in any form without the prior consent of Actuate.

Copyright © 2014 Actuate Corporation. All rights reserved. Actuate, legodo, BIRT iHub, BIRT iHub F-Type, BIRT Analytics, Actuate Customer Communications Suite, The Actuate Document Accessibility Appliance, BIRT onDemand, BIRT Viewer Toolkit, and the Actuate logo are trademarks or registered trademarks of Actuate Corporation and/or its affiliates in the U.S. and certain other countries. The use of the word "partner" or "partnership" does not imply a legal partnership relationship between Actuate and any other company. All other brands, names or trademarks mentioned may be trademarks of their respective owners. 208800