USING AZURE POLICY, BLUEPRINTS, AND ARM TO DEPLOY LARGE IAAS WORKLOADS

Guidelines for the Financial Services Industry

David Auslander – Principal Program Manager, Azure Global Financial Services Industry Team

Contents

Introduction

As more and more enterprises, especially Financial Services Institutions (FSIs), migrate workloads to the Microsoft Azure Cloud, a big focus has been on moving these workloads to Platform-as-a-Service (PaaS) solutions. Migration to PaaS lessens the need for management of various layers of the platform, as the Azure cloud will handle these. In some circumstances, enterprises need more fine-grained control of the environment and more flexibility when building in non-functional requirements. In these cases the large workloads can be and are often deployed into Infrastructure-as-a-Service (IaaS) environments. When deploying in workloads into IaaS the user can choose options for increased Performance, Security, Scaling, Availability, and many other factors.

An example of a large workload often deployed into IaaS environments is SAP and the SAP Hana in-memory data platform. SAP is one of the world's most installed software platforms used by enterprises for a range of business functions. Installing SAP in high performance and also highly restricted environments such as in Financial Services can present a number of technical challenges, which the Azure Cloud is uniquely capable of satisfying. Azure features such as High performance to support high transaction rates and tools to provide high levels of Data Governance, Security and Regulatory compliance, are key to Financial Services deployments of SAP and other large enterprise workloads.

About this guide

This Guide is provided as a head start for creating an Azure Cloud environment suitable for installation of large workloads with SAP as the example, in Financial Services. Towards that end we have documented guidance for proper configuration of:

- Virtual machines, including Azure Resource Manager (ARM) templates.
- Azure Blueprints for assignment of basic policies based on ISO27001
    - Additional Azure Policy Definitions for security and management
- Azure Blueprints mapped to standards to be applied as necessary, including:
    - NIST
    - PCI
- Other Azure Products and services that can be used to secure and manage the environment.

Assumptions – Look this over and generalize

Platform Selection

HANA Database Servers

This guide will assume the use of M-Series virtual machines as opposed to bare metal HANA Large Instances, for HANA Database nodes. While smaller Virtual Machines can be used the M-Series provide the performance and Memory capacity that most large enterprises are looking for.

Application Servers

For most application servers D-Series servers are utilized in the examples below. D-Series is sufficiently cost-effective while allowing for the advanced settings that we will discuss.

Additionally, any platform decisions need to be checked against SAP official list of supported platforms in Azure.

OS Versions and Images

For production systems, OS Version and VM type must appear on the SAP list of approved and certified configurations. Non-certified VMs are often used by organizations for pre-production systems.

For the purposes of this exercise we will utilize standard Linux Distributions available on Azure. If custom images are necessary, it is suggested that the standard image be modified and snapshotted.

SAP System Provisioning

Since each enterprise utilizes various technologies for system provisioning (eg. ARM templates, TerraForm, Ansible, etc.), this document will not make specific recommendations for provisioning and automation.

SAP Software Installation

Since each enterprise utilizes different components of the SAP system this guide will not specify guidance for SAP installation. Please see SAP's Launchpad site for information on installing SAP.

Target Enterprise Types

- The financial services industry encompasses many types of enterprises ranging from small brokerages and insurance companies to Globally Significant Banking and Insurance institutions. While this guide might be applicable to all of these, we have chosen to explore the needs of the largest institutions. In this way the most restrictive environments are covered, and less restrictive infrastructures can be served by removing some of the configurations and policies. The high-level needs of these Globally Significant Financial Services Institutions are, generally, as follows:
    - High performance to support high transaction rate mission critical processes
    - Data Governance and security
    - Regulatory compliance

