

An Oracle Business White Paper
July 2013

Solving Data Management and Scalability Challenges with Oracle Coherence

Superior Customer Experiences through Smarter Data Management

Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

Table of Contents

Executive Overview	2
The Problem: Data Management Challenges in the Era of Big Data, Mobile and Cloud Computing	4
The Solution: In-Memory Data Grid.....	5
Introducing Oracle Coherence 12c.....	6
Case Study #1: Managing Growth.....	8
Discount Travel Site Conquers Mobile Onslaught	9
Case Study #2: Improving Website Availability	10
E-Tailer Manages Traffic Spikes	11
Case Study #3: Improved Performance for Internal Applications	11
Leading Financial Services Firm Improves Response Time.....	12
The Oracle Coherence Advantage.....	13
Fundamental Capabilities of Oracle Coherence	14
Conclusion	16

Executive Overview

Customer-facing websites and applications now handle a sizeable portion of the world's business via self-service portals, e-commerce engines, and every conceivable type of information service. Cloud and mobile computing paradigms have taken hold, placing an additional burden on these data-driven assets. Many companies are forced to recognize that their IT infrastructure can't scale to meet today's mobile and cloud demands. Business and consumer interactions that were formerly confined to desktop computers are now initiated continuously from billions of mobile devices, resulting in much more traffic into public-facing websites and applications. For example, instead of simply confirming a flight from home, today's travelers are likely to check flight status multiple times while enroute to the airport—and even from the waiting room of the airport terminal. This escalating volume of requests places crushing demands on data center resources including servers, applications, and databases.



How are *your* information systems faring? Just because your applications are up and running does not mean they are appealing and effective for your consumers. Once an application is in place and the user base starts growing rapidly, the technical infrastructure backing it becomes even more important. Site uptime, quick responsiveness, and reliability are critical to success—as is the ability to add enhancements.

With the Oracle Fusion Middleware family, Oracle brings a robust set of offerings to ensure guaranteed reliability and performance for your applications—and to your data center as a whole. As an integral component of Oracle Fusion Middleware, Oracle Cloud Application Foundation uniquely provides businesses with a mission-critical Cloud platform to expand and manage their business. As the #1 in-memory data grid and caching solution across conventional and cloud environments, Oracle Coherence is a critical part of Cloud Application Foundation's technology portfolio. Enterprises around the world depend on this unique data grid software for immediate access to frequently used data. With support for Java, .NET, C++ and REST clients, Coherence provides data caching, data replication, real-time event

processing and distributed computing services to ensure scalability, availability, reliability, and performance. It relieves overloaded data sources, simplifying management and improving performance for queries, transactions and events. By automatically backing up partitioned data across clusters, Oracle Coherence provides site-safe and rack-safe guarantees, giving you continuous data availability and transactional integrity, even in large, multi-data center environments. With strategic integration between Coherence and other Oracle offerings such as WebLogic Server, Database 12c and Oracle's robust development tools, businesses are ensured a seamless, efficient adoption of and interoperability with existing systems.

Oracle Coherence is a highly-scalable and fault-tolerant distributed cache engine. Coherence can harness commodity servers and/or Oracle Exalogic Elastic Cloud to address the most demanding applications, and create a cluster that can be seamlessly expanded to add more memory, processing power or both. Coherence has no single point of failure and transparently fails over if a cluster member fails. Additionally, by offloading more expensive shared services, Coherence helps customers save costs and has been proven to achieve ROI in as little as two months for customers. (Ref: OP-Pohjola)

In this paper, we will explain how Oracle Coherence can bring your information systems into the modern age and position you for future growth.

The Problem: Data Management Challenges in the Era of Big Data, Mobile and Cloud Computing

Technology itself cannot make a business application successful. However, lack of robust technology can create a make-or-break experience for your customers. Having solid backend technology is a necessary condition for success for mission-critical applications no matter their intended interface. The enabling technologies typically include a web portal, application window or another user interface (UI) tool for the presentation of the content, an application server that hosts the application logic and powers the user experience, and a database server to support persistent create, read, update, and delete (CRUD) operations.

