Oracle Disaster Recovery on IBM Power Virtual Server

Insights and experiences: Implementing the Oracle Data Guard on Power Virtual Server LPARs hosted in different data centers.

IBM

Table of contents

Overview

The objective of this white paper is to share the experience of implementing an Oracle Data Guard solution across two IBM® Power® Virtual Server data centers. This paper will assist customers seeking sizing guidance and best practices for the Oracle Data Guard Solution on IBM Power.

The test team configured two Power Virtual Server logical partitions (LPARs) in two regions, namely Montreal and Toronto, which are approximately 540 km apart. The Oracle Data Guard solution was implemented across these LPARs. The test team selected an Online Transaction Processing (OLTP) workload and performed multiple runs, modifying the transaction mix to evaluate various read/write ratios.

Disclaimer

This paper presents results for educational purposes only. The results do not accurately portray the full potential capability of IBM Power processor based systems and Oracle Database 19c. The configurations used default values and generally accepted best practices without undergoing any intense tuning on AIX or Oracle Database 19c.

Introduction

This section provides a concise overview of key components, including IBM AIX, IBM Power Virtual Server, Oracle Data Guard, and the OLTP workload.

IBM Power Virtual Server

IBM Power Virtual Server is an Infrastructure as a Service (IaaS) offering provided by IBM. The virtual machines deployed on Power Virtual Server are known as LPARs and align with Oracle's certified on-premises LPAR configuration. Power Virtual Server instances are hosted in IBM Data Centers, separate from the IBM Cloud servers, and they have their own distinct networks and direct-attached storage.

For more information on IBM Power Virtual Server, refer to Getting started with IBM Power Systems Virtual Servers on cloud.ibm.com.

For detailed steps on creating a Power Virtual Server, refer to Creating a Power Virtual Server guide on cloud.ibm.com.

IBM AIX

IBM AIX is a UNIX® operating system built on open standards. When combined with Power Virtual Server, AIX offers enhanced flexibility and performance, enabling workload consolidation on fewer servers. This consolidation can lead to increased efficiency and energy conservation. AIX provides a high level of security, integration, flexibility, scalability, and reliability. It is designed to operate on IBM Power architecture.

For more information on AIX, refer to IBM AIX on ibm.com.

For accessing documentation resources specifically for AIX, refer to the following on ibm.com.

- IBM Power Systems and AIX Information Center.
- IBM Developer for AIX.

Oracle Database 19c

This section contains technical information for Oracle Database 19c that needs to be considered in AIX installation. In Oracle Database 19c, Grid infrastructure is an option that includes Automatic Storage Management (ASM) and Oracle Restart.

Oracle released Oracle Database 19c for IBM AIX in May 2019. Oracle Database 19c is a long-term support release for the Oracle 12.2 family of releases.

For more information refer to the following Oracle Database 19c documentation on docs.oracle.com.

- Oracle Database 19c Documentation.
- Administrators Reference for Linux and UNIX System-Based Operating Systems.

For more information on automating Oracle database using Ansible, refer to the following Galaxy URLs.

- Ansible Galaxy - IBM Power AIX Oracle.
- GitHub - IBM Ansible Power AIX Oracle.

Oracle Data Guard

Oracle Data Guard ensures high availability, data protection, and disaster recovery for enterprise data. Data Guard provides a comprehensive set of services that create, maintain, manage, and monitor one or more standby databases to enable production Oracle databases to survive disasters and data corruptions.

What is Oracle redo Transport?

Oracle redo transport refers to the mechanism used by Oracle database to move redo data from the primary database (source) to one or more standby databases (destinations). Redo data consists of a record of changes made to the database, such as inserts, updates, and deletes. This transport is a crucial component of Oracle Data Guard, which is a high-availability and disaster recovery solution. Redo transport ensures that changes made on the primary database are replicated to the standby databases in real-time or near-real-time, providing data protection and continuity in case of a primary database failure.

Modes of redo transport

Synchronous redo transport

Synchronous redo transport is also called the 'zero data loss' method as the Log Writer process (LGWR) is not allowed to acknowledge a commit has succeeded until the Log Network Server (LNS) confirms that the redo needed to recover the transaction has been written to disk at the standby site.