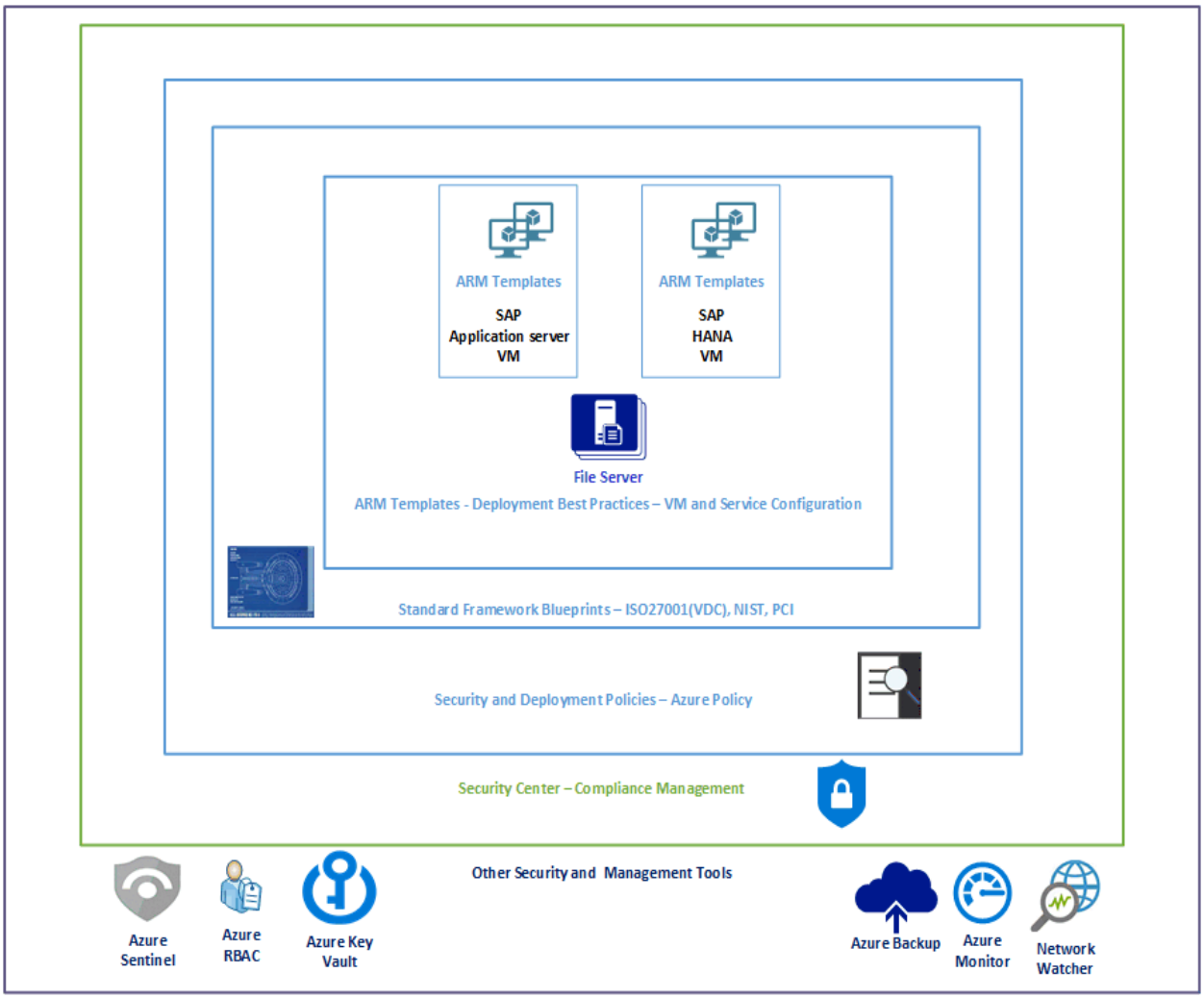
Financial Services Industry Specific Requirements

Beyond the basic infrastructure and environmental requirements ("ilities") Financial services has a list of requirements due to the highly regulated and high-performance nature of the industry.

