egenera®

Product Brief

Oracle RAC on the Egenera® BladeFrame® System

Executive Summary

Oracle Real Application Clusters (RAC) deployed on the Egenera BladeFrame is a key component to lowering total cost of ownership (TCO) for high-performance database systems. The Egenera BladeFrame and its Processing Area Network (PAN) architecture provide an unbeatable platform for Oracle Real Application Clusters. The physical features of the BladeFrame's hardware and the virtualization capabilities of its software provide a natural and optimized synergy with Oracle RAC. The resulting benefit is a significantly lower TCO while providing unbeatable scalability, availability, performance and manageability. Any IT organization looking to lower TCO for high-performance database systems should investigate Oracle RAC running on the Egenera BladeFrame.

BladeFrame Overview

The Egenera BladeFrame was purpose-built to solve the multiple pain points of enterprise-class datacenters by delivering the next generation of computing infrastructure at a lower TCO than current architectures. By integrating processing, networking and management functionality currently dispersed across server hardware, data networks and operating systems, Egenera creates a new architecture, the PAN, which consolidates and simplifies the allocation and management of computing power.

The Egenera PAN architecture is a group of autonomous, stateless Processing Blade™ resources (pBlade™) redundantly connected via a high-speed, low-latency switched fabric. With features such as built-in high availability and dynamic resource repurposing, the system was designed for mission-critical applications such as databases, application servers and transaction processing.

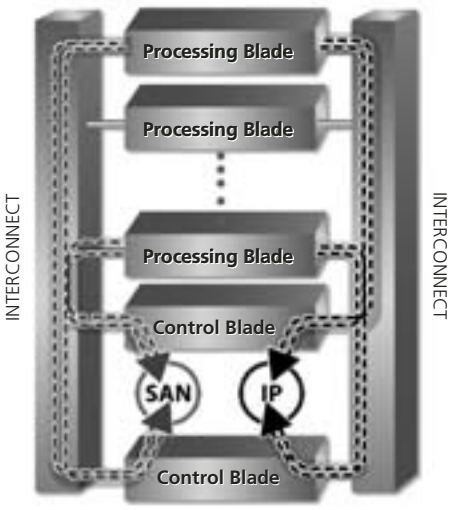


Figure 1: The architecture of the BladeFrame is composed of up to 24 Processing Blades connected by dual high-speed, low-latency interconnects.

Oracle Real Application Clusters Overview

Over the past decade, the use of clustering has grown in response to the need for high availability systems. Database servers, however, have typically been unable to take advantage of this technology due to their memory architectures. High-performance databases have therefore required expensive, high-capacity SMP systems to meet end user performance standards. While powerful, these large SMP servers have a very high TCO when compared to x86-based systems.

With Oracle RAC, the hardware requirements for highly scalable and available systems change dramatically. Oracle's development of a new software technology called Cache Fusion reduces the requirement to run high-performance databases exclusively on high-end RISC SMP machines. Through implementation of a "shared cache" architecture, also known as "shared disk" or "shared everything," each server in a cluster has full and complete access to all data in the database. Unlike previous cache-sharing schemes, Cache Fusion shares the server cache with other physically separate servers using a private network interconnect. This frees database developers from the restriction of having the server cache be located in memory on the same device as the CPU.

With Cache Fusion, a server's cache can be shared across multiple servers. Scaling of a database sever can now happen in two directions. Servers can be scaled up by adding additional CPUs to a machine or out by adding servers to the database cluster. No longer must expensive, proprietary SMP machines be used in high-performance database applications. With Oracle RAC, these applications can now be run on much smaller SMP systems.

Scalability is not the only advantage of Oracle RAC. Dramatically improved availability is another key feature. Today, large SMP-based databases typically use some form of mirroring to maintain high availability. Although somewhat effective, mirroring tends to be complex and expensive to implement.

Cache Fusion changes the availability equation. Since Cache Fusion shares a server cache with other physically separate servers, a loosely coupled clustered database is created. Since the servers are loosely coupled, the loss of a single server does not shut down the entire database.

