hazelcast®

In-Memory Databases Versus In-Memory Data Grids

A Business-Level Perspective on In-Memory Computing

White Paper

You might be familiar with databases like Redis and Aerospike, or data grids like Hazelcast and Coherence. But how are the two classes of technologies different? It is first important to review why these in-memory technologies even matter today.

Businesses are operating in an era where technology is advancing and evolving at ever faster rates. For business and IT leaders, it can be overwhelming to keep up with new developments and determine what is critical to implement immediately, and what technologies can or should wait. This situation is often complicated by passionate IT staff and architects promoting the latest and greatest new options.

Companies that follow the bright and shiny objects and respond to the voices that make the most noise are likely to spend more money and time than they need to and may not achieve the desired results. Seasoned IT leaders know that **technology decisions should be grounded in two things: customer experience and business needs**. They also understand that implementing new technology can be both costly and disruptive – starting over with a new architecture or new set of IT solutions isn't typically feasible. New business needs and customer experiences often must be enabled on top of existing systems and solutions.

> *In-Memory Databases (IMDB) and In-Memory Data Grids (IMDG) are two technologies available to IT teams to address real-time computing and big data needs without having to start over with an entirely new set of IT systems.*

The Business Need for Real-Time Data Processing

Increasingly, customer experience is the primary goal driving technology decisions in modern companies because it drives competitive advantage. Customer service consumption models have changed due to the increasing availability of data and the proliferation of different engagement opportunities. Customers expect the same level of service, speed, and information whether it be through face-to-face, online or mobile-first interactions. **The experiences companies create for customers are what keeps them coming back for more and recommending the company to friends.**

Customers abandoning a transaction because a company's website isn't responding fast enough is more than a technical issue – it's a business issue.

The underpinning thing customers are looking for in these experiences is speed. They understand a company's IT environments are complex and data is complicated, but they really don't care. They expect interactions to be immediate and get frustrated when this expectation isn't met. Customers abandoning a transaction because a company's website isn't responding fast enough is more than a technical issue – it's a business issue. **Customers expect interactions to be digital, connected, always available, and immediate.**

Customers want interactions immediately, but it doesn't change the reality that there are often complex data processing tasks that need to take place behind the scenes for those interactions to be effective. The challenge for companies is how to make these modern interactions happen, not just for a single customer's experience, but for potentially tens of thousands of customers who all want immediate interactions at the same time. The situation is complicated further when the experience includes multiple IT components and data validation/processing tasks are involved. Companies need to provide applications with limitless scale that are both experience-oriented and provide in-the-moment data.

The need for speed isn't confined to customer interactions. Modern businesses don't have the luxury of processing delays internally either. Streaming data from IoT devices,e-commerce interactions, monitoring market changes, and real-time fraud detection are examples of scenarios where business processes and confident decision-making are dependent on complete and current data. The past few years have seen the rise of billions of devices and services streaming in data to companies' IT systems non-stop. In order to take advantage of this data, use it to make informed decisions, and enable customer interactions, it needs to be processed in real-time.

Speed Is Great, but Speed at *Scale* Is What You Need

Customer experiences and business workflows require data to be processed quickly, but speed isn't the only factor that must be considered. A balanced solution that enables speed, scalability, and stability is essential.

Data processing delays slow down business activities, frustrate customers, and impede responsiveness to market opportunities.

Speed – Streaming data and customer experience need highly efficient real-time data processing that provides results within a fraction of a second. Data processing delays slow down business activities, frustrate customers, and impede responsiveness to market opportunities.

Scalability – Speed must be enabled for activities where tens of thousands of transactions take place at the same time. For some businesses this will be continuous volume, for others flexibility is required to scale up for peak processing times, such as Black Friday or a product launch, and then scaling down to avoid under-utilized capacity.

Stability – It's great to have systems that are fast and scalable when everything is working perfectly, but technology breaks. Stability and dependability are essential to avoiding business disruptions and instilling customer confidence.

Continuous Intelligence – Decision-makers need access to data in real-time from all available data sources in order to achieve the business agility required to excel in a rapidly changing marketplace. Databases alone aren't enough if continuous intelligence is required.

The Differences Between IMDB vs. IMDG

In-memory databases (IMDBs) and **in-memory data grids (IMDGs)** are the two primary in-memory computing solutions available today. Redis and Aerospike are popular IMDB technologies, and Hazelcast is a leading IMDG technology. While both types of technologies perform data processing in-memory instead of storing data on disks, the way they operate is different and each has its own advantages and limitations.

