Performance Pack

Benchmarking with
PlanetPress Connect and
PReS Connect

Contents

Introduction

Every job OL Connect handles goes through several stages: data mapping, content creation, job creation and, finally, output creation. The time needed from data file to output file is called "throughput speed."[1] It depends on many factors, like hardware, the kind of input data, the template, and the workflow complexity. Although throughput is a more representative metric, and is what customers often request or expect, it's very difficult to calculate for each customer's unique combination of factors.

In this document, we compare the throughput speeds of PlanetPress Connect and PReS Connect and explain where it makes sense to use a performance pack to scale their output performance.

To clarify our licensing, OL Connect products have a limit on the last stage, that of output creation.

---

[1] Throughput is generally the number of transactions per second supported by a system. Specifically, in computing it's a measure of a computer system's overall performance in sending data through all its components, such as the processor, buses and storage devices. Throughput is a more meaningful indicator of system performance than the raw clock speed (now measured in gigahertz) advertised by computer vendors, also known as the Megahertz or Gigahertz Myth.
We can calculate throughput using Little's Law:

$$Throughput = \frac{P_n}{DP_t + CC_t + OC_t} \qquad Output\ Speed = \frac{P_n}{OC_t}$$

Where P_n = *No. of Pages*, DP_t = *Data Processing time*, CC_t = *Content Creation time*, OC_t = *Output Creation time*

To determine if you need a performance pack, you have to identify where your bottleneck could lie, in the hardware or the software part of your system. If the software is the bottleneck in your system, then you need a performance pack.
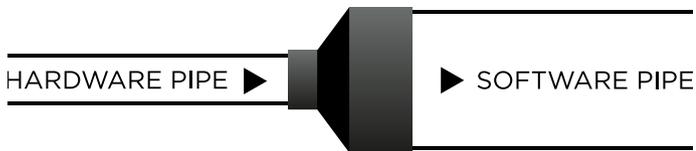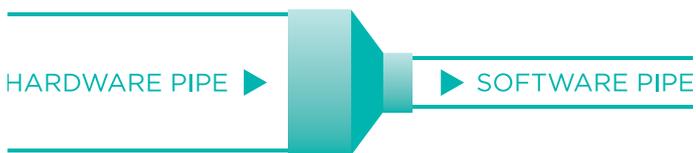

Figure 1 Hardware pipe becomes a bottleneck

Figure 2 Software pipe becomes a bottleneck

Performance packs allow OL Connect products to increase their output speed[2]

	Max. Parallel Instances	x	Speed Limit per Instance	=	Max. Output Speed
PlanetPress Connect	6	x	500 ppm	=	3,000 ppm
PlanetPress Connect + 1 Performance Pack	7	x	1,000 ppm	=	7,000 ppm
PReS Connect	8	x	1,250 ppm	=	10,000 ppm
PReS Connect + 1 Performance Pack	16	x	1,250 ppm	=	20,000 ppm
PReS Connect + 2 Performance Packs	24	x	1,250 ppm	=	30,000 ppm
PReS Connect + 3 Performance Packs	32	x	1,250 ppm	=	40,000 ppm

[2] This output speed limit is a multiple of the number of parallel instances and its basic speed limit. A performance pack will increase the maximum number of parallels and, in some cases, the speed unit, to get to a higher maximum output speed limit.

Maximum Output Speed Limit = Maximum No.Parallels x Speed Unit

Speed is measured in pages per minute (ppm); for print output, this means actual print-ready pages (Postscript, PCL, etc.). For web content creation, this refers to web pages per minute, and for email content creation, emails per minute.

Benchmarking results

We compared the throughput speeds that can be achieved with PlanetPress Connect and PReS Connect and their various performance pack options. In order to make the results of our benchmarking meaningful, we made sure that the use case, data templates, workflows and hardware used were all representative of a realistic scenario. We used a Microsoft Dynamics AX invoice template created in an optimized fashion, XML data input,[3] data to print with All-in-One workflow configuration and hardware commonly available and reasonably priced.[4]

So, with the data, templates and hardware in place, we performed several tests. The average of these tests provide a good indication of the relative impact of a performance pack on **throughput**.