Oracle RAC fundamentally changes the availability and scalability equations for high-performance databases. The Egenera BladeFrame is uniquely suited to take advantage of these advances in database design.

Oracle RAC on the Egenera BladeFrame

The Egenera BladeFrame is an unbeatable platform for Oracle RAC. The physical features of the BladeFrame's hardware and the virtualization capabilities of its software provide a natural and optimized synergy with Oracle RAC. The Egenera BladeFrame coupled with Oracle RAC significantly lowers the TCO for large databases. This cost reduction comes from unbeatable:

- Consolidation
- Availability
- Manageability
- Performance
- Scalability
- Flexibility

Unbeatable Consolidation

With Oracle RAC, companies can undertake software consolidation by placing multiple databases onto a single installation. Through software consolidation, database administration costs are dramatically lowered.

But software consolidation is just half of the consolidation story. Egenera improves on the consolidation equation by also consolidating the hardware. Due to the clustered nature of RAC, consolidating databases onto a traditional hardware architecture actually multiplies the number of hardware components. For example, taking a 32-way SMP machine and porting to Oracle RAC on traditional x86-based hardware would require at least eight four-way servers plus the interconnect switches. The number of hardware components has actually increased over the original SMP configuration.

With the BladeFrame, these same servers consolidate to just eight Processing Blades inside a BladeFrame. No additional networking switches or interconnects are needed. This hardware consolidation helps to dramatically lower capital equipment costs and ongoing maintenance costs.

The ability to operate Oracle RAC in a bladed environment is unique to Egenera. Other bladed products lack the required network switch redundancy, SAN connectivity and/or high-end blade availability.

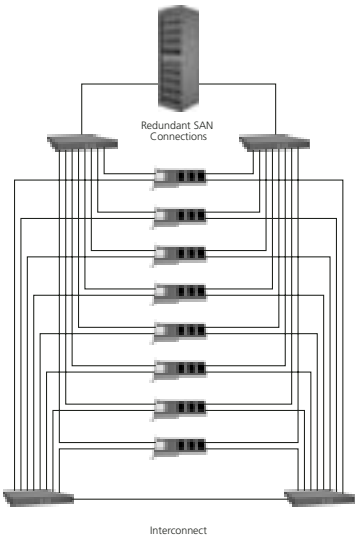


Figure 3:
Running Oracle RAC on traditional x86 server architectures requires a significant amount of hardware complexity. Each node requires traditional storage and network connections plus redundant interconnects.

Unbeatable Scalability

Scaling a database using SMP servers creates a very high TCO. High-end SMP machines have finite limits in their ability to add processing capability, making "forklift" and processing board upgrades common. In addition, their reliance on proprietary RISC has made them unable to take advantage of x86 price/performance.

The BladeFrame, by contrast, creates a pool of up to 192 x86 processors deployable entirely through software and without physical intervention. The processors are arranged on 24 Processing Blades. Hence, a single BladeFrame could scale to 24 Oracle RAC processing nodes bringing a total of 192 processors and 768 GB of memory to work against the database. A similar RISC server generally costs at least twice as much and is populated with processors running at 50 percent the speed of the x86 CPUs.

Upfront hardware cost savings is not the only TCO advantage for the BladeFrame. Today, many enterprises have utilization rates around 20 percent. By increasing utilization rates, organizations can eliminate upfront server capital costs; eliminate the annual, recurring expense of managing and monitoring hardware devices; cut datacenter space requirements in half; and cut associated I/O (storage, network) expense in half.

The BladeFrame increases server utilization by correctly scaling both the type and number of Processing Blades allocated to an application. The enterprise no longer has to overprovision servers to accommodate projected peak demand. Processing Blades from a pool can be brought into service when needed, and repurposed again when no longer required. In addition, applications can be easily moved from one-size blade to another to match the processing requirements to the hardware.

Finally, for every server in production, there are usually accompanying servers in development, disaster recovery, clustering for high availability, or in a UAT environment. With the BladeFrame, these servers are transformed from static, isolated servers to a pool of flexible processing resources that can be deployed and redeployed as needed.

For Oracle RAC databases, the BladeFrame's ability to dynamically scale processing nodes will dramatically improve server utilization. By improving server utilization, organizations enjoy lower TCO.

