F5 Networks, Inc.

# F5 Wan Optimization for Oracle Database Replication Services

## Faster Replication across the WAN

Chris Akker – Solution Architect - F5 Networks

2011

# Executive Overview

Protection of business data is important to companies all over the world. Oracle and F5 have partnered together to create a compelling solution with Oracle Database 11g Database Replication Services ( Data Guard, GoldenGate, Recovery Manager, Streams ) and F5 Wan Optimization technologies. By using these technologies together, mission critical data can be replicated across Wide Area Networks between data centers, in less time while using less network bandwidth. Combining the best practices and latest products from both Oracle and F5 Networks gives your company the database and network needed to run your business non-stop. This technical brief will detail the business problem, testing, and best practices of this collaborative solution, providing the IT Agility your business needs.

# Introduction

### Oracle Database Replication Services

Oracle offers several replication technologies, each optimized for particular purposes.

Oracle Data Guard is Oracle's strategic solution for Oracle database protection and data availability. Data Guard provides the management, monitoring, and automation software to create and maintain one or more standby databases to protect Oracle data from failures, disasters, human error, and data corruptions while providing high availability for mission critical applications. Data Guard is unique among Oracle replication solutions in offering both synchronous zero data loss and asynchronous replication options.

Oracle GoldenGate is Oracle's strategic solution for advanced replication requirements and for best-in-class real-time heterogeneous data integration and continuous data availability. GoldenGate captures and delivers updates of critical information as the changes occur and provides continuous data synchronization.

Recovery Manager (RMAN) is fundamental to every Oracle Database installation and is used to backup and restore databases, and make additional static duplicates of production data when needed for various purposes (e.g. Data Guard standby database instantiation, initial GoldenGate replica creation, or clone databases for development and test)

Oracle Streams is a legacy Oracle replication product for which Oracle continues to provide support for current and future versions of the Oracle Database to protect customer investment in applications built using this technology.

## The Challenge

The use of these Oracle replication services is often limited by Wide Area Network bandwidth, latency, and packet loss problems. The transfer of large amounts of data over the WAN has always been a battle, which can create a nightmare for the DBA responsible for meeting Recovery Point and Time Objectives (RPO/RTO). Whether you need copies of databases for business continuity and disaster recovery, compliance and reporting, performance and scaling, or other business needs, the WAN needed to handle all this data is expensive. Often, the speed of the WAN is not fast enough to replicate the volume of data in the time window needed. If you do have enough bandwidth, you need to make sure you make the most efficient use of it. Upgrading existing WAN bandwidth is very expensive and the recurring costs can quickly consume IT network budgets.

## The Solution

The solution to these WAN bandwidth challenges and associated costs is the F5 Networks Wan Optimization Module (WOM). WOM can accelerate these Oracle database replication technologies across a Wide Area Network, securely transmitting more data while using less bandwidth, and be less susceptible to latency and packet loss. Data Guard, GoldenGate, Recovery Manager, and Streams can be run across the WAN more efficiently, while reducing network load, reducing time, and offloading CPU intensive compression and encryption from the primary database server. F5's BIG-IP platforms provide the necessary processing power to handle network services like SSL encryption, data compression, de-duplication, and TCP/IP network optimizations.

Because these CPU hungry network services are running off-host, this saves valuable computing power on the database, freeing resources for what it does best – process the database needs of the organization. And for the DBA, this will help lower and maintain established RPO/RTO requirements for mission critical data. Using these two technologies together provides a solid foundation for your Oracle database infrastructure, which can now be higher performance, provide faster recovery objectives, move more data – all while saving money on expensive bandwidth.

## F5 Networks Wan Optimization Services

F5 Networks Wan Optimization Module (WOM) brings state of the art networking to the Wide Area Network. Using advanced TCP/IP enhancements, compression, de-duplication, and SSL acceleration, the BIG-IP LTM creates a secure iSession tunnel between data centers, and provides LAN like performance across the WAN. As the cost of bandwidth and the need for data transmission increases, an efficient network transport is required to run applications and move data between data centers or in the cloud.

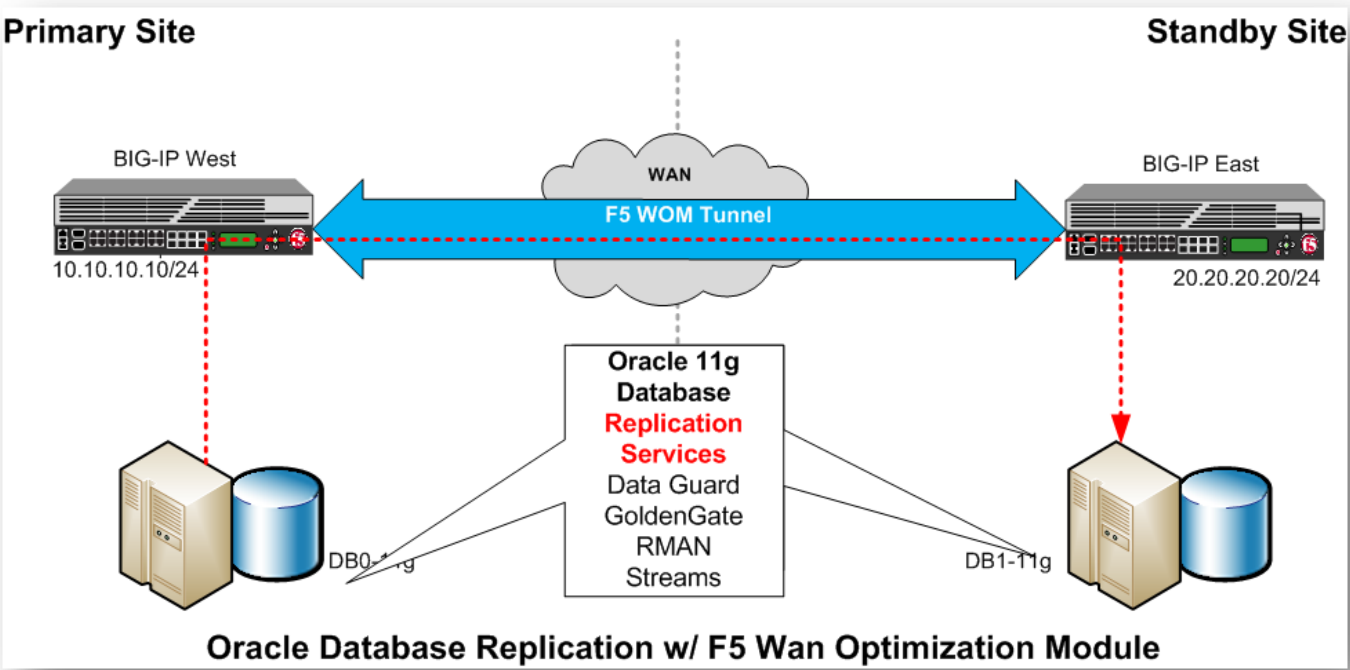
The remainder of this technical brief will outline the testing and results of using both the Oracle technologies and F5 Networks Wan Optimization technologies together. This will provide a solution for DBAs, Network Architects, and IT Managers that are challenged to meet the needs of the organization and control costs at the same time.

