Document to Knowledge™
A CANON COMPANY

IRISPowerscan 11

Watched Folder, Unattended Mode & Server

Revision History

Date	Description	Author
20171016	V1.0 – Creation	TF
20171117	V1.1 – Recommended max number of batches	TF
20180115	V1.2 – Release	TW
20190827	V1.3 – Update for IPS11	TF

I.R.I.S. SA/NV
Rue du Bosquet 10
B-1348 Louvain-la-Neuve

P: +32 (0) 10 45 13 64
F: +32 (0) 10 45 34 43

www.irislink.com

CANON GROUP

Document to Knowledge™
A CANON COMPANY

Table of Contents

1. Introduction 3
2. Definitions 3
2.1 Activities: 3
2.2 Credentials: 3
2.3 Watched Folder: 4
2.4 Unattended Mode: 4
3. Rules of Thumbs 4
3.1 One Core is not enough 4
3.2 The more you put the more you get 5
3.3 One Watched Folder per project 5
3.4 IPS does not transform an old computer into a super server 5
3.5 Size Matters 6
3.6 Key Numbers 7
3.7 Useful Facts 7
4. Legal Notice 9
4.1 Copyrights 9

I.R.I.S. SA/NV
Rue du Bosquet 10
B-1348 Louvain-la-Neuve

P: +32 (0) 10 45 13 64
F: +32 (0) 10 45 34 43

www.irislink.com

CANON GROUP

Document to Knowledge™
A CANON COMPANY

1. Introduction

This technical note summarizes the general rules of thumbs and good practices when using IRISPowerscan (IPS) in Watched Folder mode (WF) or Unattended Mode (UM) as well as with the IPS Server application.

Please read carefully the installation and activation as well as the user guides for correct installation and setup of the software. Please also note that the hardware running it and the anti-virus exclusions are key for proper WF and UM operations.

2. Definitions

2.1 Activities:

IRISPowerscan has 3 main activities, Capture, Validation and Export. These are the main tasks that IPS is carrying out during its execution flow. These activities can be executed manually or autonomously. An activity can be either executed by a Group of Users, the Service or the Sever (in IPS Server App only).

Activities ran by a Group will be done by the IPS app called IRISPowerscan.exe.

Activities ran by the Service will be done by an IPS companion app called IPSxService.exe.

These 2 applications use 2 different credentials, one for IPS and one for the Service (or the Server).

2.2 Credentials:

Credentials are a set of login and password information to enable the use of a computer, an application or a specific target (DMS, Network, Folder…).

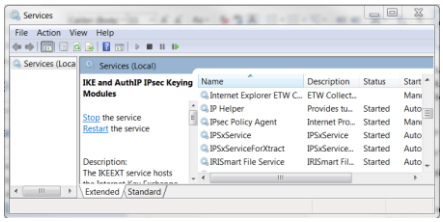
Correct credentials must be given to the IPS and to the Service.

In WF or UM modes, in some cases, you need to change the default IPSxService credentials in order to allow it accessing a target. By default, IPSxService uses Local Admin credentials. And for some targets, like Sharepoint, Therefore or private network accesses, these credentials won't be sufficient.

This applies to the Server as well. You might want to change the IPSxServer.exe to add some specific credentials.

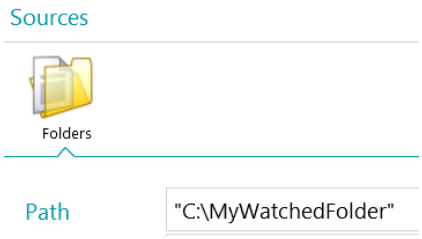
It is advised to the IT department to create a specific user account for these applications. This will avoid having concurrent accesses if one gives his/her own credentials to a service. Doing so will allow IT to guarantee that only the service is accessing the target location.

To change these credentials, simply go to the Services window in Windows, find the correct service and add the correct credentials.