Asynchronous redo transport

Asynchronous transport (ASYNC) LGWR process does not wait for acknowledgment from the LNS. This creates a near-zero performance impact on the primary database regardless of the distance between primary and standby locations.

For more information on Oracle Data Guard, refer to the following Oracle documentations.

- Oracle Data Guard Documentation.
- Oracle Data Guard Technical Overview.

Brokerage – OLTP workload

The OLTP workload simulated the transactions and database of a stock brokerage firm. The test team created a database size of approximately 200GB using 10,000 customers. The team modified the transaction mix and conducted multiple runs by varying the read/write ratio.

On-premises network validation

In an on-premises environment, the test team configured two Power8 LPARs connected by a 1 Gbps network and conducted tests to assess the 1 Gbps network interface between the primary and standby databases. To validate the configuration, the test team initially employed the iPerf tool to execute a microbenchmark test. A microbenchmark is a program that measures the performance of a single, well-defined task, such as elapsed time, the rate of operation, bandwidth, etc.

For more information about on iPerf tool, refer to [IBM AIX Performance Analysis Using iperf](#).

Figure 1 illustrates the comparison between the workload run network throughput and the microbenchmark iPerf tool results. We established a baseline packet size of 2048 and normalized the data for other packet sizes in both cases. These packet sizes were chosen as representatives of the sizes commonly used by Data Guard for transmitting data between database instances. During the microbenchmark test on the 1 Gbps network, we observed significantly higher bandwidth, up to 3 times, compared to the network throughput of the OLTP workload. This indicated that the 1 Gbps network was sufficient for handling the workload, as the network packet transfer rate remained below the capabilities of the network switch. However, it is worth noting that in the case of connecting remote datacenters, latency and bandwidth should be carefully evaluated. Increased distance tends to result in higher latency, which can significantly impact the workload.

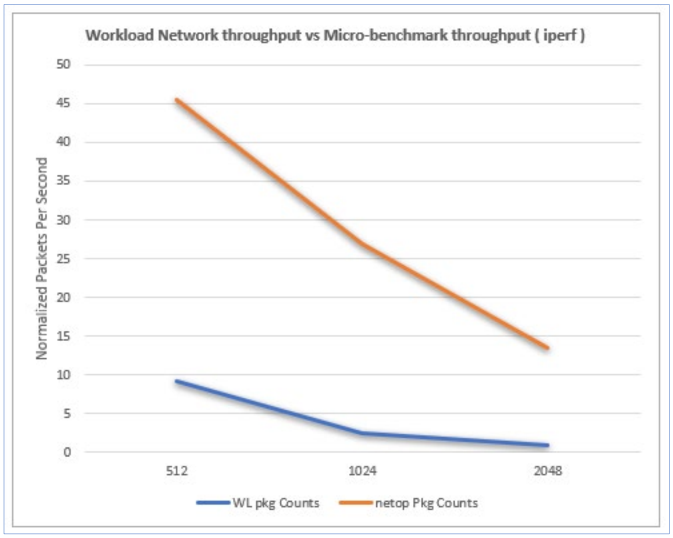

Figure 1: Workload network throughput versus Micro-benchmark throughput.

Based on the results, the test team decided to use a 1 Gbps network connection between the primary database and standby database.

Feasibility of High Availability Disaster Recovery (HADR) with multiple Power Virtual Server sites

The test team implemented the Disaster Recovery solution using Oracle Active Data Guard of Oracle Database version 19c on IBM Power Virtual Server LPARs running the AIX operating system. The Power Virtual Server LPAR was created in two different regions, one in Montreal and the other in Toronto, with approximately 540 km between the sites, establishing a true DR setup across two regions with an expected 8.43 ms latency from Toronto to Montreal data centers, as calculated by Megaport network latency calculator.

Megaport provides flexible and scalable bandwidth solutions to meet a wide range of networking requirements. Their services encompass public and private cloud connections, metro ethernet, data center backhaul, and internet exchange services.

