



Scaling Database Performance in Azure

Results of Microsoft-funded Testing

Q1 2015

Test Goals and Background Info

Test Goals and Setup

- Test goals
 - Microsoft commissioned testing to compare SQL Server performance on prem vs. paired with ScaleArc database load balancing software in Azure
 - Testing run with and without ScaleArc caching enabled
 - Simulate realistic, heavy loads that stress compute, disk I/O, and network evenly

On-premise setup

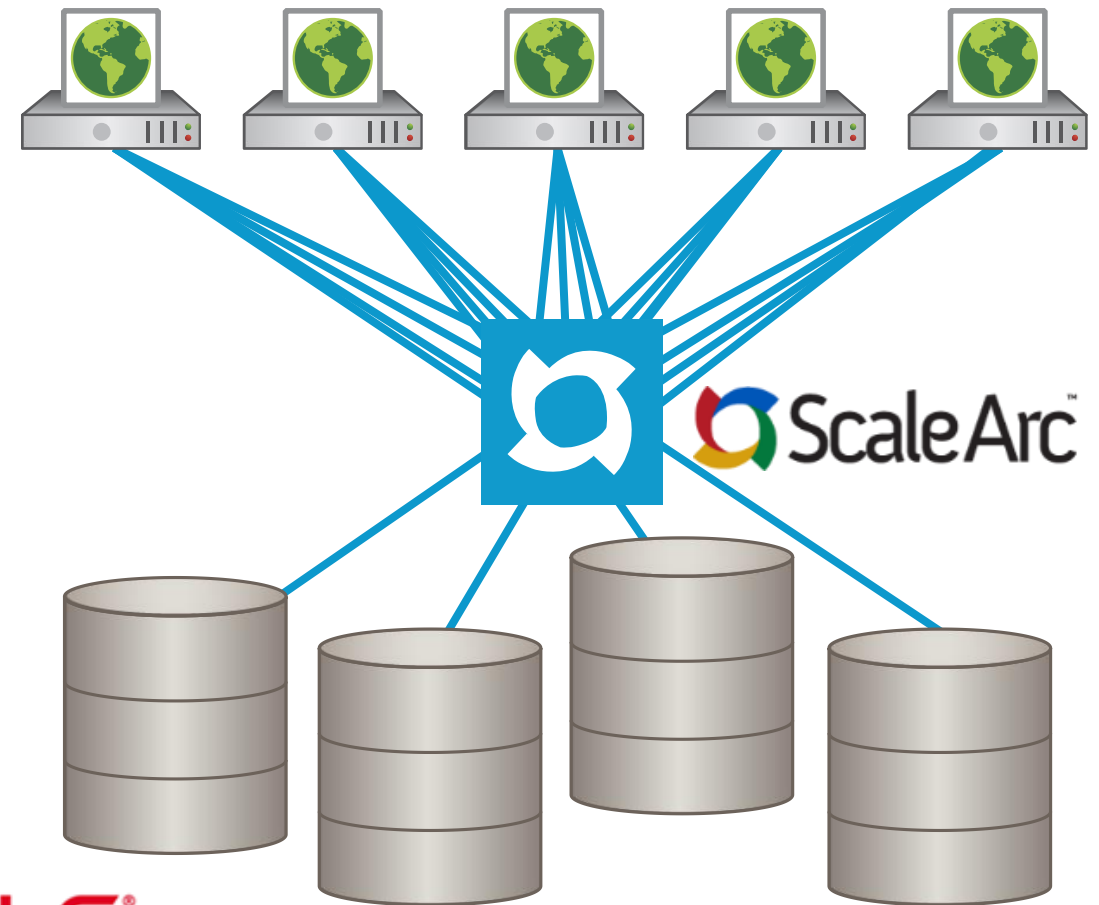
- Bare metal, 24 hyper-threaded Ivy Bridge 2.6 GHz Xeon processor cores, 64 GB RAM, >300k IOPS PCIe SSD
- 10 Gbps LAN with ping pong latency of <11 μ s @ 64K frames

Azure setup

- SQL Server on Azure (A3) instances with 4 virtual cores and 7 GB RAM
- SQL Server 2014 with both synchronous and asynchronous replication
- ScaleArc running on A3 and A6 nodes

What is ScaleArc?

