SQL Server on Linux and Containers
A Brave New World

Speaker Name
Principal Architect
Microsoft
bobward@microsoft.com
@bobwardms
linkedin.com/in/bobwardms

Get this deck at http://aka.ms/bobwardms
Try out the demos yourself at https://github.com/Microsoft/sqllinuxlabs
The response to open source demand

Businesses are embracing choice
- Heterogeneous environments
- Multiple data types
- Different development languages
- On-premises, cloud, and hybrid environments

Microsoft is delivering on choice
- HDInsight on Linux
- Machine Learning Server on Linux
- Linux in Azure
- SQL Server drivers and connectivity
- SQL Server Ops Studio open source, cross-platform administration

The world is demanding SQL Server on Linux
- 6M+
  Docker pulls of SQL Server on Linux image
- 36%
  Enterprise DB market runs on Linux
SQL Server is the platform of choice

Support for RedHat Enterprise Linux (RHEL), Ubuntu, and SUSE Enterprise Linux (SLES)

Linux and Windows Docker containers

Windows Server / Windows 10

Package-based installation: Yum Install, Apt-Get, and Zypper

Database and tools compatibility on all platforms
SQL Server on Linux - What we have learned

 Customers are....

- Standardizing on an Operating System (Mixed OS and Hybrid Cloud short-term)
- Re-evaluating the preferred database for new applications
- Migrating from ORACLE, MySQL, or PostgreSQL

<table>
<thead>
<tr>
<th>Is it the same as Windows?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is the performance the same as on Windows?</td>
</tr>
<tr>
<td>Is the licensing the same as on Windows?</td>
</tr>
<tr>
<td>Do I have to know Linux?</td>
</tr>
<tr>
<td>What Linux are most customers using?</td>
</tr>
<tr>
<td>Is SQL Server different in a container?</td>
</tr>
</tbody>
</table>
## What’s in SQL Server on Linux

<table>
<thead>
<tr>
<th>Editions</th>
<th>Windows</th>
<th>Linux</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developer, Express, Web, Standard, Enterprise</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Services</td>
<td>Database Engine, Integration Services, SQL Server Agent</td>
<td>●</td>
</tr>
<tr>
<td>Analysis Services, Reporting Services, MDS, DQS</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Maximum number of cores</td>
<td>Unlimited</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Maximum memory utilized per instance</td>
<td>24 TB</td>
<td>12 TB</td>
</tr>
<tr>
<td>Maximum database size</td>
<td>524 PB</td>
<td>524 PB</td>
</tr>
<tr>
<td>Mission critical performance and HADR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic OLTP (Basic In-Memory OLTP, Basic operational analytics)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Advanced OLTP (Advanced In-Memory OLTP, Advanced operational analytics, adaptive query processing)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>SQL Server Replication</td>
<td>●</td>
<td>● NEW</td>
</tr>
<tr>
<td>Basic high availability (2-node single database failover, non-readable secondary)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Advanced HA (Always On - multi-node, multi-db failover, readable secondaries)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Security</td>
<td>Basic security (Basic auditing, Row-level security, Data masking, Always Encrypted, Active Directory Authentication)</td>
<td>●</td>
</tr>
<tr>
<td>Advanced security (Transparent Data Encryption)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Data warehousing</td>
<td>PolyBase</td>
<td>Planned</td>
</tr>
<tr>
<td>Basic data warehousing/data marts (Basic In-Memory ColumnStore, Partitioning, Compression)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Advanced data warehousing (Advanced In-Memory ColumnStore)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Advanced data integration (Fuzzy grouping and look ups)</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Tools</td>
<td>Windows ecosystem: Full-fidelity Management &amp; Dev Tool (SSMS &amp; SSDT), command line tools</td>
<td>●</td>
</tr>
<tr>
<td>Linux/OSX/Windows ecosystem: Dev tools (VS Code), DB Admin GUI tool, command line tools</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Developer</td>
<td>Programmability (T-SQL, CLR, Data Types, JSON, Graph)</td>
<td>●</td>
</tr>
<tr>
<td>Distributed Transactions</td>
<td>●</td>
<td>● NEW</td>
</tr>
<tr>
<td>Machine Learning Services</td>
<td>●</td>
<td>● NEW</td>
</tr>
</tbody>
</table>
Linux loves packages

• Major Linux distributions follow the pattern of using a package manager.
  • RHEL: RPM and yum.
  • Ubuntu: Debian and apt-get
  • SUSE: RPM and zypper.

• Package: binaries + instructions to install.