Figure 2 depicts the system architecture created in two different regions. The architecture consists of LPARs configured in Montreal (site A) and Toronto (site B). In site A, the test team created the LPAR where the primary database was installed along with a client LPAR, while in site B, the test team created the LPAR where the secondary database (standby database) was installed and configured. Both site A and site B LPARs have the same configuration, but the client LPAR in site A has a minimal configuration. It is worth noting that the standby database system can have a configuration equal to or less than the primary database system.

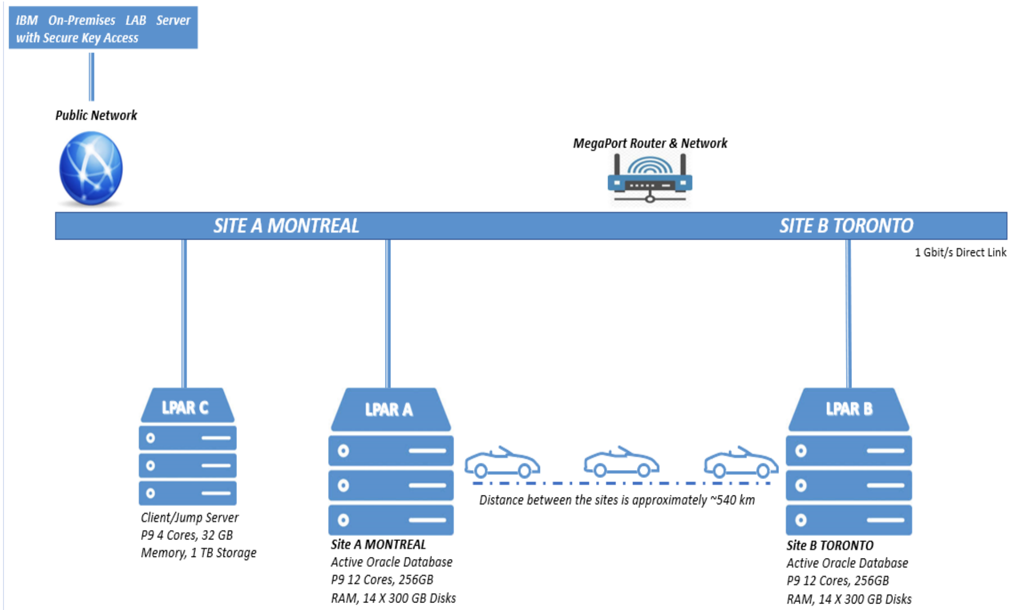

Figure 2: System architecture created in two different regions.

Figure 3 shows the connectivity between the site A and site B Power Virtual Server LPAR using IBM-owned Megaport at each site, with the two ports connected by the VXC network provided by Megaport. Both Power Virtual Server and Megaport confirmed the network latency, which is largely dependent on Megaport's network speed and the distance between the two sites.

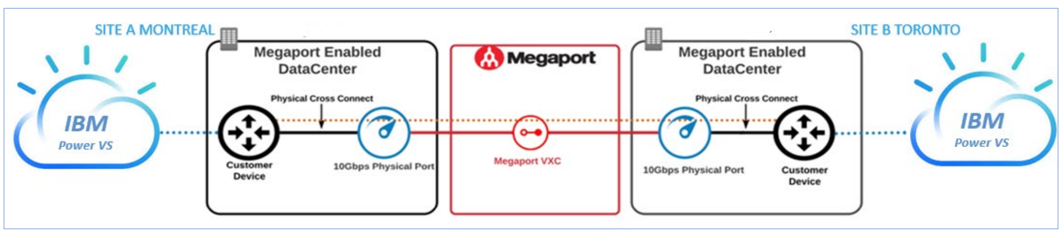

Figure 3: Connectivity between two sites.

The following table provides the hardware and software configuration details for one LPAR, while the client LPAR has the minimal configuration.

Configuration	Power9
System model	IBM Power
Processor type	PowerPC_Power9
Number of cores	12
Memory	256 GB

Firmware version	IBM, FW950.11 (VL950_075)
SMT	4
OS level	AIX 72 TL5 SP2
Oracle Grid Home Version	19.11
Oracle DB home version	19.11

Table 1: Hardware and software configurations of LPAR.