## Test Network Architecture Overview

In order to properly test these Oracle services with WOM, a test network with typical WAN link speeds, latency values, and packet loss was needed. A LANforge WAN simulation appliance was used to create the Wide Area Network. Also, a test tool was needed to generate load on the Primary database server, to create the data replication workloads.

A Test Network was built, using the following equipment and tools.

- Two Oracle 11gR1 Database Servers, one Primary, one Secondary. Stand-alone database servers were used, RAC enabled databases are beyond the scope of this paper.
- Two F5 BIG-IP Model 3900 LTM devices, running Version 10.2 software.
- One LANforge 500 WAN simulation device.




Oracle Database Servers were configured as follows:

| Hostname | Software | Hardware | OS | DB Role |
|---|---|---|---|---|
| DB0-11g | Oracle 11gR1 | VMs | Oracle Ent Linux | Primary |
| DB1-11g | Oracle 11gR1 | VMs | Oracle Ent Linux | Standby |

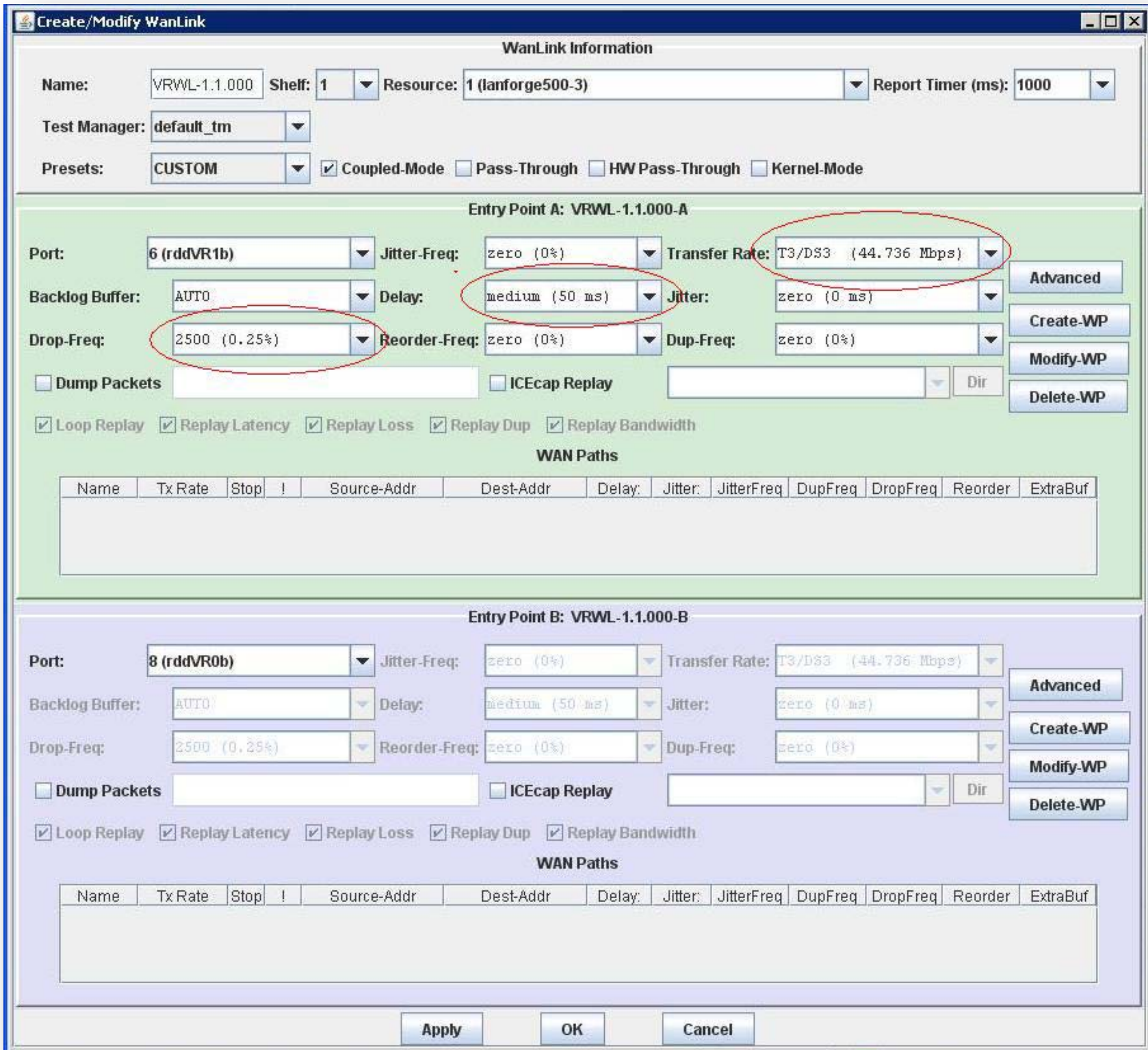
F5 BIG-IP LTM devices were configured as follows:

| Hostname | OS | Module | Location |
|---|---|---|---|
| BIG-West | LTM v10.2, RTM Build | WOM enabled | Primary Site |
| BIG-East | LTM v10.2, RTM Build | WOM enabled | Standby Site |

## LANforge WAN Configuration

A LANforge 500 WAN simulation device was placed into the test network, and configured as follows:

- 45mb/s Bandwidth Network Link
- 100ms RTT delay ( 50ms each direction )
- .5% Packet Loss ( .25% each direction )




## Oracle Net Configuration

When performing the test cases, it is important to document the calculations and settings for the Oracle Net TCP/IP stack, also commonly called SQL.NET. The calculation used is called the Bandwidth Delay Product (BDP), and is used to calculate how large the send and receive buffers need to be, in order to achieve optimal TCP/IP performance.

Per Oracle best practices, the optimal socket buffer size is three times the size of the BDP (as of Oracle Database 11g, this best practice is updated to set the socket buffer size at the larger of three times BDP or 10Mbytes) To find the BDP, the bandwidth of the link and the network Round Trip Time (RTT) are required. RTT is the time required for a packet to travel from the Primary database to the Standby database and back, expressed in milliseconds (ms). The response time value was taken from a series of PING packets done over 60 seconds, and then using the average millisecond value.

First, we calculate the BDP, as follows:

**BDP= Link Speed * RTT**

Second, we note the SQL.net `RECV_BUF_SIZE` and `SEND_BUF_SIZE` parameters equal to 3 times the *Bandwidth Delay Product* (BDP). This will produce the largest increase in network throughput.

