

# Reference Architecture for SAP HANA

---

- [Reference Architecture for SAP HANA](#)
  - [Objective](#)
  - [Approach](#)
  - [Table of Content](#)
  - [Contributing](#)

## Objective

SAP HANA database is offering many different options how to design the infrastructure.

There are different ways how to implement High Availability (HA) and Disaster Recovery (DR) and there are many optional SAP HANA extensions (like Extension Nodes, Dynamic Tiering, XSA, etc.) that can be deployed.

There are various considerations that must be taken into account when designing infrastructure - for example ability to seamlessly move tenant (tenant portability) or whole instance (instance portability) without breaking external connectivity to the component.

Additional challenge is how to configure hostname resolution for individual virtual IPs to enable support for certificates and ensure their validity in relation to tenant or instance portability.

As result of this there are too many ways how to deploy SAP HANA and therefore almost all published Reference Architectures are very high-level showing only the generic concepts and they lack important details how to correctly implement the solution.

Goal of this project is to provide **one standardized multi-cloud and on-premise architecture** for SAP HANA that is able to support as many optional extensions as possible and to offer implementation details for different platforms including Public Cloud vendors (AWS, Azure, GCP, etc.) as well as on-premise implementations (VMware, Bare Metal).

It is important to state that other architectures are still valid (as long as formally supported by SAP) and can be used for specific requirements or use cases.

## Approach

The approach taken by the team is driven by the opinion that it is simpler to remove the features rather than to add them and make them work in harmony with the rest of the design.

Basic steps are following:

1. Define complex requirements including most common optional features
2. Make Architectural Decisions (ADs) to reduce the amount of deployment options
3. Design infrastructure architecture meeting as many requirements as possible (one standardized architecture)
4. Derive simplified versions of the architecture by removing specific requirements
5. Modify the infrastructure architecture for different platforms by introducing platform-specific details

## Table of Content

- [Change Log](#)
  - [How to Contribute](#)
1. [Requirements](#)
  2. [Architectural Decisions](#)
  3. Generic SAP HANA Architecture
    - [Overall Architecture and Modularity](#)
    - [Module: Basic Architecture](#)
    - [Module: Virtual Hostname/IP](#)
    - [Module: High Availability](#)
    - [Module: Disaster Recovery](#)
    - [Module: Data Tiering Options](#)
    - [Module: SAP XSA](#)
    - [Alternative Implementations](#)
  4. Platform Specific Architecture
    - [Cloud IaaS: AWS](#)
    - [Cloud IaaS: Azure](#)
    - Cloud IaaS: IBM Cloud
    - Cloud IaaS: Google
    - On-premise: VMware
    - On-premise: IBM Power
  5. Operational Procedures
    - High Availability (HA) Operation
    - Disaster Recovery (DR) Operation
    - SAP HANA Instance Move
    - SAP HANA Tenant Move
  6. Additional Information
    - SAP HANA: Network Latency Requirements
    - SAP HANA: Stacking Options (MCOD, MCOS, MDC)
    - SAP HANA: Certificate setup

## Contributing

Please refer to [How to Contribute](#) to understand how to contribute to this project.

## Change Log

---

### 2020-09-06

- [Tomas Krojzl] Updated final content for following sections:

- 4. Platform Specific Architecture
  - Cloud IaaS: AWS
  - Cloud IaaS: Azure

- [Tomas Krojzl] Removed following sections (will be published in next release):

- 4. Platform Specific Architecture

- Cloud IaaS: IBM Cloud
- 5. Operational Procedures
  - High Availability (HA) Operation
  - Disaster Recovery (DR) Operation
  - SAP HANA Instance Move
  - SAP HANA Tenant Move

## 2020-05-13

- [Tomas Krojzl] Written initial content for following sections (Ready for Review):

- 4. Platform Specific Architecture
  - IaaS Cloud: AWS
  - IaaS Cloud: Azure

- [Tomas Krojzl] Updated initial content for following sections (Ready for Review):

- 3. Generic SAP HANA Architecture
  - Module: Data Tiering Options

## 2020-05-12

- [Vladimir Kovarik] Updated sections *Azure: Virtual Hostname/IP* in [pages/platform\\_specific\\_architecture/cloud\\_iaas\\_azure.md](#)

## 2020-04-23

- [Vladimir Kovarik] Updated sections *AWS: Virtual Hostname/IP* in [pages/platform\\_specific\\_architecture/cloud\\_iaas\\_aws.md](#)

## 2019-12-19

- [Tomas Krojzl] Written initial content for following sections (Ready for Review):

- 1. Requirements
- 2. Architectural Decisions
- 3. Generic SAP HANA Architecture
  - Overall Architecture and Modularity
  - Module: Basic Architecture
  - Module: Virtual Hostname/IP
  - Module: High Availability
  - Module: Disaster Recovery
  - Module: Data Tiering Options
  - Module: SAP XSA
  - Alternative Implementations

- [Tomas Krojzl] Created basic structure for following sections

- 4. Platform Specific Architecture
  - IaaS Cloud: AWS

- IaaS Cloud: Azure
- IaaS Cloud: IBM Cloud
- 5. Operational Procedures
  - High Availability (HA) Operation
  - Disaster Recovery (DR) Operation
  - SAP HANA Instance Move
  - SAP HANA Tenant Move

- [Tomas Krojzl] Updated content of **CONTRIBUTING.md** file

2019-08-08

- [Tomas Krojzl] Written initial content of **CONTRIBUTING.md** file
- [Tomas Krojzl] Written initial content of **README.md** file

2019-08-07

- [Tomas Krojzl] Repository created

## How to Contribute

---

If you want to contribute to a project and make it better, your help is very welcome. Please see below instructions how to contribute.

- [How to Contribute](#)
  - [1. Initial Setup](#)
    - [1.1. Create GitHub user](#)
    - [1.2. Add SSH key](#)
    - [1.3. Fork the repository](#)
    - [1.4. Clone the forked repository](#)
    - [1.5. Configure synchronization](#)
  - [2. Recurrent synchronization](#)
    - [2.1. Download new content from main project repository](#)
    - [2.2. Upload merged content to your GitHub repository](#)
  - [3. Add new Content and Commit](#)
    - [3.1. Add or edit the documentation](#)
    - [3.2. Commit the content to your local repository](#)
    - [3.3. Push changes to GitHub](#)
  - [4. Upload to main project repository](#)
    - [4.1. Synchronize your content with main project](#)
    - [4.2. Update CHANGELOG.md file](#)
    - [4.3. Create Pull Request](#)

### 1. Initial Setup

Perform following section only once at the beginning of your contribution.

#### 1.1. Create GitHub user

Detailed instructions are here: <https://help.github.com/en/articles/signing-up-for-a-new-github-account>

## 1.2. Add SSH key

Adding SSH key will enable password-less connectivity to GitHub.

Detailed instructions are here: <https://help.github.com/en/articles/adding-a-new-ssh-key-to-your-github-account>

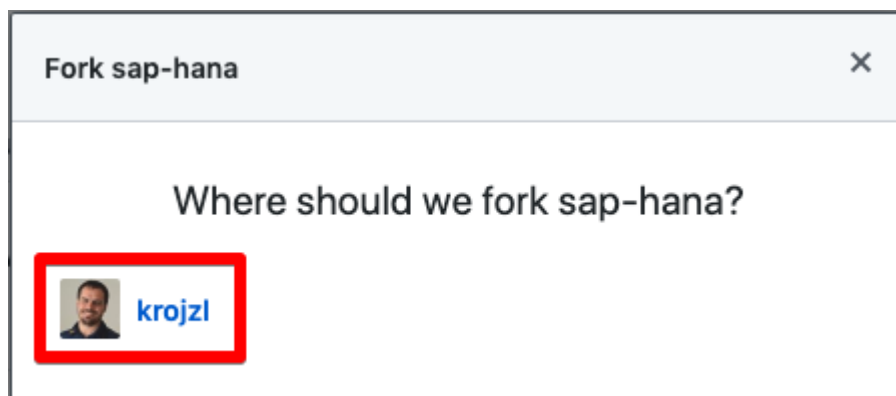
## 1.3. Fork the repository

Forking the `devel` repository will create your own personal copy of the repository.

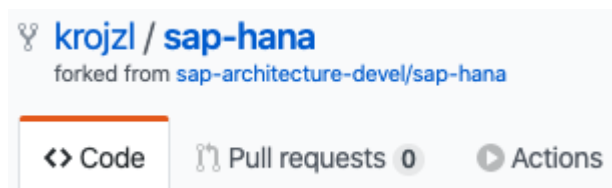
1. Navigate to the project repository: <https://github.com/sap-architecture-devel/sap-hana>
2. Click on **Fork** button in upper-right corner of the page



3. Select your user to create copy in your private space



4. As result you should see that repository was forked to your personal space:

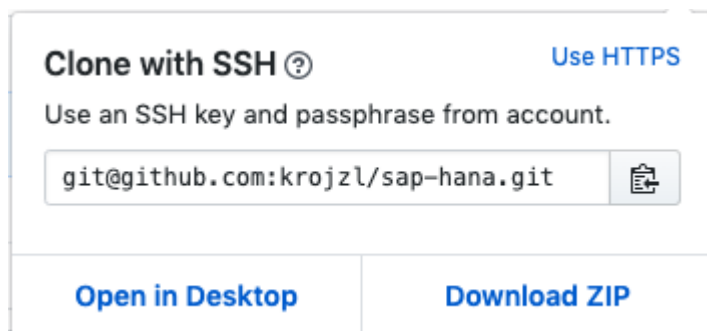


Detailed instructions are here: <https://help.github.com/en/articles/fork-a-repo>

## 1.4. Clone the forked repository

Clone your personal copy of the repository to your workstation.

1. Navigate to the forked repository: <https://github.com/<YOUR-USER>/sap-hana>
2. Click **Clone or download** button in upper-right corner of the page



Note: In order to use SSH to connect to GitHub you need to click on **Use SSH** in upper-right corner of the panel

3. Copy the URL from the panel: `git@github.com:<YOUR-USER>/sap-hana.git`
4. Open Terminal and change directory to desired location
5. Run `git` command to clone the repository

```
# git clone git@github.com:<YOUR-USER>/sap-hana.git
Cloning into 'sap-hana'...
remote: Enumerating objects: 7, done.
remote: Counting objects: 100% (7/7), done.
remote: Compressing objects: 100% (6/6), done.
remote: Total 7 (delta 0), reused 3 (delta 0), pack-reused 0
Receiving objects: 100% (7/7), done.
```

Detailed instructions are here: <https://help.github.com/en/articles/fork-a-repo>

## 1.5. Configure synchronization

Configure forked repository synchronization with main project repository.

1. Open Terminal and change directory to location of your local copy of the repository: `cd /path-to-your-repository/sap-hana`
2. List currently defined remote repositories:

```
# git remote -v
origin    git@github.com:<YOUR-USER>/sap-hana.git (fetch)
origin    git@github.com:<YOUR-USER>/sap-hana.git (push)
```

3. Add link to main project repository: `git remote add upstream git@github.com:sap-architecture-devel/sap-hana.git`
4. List again defined remote repositories:

```
# git remote -v
origin    git@github.com:<YOUR-USER>/sap-hana.git (fetch)
```

```
origin    git@github.com:<YOUR-USER>/sap-hana.git (push)
upstream  git@github.com:sap-architecture-devel/sap-hana.git (fetch)
upstream  git@github.com:sap-architecture-devel/sap-hana.git (push)
```

Detailed instructions are here: <https://help.github.com/en/articles/fork-a-repo>

## 2. Recurrent synchronization

Perform following section on regular basis and always before you create Pull Request to main project.

### 2.1. Download new content from main project repository

Download and merge new updates from main project repository into your local repository on your workstation.

1. Open Terminal and change directory to location of your local copy of the repository: `cd /path-to-your-repository/sap-hana`
2. Fetch the branches and their content from the main project repository:

```
# git fetch upstream
remote: Enumerating objects: 13, done.
remote: Counting objects: 100% (13/13), done.
remote: Compressing objects: 100% (10/10), done.
remote: Total 11 (delta 0), reused 10 (delta 0), pack-reused 0
Unpacking objects: 100% (11/11), done.
From github.com:sap-architecture-devel/sap-hana
* [new branch]      master      -> upstream/master
```

3. Change local repository branch to `master`

```
# git checkout master
Already on 'master'
Your branch is up to date with 'origin/master'.
```

4. Merge changes from `upstream/master` into your local `master` branch

```
# git merge upstream/master
Updating 11569ab..469f8ab
Fast-forward
CHANGELOG.md          | 5 +++++
CONTRIBUTING.md      | 46
+++++
README.md             | 3 ++-
images/contributing-clone-1.png | Bin 0 -> 18798 bytes
images/contributing-fork-1.png  | Bin 0 -> 6350 bytes
images/contributing-fork-2.png  | Bin 0 -> 15140 bytes
images/contributing-fork-3.png  | Bin 0 -> 12140 bytes
```

```
7 files changed, 53 insertions(+), 1 deletion(-)
create mode 100644 CHANGELOG.md
create mode 100644 CONTRIBUTING.md
create mode 100644 images/contributing-clone-1.png
create mode 100644 images/contributing-fork-1.png
create mode 100644 images/contributing-fork-2.png
create mode 100644 images/contributing-fork-3.png
```

5. In case there were merge conflicts you need to resolve them as described in

<https://help.github.com/en/articles/resolving-a-merge-conflict-using-the-command-line>

Detailed instructions are here: <https://help.github.com/en/articles/syncing-a-fork>

## 2.2. Upload merged content to your GitHub repository

Upload merged content from your local repository on your workstation to your GitHub repository.

1. Open Terminal and change directory to location of your local copy of the repository: `cd /path-to-your-repository/sap-hana`
2. Upload the merged content in active branch to your GitHub repository

```
# git push
Total 0 (delta 0), reused 0 (delta 0)
To github.com:<YOUR-USER>/sap-hana.git
11569ab..469f8ab master -> master
```

3. Now you can see the content online in your own personal copy of the repository in GitHub:

<https://github.com/<YOUR-USER>/sap-hana>

Detailed instructions are here: <https://help.github.com/en/articles/pushing-commits-to-a-remote-repository>

## 3. Add new Content and Commit

Deliver new content by editing files on your local workstation, push to online GitHub and review.

### 3.1. Add or edit the documentation

The documentation is written in Markdown language.

Additional information about Markdown:

- [Basic writing and formatting syntax](#)
- [Working with advanced formatting](#)
- [Mastering Markdown](#)
- [GitHub Flavored Markdown Spec](#)

Images should be uploaded into subdirectory `images` below directory where the page referencing the image is located.



Files should be uploaded into subdirectory **files** below directory where the page referencing the image is located.

Make sure you name all images and files in clear way.

All links to pages, images or files located in this repository should be based on relative paths - avoid using absolute paths as this will break the link functionality in forked repositories.

Use your favorite editor to add new content. Use Linting function to deliver clean and well-structured documentation.

Recommended editors:

- [Visual Studio Code](#)
  - plugin [Markdown All in One](#)
  - plugin [Markdown Lint](#)
  - plugin [Auto Markdown TOC](#)
- [Atom Editor](#)

### 3.2. Commit the content to your local repository

When unit of work is completed commit the changes to your local repository on your local workstation.

Use either your editor to commit the changes (recommended) or perform following commands:

1. Open Terminal and change directory to location of your local copy of the repository: `cd /path-to-your-repository/sap-hana`
2. Check the status

```
# git status
On branch master
Your branch is up to date with 'origin/master'.

Changes not staged for commit:
  (use "git add <file>..." to update what will be committed)
  (use "git checkout -- <file>..." to discard changes in working
  directory)

        modified:   CONTRIBUTING.md

no changes added to commit (use "git add" and/or "git commit -a")
```

Note: In example above we can see file **CONTRIBUTING.md** is NOT part of commit.

3. Add files that should be included in commit

- Add individual files: `git add <file_name>`
- Add directory: `git add <directory_name>`

In order to add all files, perform `git add .`

## 4. Check the status again

```
# git status
On branch master
Your branch is up to date with 'origin/master'.

Changes to be committed:
  (use "git reset HEAD <file>..." to unstage)

        modified:   CONTRIBUTING.md
```

## 5. Commit changes to active branch

```
# git commit -m "Update to CONTRIBUTING.md"
[master d6a2568] Update to CONTRIBUTING.md
 1 file changed, 1 insertion(+), 1 deletion(-)
```

Detailed instructions are here: <https://help.github.com/en/articles/adding-a-file-to-a-repository-using-the-command-line>

## 3.3. Push changes to GitHub

Push your local changes from your local repository on your workstation to your GitHub repository.

1. Open Terminal and change directory to location of your local copy of the repository: `cd /path-to-your-repository/sap-hana`
2. Push changes from active branch to your GitHub repository

```
# git push
Enumerating objects: 8, done.
Counting objects: 100% (8/8), done.
Delta compression using up to 12 threads
Compressing objects: 100% (6/6), done.
Writing objects: 100% (6/6), 3.23 KiB | 3.23 MiB/s, done.
Total 6 (delta 3), reused 0 (delta 0)
remote: Resolving deltas: 100% (3/3), completed with 1 local object.
To github.com:<YOUR-USER>/sap-hana.git
 469f8ab..d6a2568  master -> master
```

3. Review your changes online in your own personal copy of the repository in GitHub:

<https://github.com/<YOUR-USER>/sap-hana>

Detailed instructions are here: <https://help.github.com/en/articles/adding-a-file-to-a-repository-using-the-command-line>

## 4. Upload to main project repository

Create Pull Request (PR) from your own GitHub repository against main project repository.

#### 4.1. Synchronize your content with main project

Very likely your updates took some time. Meanwhile the content in main project repository changed and might be out of sync with your own repository.

Follow procedure described in section 2. [Recurrent synchronization](#).

#### 4.2. Update **CHANGELOG.md** file

Make sure you document what was changed in **CHANGELOG.md** file. Make sure this very last change to avoid merge conflicts.

Follow procedure described in section 3. [Add new Content and Commit](#) to adjust the file.

#### 4.3. Create Pull Request

Before you create Pull Request (PR) make sure that:

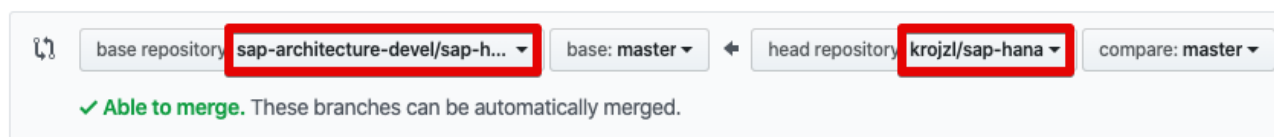
- you downloaded and merged latest content from main project repository
- you resolved all merge conflicts
- you pushed all changes from your local repository on your workstation to online GitHub repository
- you reviewed your content online in GitHub <https://github.com/<YOUR-USER>/sap-hana> and confirmed the content is rendering correctly - that includes:
  - all images are properly displayed (and are relative)
  - all links are working properly (and are relative is pointing to page in same repository)

1. Navigate to the forked repository: <https://github.com/<YOUR-USER>/sap-hana>

2. Click **New pull request** button in upper-left part of the page



3. Review the Pull Request (PR) details



Note: On left side you see destination (main project repository) and on left side you see source (your own forked repository).