Data availability and management is also critical for any application. The user experience depends in large part on the availability and management of session data, including its persistence and management. State-of the art data management technologies enable the following:

- Constant customer connections via social networks and mobile devices
- Outstanding performance, uptime, responsiveness, and reliability
- Cohesive, consistent, well-integrated experiences for customers
- The ability to introduce new offers and services in real time; on-premise and in the cloud
- Support for business growth, fluctuating customer demands, and an unpredictable supply chain

For many applications, the pertinent data resides in the database or a backend operational data store, often on a mainframe computer that was never intended to handle the increased demand it is now tasked with. One of the leading problems is shared services overloading business infrastructures, creating latency issues, unpredictable outages and poor customer experiences. Further, these back-end services cannot scale easily and cost effectively as the business grows and buying patterns change. All of these issues can lead to a poor user experience and subsequent loss of business.

The Solution: In-Memory Data Grid

Data grid software is middleware that reliably manages data objects in memory across many servers. Oracle Coherence is an in-memory data grid solution that enables organizations to predictably scale mission-critical applications by providing fast access to frequently used data, and process that data in real time. It provides a robust scale-out data abstraction layer that brokers the supply and demand of data between applications and data sources, on premise or in the cloud.

Oracle Coherence caches data in memory in a distributed manner. When a Coherence server is added or removed, the associated cluster automatically re-balances the workload among all available nodes. As a result, Coherence also provides a horizontally scalable infrastructure for managing application data to protect individual machines, racks and data centers. This sophisticated data management solution for mission-critical applications lets you store data reliably in the application tier, in a form that can be readily consumed by the application. Data in a Coherence cache is synced continually with the primary database so it is always up to date.

An in-memory data grid improves reliability, scalability and performance compared to traditional relational database management systems. But not all in-memory data grids are created equal. Oracle Coherence has a fault tolerant architecture that ensures reliability, predictable and linear scalability. By locating data side-by-side with applications, data can be accessed much faster. Some analysts use the term *fast data* (which we will refer to later in this paper) or *live data* to indicate its up-to-the-moment relevance. But caching alone is no longer enough for today's real-time application scenarios. That's why Oracle developed GoldenGate HotCache for Coherence that keeps data grids always-accurate and up to date—a constant reflection of real-time activities. This is an especially important goal for mobile apps that involve context-specific information that must be acted upon immediately.

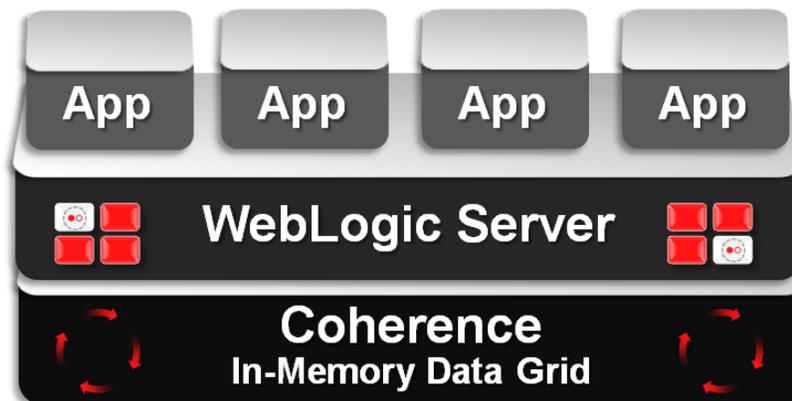
Oracle Coherence is the solution for data that needs to be accessed and updated instantly by the associated applications. As a shared infrastructure, Coherence combines data locality with local processing power to perform real-time data analysis, in memory grid computations, and parallel transaction processing. The load, risk, and reliance on backend systems are diminished, and data can be written to the back-end systems either synchronously or asynchronously, giving corporations flexible options for data management.

