AI Transforms Oracle Exadata X8 by, David Floyer

June 17th, 2019

Wikibon believes that integrating AI into Oracle Database 19c and the Exadata platform changes the game. As a result, Wikibon recommends that Exadata X8 becomes the default x86 platform for mission-critical workloads using Oracle Database.

Senior executives should regard Exadata X8 and beyond as the default platform for Oracle. It is time for database practitioners and their executive management to let go of traditional roles of database design and optimization of hardware and software. Acceptance of Oracle Cloud services for the appropriate mission-critical workloads will release headcount and reduce business risk.

Practitioners will need new skills to design hybrid databases and support hybrid applications. These applications will exploit hybrid data sources across hybrid clouds and be far more valuable to the enterprise line-of-business. They will be creating the core of enterprise digitization services.

Premise

Wikibon believes that integrating AI into Oracle Database 19c and the Exadata platform changes the game. As a result, Wikibon recommends that Exadata X8 becomes the default x86 platform for mission-critical workloads using Oracle Database.

Executive Summary of Exadata X8 as a Converged Platform

Oracle Database offers premium functionality for deterministic[1] mission-critical workloads. In order to improve performance and ease-of-operation, Oracle has integrated the database with the Exadata platform for the past 10 years.

The latest announcement of Exadata X8 includes the latest x86-based hardware optimized for Oracle workloads. This lowers latency, particularly IO and memory latency, important for database response time and throughput.

More importantly, Oracle is also integrating Exadata X8 and Oracle Database 19c with Artificial Intelligence (AI) and Machine Learning (ML). As a result, this deepens the integrations and enables quicker development of real-time operational and functional improvements.

More operational system data available to AI & ML translates to quicker and more effective development of AI & ML solutions. AI needs sustainable access to as much operational system data as possible. Oracle is achieving this by making Exadata the platform for Oracle Cloud, Cloud at Customer and Cloud Applications. The sum of the pieces is far higher than the piece parts.

Open source practitioners have developed a flood of alternative databases recently. However, Wikibon concludes these databases support different workload types. These workload types are growing faster, but are not replacing the traditional systems of record workloads and databases.

There are alternative public cloud databases that can address the systems of record workloads. However, Wikibon concludes that they do not have the same functionality as Oracle Database. Wikibon also concludes that the IT cost of conversion, the disruption costs to the lines of businesses, and the risk of serious business disruption overwhelm any benefits. As a result of this analysis, Wikibon continues to strongly advise against any strategy that involves database conversion.

In summary, Wikibon recommends deploying Oracle Database applications on the Oracle Exadata platform, starting with Exadata X8 with Oracle Database 19c.

Note[1]: See Note[1] in the Footnotes section below.