The major benefit of this technology is that data access is much faster since data can be retrieved from memory significantly faster than from disk drives or solid-state drives.

In-Memory Databases

In-memory databases are essentially a new approach to traditional relational database management systems with the difference being that they store data in-memory instead of on drives. **The major benefit of this technology is that data access is much faster since data can be retrieved from memory significantly faster than from disk drives or solid-state drives.**

The limitation of IMDB technology is that it only provides fast data retrieval, and not faster computation. And the fast retrieval only comes from memory-based versus disk-based access, as the network connection to the IMDB remains a potential bottleneck. All of this means that only one aspect of business applications can run faster with the integration of IMDBs. All CPU-based activities will not benefit from the deployment of an IMDB across a cluster of hardware servers.

IMDBs work well for smaller applications where the data is relatively contained and the processing on the data is relatively limited. IMDBs are therefore great for situations where fast data lookups, like for purchase history or customer information, are essential. IMDBs are popular for basic caching deployments where most of the data is needed for read-only uses.

In-Memory Data Grids

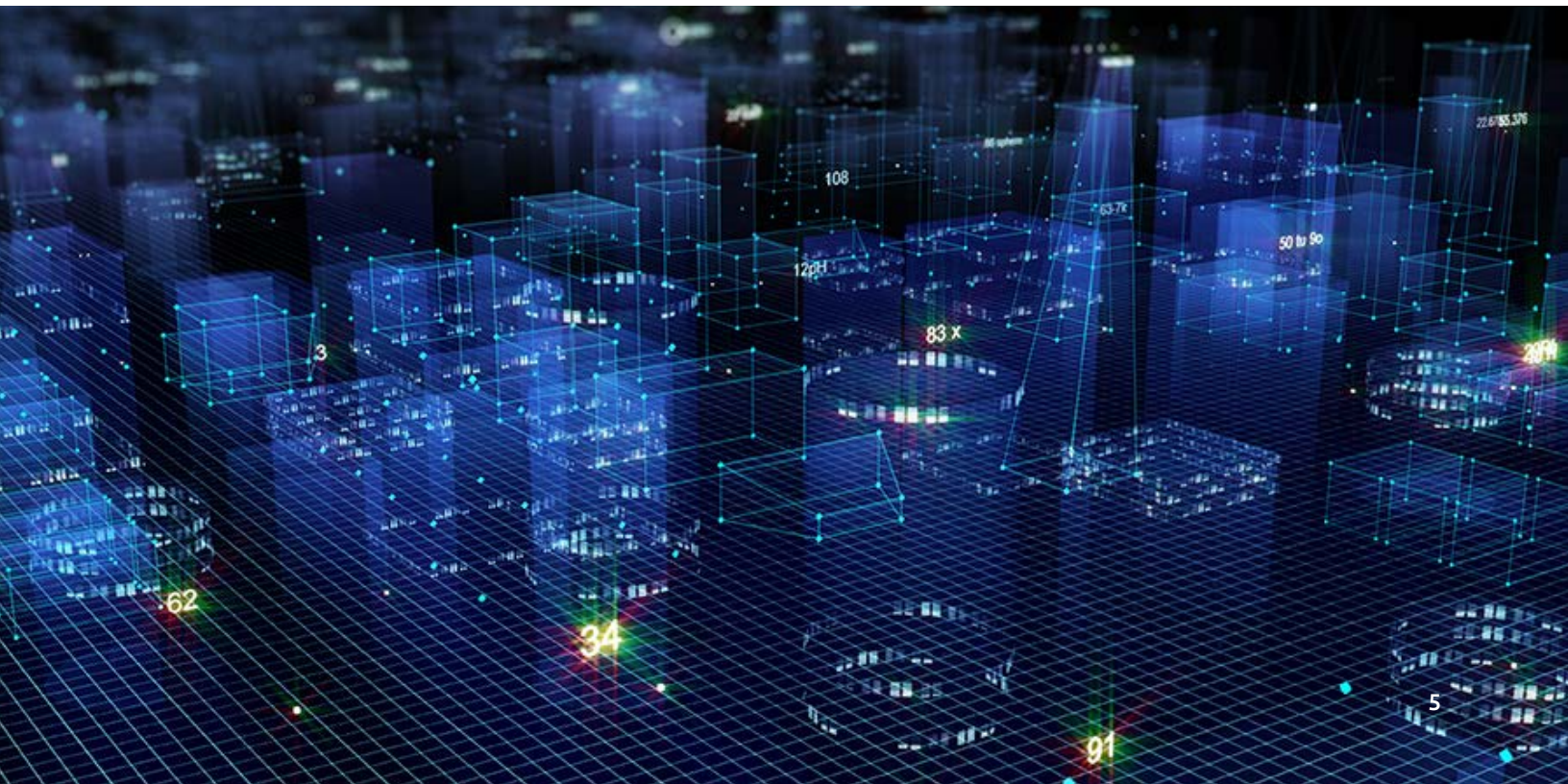
In-memory data grids take a different approach than IMDBs in that IMDGs are about more than just data storage and retrieval. They provide distributed computation to your applications on top of in-memory storage. This means you get fast data access just like IMDBs, but you also can take advantage of the collective CPU processing power of all the hardware servers in your cluster. In an IMDG, your application code is delivered as "jobs" to each server, and each job (i.e., "application instance") is then executed on those servers. Your application is designed so that each instance only needs to access the in-memory data on the server in which the instance is running, which improves performance even further because no network accesses are required. In cases where data from another server is needed, the "near-cache" feature makes a copy of that data to the local server so that future accesses to that data do not require a network call.