Oracle Coherence is bolstered by tight integration with the Oracle Cloud Application Foundation portfolio, including Oracle WebLogic Server, the application server of choice for many business and cloud applications. Coherence has also specifically been optimized to work tightly with Oracle Exalogic Elastic Cloud to achieve up to six times lower latency with half the hardware required.

Introducing Oracle Coherence 12c

Oracle Coherence 12c allows you to deliver next-generation applications on your mission-critical platform while dynamically scaling your applications to meet increased cloud and mobile demands.

Coherence Live Events allows you keep customer experiences rich and lag-free by powering application with always-accurate, real-time data. GoldenGate HotCache keeps your database and Coherence cached data in sync so there's never "stale data" being delivered to your applications from an outdated cache. Managed Coherence Servers integrate seamlessly with Oracle WebLogic to provide a managed, templated environment for dynamic scaling. Coherence also automatically adapts to changes in traffic or unpredictable user spikes to keep customer experiences consistent. Applications can be deployed on Oracle Exalogic Elastic Cloud to obtain cost and performance benefits of engineered systems.



Coherence 12c ensures your data is up to date in real-time, and integrates seamlessly with Oracle WebLogic Server 12c to deliver dynamic applications.

Coherence 12c also provides overarching simplification of operations through native cloud management.

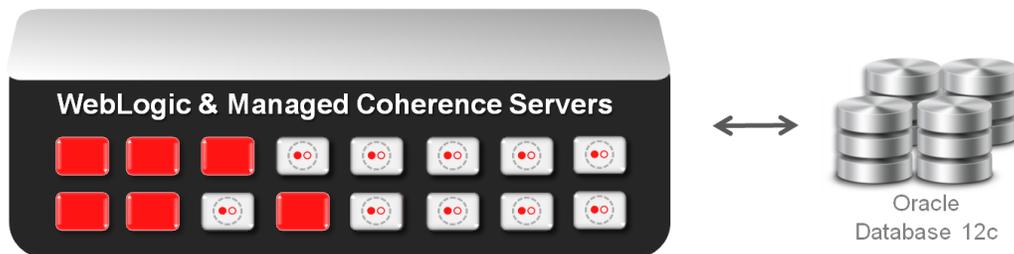
Broad integration with other Fusion Middleware offerings will help you reduce costs and streamline operational activities with a range of new features. For example, you can consolidate middleware software instances to a common framework for installing, upgrading, and patching software. The WebLogic Management Framework streamlines cluster lifecycle management and provides a single stop for management of both WebLogic and Coherence instances. And adherence to industry standards simplifies integration with Maven, GitHub and other open source software and third-party frameworks.

Further, Coherence 12c provides enhanced deployment options for high availability.

If you have a Coherence cluster that spans multiple racks within a data center you can partition the load to consider the entire topology of the cluster when backing up data—from machine to rack to site. Large installations can guarantee performance throughout multiple data centers with these site-safe and rack-safe enhancements.

The Coherence GoldenGate HotCache feature ensures that you will always deliver real-time, always-accurate data to your applications without data loss due to a "stale cache."

Overall, Coherence 12c is designed to deliver efficiencies in both performance and management, while allowing you to scale your applications to meet increasing mobile and cloud demands on your infrastructure.



Cloud Application Foundation 12c allows you to manage both WebLogic Server clusters and Coherence Servers in a unified console.

Case Study #1: Managing Growth

Many businesses find their applications are growing faster than their supporting technology can keep up with. Although it's a welcome problem, the challenges of growth are numerous:

- Growth is unpredictable. Businesses assume that they will experience a certain amount of growth within a certain amount of time in certain areas. Nobody really knows if it will happen exactly as planned. This leads to uncertainty in the investments that the business is willing to make to support that growth. Too much will lead to high cost and excess capacity. Too little would lead to risk of underperforming.
- Enterprises cannot ignore business continuity. It's important to stay focused on supporting customers with a set of stalwart information systems and the operational capabilities to support them.