**TCP BuffSize = Link Speed * RTT / 8 * 3**

Third, we use the Oracle Net Session Data Unit (SDU) size of 32767.

In the `sqlnet.ora` file for our test harness, the following was changed:

```
DEFAULT_SDU_SIZE=32767
RECV_BUF_SIZE=N
SEND_BUF_SIZE=N
Where N = WOM TCP Buffer Settings in the following Table.
```

There were different TCP/IP Profiles used for the testing. Some were used for the baselines, and some were used with the F5 WOM configurations. All of these TCP profile calculations were based on the formulas above as an Oracle Best Practice, as documented in the Oracle whitepaper " Data Guard Redo Transport & Network Best Practices." The TCP/IP settings were changed on both the Primary and Standby database servers. This is also considered a best practice, in case there is a Data Guard role change.

During additional testing at F5, it was discovered that when using the WOM transport, the TCP buffer settings could be increased even further to sustain higher levels of throughput. When using WOM, it is recommended that you double the Oracle best practice recommended value.

The following table summarizes the calculations for the 100ms test cases.

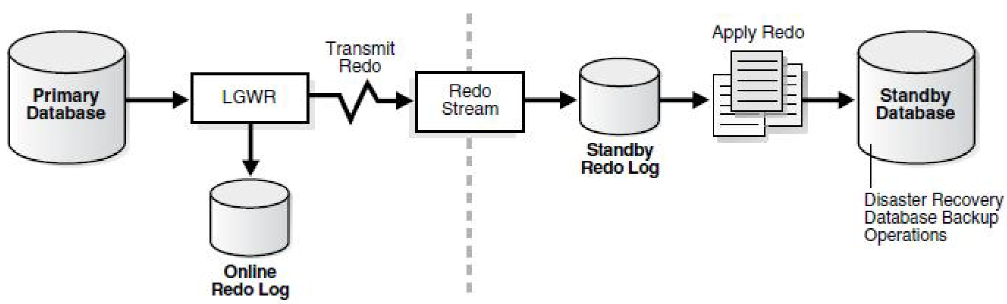
| Profile | Link | Latency | BDP Bytes | Oracle Buffer Setting | WOM TCP Buffer Setting |
|---|---|---|---|---|---|
| LAN-GigE-BDP | 1,000,000,000 | 0.0002 | 25,000 | 75,000 | 150,000 |
| Oracle Tuned-BDP | 45,000,000 | 0.1000 | 562,500 | 1,687,500 | 3,375,000 |
| F5-WOM-2xBDP | 90,000,000 | 0.1000 | 1,125,000 | 3,375,000 | 6,750,000 |

*NOTE: You must stop and restart both the Listeners and Databases, and perhaps others services, in order for the new settings to take effect.*

## Oracle DataGuard Overview and Configuration

Data Guard Zero Data Loss using Maximum Availability protection mode and synchronous redo transport was used exclusively in this test case. The design of Zero Data Loss is to ensure that any database change – insert, delete, or modify commit transactions, DDL operations, etc, are safely written to log files at BOTH the Primary and Standby Databases before commit acknowledgement is returned to the application. Data Guard insures this by synchronously transmitting database redo (the information used to recover any Oracle database transaction), directly from the primary database log buffer to the standby database, where it is applied using Oracle Managed Recovery (MRP). Data Guard is the only Oracle replication technology that can guarantee zero data loss protection. However, as the round trip network latency between the Primary and Standby increases, the acknowledgement response time from the Standby Database will also increase. This will increase the response time and reduce the throughput of whatever application or client is connected to the Primary Database. See the Figure below from Oracle for an example of a Data Guard Zero Data Loss configuration.

*Figure 1–1 Typical Data Guard Configuration*




In most cases, round trip ( RTT ) network latency between the primary and secondary databases in excess of 20 milliseconds is considered too excessive to deploy Data Guard in Zero Data Loss Synchronous mode. When network latency is greater than 20ms, most deployments will have to use the Asynchronous mode of Data Guard, also called Maximum Performance Mode. The testing of Asynchronous Mode is outside the scope of this paper.

The goal of the test case was to determine if F5's WOM technology could provide enough network improvement, that Data Guard Synchronous Mode could be used on a WAN network with RTT latencies higher than 20ms. Three test cases were chosen, 20ms, 40ms, and 100ms. Additional tests were done with 0ms, as the baseline for Data Guard with a Local LAN connected Standby. 20ms was chosen as the upper limit of what could be expected on a Metropolitan or Regional Network. 40ms was chosen to be representative of a short haul Wide Area Network. And 100ms was chosen as a long-haul WAN, as this is a practical RTT for data center to data center replication using standard WAN transports. It also represents a factor of 5 times the current tolerance level for Data Guard Synchronous Mode. And, it represents a realistic WAN RTT when trying to replicate data from coast to coast across the U.S., from the U.S. to Europe, or from the U.S. to EMEA, using commercially available WAN services from carriers.

## Swingbench Overview and Configuration

In order to generate DataGuard traffic over the WAN, a tool is needed. The Swingbench database load generation tool was used to create a workload on the Primary database, thereby causing the Data Guard process to send database redo from the Primary to the Standby instances of the database. The Swingbench software version used was 2.3.0.422. The Primary database was pre-populated with the Swingbench schema, using the defaults. As the purpose of this testing was to determine the Data Guard Zero Data Loss Synchronous performance characteristics, the Swingbench tool was configured for the worst case scenario, performing the maximum number of database writes as possible. The test profile was changed to make heavy use of the "new order process" functions of the tool, which is designed to insert as many records into the order entry table as possible. This creates a large number of database transactions, which generates redo that Data Guard will replicate to the Standby Database. When Synchronous transport is used, commit success is not signaled to the application until after the Standby Database acknowledges receipt of the redo and confirms it has been written to a log file on disk.

There were only a few parameters that were changed from the defaults.