- Database load balancing software
 - Inserts transparently between apps, databases
 - Enables an Agile Data Tier – abstraction layer
- Available on Azure Marketplace
- Enables zero downtime environment and provides instant database scalability
- Supports SQL Server, MySQL, Oracle
 - SQL Server 2005, 2008/R2, 2012, 2014
 - AlwaysOn Integration



Microsoft®
SQL Server®



ORACLE®

Without ScaleArc: Limitations of SQL Server on Azure

- Virtual machine size
 - Large VMs with a lot of compute resources are unavailable / expensive
- I/O limitations
 - Storage is shared / SAN-based – IOPS and storage latency are a challenge vs. SSD storage
- Hypervisor-related losses
 - Hypervisor layers add overhead and reduce effective resources – also, bad neighbor effect
- Network latency
 - Cloud machines have no guarantee of proximity
- Added resource utilization due to replication
 - AlwaysOn replication needed for capacity, which may cause further performance degradation

Benefits of Deploying ScaleArc in Azure with AlwaysOn

- Enables simple app migration with no performance compromise
 - Move enterprise workloads to Azure with no app changes and no performance loss
 - Aggregate smaller cloud database instances to behave like one large server
- Provides automatic high availability
 - Auto failover with no app restarts – within one locale or across regions
 - Eliminates the need for maintenance windows to patch SQL databases
- Solves future scaling issues
 - Scale up and scale out without any app modifications
- Delivers high-performance hybrid deployments
 - App in Azure, with database on prem

How ScaleArc Augments SQL Server AlwaysOn

- Aggregate capacity across VMs to get on premise-like performance
 - Combine as many as nine database VMs
- Linear scalability
 - Add machines as needed – ramp up and down for traffic surges, with no downtime
- Higher availability in Azure vs. on prem, and at lower cost
 - Auto failover + queueing supports cross-region HA, at much lower cost than two data centers
- Zero-downtime maintenance
 - Perform maintenance per server, cluster or Azure zone with no application downtime
- Deep SQL analytics
 - SQL query patterns and historical analytics enabling troubleshooting in minutes vs. hours

Test Methodology and Results

Test Methodology

- Generate SQL load similar to SaaS, content management, and eCommerce applications
 - Stress compute, memory, I/O, network evenly
 - Simplistic queries so as not to overstress CPU
 - 10% writes – twice typical content loads
- Load profile
 - 10:1 read:write ratio
 - Reads = 70% point queries, 30% top/range queries
 - Writes = 50% updates, 25% deletes, 25% inserts

```
CREATE TABLE [dbo].[oltp1]
(
    [id] [int] NOT NULL,
    [k] [int] NOT NULL,
    [c] [varchar](120) COLLATE SQL_Latin1_General_CP1_CI_AS NOT
NULL,
    [pad] [varchar](60) COLLATE SQL_Latin1_General_CP1_CI_AS NOT
NULL,
    PRIMARY KEY NONCLUSTERED
(
    [id] ASC
)
) ON [ PRIMARY ]
```

Sample Read Queries

```
SELECT c FROM oltp11 WHERE id=401

SELECT c FROM oltp13 WHERE id BETWEEN 401 AND 410

SELECT DISTINCT c FROM oltp33 WHERE id BETWEEN 401 AND 410 ORDER BY c
```