- Security
    - Network Security - Networking must meet a rigorous standard to prevent infiltration and data exfiltration, including:
        - Perimeter Security configuration – Azure provides various perimeter security solutions including: Azure Firewall, Azure Virtual WAN and Azure Front Door.
        - Access controls through Network Security Groups (NSGs) – As part of the Virtual LAN configuration NSGs provide the ability to control traffic flow into and out of the environment.
        - Encrypted data transmission - Azure uses Transport Layer Security (TLS) protocol to protect your data when its moving between cloud services and customers, HTTPS while the data is moving between your on-premises infrastructure and Azure
    - Data Security – Data needs to be protected from exfiltration, from unwanted changes as well as loss. The Azure Cloud provides capabilities for:
        - Data at rest encryption – Azure uses symmetric encryption to encrypt and decrypt the data. By keeping the data encrypted on disk, it prevents compromise of data by a potential attacker.
        - Data in use – Azure protects data from being exposed while in use, example while being processed or in memory, by employing Azure Confidential Computing through Trusted Executed Environments (TEE).
- Standards Compliance
    - The solution must be compliant with Internationally recognized standards for infrastructure setup, security, and information protection.
- Compliance Checking
    - Compliance with standards and policies needs to be easily checked and the results confirmed as compliant or marked for remediation
- High Performance configuration –

    - The virtual machines must be configured to provide high performance. While this is relatively vague, the requirement is to provide configurations that will increase overall performance.
- Only Generally Available components and services
    - To be assured of proper support for configurations FSIs require generally available or minimally public preview services and components.

Solution Features



The following is a list of features that are present in the solution. A customer can choose to utilize all of these policies and configurations or to eliminate some of them based on internal governance rules.

International Standards Applied

ISO27001

ISO/IEC 27001 is an information security standard, part of the ISO/IEC 27000 family of standards. ISO/IEC 27001 specifies a management system that is intended to bring information security under management control and gives specific requirements. (source Wikipedia)

- ISO27001 Blueprint Overview
- Article on ISO27001 Blueprint control mapping

Sample Additional Policies

The following sample policies can be applied to the ISO27001 Blueprint to increase security and control. Please see the policy descriptions for deployment instructions. Many additional policies are available in the Azure Policy GitHub Repository.

- No Public facing IP addresses other than pre-approved and protected subnets
- Data at rest encryption using a customer supplied key on virtual machines
    - Data Lake Store
    - Storage Account Files
- Create VM using Managed Disk
- Audit if Key Vault has no virtual network rules

International Standards available

NIST SP 800-53 R4

The **NIST Cybersecurity Framework** provides a policy framework of computer security guidance for how private sector organizations in the United States and around the world can assess and improve their ability to prevent, detect, and respond to cyber-attacks. (source Wikipedia)

- NIST Blueprint Overview
- Article on NIST Control Mapping
- Article on NIST Blueprint Deployment

PCI-DSS-3.2.1

The **Payment Card Industry Data Security Standard** (**PCI DSS**) is an information security standard for organizations that handle branded credit cards. The standard was created to increase controls around cardholder data to reduce credit card fraud. (source Wikipedia)

- PCI Blueprint Overview
- Article on PCI Blueprint Control Mapping
- Article on PCI Blueprint Deployment Steps

Azure Products and features

Azure Sentinel

Azure Sentinel is a scalable, cloud-native, security information event management (SIEM) and security orchestration automated response (SOAR) solution. Azure Sentinel delivers intelligent security analytics and threat intelligence across the enterprise, providing a single solution for alert detection, threat visibility, proactive hunting, and threat response.

Azure Key Vault

Azure Key Vault helps safeguard cryptographic keys and secrets that cloud applications and services use. Key Vault streamlines the key management process and enables you to maintain control of keys that access and encrypt your data. Developers can create keys for development and testing in minutes, and then migrate them to production keys. Security administrators can grant (and revoke) permission to keys, as needed.

Role Based Access Control

Access management for cloud resources is a critical function for any organization that is using the cloud. Role Based Access Control (RBAC) helps you manage who has access to Azure resources, what they can do with those resources, and what areas they have access to. RBAC is an authorization system built on Azure Resource Manager that provides fine-grained access management of Azure resources.

Azure Security Center

Azure Security Center is a unified infrastructure security management system that strengthens the security posture of your data centers, and provides advanced threat protection across your hybrid workloads in the cloud - whether they're in Azure or not - as well as on premises.

Azure Backup

Azure Backup is a cost effective, one-click backup solution capable of data protection for on-premises servers, virtual machines, virtualized workloads, SQL server, SharePoint server, and more.