IMDGs make it easy to write distributed applications, so you don't have to reinvent the distributed logic for every application. And the advantage over IMDBs is clear -- instead of a single application running all the calculations on a single machine, you get multiple machines each doing a subset of the overall work. This divide-and-conquer approach leverages parallelism that can accomplish tasks much quicker. Since IMDBs do not have the distributed processing component, they are relegated to a limited set of tasks.

IMDGs are useful for environments where a significant amount of data processing or computation is required. Large-scale number crunching, simulations, data enrichment, data transformation, etc., require the collective processing power of IMDGs. For these use cases, the bottleneck is not only the data retrieval, but also the data processing, and IMDGs can help alleviate both of those chokepoints.

IMDGs tend to be limited by data set size, as they are often intended for fast processing of data that is otherwise stored long-term in other databases. This means that you will generally deal with terabytes of data, but not necessarily hundreds of terabytes. If you need larger volumes of data, then you would store the data in a database as the system of record, but do the fast processing in an IMDG.

IMDGs are about more than just data storage and retrieval. They provide distributed computation to your applications on top of in-memory storage.



Why You Need an IMDG

There are many instances where data needs to be processed rapidly, requires a lot of scale, and can be done as independent data groupings. Once the processing is done and the results are returned, the action is complete. IMDGs are ideal for this type of scenario. There are other instances where data needs to be recorded and used later, either as part of a system of record or to enable future activities in a workflow (persistence). Traditional databases are well suited for this type of scenario, but most customer experiences and business processes involve both scenarios. **Having a relational database coupled with an IMDG enables you to optimize your end-to-end transaction process.**

Deploying IMDG in Different Environments

Most companies today don't have a homogeneous IT environment that is entirely on-premises, entirely in the cloud, or even a hybrid. They have a complex and constantly changing landscape which includes legacy systems and new capabilities. When it comes to providing services and solutions to customers and business teams, companies need capabilities that work in all these environments because at the end of the day, transactions just need to flow quickly and smoothly throughout the organization.

An IMDG can be used to enable individual applications for in-memory processing – a focused technique for addressing specific performance and scalability pain points. **This low-cost solution is where many IT leaders start to demonstrate the value of in-memory computing to their executives.** An IMDG can also be used in an enterprise context to provide performance optimization for end-to-end transactions, manage streaming data, and integrate data sources for analytics. Enterprise solutions enable IT teams to manage IMDG implementations across the organization, helping resolve points of cross-functional and cross-system process constriction and improve infrastructure utilization to control costs.

Deliver today, enable tomorrow

IT leaders aren't just concerned with the technology needs of the company today, they are also continuously surveying the environment to understand the business needs of the future and how technology evolution is likely to impact them. All signals indicate that **over the next few years there will be a rapid expansion in the amount of real-time data that businesses need to manage effectively**.