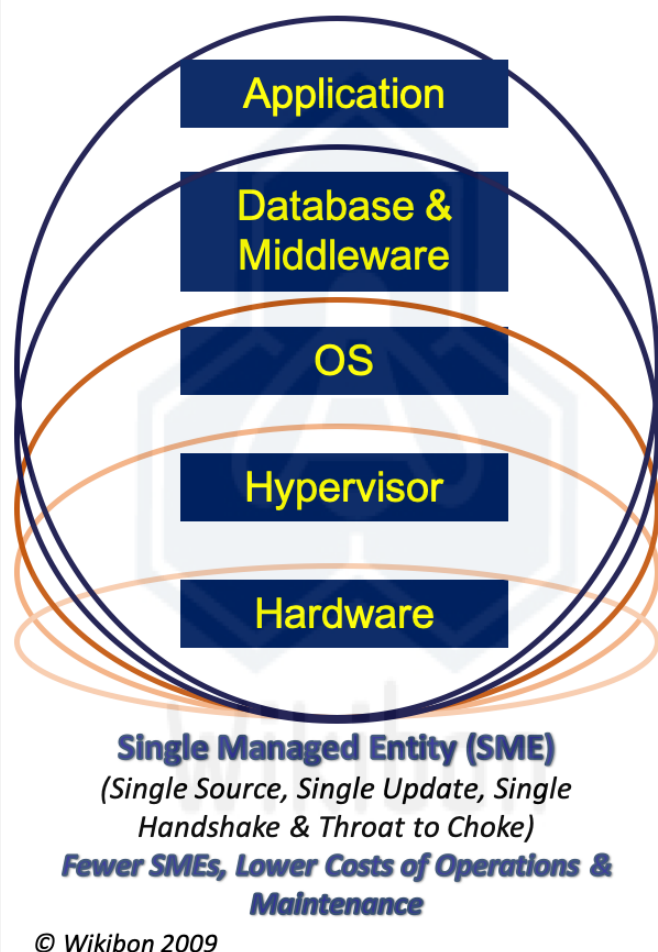## Converged Infrastructure Fundamentals

Figure 1 – Converged Full Stack Infrastructure Topology
Source © Wikibon 2009.

Wikibon has been discussing the benefits of converged infrastructure for the last decade. Figure 1 shows an established diagram illustrating the principle of combining a "Full Stack" of infrastructure and software components together, with a single vendor source responsible for all the technology.

This source is also responsible for creating all service level agreements, including updates and security patches. And the same single source is responsible for all aspects of maintenance.

This approach results in lower operational costs and reduced infrastructure. The vendor or integrator takes greater responsibility for delivering improvements and security as a continuous service.

However, the most important benefits found by practitioners is reducing the inherent friction from unique infrastructure, processes and procedures created by enterprise IT. Converged infrastructure enables faster deployment of new applications and application improvements. It is inherently more secure and easier to adapt to changing circumstances.

The success of public clouds comes from building a set of IT services on converged and hyperconverged infrastructure. They are now the foundation of what Wikibon calls "True Private Clouds," and "True Hybrid Clouds".

In order to take full advantage of converged and hyper-converged platforms such as Exadata X8, IT usually needs to change the organization. Making these changes is often difficult, and results in friction to achieving all the benefits of converged and hyperconverged infrastructure.

Wikibon believes that Oracle will continue to push AI inference code to the on-premises Exadata environments, because the compute has to be next the data. The Autonomous Database Platform will be a combination of Oracle Database and Exadata. Wikibon believes that Oracle will need to move to a continuous improvement and deployment model and move to a cloud pay-per-drink financial model over time.

Exadata X8 Highlights

Automated Datacenter Operations

The Oracle Cloud with Oracle Database 19c is the start of a journey to the Autonomous Database. This leads to tighter integration, enhanced orchestration, and automated data center operations. The initial automated data center operations include:

- Provisioning, patching, and online upgrading
- Automated Indexing
- Backup monitoring, application scaling, performance diagnostics, automated tuning and query optimization
- Bug identification and prioritization
- Testing and change management of complex applications and workloads;
- Maintenance Slot identification
- Automatic hang management, anomaly detection and handling of failures and errors, particularly alerts

The ability to extend and improve this automation will increase over time as Oracle Database Cloud services continue to mature.

Exadata X8 Platform Innovations

Exadata X8 hardware enhancements include the latest Intel Xeon processors with two or eight 24-core & 48-thread Intel Xeon 8260 processors in database servers and two 16-core/32-thread Intel Xeon 8268 processors in the storage servers. The use of PCIe NVME flash technology and the availability of 60-percent more cores in the storage servers are the reason for the Oracle claim of a 60-percent increase in I/O throughput for all-Flash storage compared to Exadata X7.

The Exadata X8 storage servers offload Oracle Database processing. They use the latest Intel chip to provide 60 percent more cores. The higher capacity HDD disk drives are now 40% bigger.

Wikibon expects Oracle will support non-volatile memory (NVDIMMs), which will allow much larger in-memory databases and applications.

Additionally, a new, much lower cost XT extended storage server is available for storing infrequently accessed, older, and regulatory data. This data layer is application transparent, is consistent with the operational models, supports hybrid columnar compression (HCC), and deploys the same encryption security model.

At the core of Exadata is the low-latency low overhead network which integrates the Exadata system. This is still based on InfiniBand using an RDMA protocol.

