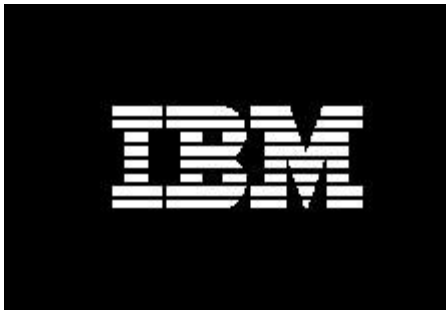

Using the IBM Cloud App Management Database Load Projections Spreadsheet

Document version 2019.4.0

IBM Cloud App Management

CONTENTS

LIST OF FIGURES

REVISION HISTORY

Date	Version	Revised By	Comments
2018/06/27	1.0	JTJ	Original document
2018/11/08	2018.4	JTJ	Updated to include Db2 agent and Cloud Resources data collector
2018/12/14	2018.4.1	JTJ	Updated to include Oracle agent
2019/04/22	2019.2.0	JTJ	Updated to include Linux KVM, Tomcat, VMware VI and Microsoft Hyper-V Server, IIS .NET and SQL Server agents
2019/05/27	2019.2.0	JTJ	Updated to include new runtime data collectors
2019/06/30	2019.2.1	JTJ	Updated to include new agents and Python runtime data collector
2019/07/31	2019.2.1	JTJ	Updated formulas and default values
2019/10/03	2019.3.0	JTJ	Updated to include summarization and new agent types
2019/12/17	2019.4.0	JTJ	Updated to include new agent and data collector types

1 Introduction

This document describes how to use the IBM Cloud App Management Database Load Projections spreadsheet.

The Cloud App Management Database Load Projections spreadsheet was created to estimate the volume of metrics processed by the Cloud App Management server, and to simplify the task of producing a disk space estimate for the metric database. This spreadsheet includes the metric group information for the supported agent types for Cloud App Management 2019.4.0, and allows the user to perform “what-if” exercises to see the database load for different agent environments.

The “metrics per minute” value shown on the *Metric Summary* worksheet shows the projected metric rate based on the spreadsheet inputs.

Monitoring for some agent types causes multiple resources to be created within Cloud App Management. As the number of resources of a given resource type increases, performance can start to degrade in the Cloud App Management 2019.4.0 release. The *Resource Counts* worksheet shows the projected number of resources for each resource type based on the spreadsheet inputs. The *Metric Summary* worksheet shows the resource type with the highest count. When planning the agent environment for a Cloud App Management server, you should try to keep the highest count for any agent resource type below 4,000.

Projections produced by this spreadsheet should be viewed as rough estimates, but should be useful in making configuration planning decisions and in performing sensitivity analysis and what-if exercises. The actual disk storage required for a given monitoring configuration will depend on complex interrelationships among many variables, not all of which have been, or could be, modeled. It is the responsibility of the user to validate the spreadsheet inputs and outputs.

THE TOOL IS PROVIDED ON AN "AS IS" BASIS. IBM CORP. DOES NOT GUARANTEE THE PERFORMANCE OF THE TOOL OR THE RESULTS CALCULATED BY THE TOOL.

2 How the spreadsheet works

The spreadsheet is made up of four worksheets. For all of the worksheets within the spreadsheet, input cells are shown with a green background. Cells showing calculations based on the input parameters are shown with a yellow background.

The spreadsheet consists of the following worksheets:

- The *Read Me* worksheet describes the spreadsheet and the limitations of its use.

- The *Metric Summary* worksheet is the main worksheet, showing a list of agent and data collector types. The user enters the number of managed resources in their environment for each agent type. For certain agent and data collector types where the number of rows written is highly variable, additional Size parameters can be specified to reflect the size of the monitored environment. For example, for the MQ agent, the Size parameter is used to specify the expected number of queues monitored by each queue manager. Yellow cells show summary calculations based on input parameters on the Summary and Details worksheets.