- Number of Users was set to 10.
- MinDelay was set to 10.
- MaxDelay we set to 20.
- To make the database create more log writes, the XLM Parameters for the Transaction Parameters were changed as follows:

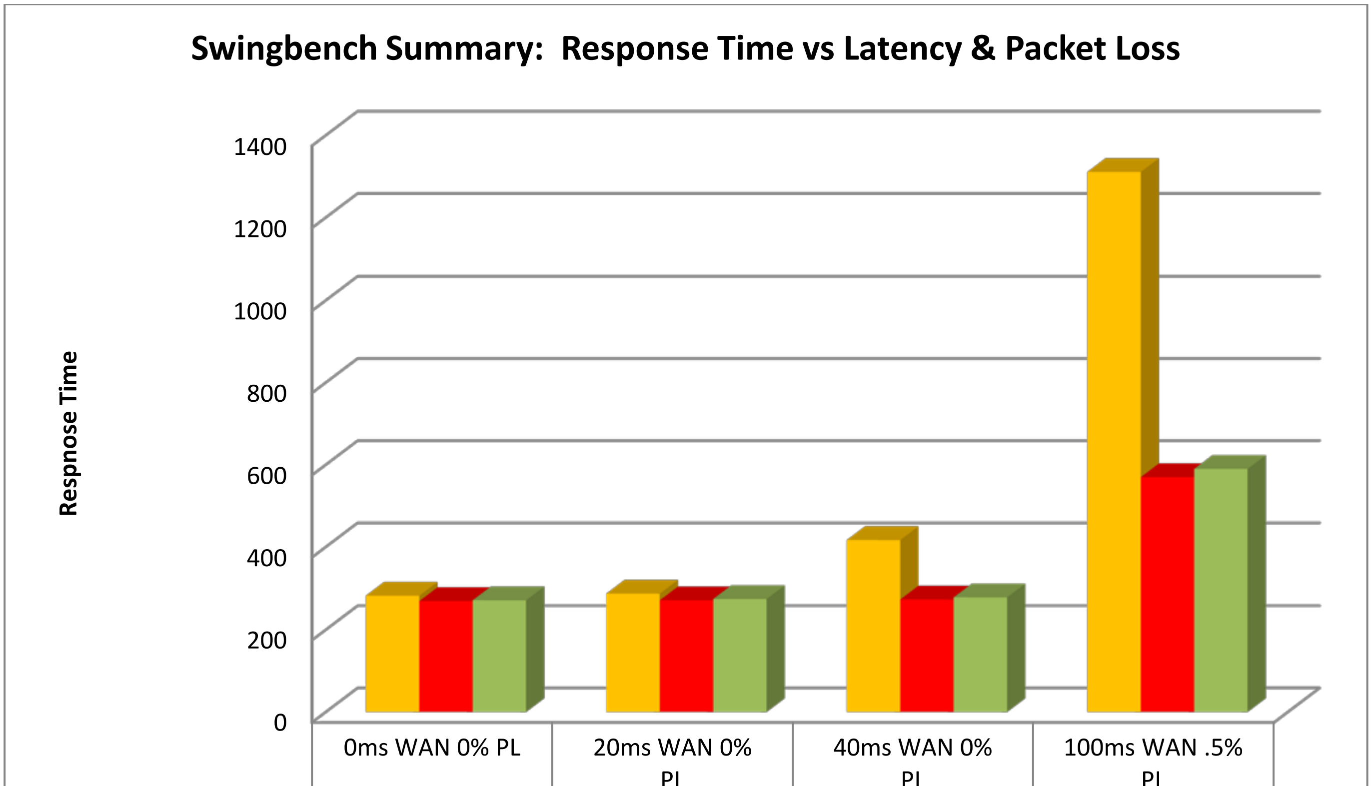
| **Transaction Parameter** | **Value** |
|---|---|
| `NewCustomerProcess` | `False` |
| `BrowseProducts` | `False` |
| `NewOrderProcess` | `True` |
| `ProcessOrders` | `False` |
| `BrowseAndUpdateOrders` | `False` |

When the different test cases were completed, the results files from the Swingbench load generator machines were analyzed, and the "Average Response Time" value was used to create the charts for the Results. In addition, the Transactions Per Minute value was collected for the 100ms worst case test, for comparison as well. A second chart for transactions per minute is shown in the 100ms Results section below.

Note: These test results are only representative of a sample application, Swingbench, on a test harness and test cases that were created in an engineering test facility. Every effort was made to ensure consistent and reproducible results. Testing on production systems was not done, and is outside the scope of this paper.

## Oracle DataGuard Results

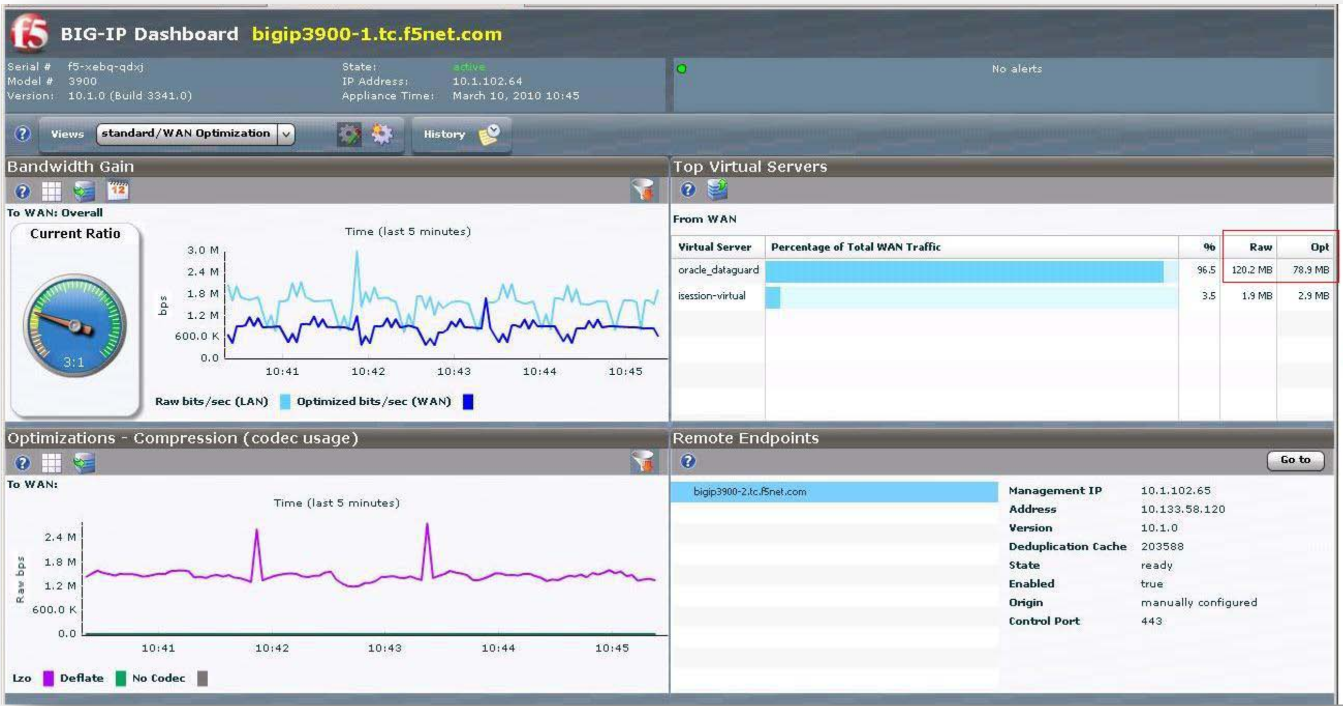
After the test cases were completed, the data was collected and summarized. As you can see, the F5 with WOM can provide LAN like response time for Data Guard over a WAN with high latency and packet loss. As the latency and packet loss increase, the Wan Optimization technologies are able to overcome these inefficiencies, and provide a higher level of performance and throughput than Data Guard can provide on its own.