Exadata X8 Software Innovations

Wikibon believes that some of the most important Exadata improvements will come from AI and ML. The foundation is a common platform and common components across all the Exadata instances. This allows data to be gathered across many different workloads and environments. These streams of data from public and private clouds can allow Oracle to develop sophisticated ways of improving data services and improve automation of operational procedures.

Of course, there needs to be an open and clear understanding of what data is collected and when. However, for the most part the objectives of Oracle and its customers are aligned, which is better performance, greater reliability, and greater compliance for mission-critical application databases. The number of Exadata volumes deployed is very important in establishing a critical mass of data streams from which machine learning can be useful. Improved inference code can then be developed and deployed to Exadata customers.

Some of the machine learning and inference code will generate general improvements across most or all environments. An example is a specialized cache fusion and RDMA algorithm for communicating transaction information. This is developed centrally, and the inference code enabled locally.

Other learning will be more specific to interactions between the database software and the infrastructure

for a particular workload. An example might be tuning a specific important database with index creation. This require deep real-time interaction between Oracle Database 19c and the Exadata X8 components and iterate using reinforcement learning to an optimum state.

Trust betrayed in the development of AI has severe repercussions, as we learned in the tragic problems between airlines, pilots, the FAA, other aviation authorities, and Boeing. Trust rewarded can lead to a virtuous cycle where practitioners can contribute and lead development efforts that help themselves and other practitioners and enterprises.

See "Exadata AI & Machine Learning" section below for additional details and more discussion. In particular, trust must be established and maintained between Oracle and its Exadata customers and practitioners.

Oracle Zero Data Loss Recovery Appliance X8

Oracle also offers an upgraded Zero Data Loss Recovery Appliance X8 to monitor, protect and recover Oracle Databases in an enterprise. This solution utilizes the Oracle Exadata X8-2 hardware to increase data protection performance and capacity. It can protect Oracle Database information of a very large number of databases that can be running on Oracle Exadata or virtually any traditional server platforms. Oracle's Recovery Appliance enables companies to protect databases from across their enterprise against system failures, natural disasters, malicious activity and ransomware and helps companies retain information for compliance purposes.

The Recovery Appliance was designed and integrated into Oracle Database by the Oracle Database Engineering Team and uses database technologies such as Data Guard transport mechanisms under the covers. This provides faster and more integrated data recovery of Oracle Databases than traditional backup appliances. Oracle practitioners have expressed confidence that they can achieve sub-second recoverability of transactions.

Exadata X8 Integration with Microsoft Azure Cloud

Oracle have recently extended their cloud offering by joining forces with Microsoft to offer a true hybrid cloud integration between Microsoft Azure and the Oracle Cloud. Some of the capabilities include:

- Connecting Azure and Oracle Cloud seamlessly, allowing customers to extend their on-premises data centers to both clouds using an Equinix low-latency direct interconnect. The Oracle Cloud is hosted on Exadata.
- A single sign-on experience and automated user provisioning, to manage resources across Azure and Oracle Cloud.
- Oracle applications can use Azure Active Directory as the identity provider and for conditional access.
- Support for custom applications and packaged Oracle applications (JD Edwards EnterpriseOne, E-Business Suite, PeopleSoft, Oracle Retail, Hyperion) on Azure connected to Oracle databases deployed in the Oracle Cloud.
- The same Oracle applications are also certified to run on Azure with Oracle databases in an Oracle Cloud.
- A collaborative support model between Oracle & Microsoft is being developed.
- Azure will continue as certified to run Oracle Database on various operating systems, including Windows Server and Oracle Linux.