- The *Metric Details* worksheet lists all of the agent tables (or metric groups) in alphabetical order by agent product code (which is shown in the first column on the *Metric Summary* worksheet). The "Metric group rows per interval per agent/data collector" column can be used to specify the expected number of rows per collection sample for each metric group. Default values are provided for most metric groups, based on values observed in test environments or expected values. Yellow cells show Size parameter values referenced from the *Metric Summary* worksheet, as well as calculated values for each agent table.

- The *Resource Counts* worksheet shows the expected number of resources created in the Cloud App Management server for each agent and data collector type. As the number of resources of a given type grows large, performance can start to degrade. On the *Resource Counts* worksheet, if more than 4,000 resources are expected for a given resource type, the resource count is shown in red and a warning message appears.

To move from one worksheet to another, click one of the tabs along the bottom of the spreadsheet, which are shown in Figure 1 below.

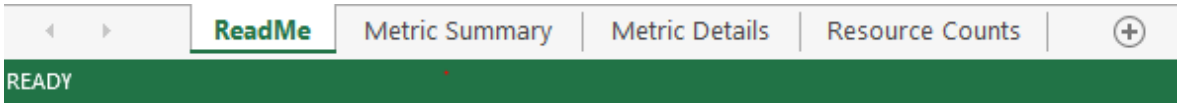

Figure 1 Worksheet tabs appear at the bottom of the spreadsheet

2.1 Typical Usage Scenario

In a typical usage scenario, the user brings up the *Metric Summary* worksheet. The user specifies the expected number of agents or data collectors for each type. For types that have Size parameters, the user specifies a value that reflects the size of their monitored environment.

No further input is required. If desired, the user can use the *Metric Details* worksheet to specify an expected number of rows written per collection sample for each table or metric group.

The yellow cells show calculated values based on the input values specifying the number of agents and the Size parameters.

On the *Metric Summary* worksheet, the three most important estimated values are shown at the top:

- *metrics inserted per minute*

- *GB disk usage (raw metrics)*

- *GB disk usage (summarized metrics)*

These values are important to consider in planning the hardware for running Cloud App Management. If there are more than 4,000 resources of any resource type, the resource count is shown in red and a warning message appears.

2.2 Explanation of Metric Summary worksheet

The *Metric Summary* worksheet has input values and calculated values for each agent and data collector type supported in Cloud App Management 2019.3.0, and summary calculations for the estimated data volume.

2.2.1 Summary results by agent type

Figure 2 shows the table of agent types in the *Metric Summary* worksheet.