4. In case all visualized changes are ok press **Create pull request** button



5. Add title and meaningful description and click **Create pull request** button

The screenshot shows the GitHub 'Changes to CONTRIBUTING.md file' interface. At the top, there's a title bar 'Changes to CONTRIBUTING.md file'. Below it are two tabs: 'Write' (active) and 'Preview'. To the right of the tabs is a rich text editor toolbar with icons for bold (AA), italic (i), quote ("), code (<>), link, list, table, checkmark, mention (@), and other actions. The main text area contains the text 'Initial content of CONTRIBUTING.md file'. Below the text area is a dashed line with the text 'Attach files by dragging & dropping, selecting or pasting them.' and a small 'M4' icon. At the bottom left, there's a checkbox 'Allow edits from maintainers.' which is checked, followed by a blue link 'Learn more'. At the bottom right is a green button 'Create pull request' with a dropdown arrow.

6. Wait for your Pull Request to be reviewed and respond to any request for changes

Note: Until your Pull Request is accepted all your commits that are pushed to your forked repository in GitHub will be automatically included in given Pull Request.

If you want to make commits that are not included in Pull Request, consider using separate branch for your updates.

Detailed instructions are here: <https://help.github.com/en/articles/creating-a-pull-request-from-a-fork>

## Requirements

---

Following requirements were taken into account for this Reference Architecture. Different requirements might lead to different Reference Architectures.

- [Requirements](#)
  - [REQ1: Scale-out Driven Architecture](#)
  - [REQ2: High Availability \(HA\) Across Availability Zones \(AZs\)](#)
  - [REQ3: Disaster Recovery \(DR\) into Remote Location](#)
  - [REQ4: SAP HANA Multitenant Database Containers \(MDC\) Implementation](#)
  - [REQ5: Virtual IP and Hostname for SAP HANA System Relocation](#)
  - [REQ6: Support for Data Tiering Options \(related to REQ2 and REQ4\)](#)
  - [REQ7: Enabled for Instance move \(related to REQ5\)](#)
  - [REQ8: Enabled for Tenant move \(related to REQ2 and REQ5\)](#)
  - [REQ9: Fully TLS enabled \(related to REQ4, REQ7 and REQ8\)](#)

### REQ1: Scale-out Driven Architecture

SAP HANA database is having ability to run in distributed way across multiple VMs (or physical servers). The Reference Architecture should take into account scale-out deployment option to ensure that single-node

design can be easily extended into scale-out without major architectural redesign.

## REQ2: High Availability (HA) Across Availability Zones (AZs)

Most IaaS offerings are supporting concept of Availability Zones - these are physically separated infrastructure segments that should not share any Single Point of Failure (SPOF), however are in close proximity to deliver low latencies required for Synchronous Replication. The Reference Architecture should take advantage of concept of Availability Zones to maximize the resiliency.

## REQ3: Disaster Recovery (DR) into Remote Location

Due to close proximity of Availability Zones these might not be seen as sufficient for Disaster Recovery purpose. Therefore, the Reference Architecture should support Disaster Recovery function by shipping the data into Remote Location.

## REQ4: SAP HANA Multitenant Database Containers (MDC) Implementation

SAP HANA is capable to run multiple independent database containers as part of one SAP HANA system. The Reference Architecture should support ability to run more than one database container.

## REQ5: Virtual IP and Hostname for SAP HANA System Relocation

To enable SAP HANA portability the SAP HANA instance should be decoupled from underlying Operating System by using its own Virtual IP and Virtual Hostname. This will simplify the SAP HANA relocation and will prevent the need to change hostnames or IPs during the migrations. The Reference Architecture should support this capability.

## REQ6: Support for Data Tiering Options (related to REQ2 and REQ4)

SAP invented various options how to distribute the data based on frequency of usage. The Reference Architecture should support following Data Tiering options (if technically viable):

- SAP HANA Native Storage Extensions (NSE)
- SAP HANA Extension Nodes
- SAP HANA Dynamic Tiering (DT)

**Note:** This requirement might be potentially conflicting with other requirements (in particular with REQ2 and REQ4).

## REQ7: Enabled for Instance move (related to REQ5)

In certain cases, it might be required to move instance of SAP HANA database to new VM (for example move from VM to Physical Server). The Reference Architecture should support such relocation without the need to change any configuration on connecting applications by ensuring that IP address and Hostname will be preserved.

**Note:** This requirement is related to requirement REQ5.

## REQ8: Enabled for Tenant move (related to REQ2 and REQ5)

SAP HANA database is supporting ability to relocate the database tenant into another SAP HANA database. The Reference Architecture should support such tenant relocation without the need to change any configuration on connecting applications by ensuring that IP address and Hostname will be preserved.

**Note:** This requirement is related to requirement REQ2 and REQ5.

## REQ9: Fully TLS enabled (related to REQ4, REQ7 and REQ8)

SAP HANA is supporting ability to encrypt the database communication by using Transport Layer Secure (TLS) / Secure Sockets Layer (SSL) protocol. Since Fully Qualified Domain Name (FQDN) is part of the TLS/SSL configuration the Reference Architecture should properly define usage of FQDNs for individual database containers (related to REQ4) and minimize the need to recreate the certificates as result of Instance Move (REQ7) and/or Tenant Move (REQ8).

## Architectural Decisions

Following Architectural Decisions (ADs) were made as part of this Reference Architecture. Different ADs might lead to different Reference Architectures.

- [Architectural Decisions](#)
  - [AD1: High Availability Concept](#)
  - [AD2: Disaster Recovery Concept](#)
  - [AD3: High Availability Takeover Automation](#)

### AD1: High Availability Concept

ID	AD1
<b>Name</b>	High Availability Concept
<b>Description</b>	SAP HANA is supporting multiple fundamentally different High Availability concepts. One needs to be selected.
<b>Assumptions</b>	Overall design should support both single-node and scale-out in parallel next to each other. Objective is to minimize the Recovery Time Objective (RTO).
<b>Options</b>	1. <a href="#">SAP HANA Host Auto-Failover (HAF)</a> 2. <a href="#">SAP HANA System Replication (synchronous)</a>
<b>Decision</b>	2. SAP HANA System Replication (synchronous)
<b>Justification</b>	- this is the only option that is supporting <a href="#">REQ2</a> for scale-out systems - the Recovery Time Objective (RTO) values are significantly smaller compared to 1. - this option is supporting additional features like <a href="#">Active/Active (Read Enabled)</a> or <a href="#">Secondary Time Travel</a>
<b>Comment</b>	Recommended Replication Mode is <a href="#">SYNC</a> in case there is possible shared Single Point of Failure (SPOF) or <a href="#">SYNCMEM</a> in case of two physically separated infrastructures. Recommended Operation Mode is <a href="#">logreplay</a> (or <a href="#">logreplay_readaccess</a> ).

## AD2: Disaster Recovery Concept

ID	AD2
<b>Name</b>	Disaster Recovery Concept
<b>Description</b>	SAP HANA is supporting multiple fundamentally different Disaster Recovery concepts. One needs to be selected.
<b>Assumptions</b>	-
<b>Options</b>	1. <a href="#">Storage Replication</a> 2. <a href="#">SAP HANA System Replication (asynchronous)</a>
<b>Decision</b>	2. SAP HANA System Replication (asynchronous)
<b>Justification</b>	<ul style="list-style-type: none"> <li>- option 1. might or might not be available and is unlikely to work cross-platform</li> <li>- option 2. is part of the product and therefore always available, it is platform independent and will work even cross-platform</li> <li>- as part of option 2. all data pages are checked for consistency during the transfer to secondary site</li> </ul>
<b>Comment</b>	<p>Replication Mode must be <b>ASYNC</b> to avoid performance impact.</p> <p>Operation Mode must be same for all tiers (either <b>delta_datashipping</b> or <b>logreplay/logreplay_readaccess</b>), combining Operations Modes is not supported.</p> <p>Operation mode <b>logreplay_readaccess</b> is available only between primary and secondary system.</p>

Note: Combination of AD1 and AD2 will lead to usage of [SAP HANA Multitarget System Replication](#) (or [SAP HANA Multitier System Replication](#) in case of SAP HANA 1.0).

## AD3: High Availability Takeover Automation

ID	AD3
<b>Name</b>	High Availability Takeover Automation
<b>Description</b>	There are different options/products how High Availability Takeover can be executed. One needs to be selected.
<b>Assumptions</b>	Objective is to minimize the Recovery Time Objective (RTO).
<b>Options</b>	1. Manual Takeover (no automation) 2. Pacemaker Cluster (Linux native solution) 3. 3rd Party Clustering Solution
<b>Decision</b>	2. Pacemaker Cluster (Linux native solution)
<b>Justification</b>	<ul style="list-style-type: none"> <li>- option 1. is not satisfying requirement to minimize the Recovery Time Objective (RTO) value</li> <li>- option 2. is seen as recommended option by both OS vendor and SAP and is most common HA solution</li> <li>- option 3. might not be available across all platforms</li> </ul>

## Overall Architecture and Modularity

---

This section is explaining overall concept and how to use this Reference Architecture.

- [Overall Architecture and Modularity](#)
  - [Modular Concept](#)
  - [Resulting Architecture Is Too Complex](#)
  - [How to Read This Reference Architecture](#)

### Modular Concept

Overall approach how this Reference Architecture was created is briefly described [here](#).

Individual customers are having different requirements. These requirements are dictating how the resulting architecture will look like. Unfortunately, initial requirements are quite often not final - customers will need to enable new functions and features and this might subsequently drive the need to significantly change the architecture and reimplement SAP HANA database.

This Reference Architecture is trying to minimize the changed to SAP HANA architecture by offering individual "modules" which are pre-integrated together. These modules are optional and solution architects should choose only those modules which are required in given scenario. The benefit is that additional modules can be enabled later without the need to redeploy the SAP HANA database.

### Resulting Architecture Is Too Complex

If all modules are selected, then resulting Reference Architecture is quite complex. This is because of complex requirements which are to be satisfied. If you consider the resulting architecture to be way too complicated, then try removing some optional modules or use simplified version of the modules (for example in area of Pacemaker configuration).

### How to Read This Reference Architecture

Next sections are started by describing the basic setup - simple Single-Node and Scale-Out deployment in single Availability Zone. Subsequent modules are then always building up on top of each other - each module describing how to include one additional function until we end up with full scenario having all options enabled.

Since all modules are optional - it is technically impossible to document all permutations of choices as separate diagrams. Therefore, it is assumed that Solution Architect working with this Reference Architecture is familiar with SAP HANA technology and is able to properly combine available modules creating his own architecture based on preferred choices.

## Module: Basic Architecture

---



This is foundational module for SAP HANA Reference Architecture. Two basic SAP HANA deployment options are explained - Single-Node and Scale-Out.

- [Module: Basic Architecture](#)
  - [SAP HANA Multitenant Database Containers \(MDC\) vs Single-tenant Implementation](#)
  - [SAP HANA Stacking Options \(MCOD, MCOS, MDC\)](#)
  - [Single-Node SAP HANA System \(in single Availability Zone\)](#)
  - [Scale-Out SAP HANA System \(in single Availability Zone\)](#)

## SAP HANA Multitenant Database Containers (MDC) vs Single-tenant Implementation

SAP HANA System can be implemented using one of two basic deployment types:

- SAP HANA Multitenant Database Containers
- SAP HANA Single-tenant Implementation

SAP HANA Multitenant Database Containers (MDC) were introduced in SAP HANA 1.0 SP09 and as of SAP HANA 2.0 SP01 the MDC concept is the only supported deployment option. Therefore, single-tenant implementation is considered outdated and is not recommended to be used.

This Reference Architecture is based on using Multitenant Database Containers deployment option however, it could be modified also for Single-tenant Implementation.

## SAP HANA Stacking Options (MCOD, MCOS, MDC)

There are different techniques how to stack multiple SAP HANA Databases on same Virtual Machine (VM):

- Multiple Components on One Database (MCOD)
- Multiple Components on One System (MCOS)
- Multitenant Database Containers (MDC)

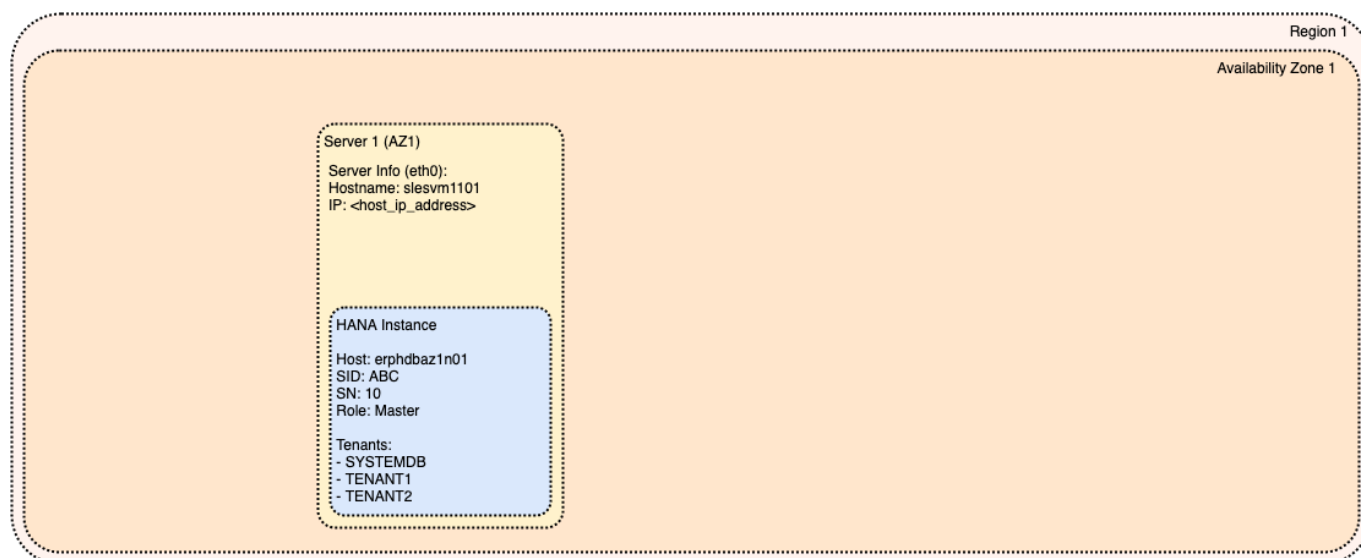
Recommended stacking option is MDC which is fully compatible with this Reference Architecture.

Stacking option MCODE is intended to be used only in specific scenarios wherever explicitly recommended by SAP.

Stacking option MCOS is not recommended and is not aligned with this Reference Architecture.

Additional information is available in section SAP HANA: Stacking Options (MCOD, MCOS, MDC) {TODO}.

## Single-Node SAP HANA System (in single Availability Zone)



This is the simplest and most traditional deployment option. All components of SAP HANA database are running together on one single Virtual Machine (VM) having one Operating System (OS).

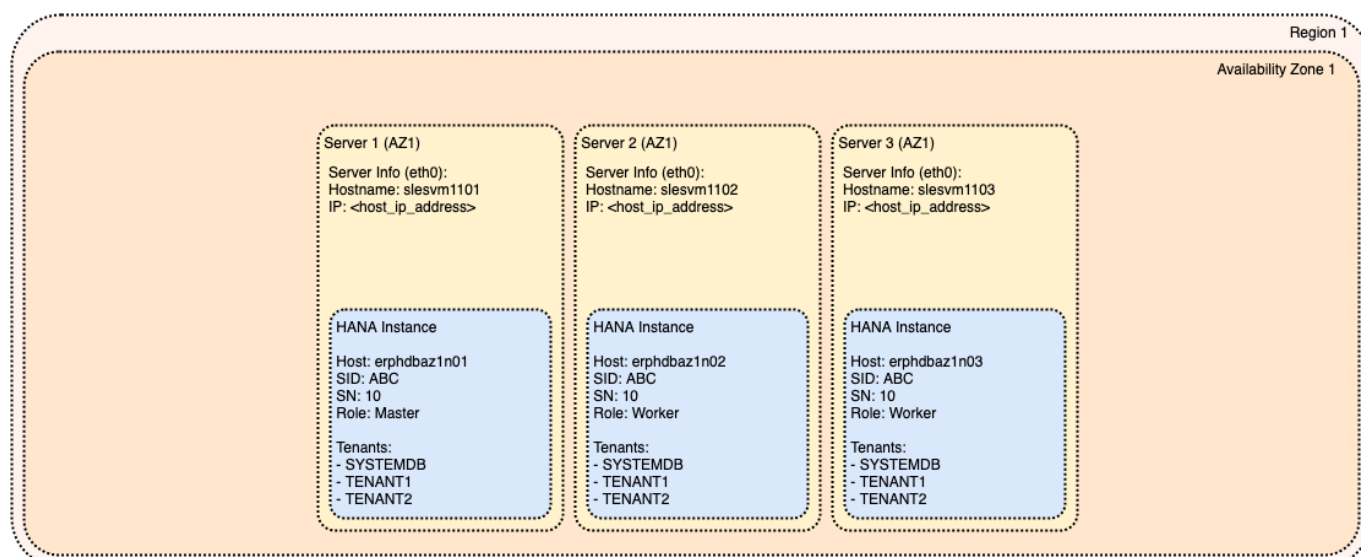
This deployment option is vulnerable to all failure scenarios:

- Regional failure (Disaster Event)
- Availability Zone failure
- VM failure / OS failure
- SAP HANA instance failure

Additional information:

- [Master Guide: Single-Host System](#)

## Scale-Out SAP HANA System (in single Availability Zone)



This deployment option is running one SAP HANA database as multiple components distributed across multiple Virtual Machines (VMs) each having its own Operating System (OS).

SAP HANA internal architecture is based on "shared nothing" principle to allow almost linear scalability - this means that each instance (set of processes running on one Operating System) is having its own Data Files and Log Files which are NOT accessible by other instances (running on other VMs).

It is always recommended to prefer Single-Node deployment option over Scale-Out due to a performance reason. Scale-Out option is positioned to overcome the limitations of single VM (or physical server) in terms of maximal memory or CPU power.

This deployment option is not available for all SAP products - for more information check Master Guide and Installation Guide for given SAP product.

This Reference Architecture will use Scale-Out deployment option as base for all other modules. Single-Node architecture is seen as simplified version of Scale-Out deployment option. Any potential differences between Single-Node and Scale-Out designs will be explicitly highlighted.

This deployment option is vulnerable to all failure scenarios:

- Regional failure (Disaster Event)
- Availability Zone failure
- VM failure / OS failure
- SAP HANA instance failure

Note: Multiple nodes are NOT offering increased availability. Due to a "shared nothing" approach, failure on any of the nodes will impact whole SAP HANA database system.

