Final report on preservation scenarios deployed and integrated with data centres

Authors

Ondrej Klima, Pavel Smrz (Brno University of Technology), Tomasz Parkola (Poznań Supercomputing and Networking Center)

September 2014

Executive Summary

The deliverable contains two main sections. Each section corresponds to one experimental context. The first focuses on large-scale visual geo-localisation in natural environments and on annotation of natural environments. The experiments reported in this section investigate spherical cross-correlation aiming at selecting candidate solutions for camera pose estimation. The preservation scenario collects various parameters characterizing actual runs of the experiments and particular characteristics of the external data centre. Memory consumption, computation time, and quality of camera pose estimation are reported for different combinations of input and output resolution of the spherical cross-correlation. It is shown that the consumption of computation time and system memory grow dramatically with the increase of output resolution. Having fewer resources available during a particular run of the experiment in a different data centre, one can expect completely different quality of the output. Thus, it is necessary to preserve the profiling information together with other experimental characteristics.

The second part refers to medical data scenarios that were deployed and integrated with the PSCN data centre. The medical data processing involves procedures of data ingestion, access, anonymization, and analysis. Key features of individual procedures are discussed together with overall performance figures.

Table of Contents

1   Introduction

This deliverable presents results of large-scale experiments running in external data centres. We apply the preservation methodology developed in deliverable D23.1 that addresses reproducibility of results and preservation of experiment environments and circumstances. Using the software tools that take advantage of the SCAPE platform, we collect various characteristics and show their relations to the overall results in particular experiments.

Executable workflows for the first set of experiments that deal with large-scale video semantic analysis were introduced in deliverable D23.1. A general workflow characterizing the use of the experiment preservation toolkit is reproduced in **Figure 1**. Readers are referred to specific sections of the previous deliverables to find workflows corresponding to individual experiments analysed in the following sections.
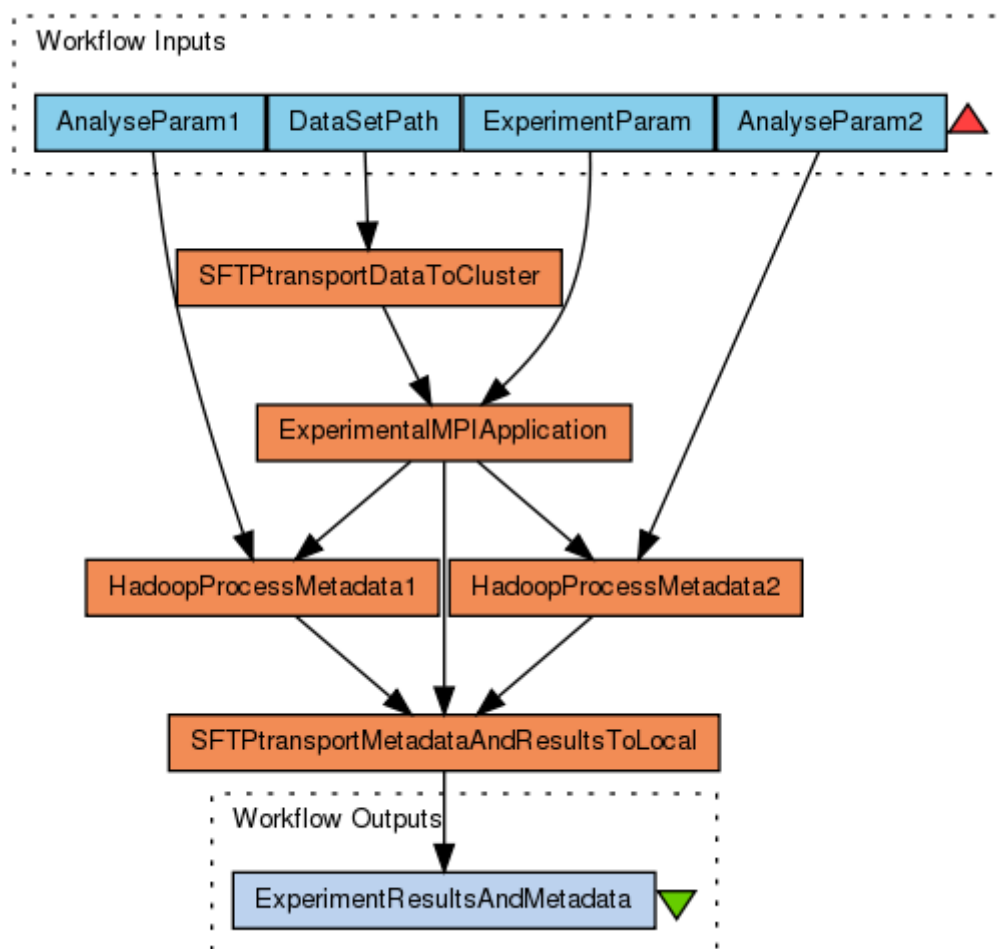

Figure 1: General workflow using the experiment preservation toolkit.

Reported large-scale video processing experiments aim at deriving 3D information from individual frames. Experimental applications were created to validate previously proposed executable workflows. This deliverable presents results based on an application that geo-localises the camera in natural environments and annotates known points of interest (e.g., mountains) in images.

The second part demonstrates preservation processes applied in the context of medical data. It shows critical components of the deployed solution in four specific steps – data ingestion, access, anonymization, and analysis. Reported results prove that the extended SCAPE preservation platform can cope with requirements typical for the medical domain.

The rest of this document is organized as follows: The second section focuses on the video annotation experiment. The main goal is to demonstrate that the experiment preservation toolkit is able to collect relevant profiling data that significantly influence results. In particular, we investigate relations between specific parameters of the geo-localisation method, the consumption of the HPC resources and the quality of localisation results. The beginning of the section introduces the video annotation application and discusses its relation to large-scale geo-localisation. A brief introduction to the methodology of camera localisation is presented next. Particular characteristics computed from the data collected by the preservation tool are divided into three subsections. The first summarizes values of the memory consumption w.r.t. parameters of the localisation method. The measurements are provided for seven different stages of the experimental application. The second subsection characterizes computation times of five specific parts of the algorithm. The last subsection focuses on the quality of the geo-localisation results depending on various input parameters. A summary of observed relations between various parameters and the quality of results concludes the section.

Section 3 presents medical data scenarios deployed and evaluated as a part of the joint work of WCPT and PSNC. Individual subsections correspond to key identified problems and related requirements – medical data ingestion necessary to store the data externally, data access through defined protocols, access to anonymized data for experts and students, and medical data analysis supporting research activities based on the stored data.

The last section summarizes all the results and concludes the deliverable.

2   Video annotation

Video annotation represents one of newly added SCAPE testbeds stressing the fact that some experiments and information processing run in external data centres that need to be included in relevant preservation scenarios. Obviously, not all graphics research institutions can have and run own high-performance computing (HPC) systems. To deal with large-scale processing, they use services of external GPU clusters provided by other institutions.

Hardware and software configuration of the HPC system used, as well as its current workload and other processing parameters may directly affect results of computer vision experiments performed. Configuration differences can make it impossible to repeat previous experiments or seriously impact their results. To guarantee preservation of data centre experiments, it is necessary to pay attention to a large set of characteristics of the cluster environment during experiment executions.

Parameters of the environment which have been identified as critical for the experiments reported in this section are measured and stored as the experiment metadata linked to preserved input and output data. Critical parameters include the amount of available memory and the number of allocated computation nodes, but also various characteristics of each current run of an experimental application, such as the memory usage and the computation time related to particular input and output parameters. The experimental metadata is stored in the form of structured files and its processing is performed using the SCAPE platform.

Large-scale experiments in external data centres involve several steps. Initially, an experimental application is installed on an external cluster and an experimental data set is uploaded to the HPC. The experimental application is executed and critical parameters are measured. Resulting metadata are then processed on the SCAPE platform, linked to the input and results and transferred back to the involved research institution, where they are archived in a local storage facility. In the case of the video annotation scenario, this process was examined using a special purpose application implemented by BUT using the infrastructure of the UVT HPC.

2.1   Image-based geo-localization

The experimental application focuses mainly on localization of captured images in mountain environments. Geo-localisation of images determines the position and orientation of a camera used to take the images. The localization method was introduced by L. Laboud and M. Čadík in their paper "Automatic Photo-to-Terrain Alignment for the Annotation of Mountain Pictures". It is based on fitting a real photo to a digital elevation model of mountains. The original paper assumes known longitude and latitude. The experiment reported in this deliverable extends and improves the previous solution, recovering the longitude and latitude automatically from the captured scene. Once the camera position and orientation is recovered, annotation of the image (for example adding mountain names as captions to the photo) or image enhancements can be performed.

The geo-localization is performed by comparing a real world image with a database of panorama images synthetically rendered from a digital elevation model. Once a corresponding image is found in the database, the localization of a query image is recovered. Comparing images of two different modalities (a real world photo and a rendered panorama image) presents a challenging task due to lack of similar features in both images. The method employs edges as the compared features. There

is typically a huge amount of noise in the real world image so that special metrics which account for the noise need to be applied.

In particular, two different metrics are brought into play. The first one is a special spherical cross-correlation, which can be computed relatively quickly in the Fourier spectrum. As it provides only approximate results, it is used for initial fast selection of a set of candidate solutions. Once they are known, a precise but slow metrics is applied for the fine alignment of a real world photo to a synthetic panorama.

Computation of the spherical cross-correlation depends mainly on two parameters – an input and an output resolution. The input resolution determines sampling frequency of edges in both input images. The result of the cross-correlation is a 3D raster block of values. Each point in the block contains a value for one concrete camera pose, represented by three angles. Coordinates of a point in the 3D block correspond to values of camera yaw, pitch and roll. The output resolution then determines how many points are present in the resulting 3D block (it cannot be higher than the input resolution).

As shown in the following section, the mentioned characteristics have a significant effect on resource consumption and the quality of results. The reported experiments were performed for varying combination of resolutions as shown in Table 1 and Table 2.

No.	1	2	3	4	5	6	7
Input resolution	64	128	128	256	256	256	512
Output resolution	64	64	128	64	128	256	64

Table 1: Combinations of input and output resolution used in experiments

No.	8	9	10	11	12	13	14
Input resolution	512	512	512	1024	1024	1024	1024
Output resolution	128	256	512	64	128	256	512

Table 2: Combinations of input and output resolution used in experiments - the second part

2.2   Scenario implementation

Taverna workflow was used for modelling the video annotation scenario (see Figure 2). The scheme can be divided into two main parts. An experimental application performing localization and annotation is executed first. Input and output parameters as well as metadata identified as critical for preservation are gathered in this part of the workflow and stored as a set of structured files. The first part is executed on a GPU cluster as the experimental application takes advantage of the GPU-accelerated OpenGL.
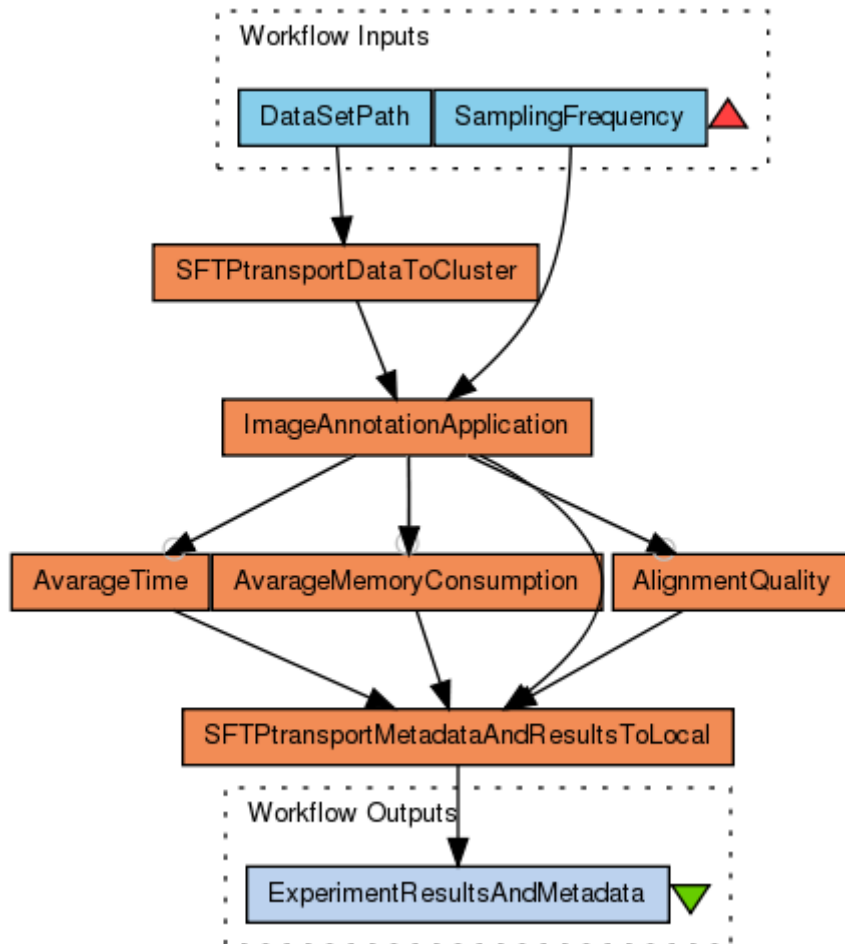
Figure 2: Taverna workflow for video annotation scenario.

