

# Sizing / Capacity Modeling for applications on the SAP Business Technology Platform

Detlef Thoms, SAP SE

**SAP**insider  
2023

# What We Will Cover

---

- Sizing / Capacity Modeling in the cloud – why?
- Architecture of a BTP application
- How to: Ensure that the selected infrastructure meets the requirements of the application?
- How to: Application Sizing / Capacity Modeling that fits to the capacity units of the service provider?
- Capacity Unit and Sizing
- Sizing BTP applications
- Excursus: BTP applications in the ABAP environment
- Example: The “Business Partner Validation” application - Sizing approach based on load tests on BTP CF

# Sizing / Capacity Modeling in the cloud – why?



# Sizing / Capacity Modeling in the cloud – why?

## Stakeholder

---

Nowadays, software is often deployed in the cloud. You may ask yourself why sizing in the cloud is required at all. Customers usually subscribe for specific cloud offerings and in general they are not interested to know how many instances they need or how much memory they have to allocate.

**But sizing directly impacts the pricing calculation.**

There are several stakeholders who need to perform a Capacity Modeling:

- SAP is developing and operating an SAP BTP application for its customers
- You are a Partner and are developing and operating an SAP BTP application for your customer
- You are Company and are developing and operating an SAP BTP application for your internal users

# Sizing / Capacity Modeling in the cloud – why?

PaaS - SaaS

*“External” Sizing:  
How much should I buy?*

## PaaS: Develop and Operate a BTP application

**ABAP environment**  
Develop ABAP Cloud apps and extensions, leveraging innovations of SAP HANA.  
**FREE TIER**

**Application Autoscaler**  
Automatically scale your applications to meet their dynamic resource needs.  
**ALWAYS FREE**

**Cloud Foundry Runtime**  
Operate polyglot applications.  
**FREE TIER**

**SAP HANA Cloud**  
Store, process, and federate data in a cloud infrastructure.  
**FREE TIER**

## SaaS: Sell a BTP application

**Supply Chain Planning**  
Strengthen supply chain resilience with SAP.  
Our product:  
[Explore supply chain planning solutions >](#)

**SAP HANA Cloud**  
Modernize your data foundation.  
Our product:  
[SAP HANA Cloud >](#)

**SAP Datasphere**  
SAP Datasphere is a unified service for data integration, cataloging, semantic modeling, data warehousing, and virtualizing workloads across SAP and non-SAP data.  
Our product:  
[SAP Datasphere >](#)

**SAP Integration Suite**  
Connect and automate your business processes with integrations, connectors, APIs, and best practices for a faster ROI.  
Our product:  
[SAP Integration Suite >](#)

**Climate action solutions**  
Reduce the carbon footprint of your entire value chain (scope 1 to 3) based on actual business transactions.  
[SAP Sustainability Footprint Management >](#)

**SAP Analytics Cloud**  
Make decisions without doubt.  
Our product:  
[SAP Analytics Cloud >](#)

*“Internal” Sizing:  
How much does it cost?*

# Sizing / Capacity Modeling in the cloud – why?

## Roles

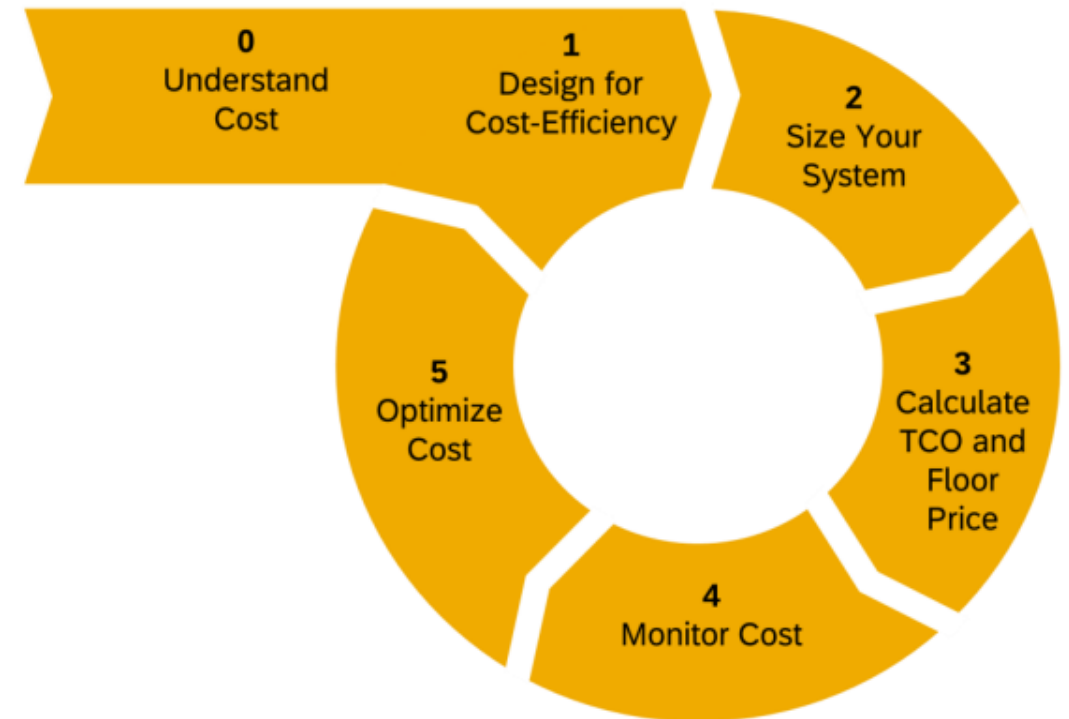
---

**Target: Determine the capacity required by an application within the customer's business context.**

- **Product Owners** need information of the operation costs per user, per service call, etc. for the TCO calculations (respectively license costs prediction/Floorplan Price).  
Please consider the costs associated with the resource consumption of your application and used SAP BTP services early e.g., in your architectural decisions. Only with a competitive floor price will an application be bought by customers.
- **DevOps** requires a model/procedure to predict and provision the resource consumption of the solution based on a reasonable number of input parameters and assumptions.

# Sizing is essential: The Cost Optimization Process

- **Step 1:** Design for Cost Efficiency starts off with considering and minimizing cost during your application's design and redesigns. Cost depends on the size of your application and the resources it uses. Therefore, the result of
- **Step 2:** Size Your System is an input to
- **Step 3:** Calculate TCO and Floor Price.
- **Step 4:** Monitor Cost prompts you to regularly review charged costs and the resource utilization of reserved instances and service plans you pay for and to look out for underused capacity. Finally,
- **Step 5:** Optimize Cost recommends cost reduction measures that should be applied during the operation of your application.



# Sizing is essential: The Cost Optimization Process

## 2: Size your system

### Resource capacity model

Determine the capacity required by an application within the customer's business context (e.g., compute, memory, storage) **based on measurements**

- Expected monthly service usage by typical customer
- Measured resource consumption of service execution + Buffer

Tier	Max Throughput/sec	# BP validations per month	Cloud Foundry Runtime [GB]	Connectivity service [1 Request]	Destination Service [1 API call]	Application Logging service [Tenant]	Identity authentication service [1 log on]	Event Mesh service [GB]	Cloud Logging service [Plan]	Hana Cloud Service [Capacity Unit]
1	12	750000	4	750000	750000	1 Standard Tenant	220	0,285	1 Standard Plan	843
2	24	1500000	8	1500000	1500000	1 Standard Tenant	440	0,57	1 Standard Plan	843
3	36	2250000	12	2250000	2250000	1 Standard Tenant	660	0,855	1 Standard Plan	1252
4	48	3000000	16	3000000	3000000	1 Standard Tenant	880	1,14	1 Standard Plan	1252

## 3: Calculate Floorprice

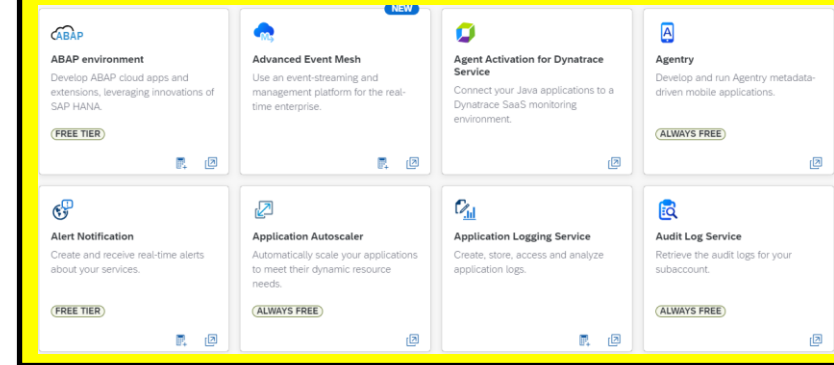
### Costs

Determine the costs for the required capacities for the floor price

- Expected monthly infrastructure cost per tier for a typical customer

### Operations

- Predict and plan for the resource consumption of the solution.