Handle More Transactions When Processing Takes Microseconds Instead of Minutes

The immediate focus of many IT departments is enabling real-time customer interactions – because this is a pain point today. Once the interaction issue is resolved, the focus will shift to transactional efficiency. Think of it like water flowing through a series of pipes. A constriction slows everything down. Once that constriction is removed, the water will flow until it hits the next point of constriction where it will again slow down. **Transactional efficiency is all about increasing the flow of business through the end-to-end system.** IMDGs can enable you to achieve speed and scale not only within a single application but across an entire workflow.

Enable AI Interactions Spanning Old and New Systems

Artificial Intelligence is changing the way users interact with technology and transforming the way customers interact with companies. Increasingly, AI systems are replacing human interactions for routine requests and activities that require intense (behind-the-scenes) data processing to provide an informed response to the user. AI interactions not only require real-time data processing to enable user interactions, but they also generate a tremendous amount of data that is useful for other purposes. **IMDGs provide AI systems a means of accessing the data they need to provide fast performance, even when the underlying source data systems are built on legacy technology.**

Process Data from IoT Devices in Real-Time

IoT devices, embedded sensors in machinery, and GPS-enabled smart devices are projected to be the largest sources of new data over the next few years. With low-cost data collection devices connected to the internet and streaming data to cloud services, companies can access rich data about nearly all facets of business operations. The challenge is how to manage this influx of streaming data in real-time so it can be used to generate relevant information and make business decisions. **The distributed architecture of IMDG enables streaming data to be collected and processed at scale.** This means that companies can continue to grow IoT portfolios without concern for data overload.

Security Threat Monitoring

Information security is a critical issue for digital businesses. The challenge that IT leaders face is how to enable effective threat monitoring and detection without slowing down transactional processes with overhead processing activities. In-memory computing enables you to do this in two ways. First, it **drastically speeds up transaction processing**, so the addition of monitoring algorithms does not adversely impact the customer experiences. Second, the parallel computing architecture of **IMDG enables you to have one set of functions performing threat monitoring while other functions are processing the business transaction**.

Don’t Think about It When the System Just Works

Everyone wants speed, there are really no exceptions. The question is how to achieve it. Infrastructure isn’t cheap, which means apps and services need to scale up and down to match demand and scale-out to support diverse business needs. **Too much capacity is wasteful and too little capacity causes issues during peak business periods.** Systems must be secure and reliable without sacrificing ease of use (or speed).

An in-memory computing solution delivers the speed, scale, and stability needed to enable your business to grow and operate with confidence. Once the mechanics of the solution are out of the way, the company can focus on the true impact opportunities. Confidence enables leaders and staff to focus on core business competencies – developing modern customer experiences and highly efficient business processes that will enable the company to thrive.

The Technical Solution to Help Companies Move Forward

Companies that want to succeed in enabling real-time customer interactions and harnessing the power of streaming data are going to have to address the constraints of IT infrastructure. **Hazelcast IMDG provides the speed, scale, and stability to enable the business performance required while leveraging the infrastructure investments already made.** Implementing in-memory computing doesn't require overhauling the IT environment.

Flexibility in implementation across different types of environments (on-premises, cloud, and hybrid) and the ability to insert IMDG as a layer between existing relational databases and applications means companies can augment existing infrastructure with as little disruption as possible.

IMDG Use Cases

E-Commerce

One of the world's largest e-commerce sites implemented an IMDG solution to achieve a 25,000% decrease in time-to-action. When customers engage with the site, they expect a superb technology-enabled experience. The IMDG solution provides no noticeable delay in complex customer interactions, even during volume spikes.

Customer Service

A leading media company is using in-memory computing to capture customer viewing history, account history, customer service interactions, and location data to drive an integrated 360-degree view of the customer. This data is used to power AI-driven customer service experiences when a customer calls for support. Callers don't realize they are speaking with a computer, which is the true indicator of AI success.

Fraud Protection

One of the world's leading consumer banks is using an IMDG based solution to continuously execute multiple fraud detection algorithms in .003 seconds, millions of times per second. This enables them to identify suspicious credit card transactions in real-time to avoid unauthorized use without slowing down payment processing for customers and merchants. Customers don't realize the fraud protection they are getting automatically – the solution just works.

hazelcast®

2 West 5th Ave., San Mateo CA 94402 USA
Email: sales@hazelcast.com Phone: +1 (650) 521-5453
Visit us at www.hazelcast.com

@hazelcast
https://www.linkedin.com/company/hazelcast
https://www.facebook.com/hazelcast