Sample Write Queries

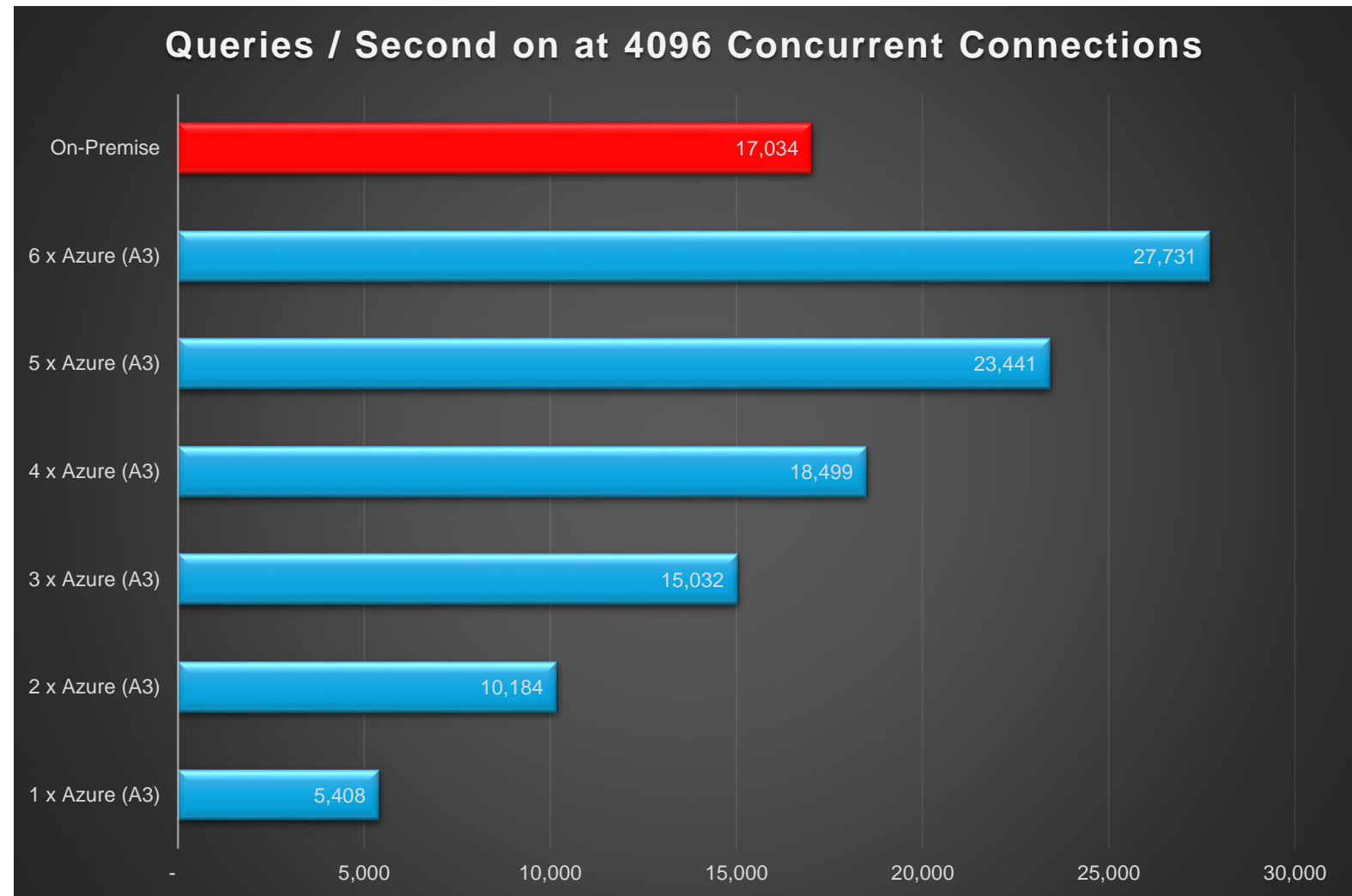
```
UPDATE oltpmem32 SET k=k+1 WHERE id=5642

UPDATE oltpmem64 SET
c='xv6pzW4COoUFVz9k6r63muulHbw4SZwADbeC6U2uvCI5nPqM1BrQ8NkjdeAca2QX
vilfsVUdOlxtl6mitLvEjgxeEXMaZKw0LHtQyT2EIOe60sGz5JevaJal' WHERE id=3454

DELETE FROM oltpmem25 WHERE id=328
```