## 5: Optimize Costs

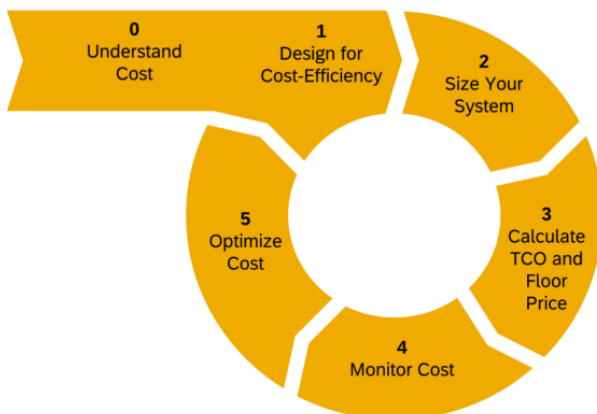
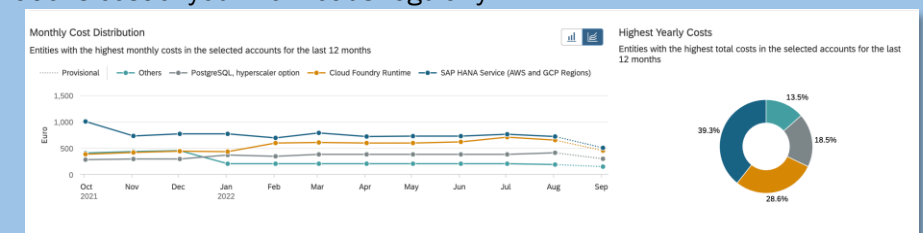
Re-calculation of infrastructure costs/

**Capacity Management: "Right-size" your infrastructure**

## 4: Monitor costs

Verify resource capacity model with productive cost distribution and provisioned capacities

- Real monthly infrastructure cost for a typical customer
- Review the resource utilization
- Monitor the operations cost of your workloads regularly
- Adopt Metering



# Architecture of an SAP BTP application



## Runtime

Cloud Foundry Runtime  
Capacity Unit:  
#instances → GB

Kyma Runtime  
Capacity Unit:  
VM

ABAP Runtime  
Capacity Unit:  
ABAP Compute Unit (ACU)

## Business Services

Connectivity Service  
Capacity Unit: #requests

Event Mesh  
Capacity Unit: # GB Bandwidth

Identity Authentication  
Capacity Unit: #Logons

Destination Service  
Capacity Unit: #API calls

Cloud Logging Service  
Capacity Unit: #Capacity Unit

Kafka  
Capacity Unit: # Node Hours

Application Logging  
Capacity Unit: #tenants

...

## Persistence Services

Mongo DB  
Capacity Unit: #tenants

...

PostgreSQL  
Capacity Unit:  
Memory In 2gb Blocks  
Memory In 4gb Blocks

PostgreSQL Storage  
Storage In 5gb Blocks  
Storage In 5gb Blocks (High  
Availability Variant)

HANA cloud  
Capacity Unit: #Capacity Unit

HANA cloud Storage

How to ensure that the selected infrastructure meets the requirements of the application?

How to do Application Sizing / Capacity Modeling that fits to the capacity units of the service provider?

**How to ensure that the selected infrastructure meets the requirements of the application?**



## Amazon EC2 Instance Types

General purpose instances provide a balance of compute, memory and networking resources, and can be used for a variety of diverse workloads. These instances are ideal for applications that use these resources in equal proportions such as web servers and code repositories.

**M5 instances** are the latest generation of General Purpose Instances powered by Intel Xeon® Platinum 8175M processors. This family provides a balance of compute, memory, and network resources, and is a good choice for many applications.

Instance Size	vCPU	Memory (GiB)	Instance Storage (GiB)	Network Bandwidth (Gbps)**	EBS Bandwidth (Mbps)
m5.large	2	8	EBS-Only	Up to 10	Up to 4,750
m5.xlarge	4	16	EBS-Only	Up to 10	Up to 4,750
m5.2xlarge	8	32	EBS-Only	Up to 10	Up to 4,750
m5.4xlarge	16	64	EBS-Only	Up to 10	4,750
m5.8xlarge	32	128	EBS-Only	10	6,800
m5.12xlarge	48	192	EBS-Only	12	9,500
m5.16xlarge	64	256	EBS-Only	20	13,600
m5.24xlarge	96	384	EBS-Only	25	19,000
m5.metal	96*	384	EBS-Only	25	19,000
m5d.large	2	8	1 x 75 NVMe SSD	Up to 10	Up to 4,750
m5d.xlarge	4	16	1 x 150 NVMe SSD	Up to 10	Up to 4,750

## Azure Machine families

General purpose VM sizes provide balanced CPU-to-memory ratio. Ideal for testing and development, small to medium databases, and low to medium traffic web servers. This article provides information about the offerings for general purpose computing.

The **Dv3-series** run on the 3rd Generation Intel® Xeon® Platinum 8370C (Ice Lake), Intel® Xeon® Platinum 8272CL (Cascade Lake), Intel® Xeon® 8171M 2.1GHz (Skylake), Intel® Xeon® E5-2673 v4 2.3 GHz (Broadwell), or the Intel® Xeon® E5-2673 v3 2.4 GHz (Haswell) processors in a hyper-threaded configuration, providing a better value proposition for most general purpose workloads.

Size	vCPU	Memory: GiB	Temp storage (SSD) GiB	Max data disks	Max temp storage throughput: IOPS/Read MBps/Write MBps	Max NICs/ Expected network bandwidth
Standard_D2_v3 <sup>1</sup>	2	8	50	4	3000/46/23	2/1000
Standard_D4_v3	4	16	100	8	6000/93/46	2/2000
Standard_D8_v3	8	32	200	16	12000/187/93	4/4000
Standard_D16_v3	16	64	400	32	24000/375/187	8/8000
Standard_D32_v3	32	128	800	32	48000/750/375	8/16000
Standard_D48_v3	48	192	1200	32	96000/1000/500	8/24000
Standard_D64_v3	64	256	1600	32	96000/1000/500	8/30000

## GCP Machine families

The general-purpose machine family has the best price-performance with the most flexible vCPU to memory ratios, and provides features that target most standard and cloud-native workloads.

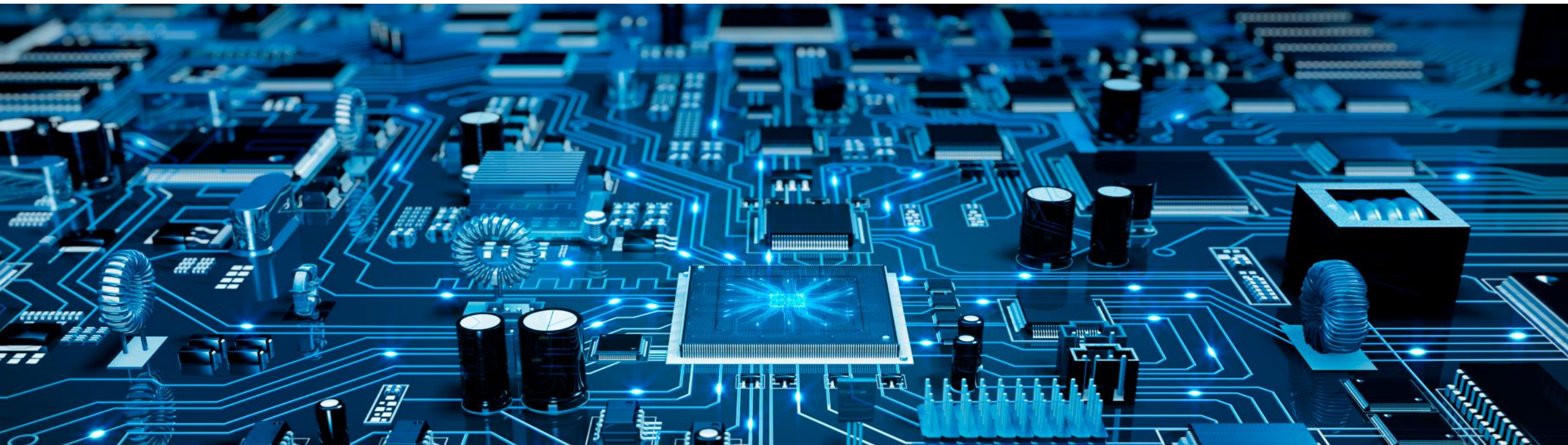
The **N2 machine series** has flexible sizing between 2 to 128 vCPUs and 0.5 to 8 GB of memory per vCPU. Machine types in this series run on the following processors:

- Ice Lake
- Cascade Lake

You can find more details about these two processors on the [CPU platforms](#) page.

Machine types	vCPUs <sup>1</sup>	Memory (GB)	Max number of (PDs) <sup>1</sup>	Max total PD size (TB)	Local SSD	Default egress bandwidth (Gbps) <sup>2</sup>	Tier 1 egress bandwidth (Gbps) <sup>2</sup>
n2-standard-2	2	8	128	257	Yes	10	N/A
n2-standard-4	4	16	128	257	Yes	10	N/A
n2-standard-8	8	32	128	257	Yes	16	N/A
n2-standard-16	16	64	128	257	Yes	32	N/A
n2-standard-32	32	128	128	257	Yes	32	50
n2-standard-48	48	192	128	257	Yes	32	50
n2-standard-64	64	256	128	257	Yes	32	75
n2-standard-80	80	320	128	257	Yes	32	100
n2-standard-96	96	384	128	257	Yes	32	100
n2-standard-128	128	512	128	257	Yes	32	100

# **How to do: Application Sizing / Capacity Modeling that fits to the capacity units of the service provider?**



# „CAPACITY UNIT” AND SIZING

---

- The Capacity Unit is a **unified metric**, used to organize the pricing of different applications and services. The semantic of a Capacity Unit with each application or service is different.
- The need to build a bridge between Capacity Unit and hardware sizing comes from the fact that the hardware resources which are used at runtime, should be known, and included in the license cost of the application/service-specific Capacity Unit.
- **It is the task of the service and application, which will use a specific service, to calculate how many Capacity Units will be sufficient for running their target customer load.**
- This means that these applications and services should provide sizing result in the corresponding metric of the service.
- For example, the Capacity Unit of “Cloud Foundry Runtime” service is sold per “GB Memory”, the Capacity Unit for “Destination Service” is “1 API Call”, the Capacity Unit for “Connectivity Service” is “1 request”, the Capacity Unit of “Identity Authentication Service” is “1 log on”, and so on.

