ABBYY

ABBYY Vantage

Performance Guide

Table of Contents

Introduction

ABBYY Vantage is installed with the minimum possible number of services and workers. Depending on the load, ABBYY Vantage will automatically scale the services and workers to optimize document processing. This guide contains information about what resources ABBYY Vantage will require depending on the load, and recommendations for the System Administrator on what needs to be done in order to correctly provide these resources to ABBYY Vantage.

Reference Configurations

Resource consumption depends on your document processing scenario — the type of documents being processed, the skill being used, and the page load (i.e. the number of pages processed within a certain time period). The tables below show resource consumption on different ABBYY Vantage configurations tested in different scenarios. For more information about the available deployment options (**Highly available** and **Without high availability** configurations), see the System Administrator's Guide.

ABBYY Vantage uses two types of machines for nodes — Standard_F8s_v2 for services and Standard_E8_v4 for workers.

Machine type	Number of virtual CPUs	RAM, GB	Ephemeral disk size, GB
Standard_F8s_v2	8	16	100
Standard_E8_v4	8	64	100

In each tested scenario,TIFF files were fed to the system via the REST API. The Process Skill with the following workflow was used:

1. Import TIFF files

2. Recognize documents

3. Classify and determine document types

4. For invoices and purchase orders, classify and determine document regions

5. Extract data from documents

6. Export data to JSON

Processing 3-page invoices in "Highly available" configuration

Tests were carried out on 500, 10,000, 40,000, 70,000, and 100,000 pages of various commercial invoices (three pages per file) using the **Highly available** configuration.

Depending on the page load, ABBYY Vantage required the following amount of resources to efficiently process documents:

Load, (number of pages/8 hours)	Nodes for services (Standard_F8s_v2)	Nodes for workers (Standard_E8_v4)
500	9	1
10,000	10	3
40,000	10	5
70,000	10	7
100,000	10	9

🟧 **Note:** When scaling ABBYY Vantage, no increase in document processing times was noted.

Processing 50-page invoices in "Highly available" configuration

Tests were carried out on 500, 20,000, 60,000, 100,000, 150,000, and 200,000 pages of various commercial invoices (50 pages per file) using the **Highly available** configuration.

Depending on the page load, ABBYY Vantage required the following amount of resources to efficiently process documents:

Load, (number of pages/8 hours)	Nodes for services (Standard_F8s_v2)	Nodes for workers (Standard_E8_v4)
500	9	1
20,000	10	8
60,000	10	13
100,000	10	15
150,000	10	19
200,000	10	24

🟧 **Note:** When scaling ABBYY Vantage, no increase in document processing time was noted.

Processing 3-page invoices in "Without high availability" configuration

Tests were carried out on 500, 10,000 and 13,000 pages of various commercial invoices (3 pages per file) using the **Without high availability** configuration.

Depending on the page load, ABBYY Vantage required the following amount of resources to efficiently process documents:

Load, (number of pages/8 hours)	Nodes for services (Standard_F8s_v2)	Nodes for workers (Standard_E8_v4)
500	5	1
10,000	7	3
13,000	7	4

Note: When scaling ABBYY Vantage, no increase in document processing time was noted.

Managing Vantage Cluster Nodes

The System Administrator can increase the maximum number of service and worker nodes available to ABBYY Vantage:

1. Sign in to your Azure account and switch to the desired subscription:

```
az login
az account set --subscription subscription_id
```

2. Increase the number of nodes using the following commands:

   o To update the maximum number of nodes in the nodepool for services:

```
az aks nodepool update --cluster-name cluster-name -g resource-group --name linuxpool
--min-count 5 --max-count new-max-count --update-cluster-autoscaler
```

   o To update the maximum number of nodes in the nodepool for workers:

```
az aks nodepool update --cluster-name cluster-name -g resource-group --name lnpool --
min-count 1 --max-count new-max-count --update-cluster-autoscaler
```

To learn more about these commands, see the [Microsoft documentation](#).

Note: The cluster will not scale beyond the number of nodes in the Azure subscription quota.