ArchivistaBox 2013/II: Up to 170,000 pages text recognition per day and ArchivistaBox

Pfaffhausen, 14 February 2013: *Rarely has Pfaffenhausen seen itself as submerged in snow as it was this year. But this has not stopped us from our rapid development of the ArchivistaBox. With our ArchivistaBox 2013/II we have been able to integrate fast A3 Fujitsu scanners, which again have a new parallel text recognition function.*

Fast A3 scanners need not be exorbitantly expensive

To date, our scan boxes have mostly been supplied with Fujitsu document scanners, which have a throughput of between 20 and 40 pages per minute. As a result of a customer enquiry, the question was raised as to whether an even faster document scanner should be used. We do indeed deploy extremely fast document scanners (e.g. Canon DR9080), but all of these devices cost a minimum of a five-figure sum. For this amount of money, we could get 3 scan stations with a 40-page scanner; in short, fast document scanners don't really fit the price/cost ratio of ArchivistaBox systems (even the **ArchivistaBox Matterhorn costs under CHF 10,000** in its basic configuration).

Fujitsu fi-6×70 series document scanners offer an ideal alternative. These devices **are priced far under CHF 10,000,** but in terms of performance, can **easily process a 10,000-sheet mountain of paper, per day of course.** The rate per minute is 80 A4 sheets (160 pages for scanners), both in black and white and colour, and at 300dpi. Thus 4800 sheets can be processed per hour. And another thing: fi-6×70 scanners can also handle A3 documents.

Parallel text recognition of up to 170,000 pages per day per Box

In order for no "backlog" to occur during the text recognition process, we have programmed the text recognition function so that the scan jobs are distributed in parallel across all of a machine's processors (CPUs). The new ArchivistaBox Rothorn has, for example, six processors. In this way, **approximately 7000 pages of text can be recognised per hour,** i.e. up to a maximum of 170,000 pages in 24 hours.

Of course, several scanning stations can be used next to each other, allowing daily volumes to be extended almost at will. And yes, of course, the text recognition jobs could also be split across the 32 or 48 CPUs of a rack machine, as required for specific projects. The **entire text recognition process occurs exclusively using the main memory, so even colour pages can be processed at high speed.**

New showroom/training room at Pfaffhausen site, displaying all our products

We are delighted to be able to demonstrate all our ArchivistaBoxes (both DMS and virtualisation) "live in action" in our new showroom/training room. We would invite you to stop by, bringing your documents. Together we will find the ArchivistaBox + document scanner which is best for you.

In our showroom/training room you will find not only the DMS products, but also all ArchivistaVM models. We'll be happy to perform the automatic construction of a DRDB cluster before your very eyes. The new showroom can also be used as a training room. Up to 8 participants can be trained at the same time. At a later date, we will be offering attractive training courses. The ArchivistaBox systems are of course extremely easy to use, but as part of our day-to-day support, we've noticed time and again that even experienced users are not using some great features because they are relatively unknown. Which is why we'd like to offer some attractive training opportunities and WorkDays. More on that in a future blog.

Whether attending a demo or training, we always look forward to welcoming you at our offices. We would **invite you to one of our OpenFridays – more information about these can be found here.** If you'd like to receive some more information about our products at this point, you can find the updated documentation regarding the ArchivistaBoxes 2013 in the "Downloads" section.