The second part focuses on preservation and further processing of the structured files prepared and stored in the previous step. The analysis of the collected information benefits from the map-reduce paradigm (using the Apache Hadoop) being a part of the SCAPE platform. This also guarantees scalability of the preservation part of the workflow.

Measurement of the resources consumption
Information about the memory consumption is obtained by parsing the standard `/proc/self/stat` file. The field containing the resident set size (RSS) – the amount of memory that is actually held in RAM – is extracted from the file and stored in the experiment metadata file.

Measurement of the computation time consumption
Time consumption is measured using the boost library, using the provided timer module. Measurements focus on the wall time, which is stored in nanoseconds (although the format deals with nanoseconds, the real precision may vary depending on the current platform).

2.3 Remote data centre infrastructure

The first part of the workflow was executed at the InfraGRID GPU cluster provided by UVT. Technical details related to the hardware infrastructure of the GPU HPC are described in Deliverable D4.3 (Final Data Centre Deployments) in Table 5. To summarize just key parameters, the GPU cluster contains 7 computation nodes and 2 head nodes. Each computation node is equipped with NVidia Tesla M2070Q (448 cores, 6GB GDDR5), 32GB RAM and 6-core CPU Intel XEON 3.46 GHz. Jobs are scheduled using the IBM Load Leveller. There is an 18TB storage system available.

The second part of the workflow is executed on UVT Hadoop platform, described in Deliverable D4.3, Table 3. Connections between these two platforms are illustrated in Figure 3.
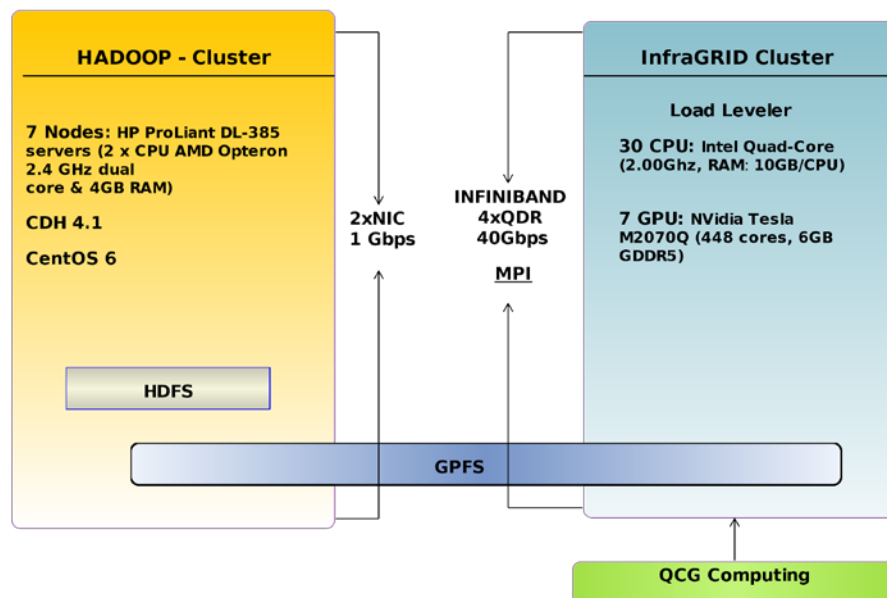

Figure 3: UVT InfraGRID GPU cluster, Hadoop platform and connections between them

The M2070Q can be used beyond the general computation even for the hardware accelerated offline graphics rendering using the OpenGL interface (the term "offline rendering" is used for rendering the graphics into pixel buffers in system memory instead of rendering on physical screen, as the M2070Q card has no output for such a device). To be able to use the offline rendering in the experimental application incorporated in the workflow, it was necessary to install the headless X servers on the computation nodes and adjust them with the IBM Load Leveller. It was also necessary to install the Qt Toolkit SDK in version 4.8.2 as the experimental application was compiled at the cluster head nodes. The scenario also required OpenMPI, ImageMagick and OpenCV libraries to be installed in the cluster environment.

2.4 Evaluation of memory and time consumption

The memory and time consumption was measured in several stages of the application which was executed in parallel on seven computation nodes equipped with NVidia M2070Q card. One instance of the application was executed on each node. Parallelization was handled by the OpenMPI framework. The following algorithm steps, described in the following subsections, were evaluated:

- Mapping and sampling of synthetic panorama image (Cylmap, cyltime)
- Sampling real-world query image (Projmap, projtime)
- Initializing data structures for Spherical Fourier Transform computation (Vscc2)
- Computation of Spherical Fourier Transform (Vscc3)
- Finalizing the SFT computation (Vscc4)
- Spherical cross-correlation computation (Vscc5)
- Finalizing the spherical cross-correlation computation (Vscc6)

2.4.1 Mapping and sampling of synthetic panorama image (Cylmap, cyltime)

The panorama images in the particular data set correspond to a cylindrical projection. It is therefore necessary to re-project the panoramas to spheres first. The re-projection is performed together with sampling the panorama image according to the input resolution. This part of the algorithm takes advantage of OpenGL – it is accelerated by the graphics card. The core of the projection uses fragment and vertex shaders. The projection and sampling is applied separately to horizontal and vertical coordinates of the cylindrical panorama. As this step cannot rely on the output resolution, the measurements were performed only for four different values of input resolution, as shown in Table 3.

Resolution	Average memory (MB)	Average time (s)
64	129.1835132841	1.06192280914
128	129.7335191362	0.79248599016
256	132.0075711025	0.882579440430
512	140.9612410756	0.933233379915
1024	176.9763284176	1.181937854604

Table 3: Average values of resources consumption for panorama sampling at different resolutions

2.4.2 Sampling real-world query image (Projmap)

The real-world images are saved in the data set with the perspective projection, so, as in the previous case, it is necessary to re-project the extracted edges with the spherical projection and to re-sample them according to the input resolution. This computation is again realized by OpenGL vertex and fragment shaders. As the real-world images are usually much smaller than the synthetic panorama images, the consumption of memory and time is significantly smaller than in the previous case, as it can be seen in Table 4**.**

Resolution	Average memory (MB)	Average time (s)
64	11.0278045997	0.116745441272
128	11.5883734199	0.078510309132
256	13.8378638030	0.172573475490
512	22.8375065835	0.329361789809
1024	58.8375193849	0.791043441075

Table 4: Average values of resources consumption for real-world image sampling

2.4.3 Initializing data structures for Spherical Fourier Transform computation (Vscc2)

Objects necessary for the Spherical Fourier Transformation are created and corresponding amounts of memory are allocated in this stage of the algorithm. The spherical cross-correlation is computed in the Fourier spectrum to reduce the time of computation. The initialization depends on the input resolution only. Results of measurements are listed in Table 5.

Resolution	Average memory (MB)	Average time (s)
64	0.813831343632	0.0146066062909
128	3.722546523873	0.0564569067727

256	25.49216311645	0.3418248613030
512	187.2805550678	2.7215382263348
1024	1431.425993387	22.010801578645

Table 5: Memory (MB) and computation time (s) consumption during the SFT initialization

2.4.4 Computation of Spherical Fourier Transform (Vscc3)

The Spherical Fourier Transform (SFM) is computed next. The data structures initialized in the previous step remain allocated. Additional memory is necessary to store the result of the SFM. The total memory and time consumption is characterized in Table 6.

Resolution	Average memory (MB)	Average time (s)
64	1.153631203183	0.039010355363
128	5.852286692415	0.022284067963
256	32.48971012796	0.095363105321
512	212.2594437617	0.785934197904
1024	1528.518982619	3.831932308855

Table 6: SFT computation resources consumption

2.4.5 Finalizing the SFT computation (Vscc4)

The initial data structures for the SFT computation are released in this stage; only the SFT result remains in the memory. Table 7 proves a low profile of the resulting structure. This is the last stage depending on the input resolution.

Resolution	Average memory (MB)
64	0.307335557116
128	0.441186797752
256	2.016098041513
512	8.017483029026
1024	32.02155021067

Table 7: Memory requirements (MB) of SFT results

2.4.6 Spherical cross-correlation (Vscc5)

The cross-correlation is computed in this step. As shown in Table 8, the memory consumption and the computation time consumption significantly increases with the increasing output resolution. The same is true for the processing time.

Resolution	Average memory (MB)	Average time (s)
64	40.5792749297753	0.38361803478181
128	266.628094276685	4.93941695744545
256	2062.02104303215	55.3110679656303
512	16404.0245274462	673.899323478682

Table 8: Spherical CC computation resources consumption according to different output resolutions

2.4.7 Finalizing the spherical cross-correlation computation (Vscc6)

The memory allocated for the results of the Spherical Fourier Transform is finally released in this stage. Results of the cross-correlation are still held in the memory for the localization of maximal values. Measurements are summarized in Table 9. This is the second measured stage that depends on the output resolution.

Resolution	Average memory (MB)
64	32.0581314372659
128	256.005215648408
256	2048.00544485409
512	16384.0068249649

Table 9: Finalization of CC memory requirements (MB)

2.4.8 Comparison of memory and time consumption related to input resolution

The graph in **Figure 4** compares the memory consumption of different parts of the algorithm that depend on the input resolution. It can be seen that the consumption grows non-linearly with the output resolution. The highest memory consumption is observed in the vscc3 and the vssc2 steps.
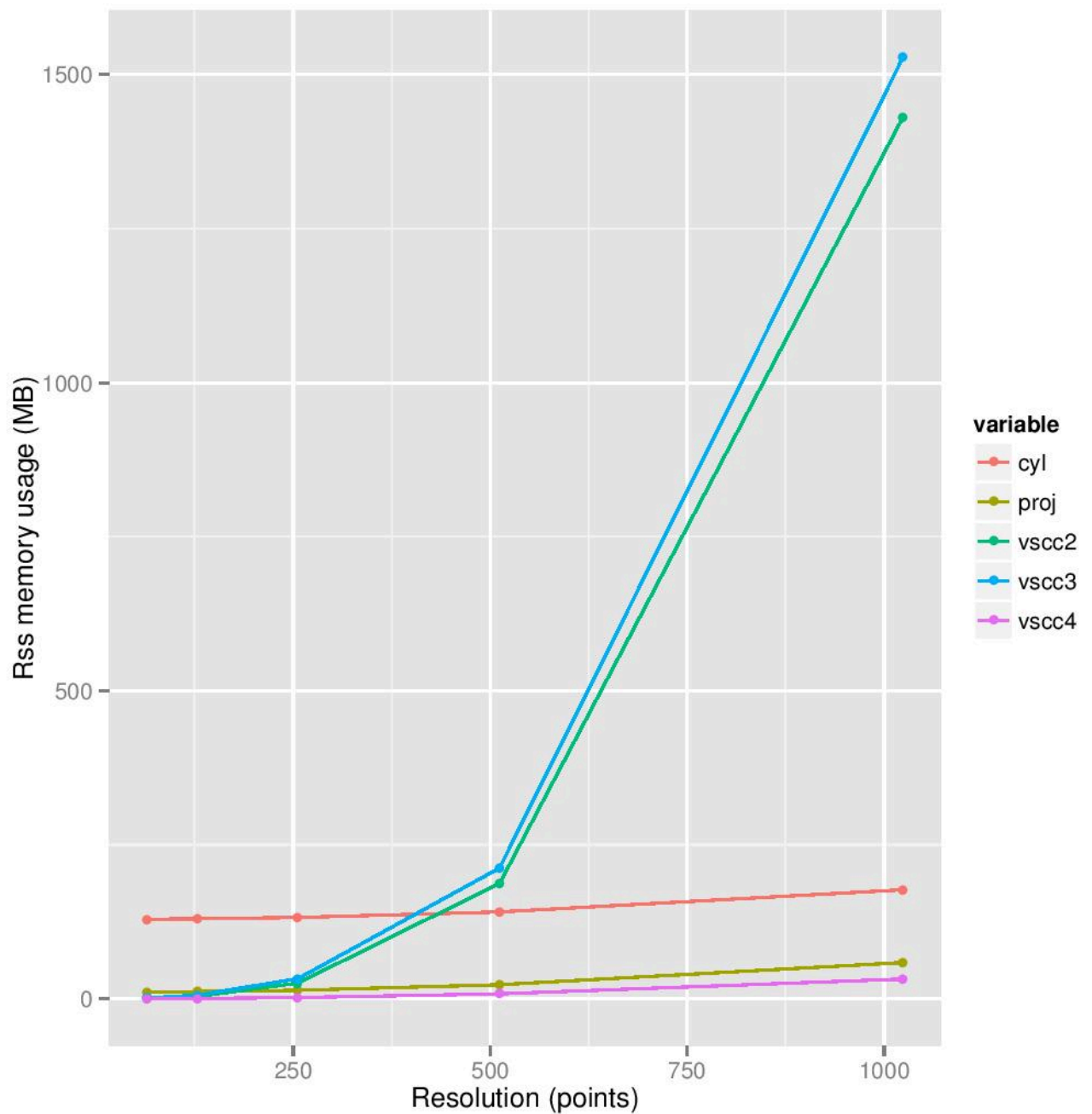
Figure 4: Comparison of memory consumption related to input resolution.