Wikibon Research

Previous Wikibon Exadata Cost Research

Wikibon looked in detail at the comparison between Oracle Exadata and a "Roll-your-own" approach in research entitled "Oracle Exadata vs. Roll Your Own: Wikibon's Take." The data in Figure 2 below is taken from this study. It is based on five years of researching converged systems with input from over 300 customers, including 40 Exadata customers. The workload analyzed is enterprise mission-critical systems of record applications, using Oracle Database.
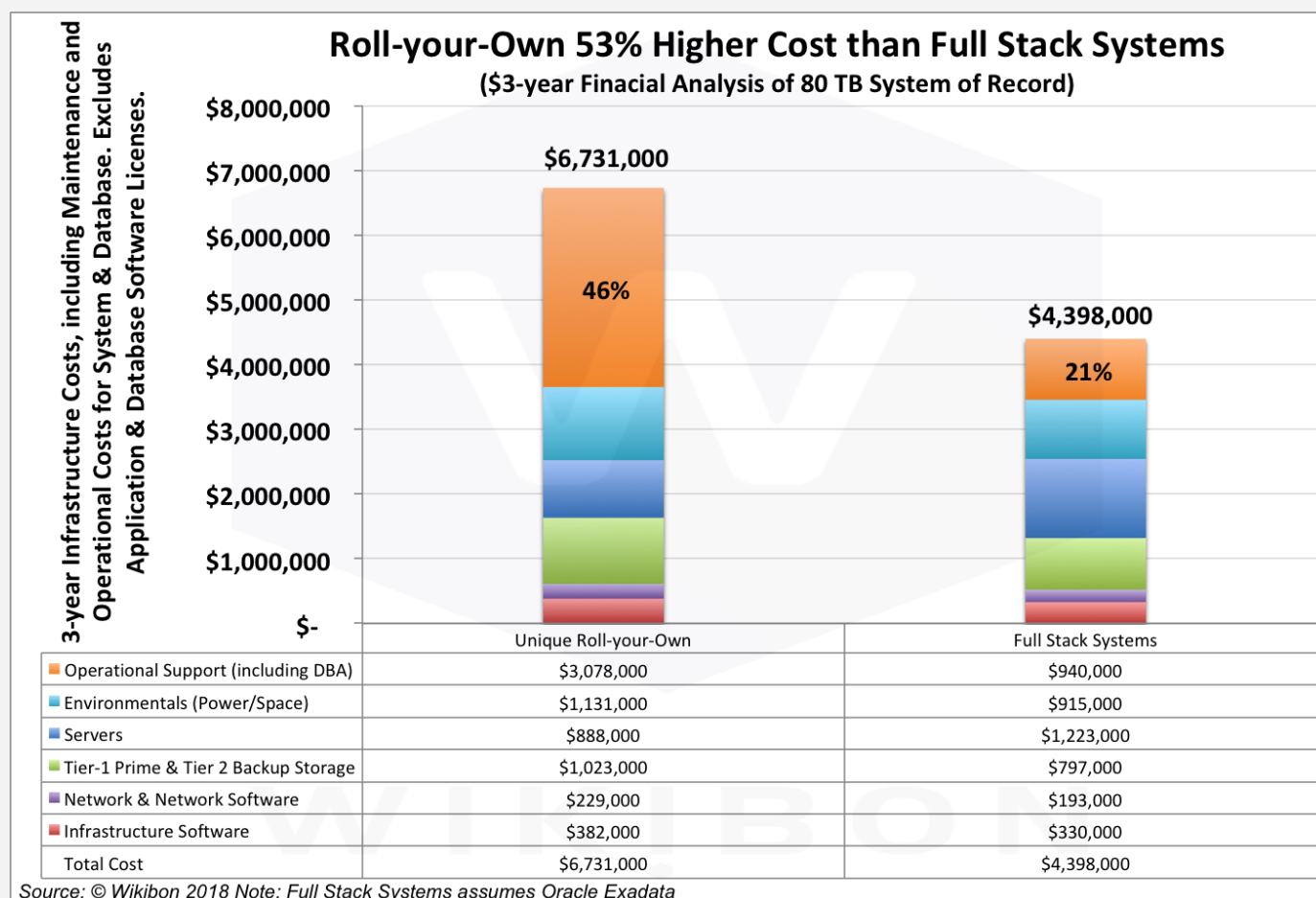
Roll-your-Own 53% Higher Cost than Full Stack Systems
($3-year Finacial Analysis of 80 TB System of Record)

	Unique Roll-your-Own	Full Stack Systems
■ Operational Support (including DBA)	$3,078,000	$940,000
■ Environmentals (Power/Space)	$1,131,000	$915,000
■ Servers	$888,000	$1,223,000
■ Tier-1 Prime & Tier 2 Backup Storage	$1,023,000	$797,000
■ Network & Network Software	$229,000	$193,000
■ Infrastructure Software	$382,000	$330,000
Total Cost	$6,731,000	$4,398,000

Source: © Wikibon 2018 Note: Full Stack Systems assumes Oracle Exadata

Figure 2 – Analysis of Operational Costs for Roll-you-Own Infrastructure vs. Full-stack System running Systems of Record Workloads.