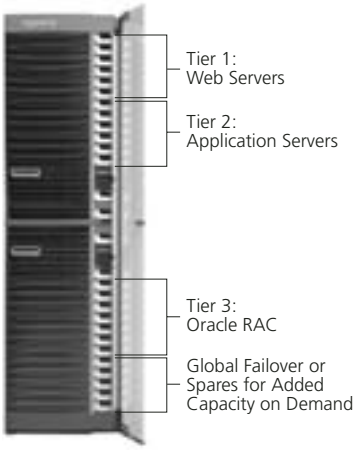


Figure 4:
The BladeFrame's 24 Processing Blades provide unbeatable scalability for Oracle RAC. All 24 Processing Blades could be allocated for Oracle RAC or allocated across different applications. Global failover blades provide high availability for all applications deployed on the BladeFrame.

Unbeatable Availability

Egenera's BladeFrame combined with Oracle RAC provides an unbeatable combination for high availability at a lower TCO than traditional RISC architectures. To improve availability in traditional database server architectures, organizations have typically employed server schemes that rely on hot-standby servers and clustering software. In these scenarios, the standby machine receives a "heartbeat" pulse from the operating server. In the event the heartbeat is lost, clustering software shifts the application over to the standby server.

These clustered server configurations are both expensive to implement and complex to maintain. A Sun 15K running16 processors, for example, would be backed up by a standby Sun 15K for a hardware cost of approximately $3M.

With Oracle RAC and the BladeFrame, the use of a "hot standby" machine is eliminated since every instance server in an Oracle RAC deployment acts as a backup to every other instance server. The failure of a single node reduces system capacity, but does not shut down the entire system. For a two-node system, the loss of one node reduces system capacity by 50 percent. Starting with a three-node system, the capacity is reduced by 33 percent.

Unlike other x86-based architectures, the BladeFrame replaces a failed node with a backup automatically within minutes. When the "new" Processing Blade is started, it has all of the network and storage configurations of the failed blade. Through virtualization, all of the server's characteristics are transferred over. Processing Blades from a pool can be brought into service when needed, and repurposed again when no longer needed. Thus the time that system capacity is compromised due to a failed node is reduced from days or hours to minutes.

Unbeatable Manageability

The amount of administrative time to manage a database system directly impacts an organization's TCO of that system. Managing an Oracle RAC installation involves:

- The initial hardware and software set up
- Ongoing software and hardware management
- Responses to system failures

The Egenera BladeFrame lowers the TCO for all these operations when compared to Oracle RAC configurations on competing x86-based hardware.

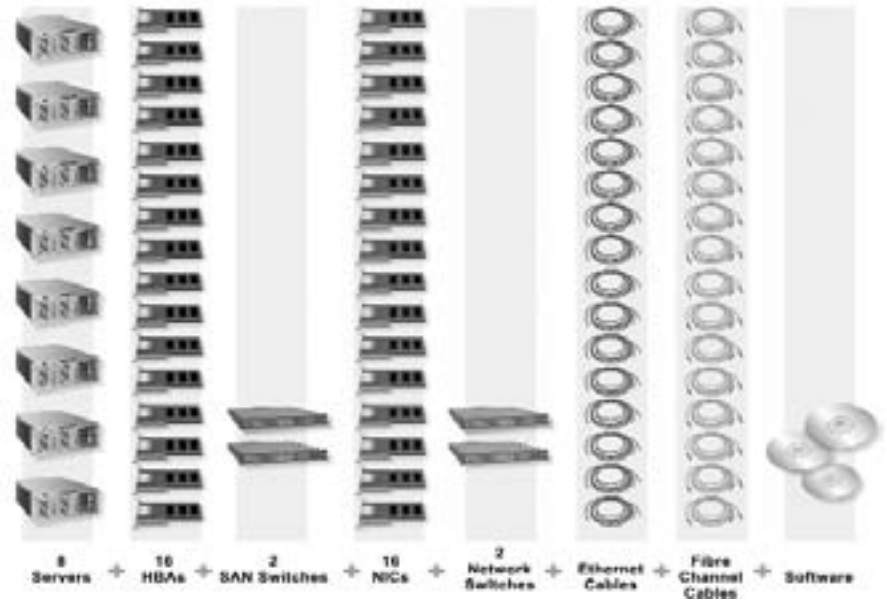


Figure 5: Hardware required for an eight-node Oracle RAC cluster on traditional x86 servers.