Figure 5 then compares the computation time consumed by different stages of the algorithm that depend on the input resolution. It can be seen that the most demanding part of the cross-correlation algorithm is initialization of data structures for Spherical Fourier Transform computation.
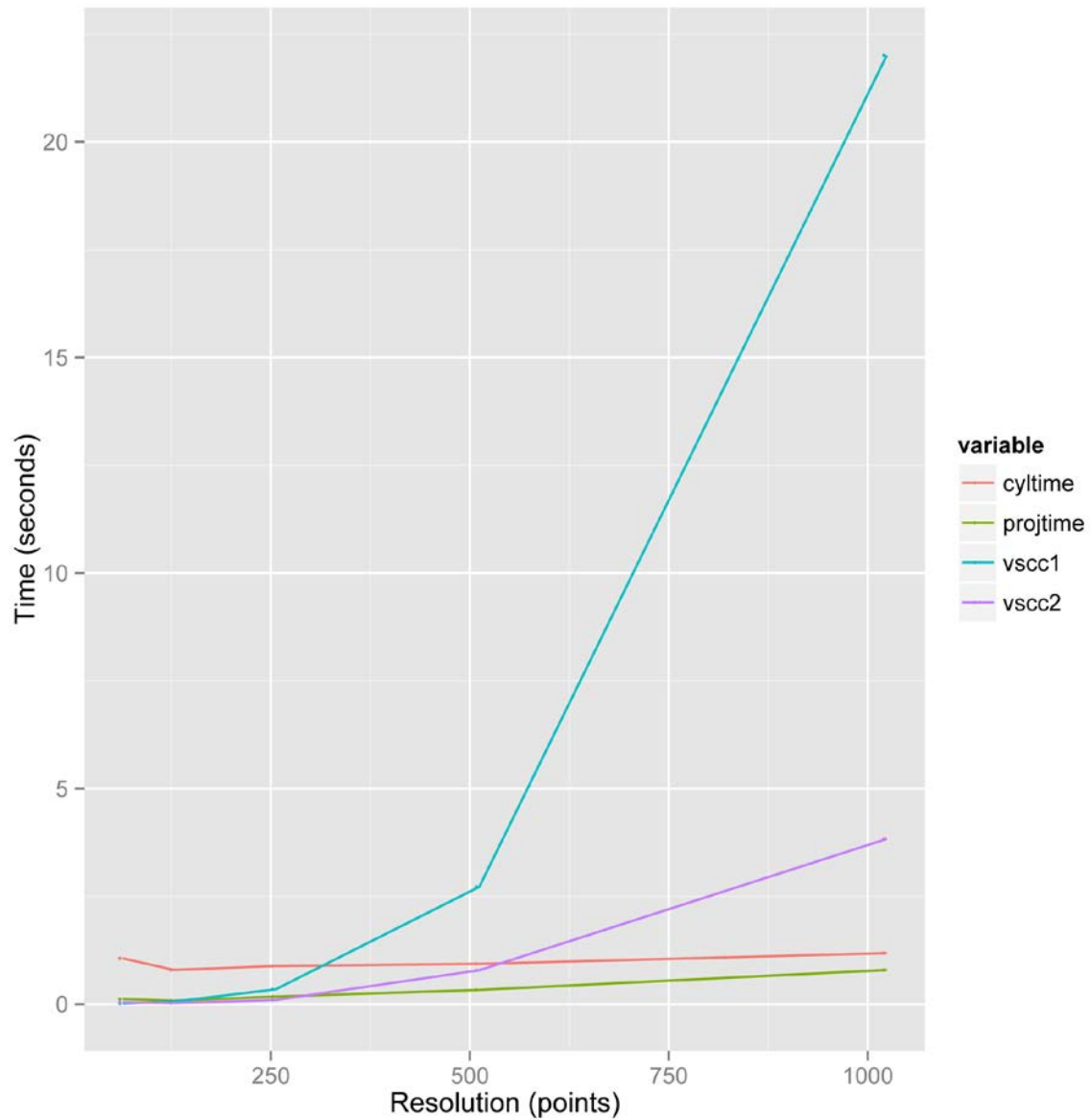
Figure 5: Comparison of computation time consumption related to input resolution.

2.4.9   Comparison of memory and time consumption related to output resolution

The graph in **Figure 6** compares the memory consumption of different parts of the algorithm that depend on the output resolution. The vscc6 line covers the vscc5 one as the measurements of these parts are very close to each other.
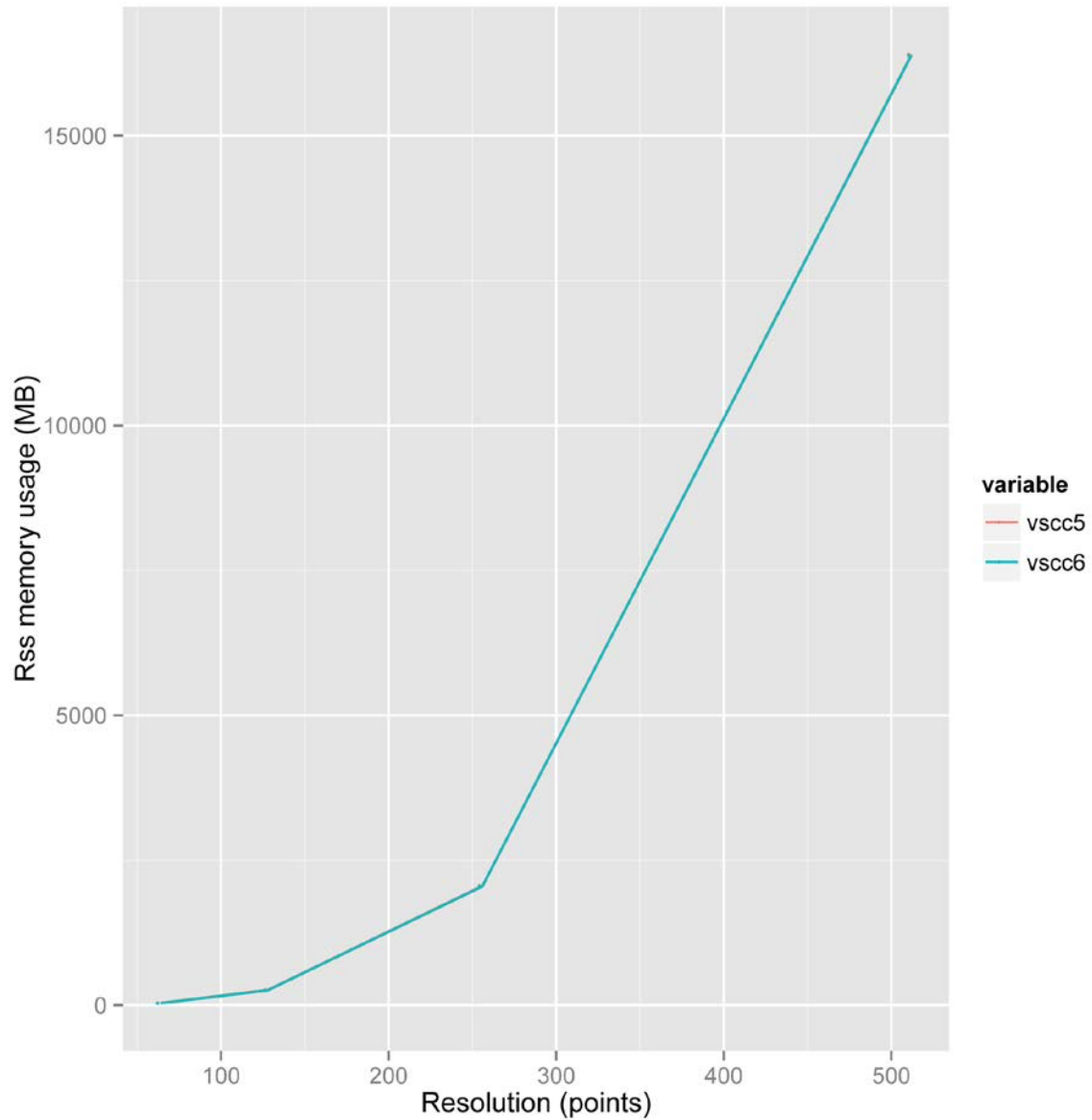
Figure 6: Comparison of memory consumption related to output resolution.

Similarly, the graph in **Figure 7** shows the computation time requirements of the algorithm steps that depend on the output resolution (vscc5).
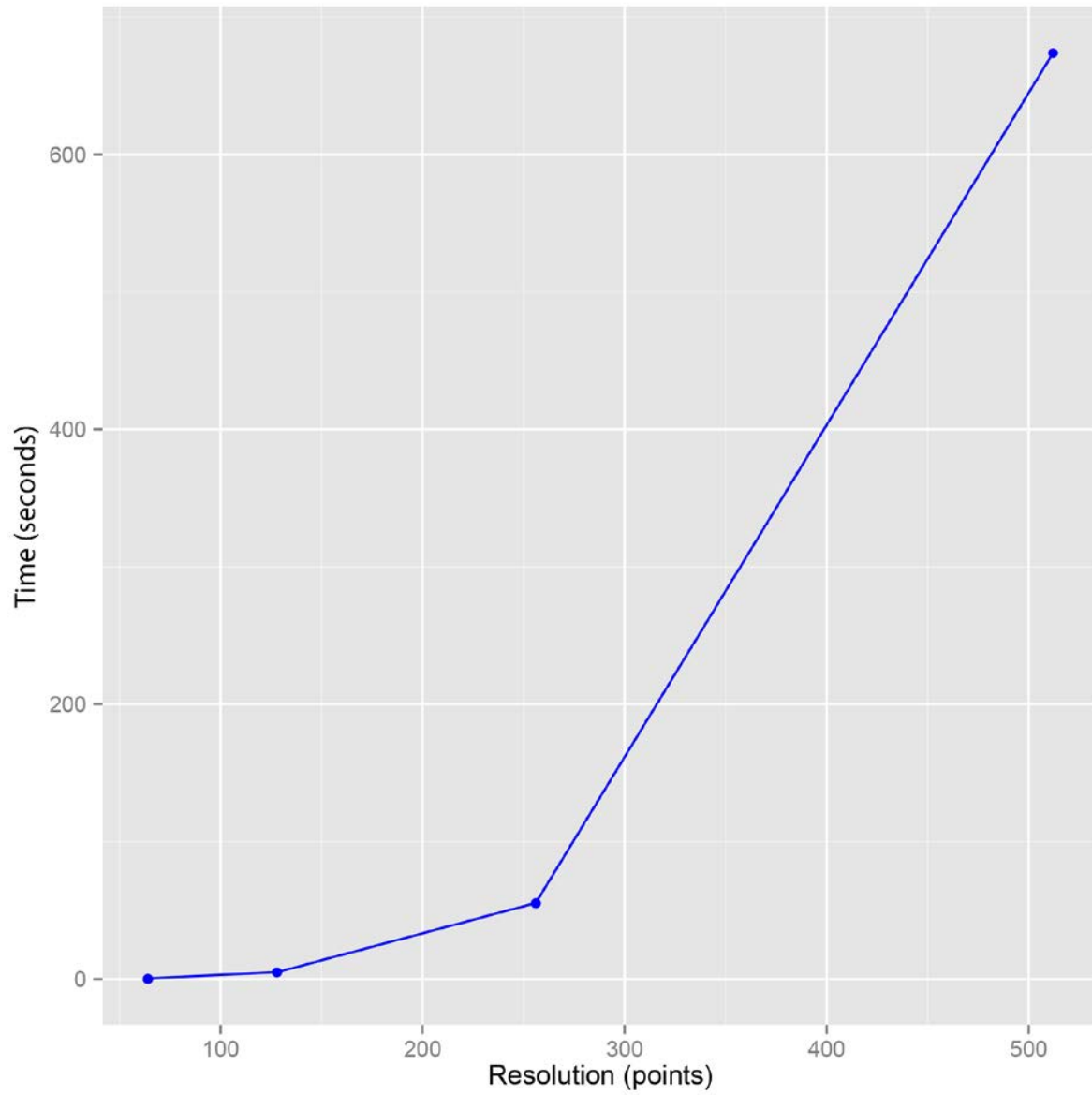
Figure 7: Time consumption related to output resolution.