		Input parameters			Results by agent/data collector type							
Product Code	Agent / data collector type	Number of agents / data collectors	Size parameter	Size parameter description	% of total metrics	Upload bytes per second per agent	Upload bytes per second total	Raw metrics inserted per minute	Summarized metrics inserted per hour	Total raw metrics retained in DB	Total summarized metrics retained in DB	Total resources
Agents												
AK	Azure Compute		100	Number of virtual machines	0.0%	0	0	0	0	0	0	0
AL	Amazon ELB		30	Number of load balancers	0.0%	0	0	0	0	0	0	0
B5	Amazon EC2		100	Number of instances monitored in configured region	0.0%	0	0	0	0	0	0	0
V6	Cisco UCS				0.0%	0	0	0	0	0	0	0
VD	Citrix VDI				0.0%	0	0	0	0	0	0	0
BN	DataPower		42	Domains monitored by agent	0.0%	0	0	0	0	0	0	0
DT	DataStage				0.0%	0	0	0	0	0	0	0
UD	Db2		200	Tables monitored by agent	0.0%	0	0	0	0	0	0	0
H8	Hadoop				0.0%	0	0	0	0	0	0	0
HU	IBM HTTP Server				0.0%	0	0	0	0	0	0	0
JE	JBoss		50	Servlets monitored by agent	0.0%	0	0	0	0	0	0	0
QI	IBM Integration Bus		200	Message flows for all brokers monitored by agent	0.0%	0	0	0	0	0	0	0
V1	Linux KVM		30	VMs monitored by agent	0.0%	0	0	0	0	0	0	0
LZ	Linux OS	10	8	Filesystems monitored by agent	9.6%	2,904	29,035	2,568	0	29,583,360	0	10
				Average number of docker containers per Linux OS								
MJ	MariaDB		100	Tables monitored by agent	0.0%	0	0	0	0	0	0	0
3Z	Microsoft Active Directory				0.0%	0	0	0	0	0	0	0
Q5	Microsoft Cluster Server				0.0%	0	0	0	0	0	0	0
EX	Microsoft Exchange Server				0.0%	0	0	0	0	0	0	0
HV	Microsoft Hyper-V Server		100	VMs monitored by agent	0.0%	0	0	0	0	0	0	0
Q7	Microsoft IIS		10	IIS Applications	0.0%	0	0	0	0	0	0	0
QE	Microsoft .NET				0.0%	0	0	0	0	0	0	0
MO	Microsoft Office 365				0.0%	0	0	0	0	0	0	0
QP	Microsoft SharePoint Server				0.0%	0	0	0	0	0	0	0
OQ	Microsoft SQL Server		200	Tables monitored by agent	0.0%	0	0	0	0	0	0	0
KJ	MongoDB				0.0%	0	0	0	0	0	0	0
SE	MySQL				0.0%	0	0	0	0	0	0	0
NU	NetApp Storage				0.0%	0	0	0	0	0	0	0
RZ	Oracle Database		200	Tables monitored by agent	0.0%	0	0	0	0	0	0	0
PN	PostgreSQL				0.0%	0	0	0	0	0	0	0
ZR	RabbitMQ		200	Queues	0.0%	0	0	0	0	0	0	0
S7	SAP HANA				0.0%	0	0	0	0	0	0	0
SA	SAP Applications		50	RFC destinations monitored	0.0%	0	0	0	0	0	0	0
SV	SAP NetWeaver Java Stack				0.0%	0	0	0	0	0	0	0
UY	Siebel		1000	Log messages captured	0.0%	0	0	0	0	0	0	0
QL	Skype for Business Server				0.0%	0	0	0	0	0	0	0
OT	Tomcat		30	J2EE applications	0.0%	0	0	0	0	0	0	0
OY	Sybase				0.0%	0	0	0	0	0	0	0
UX	Unix OS	10	15	Physical disks monitored by agent	23.3%	13,875	138,746	6,230	0	71,769,600	0	10
VM	VMware VI		80	VMs monitored by agent	0.0%	0	0	0	0	0	0	0
WB	WebLogic				0.0%	0	0	0	0	0	0	0
YN	WebSphere Applications		1	Application servers monitored by agent	0.0%	0	0	0	0	0	0	0
			53	EJBs monitored by agent								
			80	Servlets/JSPs monitored by agent								
D0	WebSphere Infrastructure Manager				0.0%	0	0	0	0	0	0	0
MQ	WebSphere MQ		250	Queues per queue manager (small <=200, medium < 2000, large < 5000, extra large > 5000)	0.0%	0	0	0	0	0	0	0
NT	Windows OS	10			33.8%	6,756	67,561	9,032	0	104,048,640	0	10
Data collectors												
K8	Kubernetes	1	195	Kubernetes pods monitored by data collector	33.3%	12,495	12,495	8,912	0	102,666,240	0	492
J2	J2SE				0.0%	0	0	0	0	0	0	0
LD	Liberty		600	Total requests per minute	0.0%	0	0	0	0	0	0	0
NJ	Node.js		600	Total requests per minute	0.0%	0	0	0	0	0	0	0
PY	Python		600	Total requests per minute	0.0%	0	0	0	0	0	0	0
KM	Ruby		600	Total requests per minute	0.0%	0	0	0	0	0	0	0

Figure 2 Sample summary results by agent / data collector type

- *Product Code* (column A) is the two digit code for the agent type.
- *Agent / data collector type* (column B).
- *Number of agents / data collectors* (column C) is an input value that specifies the expected number of agents or data collectors of that type in the monitored environment.