# SAP credits $\neq$ SAP capacity unit

---

## GENERAL COMMERCIAL MODEL

---

**What is the consumption-based model in a nutshell?**

The objective of the model is to create one rich and simple customer experience for finding, trying, buying, and consuming cloud services from SAP and partners for net-new and installed base customers. It allows customers to integrate, extend and innovate own applications with business and technical functionality, offered as cloud services in a flexible way.

There is no need to license a SAP BTP service individually anymore. Cloud services can be used when demand arises and be retired if not needed anymore. Charges incur on actual usage and will be charged monthly against the prior purchased cloud credit.

# SAP Discovery Center

SAP Discovery Center Log On

Home / Services

## Services

Integrate and extend your solutions, optimize your business processes, and create an engaging digital experience using SAP Business Technology Platform services.

Search for SAP Business Technology Platform services Show Filters

Showing 98 services Sort By: A - Z

- Categories
  - All
    - New & Featured
    - Free Tier Services
    - Retiring Services
  - By Suite
    - Extension Suite - Development Efficiency
    - Extension Suite - Digital Experience
    - Extension Suite - Digital Process Automation
    - Integration Suite
  - By Capability
    - AI & Machine Learning
    - Analytics
    - Database and Data Management
    - DevOps
    - Developer Productivity
    - Digital Process Automation
    - Integration
    - Internet of Things
    - LOBs & Industries

<b>ABAP environment</b> Develop ABAP cloud apps and extensions, leveraging innovations of SAP HANA. <b>FREE TIER</b>	<b>Advanced Event Mesh</b> Use an event-streaming and management platform for the real-time enterprise. <b>NEW</b>	<b>Agent Activation for Dynatrace Service</b> Connect your Java applications to a Dynatrace SaaS monitoring environment.	<b>Agentry</b> Develop and run Agentry metadata-driven mobile applications.	<b>Alert Notification</b> Create and receive real-time alerts about your services. <b>FREE TIER</b>
<b>Application Autoscaler</b> Automatically scale your applications to meet their dynamic resource needs.	<b>Application Logging Service</b> Create, store, access and analyze application logs.	<b>Audit Log Service</b> Retrieve the audit logs for your subaccount.	<b>Authorization and Trust Management Service</b> Manage application authorizations and connections to identity providers.	<b>Automation Pilot</b> Simplify the operational effort behind any cloud solution in the SAP BTP. <b>FREE TIER</b>
<b>Bandwidth</b>	<b>Business Application Studio</b> <b>FEATURED</b>	<b>Business Entity Recognition</b>	<b>Cloud Foundry Runtime</b>	<b>Cloud Integration Automation</b>

## Example: Cloud Foundry Runtime

---

The “Cloud Foundry Runtime” service or the Kyma Runtime on its own cannot provide sizing because it does not have any business scenarios on its own.

**It is the task of the service and application, which will run on “Cloud Foundry Runtime”, to calculate how many “Cloud Foundry Runtime” Capacity Units will be sufficient for running their target customer load.**

This means that these applications and services should provide sizing result in metric “Cloud Foundry Runtime Capacity Units”.

# Example: Cloud Foundry Runtime

SAP BTP, Cloud Foundry Runtime  
Operate polyglot applications.

Overview Pricing Related Missions Roadmap

CPEA Pay-As-You-Go Subscription Trial

Service plans for CPEA (Cloud Platform Enterprise Agreement) is a consumption based license model.

Standard  
Standard Plan for Subscription

Prices:

For the following regions:

AWS: Australia (Sydney), Brazil (São Paulo), Canada (Montreal), Europe (Frankfurt), Japan (Tokyo), Singapore, So  
Microsoft Azure: Australia (Sydney), Europe (Netherlands), Japan (Tokyo), Singapore, US East (VA), US West (WA)  
Google Cloud: Europe (Frankfurt), US Central (IA)

Provider: All SAP-managed - 3  
SAP-managed  
All SAP-managed - 3  
AWS  
Microsoft Azure  
Google Cloud  
Partner-managed  
Alibaba

Service Plan: Standard  
Regions: All - 12  
Change Currency

Metric	Billing Block Size	Unit Price per Month	Description
GB Memory	1	70,00 EUR	GB of memory is defined as an amount of memory.

SAP BTP, Cloud Foundry Runtime  
Operate polyglot applications.

Service Plan: Standard  
Region / Provider: Europe (Frankfurt) AWS  
GB Memory: 4

Your Note:

Price Details

Metric	Unit Price per Month	Quantity	Amount
GB Memory	70.00 EUR	4	280,00 EUR

## Metering

SAP BTP Control Center

Manage Accounts Register Content Analyze Costs Manage Tasks Knowledge Center

Landscape: Canary - Public Cloud [Feature Set B]

Charges and usage information for resources consumed in the global accounts of which you are an admin.

Global Account: P&S-Technology Subaccount: All (CF & NEO) Service / App: Cloud Foundry Ru... Service Plans: Standard Metric: GB Memory

Monthly Costs  
Service plans with highest costs per month for the selected service or app for last 12 months.

Monthly Usage  
Service plans with highest usage per month for the selected service or app for last 12 months.

266 EUR average monthly costs (last month) 266 EUR total costs (last 12 months)

## Example: HANA Cloud Service

---

Similar is the case with “HANA Cloud Service” – here, again the applications and services, which will use it, must define the most typical and frequently used business processes, and prepare sizing.

- This sizing will be used to determine how many “HANA Cloud Service Capacity Units” will be required to run the expected customer load.
- Unlike the pure [GB] unit of “Cloud Foundry Runtime”, the capacity unit, called “Capacity Unit” of HANA Cloud Service is a combination of required vCPU and Memory.
- To determine number of “Capacity Units” of the HANA cloud service use the [SAP HANA Cloud Capacity Unit Estimator \(ondemand.com\)](https://ondemand.com).

# Example: HANA Cloud Service

SAP HANA Cloud

Overview **Pricing** Related Missions Roadmap Customer Stories

HANA  
An in-memory database supporting federation and replication [Add to Estimator](#)

Prices:

For the following regions:  
 AWS: Australia (Sydney), Brazil (São Paulo), Canada (Montreal), Europe (Frankfurt), Japan (Tokyo), Singapore, South Korea (Seoul), US East (VA)  
 Microsoft Azure: Australia (Sydney), Europe (Netherlands), Japan (Tokyo), Singapore, US East (VA), US West (WA)  
 Google Cloud: Europe (Frankfurt), US Central (IA)

Metric	Billing Block Size	Unit Price per Month	Description
Capacity Units	1	0.80 EUR	Capacity Units are the number of units consumed by the usage of the services as outlined in the solution specific product supplement. <a href="#">Calculator</a>

Prices:

For the following regions:  
 AWS: Europe (Frankfurt) EU Access

**SAP is not required to provide EU Access unless explicitly agreed in the Order Form between Customer and SAP. EU Access eligible Cloud Services are subject to the EU Access provisions in the Order Form**

Metric	Billing Block Size	Unit Price per Month	Description
Capacity Units	1	0.92 EUR	Capacity Units are the number of units consumed by the usage of the services as outlined in the solution specific product supplement. <a href="#">Calculator</a>

SAP HANA Cloud Capacity Unit Estimator

Selected Hyperscaler: **Microsoft Azure** Service Types: **SAP HANA Database** Activity Per Month: **100** **Percent** **Total Capacity Units: 3.691 per month**

PRE-CONFIGURATION **CUSTOM CONFIGURATION**

SAP HANA Database

SAP HANA Database Configurations

Name	Amount	Rate	Activity Hours	CU (per month)
<b>Memory</b> Memory size of a SAP HANA database	128 GB	0.023 CU / GB / Hour	730	2,149.12 CU
<b>Compute</b> The number of vCPUs of a SAP HANA database	8 vCPUs	0.160 CU / vCPU / Hour	730	934.4 CU
<b>Storage</b> The disk storage space of a SAP HANA database	360 GB	0.013 CU / 16 GB / Hour	730 (fixed per month)	213.52 CU
<b>Backup Storage</b> The backup storage space of a SAP HANA database It depends on the size of your hot and warm storage. It is calculated as a fixed amount of 14 times the size of the hot and warm storage combined.	2688 GB	0.011 CU / 64 GB / Hour	730 (fixed per month)	337.26 CU
<b>Network Data Transfer</b> The amount of network traffic generated by reading from the system (egress)	128 GB	0.438 CU / GB	N/A	56.06 CU

SAP HANA Cloud

Store, process, and federate data in a cloud infrastructure.

Service Plan: HANA Your Note:

Region / Provider: Europe (Frankfurt) AWS

Capacity Units: 3691

Price Details

Metric	Unit Price per Month	Quantity	Amount
Capacity Units	0.80 EUR	3691	2,952.80 EUR

2,952.80 EUR / MONTH

## Metering

CFS | SAP BTP Control Center

Landscape: Canary - Public Cloud [Feature Set B]

Charges and usage information for resources consumed in the global accounts of which you are an admin.