2.4.10 Comparison of memory and time consumption between input and output resolution

Figure 8 shows compares maximum memory consumption of the input and output resolutions. The axes of the graph are in logarithmic scales; the vertical axis in log 10; the horizontal on in log 2. The blue line in the plot shows the maximum memory consumption related to the output resolution. The red line shows the maximum memory consumption related to the input resolution, which is dominated by the vscc3 step. The plot shows that the influence of the output resolution is much more significant.
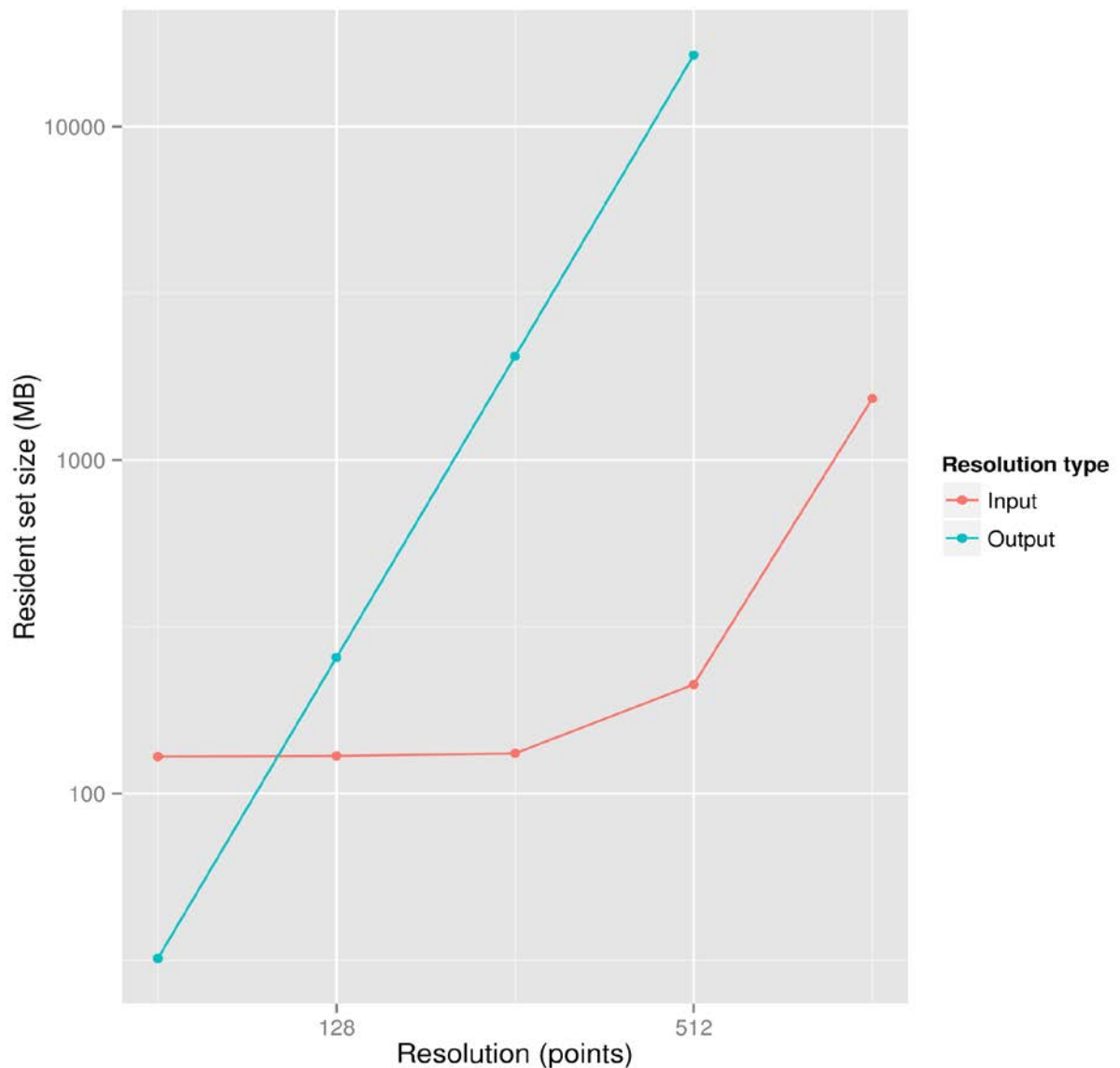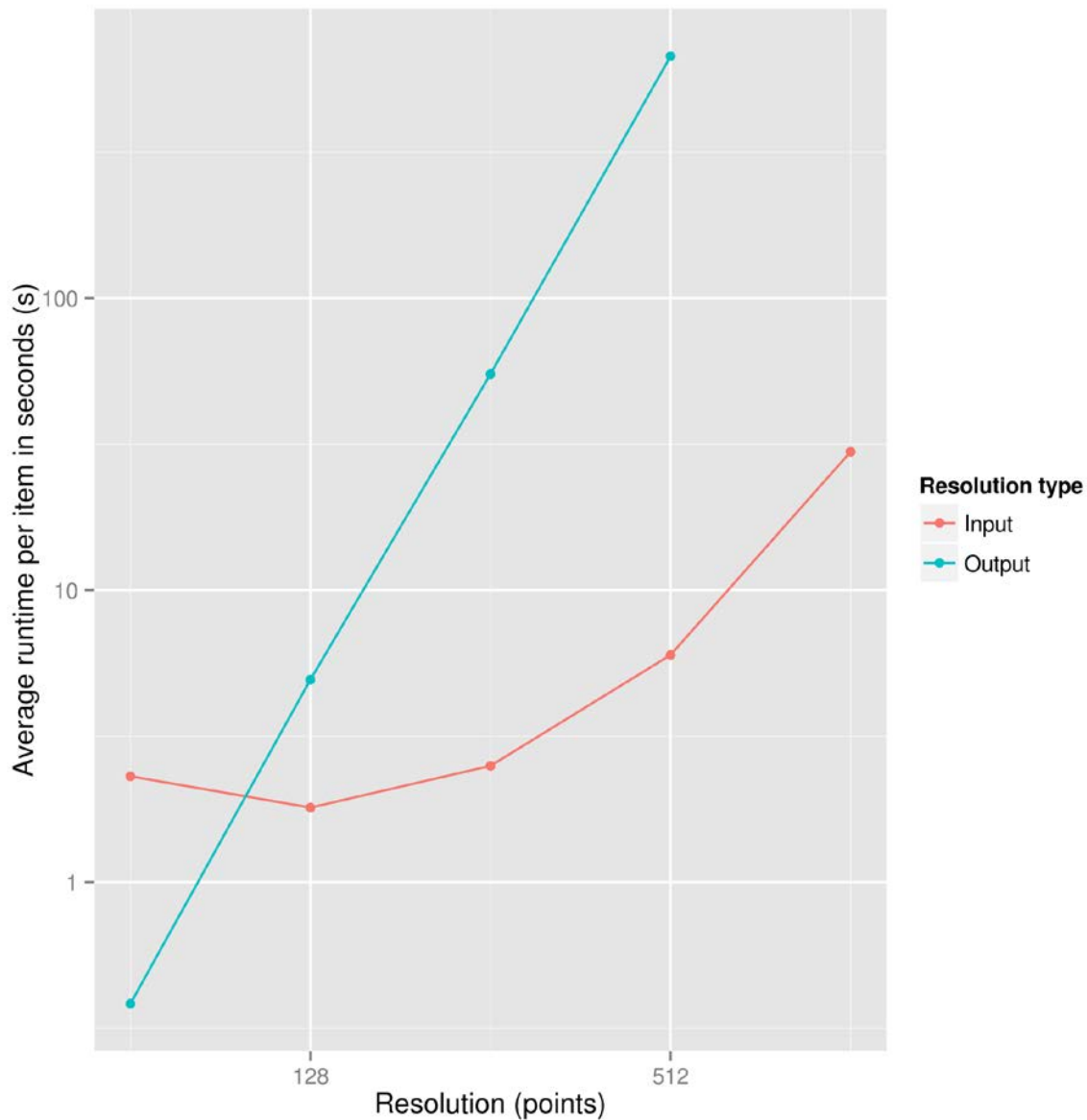

Figure 8: Memory consumption related to input and output resolution.

The dependency of the memory consumption on the output resolution can also be easily derived. The dependence is described by the following equation:

$$memory(MB) = 10^{0,9029 \log_2 resolution(pts) - 3,9114}$$

20

Figure 9 compares maximum time consumptions between the input and the output resolutions. The axes of the plot are in logarithmic scale; the vertical axis is scaled by log 10, the horizontal one is scaled by log 2. The blue line in the plot shows the maximum time consumption related to the output resolution, which corresponds to the vscc3 stage of the algorithm. The red line shows the maximum time consumption related to the input resolution, corresponding to vscc1. It is clear that the time consumption related to the output resolution is much more significant than the consumption related to the input resolution.

Figure 9: Comparison of time consumption between input and output resolution.

The dependency of the memory consumption on the output resolution can be described by the following equation:

$$time(s) = 10^{1,078 \log_2 resolution(pts) - 6.875}$$

2.5 Evaluation of result quality

It can be expected that the quality of results will depend on the input and the output resolutions – the higher they will be, the better quality will be achieved. However, higher resolutions bring higher memory and time demands. One can therefore search for a trade-off between the quality and the amount of resources available. It is also critical to preserve the relevant data to guarantee the same quality in repeated experiments.

2.5.1 False candidates as the measure of the cross-correlation quality

Section 2.1 discussed two metrics to match a real world photo to a rendered panorama image. The explored demonstration algorithm involves computation of a spherical cross-correlation that efficiently generates a set of candidate solutions (that are later checked by a computational intensive fine tuning process). The quality of candidate selection can be described by the amount of false candidate solutions that have a better score (assigned by the cross-correlation) than a correct solution. The lower the number of false candidates is, the better final results can be obtained. The number of false candidates was measured in a relative manner – as a ratio of the number of false candidates to the count of all possible solutions.

When measuring the quality of results for particular combination of an input and an output resolution, the result can be expressed by certain amount quintile values. A high number of the quintiles gives a precise description of the measurement results. Median relative errors for each resolution combination are given in the following table.