• Benefits: dependency management, automatic versioning, one-tool-to-rule-them-all, architecture control.
mssql-conf

- **mssql-conf** is a configuration script that installs with SQL Server 2017 on RHEL, SUSE, and Ubuntu. This utility can be used to set:
  - Enable SQL Server Agent
  - TCP port
  - Default data directory
  - Default log directory
  - Default dump directory
  - Default backup directory
  - Core dump type
  - Enable Availability Groups
  - Set startup traceflags
  - Set collation

SQL Server also supports config through environment variables
SQL Server Linux Architecture

- **LibOS (Win API and Kernel)**
- **Host Extension mapping to OS system calls (IO, Memory, CPU scheduling)**
- **API**
  - LibOS (Win API and Kernel)
  - SQLOS (SQLPAL)

**System Resource & Latency Sensitive Code Paths**

**Everything else**

**SQL Platform Abstraction Layer (SQLPAL)**

**Linux APIs (mmap, pthread_create, ...)**

**Based on Microsoft Research Drawbridge Project**

**Linux Process (Ring 3)**

**Ring 0**

**Linux Kernel**

**Parent “Watchdog” process**

**fork**
Demo
Deploy and explore SQL Server on Linux
Performance on SQL Server 2017 on Linux

1. #1 TPC-H Benchmark
   1,009,065 QphH with $0.47 per QphH

2. World’s first enterprise-class diskless database

3. Columnstore Indexes and In-Memory OLTP capabilities

4. 180 billion rows scanned in <20 seconds with 480 CPUs

5. Adaptive Query Processing and Automatic Tuning

6. Linux performance best practices guidance
Layers of security for enterprise-grade protection

Defense in depth:

- **Information Protection**
  - Encryption-at-rest
    - Transparent data encryption (TDE)
    - Backup encryption
    - Cell-level encryption
  - Encryption in transit
    - Transport Layer Security (TLS 1.2)
    - Client-side encryption
    - Always Encrypted

- **Access Management**
  - Database access
    - SQL Server Logins
    - Active Directory Authentication
    - Roles and permissions
    - Schemas
  - Application access
    - Row-Level Security
    - Dynamic Data Masking

- **Threat Protection**
  - Tracking & investigation
    - Fine-grained audit
    - Vulnerability Assessment
    - Data Discovery & Classification

Trusted: most secure over last 8 years
# HADR with SQL Server on Linux

**Failover Cluster Instance**
- Pacemaker and Corosync
- Single SQL Server instance
- SQL resource agent (mssql-server-ha)

- Open-source resource agent is available on GitHub Repository\(^1\)

**Shared Storage**
- ISCI, NFS, SMB

**High Availability Group**
- CLUSTER_TYPE = EXTERNAL
- Auto failover
- Pacemaker
- SQL resource agent (mssql-server-ha)

- 3 replicas required
- Configuration only replica for metadata
- Full AG capabilities

**Read-Scale Availability Group**
- CLUSTER_TYPE=NONE
- No clustering required
- Manual or forced failover
- Sync or async replicas
- Read scale routing
- Cross-platform

---

1. [https://github.com/Microsoft/mssql-server-ha](https://github.com/Microsoft/mssql-server-ha)

**Industry-leading solutions available from HA/DR partners**

---

---
Existing command line tools
• sqlcmd, bcp, sqlpackage

Visual Studio Code mssql extension

Azure Data Studio

mssql-cli

mssql-scripter

Windows SQL Server Management Studio (SSMS)
Windows SQL Server Data Tools (SSDT)

3rd party tools continue to work

Existing drivers/frameworks supported

SQL Server 2017 and 2019 on Linux

Tools and programmability
SQL Server on containers

**Portable**
Run anywhere Docker is supported

**Lightweight**
Reduced disk, CPU, and memory footprint

**Consistent**
Consistent image of SQL Server, scripts, and tools

**Efficient**
Faster deployment, reduced patching, and less downtime
Shared storage HA in Kubernetes

Built-in HADR orchestration with no clustering required

Persistent Volume Storage
Mission critical availability
Keep SQL Server running

- Always On availability group enhancements
- Resumable online index creation
- Online Clustered Columnstore index creation and rebuild
- Availability groups on Kubernetes
Demo

SQL Server and Docker Containers
SQL Server 2019 on Linux CTP 2.0 New Capabilities

- Replication
- Availability Groups with Kubernetes
- OpenLDAP support for 3rd party providers
- Persisted Memory Optimized I/O Memory Pressure Notifications
- Machine Learning Services with R/Python
- Distributed Transactions
- Microsoft Container Registry
- RHEL Container Image

- Database
- SQL
- Persistence
- Machine Learning
- Database Languages: T-SQL, Java, C/C++, C#/.NET, PHP, Node.js, Python, Ruby
Next steps

Evaluate SQL Server 2017 at http://aka.ms/sql2017

Leverage resources
http://aka.ms/sqllinux
http://aka.ms/sqlblog
https://github.com/Microsoft/bobsql
https://github.com/Microsoft/sqllinuxlabs
http://aka.ms/sqlchannel
http://aka.ms/sqldev

Join SQL Server Early Adoption Program
https://aka.ms/eapsignup