Source © Wikibon 2018. Infrastructure costs exclude Database and Application License Costs. Operational Costs include infrastructure and database support.

Figure 2 shows a traditional unique Roll-your-Own environment is more than 50% higher costs compared to a Full Stack Exadata-based system. Figure 2 also shows clearly that RYO operational costs are 46% of the total, and Exadata system costs are only 21%.

Wikibon believes that the results of this study would apply to Exadata X8, and the differences would likely be higher. One reason is improved patching, which is addressed in the next section.

Patching

In Wikibon research entitled "Why IT Disasters Happen", Wikibon examined the Equifax disaster. Equifax

was the victim of a cyber-attack in 2017, affecting the highly personal data of 145 million people in the USA. Equifax IT failed to apply a patch to an open-source web server, which allowed the breach. As a result, Moody's is planning to include the risk of a business-critical cyber-attack as part of its credit rating, and recently cited the cyber-attack as a factor in downgrading Equifax. Moody's estimates the cost of the Equifax cyber-attack as over **$1 Billion over 3 years,** as reported by Forbes. If Equifax had been operating in Europe, the GDPR legislation would result in fines of multiple billions.

As a result of this, ***patching is a board-level issue***. Wikibon reviewed a "best-of-breed" enterprise environment with multiple vendors applying over 1,000 patches every year. The introduction of Exadata reduced the number of patches to about 16 applied by Oracle. Of course, practitioners will always need to patch in an emergency, but much less frequently.

Proactive efforts to protect data and maintain compliance with data privacy and governance regulations is much easier to present to the board than explaining why the company's reputation has been tarnished and bottom line effected by a cyber-attack.

Exadata X8 AI & Machine Learning

Exadata X8 now runs the same code as Autonomous Database for the Automatic Indexing feature, and delivers the same machine learning capabilities for on-premises machines.

- Autonomous Transaction Processing (ATP) includes Automatic Indexing
  - Using machine learning, Automatic Indexing continuously learns and tunes the database as usage patterns change. It will identify index candidates based on the specific Database workloads, verify and validate their effectiveness, and create the indexes. The entire process is automatic and improves database performance and eliminates manual index-tuning tasks for critical applications.
  - Experts had tuned a complex Netsuite workload (17.5K SQL statement, 1.8K tables & 8K indexes) over 15 years.
  - ATP achieved the same level of performance in 30 minutes from scratch, using only 1.7K indexes
  - This was achieved in real-time using Reinforcement Learning, which allows the algorithm to learn from its own actions (same principle as Leela Chess!)
  - Needs the Exadata X8 to work directly with Oracle DB 19c in real-time.
  - Practitioners can stop creating indexes for applications running on Oracle Database and focus on more strategic business objectives.
- DB operations implements automated detection of hang management, anomaly detection, maintenance slot identification and bug identification and prioritization.
- Autonomous Data Warehouse (ADW) workload optimizations includes an advanced query optimizer and real-time statistics.
- Oracle Database 19C allows real-time detection of runaway SQL, and automatic quarantine of the code.

Wikibon expects solid development of AI/ML solutions to real operational issues. Top of our list would be real-time analysis and synthesis of machine alerts in context.

Alternative Databases for Mission-critical Deterministic[1] Workloads

Senior IT executives should know if there is an alternative database platform to Oracle.