Additional information:

- [Master Guide: Distributed System \(Multiple-Host System\)](#)
- [Administration Guide: Multiple-Host \(Distributed\) Systems](#)

## Module: Virtual Hostname/IP

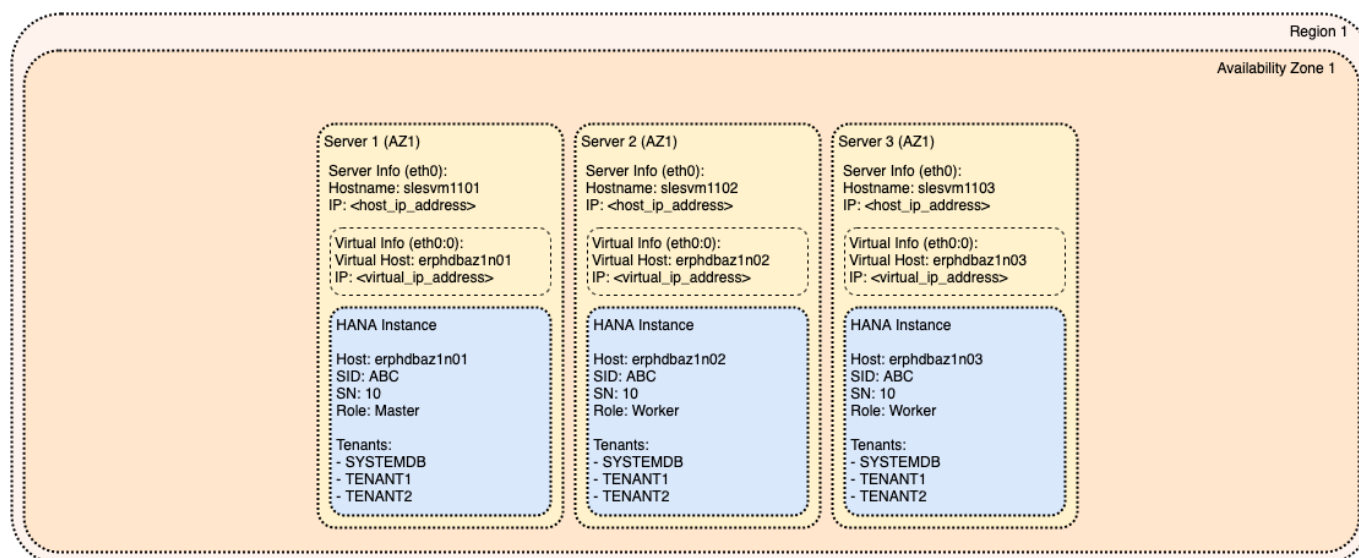
---

In this module the basic architecture is extended by decoupling SAP HANA installation from OS installation by using Virtual Hostname and Virtual IP address dedicated to SAP HANA instance.

This module is prerequisite to support Instance Move as documented in [SAP HANA Instance Move](#).

- [Module: Virtual Hostname/IP](#)
  - [Concept of Virtual Hostname and Virtual IP](#)

### Concept of Virtual Hostname and Virtual IP



Hostname and IP address is critical attribute uniquely identifying Operating System running on given Virtual Machine (VM). It is registered in many backend Systems of Management (including Active Directory integration, DNS integration, OS monitoring, etc.) and Systems of Records. Since is tightly coupled with the identity of given Operating System quite often it cannot be easily changed.

However, SAP HANA instance must be installed against some Hostname (see [Administration Guide: Default Host Names and Virtual Host Names](#) for additional information). Unless specific Hostname is provided then Hostname of Operating System is used.

Hostname used to install SAP HANA and its Fully Qualified Domain Name (FQDN) form is playing critical role in regard to connected applications and usage of Certificates to encrypt network communication (see [Security Guide: TLS/SSL Configuration on the SAP HANA Server](#) for additional information).

Therefore, many customers are having requirement to preserve the Hostname used during SAP HANA installation so that external connectivity is not disrupted following the migration.

Unfortunately, this is in direct contradiction with requirement to have different Hostname during SAP HANA System Replication which is recommended option how to minimize the business downtime during homogeneous migration of SAP HANA system (see [Administration Guide: General Prerequisites for Configuring SAP HANA System Replication](#) for additional information).

Convenient solution is usage of Virtual Hostname and Virtual IP for SAP HANA installation that is always following given SAP HANA instance. The advantage is that SAP HANA installation is decoupled from Operating System and can be easily relocated to new Operating System.

The procedure how to relocate SAP HANA system using Virtual Hostname/IP is described in [SAP HANA Instance Move](#).

The real implementation of Virtual Hostname is platform specific and is described in detail in Platform Specific Architecture part of the documentation.

Note the difference between the Virtual Hostname/IP and Cluster Hostname/IP. Purpose of Cluster Hostname/IP is to always to follow "active" instance of SAP HANA High Availability (HA) solution while Virtual Hostname/IP is static and is always following given instance regardless of its role in HA solution.

# Module: High Availability

---

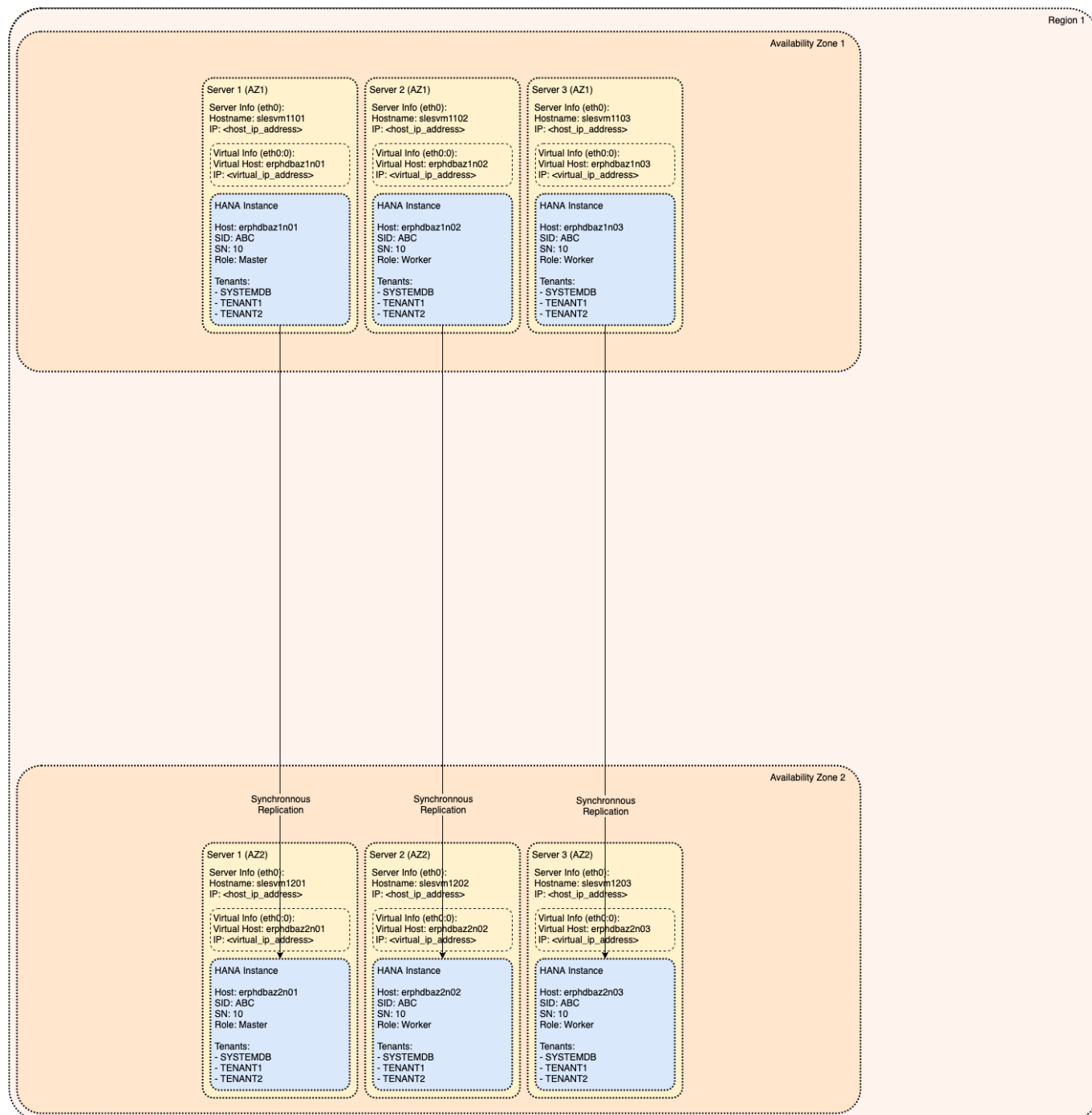
This module is enhancing the architecture by additional SAP HANA system increasing the Availability.

High Availability scenario is based on the Pacemaker cluster automating the takeover to secondary SAP HANA system.

Different options how Cluster IP can be configured are presented - each having its own advantages and disadvantages.

- [Module: High Availability](#)
  - [Enhanced Availability without Clustering \(manually operated\)](#)
  - [High Availability with Pacemaker Cluster](#)
    - [IPMI-like Fencing](#)
    - [SBD \(Storage Based Death\) Fencing](#)
    - [Majority Maker Node](#)
      - [SAP HANA Scale-Out Scenario](#)
      - [SAP HANA Single-Node Scenario](#)
    - [Cluster IP Design](#)
      - [Typical Cluster IP Implementation](#)
  - [Active/Active High Availability with Pacemaker Cluster](#)
  - [Active/Active High Availability with Pacemaker Cluster \(enabled for Tenant Move\)](#)

## Enhanced Availability without Clustering (manually operated)



This is the very basic option how to increase SAP HANA Availability by adding secondary SAP HANA system in separate Availability Zone and configuring synchronous SAP HANA System Replication (see [Administration Guide: Replication Modes for SAP HANA System Replication](#) for additional information).

Following two Replication Modes are acceptable for Availability management:

- Synchronous on disk (**SYNC**)
- Synchronous in-memory (**SYNMEM**)

Synchronous on disk (**SYNC**) Replication Mode is having higher latency impact because it waits for disk write operation on secondary SAP HANA system to complete. The advantage is that Recovery Point Objective (RPO) is guaranteed to be zero (no data loss possible as long as secondary system is connected). This option is recommended in situations where we have potential Single Point of Failure (SPOF) shared between both primary and secondary SAP HANA system.

Synchronous in-memory (**SYNCMEM**) Replication Mode is having Recovery Point Objective (RPO) only "close to zero" because information on secondary SAP HANA database is written to disk asynchronously. The advantage is improved performance because the latency impact is reduced by disk write operation. However, this Replication Mode can lead to data loss in case that both primary and secondary SAP HANA system will fail at the same time - therefore it is recommended only in scenarios where there is no Single Point of Failure (SPOF) shared between both primary and secondary system - for example in combination with Availability Zones.

Note that **Full Sync Option** as described in [Administration Guide: Full Sync Option for SAP HANA System Replication](#) is not suitable for any High Availability usage. This is because any failure (either of primary or secondary SAP HANA System) will result in remaining SAP HANA System to be blocked.

Asynchronous (**ASYNC**) Replication Mode is not acceptable because it cannot guarantee Recovery Point Objective (RPO) to be zero or "close to zero".

In this scenario the takeover is executed manually by SAP HANA administrator (see [Administration Guide: Performing a Takeover](#) for additional information) and therefore the Recovery Time Objective (RTO) depends mainly on monitoring lag and reaction time of support teams.

There are two techniques how to ensure that application connectivity to SAP HANA database is not disrupted following the takeover operation (see [Administration Guide: Client Connection Recovery After Takeover](#) for additional information):

- IP redirection (referred as Cluster IP)
- DNS redirection

IP redirection (repointing Cluster IP to new primary SAP HANA system after takeover) or DNS redirection is executed manually following the takeover action.

The implementation details for Cluster IP are platform specific and are described in Platform Specific Architecture part of the documentation.

Additional Information:

- [How To SAP HANA System Replication Whitepaper](#)
- [SAP Note 2407186: How-To Guides & Whitepapers For SAP HANA High Availability](#)

## High Availability with Pacemaker Cluster

In order to decrease the Recovery Time Objective (RTO) the takeover process must be automated. This can be done using Pacemaker as cluster management solution.

Additional Information:

- [ClusterLabs: Pacemaker Documentation](#)
- [SLES12 SP4: High Availability Extension - Administration Guide](#)
- [SLES15 GA: High Availability Extension - Administration Guide](#)
- [RHEL7: High Availability Add-On Reference](#)

When dealing with Pacemaker cluster there are two topics that are impacting the final architecture:

- Implementation of Cluster IP

- Fencing Mechanism

Technical implementation of Cluster IP is specific to given infrastructure and therefore is in detail described in Platform Specific Architecture part of the documentation. In this section of the Reference Architecture we will only discuss generic concepts how to design Cluster IP configuration.

Fencing is critical mechanism protecting data from being corrupted. What is fencing and how it works is explained here:

- [Split-brain, Quorum, and Fencing](#)
- [ClusterLabs: What is Fencing?](#)

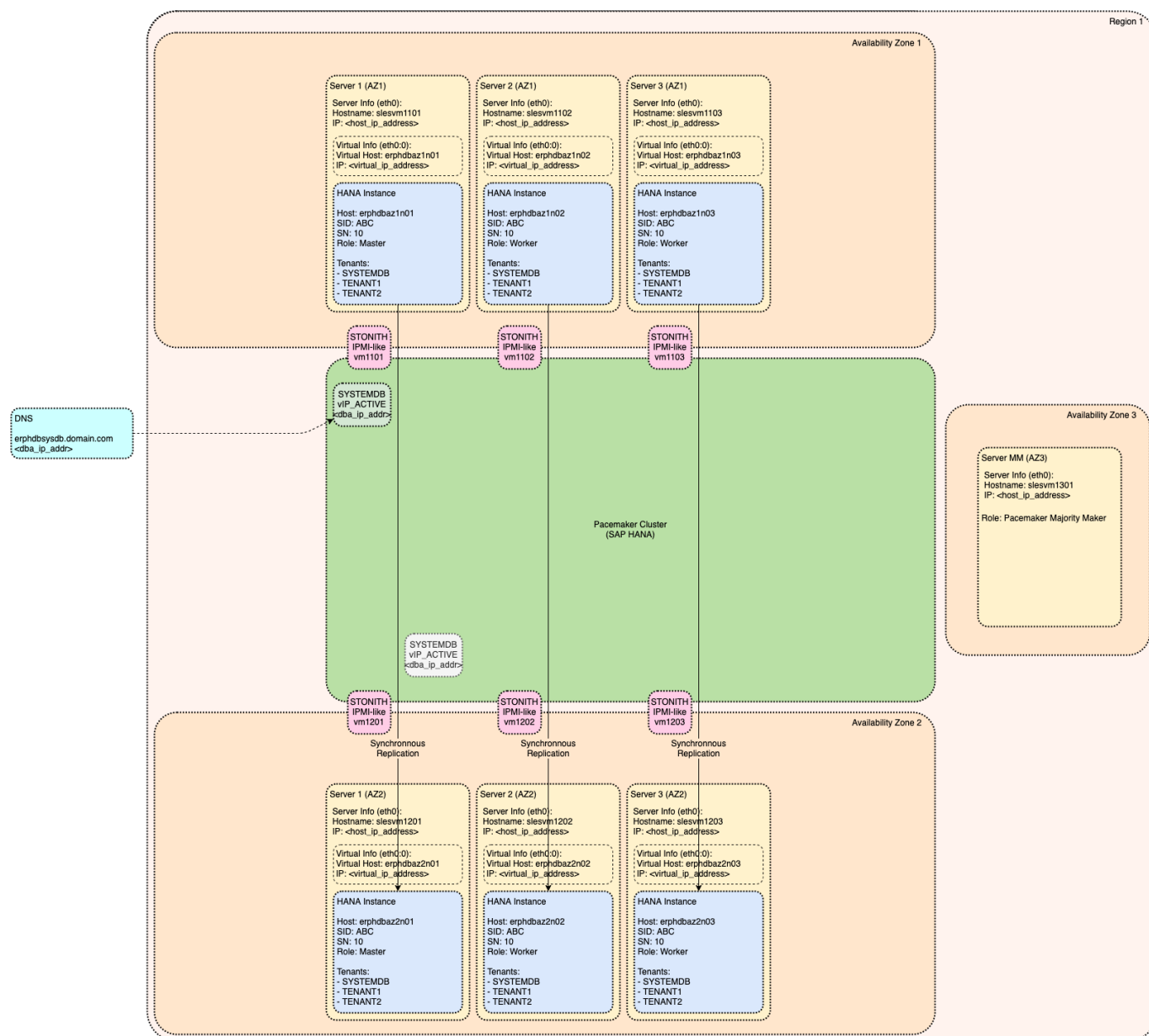
Sections below are explaining two basic options how to implement Fencing Mechanism:

- IPMI-like Fencing
- SBD (Storage Based Death) Fencing

Recommendation which option to use on each platform is described in Platform Specific Architecture part of the documentation.

## IPMI-like Fencing





IPMI-like fencing approach is based on direct access to Management Interface of given server which is also called IPMI ([Intelligent Platform Management Interface](#)) and which is having ability to power down the given server.