- *Size parameter* (column D) is an input value that is used by certain agent and data collector types that can generate a potentially large number of agent table rows for each collection sample. For example, the MQ agent generates a row for every queue managed by the monitored queue manager. Sample values are given for the agent / data collector types that have a Size parameter specified. The user should modify these values to match the expected values for their monitored environment.

- *Size parameter description* (column E) applies to the *Size parameter* (if used) for the agent / data collector type.

- *% of total metrics* (column F) is a calculated value showing the relative percentage of *Metrics inserted per minute total* (column I) for this agent / data collector type vs. the total for all agent / data collector types.

- *Upload bytes per second per agent* (column G) is a calculated value based on the sum across all tables for the agent type of *Upload bytes per second per agent* (column G) from the *Metric Details* worksheet. Additional network activity for heartbeat and protocol overhead is also included. **This value represents the estimated average network overhead due to monitoring on the agent machine**. In this version of the spreadsheet, network activity is not shown for data collectors, but it will be added in future versions.

- *Upload bytes per second total* (column H) is a calculated value for the total network overhead across all agents of the agent type, and is the product of *Upload bytes per second per agent* (column G) and *Number of agents* (column C).

- *Raw metrics inserted per minute* (column I) is a calculated value based on the sum across all tables for the agent / data collector type of *Raw metrics inserted per minute total* (column J) from the *Metric Details* worksheet.

- *Summarized metrics inserted per hour* (column J) is a calculated value based on the sum across all tables for the agent / data collector type of *Hourly summarized metrics inserted* (column L) and *Daily summarized metrics inserted* (column M) from the *Metric Details* worksheet.

- *Total raw metrics retained in DB* (column K) is a calculated value based on the sum across all tables for the agent / data collector type of *Total raw metrics retained in DB* (column K) from the *Metric Details* worksheet.

- *Total summarized metrics retained in DB* (column L) is a calculated value based on the sum across all tables for the agent / data collector type of *Total summarized metrics retained in DB* (column M) from the *Metric Details* worksheet.

- *Total resources* (column M) is a calculated value which is the sum of the resource counts for the resource types created for the agent / data collector from the *Resource Counts* worksheet.

2.2.2 Results summary (total across agent types)

Figure 3 shows sample results of the estimated data volume:

Results summary (total across agent types)			
26,742	metrics inserted per minute		
6.6	GB disk usage (raw metrics)		
0.0	GB disk usage (summarized metrics)		
242	KB per second total network traffic from agents to server		
31	Total agents & data collectors		
522	Total resources		
	Resource type with highest count	Highest count	
	Kubernetes Pod	195	

Figure 3 Sample summary results – estimated data volume

- *Metrics inserted per minute* shows the estimated total metric insert rate across all agent / data collector types, including both raw and summarized metrics. This value is useful in determining the hardware requirements for Cloud App Management.
- *GB disk usage (raw metrics)* shows the estimated total disk space usage for raw metrics across all agent types. This value is the sum of the *Total raw metrics retained in DB* (column M) values times the average space usage per metric (23 bytes), converted to gigabytes.
- *GB disk usage (summarized metrics)* shows the estimated total disk space usage for summarized metrics across all agent types. This value is the sum of the *Total summarized metrics retained in DB* (column N) values times the average space usage per metric (23 bytes), converted to gigabytes.
- *KB per second total network traffic from agents to server* shows the estimated network bandwidth usage (received and transmitted) between Cloud App Management and the agents and data collectors being managed. This is calculated by summing the *Upload bytes per second total* (column H) values, and expressed as KB per second.
- *Total agents & data collectors* is the sum of the *Number of agents / data collectors* (column C) across all types.
- *Total resources* is the sum of the *Total resources* (column C) across all types in the *Resource Counts* worksheet.
- *Resource type with highest count* shows the resource type with the largest expected number of instances as calculated in the *Resource Counts* worksheet. If this value is greater than 4,000, the number is shown in red and the following warning message appears:

  WARNING: more than 4,000 resources of one type not supported in this release

2.2.3 Data retention and summarization settings

Figure 4 shows the input parameter settings for data retention and summarization:

Data retention and summarization settings	
8	Raw data retention (days)
0	Summarization enabled (0=false, 1=true)
60	Hourly summarization retention (days)
0	Daily summarization enabled (0=false, 1=true)
183	Daily summarization retention (days)

Figure 4 Data retention and summarization settings

- *Raw data retention (days)* is the number of days that raw metrics will be kept in the database (the rawMaxDays value). Valid values are between 2 and 32 days, and the default is 8 days. Any value beyond 32 days is not recommended and can degrade Cloud App Management performance.
- *Summarization enabled* specifies whether to perform data summarization.
- *Hourly summarization retention (days)* is the number of days that hourly summarization metrics will be kept in the database. This is the SHORT_TERM_SUMMARY_TTL value, and the default value is 60 days.
- *Daily summarization enabled* specifies whether to perform data summarization into daily values and kept for long-term history. This is the DAILY_SUMMARIZATION_ENABLED value.
- *Daily summarization retention (days)* is the number of days that daily summarization metrics will be kept in the database. This is the LONG_TERM_SUMMARY_TTL value, and the default value is 6 months. In this spreadsheet, we specify the value in days to simplify the calculations.

2.3 Explanation of Metric Details worksheet

The *Metric Details* worksheet has calculated values for each agent table (metric group). Figure 5 shows sample calculated values for Linux OS agent tables (product code LZ).

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Product Code	Agent / data collector type	Table Name (Metric Group)	Metric group rows per interval per agent / data collector	Description of expected metric group rows per interval	Number of agents / data collectors	Upload bytes per second per agent / data collector	Metric group rows uploaded per minute per agent / data collector	Total metric group rows uploaded per minute	Raw metrics inserted per minute	Total raw metrics retained in DB	Hourly summarized metrics inserted	Daily summarized metrics inserted	Total summarized metrics retained in DB
LZ	Linux OS	KLZ_CPU	5	Number of CPU IDs plus 1	4000	112.4	5.00	20,000	140,000	1,612,800,000	140,000	140,000	227,220,000
LZ	Linux OS	KLZ_Custom_Scripts	0		4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Custom_Scripts_Rtm_Smp	0		4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Disk	8	Number of file systems	4000	39.8	1.60	6,400	19,200	221,184,000	96,000	96,000	155,808,000
LZ	Linux OS	KLZ_Disk_IO	5	Number of disk devices in /dev directory	4000	108.4	5.00	20,000	100,000	1,152,000,000	100,000	100,000	162,300,000
LZ	Linux OS	KLZ_Docker_CPU	0	Number of Docker container CPUs	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Info	1	One row per interval	4000	1.3	0.20	800	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_IO	0	Number of Docker container I/O devices	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Memory	0	Number of Docker containers	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Network	0	Number of Docker container network interfaces	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Processes	0	Number of Docker container processes	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Stat	0	Number of Docker containers	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Docker_Version	0	Number of Docker containers	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_IO_Ext	5	Number of disk devices in /dev directory	4000	67.3	5.00	20,000	40,000	460,800,000	0	0	0
LZ	Linux OS	Klz_LFAProfiles	0	Number of LFA profiles	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	Klz_Log_File_Status	0	Number of log files being monitored	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	Klz_LogfileProfileEvents	0	Number of new events	4000	0.0	0.00	0	0	0	0	0	0
LZ	Linux OS	KLZ_Network	5	Number of network interfaces	4000	118.0	5.00	20,000	140,000	1,612,800,000	100,000	100,000	162,300,000
LZ	Linux OS	KLZ_Process	20	Top 5 processes for 4 metrics (up to 24)	4000	1855.5	20.00	80,000	480,000	5,529,600,000	80,000	80,000	129,840,000
LZ	Linux OS	KLZ_System_Statistics	1	Always one row per interval	4000	11.7	1.00	4,000	44,000	506,880,000	8,000	8,000	12,984,000
LZ	Linux OS	KLZ_Vm_Stats	1	Always one row per interval	4000	15.4	1.00	4,000	64,000	737,280,000	60,000	60,000	97,380,000
LZ	Linux OS	Linux_CPU_Config	2	One row per interval	4000	3.3	0.25	1,000	0	0	0	0	0
LZ	Linux OS	Linux_IP_Address	5	Number of network interfaces	4000	9.3	0.63	2,500	0	0	0	0	0
LZ	Linux OS	Linux_Machine_Information	1	Always one row per interval	4000	1.2	0.13	500	0	0	0	0	0
LZ	Linux OS	Linux_OS_Config	1	One row per interval	4000	1.0	0.13	500	0	0	0	0	0