Steps for setting up an Oracle DR solution in Power Virtual Server

This section explains the high-level steps to implement the Oracle Data Guard solution in Power Virtual Server, creating the IBM cloud account, creating the Power LPARs, setting up Active Data Guard between the primary database and secondary database.

1. Create an IBM Cloud account.
2. Review the Identity and Access Management (IAM) information.
3. Create a public and private SSH key that you can use to securely connect to your Power Virtual Server.
4. Create a Power Virtual Server.
5. Create a private subnet for the Oracle DB instance, which will also be used to connect through Megaport VxC to the database instance.
6. Set up Megaport connection between 2 sites – get the service key.
7. Order and setup Direct Link Connect using the service key with Megaport.
8. Setup Active Data Guard between the primary database and secondary database after the network connection is established.

Test analysis

The test team evaluated the workload capability with DR configuration using SYNC and ASYNC modes and without DR configuration.

OLTP brokerage workload without DR versus with DR using async configuration.

Figure 4 depicts the results of the comparison of OLTP Brokerage workload throughput between 'with DR using ASYNC Mode' and 'without DR' scenarios, where 'No DR' or 'without DR' means the transport of redo logs to standby is stopped or suspended. As the number of write operations increases, the workload throughput (TPS) also increases since there is no storage bottleneck. It is observed that the workload throughput is almost the same in both scenarios, with and without DR using ASYNC mode.

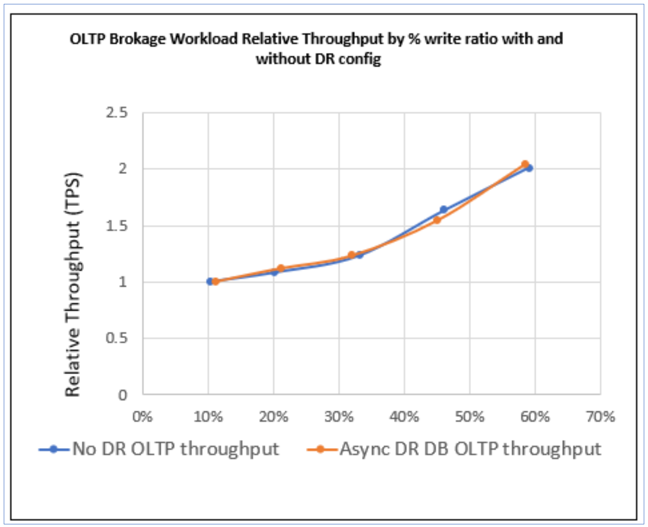

Figure 4: OLTP brokerage workload throughput comparison results.

Figure 5 depicts the comparison results of OLTP brokerage workload storage and network throughput between 'with DR using ASYNC Mode' and 'without DR' scenarios. The test team observed that as the write ratio increased, storage throughput in both scenarios performed similarly. In ASYNC Mode, it was observed that the distance between the DR sites (site A and site B) did not affect the network and database performance. Both cases 'without DR' and 'with DR using ASYNC' were performing similarly. From the network throughput chart in the 'without DR' case, it can be noticed that there is a slight increase in the number of packets being transferred over the network. This is because the client is generating the workload on a different LPAR on site A.