PlanetPress Connect	Throughput (ppm)	Improvement
Standard	2,584	*Baseline*
With a Performance Pack	6,378	147%

PReS Connect	Throughput (ppm)	Improvement
Standard	5,714	*Baseline*
With 1 Performance Pack	11,413	100%
With 2 Performance Packs	14,127	147%
With 3 Performance Packs	16,110	182%

Version 1.5 of OL Connect was used for these benchmarks.

---

[1] This benchmark used our sample XML data and template. If we used PDF as a data source, then times would be much slower due to the additional processing required to extract data from them.
[2] See the appendix for our benchmarking details.

First scenario: Print production on demand

Throughput vs. Output Speed

Printing batches on demand is all about throughput speed because you want the job to arrive **at a specific** printer as soon as possible from the point of job submission. Of course, when you start the job, data mapping, composition and job creation need to occur before the final print stage.

So let's take a look at what that **might** look like using the PlanetPress Connect license and some possible ways to increase the throughput speed.

- Step 3 and step 4 are **not limited** by the license, so the speed is a factor of hardware, data mapping, composition, and possibly even workflow overhead. If these remain constant, then the time needed would be the same **with or without** a performance pack for PlanetPress Connect and PReS Connect (scenarios A, B and C).

- What changes with a performance pack is the **Output Creation Time**.

- Output creation time also decreases if we upgrade to PReS Connect (10,000 ppm). But this results in only a small increase in throughput.

- What could really make a difference is combining a performance pack **with** a review of the computer hardware. In addition, the data mapping, composition and scheduling settings could be optimized to reduce times for data mapping and content creation (Scenario D).

Print On Demand PlanetPress Connect	A. Standard (3,000 ppm)	B. With 1x P. P. (7,000 ppm)	C. If we upgrade to PReS Connect (10,000 ppm)	D. With 1x P. P. (7,000 ppm) & Content creation process *
1. No. of records	5,500	5,500	5,500	5,500
2. No. of pages	6,000	6,000	6,000	6,000
3. Time for data mapping (**not** limited by license)	0.25 minutes	0.25 minutes	0.25 minutes	0.25 minutes
4. Time for content creation (**not** limited by license)	1.75 minutes	1.75 minutes	1.75 minutes	**0.75 minutes**
5. Time for output creation (**license-depended**)	2 minutes	**0.85 minutes**	**0.60 minutes**	0.85 minutes
6. Throughput time (3 + 4 + 5)	4 minutes (1,500 ppm)	**2.85 minutes** (2,105 ppm)	**2.6 minutes** (2,307 ppm)	**1.85 minutes** (3,243 ppm)

* PlanetPress Connect with 1x Performance Pack (7,000 ppm) and content creation process configured to use double the number of merge engines (assumes we had good hardware)

Second scenario: Batch printing

Throughput vs. Output Speed

Generally, batch printing assumes that data mapping and composition are done and accumulated over time, in a lights-out automated process. The final moment to create print output is determined at the last minute or at a predetermined moment (e.g. every Friday at 3 p.m.). This kind of processing might be done in a print room with several printers and the decision to use one printer over another is a function of availability at the time of need.

In this kind of environment, the task is to convert content to print output at the last minute and at the fastest possible speed. This is where the Connect output speed limit really makes sense.

Connect can apply one or more output presets that convert the content into a format compatible with a specific printer. When the output preset is triggered (as part of a workflow), the performance is no longer dependent on the previous steps. So the speed of output creation—assuming the computer hardware is capable—is more likely to be representative.

Batch Printing PlanetPress Connect	A. With 1x P. P. (7,000 ppm)	B. Same as A but larger job	C. If we upgrade to PReS Connect (10,000 ppm)	D. PReS Connect with 1x P. P. (20,000 ppm)	E. PReS Connect with 2x P. P. (40,000 ppm)
1. No. of records	5,500	200,000	200,000	200,000	200,000
2. No. of pages	6,000	250,000	250,000	250,000	250,000
3. Time for data mapping (not limited by license)	Automated and done earlier	Automated and done earlier	Automated and done earlier	Automated and done earlier	Automated and done earlier
4. Time for content creation (not limited by license)	Automated and done earlier	Automated and done earlier	Automated and done earlier	Automated and done earlier	Automated and done earlier
5. Time for output creation (license-dependent)	0.85 minutes	36 minutes	25 minutes	12.5 minutes	6.25 minutes
6. Maximum output speed (limited by license)	7,000 ppm	7,000 ppm	10,000 ppm	20,000 ppm	40,000 ppm

Now, these numbers show maximum output creation limits to make the point and illustrate why we chose to use them. As you can see, performance is easier to predict in this "last-minute" printer environment because we eliminate the overhead of data processing and content creation.