Resol.	64:64	128:64	256:64	512:64	1024:64	128:128	256:128
Median relative error	0.0094385	0.006380	0.00514	0.003871	0.003904	0.0030366	0.0017562

Table 10: Median relative error of tested resolution combinations

Resol.	512:128	1024:128	256:256	512:256	1024:256	512:512	1024:512
Median relative error	0.0011149	0.001056	0.00072	0.000208	0.000154	3.649e-05	1.1826e-05

Table 11: Median relative error of tested resolution combinations (second part)

2.5.2   Dependence on resolution

Results of individual measurements are shown as Q-plots in **Figure 10**. Each combination of input and output resolutions is represented by an individual colour. Points in the plots correspond to 50-quantiles. The lower a curve is, the better results were obtained.
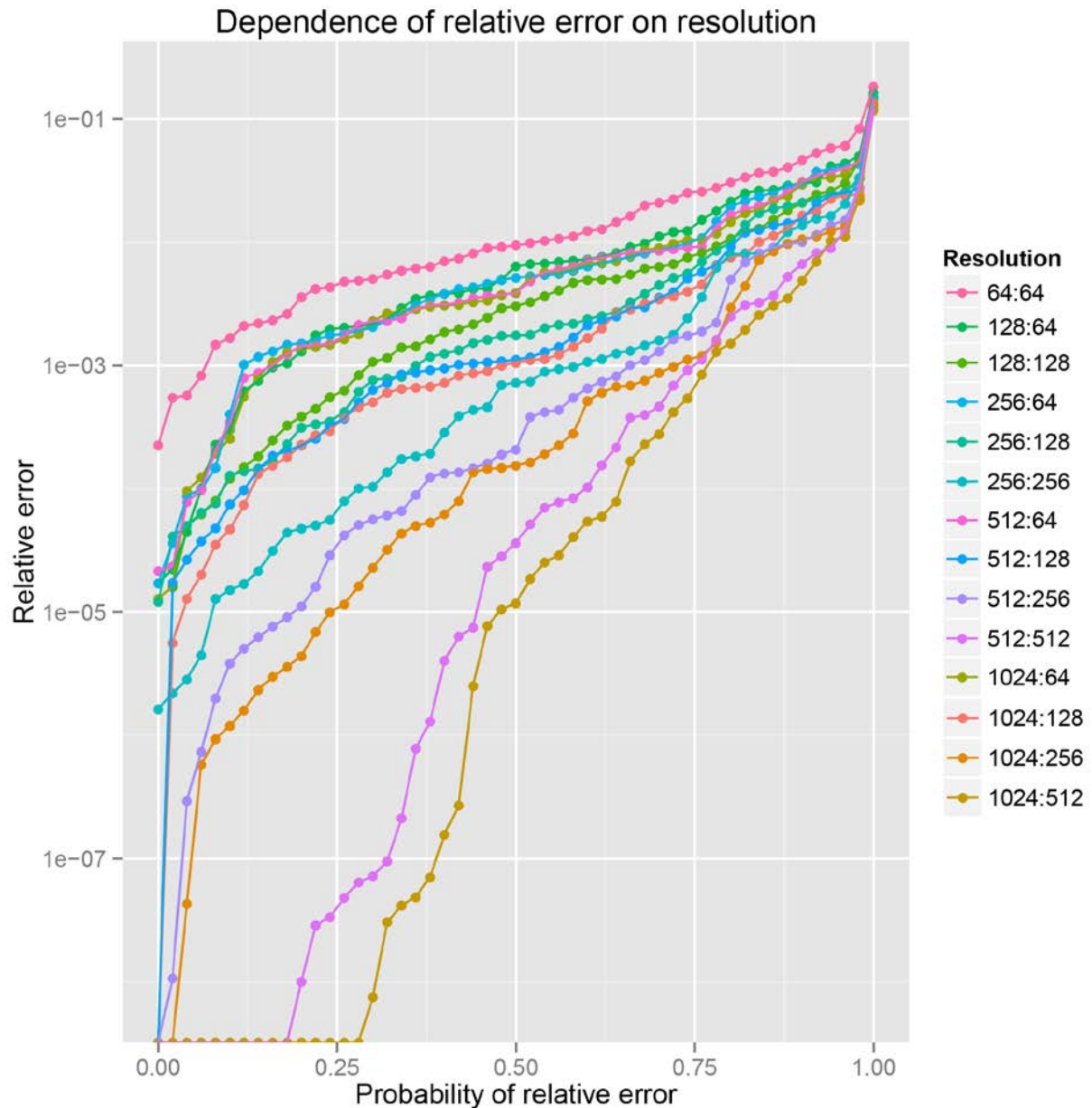

Figure 10: 50-quantiles relative error for each possible combination of input and output resolution.

As expected, the combination of the highest input and output resolutions – 1024:512 – leads to the best results, while the worst results correspond to the resolution combination 64:64. In some cases, the relative error is equal to zero (values that lie at the bottom horizontal axe of the plot) – the real solution was found by the cross-correlation.

It can be also observed that the output resolution has a greater effect on the quality than the input resolution. This is stressed by the graph in **Figure 11** which shows the same results as the previous one, but uses the same colour for values corresponding to experiments with the same output resolutions.
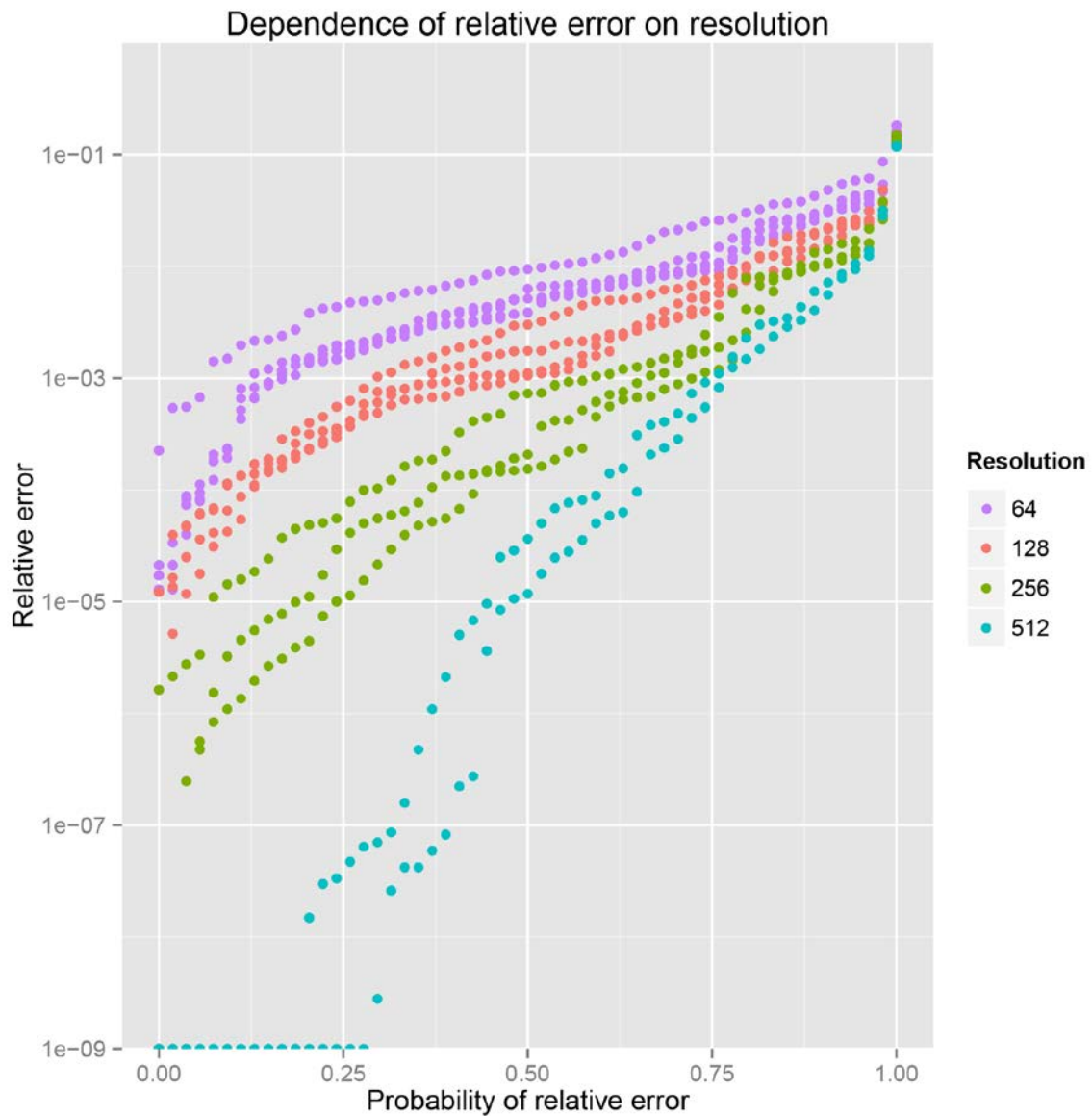
Figure 11: Relative errors grouped by output resolution.

2.6 Scenario scalability

Scalability of the scenario is easy to determine – it mainly depends on the number of involved computation nodes and the size of the output given by its required resolution. Table 12 shows the average time (in seconds) needed to align one input image to one rendered panorama for different numbers of computation nodes and varying output resolution.

Resolution/Nodes	1	3	5	7
64	0.383618	0.127873	0.076724	0.054803
128	4.939417	1.646472	0.987883	0.705631
256	55.31107	18.43702	11.06221	7.901581
512	673.8993	224.6331	134.7799	96.27133

Table 12: Average time to process one input image

We have found out that the biggest bottleneck of the workflow is the SFTP upload/download of the input/output data to/from the remote cluster. Pre-processed real-world images and rendered panoramas take together 132 MB. As the throughput between the BUT local storage and the UVT InfraGRID cluster showed to be rather low, the slow network connection has been identified as a critical factor in the scenario.

3   Medical data processing

WCPT and PSNC joined SCAPE project in order to validate scalable technologies in the context of medical data preservation. One of the first steps in this process was to analyse challenges that WCPT faces in a day-to-day duties. As a result of the analysis the following key issues at WCPT have been identified[1]:

- Limited storage facilities (WCPT is facing ~20TB of produced data each year and cannot preserve all data).
- No real-time access to full patient's treatment history (real-time access is available for 2 years only).
- Need to improve educational activities via medical data exposure, using simple web interface.
- Need to perform large-scale analysis of the medical data recorded in the textual format to obtain statistics and analysis at a general level.
- Medical data stored in the external facilities, such as data centres, cannot contain personal sensitive data.
- The communication between WCPT and PSNC needs to be encrypted.

From the above limitations and challenges specific conclusions have been drafted and summarized in four main scenarios that were further investigated:

- Medical data ingest – WCPT needs external data storage due to limited resources onsite. WCPT policy forbids transferring sensitive personal data outside its premises. Therefore before ingesting medical data to external data storage (PSNC data centre) WCPT needs to anonymise specific data.
- Medical data access – WCPT user needs full access to patient's treatment history, including data coming from PSNC's data centre and sensitive personal data kept only at WCPT. In order to achieve this goal it is necessary to involve DICOM files personalization component within the WCPT environment. This component adds personal data to DICOM files, which are afterwards presented to WCPT user.
- Access to anonymised medical data – teachers at medical universities need a simple and straightforward way to showcase specific diseases together with their description prepared by professionals in the field (e.g. radiologists). PSNC's data centre stores anonymised data that can be shared with the teachers at universities as well as with students.
- Medical data analysis – to support research activities it is required to calculate statistics on textual data provided by WCPT. This scenario involves large scale processing of textual data with the use of technologies and paradigms such as Map-Reduce approach or NoSQL database.

Above four medical data scenarios are depicted at a very general level on Figure 12. The general idea is that there are two main technical elements working and communicating with each other:

- WCPT is visible on the right side. WCPT has existing software components (visible at the bottom) such as PACS (Picture Archiving and Communication System) and HIS (Hospital Information System). WCPT hosts so called Hospital Archive System (HAS) for handling medical data before transferring them to PSNC's Medical Data Center.
- PSNC is visible on the left side. PSNC hosts so called Medical Data Center (MDC), which leverages SCAPE-identified software components (e.g. Hadoop, HBase) as well as

---

[1] Please refer to D8.3 Initial anonymised ingestion prototype for full discussion of the identified challenges

technologies used in the context of cloud and data center environment (e.g. OpenStack, dcm4che adjusted to cloud needs).

The HAS system is composed of the following components:

- Transfer Service (TS) – responsible for HAS-MDC communication. It acts as a client of the MDC. It can send data to MDC and access data from MDC. For example, it uses dcm4che tool[2] (dcmsnd command) to send anonymized DICOM files to MDC.
- DICOM anonymization component – used for anonymization of the medical data before transferring them to MDC via the TS.
- DICOM personalisation component – used for personalisation of the medical data before giving access to WCPT users. The idea is that whenever the WCPT user (granted permission to see sensitive personal data) wants to access patient's treatment history, the DICOM files coming from the MDC (which are anonymised) will be supplemented with the sensitive personal data. In this respect, the anonymization and personalization process is transparent for the WCPT users.
- Encryption component – used for encrypting and decrypting those parts of medical data that are most sensitive. Currently this is the identifier of the patient's visit at the hospital.
- HIS-MDC – component which keeps track of references between HAS data identifiers and MDC data identifiers.

The Medical Data Center covers the following components:

- HDFS PACS – a server application, which is a customized version of the Picture Archiving and Communication System based on dcm4che toolset. The tool was modified in a way that it can store the data (DICOM files) in HDFS (for processing purposes) and PLATON U4[3] cloud storage (for archiving purposes).
- HDFS HL7 – a server application, which is a HL7 gateway. It receives textual data encoded in HL7 version 3 format (XML-based) and stores it on HDFS (for further processing) and cloud storage (for archiving purposes).
- WWW – it is a web-based component for accessing anonymised data. The main idea behind it is to give access to medical cases of the WCPT hospital for university teachers or students.
- A set of tools which are treated as backend technologies, such as OpenStack cloud environment, Cloudera components (Hadoop and HBase) as well as dcm4che toolset for handling medical data.