| | 0ms WAN 0% PL | 20ms WAN 0% PL | 40ms WAN 0% PL | 100ms WAN .5% PL |
|---|---|---|---|---|
| DG Direct 45mb | 283 | 288 | 417 | 1311 |
| F5 WOM Lite 45mb | 270 | 272 | 274 | 570 |
| F5 WOM All 45mb | 272 | 275 | 279 | 590 |

You can see that the performance for both the 0ms and 20ms networks is almost the same, with and without WOM. However, as the latency and packet loss increase in the 40ms and 100ms networks, the performance improved with WOM is substantial. *The higher the latency the packet loss, the more benefit the F5 WAN Optimization technology can provide.* In addition, the data was transported within the encrypted iSession Tunnel.

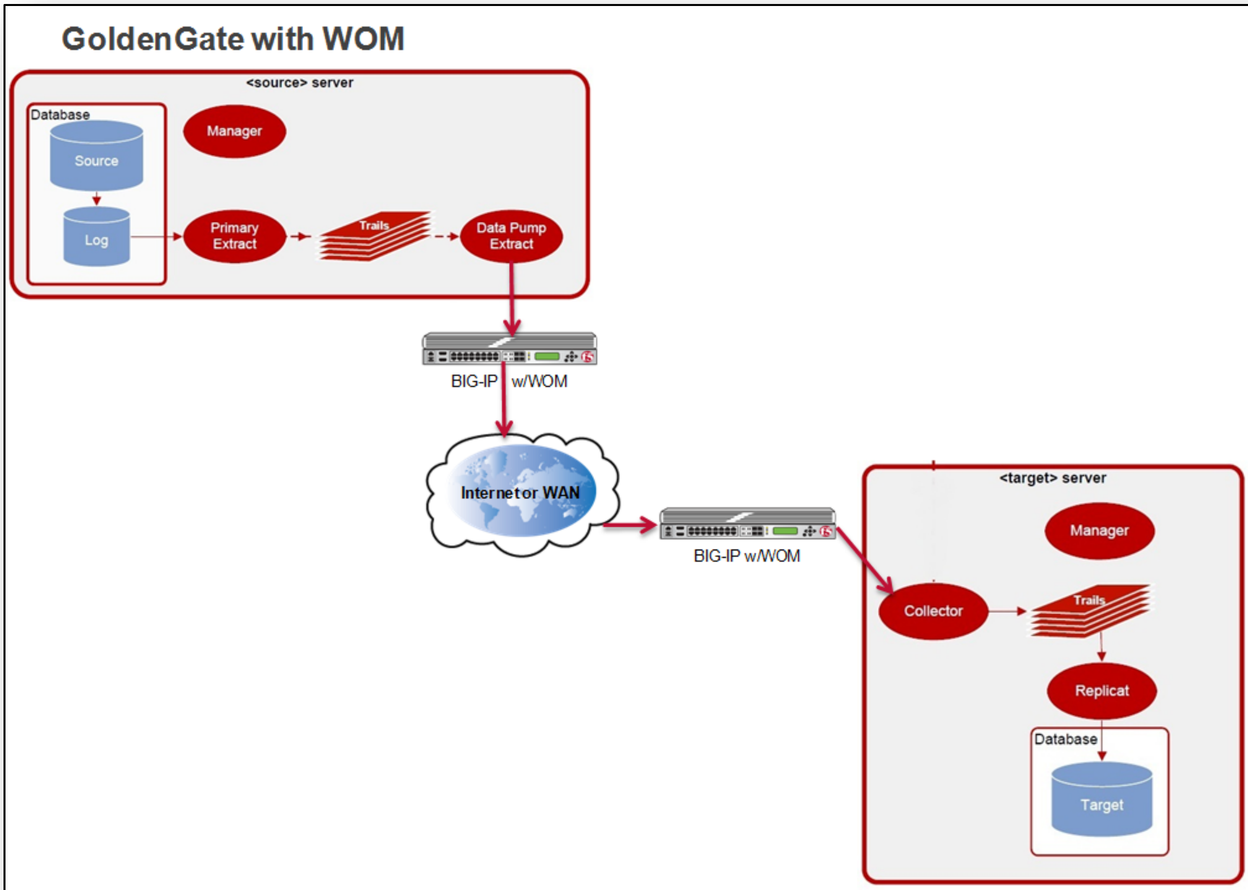
When running a test for Data Guard in Synchronous mode, you can see from the F5 WOM Dashboard, that the WOM module is compressing the data. During the test pass, the Virtual Server "Oracle_Data Guard" - Raw bytes is approximately 120 MB, and the Optimized bytes is approximately 79 MB ( red square in upper right corner.), close to a 50% savings. In looking at the upper left Bandwidth Gain window, you will see approximately a 2:1 ratio. The lower left panel is showing the LZO codec is active.