Global Account: P&S-Technology Subaccount: All (CF & NEO) Service / App: SAP HANA Cloud Service Plans: HANA Metric: Capacity Units

Monthly Costs: Service plans with highest costs per month for the selected service or app for last 12 months.

Monthly Usage: Service plans with highest usage per month for the selected service or app for last 12 months.

97 EUR average monthly costs (last 6 months) 582 EUR total costs (last 12 months)

# SAP HANA Cloud Capacity Unit Estimator

 Disclaimer

~~Sizing~~ information shown are estimates only and do not constitute a quote or an offer by SAP. Whilst every care is taken to ensure the sizing information is correct, no responsibility is accepted by SAP for its accuracy.

Contact SAP if you need assistance or additional information on sizing.

OK

SAP HANA Cloud Capacity Unit Estimator is a tool to calculate the required capacity units for HANA based on the capacity model.

SAP HANA Cloud Capacity Unit Estimator

Selected Hyperscaler

Microsoft Azure

Service Types

SAP HANA Database

Activity Per Month

100 Percent

Total Capacity Units

1,076 per month

PRE-CONFIGURATION

CUSTOM CONFIGURATION

CUSTOM CONFIGURATION

SAP HANA Database

SAP HANA Database Configurations

Name	Amount	Rate	Activity Hours	CU (per month)
<b>Performance Class</b>				
Choose a performance class that will provide an optimal performance-cost balance for this instance with its expected workload	Compute			
<b>Memory</b>				
Memory size of a SAP HANA database	32 GB	0.023 CU / GB / Hour	730	537.28 CU
<b>Compute</b>				
Number of vCPUs of a SAP HANA database	4 vCPUs	0.16 CU / vCPU / Hour	730	467.2 CU
<b>Storage</b>				
Disk storage space of a SAP HANA database	120 GB	0.013 CU / 16 GB / Hour	730 (fixed per month)	71.17 CU

High Compute


Compute

Memory

High Memory

# Kyma Runtime Estimator

[Kyma Runtime Estimator](#) is a tool to calculate the required capacity units for Kyma based on the capacity model.

 **Kyma Runtime Estimator**  
Size your scenario to calculate the required capacity units

↓ ↺

**RUNTIME** DETAILS

### Runtime

Name	Description	Service Consumption	Time Consumption	Capacity Units
Base Configuration (minimum requirements)	<ul style="list-style-type: none"><li>Virtual Machines of the selected size</li><li>Auto Scaler Min denotes the minimum number of available Virtual Machines</li><li>224 GB of storage</li></ul>	4 CPU - 16GB RAM 3	Virtual Machine Size Auto Scaler Min 720 hrs	1103,00
<b>Additional resources for custom workload</b>				
		0	Virtual Machines Based on choosen VM size	
Node	<b>Number of Additional Nodes for the Kyma cluster.</b> The term 'Node' refers to the commercial unit of 2vCPUs and 8GB RAM. It is not referring to a 'Kubernetes Node'.	0	Nodes 2vCPUs & 8GB RAM 720 hrs	0,00
Storage	<b>Number of GB</b> Increase steps are 32GB.	0	GB 720 hrs	0,00
<b>Total Estimated Capacity Units per month</b>				<b>1103,00</b>

# Sizing BTP applications

---

The following steps are required to come to an initial capacity model.

1. Description of the main business functions
2. Description of the architecture
3. Description of the underlying business processes
4. Description of the setup of the Single user/ load test and the test results

**The tests must be structured in such a way that they match the pricing tier (e.g. user, transaction, document, record,...).**

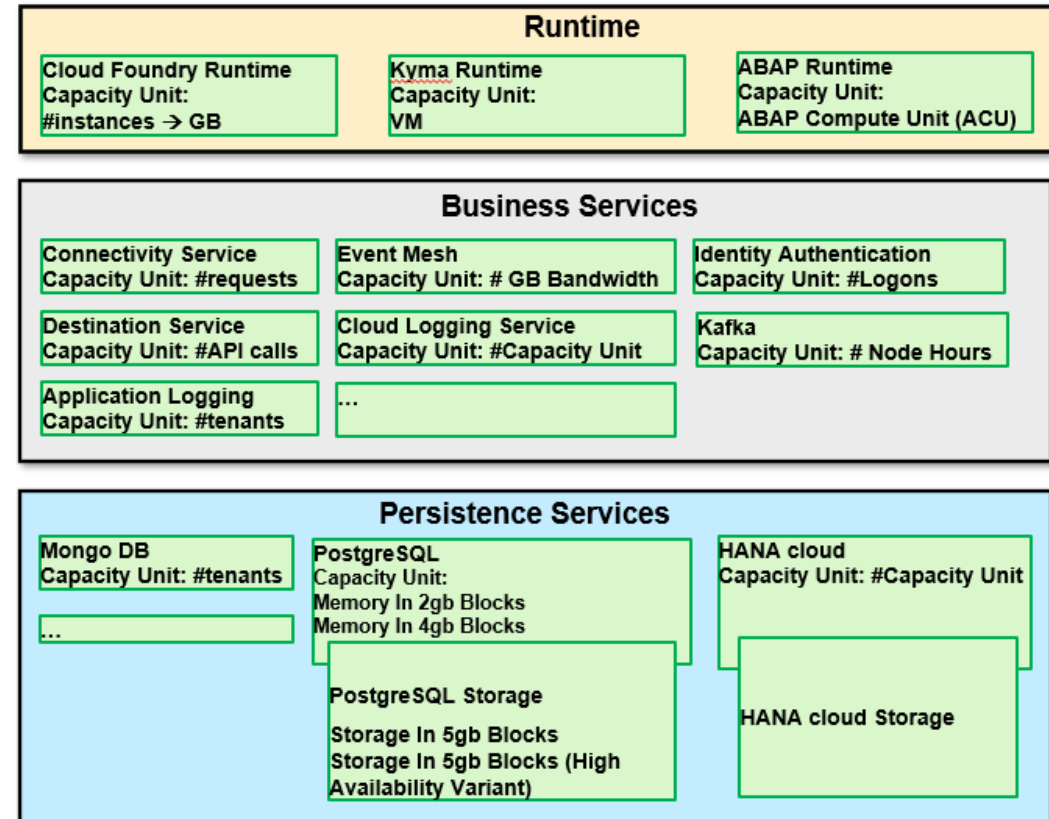
**Ultimately, the question is how high are the infrastructure costs per pricing tier.**

**This is a different focus compared to scalability tests that prove linearity.**

# Sizing SAP BTP applications

5. Usually, the SAP BTP applications and services have multiple references to other applications and services, and they reuse them at runtime. Provide a capacity model, for the target customer load, for the used services of the three main blocks:

- Runtime
- Business services (If possible, measure the usage of the business services during the tests. Otherwise do assumptions and an educated guess by reviewing your architecture)
- Persistency services



# Two Options for Sizing SAP BTP applications

A BTP application needs to deliver sizing results in capacity units for all the used services based on either single user tests and or multiuser load tests.

## Sizing approach: Single user tests

Sizing test cases should be defined for each of the sizing-relevant business processes. Different values of the input parameters should be used to analyze the dependence of resource consumption on concrete usage.

The best practice is to measure the sizing KPIs while executing the sizing scenario with single user for creating a mathematical sizing model.

## Sizing approach: Load tests

The sizing approach is based on load test results, to create a sizing formula/capacity model describing how much throughput can be achieved with an application instance with the tested memory configuration.

Sizing results are verified by scaling out: Instantiate a second instance with same size and verify whether twice as much throughput with acceptable response time can be achieved.

$$\#vCPU = \frac{N_{scenarios} * CPU [s]}{tP[s] * targetUtil} = \frac{\text{Required CPU [s]}}{\text{Available CPU [s] per vCPU} * targetUtil}$$

$$\text{Memory [GB]} = \text{roundup} \left( N_{scenarios} * \frac{D[s]}{tP[s]} \right) * \text{Memory [MB]}$$

$N_{scenarios} * \frac{D[s]}{tP[s]}$  ■ Determines how many scenarios will consume memory at the same time. Value should be rounded up to integer.



# Options for Sizing SAP BTP applications – Single User Test

---

Option number one is the **sizing approach based on single user tests**. It comprises the following steps:

- Proven linearity via separate load tests as precondition
- Scenario definition
- Single user tests to measure resource consumption metrics per user interaction step (UIS)
- The tests must be structured in such a way that they match the pricing tier. Ultimately, the question is how high are the infrastructure costs per pricing tier.
- Sizing formula/capacity model creation → mathematical model (extrapolation based on single user measurement)
- Verification of mathematical model with load test for a reduced set of scenarios

# Options for Sizing SAP BTP applications – Load Tests

---

The **load test sizing approach** comprises the following steps.