Setting up an eight-node cluster on traditional x86 hardware involves setting up and configuring eight servers, two interconnect switches, 24 NICs, 16 HBAs, two SAN switches, cables and drivers for all interconnections and devices.

Setting up the same configuration on the BladeFrame is significantly easier. Egenera's network virtualization reduces by 80 percent the hardware components required for standard server configurations. There are no NICs, HBAs or SAN switches to configure and cable. All configuration is done through software. By transforming manual, physical tasks into automated, software-based tasks the BladeFrame allows system administrators to significantly increase the span of servers under their control and rapidly set up an Oracle RAC installation.

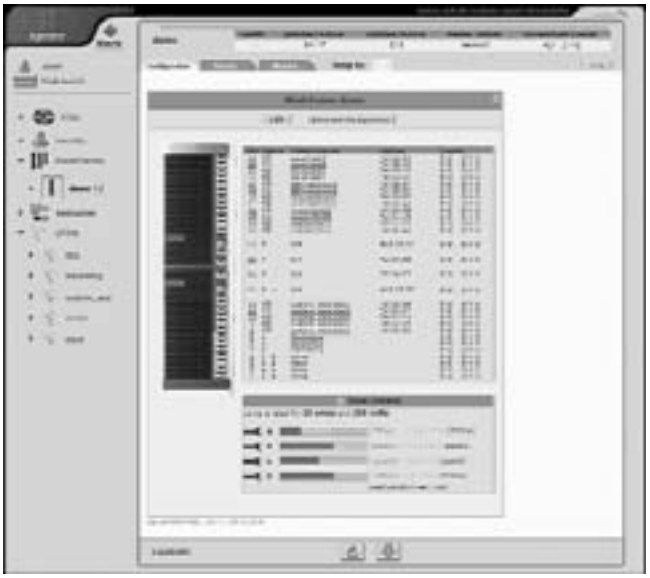


Figure 6: Egenera PAN Manager software manages all the virtualized server resources.

The BladeFrame also excels at ongoing system management. Its virtualization technology enables a customer to move an Oracle RAC database to a higher-performing blade, or add a node to a cluster, on-the-fly. In the case of sudden demand on the database, a customer using the BladeFrame can shift processing resources from less critical applications to the database in minutes.

In addition to keeping up with processor demand, keeping up with storage demands is also critical. As the database grows, additional disk resources are required. Egenera PAN Manager™ software enables users to easily reconfigure processing resources with additional storage resources without shutting down the database.

Finally, in response to failures, managing an Oracle RAC installation using a BladeFrame is simpler and easier than traditional server architectures. Because server resources are assigned through software, a failed Oracle RAC node can be easily reassigned to another Processing Blade in minutes.

This is not necessarily the case for Oracle RAC installations on traditional platforms. For example, if a two-node Oracle RAC configuration suffers a failure, the database's capacity is reduced by 50 percent. Assuming a hardware failure, returning the system to full capacity in a traditional environment would involve troubleshooting and replacing equipment to return the system to operation. All of these operations would take significant time to complete.

With the BladeFrame, the database could be restored to full capacity in minutes automatically. Utilizing the BladeFrame's global failover capability, the Control Blade™ modules can move an application and its associated network address and storage mappings to a new Processing Blade in seconds. Thus the database can be returned to full capacity in minutes, not hours, all through software. This dramatically reduces the TCO of managing the database.

Unbeatable Flexibility

The move to blade servers is a key component to lower TCO. Blade servers help lower TCO by consolidating power, network and storage connections, and by reducing real estate and power requirements.