Benchmarks: On Premise vs. Azure Cloud Performance

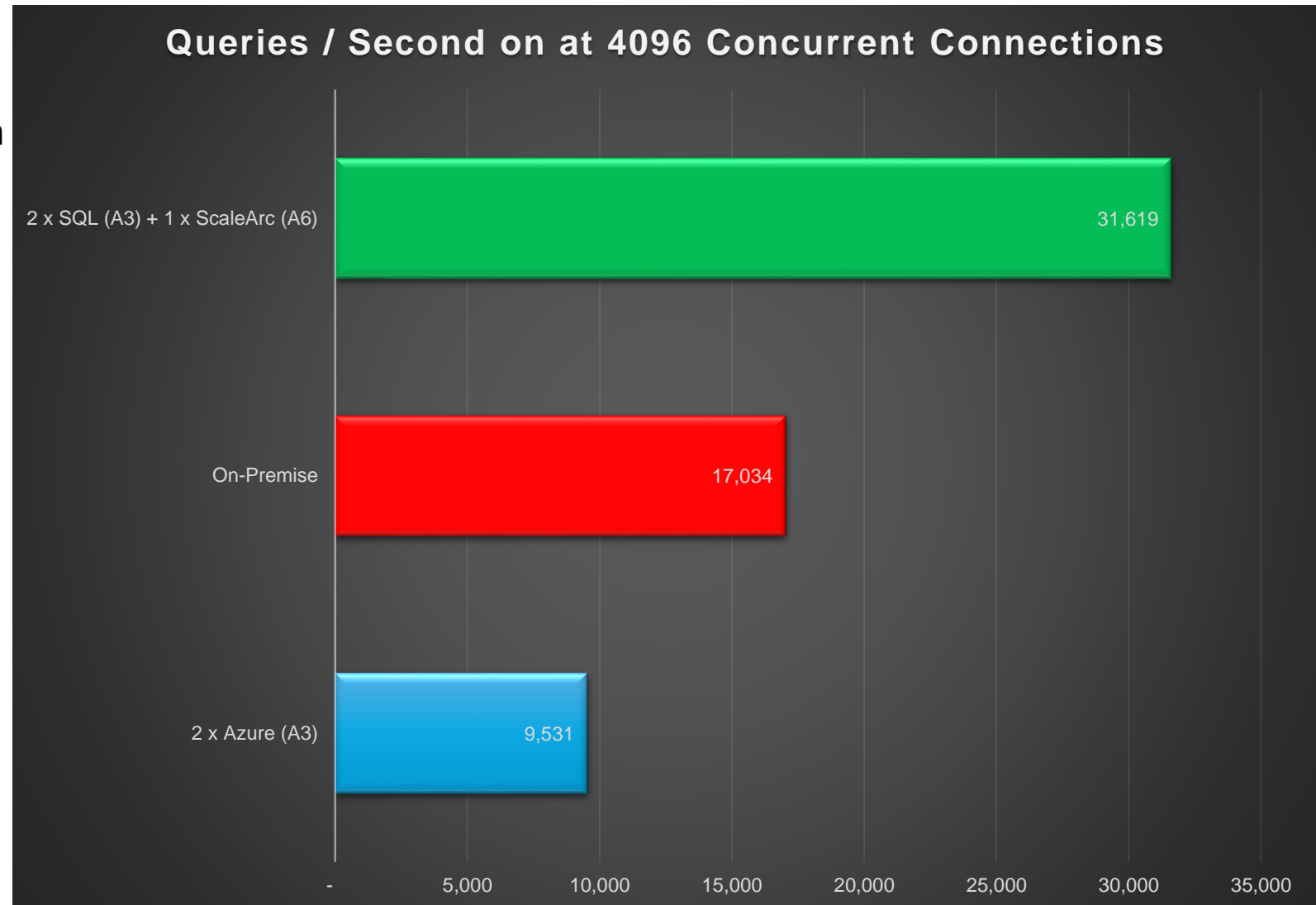
- On premise (12 physical cores):
 - 17K QPS, 1416 queries / physical core
- 3 x SQL on A3 instances + 2 x ScaleArc (A3) (20 virtual cores)
 - 15K QPS, 750 queries / virtual core
- 4 x SQL on A3 instances + 2 x ScaleArc (A3) (24 virtual cores)
 - 18.5K QPS, 770 queries / virtual core
- 6 x SQL on A3 instances + 2 x ScaleArc (A3) (32 virtual cores)
 - 28K QPS, 866 queries / virtual core