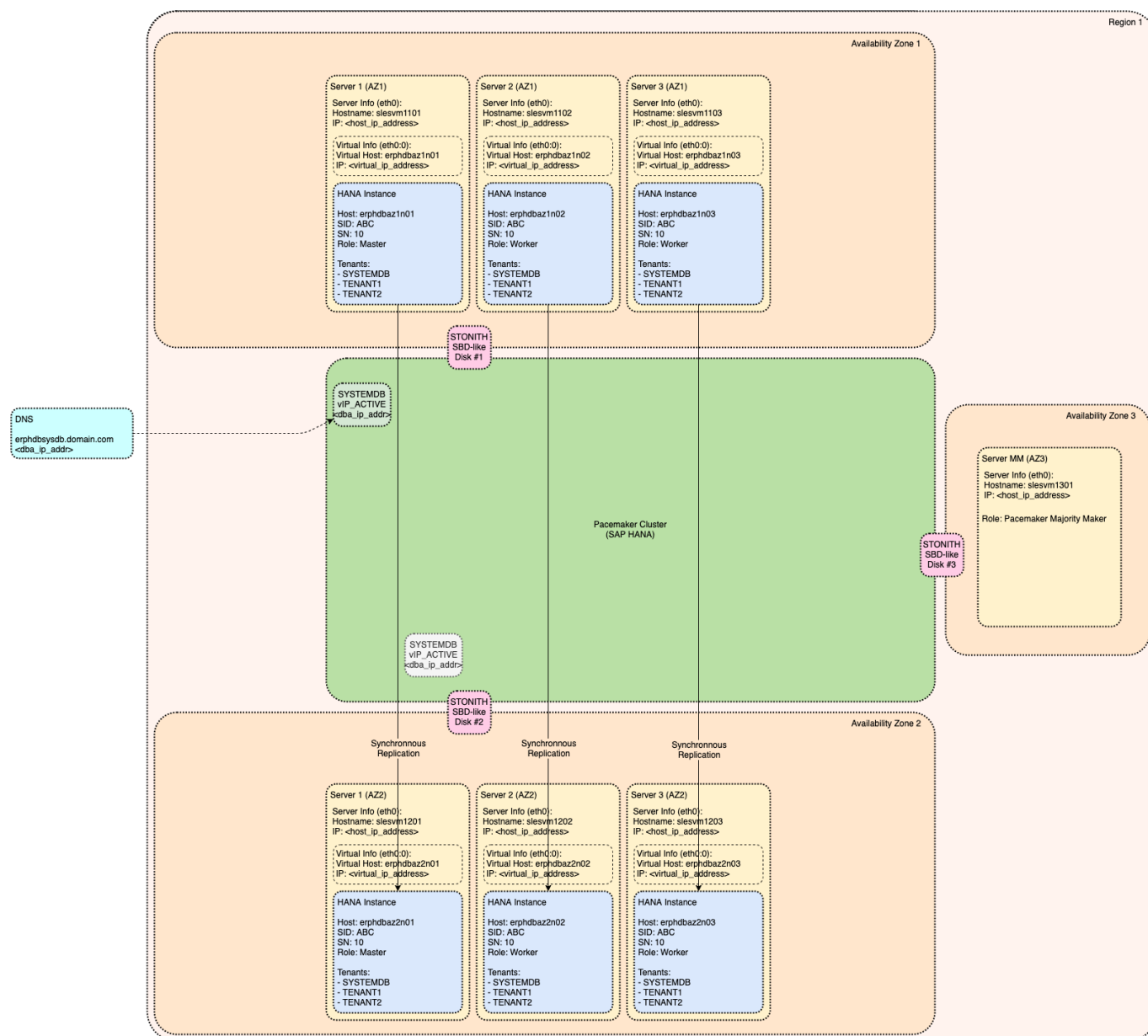
Here are some example implementations of IPMI-like agents:

- Amazon Web Services (AWS): [external/ec2](#)
- Microsoft Azure: [fence\\_azure\\_arm](#)
- Google Cloud Platform (GCP): [external/gcpstonith](#)
- On-premise Bare Metal: [external/ipmi](#)

All of these agents are having same purpose - to kill the Virtual Machine or Bare Metal Server as soon as technically possible. The goal is not to perform graceful shutdown but to immediately terminate the server to ensure that it is down before secondary server can takeover.

In cluster configuration each SAP HANA server needs to have its own IPMI-like fencing agent configured and fully operational. The IPMI-like fencing agent is always called from the remote side (for example secondary systems are fencing off primary systems).

## SBD (Storage Based Death) Fencing



SBD (Storage Based Death) fencing is based on different approach.

The SBD device is shared raw disk (can be connected via Fibre Channel, Fibre Channel over Ethernet or iSCSI) that is used to send messages to other nodes.

When SBD device is initiated it will overwrite the beginning of the disk device with messaging slot structure. This structure is used by individual nodes to send messages to other nodes. Each cluster node is running SBD daemon that is watching the slot dedicated to given node and performing associated actions.

In case of fencing event the node that is triggering the fencing will write the "poison pill" message to the slot associated with target system. SBD daemon on target system is monitoring given slot and once it will read the "poison pill" message it will execute suicide action (self-fencing itself from the cluster by instant powering off).

Here is implementation of SBD agent:

- external/sbd

Recommended amount of SBD devices is either:

- Three SBD devices - each in separate Availability Zone (visualized on picture above)
- One SBD device - in 3rd Availability Zone (other than used by SAP HANA VMs)

Additional Information:

- [Linux-HA: SBD Fencing](#)
- [sbd - STONITH Block Device daemon](#)

## Majority Maker Node

As said above Fencing concept is vital to protect the data from corruption. However, in case of communication issues between cluster nodes there is risk that individual nodes will continuously keep fencing each other by taking over the role of primary system.

Therefore, there is second equally important concept called Quorum that is deciding which subgroup of cluster nodes is entitled to become primary.

## SAP HANA Scale-Out Scenario

Most simple implementation of Quorum is to use odd number of nodes and base the Quorum logic on majority of nodes in subgroup.

This method is applicable to SAP HANA Scale-Out configurations where we have same number of nodes on each Availability Zone and we need one additional VM in 3rd Availability Zone to act as "Majority Maker" helping to decide which side will become primary.

Recommended settings are described in [SLES12 SP2: SAP HANA System Replication Scale-Out - Cluster Bootstrap and more](#).

## SAP HANA Single-Node Scenario

Concept of using "Majority Maker" is also applicable to SAP HANA Single-Node implementations (two node clusters). However, such small clusters can be also implemented using special `two_node` Quorum approach as described in [SLES12 SP4: High Availability Extension - Corosync Configuration for Two-Node Clusters](#):

```
quorum {
  provider: corosync_votequorum
  expected_votes: 2
  two_node: 1
}
```

Note that setting `two_node: 1` value will implicitly configure `wait_for_all: 1`.

The configuration is explained in `votequorum` man pages [Man Pages: VOTEQUORUM\(5\)](#) and in following document [New quorum features in Corosync 2](#).

Effectively the configuration setting is adjusting how Quorum is calculated. When cluster is started for the first time (both nodes down) then `wait_for_all: 1` parameter is ensuring that both nodes need to be available to achieve the quorum. This is critical to protect the consistency of data as explained in following blogs:

- [Be Prepared for Using Pacemaker Cluster for SAP HANA – Part 1: Basics](#)
- [Be Prepared for Using Pacemaker Cluster for SAP HANA – Part 2: Failure of Both Nodes](#)

However, if the cluster is already active (and Quorum was achieved) then parameter `two_node: 1` will ensure that in case that one node will fail the other node is still having Quorum even if it is not having majority.

In case of split-brain situation (when both nodes are active however unable to communicate), both nodes will have Quorum and both nodes will race to fence the other node. The node that will win the race will be primary.

This option is applicable only to Single-Node scenario where we have one SAP HANA System in each Availability Zone. In such case "Majority Maker" VM in 3rd Availability Zone is not required (although possible).

## Cluster IP Design

The simplest SAP HANA High Availability scenario needs only one Cluster IP that is following Active Nameserver of primary SAP HANA system (which is where System Database is available).

Each tenant in SAP HANA system is having its own ports that can be used to directly connect to given Tenant Database. Although this direct connection is possible it is recommended to connect indirectly by specifying the port for the System Database (`3xx13` for ODBC/JDBC/SQLDBC access) and the Tenant Database name. SAP HANA system will ensure that connection is internally rerouted to target Tenant Database.

In case of takeover event the Pacemaker cluster will ensure that Cluster IP is moved to new primary SAP HANA system.

As explained above the technical implementation of Cluster IP is in detail covered in Platform Specific Architecture part of the documentation.

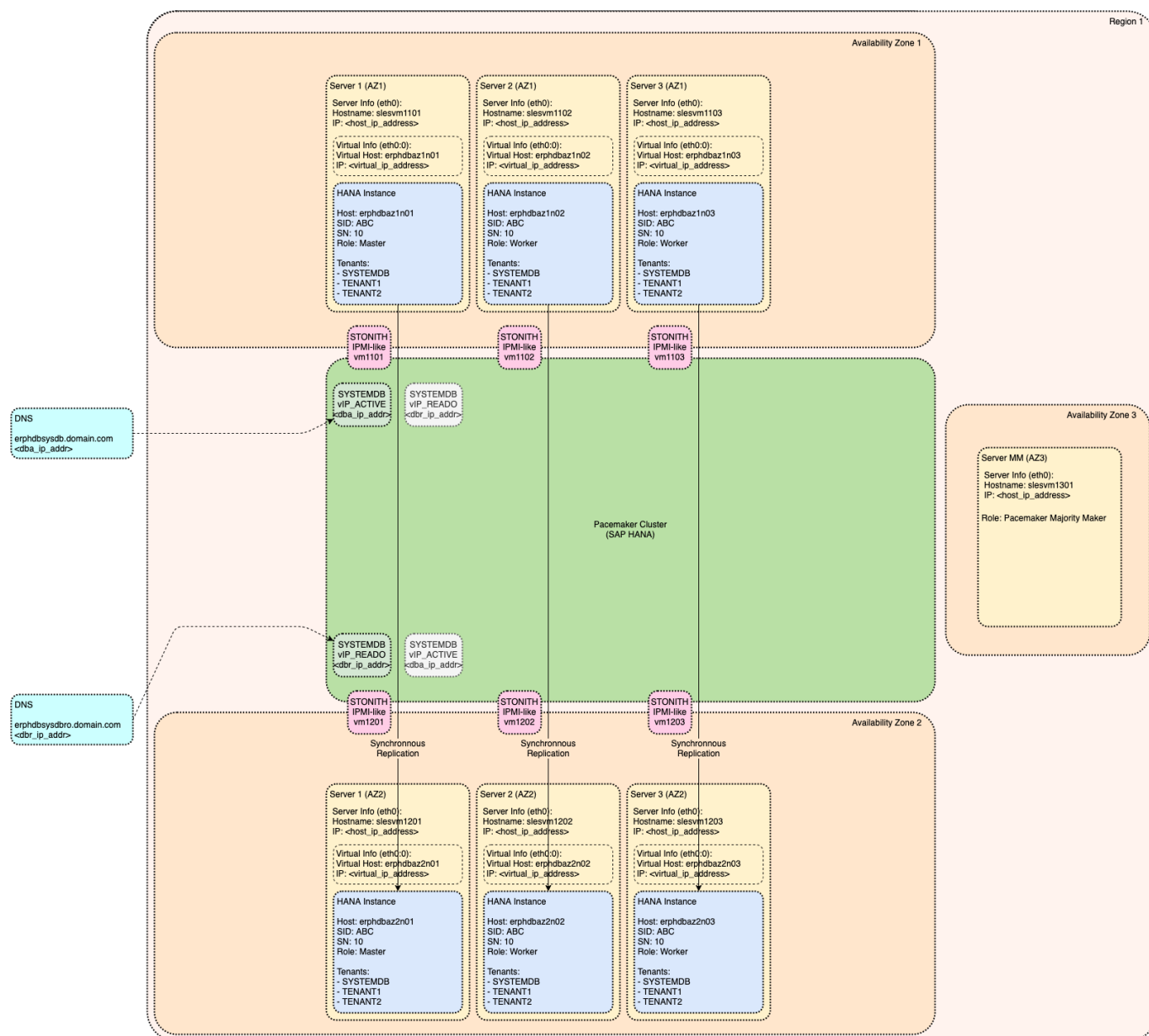
Additional Information:

- [Administration Guide: Server Components of the SAP HANA Database](#)
- [Tenant Databases: Connections for Tenant Databases](#)
- [Tenant Databases: Scale-Out Architecture of Tenant Databases](#)
- [Administration Guide: Connections from Database Clients and Web Clients to SAP HANA](#)
- [Administration Guide: Connections for Distributed SAP HANA Systems](#)
- [TCP/IP Ports of All SAP Products](#)

## Typical Cluster IP Implementation

Traditional implementation of Cluster IP is based on ARP cache invalidation. On primary server Pacemaker Cluster will define Cluster IP address by using command `ip addr add` combined with ARP cache invalidation via `arping` (see [ClusterLabs / resource-agents / heartbeat / IPaddr2](#)). During the takeover the Pacemaker Cluster will remove Cluster IP address from old primary server by using command `ip addr del` and it will recreate it on new primary server using commands mentioned above. The key requirement here is that both primary and secondary server are in same subnet so that Cluster IP address can be moved between them.

## Active/Active High Availability with Pacemaker Cluster



Historically secondary SAP HANA System was closed, and connection attempts were rejected (this is still valid for Operation Modes **delta\_datashipping** or **logreplay**).

Since SAP HANA 2.0 new operation mode **logreplay\_readaccess** is available which is offering capability to open secondary SAP HANA System for read-only access.

As explained in SAP HANA [Administration Guide: Connection Types](#) and in [Administration Guide: Virtual IP Address Handling](#) secondary Cluster IP following Active Nameserver of secondary SAP HANA system is required.

During normal operation both Cluster IP addresses are anti-located to each other - primary Cluster IP address is following primary SAP HANA System and secondary Cluster IP address is following secondary SAP HANA System.

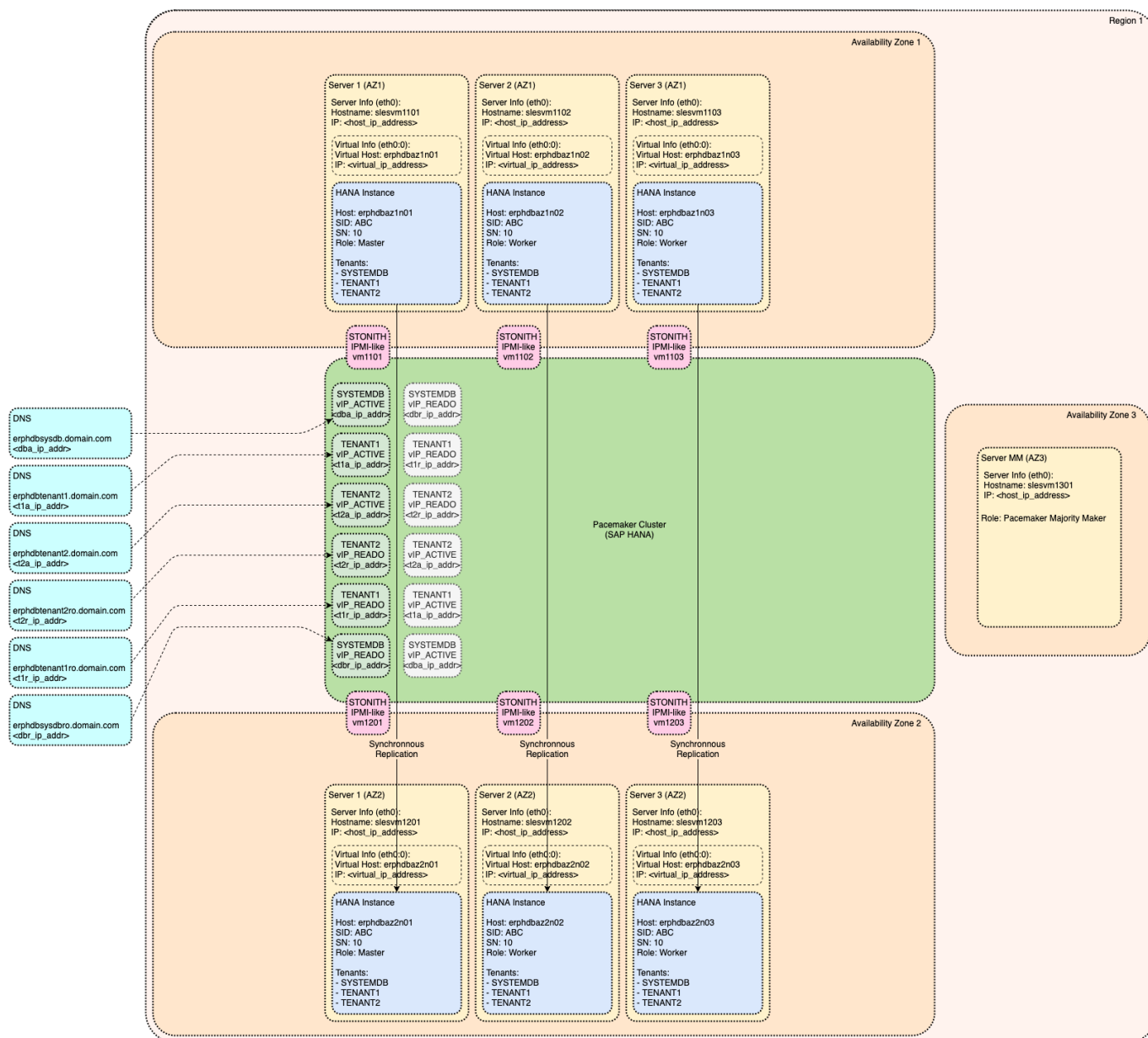
As part of takeover event the Pacemaker cluster will switch location of both IP addresses along with change of primary and secondary roles of SAP HANA Systems.

Additional Information:

- [Administration Guide: Active/Active \(Read Enabled\)](#)

- Administration Guide: Connection Types
- Administration Guide: Virtual IP Address Handling

## Active/Active High Availability with Pacemaker Cluster (enabled for Tenant Move)



SAP HANA is offering option to move Tenant Database from existing SAP HANA System to new SAP HANA System having different **SID** and **system\_number**.

Architecture documented in previous section is having one big limitation related to Tenant Move operation. The design is supporting multiple Tenant Databases on one SAP HANA cluster however, all tenants are accessed over one shared Cluster IP.

In such configuration when Tenant Database is moved, all applications connecting to this Tenant Database must be reconfigured to use Cluster IP of target SAP HANA cluster.

To make Tenant Move operation as seamless as possible each tenant needs to have its own Cluster IP that will be moved to target SAP HANA cluster along with given tenant.



All tenant-specific Cluster IPs are implemented in same way as System Database Cluster IP, they are following Active Nameserver of primary SAP HANA system - which is where System Database, used to connect to individual tenants, is available.

Second challenge that needs to be addressed is port used for connecting to System Database (3xx13 for ODBC/JDBC/SQLDBC access). This port is dependent on `system_number` of given SAP HANA System and therefore can differ. Solution to this problem is to allocate additional port (same across all SAP HANA Systems) on which System Database Tenant will listen. The procedure is described in [Administration Guide: Configure Host-Independent Tenant Addresses](#).

The procedure how to relocate Tenant Database to new SAP HANA System is described in [SAP HANA Tenant Move](#).