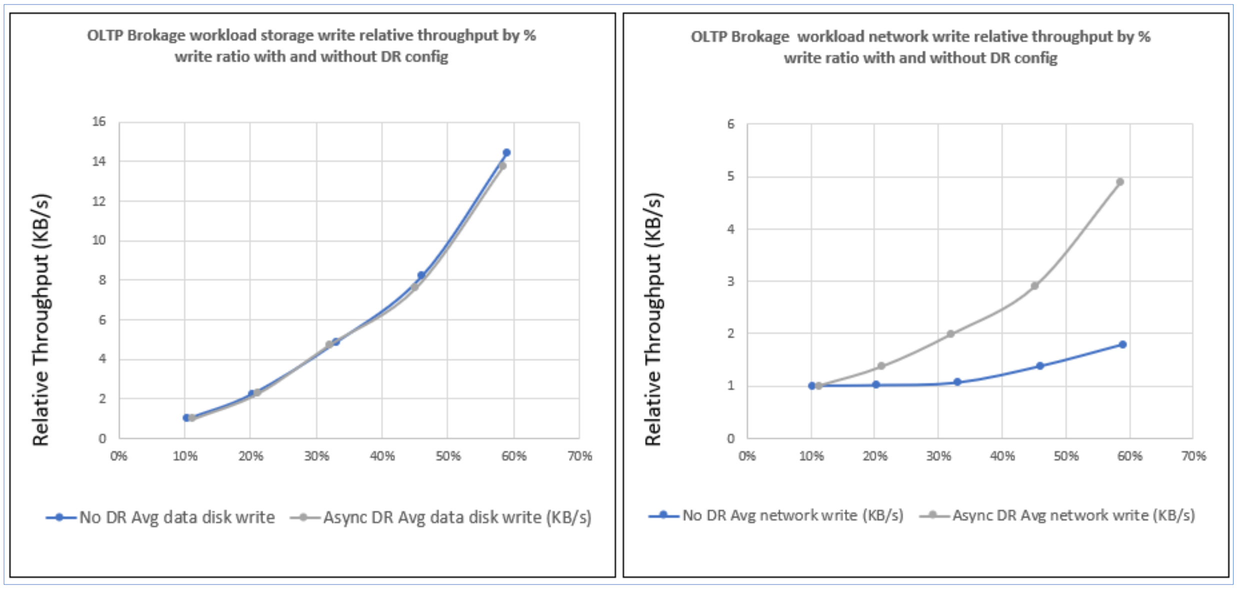

Figure 5: OLTP brokerage workload storage and network throughput comparison results.

As the results were similar between 'with DR using ASYNC' and 'without DR' scenarios, it is recommended to configure ASYNC DR if the two regions are in long distance.

OLTP brokerage workload with DR ASYNC versus SYNC configuration

ASYNC allows the LGWR process to continue without waiting for confirmation from the LNS. This means that the primary database's performance is not significantly affected, regardless of the distance between the primary and standby sites. On the other hand, SYNC, also known as the 'zero data loss' method, requires the LGWR process to wait for confirmation from the LNS before acknowledging a successful commit. This ensures that the necessary redo information for transaction recovery is written to disk at the standby site.

The test team performed tests by varying the write ratio of the workload using Oracle Data Guard's ASYNC and SYNC modes. Figure 6 illustrates the comparative results of OLTP Brokerage workload throughput between DR utilizing ASYNC and SYNC Mode. When the number of write transactions is low, SYNC mode exhibits similar performance to ASYNC mode. However, as the number of write transactions increases in SYNC mode, there is a decrease in TPS, as depicted in Figure 6. At the database level, an increase in waits for the 'log file sync' wait event has been observed.

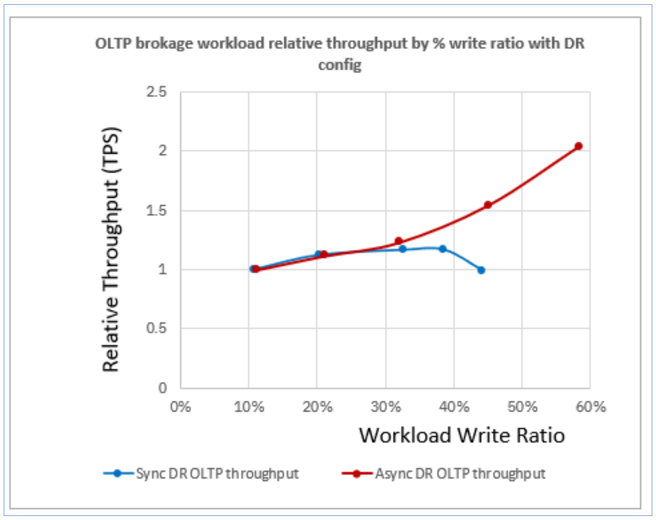

Figure 6: OLTP brokerage workload throughput by % write ratio with DR configuration.