- Scenario definition
- Multiuser tests for measuring the achievable throughput of an application instance with a defined memory configuration and number of vCPU
- Multiuser test to validate linearity and scalability and sizing/capacity model
- The tests must be structured in such a way that they match the pricing tier. Ultimately, the question is how high are the infrastructure costs per pricing tier.
- Sizing formula/capacity model creation based on load test results and applying scale out factor

# Guidance on How to Size CPU and Memory

## Preconditions

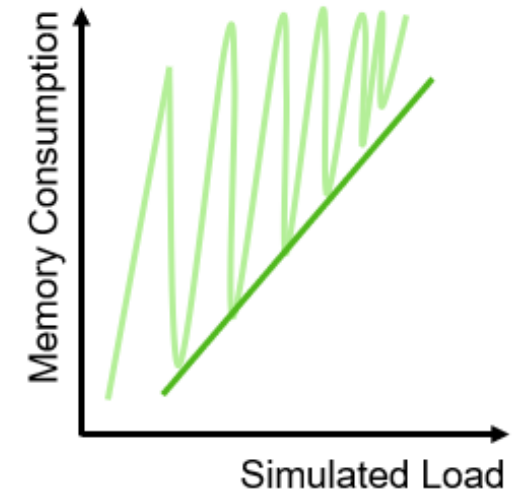
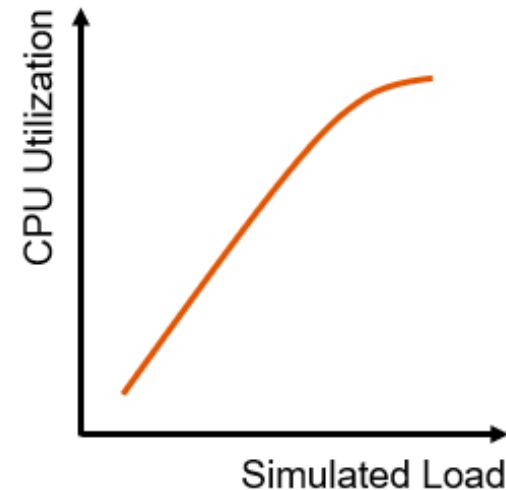
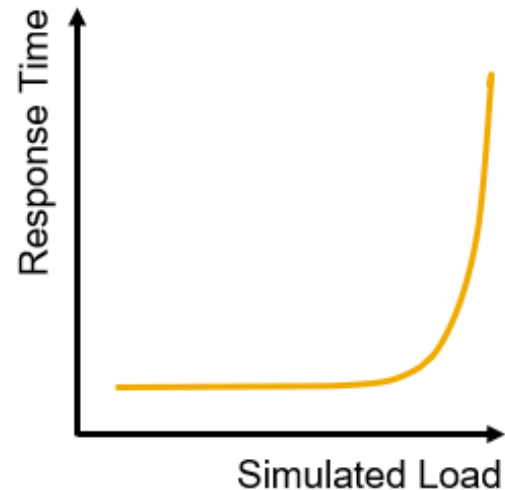
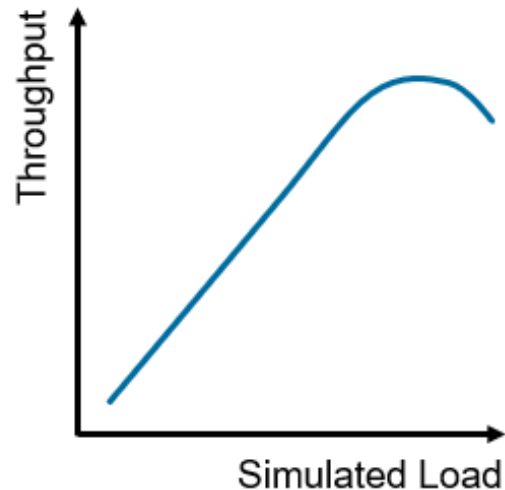
---

To follow the guidance, the following preconditions must be fulfilled:

- A representative error-free test scenario is identified
- Some experience with running load tests with the load test tool of your choice is available
- A load test script simulating the test scenario is in place
- Monitorability of simulated throughput and request response time is available
- Monitorability of memory consumption & CPU utilization for all components (e.g., compute node, persistence) of the application instance under test is available

# Guidance on How to Size CPU and Memory

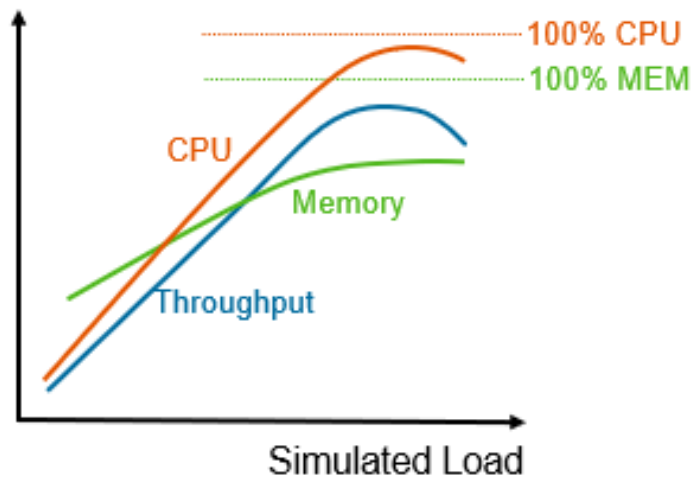
Collect throughput, response time, CPU utilization and memory consumption metrics for the test.



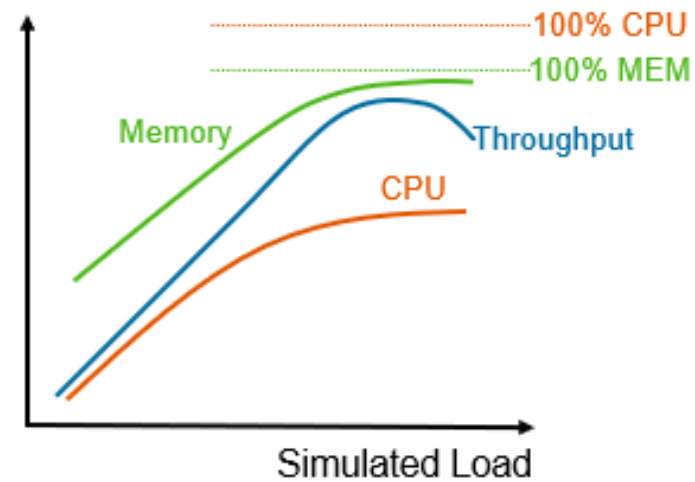
# Guidance on How to Size CPU and Memory

Identify the limiting resource: is your application CPU or memory bound? If your application does not hit the CPU or memory limit (e.g., due to IO bottlenecks or wait situations) you cannot continue with sizing. Find and solve the bottleneck first.

Example for a CPU-bound application



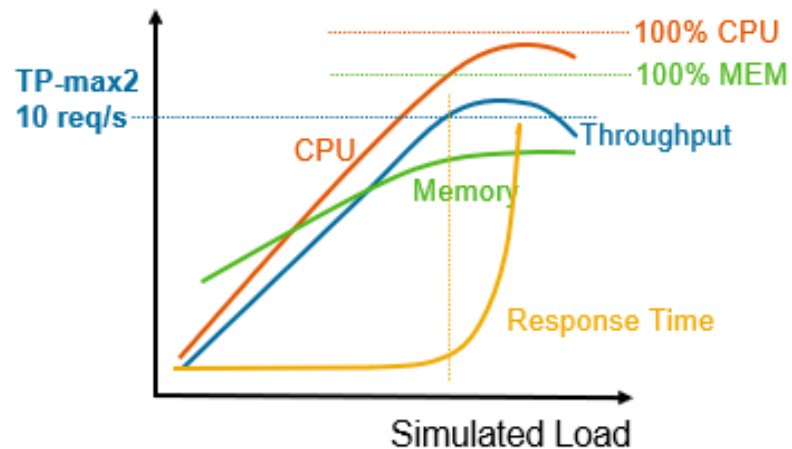
Example for a memory-bound application



# Guidance on How to Size CPU and Memory

Based on the results, create a sizing formula/capacity model describing how much throughput can be achieved with an application instance with the tested memory configuration.

Example: 10 requests / sec can be processed on an application instance with 4GB



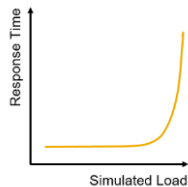
$$\text{Min. Required Memory (GB)} = \frac{TP * 4GB}{10}$$

*TP = Desired requests per sec*

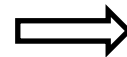
Desired Requests / sec	Min. Required Memory	# App Instances
8	3,2GB	1 (* 4GB)
22	8,8GB	3 (* 4GB)
54	21,6GB	6 (* 4GB)
89	35,6GB	9 (* 4GB)

# Guidance on How to Size CPU and Memory

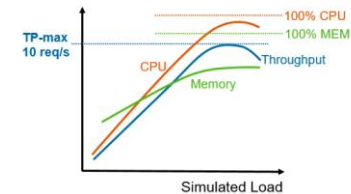
Verify the sizing by scaling out. Instantiate a second instance with same size and verify whether twice as much throughput with acceptable response time can be achieved. If yes, you're done. If not, stop here, solve the issue and run the test again.



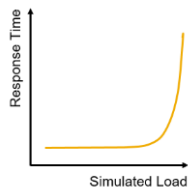
10 requests / sec



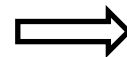
4GB App.  
Instance



Scale Out

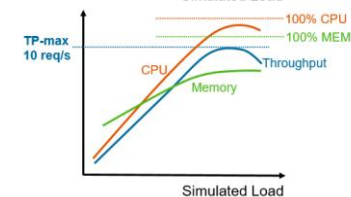
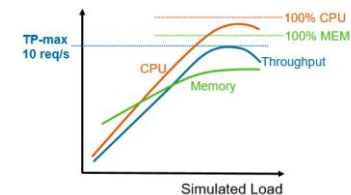


20 requests / sec



4GB App.  
Instance

4GB App.  
Instance



# Central services

---

In case the application instance to be sized is using central services provided e.g., by SAP BTP, identify those who do not scale elastically.