Additional Information:

- [Administration Guide: Copying and Moving Tenant Databases](#)
- [Administration Guide: Configure Host-Independent Tenant Addresses](#)

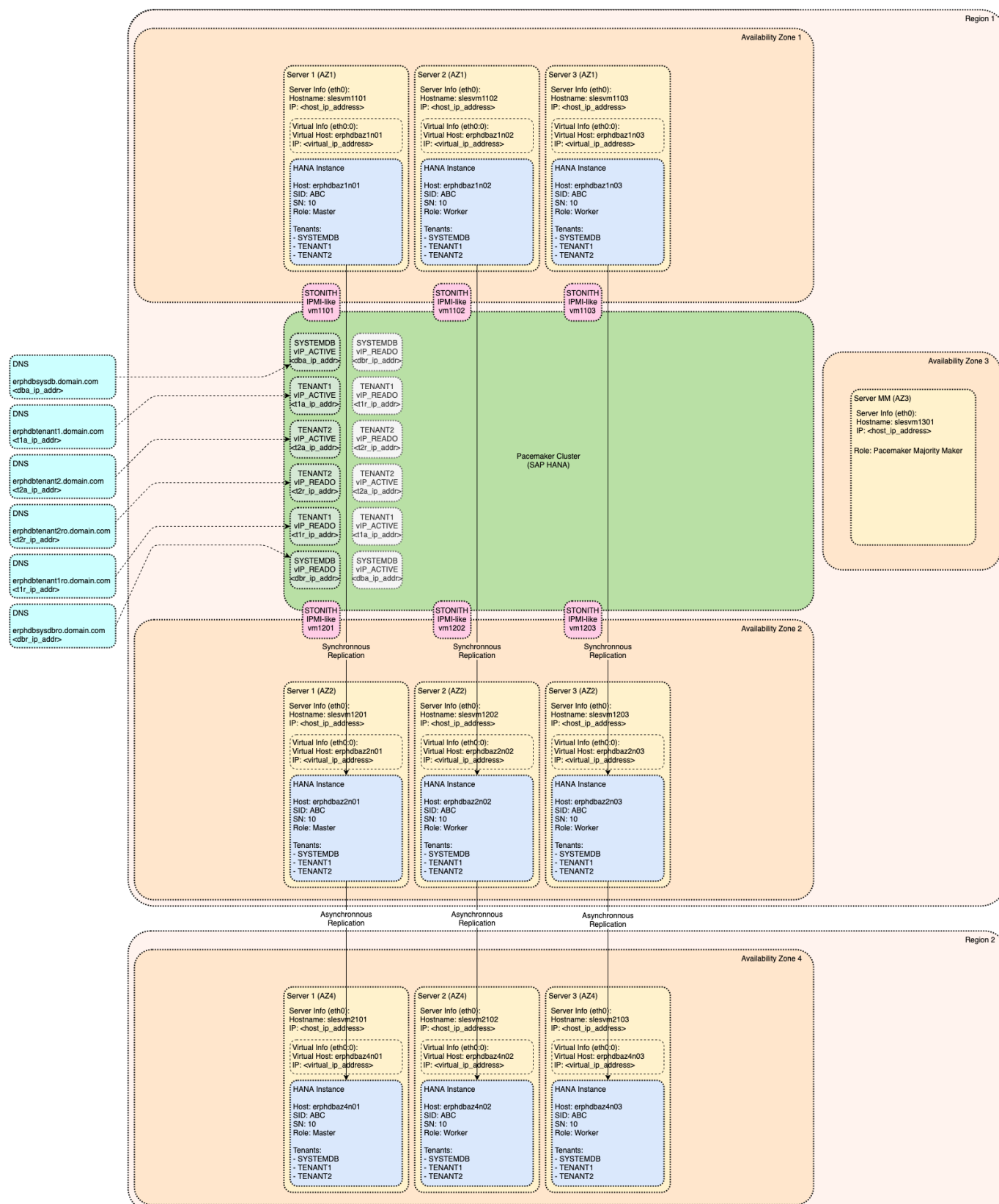
## Module: Disaster Recovery

---

Disaster Recovery module is enhancing the design by adding protection against Regional disaster. Such disaster events can be caused by nature (floods, tornadoes, hurricanes or earthquakes), by humans (strikes, terrorism) or by technology (power blackout).

- [Module: Disaster Recovery](#)
  - [Disaster Recovery Concept](#)

### Disaster Recovery Concept



All events mentioned above are typically having wide area of impact. Since High Availability design is based on synchronous replication between two Availability Zones of same Region, there are chances that both primary and secondary system will be impacted at the same time by Regional disaster. Therefore, there is need to replicate the data outside the impacted Region.

SAP HANA Asynchronous (ASYNC) Replication is recommended approach how to ship the data to remote Disaster Recovery location (independent Region). The advantage is that Asynchronous Replication Mode is not susceptible to increased latency because the replication is happening on background.



As explained in [System Replication Guide: Replication Performance Problems](#) the network bandwidth is still critical even for Asynchronous Replication Mode.

Number of active nodes of the target SAP HANA System in the Disaster Recovery location must be same as on source system (see [Administration Guide: General Prerequisites for Configuring SAP HANA System Replication](#) for additional information).

During Disaster Recovery Event both SAP Application Servers and SAP HANA Database is being failed over to Disaster Recovery location. All applications that were subject of failover are reconfigured and tested as part of failover procedure. For all external applications that need connectivity to SAP HANA System it is strongly recommended to connect via DNS, so that application connectivity can be restored by adjusting single DNS entry.

Additional Information:

- [System Replication Guide: SAP HANA System Replication](#)
- [Administration Guide: SAP HANA System Replication](#)
- [Administration Guide: SAP HANA Multitier System Replication](#)
- [Administration Guide: SAP HANA Multitarget System Replication](#)
- [Administration Guide: Disaster Recovery Scenarios for Multitarget System Replication](#)
- [How To SAP HANA System Replication Whitepaper](#)
- [SAP Note 2407186: How-To Guides & Whitepapers For SAP HANA High Availability](#)

## Module: Data Tiering Options

---

SAP is offering range of capabilities how to optimize costs by segregating data into different storage and processing tiers. This module is discussing how individual Data Tiering options can be implemented as part of this Reference Architecture.

- [Module: Data Tiering Options](#)
  - [Overview of Data Tiering Options for SAP HANA](#)
  - [Persistent Memory \(NVRAM\)](#)
  - [SAP HANA Native Storage Extensions \(NSE\)](#)
  - [SAP HANA Extension Nodes](#)
  - [SAP HANA Dynamic Tiering \(DT\)](#)

## Overview of Data Tiering Options for SAP HANA

SAP is dividing the data based on the aging characteristics of the data and frequency of usage. Following data temperature tiers and tiering options are available:

Data Tiering Option	Native HANA	BW on HANA BW/4HANA	Suite on HANA S/4HANA
<b>Hot Data</b>			
- Dynamic Random Access Memory (DRAM)	Yes	Yes	Yes
- <a href="#">Persistent Memory (NVRAM)</a>	Yes	Yes	Yes

Data Tiering Option	Native HANA	BW on HANA BW/4HANA	Suite on HANA S/4HANA
<b>Warm Data</b>			
- <a href="#">SAP HANA Native Storage Extensions (NSE)</a>	Yes	Yes	Yes
- <a href="#">SAP HANA Extension Nodes</a>	Yes	Yes	-
- <a href="#">SAP HANA Dynamic Tiering (DT)</a>	Yes	-	-
<b>Cold Data</b>			
- SAP Near Line Storage			
- SAP Data Hub / SAP Data Intelligence			
- SAP HANA Spark Controller (Hadoop)			

Selected Data Tiering Options (those impacting SAP HANA Reference Architecture) are discussed in sections below.

Additional Information:

- [Administration Guide: Data Tiering](#)

## Persistent Memory (NVRAM)

SAP HANA in-memory data can be divided into following usage types:

- Main Data fragments (the in-memory copy of tables; infrequent changes)
- Delta Data fragments (update information; frequent changes)
- Temporary Data fragments (computational data; very frequent changes)

Server memory must be combination of Traditional RAM (**DRAM**) and Persistent Memory (Non-Volatile Random Access Memory - **NVRAM**). Traditional RAM (**DRAM**) is required during Operating System start and is also offering better performance for write operations. On the other hand, Persistent RAM (**NVRAM**) is cheaper and bigger and almost as fast as **DRAM** for read operations.

Therefore, Persistent Memory is intended only for Main Data fragments of Column Store tables that are changed very infrequently (only during Delta Merge operation).

Persistent Memory is supported since SAP HANA 2.0 SP03 (revision 35 and higher) and SAP HANA 2.0 SP04.

Usage of Persistent Memory can be activated on the level of whole SAP HANA Database, selected Tables, selected Table Partitions or only selected Table Columns.

Since Persistent Memory DIMMs are having much bigger capacity compared to Traditional **DRAM** DIMMs, SAP HANA System with Persistent Memory will have increased overall RAM capacity able to host higher amount of data.

Although replication between SAP HANA System having Persistent Memory and SAP HANA System without Persistent Memory is supported, it is not recommended for High Availability purpose. In any case proper memory sizing must be ensured to avoid out-of-memory situations after takeover.

Additional Information:

- [Administration Guide: The Delta Merge Operation](#)
- [Administration Guide: Persistent Memory](#)
- [SAP Note 2618154: SAP HANA Persistent Memory - Release Information](#)
- [SAP Note 2700084: FAQ: SAP HANA Persistent Memory](#)

## SAP HANA Native Storage Extensions (NSE)

SAP HANA is offering native option how to manage less frequently accessed data using built-in warm data store capability called Native Storage Extensions (NSE).

Data management of hot data is well described in [Administration Guide: Memory Management in the Column Store](#).

Hot data is normally stored in "In-Memory Column Store". SAP HANA is automatically loading complete data structures (Table Columns or Table Column Partitions) into the memory based on first usage and will keep all data in memory as long as possible. These data structures are unloaded from memory only in case that allocated memory has reached the maximum limit and memory is required for processing another workload. In such case least recently used data structures are unloaded first.

SAP HANA Native Storage Extensions functionality is offering different approach based on "Disk Based Column Store". It can be activated for selected database objects (tables, partitions or columns). The data structures (Table Columns or Table Column Partitions) are kept on disk and only selected data pages are loaded into memory into "Buffer Cache". This concept is well-known from all other classical databases.

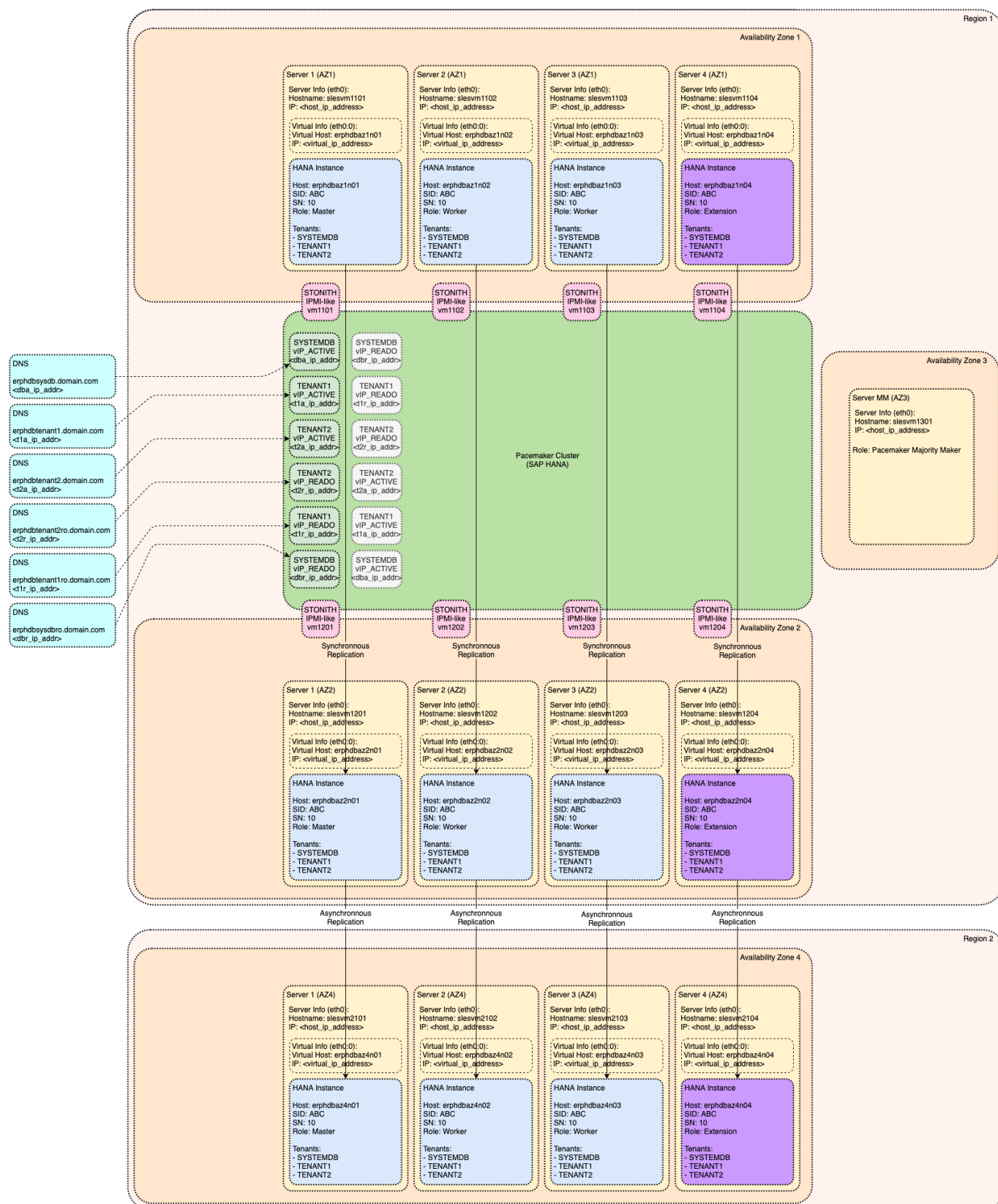
By using "Buffer Cache" that is significantly smaller than size of data in "Disk Based Column Store", Native Storage Extensions capability is increasing maximum amount of data that can be stored in SAP HANA database. Therefore, total storage requirements are also increased which needs to be reflected by infrastructure.

SAP HANA Native Storage Extension feature is supported since SAP HANA 2.0 SP04 and is limited only to Single-node SAP HANA Systems. Please note other functional restrictions as mentioned in [SAP Note 2771956: SAP HANA Native Storage Extension Functional Restrictions](#).

Additional Information:

- [Administration Guide: SAP HANA Native Storage Extension](#)
- [SAP HANA Native Storage Extension Whitepaper](#)
- [SAP Note 2799997: FAQ: SAP HANA Native Storage Extension \(NSE\)](#)

## SAP HANA Extension Nodes



SAP HANA Scale-Out Systems can leverage SAP HANA Extension Nodes capability - new type of SAP HANA node type used exclusively for warm data.

SAP HANA Extension Node (configured as a slave node, worker group value `worker_dt`) is storing the warm data in "In-Memory Column Store" like regular SAP HANA node used for hot data. Since the warm data is less frequently used, the performance for `SELECT` statements against this data is not seen as important. Therefore, we can overload this node with amount of data to be doubled or in some cases even quadrupled compared to regular SAP HANA nodes.

Because the data is stored in "In-Memory Column Store" the internal mechanics is same as described in [Administration Guide: Memory Management in the Column Store](#) and in previous section. Due to a high volume of data in given node, the data structures are loaded and unloaded much more often than hot data in other nodes. Since this is associated with performance degradation, the Extension Node must be dedicated only for warm data.

Warm data must be placed in separate Tables or in separate Table Partitions. Subsequently those Tables and Table Partitions are relocated to SAP HANA Extension Node(s). Note that as described in [Module: Basic Architecture](#) each SAP HANA node is having its own subset of data in its own data files.

SAP HANA Extension Nodes are supported since SAP HANA 1.0 SP12 (for BW scenario) and since SAP HANA 2.0 SP03 (native scenario) and are limited only to Scale-Out SAP HANA Systems.

For SAP BW scenarios the hardware used for SAP HANA Extension Nodes can differ compared to other worker nodes starting from SAP HANA 2.0 SP03. For native scenarios this is supported from SAP HANA 2.0 SP4.

Additional Information:

- [Administration Guide: Extension Node](#)
- [Administration Guide: Redistributing Tables in a Scaleout SAP HANA System](#)
- [More Details – HANA Extension Nodes for BW-on-HANA](#)
- [SAP Note 2644438: SAP HANA Extension Node – Master Release Note](#)

## SAP HANA Dynamic Tiering (DT)

SAP HANA Dynamic Tiering (DT) is optional add-on for SAP HANA database to manage warm data. Behind the scenes it is SAP IQ database that was modified and integrated into SAP HANA to act as new type of SAP HANA database process **esserver**. Since SAP IQ is based on "Disk Based Column Store", SAP HANA Dynamic Tiering database process (**esserver**) is offering very similar capabilities like SAP HANA in-memory database process (**indexserver**) with reduced memory requirements and reduced performance.

Although SAP HANA Dynamic Tiering is based on different database, the integration effort is ensuring that Dynamic Tiering node is embedded into SAP HANA operational processes like Start/Stop, Backup/Recovery and System Replication making it sub-component of SAP HANA database. Therefore, SAP HANA Dynamic Tiering cannot be operated independently from SAP HANA.

SAP HANA Dynamic Tiering database process (**esserver**) is typically deployed on dedicated host as part of SAP HANA Scale-Out scenario, however, there is option to co-deploy it on same host as SAP HANA in-memory database process (**indexserver**) for Single-Node scenario.

However, there are still several limitations that need to be taken into consideration when deploying SAP HANA Dynamic Tiering. These limitations include:

- Since SAP IQ is not having concept of Multitenant Database Containers (MDC), each SAP HANA Tenant Database using Dynamic Tiering needs its own dedicated SAP HANA Dynamic Tiering database process (**esserver**)
- However, for Single-Node deployment scenario only, just one Dynamic Tiering database process (**esserver**) can be co-deployed on same host as SAP HANA in-memory database process (**indexserver**). Additional Dynamic Tiering database processes (associated with additional Tenant

Databases) must be deployed on dedicated hosts - this is effectively turning the architecture into Scale-Out like (multi-host) deployment

- SAP HANA Dynamic Tiering component itself is not "Scale-Out enabled", that means one SAP HANA Tenant Database cannot distribute data across multiple Dynamic Tiering database processes (**esserver**) - the Dynamic Tiering host must be sized properly to be able to support complete volume of the warm data for given Database Tenant
- Be sure to also review following restrictions when using Dynamic Tiering:
  - [Administration Guide: Extended Store Table Functional Restrictions](#)
  - [Administration Guide: Multistore Table Functional Restriction](#)
  - [Installation and Update Guide: SAP HANA Landscape Functional Restrictions](#)
  - PDF document attached to [SAP Note 2767107: SAP HANA Dynamic Tiering Support for SAP HANA System Replication](#)
  - [SAP Note 2375865: SAP HANA Dynamic Tiering 2.0: Backup and Recovery Functional Restriction](#)

Basic deployment options are described in [Installation and Update Guide: SAP HANA Dynamic Tiering Architecture](#).

Until SAP HANA Dynamic Tiering 2.0 SP04 there was no support for Cluster Manager - therefore High Availability using Pacemaker Cluster was not supported. Because this is recently released feature, this version of SAP HANA Reference Architecture is not yet supporting SAP HANA Dynamic Tiering. The support will be included in future versions based on detailed examination of functional restrictions mentioned above.