Businesses are looking for infrastructure technology that will scale and grow easily and cost effectively. They also want an architecture and supporting infrastructure that will support their requirements for website performance, uptime, and reliability as the business grows.

Customer-facing websites and applications are particularly prone to scalability issues. For example, in the travel industry, for every actual booking there are hundreds of sessions where people merely consider potential routes, costs, and possibilities. Many people interact with airlines and hotels, yet only a fraction of those sessions result in a booking and subsequent revenue. Due to this high "look to book" ratio, the *volume* of traffic is going up dramatically while the *value* of each transaction is going down. Thus, the cost per transaction must go down, or IT costs will increase dramatically.

This phenomenon began with the Web but the proliferation of mobile applications is accelerating the problem. Websites and Web-based applications that easily supported yesterday's traditional Internet users are being overwhelmed by the growth of today's mobile users. For example, a user of yesteryear may have gone online to create a diet and exercise plan, entering in their own data on a simple website. A user of today will set timers, goals and provide data on a smartphone that logs their performance and communicates constantly with the exercise application's back-end. This type of phenomenon will only increase as smartphone and tablet sales continue to accelerate. There will soon be 7.4 billion wireless devices in the market and 1.2 billion

smartphones will enter the market over the next five years, according to ABI Research.¹

Discount Travel Site Conquers Mobile Data Growth

Systems that meet customer expectations of fast response times struggle to scale to meet those expectations as more people use IT services. This leads to the perception that the system has performance issues, when in fact it has scalability issues. To address this, scalability must be baked into the architecture. With the boom in mobile devices, etc, this is an imperative today.

To solve this growth issue, you need an application server and load-balancing environment that can scale to meet these requirements. Such was the conclusion for one of the world's largest travel websites, which provides technology and services to hotel companies and travel distributors. As its customer base and search volumes increased, the application infrastructure needed a more scalable solution to support its continued growth. These problems were exacerbated with the rise of mobile users and the ease of perusing travel options, raising the look-to-book ratio from 4:1 to 4500:1. After evaluating multiple alternatives and performing extensive testing, IT leaders at this travel company chose to implement Oracle Coherence and other Oracle Fusion Middleware components including Oracle WebLogic Server. Their goal was to provide a more cost-effective, scalable and responsive application infrastructure to support the 24x7 demands being placed on their infrastructure.

In less than a month, the company upgraded its application infrastructure to accommodate a huge surge in queries from mobile users. They are now able to support millions of customers performing hundreds of thousand searches per day while maintaining optimum performance during peak traffic periods. Oracle Coherence provides the in memory performance and distributed, real-time caching capabilities that alleviate bottlenecks, reduce data contention and improve application responsiveness. It also allows them to scale linearly and dynamically for more predictable performance and improved resource utilization moving forward.

Oracle Coherence maintains frequently accessed data—such as specific deals customers are investigating that were advertised in the site's weekly newsletter—in an “availability cache.” This caching architecture reduces network traffic and database loads, resulting in better overall performance across the technology stack. Additionally, Oracle Coherence helps ensure that the site remains at peak performance around the clock by automatically and transparently failing over and

¹ Forbes, “The Latest Infographics: Mobile Business Statistics For 2012” (May 2, 2012)

redistributing cached data across the grid when a server becomes inoperable or is disconnected from the network. By taking advantage of the Oracle Coherence Elastic Data features, the travel company can offload cached objects from the Java Virtual Machine heap onto commodity disks, dramatically reducing the cost of processing these transactions, both for the travel company and for its hotel partners.

Online travel is just one of many examples of an industry that grew faster than the technology supporting it. By deploying an in-memory data grid and an industrial-strength application server on commodity hardware, this company increased customer response time by 20-30x, supporting business and customer loyalty goals.