Note: Document complexity (lots of fonts, graphics and layers) and the clustering of overheads (moving files from one server to another) will make things slower.

Third scenario: Web content and email generation

The benchmarks have been carried out with print content, as those are considered to produce the heaviest load on the system. Still, performance packs are relevant to web and email output. For example:

- When the server is requested to generate too many pieces of content per minute, the next request will have to wait. If there are consistently more requests than licensed, the number of requests waiting will keep growing, increasing the wait time with each request. At some point, browsers may time out.

  The time it takes to load a page is considered the number one usability aspect of websites. So waiting times significantly impact the user experience. It is important that the license be sufficient for any peak in requests that may occur at any time.

- Email batches are often scheduled to be sent at an optimal time for the recipient. If the licensed speed, hardware and email service do not allow for submission at the speed needed to process the whole batch within the intended time frame, attention to the received email may suffer.

- The maximum theoretical capacity of a standard PlanetPress Connect license (with no other work assigned) is just over 4 million emails per full day. Any higher demand will require performance packs or an upgrade to PReS Connect. It is important that the license be sufficient for any peak in requests that may occur at any time.

Conclusion

The goal of this white paper was to explain the value of performance packs, and hopefully, that goal has been achieved. However, it is important to know that hardware, networks, and other related systems, play an important part in overall performance. And where these are the limiting factor, a performance pack may not make a difference at all.

Data, document and workflow complexity will affect data processing times and the resulting throughput, especially when printing on demand. Batch printing, where the data processing and content creation are performed before the final output is required, is more about output speed—and that's where performance packs have a more direct impact.

Appendix: Benchmarking details

Use case - Printing invoices

We focused on a standard application: printing of invoices. This is an actual use case that is more challenging than more basic use cases (like static letters, or just counters), and certainly more interesting than benchmarking empty pages.

Input data - XML from Dynamics AX

Instead of the simplest kind of data we could find, we chose data that resembled the output of Dynamics AX. This is XML data with a significant number of fields and repeating details.

For the test, 15,000 records have been generated, divided over 6 files, which the workflow automation tool can do for you. This split over multiple files allows a convenient distribution across parallels and clusters. These records resulted in 18,788 invoice pages.

Template – Invoice template from AX

We used a template that was taken from the standard template for AX. This template has some typical Connect functionality, such as a details table with page overflow, detail lines where the description can overflow to a new line, calculated subtotals and an alternating background colour.

To get the best performance, this template was optimized according to the tips mentioned in the Tech article *Speed Up your Connect template*. It is fair to assume that an optimization like this would be done in cases where performance is key. Creating templates in an optimized fashion can yield a lot of improvement. In the case of this template, we tested a less optimized version of the template as well, and the optimized version performs at least 17% better.