In order to move the data to Dynamic Tiering node SAP introduced concept of Extended Store Tables and Multistore Tables:

- [Administration Guide: Extended Store Tables](#) are tables that fully reside in Dynamic Tiering "Disk Based Column Store" (**esserver**)
- [Administration Guide: Multistore Tables](#) are tables that can have some partitions on SAP HANA "In-Memory Column Store" (**indexserver**) and other partitions on Dynamic Tiering "Disk Based Column Store" (**esserver**)

These new database table types are well explained and visualized in following blog: [A Closer Look at SAP HANA Dynamic Tiering for Warm Data Management](#).

SAP HANA Dynamic Tiering is available only for use cases documented in [Master Guide: Dynamic Tiering Use Cases](#). SAP HANA Dynamic Tiering is NOT supported for following applications:

- SAP BW/4HANA ([SAP Note 2517460](#))
- SAP S/4HANA ([SAP Note 2462641](#))
- SAP Business Suite on HANA ([SAP Note 2462641](#))
- SAP BW on HANA ([SAP Note 2462641](#))

Additional Information:

- [SAP HANA Dynamic Tiering Landing Page](#)

## Module: SAP XSA

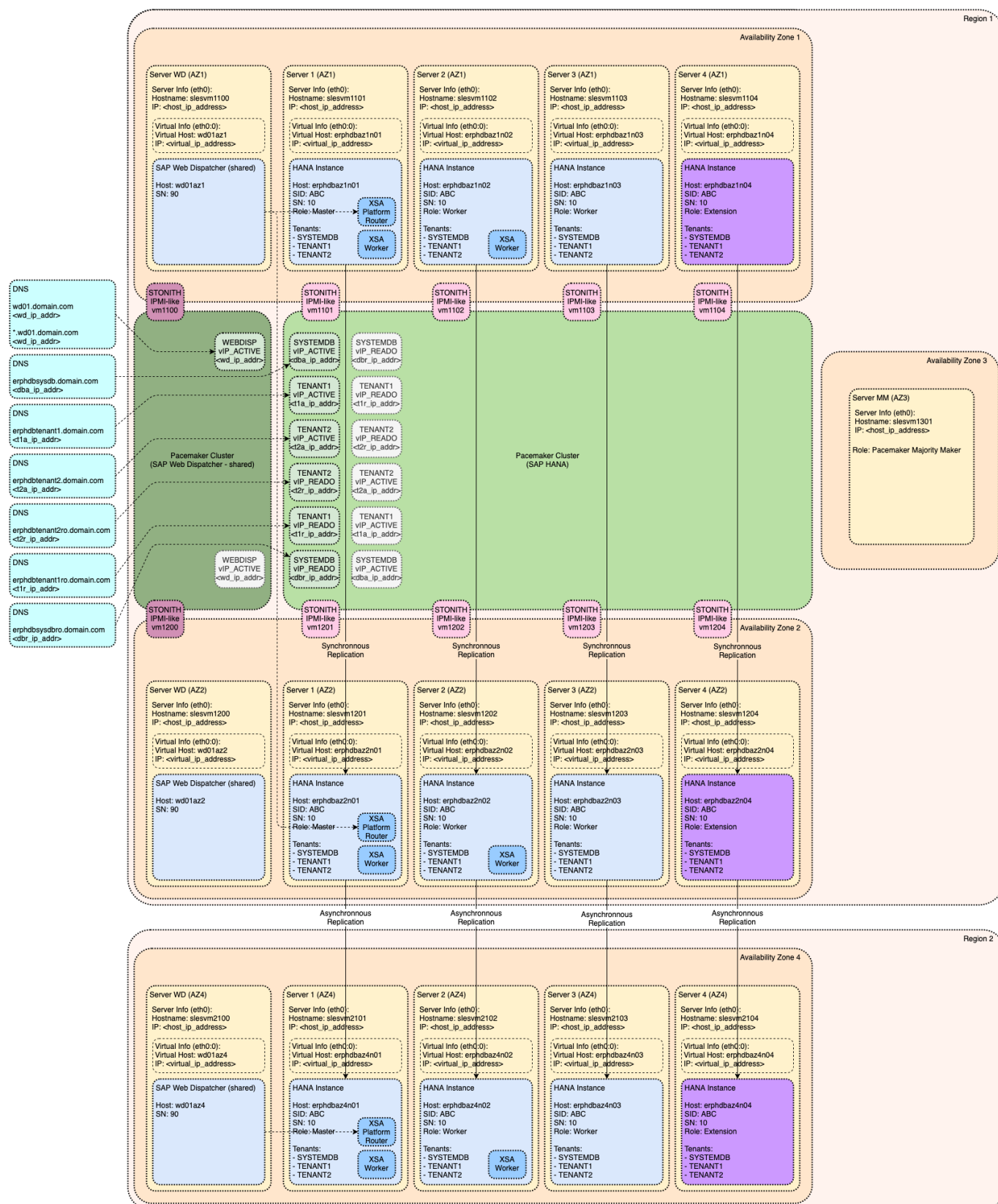
---



SAP HANA extended application services, advanced model (XSA) is application server platform used for the development and execution of native data-intensive applications. This module is explaining how to implement SAP XSA as part of this Reference Architecture.

- [Module: SAP XSA](#)
  - [Reference Architecture of SAP XSA](#)
    - [Additional Host Roles](#)
    - [Tenant Database Installation](#)
    - [Routing Mode](#)
    - [Usage of Dedicated SAP Web Dispatcher](#)

## Reference Architecture of SAP XSA



## Additional Host Roles

Architecture of SAP XSA is introducing new host roles:

- **xs\_worker** for active SAP XSA host running XSA applications
- **xs\_standby** not used as part of this Reference Architecture - it is used only in case of SAP Host Auto-Failover High Availability option (as documented in section [Alternative Implementations](#))



These roles can be assigned automatically (each SAP HANA database **worker** host is assigned SAP XSA **xs\_worker** role) or manually (only selected hosts are assigned **xs\_worker** role or there can even be dedicated **xs\_worker** host).

Additional Information:

- [Installation and Update Guide: System Concepts for XS Advanced Runtime Installations](#)

## Tenant Database Installation

As of SAP HANA 2.0 SP03 (revision 34) there is option to choose between installation of XSA Technical data into System Database Tenant or into separate Tenant Database.

XSA Platform is recognizing following data types:

- XSA Technical Data (one set of data shared for all Customer XSA Applications)
  - Users (including users for Customer XSA Applications)
  - Platform Data (applications, build packs, runtime information, etc.)
  - System Application Data (audit log service, deploy service, etc.)
- XSA Custom Application Data (data for Customer XSA Applications)

XSA Custom Application Data is always stored in Tenant Database. Additional Tenant Databases can be configured to separate data belonging to different XSA Custom Applications.

XSA Technical Data can be stored either in System Database Tenant (default option) or in Tenant Database along with XSA Customer Application Data (available since SAP HANA 2.0 SP03 (revision 34)).

Since installing XSA Technical Data in System Database Tenant is making Tenant Move operation impossible, it is recommended to install XSA Technical Data into Tenant Database wherever possible. However, please note the technical restrictions as described in [Installation and Update Guide: Installing XS Advanced in a Tenant Database](#).

Additional Information:

- [Installation and Update Guide: XS Advanced Database Setup Options](#)
- [Installation and Update Guide: Installing XS Advanced in the System Database](#)
- [Installation and Update Guide: Installing XS Advanced in a Tenant Database](#)

## Routing Mode

SAP XSA is offering option to use either "port based routing mode" or "hostname based routing mode". This Reference Architecture is based on using "hostname based routing mode" which is recommended for productive use.

Additional Information:

- [Installation and Update Guide: Setting Up the XS Advanced Runtime Behind a Reverse Proxy](#)
- [SAP Note 2245631: Routing Mode and Default Domain configuration for SAP HANA extended application services, advanced model](#)

## Usage of Dedicated SAP Web Dispatcher

XSA Platform Router is central entry point for all XSA applications. In case of SAP HANA Scale-Out system, the XSA component will automatically start and maintain its own internal SAP Web Dispatcher process that will forward the request to other hosts. This behavior is described in [Administration Guide: Multi-Host Setup with XS Advanced](#)

Note that XSA Platform Router (**xscontroller** service) can run on different host, other than host where Active Nameserver service (associated with Cluster IP) is running. So additional Cluster IP dedicated for XSA Platform Router would be required to support High Availability scenario.

Different approach is usage of Failover Router (external SAP Web Dispatcher) acting as Reverse Proxy dispatching the XSA requests to XSA Platform Router running on currently active SAP HANA System.

This Reference Architecture is based on external Highly Available SAP Web Dispatcher protected by its own Pacemaker Cluster. This SAP Web Dispatcher can be shared for multiple SAP HANA systems and since it will act as "Single Point of Entry" it will make "Tenant Move" operation easier to execute.

Additional Information:

- [Installation and Update Guide: Setting Up the XS Advanced Runtime Behind a Reverse Proxy](#)
- [SAP Note 2300936: Host Auto-Failover & System Replication Setup with SAP HANA extended application services, advanced model](#)

## Alternative Implementations

---

This section describes alternative implementations of SAP HANA System. These are not seen to be part of this Reference Architecture and are described only for informational purpose.

- [Alternative Implementations](#)
  - [SAP HANA Host Auto-Failover \(in single Availability Zone\)](#)

### SAP HANA Host Auto-Failover (in single Availability Zone)

As discussed in [AD1: High Availability Concept](#) there are two options how to implement SAP HANA High Availability. This Reference Architecture is based on SAP HANA Synchronous System Replication option. This section is discussing SAP HANA Host Auto-Failover (HAF) option which is not used in this Reference Architecture.

SAP HANA Host Auto-Failover is native function of SAP HANA System and is in detail described in [SAP HANA Host Auto-Failover Whitepaper](#).

SAP HANA Host Auto-Failover High Availability option is based on adding one or more dedicated **standby** hosts, that are passively waiting for failure of one or more of the active hosts. When such failure will happen, SAP HANA instance running on this **standby** host will takeover the data files and log files of failed host, thus replacing the failed host in its function.

This option is having both advantages and disadvantages when compared to SAP HANA Synchronous System Replication option.

Advantages:

- Lower Costs - SAP HANA Host Auto-Failover is adding 1-3 extra hosts dedicated for High Availability ( $n+m$  approach), while SAP HANA System Replication needs same number of hosts on both sides ( $n+n$  approach)
- Native Feature - SAP HANA Host Auto-Failover is native function of SAP HANA System - no external cluster software is required (therefore no additional knowledge is required)

#### Disadvantages:

- Single Failure Protection - One **standby** host is able to protect only against failure of one active host, maximum three **standby** hosts can be deployed (with cost implications)
- No Availability Zone Support - SAP HANA Host Auto-Failover architecture is unable to support multiple Availability Zones as all active hosts must be available simultaneously
- High Takeover Times - Since **standby** host cannot predict which of active hosts will fail, no pre-loading was implemented - therefore, takeover times can easily take dozens of minutes for very large systems
- No Support for Active/Active - since SAP HANA Host Auto-Failover does not have secondary system, thus there is no Active/Active capability (see [Administration Guide: Active/Active \(Read Enabled\)](#) for additional information)
- No Support for Secondary Time Travel - since SAP HANA Host Auto-Failover does not have secondary system, thus there is no Secondary Time Travel capability (see [Administration Guide: Secondary Time Travel](#) for additional information)
- No Support for Near Zero Downtime Upgrades - since SAP HANA Host Auto-Failover does not have secondary system, thus there is no Near Zero Downtime Upgrades capability (see [Administration Guide: Use SAP HANA System Replication for Near Zero Downtime Upgrades](#))
- Not Supported by all IaaS Cloud Vendors - as explained in [Administration Guide: Multiple-Host System Concepts](#) SAP HANA Host Auto-Failover must to be deployed either using Shared File Systems (like NFS) or based on Storage Connector API, therefore, not all IaaS Cloud Vendors are supporting this High Availability option

#### Additional Information:

- [Administration Guide: Host Auto-Failover](#)
- [Administration Guide: Multiple-Host System Concepts](#)
- [SAP HANA Host Auto-Failover Whitepaper](#)

## Platform Specific Architecture for AWS (Amazon Web Services)

---

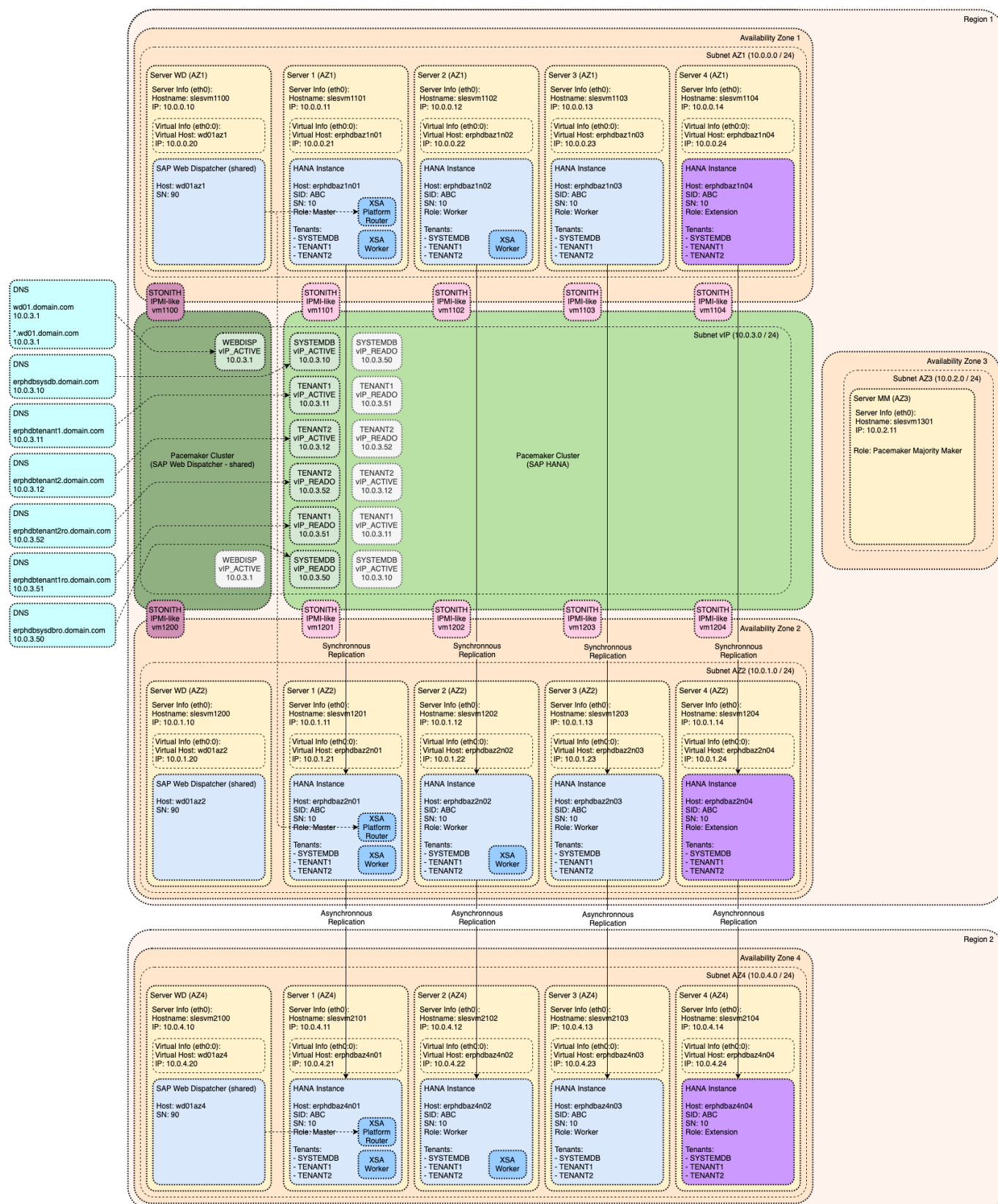
- [Platform Specific Architecture for AWS \(Amazon Web Services\)](#)
  - [AWS: Overall Architecture](#)
  - [AWS: Basic Architecture](#)
    - [AWS: Supported Instance Types for SAP HANA](#)
    - [AWS: Storage Setup for SAP HANA Implementation](#)
    - [AWS: Networking specifics for AWS Availability Zones](#)
  - [AWS: Virtual Hostname/IP](#)
    - [AWS: Generic implementation steps](#)
    - [AWS: Additional comments](#)
      - [AWS: SAP HANA inbound network communication](#)

- AWS: SAP HANA outbound network communication
- AWS: High Availability
  - AWS: High Availability across Availability Zones
    - AWS: Which Availability Zone to use
  - AWS: Available Fencing mechanism
  - AWS: Implementation of Cluster IP
- AWS: Disaster Recovery
- AWS: Data Tiering Options
  - AWS: Persistent Memory (NVRAM)
  - AWS: SAP HANA Native Storage Extensions (NSE)
  - AWS: SAP HANA Extension Nodes
  - AWS: SAP HANA Dynamic Tiering (DT)
- AWS: XSA

## AWS: Overall Architecture

Following diagram is illustrating complete version of Reference Architecture for SAP HANA tailored for AWS (Amazon Web Services).

For detailed explanation of individual modules please see individual sections in [Generic SAP HANA Architecture](#).



You can also review official AWS Reference Architecture and other documentation:

- [AWS: SAP on AWS Technical Documentation: SAP HANA](#)
- [AWS: SAP HANA on the AWS Cloud: Quick Start Reference Deployment](#)
- [AWS: Multi-AZ \(HA\), Single-Node Architecture](#)

## AWS: Basic Architecture

Link to generic content: [Module: Basic Architecture](#)

## AWS: Supported Instance Types for SAP HANA

Not every instance type is supported for productive SAP HANA usage.

Official list of SAP certified instance types is available at [SAP: The SAP HANA Hardware Directory](#). This should always be used to decide whether the particular instance type is supported for SAP HANA or not.