Customers such as this will look forward to adopting the recent Coherence 12c release, which provides increased operational efficiencies through native cloud management, including automatically adapting to spikes in user traffic and full integration with Oracle WebLogic Server and Oracle Exalogic.

Case Study #2: Improving Website Availability

In many industries, the Web has reduced customer-switching cost to nearly zero—literally a mouse click for many businesses who are being forced to design and support applications that engage and hold customer interest. If not addressed, customer attrition will be a recurring phenomenon, with a significant impact on business revenue.

“Site response and uptime are our two most important business metrics,” says the director of engineering for a leading online social network. Poor site response performance means that the page takes too long to render and visitors run out of patience. Visitors then lose interest in the site and typically go to sites with comparable or competing values. Poor uptime also means that the website fails to be continuously available. Negative customer experiences lead to a declining number of visits, a declining number of transactions per visit, and lost revenue.

There are some common causes of these IT challenges:

- Shared or partner services are sometimes unavailable due to network outages or a reliance upon overtasked shared services
- Some components of the application infrastructure are overloaded – possibly because the application is experiencing an unexpected traffic spike

E-Tailer Manages Traffic Spikes

One retailer correctly predicted that the infrastructure supporting its web application would not be able to support the tremendous volume of transactions that they foresaw on Cyber Monday. Specifically, they identified that the availability of data and data operations would be a potential source of bottlenecks and contention. The business was at risk of poor responsiveness, leading to customer dissatisfaction, lost transactions, and lost customers.

This retailer chose Oracle Coherence to provide distributed caching for its web applications. Start-to-finish, the Oracle Coherence implementation took just two months. Coherence provides a scalable caching solution, serving as a buffer between the e-tailer's presentation layer and its back end search layer (powered by Oracle Database), thus improving response times for frequently searched items. It delivers predictable performance, scalability, resilience and availability. It enables the e-commerce site to deliver fast and reliable access to frequently searched data, improving customer search response times by nearly 100 percent and guided navigation response times by over 200 percent. These improvements correlate to increased conversion rates and revenue.

Customers such as this e-tailer will be excited to see enhanced deployment options in Oracle Coherence 12c to provide high availability of their data, including site-safe and rack-safe enhancements. Further, GoldenGate HotCache will provide customers with real-time updates from their database to their Coherence cache, ensuring that their most current data is always available.

Case Study #3: Improved Performance for Internal Applications

Customers are quick to form opinions about a company based off of quick interactions with company websites. This opinion is also significantly impacted by internal operations as supported by a couple of different types of applications:

- Applications used while interacting with customers in real time such as a retailer's customer service management application, a telecom service providers' order entry application, or an insurance company's loan processing application
- Applications that provide customer service or customer fulfillment related functions such as trouble ticketing systems and order fulfillment systems

Many times, these applications are not able to keep up with the demands of the business. The hardware components and the compute resources that support these applications – the application server, the database, and the portal server – are not able to support a transaction volume that is heavier than that for which the system was originally designed. As a result, the system is often overloaded. This has a direct impact on the customer as shown by the following examples:

- A customer's order with an online book retailer is not fulfilled on time. This is because the order processing system could not support the increased volume of orders entered during this particular customer's order.
- It takes too long for customer service representative to take down an order, displeasing the prospect or even losing them.
- A loan processing application takes too long, having a negative impact on revenue (the sooner the loan is processed, the sooner the revenue stream gets started) as well as customer satisfaction.

Leading Financial Services Firm Improves Response Time

One leading global financial services firm, with assets exceeding \$2 trillion and operations in more than 60 countries, had a portal application that hosted and displayed its customers' account information. Account managers would interact with this portal application in real time while the customer was physically present in a bank branch or on the phone. The firm was facing significant challenges related to the performance of a dashboard that displayed account information resulting in long wait times for customers. The rate of transactions going through the system was lower than expected, since it took a long time to retrieve data from a mainframe database. The dashboard application, had to access the mainframe system and database for every account look-up, and session data had to be maintained.