---

[2] http://www.dcm4che.org/
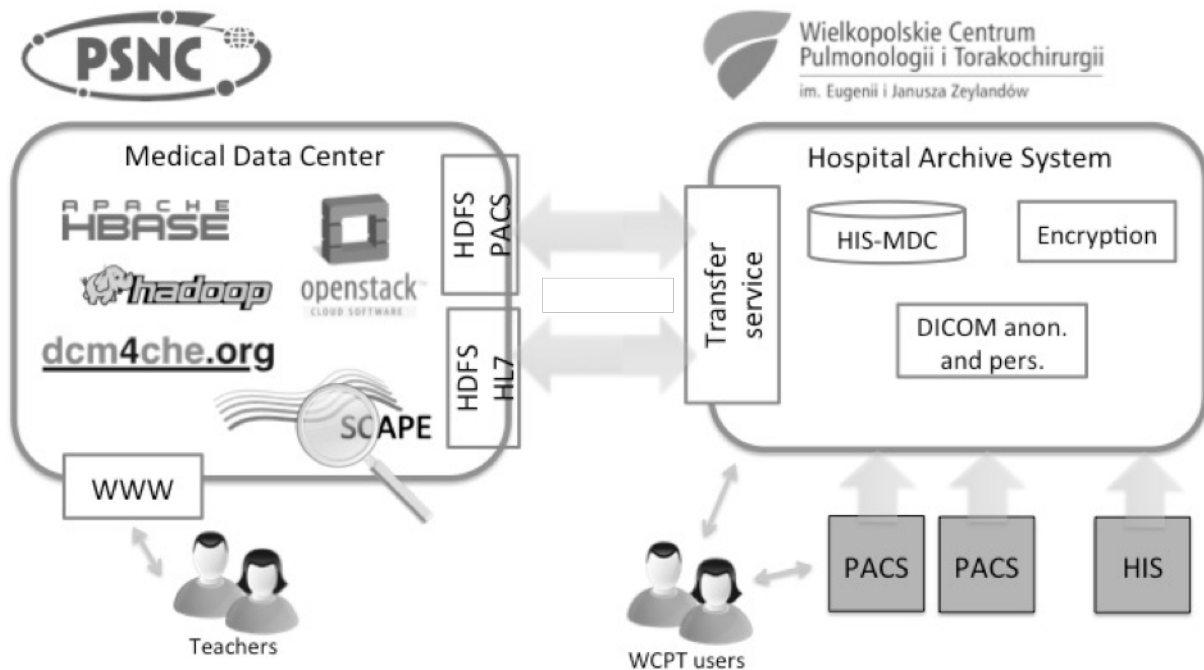[3] http://storage.pionier.net.pl/

Figure 12 Medical data scenarios - WCPT and PSNC technical components and communication

PSNC's hardware infrastructure that was used in the context of MDC has been described in Deliverable 4.3 Final Data Center Deployments[4]. Shortly the infrastructure to build MDC is composed of six server nodes, each containing 6 CPU cores (Intel® Xeon® Processor 2.83 GHz) and 13GB of RAM. The nodes have been connected with 1Gb Ethernet. The storage capacity in HDFS build on top of these servers is approx. 30TB. Apart from this MDC deployment (so called production-mode cluster) PSNC has provided development-mode cluster composed of 4 virtual servers launched in cloud powered by OpenStack.

In the following part of this chapter all medical preservation scenarios are described. All of the information related to the performance or scalability of particular scenarios has been included and described in Deliverable 18.2 SCAPE final evaluation and methodology report (see Data Center Testbed evaluation).

3.1 Medical data ingest scenario

WCPT is responsible for storing medical data and therefore needs an archiving system. Current software and hardware resources used at WCPT are limited therefore only part of the medical data produced by WCPT is preserved. The main reason for that is lack of necessary resources (mainly storage space and computing power) to store data according to the requirements enforced by law, namely to store medical data for at least 20 years and 30 years in specific cases. In order to overcome this barrier a special solution has been identified and developed by WCPT and PSNC. Before going into details, please note that the medical data is created and available in the external system such as HIS. The data can come from the registration desk (personal information), from particular diagnosis (e.g. CT, RTG) or specific observation. There are two workflows related to the medical data ingest. One is related to the patient's visit information (HL7 files) and the other one is related to the medical imaging data (DICOM files). The workflow related to ingesting of an HL7 file is composed of several core steps depicted on Figure 13 and listed below:

---

[4] https://portal.ait.ac.at/sites/Scape/Shared%20Documents/Deliverables/Final/SCAPE_D4.3_UVT_V1.0.pdf

1. First step in the HL7 file ingest scenario concerns data retrieval action performed by the HAS system. It means that HAS can access HIS and retrieve all the necessary information that need to be transferred to MDC. Basic unit of the information that is retrieved from HIS is related to a patient's visit.

2. The second step is related to encryption of the patient's visit identifier. The encryption is done using AES algorithm with 128 bytes key length. The key is stored in the HAS system.

3. The third step assures that data available at MDC is linked with the data in the hospital system (HIS). It is done by saving the link in the HAS system (HIS-MDC component), namely the link between patient's visit identifier stored in the MDC and patient's visit identifier stored in the hospital system.

4. Next, the data are sent over the internet to the MDC using HL7 client. MDC HL7 gateway receives the information. Communication is done via HTTPS protocol. As already indicated, in communication between MDC and HAS, the information related to patient's visit is represented as an XML-based file conformant with HL7 format (version 3). Sensitive personal data are not included in the HL7 file that is sent to MDC (this is assured by the HAS).

5. Finally, the information (HL7 file) is stored in the MDC both for processing (on HDFS) and archiving purposes (in the cloud storage).
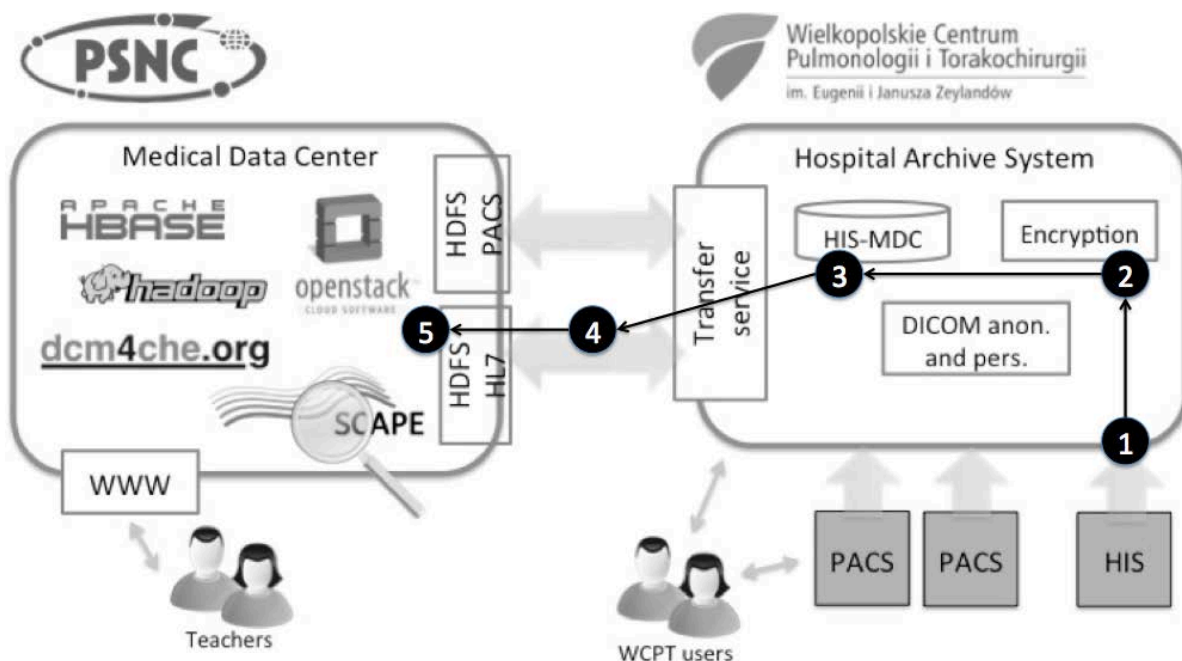
Figure 13 Workflow for ingesting information about patient's visit

HL7 files send from HAS to MDC contain information about patient's visit excluding medical imaging such as CT or RTG. The HL7 files still keep descriptive information and identifiers (accession number of related DICOM files) of the medical imaging files, so that is possible to access those, if available in the MDC. For that reason it was necessary to prepare additional ingest workflow for the medical

imaging items, namely for the DICOM files. The DICOM files ingest workflow is depicted on *Figure 14*. It is composed of four steps listed below:

1. In the first step HAS retrieves DICOM files from the PACS system at WCPT hospital. Because DICOM files contain sensitive personal data, it is necessary to anonymise them before ingest.

2. The second step is related to anonymization. A special tool called PADI[5], developed in the framework of the SCAPE project, is used to anonymize each DICOM file that is to be sent to MDC.

3. Next, the DICOM files are sent from HAS to MDC. Communication is executed using DICOM protocol and dcm4che toolset[6] with enabled encryption option. HAS uses default dcmsnd command to send data.

4. In the last step MDC stores the DICOM files in the MDC. For that purpose MDC uses modified version of the PACS server, so that it can store DICOM files in HDSF and cloud storage.
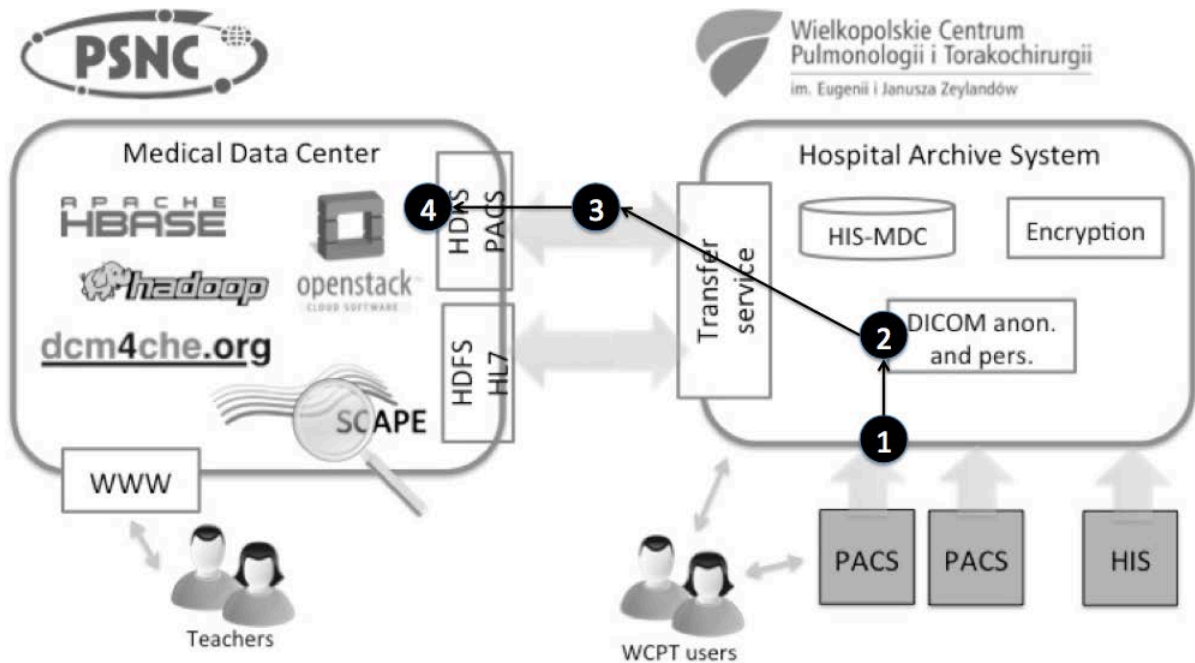
Figure 14 Medical imaging data ingest workflow

3.2 Medical data access scenario

Employees at the WCPT hospital need easy access to preserved medical data (those stored at the Medical Data Center hosted by PSNC). The goal is to have access to the history of patient's treatment including data accessible from the Medical Data Center. There are two workflows for medical data access at the hospital. One is related to patient's visit information access (HL7 file access) and the other one is related to medical imaging access (DICOM files access). The workflow related to patient's visit information access is depicted on Figure 15. It is composed of 4 steps that are listed below:

---

[5] https://github.com/openplanets/PADI
[6] http://www.dcm4che.org/

30

1. In order to access the full patient's history, a WCPT user needs to request data specific to a certain patient's visit. WCPT user does this via HAS user interface, by indicating a certain patient's visit.
2. In the second step HAS determines patient's visit identifier in the hospital systems and looks up for corresponding identifier in the MDC (using HIS-MDC component).
3. In the third step HAS communicates with the MDC and retrieves data via HTTPS protocol (for that purpose it uses MDC identifier of the patient's visit).
4. Finally, the information can be presented to the WCPT user together with sensitive personal information that is kept in the HAS and HIS systems.
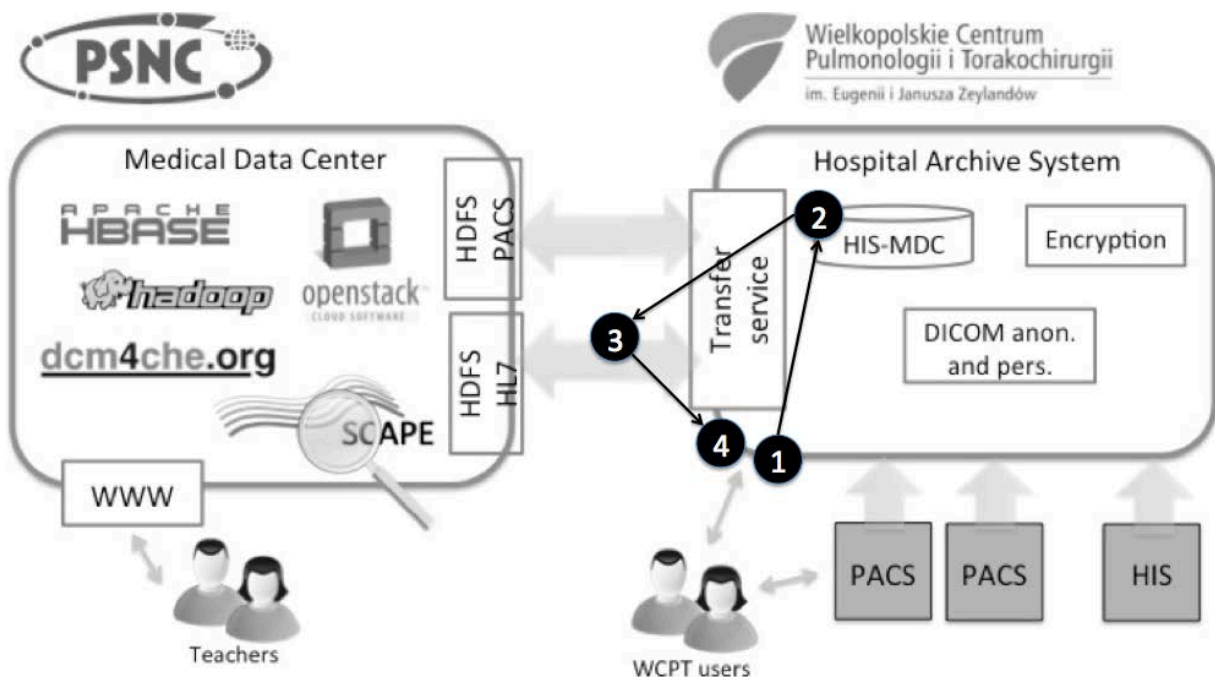
Figure 15 Workflow for accessing information about patient's visit

The workflow for medical imaging data access at WCPT hospital is depicted on Figure 16. If a WCPT user needs to access medical imaging data from the MDC he or she needs to use HAS user interface. The following steps need to be undertaken in order to access medical imaging data:

1. In the first step WCPT user requests data related to specific examination (please note that the precondition for that is to have patient's visit information already retrieved from MDC, as HL7 files are necessary to link patient's visit with medical imaging data).
2. In the second step HAS system determines examination identifier (accession number from the HL7 file) and retrieves the data from the MDC. Communication is done via HTTPS protocol.
3. In the next step HAS executes personalisation of the DICOM file with sensitive personal data (information about the patient). It means that in this step anonymised DICOM files are supplemented with sensitive personal data.
4. Finally, the resulting files are presented to the WCPT user, who can view them in DICOM files viewer.
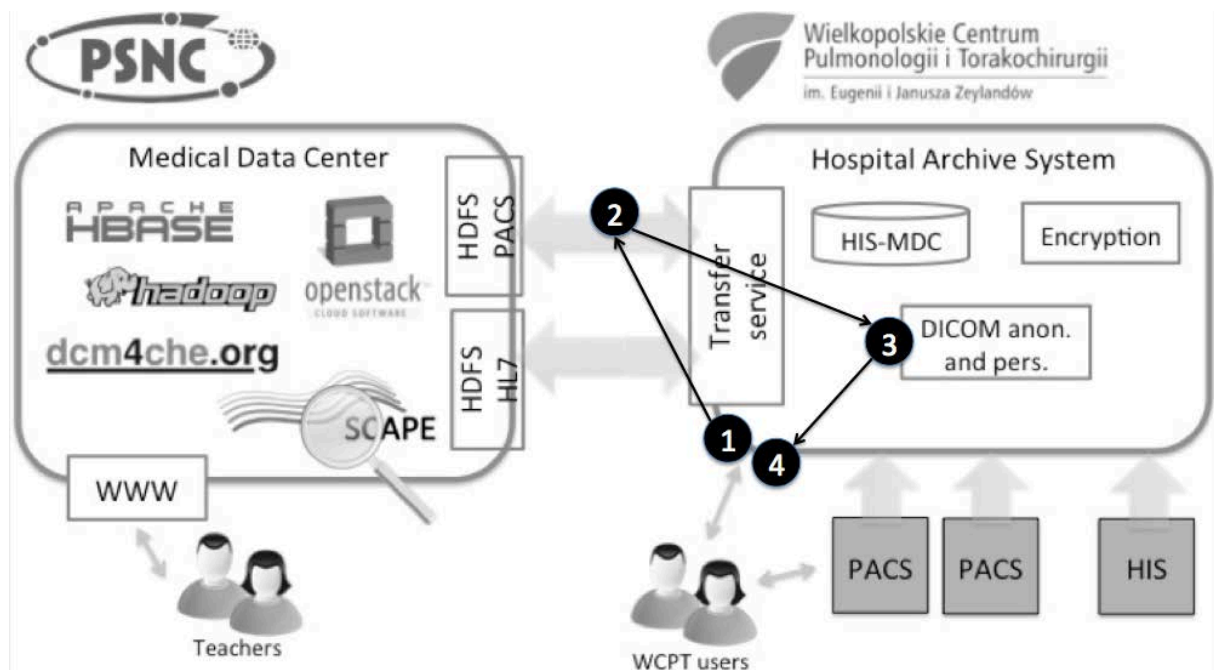
Figure 16 Medical imaging data access workflow

3.3 Access to anonymised medical data scenario

For the educational purposes it is required to have an easy access to various examples of different diseases. In this case personal sensitive data is not relevant. The important information is the professional description of the examination results as well as the examination results. In this respect, a workflow for accessing anonymised medical data has been identified, implemented and deployed in the MDC. The basic idea behind it is to have MDC access portal to all of the medical data stored in MDC. The portal is available via the internet and using a web browser. The user can search for specific cases based on the following criteria:

- ICD10 code or description
- ICD9 code or description
- Sex of patient
- Age of patient (range is possible to specify)
- Dates of patient's visit in the hospital
- City of patient's residence
- Examination results

The results of the search function are the medical cases. Each medical case is related to a certain patient and is composed of all hospital visits of this patient. Figure 17 presents an exemplary screen with search results.

Figure 17 Search results in the medical data center portal for accessing anonymized data

Having the search results the user can open the details view of the specific medical case and browse all of the relevant information. If available, medical imaging data can be also presented (a DICOM file online viewer is build-in into the MDC portal). Figure 18 presents details view of the patient's visit. In the tabs all patients' visits are available and can be accessed by the user.

Use the filters to narrow the search results

Search

Medical case details

Medical Case: male, 58 years

| Visit in hospital from 12-05-2014 to 01-07-2014 | Visit in hospital from 12-05-2014 to 01-07-2014 |

Visit in hospital from 12-05-2014 to 01-07-2014

Under. diag. ICD-10:   J85.1   Ropień płuca z zapaleniem płuc

ICD-9 procedures:
93.1910; 93.1814; 93.1813; 93.1812; 23.0105

Complaints
Chory lat 57 przekazany z Oddziału P-4 celem rehabilitacji układu odddechowego po leczeniu zachowawczym ropnia płuca lewego.Od 12.05.2014 w Oddziale P-4 do którego został przekazany z oddziału detoks

Addiction
Alkohol: N Papierosy od: 20 do: 57 roku w liczbie: 20

Blood type
B RhD+/dod

Laboratory diagnostics

Parameter	Results	Unit	Reference values	
WBC	9,07	G/L	4,23 - 9,07	
RBC	4,53	T/L	4,63 - 4,63	L
HcT	0,4010	L/L	0,40 - 0,51	
HGB	8,10	mmol/L	8,5 - 10,9	L
MCV	88,50	fL	79,0 - 92,2	
MCHC	20,20	mmol/L	20,0 - 22,7	
RDW	15,60	%	11,6 - 14,4	H
PLT	252	G/L	163 - 337	
MPV	10,50	fL	9,4 - 12,6	

RTG
30-06-2014 09:40 Show examination
.W porównaniu do zdjęć z dnia 23-05-2014 i 20.06 dalsza resorbcja przywnękowego nacieku w płucu lewym, nie zmieniła się istotnie grubościenna jama w jego obrębie . W języczku. okołooskrzelowe pasma niedodmy.Nie ma objawów obecności płynu w jamach opłucnej . Miąższ pozostałych partii płuc rozedmowy , w segP2 drobne zwłóknienia, poza tym bez zagęszczeń. POPRAWA

Epicrisis
Chory lat 57 przekazany z Oddziału P-4 celem rehabilitacji układu odddechowego po leczeniu zachowawczym ropnia płuca lewego.Od 12.05.2014 w Oddziale P-4 do którego został przekazany z oddziału detoksykacji dla uzależnionych od alkoholu w Gnieźnie.W wywiadzie alkoholizm, pourazowe zlamanie ramienia prawego., nikotynizm. Po leczeniu zachowwczym(antybiotykoterapia) z poprawą kliniczną i radiologiczną(w kontrolnym KT) przekazany do tutejszego oddziału w satnie ogólnym dość dobry. Chory zgłasza odksztuszanie dużej ilości plwociny,bóle stopy prawej od tygodnia. W badaniu przedmiotowym szmer pecherzykowy prawidłowy,czynność serca miarowa. Pacjent zgłasza odkrztuszanie podbarwionej krwią wydzieliny-po zastosowaniu leczenia poprawa. W czasie hospitalizacji stosowano kinezyterapię-ćwiczenia oddechowe.W kontrolnym rtg płuc dalsza resorbcja zagęszczeń zapalnych,nadal utrzymuje się jama. W stanie ogólnym dobrym z zaleceniem kontroli w Por.Pulmonologicznej pacjent został wypisany do domu.

Recommendations
Kontrola w Por.Pulmonologicznei za 3-4 tygodnie. Wskazane wykonanie kontrolnej KT klatki piersiowej za 6-8 tygodni. Leki: Augmentin 2x1tabl po