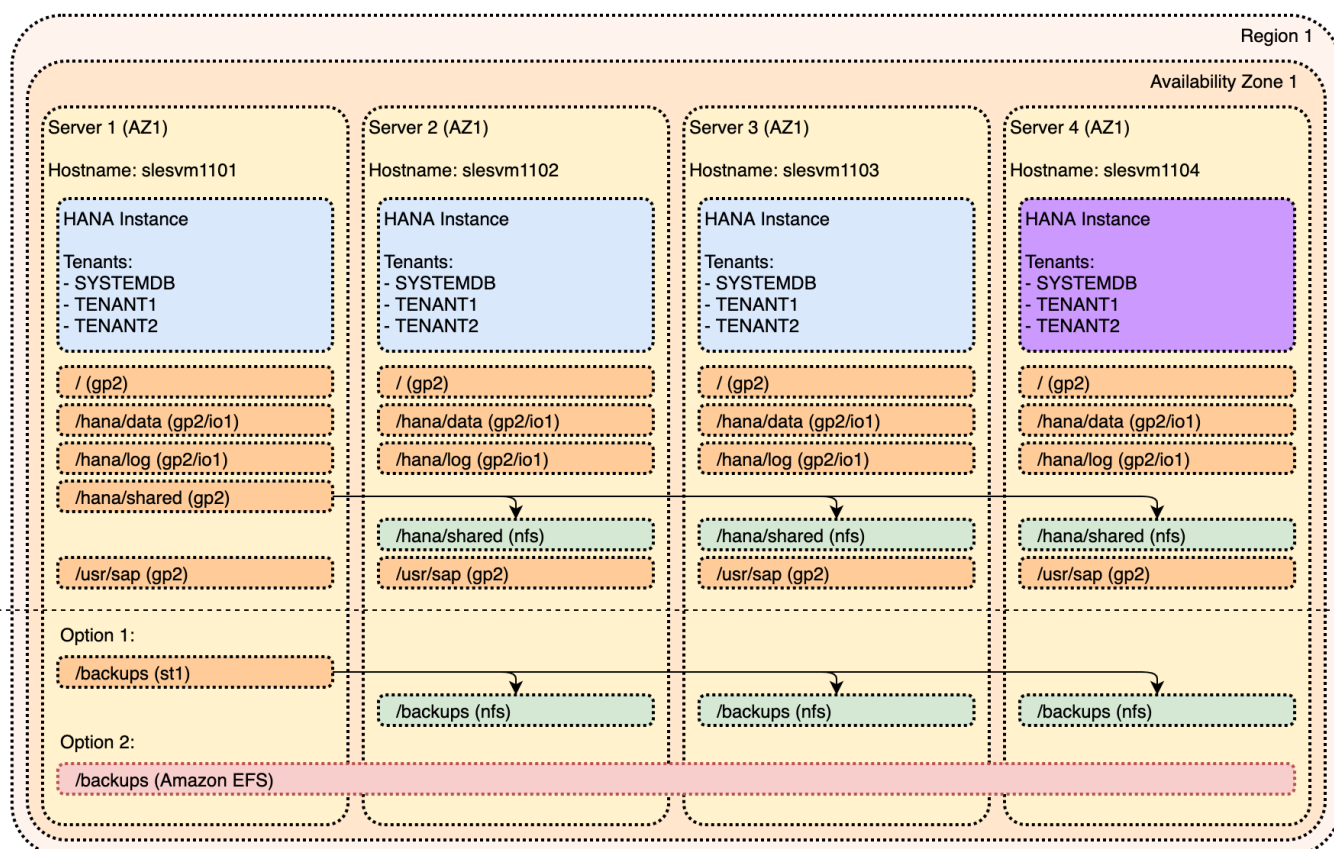
AWS specific list of certified instances with additional details can be found in [AWS: SAP HANA on AWS Quick Start Guide: AWS Instance Types for SAP HANA](#)

## AWS: Storage Setup for SAP HANA Implementation

SAP HANA Storage Configuration is coming in two flavors:

- General Purpose SSD (**gp2**) storage - cheaper storage that meets SAP KPI requirements for most of the SAP HANA workloads
- Provisioned IOPS SSD (**io1**) storage - high-performance storage intended for most demanding SAP HANA workloads

Following disk setup is recommended:



Filesystem	Name	Device type	Comment
/	Root volume	gp2	
/hana/data	SAP HANA data	gp2 / io1	
/hana/log	SAP HANA logs	gp2 / io1	



Filesystem	Name	Device type	Comment
/hana/shared	SAP HANA shared	gp2	Provisioned to the master, NFS-mounted on other nodes
/usr/sap	SAP binaries	gp2	
/backups	SAP HANA backup	st1	Provisioned to the master, NFS-mounted on other nodes
/backups	SAP HANA backup	Amazon EFS	Alternative option for SAP HANA backup filesystem

Instance specific sizing recommendations are available at [AWS: SAP HANA on AWS Quick Start Guide: Storage Configuration for SAP HANA](#).

## AWS: Networking specifics for AWS Availability Zones

As visualized on the Overall Architecture diagram - in AWS each Availability Zone is having its own subnet. It is not possible to stretch one subnet across multiple Availability Zones. This needs to be taken into consideration during deployment planning.

Impact on implementation of Cluster IP in AWS is described in [AWS: High Availability](#).

## AWS: Virtual Hostname/IP

Link to generic content: [Module: Virtual Hostname/IP](#)

This chapter describes recommended implementation of Virtual Hostname and Virtual IP for an SAP HANA installation on the AWS cloud.

The implementation is based on assigning a *Secondary private IP address* to an existing network interface of the AWS instance. A *Secondary private IP address* can be easily reassigned to another AWS instance and so it can follow SAP HANA instance relocation. For more details see [AWS: Multiple IP Addresses](#). This *Secondary private IP address* is associated with the Virtual Hostname which is used during SAP HANA instance installation.

## AWS: Generic implementation steps

- define Virtual IP (in the same subnet as the network interface) and Virtual Hostname for SAP HANA Instance
- assign *Virtual IP* as *Secondary private IP address* to existing network interface (see [AWS: assign secondary private IP](#))
- configure OS to use *Virtual IP Address* (e.g. [SUSE: Administration Guide - Configuring Multiple Addresses](#))
- make Virtual Hostname resolvable (e.g. updating `/etc/hosts`)
- install SAP HANA instance with the *Virtual Hostname* (see [SAP: Administration Guide - Default Host Names and Virtual Host Names](#))

**Note:** *Virtual IP* can be reassigned to another AWS instance thanks to option "*Allow reassignment*" of the network interface (see [AWS: assign a secondary private IPv4 address](#))

AWS: Additional comments

### **AWS: SAP HANA inbound network communication**

A network communication initiated from a remote system to the Virtual IP of SAP HANA instance is established and takes place between remote system IP and the Virtual IP (*Secondary private IP address*). The *Primary private IP address* on the same interface is not involved in this communication.

### **AWS: SAP HANA outbound network communication**

In contrast to an inbound connection, when SAP HANA instance initiates a network connection to a remote system the *Primary private IP address* is used as source IP instead of Virtual IP (*Secondary private IP address*). If there is requirement to use Virtual IP as the source IP, it could be achieved by means of linux routing. The core of the idea is to add route specifying source address like `ip route add <network/mask> via <gateway> src <virtual_ip>` (see [Routing for multiple uplinks/providers](#)).

## **AWS: High Availability**

Link to generic content: [Module: High Availability](#)

### **AWS: High Availability across Availability Zones**

Best practice for deploying SAP HANA is to stretch High Availability cluster across Availability Zones. Each AWS Availability Zone is physically separate infrastructure, therefore deploying High Availability across Availability Zones ensures that there is no shared single-point-of-failure (SPOF) between primary and secondary SAP HANA system. This approach is significantly increasing overall resiliency of the High Availability of the solution.

List of existing Availability Zones for individual AWS Regions is available here: [AWS: Regions and Availability Zones](#).

### **AWS: Which Availability Zone to use**

Most critical factor for selecting Availability Zones is network latency. Latency between individual Availability Zones can significantly differ and therefore it is important to measure network latency using SAP `niping` tool (see [SAP Note 500235: Network Diagnosis with NIPING](#) for additional information) and select Availability Zones with minimal latency.

Furthermore, it is important to note that internal numbering of Availability Zones is specific for each individual AWS account. Therefore, the network latency test must be performed in given account. For additional information please see [AWS: Regions, Availability Zones, and Local Zones - Availability Zones](#).

Last thing to consider is whether desired instance types are available inside selected Availability Zones as not all Availability Zones are offering all instance types.

### **AWS: Available Fencing mechanism**



AWS is using IPMI-like Fencing (see [Module: High Availability - IPMI-like Fencing](#) for additional details). SBD (Storage Based Death) Fencing is not available in AWS.

Fencing agent source code is available here: [external/ec2](#). Behind the scenes it is using [StopInstances](#) API call with [force](#) option. This will hard stop EC2 instance without gracefully stopping the Operating System.

Following are most important prerequisites for [external/ec2](#) fencing mechanism to work properly:

- EC2 instances need to be properly tagged - tags are used by fencing agent to find correct instance to stop (see [SUSE: Tagging the EC2 Instances](#))
- Configure AWS CLI - fencing agent is dependent on AWS CLI which needs to be configured properly (see [SUSE: Creating an AWS CLI Profile on Both EC2 Instances](#))
- Configure HTTP Proxies (or use NAT gateway) to access AWS APIs over internet (see [SUSE: Configure HTTP Proxies](#))
- Create STONITH policy and attach it as instance role (see [SUSE: STONITH Policy](#))

Please see [SUSE: SAP HANA High Availability Cluster for the AWS Cloud - Setup Guide - Prerequisites for the AWS-Specific HA Installation](#) for complete list of prerequisites.

Additional Information:

- [SUSE: SAP HANA High Availability Cluster for the AWS Cloud - Setup Guide](#)
- [Red Hat: Configure SAP HANA System Replication in Pacemaker on Amazon Web Services](#)

## AWS: Implementation of Cluster IP

Traditional implementation of Cluster IP (not applicable to AWS) is covered in section [Module: High Availability - Typical Cluster IP Implementation](#).

In AWS each Availability Zone is having its own separate subnet and it is not possible to stretch one subnet across multiple Availability Zones. Therefore, different mechanism is required. AWS implementation of Cluster IP address is based on "Overlay IP" address concept. It is defined as additional entry in VPC routing table entry that is managed directly by Pacemaker Cluster. This entry is forwarding all packets sent to Overlay IP (third IP address in its own separate subnet) to the IP address of either primary or secondary server. During cluster takeover this VPC routing table entry is updated by Pacemaker cluster via AWS Command Line Interface (CLI) utility to point to new primary server. This concept is more in detail explained here:

- [AWS: SAP on AWS High Availability with Overlay IP Address Routing](#)
- [IP Failover with Overlay IP Addresses](#)

In order to ensure that Pacemaker cluster is able to properly manage the VPC route table adjustments, it needs proper IAM access to be assigned to given VM. Technical details are explained in following documentation:

- [SUSE: SAP HANA High Availability Cluster for the AWS Cloud - Setup Guide](#)
- [Red Hat: Configure SAP HANA System Replication in Pacemaker on Amazon Web Services](#)

Third requirement for this concept to work is to disable "Source/Destination Check" for given Network Interface as explained in above mentioned guides. This is to ensure that packets are not discarded based on using Cluster IP address. For additional information please see [AWS: Disabling source/destination checks](#)).

## AWS: Disaster Recovery

Link to generic content: [Module: Disaster Recovery](#)

Disaster Recovery for SAP HANA via SAP HANA System Replication is not infrastructure dependent.

## AWS: Data Tiering Options

Link to generic content: [Module: Data Tiering Options](#)

Following data tiering options are supposed on AWS:

Data Tiering Option	Supported
<a href="#">Persistent Memory (NVRAM)</a>	No
<a href="#">SAP HANA Native Storage Extensions (NSE)</a>	Yes
<a href="#">SAP HANA Extension Nodes</a>	Yes
<a href="#">SAP HANA Dynamic Tiering (DT)</a>	Yes

Additional Information:

- [AWS: SAP Data Tiering](#)

### AWS: Persistent Memory (NVRAM)

Amazon Web Services platform does not offer any instance types having NVRAM that are supported for productive SAP HANA usage.

### AWS: SAP HANA Native Storage Extensions (NSE)

SAP HANA Native Storage Extensions only need additional disk space compared to regular SAP HANA deployments. Amazon Web Services platform does allow to provision additional disks to SAP HANA VM and add capacity into existing filesystems. There is no change to the design of the storage layout.

### AWS: SAP HANA Extension Nodes

SAP HANA Extension Node implementation is based on provisioning additional SAP HANA node (with increased storage capacity) to existing SAP HANA system. Result is SAP HANA Scale-Out system where one of the nodes is designated as SAP HANA Extension Node. Storage layout is same as for regular SAP HANA nodes and it is visualized above in section [AWS: Storage Setup for SAP HANA Implementation](#).

Additional Information:

- [AWS: Warm Data Tiering Options: SAP HANA Extension Node](#)

### AWS: SAP HANA Dynamic Tiering (DT)

SAP HANA Dynamic Tiering (DT) implementation in Amazon Web Services platform is based on provisioning additional VM for Dynamic Tiering component and connecting it to VM hosting SAP HANA instance, thus forming two-node distributed setup. Storage layout is identical to SAP HANA Scale-out setup as illustrated above in section [AWS: Storage Setup for SAP HANA Implementation](#).

Additional Information:

- [AWS: Warm Data Tiering Options: SAP HANA Dynamic Tiering](#)

## AWS: XSA

Link to generic content: [Module: SAP XSA](#)

SAP HANA extended application services, advanced model (XSA) component is not infrastructure dependent.

# Platform Specific Architecture for Azure (Microsoft Azure)

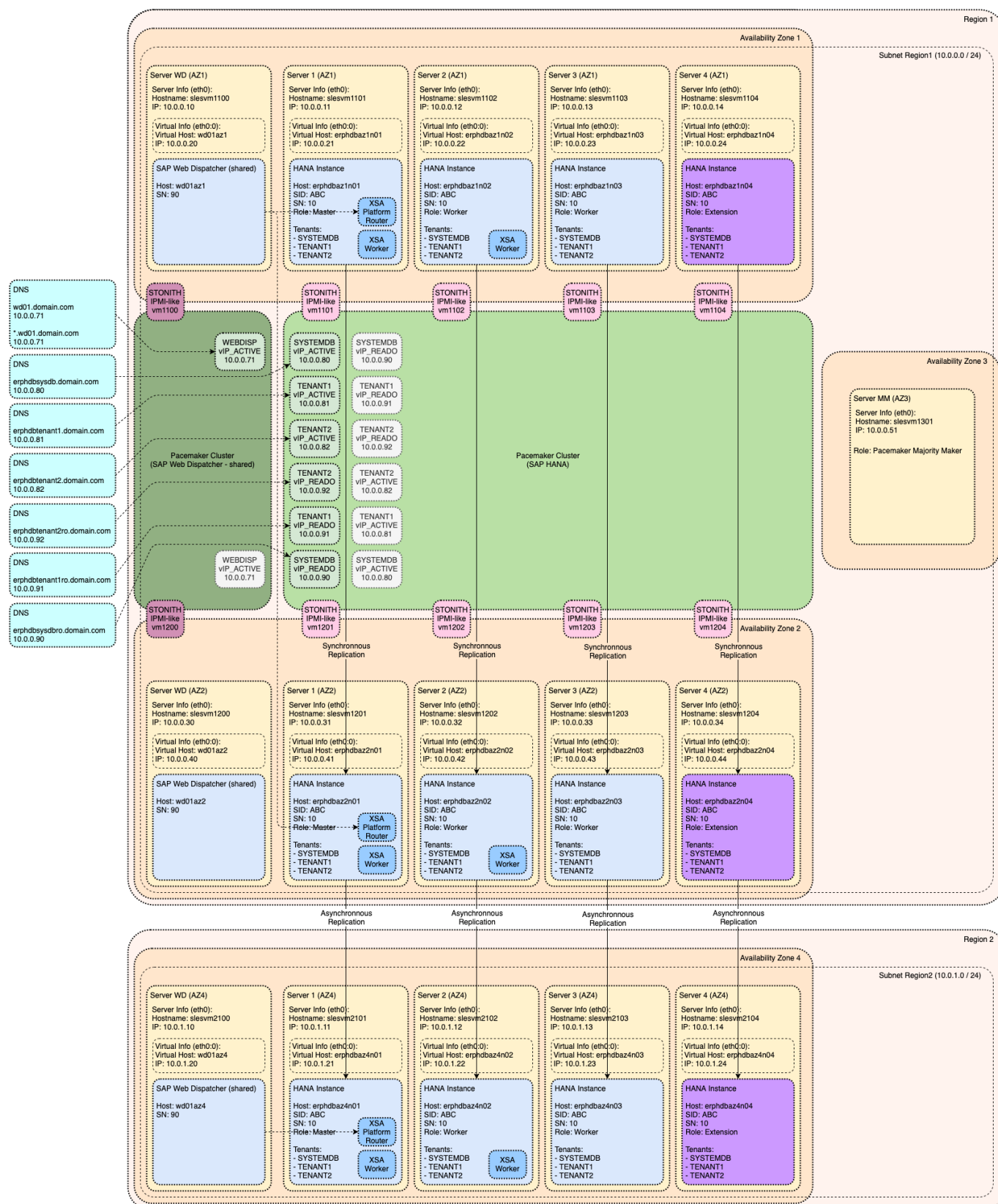
---

- [Platform Specific Architecture for Azure \(Microsoft Azure\)](#)
  - [Azure: Overall Architecture](#)
  - [Azure: Basic Architecture](#)
    - [Azure: Supported Instance Types for SAP HANA](#)
    - [Azure: Storage Setup for SAP HANA Implementation](#)
      - [Azure: Storage Setup for SAP HANA Implementation - Azure Premium SSD](#)
      - [Azure: Storage Setup for SAP HANA Implementation - Azure Ultra disk](#)
      - [Azure: Storage Setup for SAP HANA Implementation - Azure NetApp Files](#)
    - [Azure: Networking specifics for Azure Availability Zones](#)
  - [Azure: Virtual Hostname/IP](#)
    - [Azure: Generic implementation steps](#)
    - [Azure: Additional comments](#)
      - [Azure: SAP HANA inbound network communication](#)
      - [Azure: SAP HANA outbound network communication](#)
  - [Azure: High Availability](#)
    - [Azure: High Availability across Availability Zones](#)
      - [Azure: Which Availability Zone to use](#)
    - [Azure: Available Fencing mechanism](#)
    - [Azure: Implementation of Cluster IP](#)
  - [Azure: Disaster Recovery](#)
  - [Azure: Data Tiering Options](#)
    - [Azure: Persistent Memory \(NVRAM\)](#)
    - [Azure: SAP HANA Native Storage Extensions \(NSE\)](#)
    - [Azure: SAP HANA Extension Nodes](#)
    - [Azure: SAP HANA Dynamic Tiering \(DT\)](#)
  - [Azure: XSA](#)

## Azure: Overall Architecture

Following diagram is illustrating complete version of Reference Architecture for SAP HANA tailored for Azure (Microsoft Azure).

For detailed explanation of individual modules please see individual sections in [Generic SAP HANA Architecture](#).



You can also review official Azure Reference Architecture and other documentation:

- [Azure: SAP HANA infrastructure configurations and operations](#)
- [Azure: SAP HANA \(Large Instances\) architecture](#)

## Azure: Basic Architecture

Link to generic content: [Module: Basic Architecture](#)

## Azure: Supported Instance Types for SAP HANA

Not every instance type is supported for productive SAP HANA usage.

Official list of SAP certified instance types is available at [SAP: The SAP HANA Hardware Directory](#). This should always be used to decide whether the particular instance type is supported for SAP HANA or not.