2.3 Watched Folder:

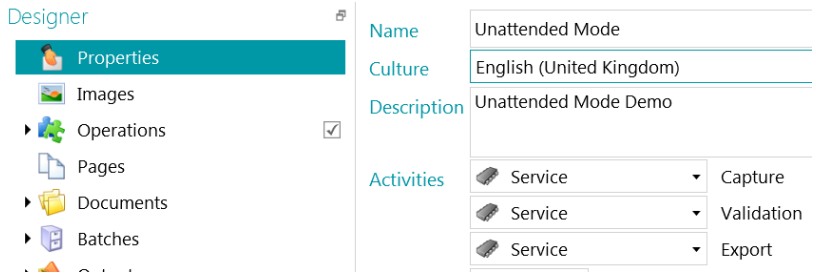
This mode is enabled when we set the Capture activity as a Service. In that way, that project will start "Watching" the folder selected by the input path.



Follow the rules of thumbs described later in this document for correct use of this feature.

2.4 Unattended Mode:

This mode is enabled when all the activities are set to Service.



This mode allows IPS to work in an unattended manner. The IPS project will execute its tasks autonomously until all the document are processed or until a fault condition occurs. Now thanks to the new feature called Warning System, an email can be sent to warn the user that something went wrong. It can be a project error, a batch error, a validation error or a software error.

Follow the rules of thumbs described later in this document for correct use of this feature.

When IPS detect an invalid document or image, the project will stop and hence the complete processing will stop as well. In order to avoid this, the SkipInvalidImages parameter can be used. Please read the documentation to implement this feature.

3. Rules of Thumbs

3.1 One Core is not enough

IRISPowerscan is a multi-threaded and a multi processes application. IPS will use the most of your computer or server. The more resources the best performance you can get.

IPS can start different companion applications or services to improve the overall speed. IPS will calculate the maximum number of such companions based on the number of Cores.

It's like a car, 4 wheels is a minimum to get a high performance car. One could work, but you most likely will decrease the balance and stability.

So the rules for WF or UM is at least 4 Cores. When it comes to virtual cores or vCPU. You have to get at least 4 vCPU. We strongly advise you to have 8 to 16 vCPU for very demanding applications and projects. When you have vSocket and vCores, you can simply multiply the 2 to get the number of vCPU.

3.2 The more you put the more you get

Importing or Exporting PDF documents requires a lot of decompression and compression effort.

If you add some image processing operations in your projects, it will also add a tremendous effort to your resources.

Not only the number of Core has a significant impact but also the computational power of your hardware. High speed processors and data bus will improve you overall performance.

Solid State Disk are also preferred, especially where the application is running. IPS is saving images to cache and on the disk. So the more SSD the faster it will read and write.

The memory access time is also key to get the best performance. The disk IO and memory IO must be as fast as possible. In some shared resources architecture, the IO bandwidth can lead to extremely slow solution and worst to unpredictable random errors. Please make sure that the IOs are fast and not virtualized. For instance, a server working with a NAS system (for RAM and ROM) must be very fast. 1 or 2Gb/s is not enough, it must be around 10 or 20Gb/s.

3.3 One Watched Folder per project

In one IPS project, you can set up to 1 watched folder. You can't put 2 different sources when you use Capture as a Service in one project.

Nevertheless, you can have multiple projects in Capture as a Service, thus allowing you to have multiple WF at the same time.

Pay attention to avoid having 2 projects watching the same folder. This will create random effects and it will most likely disturb you seeing some documents disappearing or getting processed when you think they should not.

It is recommended to not go over 5 watched folder per IRISPowerscan station.

3.4 IPS does not transform an old computer into a super server

IPS WF and UM is designed to handle an average of 10000 pages per day (24 hours).

In some cases, it can be more, in other cases it can be less. It all depends on the complexity of the projects and the number of projects.

A good rule of thumb is one core per project when you want to use IPS in UM.

The documents should not arrive at the same time. If you want to convert 10000 document plus 100 after an hour, you may have to wait for a day before these last 100 documents get processed.

IPS will loop between projects, capture and export to guarantee a continuous flow of document processing. But it also depends on the amount of documents (and their size) that you release in the WF at the same time.

If your documents come from a slow network drive and are pretty big (>200MB), not only IPS will have to process the documents but it also has to download them. When there are hundreds of document like this, it is taking a significant time.