This customer implemented Oracle Coherence for more efficient data management. Coherence also handles session management, including data sharing and preservation. The dashboard application has an unlimited amount of memory on which to persist user sessions, which reside in the Coherence cache and can be updated in real time. There is no need to access the backend database for online transactions or account access. The user, transaction, and session data are uploaded to the backend system as batch updates. The financial services company did not have

to change any of its existing applications, which accelerated the implementation and minimized business risk.

Coherence has brought unlimited scalability to this application. Even though the transaction rate has doubled, the response time has improved by 60 percent. The organization can support more data and additional traffic simply by scaling to add servers to the Coherence data grid.

In addition to improved performance, Oracle Coherence improves disaster recovery for this mission-critical application since there is no single point of failure in the entire data persistence layer. Coherence nodes can interact across local and wide-area networks. As a result, if disaster strikes a particular data center, the Coherence nodes in other geographic locations can take over and the operational data for the business continues to be available. When the affected data center comes back up, Coherence nodes can be synced up with other locations.

Oracle Coherence 12c will allow customers like this financial firm to scale to meet ongoing mobile and cloud demands. Through Live Events, their customers will be able to interact with applications that employ real-time processing. Further, deep integration with Exalogic in 12c will allow their mission-critical applications to respond up to 6x faster on as little as half the hardware.

The Oracle Coherence Advantage

Enterprises tend to invest heavily in mission-critical applications because the cost of failure is so high. However, most IT investments are carefully scrutinized. Senior managers want to see solid business cases and, for many projects, a direct ROI.

Imagine having an architecture that provides capacity on demand. Imagine protecting against the cost of expanding your shared services to meet growing business challenges. Imagine having information systems that scale with your business opportunity, whether that growth is from 1 million to 10 million users or 10 million to 100 million users and large multi-data center deployments.

With virtually unlimited scalability, Oracle Coherence is designed to support growth of this magnitude. It has a small footprint and utilizes commodity hardware - or can be paired with Oracle Exalogic for up to 6x speed increases - driving additional cost savings. As business needs change, it's easy to re-deploy Coherence nodes to support new applications and opportunities.

Fundamental Capabilities of Oracle Coherence

Performance – Oracle Coherence solves latency problems and drives dramatic increases in performance by caching and processing data in real time. In memory performance alleviates bottlenecks and reduces data contention, improving application responsiveness. Parallel query and computation also improves performance for data-based calculations.

Reliability – Oracle Coherence is fault-tolerant, providing data reliability, accuracy and consistency. Organizations can meet data availability demands in mission-critical environments with its support for data tolerance and continuous operation. The reliability of the data grid minimizes the need for applications to compensate for server and network failures, streamlining the development and deployment process.

Scalability – Oracle Coherence enables applications to scale linearly and dynamically for predictable cost and improved resource utilization. For many applications, it offers a straightforward approach to increasing the effective capacity of shared data sources. Oracle Coherence handles continually growing application loads without risking data loss or interruption of service.