Figure 5 Sample calculations for Linux OS agent tables

- *Product Code* (column A) is the two digit code for the agent type
- *Agent / data collector type* (column B)
- *Table Name (Metric Group)* (column C) is the agent long table name.

- *Metric group rows per interval per agent / data collector* (column D) is an input column. Default values are provided for most agent tables based on observed values on test systems or expected values. These values can be overridden to better reflect the monitored environment. For some agent tables, the Size parameter from the *Metric Summary* worksheet is referenced (for example, the MQ agent Queue Status table), and in those cases, the cell has a yellow background.

- *Description of expected metric group rows per interval* (column E) provides information to help in setting an appropriate value for *Metric group rows per interval per agent* (column D).

- *Number of agents / data collectors* (column F) is referenced from the *Metric Summary* worksheet.

- *Upload bytes per second per agent / data collector* (column G) is a calculated value based on the number of columns and rows of data uploaded for the metric group and the upload interval for the metric group (1 minute in most cases).

- *Metric group rows uploaded per minute per agent* (column H) is a calculated value based on the *Metric group rows per interval per agent* (column D) and the upload interval for the metric group.

- *Total metric group rows uploaded per minute* (column I) is a calculated value based on the *Metric group rows per interval per agent / data collector* (column D), the *Number of agents / data collectors* (column F) and the upload interval for the metric group.

- *Raw metrics inserted per minute total* (column J) is a calculated by multiplying the *Total metric group rows uploaded per minute* (column I) by the number of metrics per metric group row.

- *Total raw metrics retained in DB* (column K) is a calculated value based on *Metrics inserted per minute total* (column J) and the retention period for raw metrics in the metric database (default 8 days).

- *Total raw metrics retained in DB* (column K) is a calculated value based on *Metrics inserted per minute total* (column J) and the retention period for raw metrics in the metric database (default 8 days).

- *Hourly summarized metrics inserted* (column L) is a calculated value showing the average number of hourly summarized metrics inserted per hour into the database.

- *Daily summarized metrics inserted* (column M) is a calculated value showing the average number of daily summarized metrics inserted per day into the database. If daily summarization is enabled, the values in column L and column M will be the same.

- *Total summarized metrics retained in DB* (column N) is a calculated value based on the prior two columns and the data retention values for hourly and daily summarized data from the Metric Summary worksheet.

Trademarks

- Db2, IBM, and WebSphere are trademarks or registered trademarks of IBM Corporation in the United States, other countries, or both.
- Linux is a trademark of Linus Torvalds in the United States, other countries, or both.
- Excel and Windows are registered trademarks of Microsoft Corporation in the United States, other countries, or both.
- Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.
- Other company, product, and service names may be trademarks or service marks of others.

IBM United States of America

Produced in the United States of America

IBM Director of Licensing
IBM Corporation
4205 South Miami Boulevard
Research Triangle Park, NC 27709 U.S.A.