Choose from the existing service plans of the central service, so that the selection allows your application to scale well. In other words, make sure that the scalability of the application to be sized is not limited by the capacity of used central services.

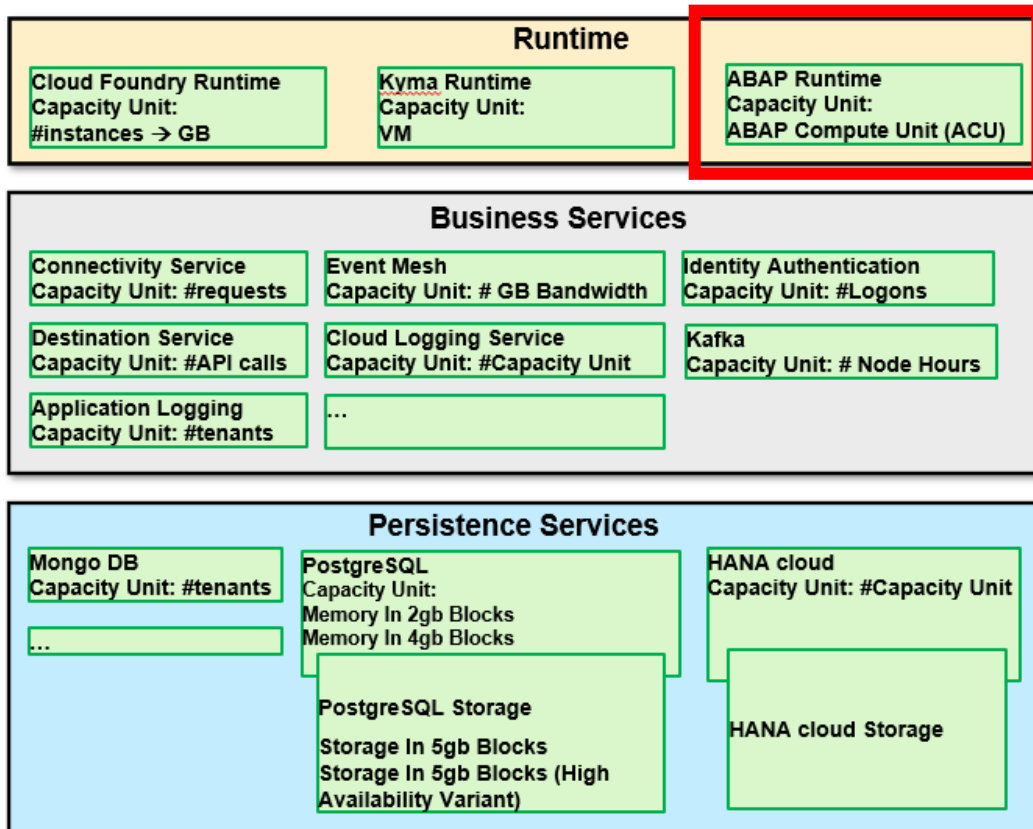
Monitor while running the sizing tests the usage of central services as well.

In your final sizing guide add a table allowing to map the throughput level of application to be sized to the required service plan of the used central services.

Example:

Requests / sec	Required Memory	# App Instances	Central Service A
8	3,2GB	1 (* 4GB)	Service Plan: Small
22	8,8GB	3 (* 4GB)	Service Plan: Small
54	21,6GB	6 (* 4GB)	Service Plan: Medium
89	35,6GB	9 (* 4GB)	Service Plan: Large

# Excursus: ABAP runtime



# SAP BTP ABAP Environment

---

The most recent innovation was the enablement of the ABAP stack for SAP BTP (project Steampunk).

With SAP BTP ABAP Environment customers and partners can build Cloud applications and loosely coupled side-by-side extensions for SAP's flagship solutions SAP S/4HANA and SAP S/4HANA Cloud.

With the Embedded Steampunk approach tightly coupled extensions can be built that run directly on top of the SAP S/4HANA Cloud technology stack using public local APIs offered by the SAP S/4HANA business objects.

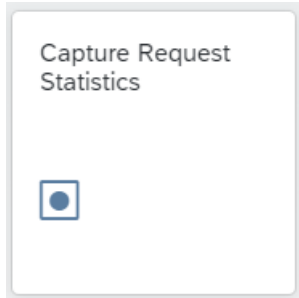
They are upgrade-safe avoiding the typical adaptation efforts that can occur in classical ABAP developments.

The ABAP RESTful Application Programming Model (RAP) is at the heart of each application development and extension project.

# BTP applications in the ABAP environment

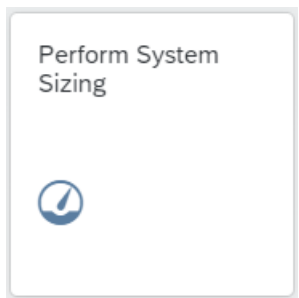
---

The documentation for [ABAP System Sizing](#) describes the principles of how to perform system sizing for a custom application operated in the ABAP environment.



With the **Capture Request Statistics** app, you can capture request statistics to find out which activities are running in your ABAP system. For sizing, you can define a capture profile that records defined activities of a business user or business process.

When you order an ABAP environment in SAP BTP, the ABAP system size is specified in **ABAP compute units (ACUs)**. One ABAP compute unit comprises the total ABAP memory usable by applications, the ABAP work process time per minute, and the ABAP CPU time per minute.



With the **Perform System Sizing** app, you can calculate the required total ABAP memory, the work process time, and the ABAP CPU time in ABAP compute units for a custom application operated in the ABAP environment. You can use these calculations as a basis for ABAP system sizing.

# SAP BTP applications in the ABAP environment

After performing a system sizing for each business process, combine the results for multiple custom applications that are planned to run in the production system. As a result, you get an idea of its required system size in ABAP compute units (ACUs).

Flight Reference Scenario: Example Sizing Results

Business Process	ABAP Total Memory	Work Process Time	ABAP CPU Time	Process Executions/Minute
Create a new travel request	0.108 ACUs	0.306 ACUs	0.414 ACUs	180
Approve a travel request	0.072 ACUs	0.162 ACUs	0.234 ACUs	180
Calculate and send metrics for reporting	0.025 ACUs	0.331 ACUs	0.092 ACUs	1



# SAP BTP applications in the ABAP environment

ABAP SAP BTP, ABAP environment

Service Plan: Standard

Region / Provider: Europe (Frankfurt) AWS

Hours: 730

Hours of Runtime Memory in 16 GB Blocks: 1

Hours of Persistent Memory in 16 GB Blocks: 2

Your Note:

Price Details

Metric	Unit Price per Hour	Quantity	Hours	Amount
Hours of Runtime Memory in 16 GB Blocks	EUR 1.923	1	730	EUR 1.403,79
Hours of Persistent Memory in 16 GB Blocks	EUR 0.549	2	730	EUR 801,54

# SAP BTP applications in the ABAP environment

**ABAP runtime size** (parameter `size_of_runtime`) refers to the size of the ABAP runtime. It's part of the quota plan `abap_compute_unit`, with **one ABAP compute unit representing 16 GB**.

The supported number of `abap_compute_unit` is 1, 2, 4, 6, 8, 16, or 32.

There are two different sizes for the App-server in place (8 GB RAM/ 32 GB RAM); If additional ACUs are required, additional App-servers will be added (scale out).

## Restriction

You only have the following options to increase the ABAP runtime size:

- From 1 to 2, 4, 6, or 8
- From 2 to 8
- From 4 to 6 or 8
- From 6 to 8

If you want to decrease the number of ABAP runtime size, create an incident using component `BC_CP_ABA`.

# SAP BTP applications in the ABAP environment

---

**HANA memory size** (parameter `size_of_persistence`) refers to the size of SAP HANA memory. It's part of the quota plan `hana_compute_unit`, with one HANA compute unit representing the suitable block size for the underlying SAP HANA Cloud instance (15 GB on AWS).

The supported number of `hana_compute_unit` per HANA instance is **2**, 4, 8, 16, 32, or 64.

The **size of persistence disk** (parameter `size_of_persistence_disk`) refers to the size of the SAP HANA Cloud storage in GB. If the parameter is set to `auto`, the SAP HANA Cloud storage size is set to the minimal value  $40 * \text{size\_of\_persistence} + 40$ . The maximum allowed value is  $120 * \text{size\_of\_persistence} + 40$ . If you set a higher value, it will consume 0.002 HANA compute units (HCU) for any GB exceeding the minimal default size of the persistence disk. Therefore, the HCU ratio of additional storage disk to RAM is 1 : 33.33 per GB on AWS (as 1 HCU = 15 GB) and 1 : 31.25 per GB on Microsoft Azure (as 1 HCU = 16 GB).