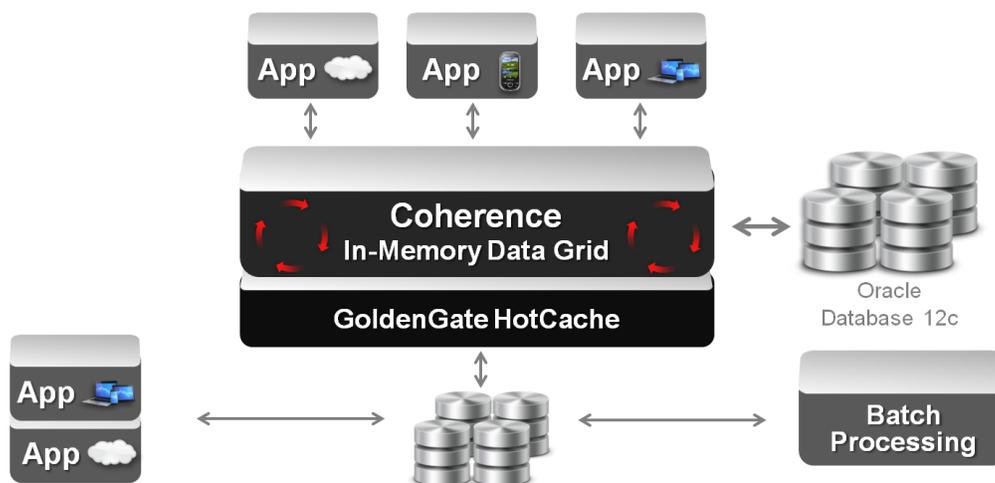
Transactional Integrity – Applications manage transactional data in memory inside the data grid. Coherence is able to manage transactions in memory until they are persisted to an external data source for archiving and reporting.

Multi-language support – Oracle Coherence can support applications developed on Java, C/C++ and .NET, as well as REST interface for other language support. Further, Coherence enables data to be shared amongst applications in written in disparate languages. This reflects the true state of customer applications that have been developed on different systems at different points in time by different organizations and now may have to ‘come together’ to support applications that depends on all of these.

Disaster recovery – With its capability to duplicate data and maintain transactional integrity, Coherence can serve as a great tool for disaster recovery. Coherence clusters maintained throughout the enterprise and across geographies constitute an automatic ‘backup store’ for organizational data. Individual machines, racks and data centers can all be flexibly backed up to ensure customer availability.

Supporting and Driving Fast Data - As mobile and the "internet of things" continue to grow in scope, the need for not just processing large volumes of data is on the rise, but also the need to process this data at high velocity. Coherence is uniquely capable of persisting and triggering the processing of data in your middle tier to allow you to keep your data moving at high velocity. As *Big Data* grows, so the need for *Fast Data* increases. A dedicated Oracle white paper on the Fast Data topic can be viewed [here](#).

Integration with Oracle Fusion Middleware Products – Oracle Coherence comes pre-integrated with many products, such as Oracle WebLogic Server, ATG, PeopleSoft and others out of the box. Coherence's management functionality is bolstered by Oracle Enterprise Manager, which supports all Oracle Fusion Middleware management. Users don't have to learn new management tools and efficiency is improved by looking at a 'single pane of glass'.



Oracle Coherence data grid provides an operational data persistence capability in the application tier. It front-ends traditional data management technologies like databases and mainframes, with GoldenGate HotCache keeping your data synced in real time. Coherence can interact with different types of applications including real time clients, web clients, Web services, and SOA clients.

Conclusion

Data management is a critical issue for nearly every business. Customers expect an accurate, well-integrated experience even as you introduce new features, offers and services. To stay competitive you need to be able to take risks, track success, and take corrective action quickly—all while supporting continual growth and new paradigms driven by mobile computing, big data, the internet of things and cloud computing.

Traditional methods of persisting and managing data in mainframe or legacy systems can prove prohibitive from a cost, risk, growth, and business continuity standpoint. In-memory data grids are smart, cost effective, and non-disruptive technologies that provide a viable data management solution for your business.

Oracle Coherence is the market's leading in-memory data grid solution. It can scale linearly - including to cloud-based deployments - as the business grows while mitigating IT costs and relieving the load and risk posed to backend data repositories. As part of the Oracle Fusion Middleware product family, Oracle Coherence naturally integrates with many other Oracle products. With Oracle Coherence, you can focus on business issues, application logic, and critical success factors rather than being burdened with the limitations of legacy technology.





Solving Data Management and Scalability
Challenges with Oracle Coherence

July 2013

Author: Jens Eckels

Contributing Authors: David Baum

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com



Oracle is committed to developing practices and products that help protect the environment

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.

This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group. 0613

Hardware and Software, Engineered to Work Together