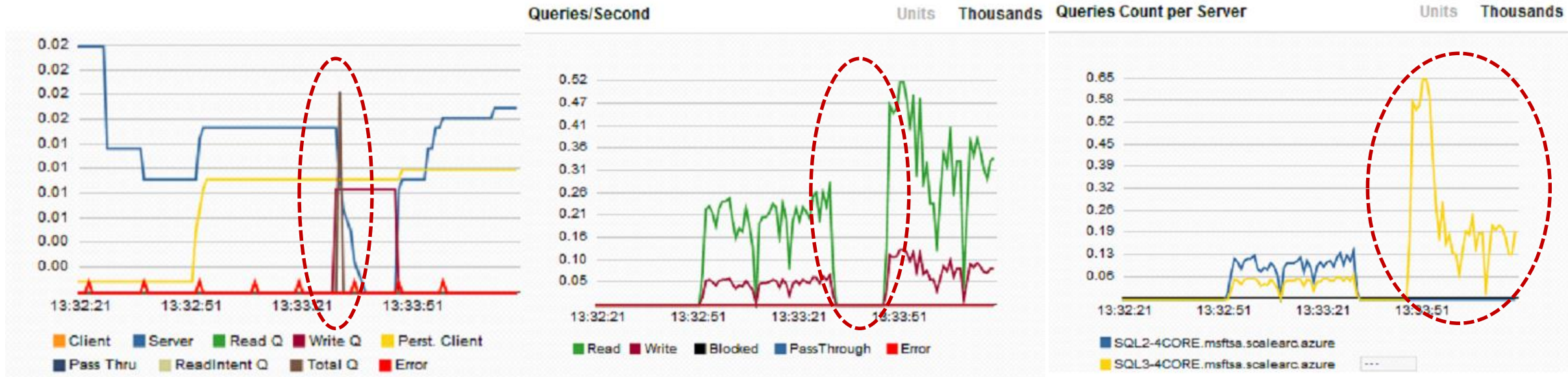
Note: ScaleArc + Azure enables linear scalability gains from adding readable secondaries with no app mods or downtime

ScaleArc Cache Performance

- 2 x SQL on A3 + 1 x ScaleArc (A6) with read cache enabled (16 virtual cores)
 - 31.6K QPS, 1976 queries / virtual core
- On-premise server with 24 hyper-threaded cores (12 physical cores)
 - 17K QPS, 1416 queries / physical core
- 2 x SQL on A3 instances + 2 x ScaleArc (A3) (16 virtual cores)
 - 9.5K QPS, 595 queries / virtual core



Automated Failover Testing



- ScaleArc detects failure of current AlwaysOn Primary
- **ScaleArc queues inbound client traffic**
- Queue shields app from errors while failover completes

- Traffic pauses briefly as ScaleArc queues traffic
- **No errors are seen by app** while failover completes in SQL AlwaysOn (other than the failed queries to the server that just died)

- Before failover, ScaleArc performed automatic read/write split, sending secondary reads and primary reads and writes
- **No config or app changes needed, no read-intent strings**
- After failover, newly promoted server serves all traffic

Additional Resources on Augmenting SQL Server

White Paper

ScaleArc

Achieving Zero Downtime for Apps in Microsoft SQL Server Environments

White Paper

ScaleArc

The Top 5 Challenges in Upgrading to SQL Server 2012/2014 – And How to Overcome Them

Solution Brief

ScaleArc

ScaleArc for SQL Server
For Deployment with SQL Server 2012/2014

Summary

SQL Server 2012/2014 offers a number of benefits, including AlwaysOn Availability Groups, but migrating to these versions of SQL Server presents a number of challenges. ScaleArc offers a seamless and transparent path for migration, along with an essential set of features that take fault tolerance, performance, scalability, and visibility to the next level.

ScaleArc can be transparently deployed into SQL Server environments, including SQL Server 2005, 2008/2008 R2, and 2012/2014. For those looking to migrate to SQL Server 2012/2014 from an older version, ScaleArc can facilitate a faster migration at lower operational cost. ScaleArc can be deployed in minutes and does not require any changes to existing applications or databases. The ScaleArc software can be deployed on bare metal, on VM, or in the cloud.

The patented ScaleArc technology provides the following benefits:

- enables automated higher availability with dynamic SQL query-level load balancing
- immediately improves performance – speeds response times up to 60X with the world's first transparent, query-level cache
- delivers real-time SQL analytics for instant troubleshooting and capacity planning
- enables scaling of your existing database infrastructure 10x or more, without making changes to databases or applications
- provides an elegant, simple SQL firewall as an added layer of protection

This solution brief outlines the benefits that ScaleArc provides when migrating to, and operating in, a Microsoft SQL Server 2012/2014 database environment.