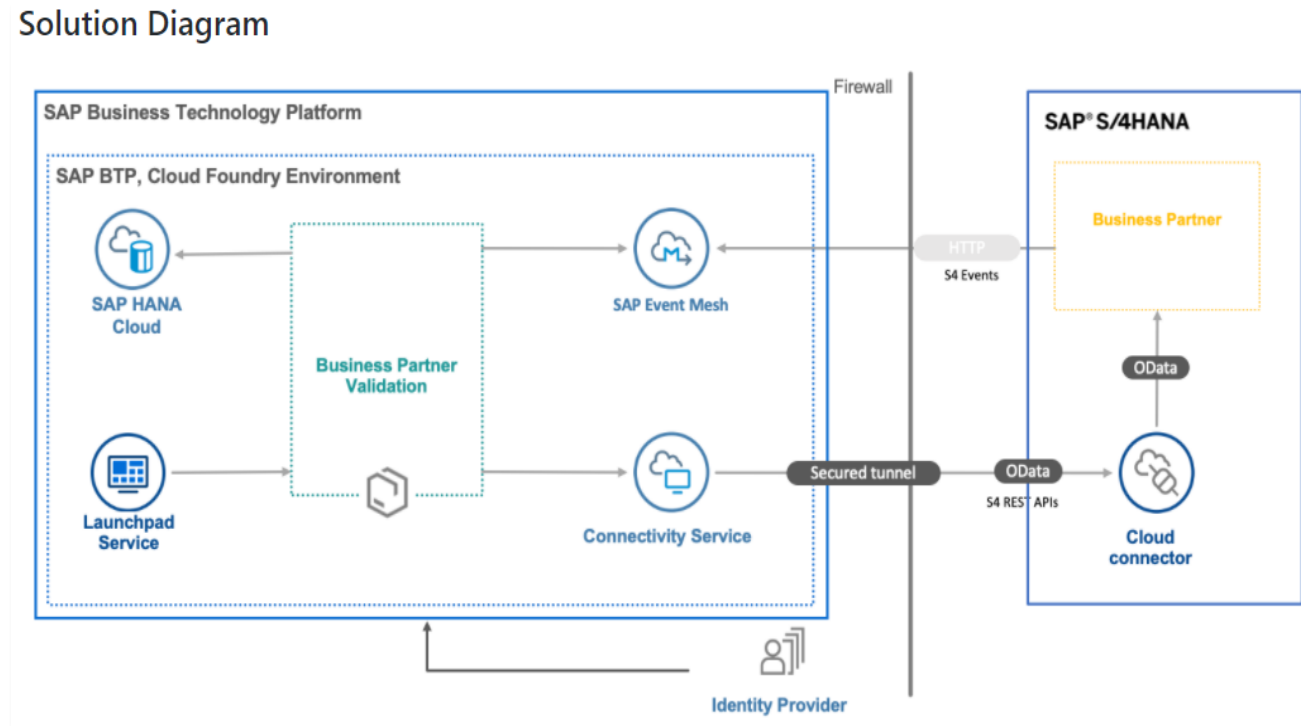
Source: <https://help.sap.com/docs/BTP/65de2977205c403bbc107264b8eccf4b/50b32f144e184154987a06e4b55ce447.html>  
<https://help.sap.com/docs/BTP/65de2977205c403bbc107264b8eccf4b/7890ffa8a7274ac1852b37ede5b773d1.html>

# **Example: The “Business Partner Validation” application - Sizing approach based on load tests on SAP BTP CF**



# “Business Partner Validation” application

A typical solution diagram of a SAP BTP application is presented below. The “Business Partner Validation” application is running in “Cloud Foundry Runtime” and is sending calls to “Event Mesh Service”, “Connectivity Service”, and “Identity Provider Service”. It stores data into “SAP HANA Cloud Service”. It stores data into “SAP HANA Cloud Service”. It stores data into “SAP HANA Cloud Service”.



# Scope of sizing

---

Provide a capacity model

- for the **“key entity”**,
- for the **target customer load** and
- for the **used services** of the three main blocks:
  - Runtime
  - Business services (**If possible, measure the usage of the business services during the tests. Otherwise do assumptions and an educated guess by reviewing your architecture**)
  - Persistency services

# Key entity

---

## Key sizing entity of business scenario

For the business scenario the “key entity” needs to be defined by the product owner. Basically, these are the pricing metrics to determine the initial floor price (infrastructure costs) of your application/ service.

- Example:

Tier	Max Throughput/sec	# events per month
1	12	750000
2	24	1500000
3	36	2250000
4	48	3000000

Tier	FUE
1	500
2	1500
3	4000
4	>6000

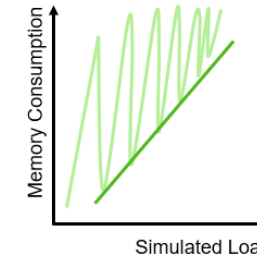
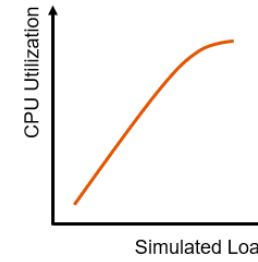
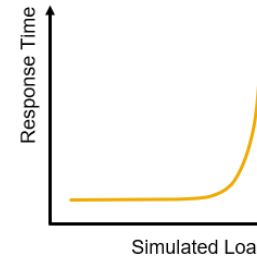
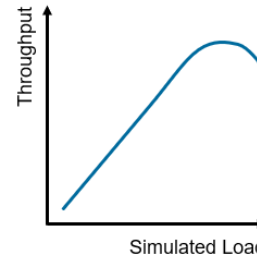
# Load Test Results

---

- The tests must be structured in such a way that they match the pricing tier. Ultimately, the question is how high are the infrastructure costs per pricing tier. This is a different focus compared to scalability tests.
- During the load tests all services including the central services need to be measured to enable the creation of a capacity model
- Make sure that the scalability of the application to be sized is not limited by the capacity of used central services
- Monitor while running the sizing tests the usage of central services as well
- During the load test you have to monitor the application instances and the resource consumption of the persistence service. If you run into resource bottlenecks with the persistence service, you need to identify if you are memory or CPU bound and adjust the configuration accordingly
- During the load scale out instances

# Load Tests

Multi User test	Only for BTP application		HANA Configuration	
Delay	100ms delay		Memory:	50GB
Rampup	1 user for every 2 seconds		CPU	2vCPUs
CF Runtime Instance	Each instance of 4Gb memory		Storage	120GB



Instance	User	Duration (mins)	Total Business Partner Created	StepName	Client Side Metrics		App Metrics		HANA Metrics		Comments	Amount of Data sent from S4H to BTP(KB) ( Each message of size 410 bytes)	Destination Service			Identity Service		
					Avg Res Time	Throughput /sec	CPU %	Memory (MB)	CPU %	Max Memory (MB)			Response time (ms)	Throughput	Total Request	Response time	Throughput	Total Request
1	10	15	10800	020_Get_Not_id	138	12	80-85	~420	30-40	6500	100ms delay between update of BP							
1	10	15	10800	030_Create_Dra	79	12	80-85	~420	30-40	6500	100ms delay between update of BP							
1	10	15	10800	040_Update_BP	84	12	80-85	~420	30-40	6500	100ms delay between update of BP							
1	10	15	10800	050_Save_BP	108	12	80-85	~420	30-40	6500	100ms delay between update of BP	4324	38	344/min	6.87K	152	0.05/min	1
2	20	15	22000	020_Get_Not_id	134	24	80-85	~420	30-50	6500	100ms delay between update of BP							
2	20	15	22000	030_Create_Dra	74	24	80-85	~420	30-50	6500	100ms delay between update of BP							
2	20	15	22000	040_Update_BP	70	24	80-85	~420	30-50	6500	100ms delay between update of BP							
2	20	15	22000	050_Save_BP	102	24	80-85	~420	30-50	6500	100ms delay between update of BP	8808	33	629/min	22K	152	0.14/min	5
3	30	15	29000	020_Get_Not_id	134	32	80-85	~750-900	50-75	6500	100ms delay between update of BP							
3	30	15	29000	030_Create_Dra	102	32	80-85	~750-900	50-75	6500	100ms delay between update of BP							
3	30	15	29000	040_Update_BP	96	32	80-85	~750-900	50-75	6500	100ms delay between update of BP							
3	30	15	29000	050_Save_BP	135	23	80-85	~750-900	50-75	6500	100ms delay between update of BP	11611	38	652/min	29K	152	0.14/min	5
4	40	15	37000	020_Get_Not_id	157	42	80-85	~750-900	70-85	6500	100ms delay between update of BP							
4	40	15	37000	030_Create_Dra	107	42	80-85	~750-900	70-85	6500	100ms delay between update of BP							
4	40	15	37000	040_Update_BP	102	42	80-85	~750-900	70-85	6500	100ms delay between update of BP							
4	40	15	37000	050_Save_BP	127	42	80-85	~750-900	70-85	6500	100ms delay between update of BP	14814	39	950/min	37K		Rarely ocured	
5	50	15	42700	020_Get_Not_id	185	48	80-90	~750-900	80-100	6500	100ms delay between update of BP							
5	50	15	42700	030_Create_Dra	125	48	80-90	~750-900	80-100	6500	100ms delay between update of BP							
5	50	15	42700	040_Update_BP	118	48	80-90	~750-900	80-100	6500	100ms delay between update of BP							
5	50	15	42700	050_Save_BP	159	48	80-90	~750-900	80-100	6500	100ms delay between update of BP	17096	39	920/min	38k	315	5/min	208
1	1	15	1200	020_Get_Not_id	138	1,3	15	~240	5	6500	100ms delay between update of BP							
1	1	15	1200	030_Create_Dra	73	1,3	15	~240	5	6500	100ms delay between update of BP							
1	1	15	1200	040_Update_BP	67	1,3	15	~240	5	6500	100ms delay between update of BP							
1	1	15	1200	050_Save_BP	90	1,3	15	~240	5	6500	100ms delay between update of BP	492	25,7	65/min	1.2K	94,8	0,07	2
1	2	15	2300	020_Get_Not_id	139	2,6	23-25	~300-400	8	6500	100ms delay between update of BP							
1	2	15	2300	030_Create_Dra	69	2,6	23-25	~300-400	8	6500	100ms delay between update of BP							
1	2	15	2300	040_Update_BP	64	2,6	23-25	~300-400	8	6500	100ms delay between update of BP							
1	2	15	2300	050_Save_BP	87	2,6	23-25	~300-400	8	6500	100ms delay between update of BP	943	26,1	77/min	2.3k	50,4	0,07	2
1	4	15	4700	020_Get_Not_id	134	5,2	45	~300-400	12	6500	100ms delay between update of BP							
1	4	15	4700	030_Create_Dra	67	5,2	45	~300-400	12	6500	100ms delay between update of BP							
1	4	15	4700	040_Update_BP	64	5,2	45	~300-400	12	6500	100ms delay between update of BP							
1	4	15	4700	050_Save_BP	86	5,2	45	~300-400	12	6500	100ms delay between update of BP							
1	8	15	9100	020_Get_Not_id	134	10,2	70	~300-400										
1	8	15	9100	030_Create_Dra	73	10,2	70	~300-400										
1	8	15	9100	040_Update_BP	67	10,2	70	~300-400										
1	8	15	9100	050_Save_BP	97	10,2	70	~300-400										