## Oracle GoldenGate Overview and Configuration

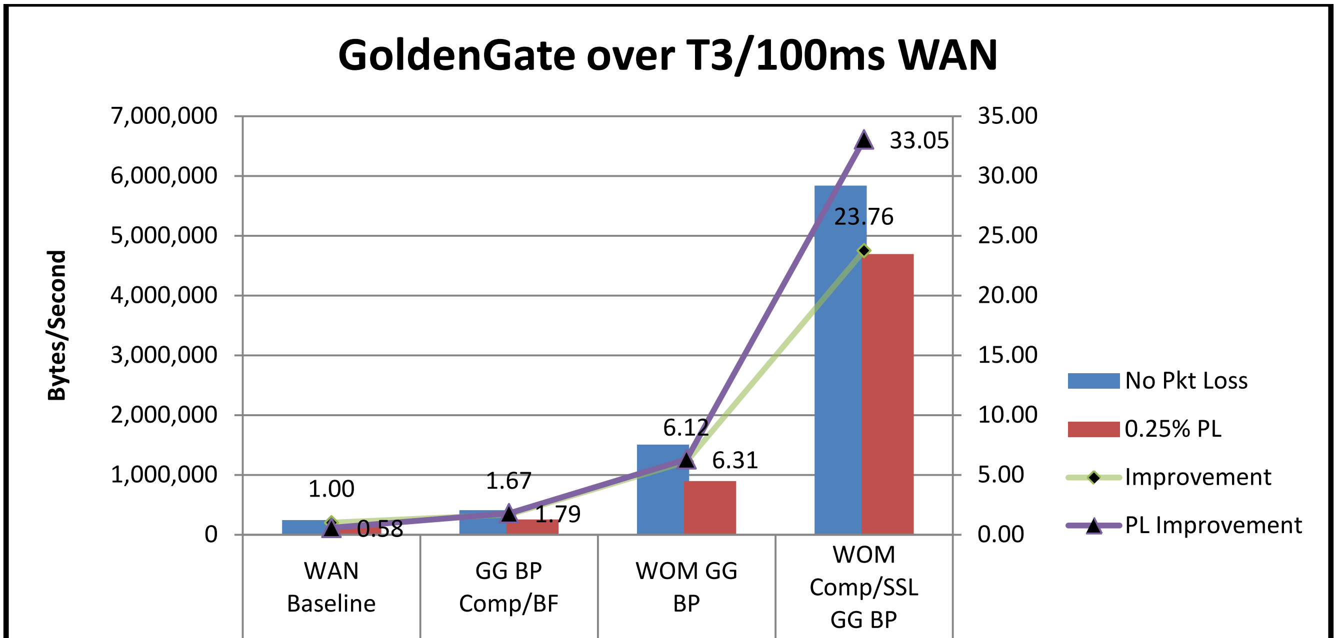
In this test, we tested the time that it takes to move data from a GoldenGate source to a target database. In our test harness, both the source and target databases were Oracle 11gR2 standalone instances. We wanted to see how much F5's WOM could speed up the "datapump" process, which is extracting data from the trail files, and sending it over the network to the collector. We used the Swingbench tool to populate the source database and trail files, and then ran the same datapump process multiple times, using the same source data files. We used the built-in GoldenGate replication tools that come with the product, no special software was used. When running these tests, the software will actually show how much data has been sent, and how long it took. We simply took these 2 values, and calculated the Bytes per Second after 10 or 15 minutes test passes. We ran 3 test passes for each, and averaged the results of 3 passes.

After tuning the GoldenGate parameters for maximum network throughput, we ran a series of tests, and enabled compression or encryption, or both. We did this for both the built-in GoldenGate zlib compression, and the blowfish encryption. Then we did this on the F5 WOM software with LZO compression or SSL encryption, or both. Trying various combinations to see what worked best. This is the network diagram of our test harness for GoldenGate:




## Oracle GoldenGate Results

Here is a summary result chart of the performance improvements.




| | WAN Baseline | GG BP Comp/BF | WOM GG BP | WOM Comp/SSL GG BP |
|---|---|---|---|---|
| No Pkt Loss | 245,765 | 411,465 | 1,504,038 | 5,839,268 |
| 0.25% PL | 141,928 | 254,563 | 895,110 | 4,691,338 |
| Improvement | 1.00 | 1.67 | 6.12 | 23.76 |
| PL Improvement | 0.58 | 1.79 | 6.31 | 33.05 |

You can see from the chart, that the best improvement you can expect from tuning the GoldenGate software is about 1.7x over the defaults. If you use the WOM technology from F5 with its TCP optimizations, compression, and SSL encryption, you can extend this improvement to over 23x on a clean network, and up to 33x on a dirty WAN with packet loss. An interesting side note is that the source database CPU usage actually increased, because the WOM tunnel allowed it to extract and send more data faster. You will also notice, that when using GoldenGate on the baseline WAN, the throughput dropped over 40% when there was packet loss in the network.

## Oracle Recovery Manager Overview and Configuration

In this test, we tested the time that it would take to instantiate a duplicate database across the WAN, creating a new Data Guard Standby database while the Primary was running. The default installation of Oracle 11gR1 Enterprise Edition was used, about 2.3 GB in size. This is an empty database. The Oracle best practice "RMAN duplicate database as standby" script was used, a link can be found in the References section at the end of this document.

Two tests were performed, one with the Primary connected directly to the Standby target over the WAN, and a second case where the Primary was connected through the F5 WOM tunnel to the Standby target. The network used was a T3, 45mb/s network, with 100ms RTT, and no packet loss. This would be roughly the equivalent of a WAN network from the West coast to the East coast of the U.S.

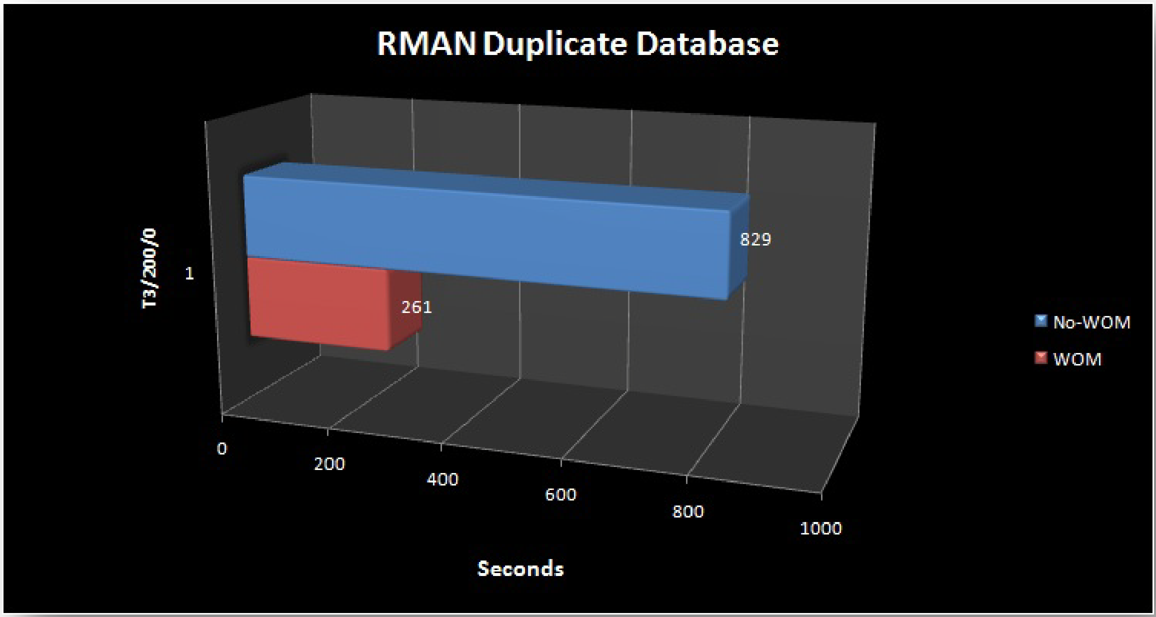
RMAN script example:

```
connect target /;
connect auxillary sys/oracle1@STANDBY_dgmgrl;
duplicate target database for standby from active database
spfile
set db_unique_name='standby'
set control_files='/u01/app/oracle/oradata/standby/control01.dbf'
set instance_number='1'
set audit_file_dest='/u01/app/oracle/admin/standby/adump'
set remote_listener='LISTENERS STANDBY'
nofilenamecheck;
```

The linux shell command "TIME" was used to measure how long the RMAN script took to execute, completing the duplication.