Open source Databases

Open source has led the development of many successful database innovations, such as Aerospike, Cassandra, CockroachDB, CouchDB, MariaDB, MongoDB, Neo4j, PostgreSQL, Redis, RethinkDB, and Timescale. However, the architecture is usually different from the traditional SQL Databases. For example, the fundamental database model[2] many of these new databases trade Partition Tolerance for

either Availability or Consistency. Large-scale **_probabilistic[1]_** workloads are the design point.

Public Cloud Databases

AWS with Aurora and Google with Spanner are attempting to develop mission-critical database alternatives in a public cloud for traditional deterministic[1] workloads.

- AWS can achieve some of the Oracle functionality with a combination of Aurora (based on the open-source MySQL) and AWS micro-services. However, this is a kluge that enterprise practitioners must develop and maintain, and therefore unsuitable as a replacement for mission-critical Oracle Database. In addition, AWS does not offer Aurora natively on its on-premises Hybrid Cloud offering called Outposts.
- Google Spanner uses a Consistent/Partition Tolerant (CP) model[2]. Google Spanner is technically superior to AWS Aurora and includes advanced features such as TrueTime. However, the fundamental model is very different from the traditional Availability/Consistency model and this makes conversion to Google Spanner very difficult. Google Spanner is also not available on-premises.

High-end mission-critical database Market

Wikibon believes that high-end mission-critical databases for enterprise deterministic database workloads form a clear, differentiated, and sustainable market. Oracle Database dominates the high-end mission-critical database market, with more function and deployment than any other platform. IBM DB2 and Microsoft SQL Server are next in line. As these databases all use a Consistency/Availability (CA) model[2], they are therefore suited for deterministic mission-critical workloads.

Oracle is the leader in this slowly shrinking market. Wikibon believes that the Exadata X8 platform and future generations, together with the Oracle Database Cloud strategy, can become increasingly cost effective in this market and increase Oracle market share and revenue.

Sustainable access to large volumes of data will be very important for the development of real-time AI/ML and sustaining market share. Oracle can capture a high percentage of the available data.

Conclusions & Recommendations

Oracle has moved aggressively to support mission-critical enterprise database-as-a-service across hybrid cloud environments. Real-time AI & ML are key to improving the converged Exadata X8 platform. Wikibon believes Oracle can establish trust with its enterprise customers and establish a sustainable and unique volume of data to drive its AI strategy. This will increase the amount of database code that will only run on the Exadata platform.

An alternative strategy for enterprises is to convert existing Oracle mission-critical applications to new cloud databases. Wikibon believes this "Lift and Shift" strategy is almost always wrong for this type of workload. Enterprises have not adopted this strategy. Conversion costs are very high. Application developers must freeze existing applications, and the business cost of freezing is even higher. The potential functionality limitations of those cloud databases once they are moved to a generic cloud environment usually impacts enterprise productivity.

Wikibon has worked with many clients evaluating conversion projects. These clients have consistently underestimated the business risk of conversion projects. Wikibon is happy to share its methodologies with enterprise executives.

This does not mean that enterprises should not convert any Oracle databases to a cloud database. Developers often deploy the wrong technology, or requirements change. Developers should deploy other specialized cloud databases. Wikibon strongly believes in "horses for courses".

Wikibon believes mission-critical systems of record will continue to be a vital business IT component which requires premium database functionality. Hybrid applications will need integration between existing systems of record and new applications using advanced analytics and AI/ML technologies.

Wikibon believes that the advanced enterprise database market can support an independent cloud ecosystem, as the market is large. Oracle and Microsoft have made complementary integrated cloud offerings that will enable flexibility of application and data placement in hybrid cloud configurations.

Wikibon also believes that Oracle should move to a cloud consumption model and move its software licensing model to a cloud-first continuous improvement model.

Wikibon's primary recommendation is that Oracle Exadata X8 & Autonomous Database should be the default platform for large-scale mission-critical traditional database workloads.

Action Item

Senior executives should regard Exadata X8 and beyond as the default platform for Oracle. It is time for database practitioners and their executive management to let go of traditional roles of database design and optimization of hardware and software. Acceptance of Oracle Cloud services for the appropriate mission-critical workloads will release headcount and reduce business risk.