CF Runtime	Client Side Metrics	Event Mesh Service	Destination Service	Identity Service	Connectivity Service
4GB Instance	Throughput/sec	Amount of Data sent from S4H to BTP(GB) ( Each message of size 410 bytes)	Total Request per BP	Total Request Based on #users Working Days per month: 22	Total Request per BP
2	24	0,00000038	1	20	1

# Capacity Model – Calculation in detail

The test results of the business scenario provide the required capacity information for the used services that can be mapped to the key entity.

For the scenario “Changes in S/4 communicated via events in real time to extension application the “key entity is the validation of Business Partner”. Defined by the product owner there will be two dimensions being offered to the external customer:

- Max Throughput (#validations of business partners)/sec
- # Business Partner validations per month

Tier	Max Throughput/sec	# BP updates per month
1	12	750,000
<b>2</b>	<b>24</b>	<b>1,500,000</b>
3	32	2,250,000
4	42	3,000,000
5	48	3,750,000

CF Runtime				Client Side Metrics	Event Mesh Service	Destination Service	Identity Service	Connectivity Service
4GB Instance	User	Duration (mins)	Business Partner	Throughput/sec	Amount of Data sent from S4H to BTP(GB) ( Each message of size 410 bytes) per BP	Total Request per BP	Total Request Based on #users Working Days per month: 22	Total Request per BP
2	20	15	22000	24	0,00000038	1	20	1

Service	Capacity Unit	Capacities per month	Calculation
Cloud Foundry Runtime	GB Memory (4 GB = 1 Instance)	8	= 2 Instances x 4GB
Connectivity Service	1 request	1500000	= 1500000 BP x 1 Request
Destination Service	1 API call	1500000	= 1500000 BP x 1 API call
Application Logging Service	1 Standard Tenant	1	Educated Guess
Identity Authentication Service	1 log on	440	=22 working days x 20 daily logons
Event Mesh Service	1 GB	0,57	=1500000 BP x 0,00000038
Cloud Logging Service	1 Standard Plan	1	Educated Guess
HANA Cloud Service	Capacity Unit	843	SAP HANA Cloud Capacity Unit Estimator

# Capacity Model

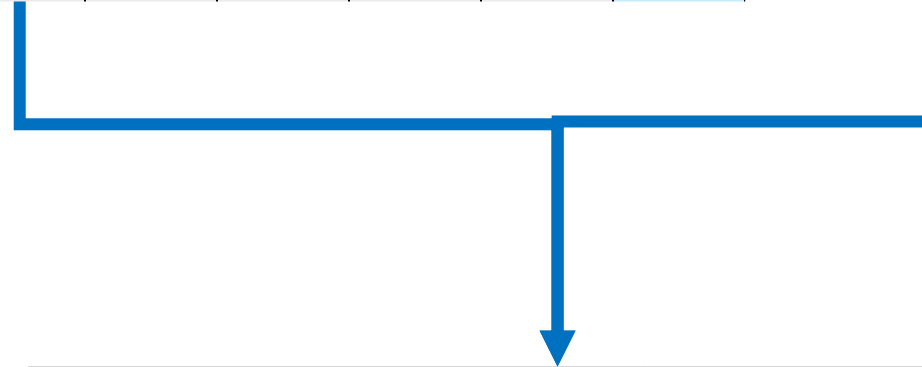
The output of correctly prepared sizing would look as follows.

Tier	Max Throughput/sec	# BP validations per month	Cloud Foundry Runtime [GB]	Connectivity service [1 Request]	Destination Service [ 1API call]	Application Logging service [Tenant]	Identity authentication service [1 log on]	Event Mesh service [GB]	Cloud Logging service [Plan]	Hana Cloud Service [Capacity Unit]
1	12	750000	4	750000	750000	1 Standard Tenant	220	0,285	1 Standard Plan	843
2	24	1500000	8	1500000	1500000	1 Standard Tenant	440	0,57	1 Standard Plan	843
3	36	2250000	12	2250000	2250000	1 Standard Tenant	660	0,855	1 Standard Plan	1252
4	48	3000000	16	3000000	3000000	1 Standard Tenant	880	1,14	1 Standard Plan	1252

# Cost calculation

## Pricing metric

Tier	Max Throughput/sec	# BP validations per month	Cloud Foundry Runtime [GB]	Connectivity service [1 Request]	Destination Service [1 API call]	Application Logging service [Tenant]	Identity authentication service [1 log on]	Event Mesh service [GB]	Cloud Logging service [Plan]	Hana Cloud Service [Capacity Unit]
1	12	750000	4	750000	750000	1 Standard Tenant	220	0,285	1 Standard Plan	843
2	24	1500000	8	1500000	1500000	1 Standard Tenant	440	0,57	1 Standard Plan	843
3	36	2250000	12	2250000	2250000	1 Standard Tenant	660	0,855	1 Standard Plan	1252
4	48	3000000	16	3000000	3000000	1 Standard Tenant	880	1,14	1 Standard Plan	1252



Service	Costs per month
Cloud Foundry Runtime	648,00
Connectivity Service	0,00
Destination Service	0,00
Application Logging Service	299,00
Identity Authentication Service	51,39
Event Mesh Service	68,73
HANA Cloud Service	744,28
	<b>1811,40</b>

## Discovery Center

The screenshot shows the SAP Discovery Center interface. At the top, there are tabs for 'CPEA', 'Pay-As-You-Go', 'Subscription', and 'Trial'. Below the tabs, there are three sections, each with a blue information icon and a title: 'Service plans for CPEA (Cloud Platform Enterprise Agreement is a consumption based commercial model)', 'Service plans for Pay-As-You-Go (Pay-As-You-Go is a consumption based commercial model)', and 'Service plans for Subscription (A subscription allows access to cloud service up to a licensed maximum, independent of actual usage)'. Below these sections, there is a search bar and a grid of service cards. The cards include: 'ABAP environment', 'Advanced Event Mesh', 'Agent Activation for Dynamic Service', 'Agency', 'Alert Notification', 'Application Autoscaler', 'Application Logging Service', 'Audit Log Service', 'Authorization and Trust Management Service', 'Automation Pilot', 'Bandwidth', 'Business Application Studio', 'Business Entity Recognition', 'Cloud Foundry Runtime', and 'Cloud Integration Automation'.

# Where to Find More Information

---

- <https://www.sap.com/germany/products/technology-platform.html>
- [SAP Discovery Center](#)
- [SAP S/4HANA Cloud ABAP Environment | SAP Community](#)
- <http://www.sap.com/sizing>
  - Sizing methodology
  - Access Quick Sizer
  - Access Sizing Guidelines
- <http://www.sap.com/benchmark>
  - Certified benchmark results
  - Descriptions of the available benchmarks
- <http://www.sap.com/performance>
  - Tips for performance improvements

# Key Points to Take Home

---

- The cloud services provider hosts, manages and maintains all the hardware and software included in the platform (PaaS)
- If you develop an application on the SAP Business Technology Platform you need to determine the capacity required by your application within the customer's business context.
- The Capacity Unit is a unified metric, used to organize the pricing of different applications and services. The semantic of a Capacity Unit with each application or service is different.
- An SAP BTP application needs to deliver sizing results in capacity units for all the used services based on either single user tests and or multiuser load tests.
- Provide a capacity model, for the target customer load, for the used services of the three main blocks:
  - Runtime
  - Business services (If possible, measure the usage of the business services during the tests. Otherwise do assumptions and an educated guess by reviewing your architecture)
  - Persistency services

---

Detlef Thoms

Chief Product Expert Performance &  
Scalability

SAP SE

[detlef.thoms@sap.com](mailto:detlef.thoms@sap.com)

<https://www.linkedin.com/in/detlef-thoms-39782723/>

Please remember to complete  
your session evaluation.



## SAPinsider.org

PO Box 982Hampstead, NH 03841  
Copyright © 2023 Wellesley Information Services.  
All rights reserved.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies. Wellesley Information Services is neither owned nor controlled by SAP SE.

---

**SAPinsider comprises the largest and fastest growing SAP membership group worldwide, with more than 750,000 global members.**

---