Azure Monitor

Azure Monitor, which now includes Log Analytics and Application Insights, provides sophisticated tools for collecting and analyzing telemetry that allow you to maximize the performance and availability of your cloud and on-premises resources and applications. It helps you understand how your applications are performing and proactively identifies issues affecting them and the resources they depend on.

Azure Network Watcher

Azure Network Watcher provides tools to monitor, diagnose, view metrics, and enable or disable logs for resources in an Azure virtual network.

VM Configuration Best Practices for SAP on Azure

Operating System Configuration

OS Versions and compatibilities

This guide and the associated configurations/templates will provide for the installation of two operating systems RedHat Linux and SUSE Linux. Differences in OS configuration will be noted below.

Waagent.conf versus cloud-init

The waagent.conf and cloud-init files serve similar purposes in setting up various cloud related and SAP related operating system configurations. It is highly recommended that only one of these configuration files be used and that the preferred file be waagent.conf. Use of waagent.conf only prevents the unexpected changing of configuration at VM boot time. A

OS Disk Configuration

Swap and temp file placement

The Linux swap file should always be placed on the “resource disk”, which is a locally attached ephemeral storage. This should be done via the waagent.conf configuration file. The following settings should be included in the waagent.conf file.

```
ResourceDisk.EnableSwap=y

ResourceDisk.SwapSizeMB=2048
```

Likewise, for best performance, temp files should also be written to the resource disk when needed. SAP application servers have a configuration for the location of temporary files – this setting should be changed to point to the resource disk.

Proximity of application servers to HANA DB servers

In order to reduce network latency between application and database servers it is a good idea to collocate these resources in the same physical complex. This can be achieved in two ways:

- Pinning resources, via ICM request, within an availability set to the same physical datacenter.
- The preferred method is to utilize Placement Groups (within VM scale sets). Part of the function of placement groups will be to keep the compute and storage for the virtual machines co-located in the same datacenter thereby eliminating latency concerns. High Availability

Virtual Machine (Infrastructure-as-a-Service IaaS) High availability can be achieved in multiple ways within the Azure cloud. Azure provides two native methods of failover:

High Availability Option: Availability Sets

An Azure Availability Set (AS) is a group of virtual machines that are deployed across fault domains and update domains within the same Azure Datacenter. Availability sets make sure that your application is not affected by single points of failure, like the network switch or the power unit of a rack of servers.

High Availability Option: Availability Zones

An Availability Zone (AZ) is like an availability set in that the virtual machines are deployed across fault and update domains. The difference is that AZs provides a higher level of availability by spreading the VMs across multiple Azure datacenters within the same region.

Storage and Disk Performance

Log and Data Volume configuration

For optimal performance, the disks attached to the HANA machines should be provisioned in accordance with the SAP on Azure documentation:

1. The log volume should be striped across two disks, with Write Accelerator enabled.
2. The Data volume should be striped across 2 or more disks.

Co-location of storage and compute resources

To achieve optimal performance, servers (Application and HANA Database) and storage nodes should be collocated in the same datacenter within a region. The standard placement algorithm for managed disks sometimes places disk resources for a VM in different datacenter, introducing significant latency in disk access. This can be corrected in one of two ways:

- The disks can be moved (processed in the background) to the same physical facility as the compute via an ICM.
- Azure managed disk will automatically re-balance the storage accounts containing the disks, to ensure the system doesn't hit storage account level limits. This may take some time as the new disks are placed using the same algorithm as above.
- The preferred method is to utilize Placement Groups (within VM scale sets). Part of the function of placement groups will be to keep the compute and storage for the virtual machines co-located in the same datacenter thereby eliminating latency concerns. This functionality should be tested and verified.

Disk Sizing

IOPS rating when using Azure premium managed disk is a function of disk size. To satisfy IOPS requirements the customer should determine the number of IOPS and then determine the size and number of disks necessary.

Storage Service Endpoints

Storage in Azure is a service as opposed to a direct connection to disk resources. In a restricted environment such as Financial Services this can lead to network delays, due to Forced-Tunneling or Hairpinning, while traversing security measures. To speed up storage access, by optimizing the network path, it is recommended that storage service endpoints should be implemented.