Doing lots of image processing (binarization, color dropout, auto rotation, OCR, BCR, etc) will also affect the speed of the processing. You should optimize your project to do image processing only when it is needed. You can use the page types to improve your project speed.

3.5 Size Matters

Big documents or documents with a high number of pages can drastically reduce the performance.

A big document is a file that is above 200MB. A very large document is above 500MB.

Anytime you are handling such documents, some speed decrease can be expected. In those cases, high speed network and high speed storage disks are more than welcome.

When it comes to page counts, a document with more than 500 pages is considered as very high. This will also decrease the processing speed. And in some cases, it will simply not work. IPS still has a limitation in the addressable fast memory and PDF generation that will prevent it from generating big and high page count document. If you need to work with 500 pages or more, we strongly advise you to work with 200dpi and black and white images. This limitation is known and we will overcome it in future releases.

Image resolution is also an important point. IPS can't handle more than 600dpi due to the image processing capabilities. If you want to process large size documents like A0 or A1, you will need to decrease the resolution.

DPI \ SIZE	A0	A1	A2	A3
600	NOK	NOK	NOK	NOK
400	NOK	NOK	OK	OK
300	NOK	NOK	OK	OK
200	OK	OK	OK	OK

If your documents have not a standard size, the other limitation is 75 Mpixels (75000000 pixels).

If your image is above this value, IPS will not be able to process it.

If one image in a multi pages document is above a threshold, the complete document will be invalid.

3.6 Key Numbers

The Watched Folder feature is designed to handle 5 to 10 watched folder with 1000 pages dropped into them per day. It is not designed to receive 1000 pages per WF per hours.

The Unattended Mode feature is designed to process an average of 10000 pages per day. This can be spread amongst 5 to 10 projects. For heavy duty projects, it is advised to multiply the amount of IPS workstations.

5 is the recommended maximum watched folders for one IPS workstation.

The client server feature or Central Management is designed for 10 clients. The overall page processed is also 10000 per day. It is not limited to that. If you work in more than one shift and have enough processing resources, you may get a higher throughput.

The total number of batches present in the manage view (IPS or IPS Server) should not go too high. For batches with lots of document, the total should be below 300. For batches with few documents, it can go up to 500. We can demonstrate that it will take ~15% more time to scan 500 batches of 10 pages after scanning a first set of 500. This does not apply if you delete the first set of 500 batches.

A good rule of thumb is to keep the total amount of pages in all batches below 5000.

3.7 Useful Facts

Note: the following examples are based on IPS10.

Example 1:

Let's take a simple project and run it on different computers to assess the performance.

Project Details:

- Manual Import of one PDF document of 500 pages with 200 dpi resolution
- Manual Export of the same document as a Searchable PDF iHQC level 1 High Quality
- We measure the duration of the import and the export

Computer Details:

- Laptop: W7, i7 4 Cores, SSD
- Computer: W7, i7 8Cores, SSD

Results:

- Laptop
  - Import in 23' 23"
  - Export in 31' 11"
- Computer
  - Import in 04' 34"
  - Export in 08' 17"

Example 2:

Let's do a UM on 3 different machines with the exact same project and same input documents to compare the speed.

Project details:

- Input from 10 Watched Folder, multiTIFF, one doc of 190 pages released in each watched folder (450 MB is the total input size)
- Output to searchable PDF 1.7.8 Maximum Quality (625MB is the total output size)
- The time it takes to capture, validate and export is measured and the average PPM is calculated.

Results:

- Server 1: Windows Server 2016, 4 Cores, SSD
  - 1h49’ (average PPM = 17.3)
- Server 2: VM Windows Server 2012 4 Cores, HDD
  - 2h21’ (average PPM = 13.5)
- Computer 1: Windows 8.1 8 Cores, SSD
  - 1h07’ (average PPM = 28.3)

Document to Knowledge™
A CANON COMPANY

4. Legal Notice

4.1 *Copyrights*

© 2019 I.R.I.S. All Rights Reserved.

I.R.I.S. SA/NV
Rue du Bosquet 10
B-1348 Louvain-la-Neuve

P: +32 (0) 10 45 13 64
F: +49 (0) 10 45 34 43

www.irislink.com