Workflow - Data to print with All-in-One

The workflow configuration is one of the simplest that could be used in real life: the All-in-One plugin is used to run the Connect configuration. It is self-replicated to allow multi-threading of the whole process, which is an easy way to run multiple output engines at the same time. Running processes in parallel implies that the use case will have multiple input files coming in that can be processed in parallel. In cases where the input is a single file, either a clustered setup is required, or the workflow process needs to take care of splitting the file for distributed processing.
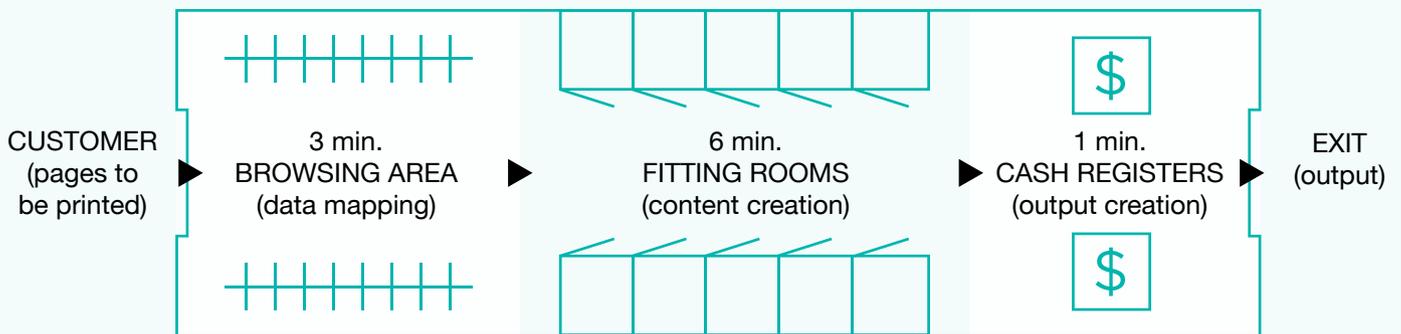
OL Connect version 1.5 was used for the benchmark numbers provided. With each version, we aim to improve performance and scalability.

Servers and hardware

These benchmarks have been carried out on good servers. Each server had the following hardware:

- **Processor: Intel Xeon CPU E5 - 1660 v3**
  This CPU has 8 cores (and hyper-threading, which gives 8 more "virtual cores." It is often written as 8c/16t) running at 3GHz.

- **Memory: 128GB RAM**
  RAM prices are such that we chose to make sure that the RAM available could not pose a bottleneck.

- **Disks: fast SSD drives, 2 400GB drives in RAID0**
  SSD drives pose little extra cost for greatly increased performance. For the benchmarks, the size is of little relevance, but for most real life scenarios, the chosen capacity is expected to be sufficient.

These are fast servers, but they were not exotic at the time of testing (December 2016). The buying price was estimated between €3,000 and €4,000 and could be rented for between €200 and €300 per month. Hardware availability and pricing is subject to continuous change, generally with dropping prices.

CUSTOMER (pages to be printed) ▶ 3 min. BROWSING AREA (data mapping) ▶ 6 min. FITTING ROOMS (content creation) ▶ 1 min. CASH REGISTERS (output creation) ▶ EXIT (output)

Analogy: The clothing store

The exact difference between throughput and output can be a difficult concept to grasp. So if the introduction raised more questions than provided answers, we will use a simple analogy of a clothing store to illustrate the different scenarios. The scenarios illustrate when it makes sense to use a performance pack, or when investment in hardware is needed.

Imagine you have a store that sells clothes. It has 2 cash registers (license-dependent output speed).

The amount of clothes each customer buys (complexity of the job) can have an effect on the time each step in the shopping process takes.

Analogy	Action	Time
Browsing	Data mapping	3 minutes
Trying on clothes	Content creation	6 minutes
Paying	Output creation	1 minute

Total time spent by each customer: 10 minutes

First scenario: Print production on demand

Need for a performance pack
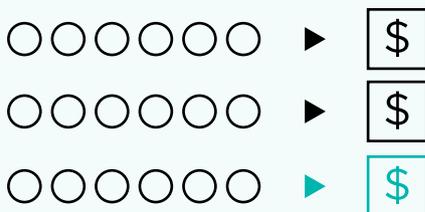
Output scenario – not enough cash registers

- 60 customers come in every hour:
throughput = 1 customer per minute.
You can reach that with your 2 cash registers.

○○○○○○ ▶ $ / $

- 120 customers come in every hour:
throughput = 2 customers per minute.
This is also within your reach with the 2 cash registers.

○○○○○○ ▶ $
○○○○○○ ▶ $

- 180 customers come in every hour:
throughput = 3 customers per minute.
Your 2 cash registers can only handle 2 customers every minute, so you need an extra cash register (performance pack).

○○○○○○ ▶ $
○○○○○○ ▶ $
○○○○○○ ▶ $

Need for faster hardware

Now imagine the store has another potential bottleneck: only 10 fitting rooms (server hardware performance).