Practitioners will need new skills to design hybrid databases and support hybrid applications. These applications will exploit hybrid data sources across hybrid clouds and be far more valuable to the enterprise line-of-business. They will be creating the core of enterprise digitization services.

Footnotes

The detailed definitions below come from Wikibon research entitled "Google Spanner Pressures Oracle".

Note [1]: Deterministic vs. Probabilistic Workloads

One useful classification of applications and workloads is as either deterministic or probabilistic. With most traditional computing workloads, if you have the same input and same code, you expect to generate exactly the same output. We can classify this as a deterministic workload. Examples of this are systems of record, such as finance, ERP, payroll, stock control, etc. Most compliance models assume deterministic outcomes.

There is also a broad set of problems that can be solved more efficiently with probabilistic methods, especially when the applications are driven by large amounts of data and compute requirements. In order to meet the constraints of elapsed time to solution, systems can dynamically choose algorithms designed for "good enough" output, which can be computed quicker. Examples of probabilistic systems could include real-time price updates or deciding the price to bid for delivering an advertisement in real-time to an end-user. Speed to solution is more important than absolute accuracy and repeatability of the assessment.

Hybrid applications will often be a combination of deterministic and probabilistic methods, and hybrid clouds will be an important architectural foundation.

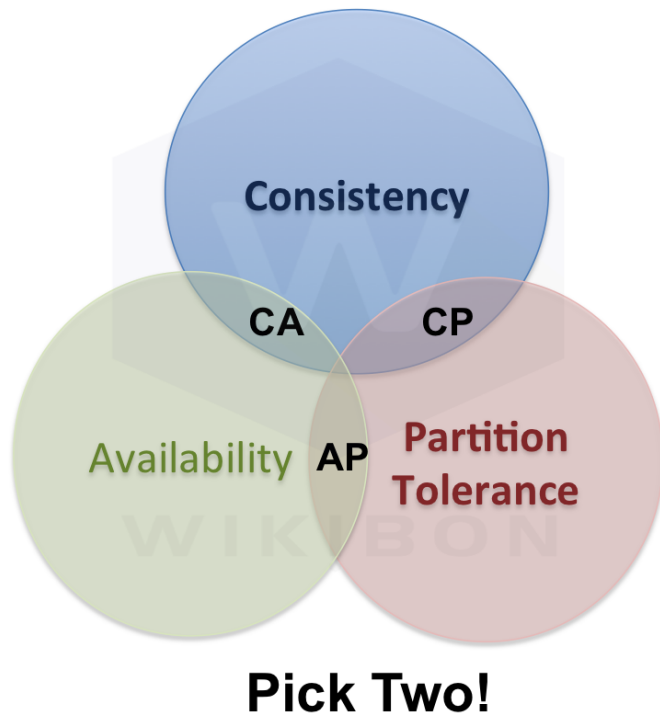### Note[2]: CAP Theorem & Database Models

Figure 3: CAP Theory
Source: (c) Wikibon 2018, based on the Theoretical work of Eric Brewer

The CAP theorem (Figure 1) dictates that a system cannot have guaranteed capabilities of being Consistent, Available, and Partition Tolerant (scalable). The CAP theorem states that a database can only guarantee only two of the three capabilities, but not all three. The Oracle, IBM DB2 and Microsoft SQL Server databases guarantee availability and consistency, a CA model. Google Spanner is a CP model.

It should be noted that it is technically difficult to provide databases that guarantee any of the database dimensions. Many databases fail with all three.

Previous wikibon research includes "Google Spanner Pressures Oracle" and "Microsoft Acquires Avere's Hybrid Cloud Data Consistency"

David Floyer spent more than 20 years at IBM, holding positions in research, sales, marketing, systems analysis and running IT operations for IBM France. He worked directly with IBM's largest European customers, including BMW, Credit Suisse, Deutsche Bank and Lloyd's Bank. Floyer was a Research Vice President at International Data Corporation (IDC) and is a recognized expert in IT strategy, economic value justification, systems architecture, performance, clustering and systems software.

David Floyer

@dfloyer
david.floyer@wikibon.org