Figure 7 depicts the comparative results of OLTP brokerage workload storage and network throughput between DR utilizing ASYNC and SYNC modes. Regarding network write throughput, SYNC mode exhibits a higher transfer of packets between the primary and secondary databases compared to ASYNC mode. However, as the number of write transactions increases, the workload performance in SYNC mode diminishes due to the primary database waiting for acknowledgment from the secondary database. This leads to a decline in workload performance and an impact on network traffic.

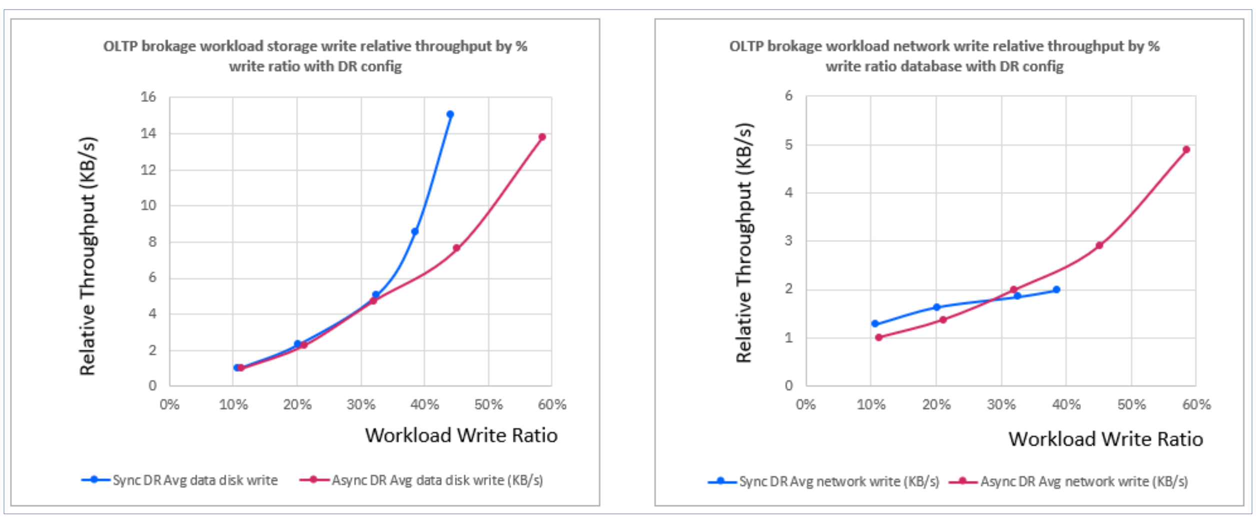

Figure 7: Comparison results of OLTP brokerage workload storage and network throughput.

Based on the above findings, SYNC mode has an impact on database performance but ensures zero data loss. It has been observed that when the workload write ratio is below 30%, the performance of SYNC and ASYNC modes is similar.

Comparing results of on-premises with Power Virtual Server

The test team implemented the Disaster Recovery solution using the Oracle Data Guard feature of Oracle Database 19c Enterprise Edition. This implementation used on-premises LPARs operating on the AIX operating system. The following system architecture diagram depicts the setup, with one of the LPARs serving as the primary Oracle database, another as the standby Oracle database, and the third LPAR dedicated to the client. All three LPARs were established in the same datacenter, featuring a dedicated 1 Gbps connection. The test team also executed the OLTP brokerage workload utilizing an Async DR configuration. These tests were conducted on the on-premises setup to verify the adequacy of the 1 Gbps network connection between the primary database and the standby database. The results indicated that the 1 Gbps network speed was sufficient for this workload, as previously depicted in Figure 1.