Content creation scenario – not enough fitting rooms

- 60 customers come in every hour:
throughput = 1 customer per minute.
A fitting room can handle 10 customers per hour, so the 10 you have can deal with up to 100 and customers will not be delayed.

○○○○○○ ▶ □□□□□□ / □□□□□
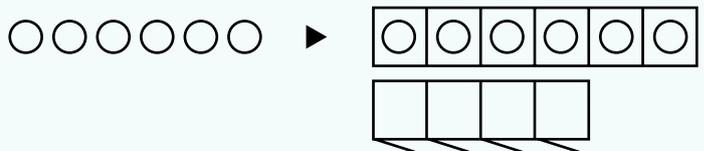
- 120 customers come in every hour:
throughput = 2 customers per minute.
A fitting room can handle 10 customers per hour, so the 10 you have can deal with up to 100, but there are 120 customers, so people will have to line up at the fitting rooms.

○○○○○○ ▶ □□□□□□
○○○○○○ ▶ □□□□
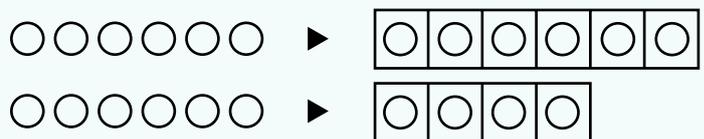
Even if your cash registers could handle the 120 customers, they would only serve 100 because of the limited fitting room capacity.

The other way around: If there are enough fitting rooms to serve 120 customers an hour, but the cash registers cannot keep up—for example, because each customer brings a huge amount of clothes—faster staff would be needed (faster hardware).

Second scenario: Batch printing

Try on clothes
anytime all week

Pay on Friday
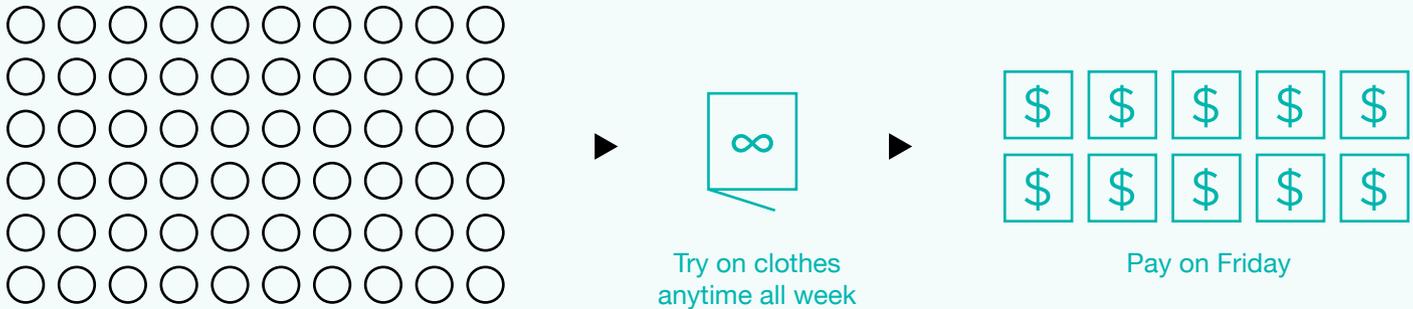
Batching production runs allow for **maximizing performance** of output at **minimum investment** in hardware: Content buffering scenario - separating the store from the cash registers.

Your store is now so popular that it attracts 600 customers every hour. You can get the 10 cash registers in place to handle the throughput, but investment in 60 fitting rooms is not desirable.

You decide to allow people to try on the clothes at any time of the week. Then every Friday at 10 a.m., you open the cash registers for everyone to come and pay. You are now maximizing the output speed of the 10 cash registers at a continuous 600 customers an hour (software license), while not investing in extra fitting rooms (hardware).

If, using this approach, the 600 customers/hour rate is not reached, the cash registers may be operating too slow (faster hardware needed).

If the 600 customers/hour rate is reached but considered insufficient, more cash registers are needed (performance pack).

OL™

objectiflune.com