Figure 18 Patient's visit detail view

Detailed view of the patient's visit contains various information including examination results such as morphology, RTG or CT. Textual data are presented on the detailed view page, but in order to get access to medical imaging data the user needs to click on a "Show examination" link visible near CT or RTG summary (as presented on **Figure 18** – near RTG summary). **Figure 19** presents medical imaging data for RTG.
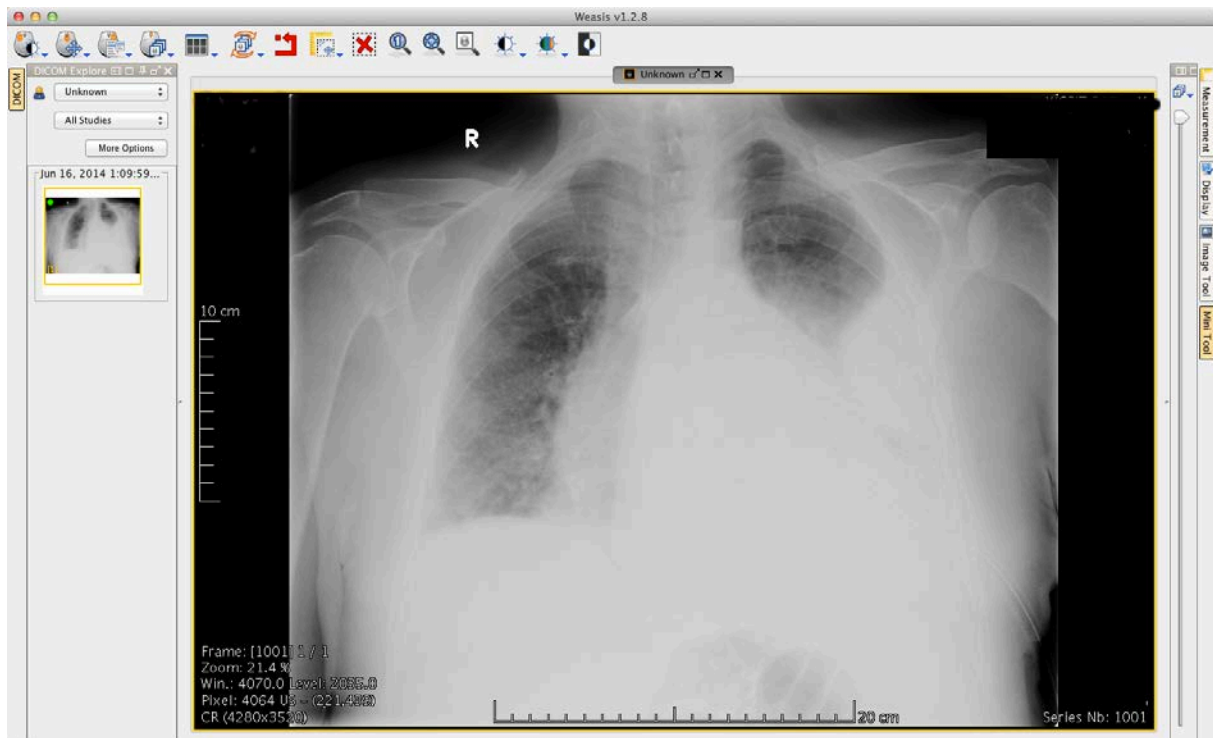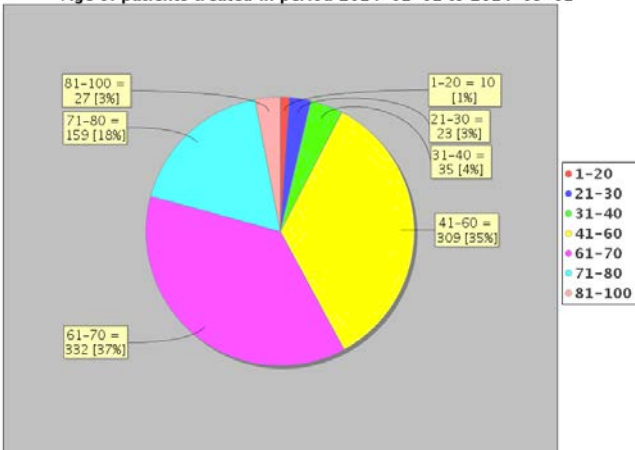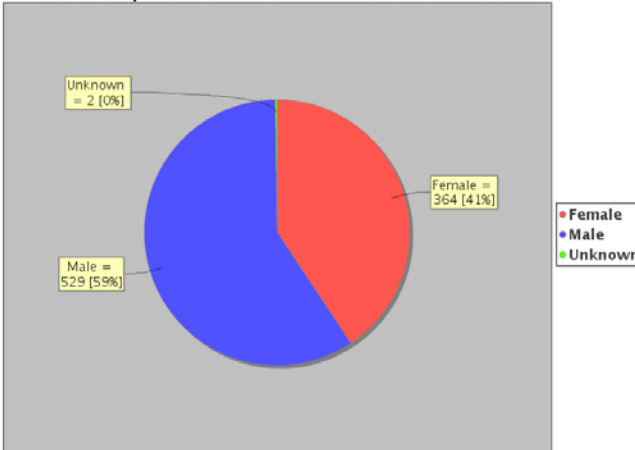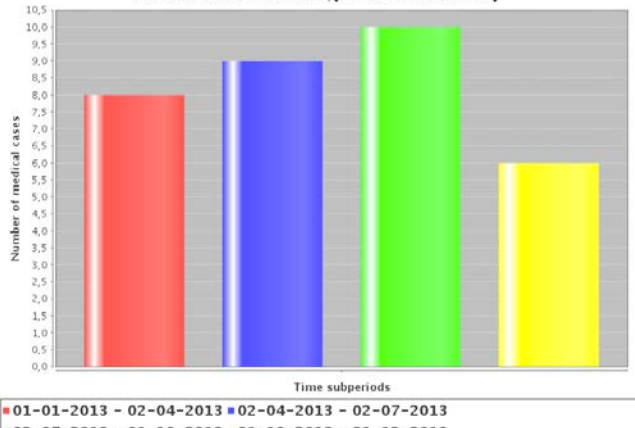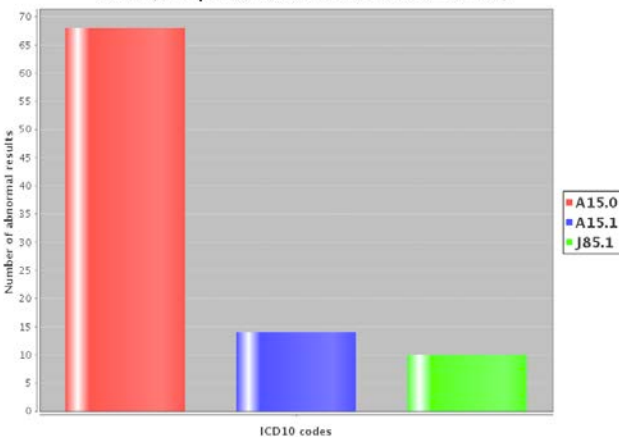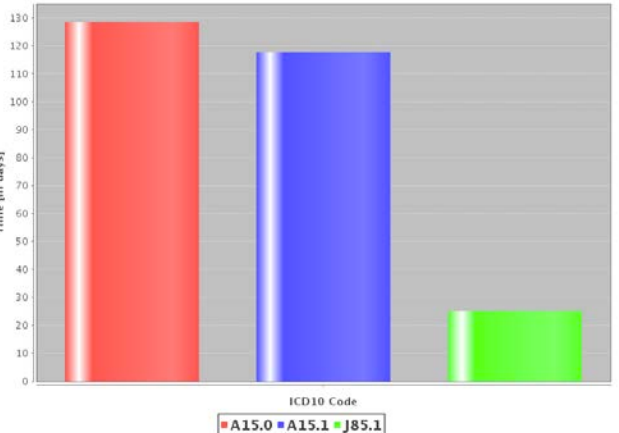
Figure 19 Medical imaging data available at MDC

3.4 Medical data analysis scenario

Researchers working at the hospital need a way to analyse large amounts of medical data related to various patients' treatment. The analysis is obviously helpful in the research activities or can be directly used by the hospital in statistical analysis. All of the relevant data is stored in the MDC on the HDFS (with Cloudera toolkit attached), which makes it is possible to process them in a large-scale manner. WCPT hospital identified five most interesting analysis that could be executed on the medical data stored at MDC. Table 13 presents a summary of the exemplary analysis. The first column indicates the analysis name, the second column contains input parameters and the third column presents an exemplary result (chart).

Table 13 Summary of exemplary analysis executed using medical data stored in the MDC

Analysis	Input parameters	Exemplary result (chart)
Age of patients treated in a given period	the period of time to be analysed, the age intervals	
Sex of patients treated in a given period	the period of time to be analysed	
Number of cases of a given disease in a given period	the period of time to be analysed, the number of sub-periods visible on the resulting chart, the ICD10 disease code	

Number of abnormal results in laboratory examinations for a given disease codes in a given period	the period of time to be analysed, the ICD10 disease codes	Number of abnormal results in laboratory examinations for specified ICD10 codes (time period: 2013-01-01 to 2013-12-31)
Average time of patient's visit for a given disease codes in a given time period	the period of time to be analysed, the ICD10 disease codes	Average time of visit for specified ICD10 codes (time period: 2013-01-01 to 2013-12-31)

Because each analysis is different, relevant algorithms needed to be implemented as separate Hadoop jobs. Each Hadoop job can be executed several times with different parameters. By default the analysis jobs are executed using the command line tool. In order to simplify access to the analysis algorithms PSNC has prepared a dedicated subpage on the MDC portal, next to the search and browse interface of the anonymised medical data. The user interface is simple, but it already showcases the possibilities of analysis that can be executed by any interested party. Figure 20 presents user interface of one of the analysis algorithms (on the left side input parameters site, on the right side results site). When the user executes the analysis, MDC portal executes a certain Hadoop job with the given parameters. Please note that if the analysis needs to cover large period of the time then calculating results can take a long time (multiple minutes, hours).
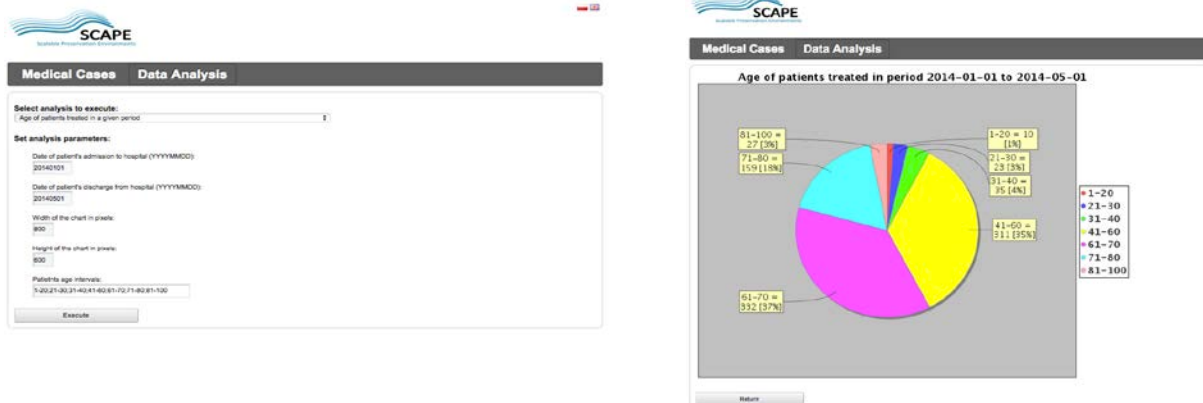
Figure 20 Medical data analysis in the MDC portal (input parameters on the left side, resulting chart on the right side)

4   Conclusions

This deliverable reported on deployment and evaluation of two diverse preservation scenarios that share the aspect of using external data centres for storing and processing data.

The first part focused on large-scale experiments and preservation actions aimed at reproducibility of results. A specific video analysis process served as an example of a complex distributed program, results of which are significantly influenced by various settings of particular computing platform used, as well as actual availability of resources on individual processing nodes (e.g., consumption of resources by other processes). Indeed, the memory and time consumption grows rapidly with increasing output resolution). Derived equations predict that the output resolution involving 2048 points would consume as much as 1 TB of the RAM memory and the computation would take almost 27 hours. This would definitely conflict with other tasks running in a data centre so that the nodes would need to be shared. Then, it is critical to preserve relevant parameters, including those of actual runs of an experiment in a particular environment. The realized SCAPE tools allow users to define a preservation strategy in such cases and to collect all necessary data.

The main benefit for the computer graphics and vision community is the possibility to perform large-scale experiments in external HPCs preserving actual characteristics of the cluster environment. The preservation contributes to repeatability of experiments and enables further analyses of previous runs of experiments, e.g., comparing scalability of newly developed methods to previous ones. The general processing and preservation mechanism explored in the SCAPE project will continue to be used in the established cooperation between BUT and UVT and will be newly applied as a contribution to the Czech national supercomputing centre IT4I.

Medical data preservation was a primary focus of PSNC and WCPT cooperation in the SCAPE project. Work carried out during 13 months of the final project phase resulted in a very promising solution for hospitals that are focused on medical data preservation and long-term accessibility.

Medical Data Center deployment has been developed to overcome existing barriers and address new needs appearing in the WCPT hospital. The resulting set of tools aim at making medical data more accessible and safe. Data accessibility is ensured by the MDC with dedicated educational portal and standardised access APIs, while the safety is assured by making copies of the medical data to nation-wide cloud storage (PLATON U4).

The prototype solution has been tested and validated by the WCPT and provided satisfactory results in terms of usefulness and required performance. PSNC was able to verify its HPC infrastructure on a real-life scenario related to medical data preservation. Two server clusters were used for data processing and several approaches have been tested for data storage and access. For details related to evaluation see Data Center Testbed evaluation in Deliverable 18.2 SCAPE final evaluation and methodology report.

Because PSNC joined the project in the last year it was crucial to benefit from original SCAPE partners experiences, hence map-reduce approach was investigated in the medical scenario. Running medical data processing as well as storage and access services combined with large-scale analysis using map-reduce approach gave a deep understanding of the advantages and limitations of such platforms. Using a toolset composed of Hadoop, HBase, dcm4che and PADI it was possible to prepare a Medical Data Center (MDC) platform that can store, analyse and provide data in a large-scale manner.

39

The MDC has been integrated with the WCPT hospital environment for smooth data transfer, anonymization and preservation. The overall solution can be re-used in other medical environments, giving an opportunity to integrate data coming from various hospitals.