## Oracle Recovery Manager Results

The results of these two tests are shown in the chart below. The RMAN script without WOM took 13 minutes, 49 seconds. The RMAN script with WOM took 4 minutes, 21 seconds. The RMAN duplicate database process was approximately 3.2 times faster using WOM than the direct case. As there was no packet loss introduced in this network, most of the improvement can be contributed to the WOM compression and de-duplication. It is impossible to predict how any given database will benefit from compression, as every database is unique. This is a worst case test scenario for RMAN, as the default installation of 11gR1 database contains very little actual user data. You will most likely achieve even better results with a production source database.



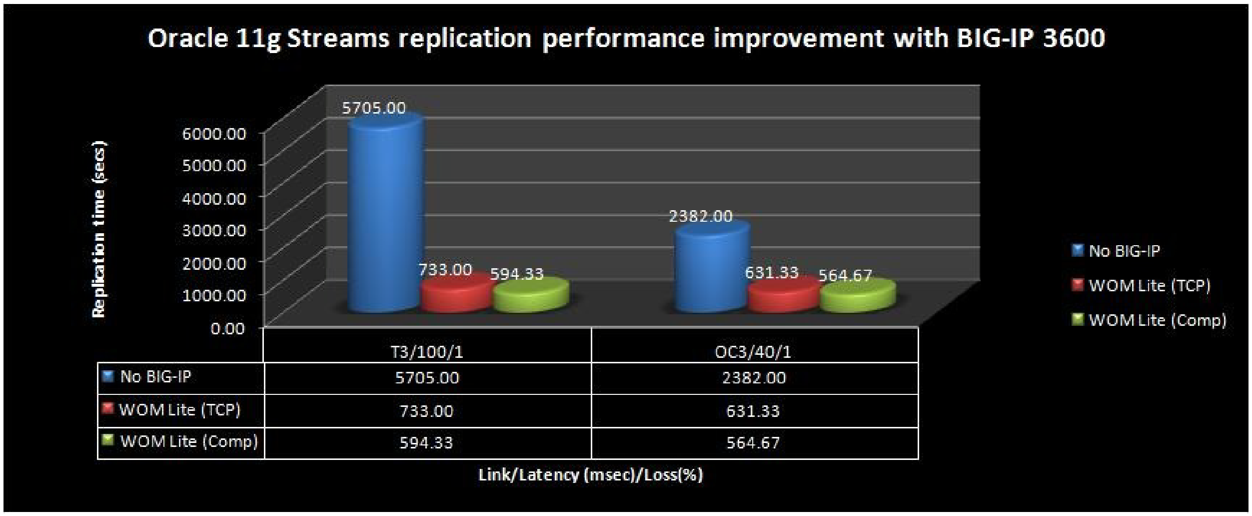

## Oracle Streams Overview and Configuration

Oracle Streams replication testing was done with the TPC-E benchmark tool. This benchmark simulates a financial trading company's transactions. In our tests, we ran the load generation tools that would yield one business day's worth of trading data. Oracle Streams was configured to replicate that set of data to a remote target database, and measured the amount of time it took by looking at the alert log files. You can find more information about the TPC-E benchmark tool in the References section at the end of this document.

The TPC-E schema was created on both the source and target database instances. Oracle Streams was then configured to enable uni-directional replication from the source to the target using Oracle stored procedures. Once the Capture, Propagate and Apply processes were successfully started, the capture process was stopped on the source and the TPC-E load generation program was executed to populate the source database instance. The capture process was then enabled, and the time to complete replication was measured. The replication was verified by checking the contents of the TPC-E tables before and after replication. This is a "gap resolution test," where we are looking at the time that it takes for the target to "catch up" and be in synch with the source. The larger the gap and the longer it takes, the more exposed your data is in the case of a failure on the source database.

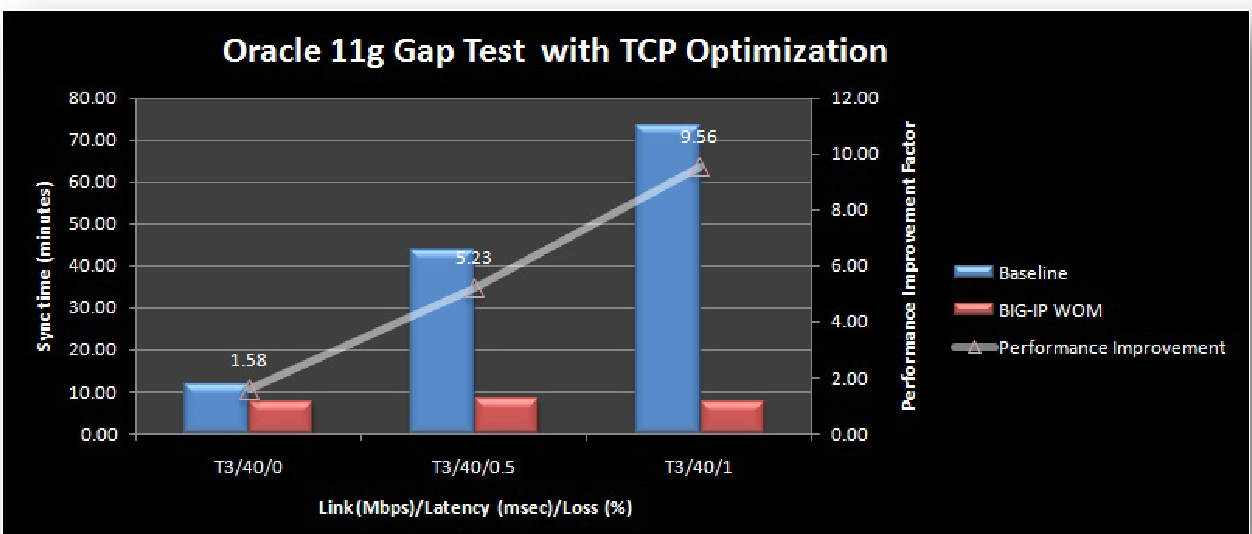
The Test Network Architecture was the same - having 2 BIG-IPs with WOM were used on either end of the WAN link. But for this test, two different LANforge WAN settings were used. One with a WAN link of T3 45mb/s, 100ms RTT, 1% packet loss, and one with an OC3 135mb/s link, 40ms RTT, and 1% packet loss.