Azure specific list of certified instances with additional details can be found in following references:

- [Azure: SAP HANA certifications](#)
- [Azure: Available SKUs for HANA Large Instances](#)

## Azure: Storage Setup for SAP HANA Implementation

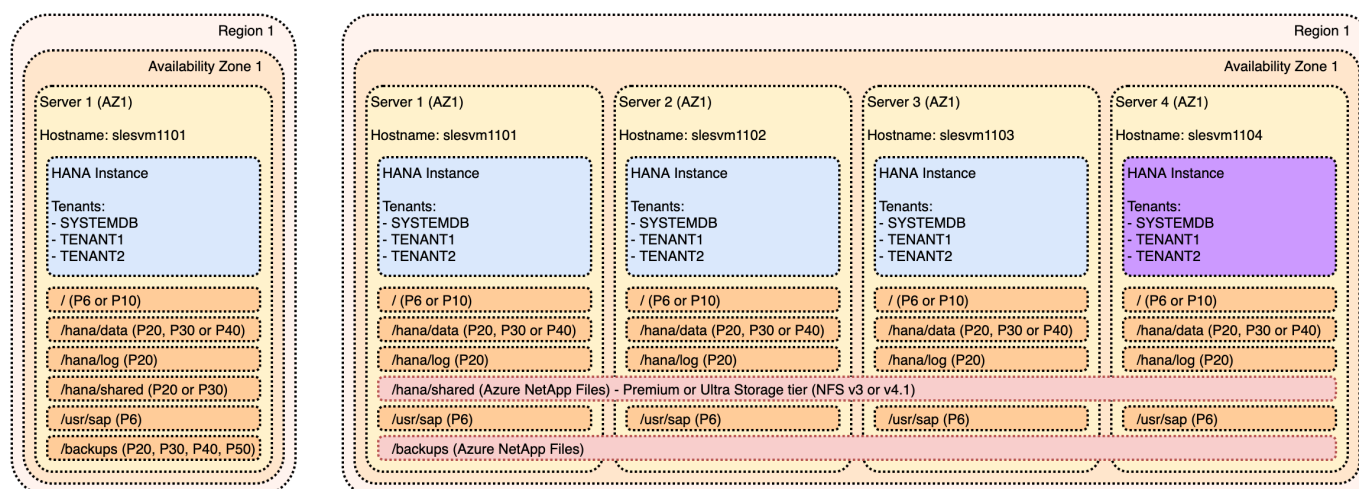
SAP HANA Storage Configuration for Virtual Machines is coming in following flavors:

- [Azure Premium SSD](#) - cheaper storage that meets SAP KPI requirements for most of the SAP HANA workloads
- [Azure Ultra disk](#) - high-performance storage intended for most demanding SAP HANA workloads
- [Azure NetApp Files](#) - special option for [SAP HANA Host Auto-Failover \(in single Availability Zone\)](#)

Recommended disk setup for each option is described in following sections.

SAP HANA Storage Configuration for HANA Large Instances is always based on NFS storage as supported deployment options are covered in [Azure: Supported scenarios for HANA Large Instances](#).

### Azure: Storage Setup for SAP HANA Implementation - Azure Premium SSD



Filesystem	Name	Device type	Comment
/	Root volume	P6 or P10	
/hana/data	SAP HANA data	P20, P30 or P40	
/hana/log	SAP HANA logs	P20	
/hana/shared	SAP HANA shared	P20 or P30	For Single-Node deployments
/hana/shared	SAP HANA shared	Azure NetApp Files	For Scale-Out deployments

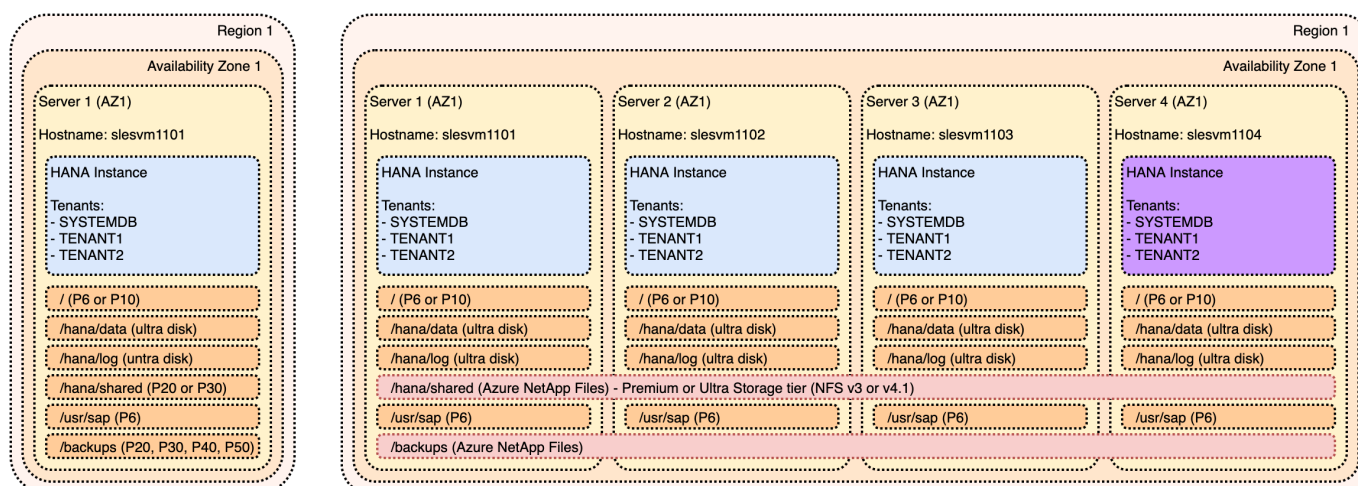


Filesystem	Name	Device type	Comment
/usr/sap	SAP binaries	P6	
/backups	SAP HANA backup	P20, P30, P40, P50	For Single-Node deployments
/backups	SAP HANA backup	Azure NetApp Files	For Scale-Out deployments

Instance specific sizing recommendations are available at [Azure: Azure Premium SSD](#).

Note: Cost conscious storage configuration is available for use, however, it is not officially supported by SAP.

### Azure: Storage Setup for SAP HANA Implementation - Azure Ultra disk

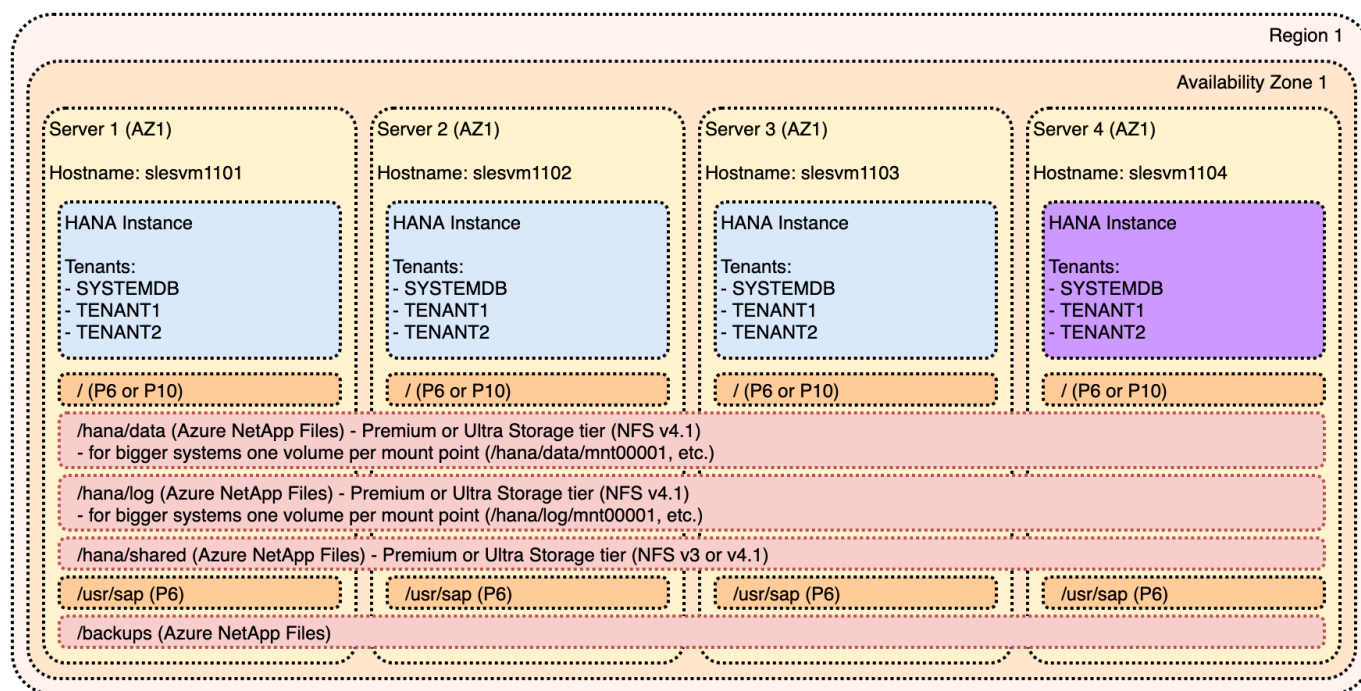


Filesystem	Name	Device type	Comment
/	Root volume	P6 or P10	
/hana/data	SAP HANA data	Ultra disk	Alternatively can be on one volume with logs
/hana/log	SAP HANA logs	Ultra disk	Alternatively can be on one volume with data
/hana/shared	SAP HANA shared	P20 or P30	For Single-Node deployments
/hana/shared	SAP HANA shared	Azure NetApp Files	For Scale-Out deployments
/usr/sap	SAP binaries	P6	
/backups	SAP HANA backup	P20, P30, P40, P50	For Single-Node deployments
/backups	SAP HANA backup	Azure NetApp Files	For Scale-Out deployments

Instance specific sizing recommendations are available at [Azure: Azure Ultra disk](#).

Note: Please be aware about Ultra disk limitations as described in [Azure: Ultra disk - GA scope and limitations](#).

### Azure: Storage Setup for SAP HANA Implementation - Azure NetApp Files



Filesystem	Name	Device type	Comment
/	Root volume	P6 or P10	
/hana/data	SAP HANA data	Azure NetApp Files	For bigger systems one volume per mount point
/hana/log	SAP HANA logs	Azure NetApp Files	For bigger systems one volume per mount point
/hana/shared	SAP HANA shared	Azure NetApp Files	
/usr/sap	SAP binaries	P6	
/backups	SAP HANA backup	Azure NetApp Files	

Instance specific sizing recommendations are available at [Azure: Azure NetApp Files](#).

Note: Please be aware about Azure NetApp Files limitations as described in [Azure: Azure NetApp Files - Important considerations](#).

Key limitation is fact that Azure NetApp Files are not Availability Zone aware and can cause performance degradation when not deployed in close proximity (for example following High Availability takeover).

## Azure: Networking specifics for Azure Availability Zones

As visualized on the Overall Architecture diagram - in Azure both Availability Zones are using one subnet stretched across multiple Availability Zones. This is achieved by Network Virtualization, however, side effect of this approach is that classical [ARP cache](#) invalidations are not working - this is having big impact on implementation of Cluster IP in Azure.

Impact on implementation of Cluster IP in Azure is described in [Azure: High Availability](#).



## Azure: Virtual Hostname/IP

Link to generic content: [Module: Virtual Hostname/IP](#)

This chapter describes recommended implementation of Virtual Hostname and Virtual IP for an SAP HANA installation on the Azure cloud.

The implementation is based on assigning a *Secondary static private IP address* to an existing network interface of the Azure Virtual Machine (VM). A *Secondary static private IP address* can be reassigned to another VM and so it can follow SAP HANA instance relocation. For more details see [Azure: Assign multiple IP addresses](#). This *Secondary static private IP address* is associated with the Virtual Hostname which is used during SAP HANA instance installation.

### Azure: Generic implementation steps

- define Virtual IP (in the same subnet as the network interface) and Virtual Hostname for SAP HANA Instance
- assign *Virtual IP* as *Secondary static private IP address* to existing network interface (see [Azure: Add IP addresses to a VM](#))
- configure OS to use *Virtual IP Address*
  - e.g. [Azure: Add IP addresses to a VM operating system](#)
  - e.g. [SUSE: Administration Guide - Configuring Multiple Addresses](#))
- make Virtual Hostname resolvable (e.g. updating `/etc/hosts`)
- install SAP HANA instance with the *Virtual Hostname* (see [SAP: Administration Guide - Default Host Names and Virtual Host Names](#))

**Note:** to reassign the *Virtual IP* to another VM you need first to remove it on the existing VM and after that to assign it on the new VM (see [Azure: Remove IP addresses](#))

### Azure: Additional comments

#### Azure: SAP HANA inbound network communication

A network communication initiated from a remote system to the Virtual IP of SAP HANA instance is established and takes place between remote system IP and the Virtual IP (*Secondary static private IP address*). The *Primary private IP address* on the same interface is not involved in this communication.

#### Azure: SAP HANA outbound network communication

In contrast to an inbound connection, when SAP HANA instance initiates a network connection to a remote system the *Primary private IP address* is used as source IP instead of Virtual IP (*Secondary static private IP address*).

If there is requirement to use Virtual IP as the source IP, it could be achieved by means of linux routing. The core of the idea is to add route specifying source address like `ip route add <network/mask> via <gateway> src <virtual_ip>` (see [Routing for multiple uplinks/providers](#)).

## Azure: High Availability

Link to generic content: [Module: High Availability](#)

## Azure: High Availability across Availability Zones

Best practice for deploying SAP HANA is to stretch High Availability cluster across Availability Zones. Each Azure Availability Zone is physically separate infrastructure, therefore deploying High Availability across Availability Zones ensures that there is no shared single-point-of-failure (SPOF) between primary and secondary SAP HANA system. This approach is significantly increasing overall resiliency of the High Availability of the solution.

List of existing Availability Zones for individual Azure Regions is available here: [Azure: Azure geographies](#).

### Azure: Which Availability Zone to use

Most critical factor for selecting Availability Zones is network latency. Latency between individual Availability Zones can significantly differ and therefore it is important to measure network latency using SAP [niping](#) tool (see [SAP Note 500235: Network Diagnosis with NIPING](#) for additional information) and select Availability Zones with minimal latency.

Furthermore, it is important to note that internal numbering of Availability Zones is specific for each individual Azure account. Therefore, the network latency test must be performed in given account. For additional information please see [Azure: Regions and Availability Zones in Azure - Availability Zones](#).

Last thing to consider is whether desired VM sizes are available inside selected Availability Zones as not all Availability Zones are offering all VM sizes.

### Azure: Available Fencing mechanism

Azure is using slightly different fencing mechanism based on Linux distribution that is used:

- Red Hat Enterprise Linux (RHEL) is using IPMI-like Fencing (see [Module: High Availability - IPMI-like Fencing](#) for additional details) and SBD (Storage Based Death) Fencing is not available
- SUSE Linux Enterprise Server (SLES) is using combination of IPMI-like Fencing and SBD (Storage Based Death) Fencing

Fencing agent source code is available here: [fence\\_azure\\_arm](#). Behind the scenes it is using Azure SDK for Python call [azure\\_fence.set\\_network\\_state](#). This will hard stop Azure VM without gracefully stopping the Operating System.

Prerequisites for fencing mechanism to work properly are documented here:

- [Azure: Setting up Pacemaker on SUSE Linux Enterprise Server in Azure](#)
- [Azure: Setting up Pacemaker on Red Hat Enterprise Linux in Azure](#)

Additional Information:

- [Azure: High availability of SAP HANA on Azure VMs on SUSE Linux Enterprise Server](#)
- [Azure: High availability of SAP HANA on Azure VMs on Red Hat Enterprise Linux](#)

### Azure: Implementation of Cluster IP

Traditional implementation of Cluster IP (not applicable to Azure) is covered in section [Module: High Availability - Typical Cluster IP Implementation](#).

Azure Software Defined Networking (SDN) is not supporting ARP cache updates, that are required for traditional implementation of Cluster IP. Therefore, although subnets are stretched across Availability Zones, different mechanism is required. Azure implementation of Cluster IP address is based on Azure Network Load Balancer, that is forwarding all packets sent to Load Balancer IP (third IP address) to the IP address of either primary or secondary server, depending on wherever "health probe port" is active.

This "health probe port" is managed by Pacemaker cluster and is implemented as `nc` (RHEL) or `socat` (SLES) command opening dummy port to signal to Load Balancer which VM is currently active (see [ClusterLabs / resource-agents / heartbeat / azure-lb](#)). During cluster takeover the Pacemaker cluster will close the "health probe port" on old primary VM and will open it on new primary VM. This will be detected by Azure Load Balancer, which will start forwarding all traffic to new primary VM.

Additional Information:

- [YouTube: TUT1134: Microsoft Azure and SUSE HAE When Availability Matters](#)
- [PDF: TUT1134: Microsoft Azure and SUSE HAE When Availability Matters](#)

## Azure: Disaster Recovery

Link to generic content: [Module: Disaster Recovery](#)

Disaster Recovery for SAP HANA via SAP HANA System Replication is not infrastructure dependent.

## Azure: Data Tiering Options

Link to generic content: [Module: Data Tiering Options](#)

Following data tiering options are supposed on Azure for traditional Virtual Machines (VMs) and HANA Large Instances (HLI):

Data Tiering Option	Supported (VMs)	Supported (HLI)
<a href="#">Persistent Memory (NVRAM)</a>	No	No
<a href="#">SAP HANA Native Storage Extensions (NSE)</a>	Yes	Yes
<a href="#">SAP HANA Extension Nodes</a>	Yes	Yes
<a href="#">SAP HANA Dynamic Tiering (DT)</a>	Yes	No

### Azure: Persistent Memory (NVRAM)

Microsoft Azure platform does not offer any instance types having NVRAM that are supported for productive SAP HANA usage.

### Azure: SAP HANA Native Storage Extensions (NSE)

SAP HANA Native Storage Extensions only need additional disk space compared to regular SAP HANA deployments. Microsoft Azure platform does allow to provision additional disks to both SAP HANA VMs and HANA Large Instances to provide additional capacity into existing filesystems. There is no change to the design of the storage layout.

## Azure: SAP HANA Extension Nodes

SAP HANA Extension Node implementation is based on provisioning additional SAP HANA node (with increased storage capacity) to existing SAP HANA system. Result is SAP HANA Scale-Out system where one of the nodes is designated as SAP HANA Extension Node. Storage layout is same as for regular SAP HANA nodes and it is visualized above in section [Azure: Storage Setup for SAP HANA Implementation](#).

Additional Information:

- [Azure: HANA Large Instances: Use SAP HANA data tiering and extension nodes](#)

## Azure: SAP HANA Dynamic Tiering (DT)

SAP HANA Dynamic Tiering (DT) implementation in Microsoft Azure platform is based on provisioning additional VM for Dynamic Tiering component and connecting it to VM hosting SAP HANA instance, thus forming two-node distributed setup. Storage layout is identical to SAP HANA Scale-out setup as illustrated above in section [Azure: Storage Setup for SAP HANA Implementation](#).

Additional Information:

- [Azure: SAP HANA Dynamic Tiering 2.0 for Azure virtual machines](#)

## Azure: XSA

Link to generic content: [Module: SAP XSA](#)

SAP HANA extended application services, advanced model (XSA) component is not infrastructure dependent.