Write Acceleration

The Azure Write Accelerator (WA) is critical for achieving good performance for the SAP HANA servers. WA should **only** be used for the log disks on all HANA machines. One word of caution write accelerated disk cannot be backed up using Azure Backup. Either another facility must be used or log disk backup must be implemented by offloading the contents periodically.

Network Performance

Accelerated Networking

The proper operation of Accelerated Networking (AN) is critical for obtaining optimal network performance. AN should be verified and tested, from within the virtual machine, for all app servers and database servers before any go-live. When tested with iperf3, the network bandwidth between the application servers and the database server(s) should be at least 13 Gbps. See the Documentation of the Accelerated Networking testing process.

Kernel Bypass/Data Plane Development Kit (DPDK)

Data Plane Development Kit (DPDK) on Azure offers a faster user-space packet processing framework for performance-intensive applications. This framework bypasses the virtual machine's kernel network stack.

Single VNET/Subnet

To increase network performance between servers all servers should be deployed into a single VNET/Subnet.

Backup/Restore

Reliable backup and restore of SAP environments, is a critical component of managing any environment. SAP on Azure is no exception. For the purposes of this guide we will document the use of Azure Backup. Please note that as of January 2019 Azure Backup has a feature to ignore disks with set with Write Accelerator. While this will enable the backup of the virtual machine it requires several additional steps during backup and restore:

- Backup
    - Backup Hana logs to a non-Write Accelerated disk
    - Snapshot log volumes
- Restore
    - Restore the VM
    - Reattach snapshotted disks
    - Start the VM with database stopped
    - Restore log files
    - Start the Database

Our recommendation is to develop and test a script to run these processes.

The Azure Backup team will soon be introducing a new feature called “backint”, which is the HANA interface directly to a backup system.

Data in transit encryption

Encrypted communication between virtual machines over an Azure VNET/Subnet cannot be controlled by policy or set during VM configuration. The use of TLS or SSL must be configured by the user during operating system configuration.

Sample ARM Templates for VM Deployment

The following are commonly used ARM templates for deployment of common SAP components. These templates are part of a larger GitHub repository for SAP environment configurations. The reader is welcome to read through and utilize other configurations in the repo. These templates are examples and should be reviewed prior to use in the readers environment.

SAP Virtual Machines and Cluster

Large and Extra Large VMs are used for HANA Database servers dependent on the size and layout of the HANA in-memory database. Medium and Small VMs are most useful for application components of the SAP environment.

- Json template for Large VM
- Json template for Extra Large VM
- Json template for Medium VM
- Json template for Small VM
- SAP HANA Cluster for deployment of a database cluster as opposed to a standalone server.

SAP Jump box

The SAP HANA Jump box is used to hold software and data to be deployed to the SAP HANA cluster.

NFS Server

NFS service for common data storage among the SAP components.

Deployment

All deployments of SAP, or any other large IaaS workload, can differ based on varying enterprise standards or requirements. Given the above example, this section provides a basic set of steps to deploy the above environment.

- Create and deploy the base Azure Blueprint
    - Create a new Blueprint using ISO27001 as the basis
    - Add Additional Policies to the new Blueprint (some suggestions given above)
    - Deploy New Blueprint into the target Azure subscriptions
- Deploy additional Azure Blueprints as required
    - NIST
    - PCI
    - Other - Azure Blueprints Samples, Azure Blueprints GitHub Repository
- Build SAP Azure VMs and DBs using the sample templates
- Install SAP Software according to enterprise standards and requirements
- Load Data utilizing SAP Native tools and NFS services

Conclusion

As Migration and modernization efforts to the Azure Cloud continue to increase many workloads require the fine control of an IaaS environment. SAP in Azure is a model for large workloads deployed to the public cloud. Azure provides capabilities that Financial Services Institutions can utilize for performance, security, and regulatory compliance, for these large mission critical workloads. For more information on running SAP and other large IaaS worloads in the Microsoft Azure Cloud, please see the Azure Documentation site.

Thank You

Thank you to Meladie Espiritu for your guidance and advice as I started looking into Policy and Blueprints.

A special thank you to Ross Sponholtz for not only partnering with me on a couple of SAP on Azure projects but also being a major contributor of SAP knowledge to this paper.

Thank you to Astha Malik and Melissa Hovis for being my very trusted reviewers.