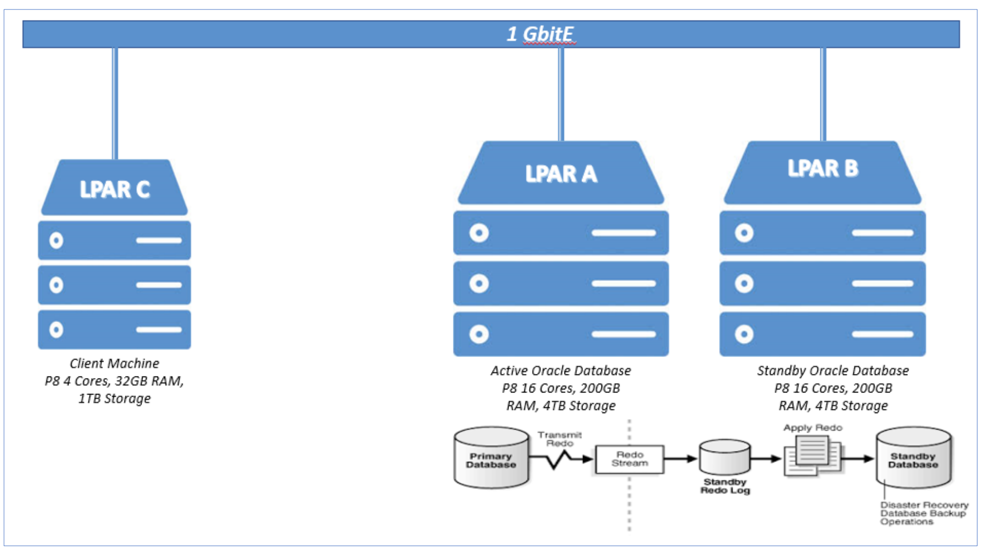

Figure 8: System architecture for on-premises LPARs.

Each LPAR has the same configuration, and the following table provides the hardware and software configuration details for the LPAR.

Configuration	Power8
System model	IBM Power
Processor type	PowerPC_Power8
Number of cores	16
Memory	256 GB
Firmware version	IBM, FW860.90 (SC860_226)
SMT	4
OS level	AIX 72 TL5 SP2
Oracle Grid Home Version	19.11
Oracle DB home version	19.11
Oracle one-off patch	32109594

Table2: Hardware and software configuration details for the on-premises LPAR.

Upon finalizing the establishment of Oracle Data Guard between the primary database and the secondary standby database within the shared data center, the test team initiated the OLTP workload via a client LPAR present in the same data center. The test team replicated the approach used in Power Virtual Server, adjusting the read versus write Ratio from 90/10 to 60/40. The test team established a database size of around 200GB using 10,000 customers on both on-premises and Power Virtual Server environments. The following chart illustrates the outcomes of the Async modes for both Power Virtual Server and the on-premises setup.

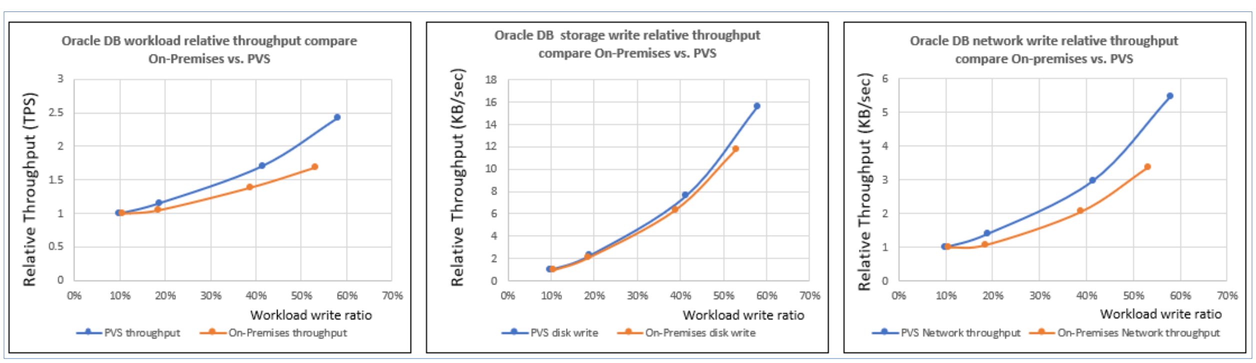

Figure 9: PVS versus on-premises Oracle DB workload, storage, network write throughput.