Unfortunately, most of today's blade servers lack the flexibility required for applications such as Oracle RAC. This inflexibility comes from several sources:

- Integrated proprietary Fibre switches that are not certified with common SAN fabrics
- A limited number of network switches that do not support sufficient redundancy for Oracle RAC's interconnect
- A limited number of NICs per blade server creating the need for cache synchronization to occur across the public network connections
- Lack of support for high-end x86 processors

Egenera's BladeFrame is the only x86-based platform that can support Oracle RAC in a bladed form factor. Egenera's advantage comes from the combination of bladed form factor, stateless and anonymous processing resources, and virtualization capabilities. Each of these capabilities contribute to the flexibility of deploying Oracle RAC on the BladeFrame. Egenera has:

- No integrated physical Fibre switch. All SAN switching is accomplished through software. To a datacenter's SAN, the servers deployed on the BladeFrame look like hosts, not a proprietary switch.
- The ability to deploy over 1,000 virtual network switches per BladeFrame. Thus, for an application like Oracle RAC that requires complex network configuration, the BladeFrame can easily accommodate and ensure there are no single points of failure.
- Removed hard drives, NICs, HBAs and other components not needed due to the BladeFrame's virtualization capabilities. By removing these components and implementing a shared nothing cooling architecture, Egenera Processing Blades are able to support high-end CPUs in a 1U form factor, a feature unique to Egenera.

Blades are an important part of an organization's plan to reduce TCO in the datacenter. Only Egenera provides the flexibility needed to deploy Oracle RAC in a bladed form factor.

Unbeatable Price/Performance

Oracle RAC performance is directly linked to:

- CPU performance
- Interconnect bandwidth and latency

The BladeFrame's combination of high-performance x86 CPUs and a high-bandwidth, low-latency backplane provides superior price/performance for Oracle RAC. Egenera is committed to delivering Processing Blade configurations that include the fastest x86 CPUs in production.

CPU performance is half the equation for Oracle RAC performance. The interconnect is the other half. The interconnect serves as the database's central nervous system, ensuring that physically separate server caches can be coordinated using Oracle's Cache Fusion. A dedicated high-speed, low-latency, switched fabric is optimal for this type of communication.

The BladeFrame's internode communication speed and latency are superior to those of standard Ethernet. Redundant direct memory access interconnects provide a secure, point-to-point network between Processing Blades. In addition to bandwidth improvements, the direct memory access lowers the latency of the TCP/IP stack by up to 60 percent.

Demonstrating Oracle Real Application Clusters (RAC) on the BladeFrame: The Calling Circle Problem

Introduction

To demonstrate the capabilities of Oracle RAC and the Egenera BladeFrame, Egenera has conducted benchmark tests using Oracle-supplied problems. The results demonstrate how the BladeFrame and Oracle RAC are an unbeatable combination.

Problem Setup

The Calling Circle application represents a self-service OLTP application developed to test the capabilities of Oracle RAC. The application was developed in SQL, PL/SQL and Java by Michael Hallas and Dominic Giles of Oracle. The application models the customers of a telecommunications company registering, updating and inquiring on a calling circle of their most frequently called numbers in order to receive discounted call pricing. The Calling Circle application generates a heavy database workload and allows demonstration of failover/failback and load balancing capabilities.

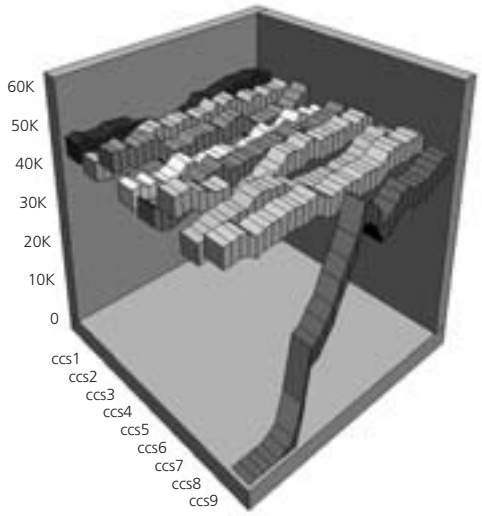
The Calling Circle application was originally developed to include aspects known to cause scalability challenges with Oracle Parallel Server. For example, the database features several keys generated from sequence numbers, resulting in contention for right growing indexes. The application also maintains activity counters and a history of changes to each Calling Circle, further increasing the proportion of insert and update statements.