### Oracle Streams Results



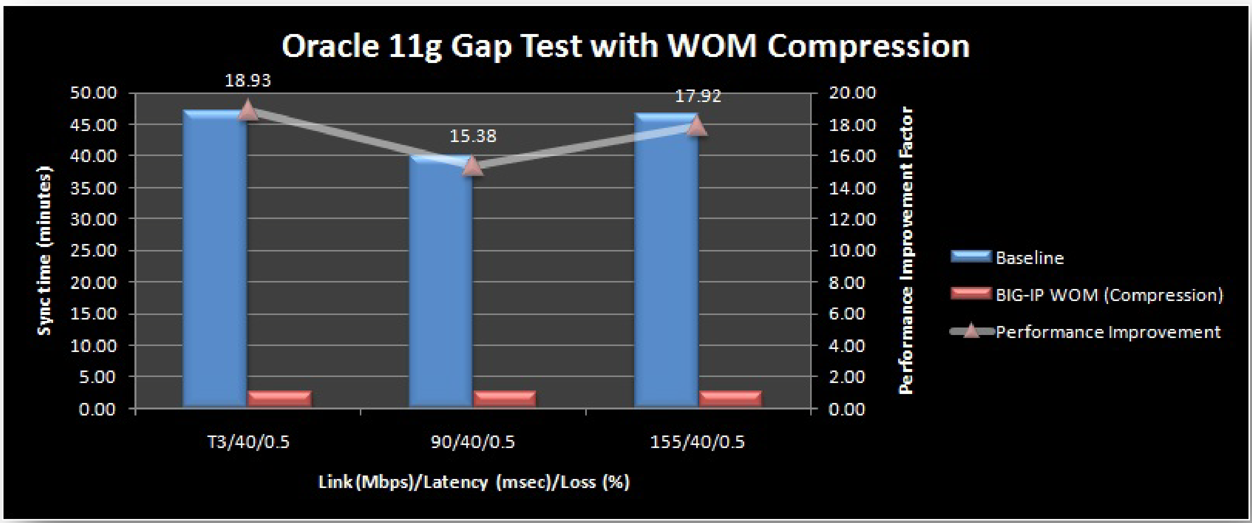

As you can see, at both the T-3 and OC3 link speeds, the WOM software can significant decrease the amount of time that it took to replicate a day's worth of trading data, about 2GB worth, from the source to the target database. On the T-3 link, without WOM, it took about 95 minutes, and with the WOM module, it took about 10 minutes. That is a 9.5x improvement. On the OC3 link, it took 40 minutes versus 9.4 minutes, which is about a 4x improvement. Streams replication across both of these networks was able to be improved by the Wan Optimization module.

To further explain how WOM is providing additional value of off-host compression and TCP optimizations, these next two charts separate these two WOM services. This first chart shows the WOM TCP/IP optimizations, and how efficiently it overcomes the detriments of packet loss. Again, using the same T-3 WAN at 40ms RTT, packet loss rates were varied to 0%, 0.5%, and 1%.



You can see, that as packet loss increases, so does the baseline gap resolution time without F5 WOM ( the blue bars ). However, with F5 WOM, the gap resolution time remains consistently at 9 minutes for this test ( the red bars ). The white line notes the percentage performance improvement. So even on a clean network with no packet loss, the WOM TCP/IP enhancements can increase throughput by as much as 50%. The "dirtier" your WAN network, the more WOM can help clean it up for your replication services.

Next, we take a look at the compression provided by WOM. Again, we used a T-3 45mb/s WAN, 40ms RTT, and 0.5% packet loss. This time, we ran the test with the WOM compression disabled, then enabled. In this chart, you see that the compression again helps resolve the gap faster. By compressing the data, WOM is in effect transmitting more data, sending the redo blocks faster.




Looking at the chart, you can see that the F5 WOM compression allowed the target to be consistently only a few minutes behind the source. Without compression, it took much longer for the gap to be resolved, approximately 15 times longer or more.

## F5 Networks Results Summary

The F5 Networks BIG-IP with Wan Optimization Module provided improvements for Database Replication across the WAN in several areas, using several F5 Networks advanced technologies, including:

- TCP Express 2.0 – built on the TMOS architecture, hundreds of TCP network improvements using both RFC based and proprietary enhancements to the TCP/IP stack, allow the BIG-IP to move more data more efficiently across the WAN. Advanced features like adaptive congestion control, selective TCP window sizing, fast recovery algorithms, and other enhancements make this possible. The benefits of these improvements are the LAN-like performance characteristics across the WAN.
- iSession Secure Tunnel – the iSession Tunnel created between the 2 BIG-IP devices can be secured with SSL, allowing for the encrypted transport of sensitive data across any network.
- Adaptive Compression – the Wan Optimization software has the ability to automatically select the best compression codec based on network conditions, BIG-IP cpu load, and different payload types. The results of this testing showed that the LZO algorithm provided the most consistent compression results for Oracle database replication services.
- Symmetric Data Deduplication - eliminates the transfer of redundant data across the WAN to improve response times and throughput while using less bandwidth. Using dedupe cache from Memory, Disk, or both modes are supported with Oracle database replication.

## Conclusion

Using Oracle Database Replication and F5 Wan Optimization technologies together will allow more customers to deploy and use these database services over a WAN, while saving resources by using off-host encryption, compression, deduplication and network optimization. WOM helps DBAs and Network Architects lower their Recovery Time and Recovery Point Objectives needed for business continuity and disaster recovery. WOM helps minimize the effects of WAN network latency and packet loss, and saves money by eliminating or reducing the large expense of upgrading Wide Area Networks.

## References


1. Oracle Maximum Availability Architecture Web site
   http://www.oracle.com/goto/maa
2. Oracle Data Guard Technical Information
   http://www.oracle.com/goto/dataguard
3. Oracle Database High Availability Overview (Part #B14210)
   http://otn.oracle.com/pls/db111/db111.to_toc?partno=b28281
4. Oracle Database High Availability Best Practices (Part B25159)
   http://otn.oracle.com/pls/db111/db111.to_toc?partno=b28282
5. Data Guard Redo Transport & Network Best Practices Oracle Database 10g Release 2
   http://www.oracle.com/technetwork/database/features/availability/maa-wp-10gr2-dataguardnetworkbestpr-134557.pdf
6. Oracle GoldenGate Overview
   http://www.oracle.com/technetwork/middleware/goldengate/overview/index.html
7. Overview of RMAN Database Duplication
   http://download.oracle.com/docs/cd/B28359_01/backup.111/b28270/rcmdupdb.htm#insertedID1
8. F5 Networks BIG-IP Product Overview
   http://www.f5.com/products/big-ip/
9. F5 Networks Wan Optimization Module Product Overview
   http://www.f5.com/products/big-ip/product-modules/wan-optimization-module.html
10. F5 Networks Wan Optimization Module Datasheet
    http://www.f5.com/pdf/products/big-ip-wan-optimization-module-ds.pdf
11. TPC-E Benchmark Tool
    http://www.tpc.org/tpce/
12. Swingbench benchmark tool
    http://www.dominicgiles.com/swingbench.html