The test team observed that the Oracle OLTP brokerage workload performed more efficiently on Power Virtual Server with Power9 systems compared to on-premises Power8 systems as per Figure 9, for the ASYNC scenario. This improvement is due to the enhanced processing capabilities of Power9 systems, driving a higher transaction rate, resulting in increased storage and network throughput as shown in Figure 9. The Power Virtual Server setup with Power9 systems benefits from a DR configuration that spans two regions, a strategic implementation does not present in the on-premises setup with Power8 systems. This distributed DR setup provides increased redundancy and availability for the Power9 systems. Based on the tests conducted by the testing team, although Power9 systems have fewer cores compared to Power8 systems, they have demonstrated better results when compared to on-premises Power8 systems.

Summary

This paper presents a step-by-step procedure for implementing Oracle Data Guard across two Power Virtual Server data centers. The tests were conducted with various Oracle Data Guard modes, revealing that when the workload involved low write transactions, the SYNC mode performed similarly to the ASYNC mode while ensuring high availability. However, as the workload increased with more writes, the database performance was affected during the SYNC mode. For high-performance environments, customers may consider using the ASYNC mode, especially when the workload involves more writes and the distance between the two sites is considerable. However, it is important to note that this choice introduces the risk of data loss if failover to the standby database occurs. Additionally, this paper compares the performance of the on-premises Power8 DR solution with the Power Virtual Server DR solution. The results indicate that the Power9 systems outperformed Power8 systems, particularly in high-performance scenarios.

Get more information

For more information on using Oracle database on IBM Systems, contact ibmoracle@us.ibm.com.

For more information on using Oracle Database 19c and Oracle Disaster Recovery, refer to the following:

- Oracle Grid Infrastructure Installation and Upgrade Guide 19c for IBM AIX on Power Systems (64- Bit) E96274-08.
- Oracle Database Installation Guide 19c for IBM AIX on Power Systems (64-bit) E96437-07.
- Oracle Database Release Notes - 19c for IBM AIX on Power Systems (64-Bit) E96072-28.

- Oracle Database Licensing Information User Manual – Oracle Database 19c E94254-46.
- Data Guard Concepts and Administration.
- Plan Disaster Recovery for Databases.
- Oracle Active Data Guard White Paper.

About the authors

Wayne Martin is the IBM Systems Technology Solutions Manager responsible for the technology relationship between IBM and the developers of Oracle Corporation's Database and Fusion Middleware for all IBM server brands. His responsibilities include driving the mutual understanding between IBM and Oracle on technology innovations that will generate benefits for mutual customers, managing the process of getting that technology implemented in products and ensuring that availability of the products to customers is timely. Wayne has held a variety of technical and management roles at IBM that have focused on driving enhancements of ISV software that uses IBM's mainframe, workstation, and scalable parallel products. You can reach Wayne Martin at wmartin@us.ibm.com.

Theresa Xu is an experienced Performance Analyst (with over 10 years of experience) leading a team focused on improving IBM Banking workloads, big data AI, and analytics workloads on the Power platform in hybrid cloud environments. Over 12 years of technical management and coding experience, specializing in IBM DB2 using C, C++, and Java in AIX and Linux. You can reach Theresa Xu at theresax@ca.ibm.com.

Bhargavaram Akula is a Technical Consultant with IBM India, Hyderabad. He collaborates with the specialists at the IBM Oracle International Competency Center based in Foster City and Redwood Shores (California, US) working on Oracle product certifications on IBM Power running AIX. He has extensive experience with Oracle products. You can reach Bhargavaram Akula at bhargaku@in.ibm.com.

Sudhindra Kumasagi is a Technical Consultant with IBM India, Bangalore, He collaborates with the specialists at the IBM Oracle International Competency Center based in Foster City and Redwood Shores (California, US) working on Oracle product certifications on IBM Power running AIX. He has extensive experience with various Oracle products. You can reach Sudhindra Kumasagi at Sudhindra.Kumasagi@ibm.com.

Shiva Laveti is a Technical Consultant with IBM India, Bangalore, He collaborates with the specialists at the IBM Oracle International Competency Center based in Foster City and Redwood Shores (California, US) working on Oracle product certifications on IBM Power running AIX. He has extensive experience with various Oracle products. You can reach Shiva Laveti at shiva.laveti@ibm.com.

IBM Corporation New Orchard Road Armonk, NY 10504

Produced in the
United States of America
January 2024