The Calling Circle application simulates a randomized workload of customer transactions and measures transaction throughput and response times. The test workload is specifically designed to perform a high proportion of database changes to reveal any contention issues. In fact approximately 97 percent of the customer transactions cause at least one database update, with well over 75 percent performing two or more updates.

Results

The Calling Circle problem validated the unbeatable combination of Oracle RAC and the Egenera BladeFrame. The application was run on two to nine four-way, 1.6Ghz, Intel-based Egenera Processing Blade™ resources running Red Hat Advanced Server.

Performance



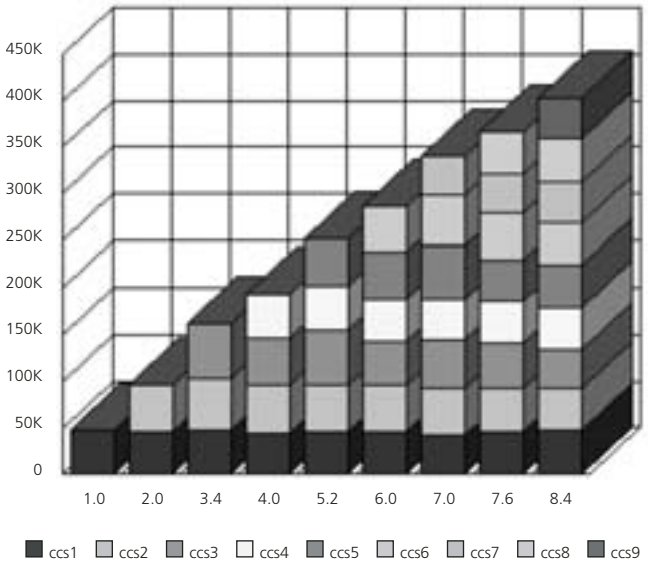
Transactions Per Minute (TPM)

Total Transactions Per Minute: 406895

Operating on nine nodes, the Calling Circle results were 414K transactions/minute.

Scalability

The application was 93 percent scalable at nine nodes. More importantly, using Egenera's virtualization services, the node count was increased and decreased on the fly, all through software. By repurposing the Processing Blades as needed, the application was supplied with the right amount of processing power at the right time.



Node Scalability

Manageability

Compared to traditional Oracle RAC implementations, the nine-node BladeFrame implementation represents an 80 percent reduction in time to manage the configuration.

Flexibility

The Calling Circle problem was run on four-way blades, SAN connected, with completely redundant interconnects, data, power and network connections. Egenera supplies the only bladed form factor system capable of meeting these configuration specifications.

Availability

To demonstrate the high availability of the BladeFrame, nodes were deliberately failed by removing the Processing Blades. The BladeFrame's virtualization services restarted the failed node of RAC on a different server by transferring the virtual network and storage configurations to the new node.

Conclusions

The Egenera BladeFrame and its PAN architecture provide an unbeatable platform for Oracle Real Application Clusters. The physical features of the BladeFrame's hardware and the virtual advantages of its software provide a natural and optimized synergy with Oracle RAC. The resulting benefit is a significantly lower total cost of ownership while providing for unbeatable scalability, availability, performance, flexibility and manageability.

egenera®

Corporate Headquarters
Egenera, Inc.
165 Forest Street
Marlboro, MA 01752
U.S.A.
Phone: 508-858-2600
Fax: 508-481-3114
www.egenera.com

European Headquarters
Egenera Ltd.
Venture House
Arlington Square
Bracknell, Berkshire RG12 1WA
United Kingdom
Phone: +44 (0)1344 475237
Fax: +44 (0)8703 305946
www.egenera.com

Asia Pacific Headquarters
Egenera K.K.
Shinjuku NS Bldg. 6F,
2-4-1 Nishishinjuku,
Shinjuku-ku
Tokyo 163-0806 Japan
Phone: +81-3-5321-7157
Fax: +81-3-5321-7158
www.egenera.com