A network performance view of a biobanking system for diagnostic images

Giusy Esposito, Giulio Pagliari, Gianluca Coppola
Bio Check Up S.r.l. (Italy)
{gesposito,gpagliari,gcoppola}@biocheckup.net

Marco Aiello, Marco Salvatore
IRCCS SDN (Italy)
marco.aiello@synlab.it,
marcosalvatore.segreteria@gmail.com

Giuseppe Aceto, Antonio Pescapé
University of Napoli Federico II (Italy)
{giuseppe.aceto,pescape}@unina.it

***Abstract*—A significant contribution of ICT to healthcare is constituted by systems automating and enhancing the management of research and clinical data. More specifically, PACS (Picture Archiving ad Communication Systems) have improved the efficiency of diagnostic images and clinical data management. Their evolution (image biobanks) are now enabling new collaborations and analysis possibilities similarly to—and beyond—the biobanks (their biologic samples analogous and complement). In this work we describe and evaluate the network performance of a biobanking system for diagnostic images, based on the XNAT open source platform, as implemented and operated by Bio Check Up Srl. The point of view of the user is adopted, in assessing the performance in three setups: local (virtual machines communicating in a single host), LAN (organization-local access), and VPN (remote secure access through the Internet). Both upload and download usage cases are considered, with both a medium-sized and big-sized set of diagnostic images. Several metrics are extracted from traffic traces captured in the experimental campaign, and discussed. Results show that the current setup is well provisioned for satisfying the planned number of concurrent users, and point to further experimental campaigns.**

***Index Terms*—biobank, PACS, network performance, e-health, imaging, diagnostics**

I. Introduction

Some of the fast-paced expanding contributions of ICT to healthcare research are the automatic management and the enhanced use of patient digital data. The wider diffusion of electronic systems in clinical practice, which include mobile devices, computers and diagnostic equipment [1], has led to the development of health informatics, defined as the science underlying the academic investigation and practical application of information and communication technology to healthcare and biomedical research. . ICT applications in health concern the design and optimization of information systems for hospitals, imaging centers and medical offices (e.g. HIS, RIS, PACS, etc.). Additionally, health informatics include medical data modelling, processing and analysis. Finally, ICT can deeply contribute to the standardization and automation of medical diagnostics, screening and consultation processes in order to reduce time and costs and increase patients' benefits [2]. Health data include digital images, patients' personal details, and reports. They can be additionally associated with information derived from other tests (clinical, genomic, oncological) or quantitative data extracted from direct processing on the images. The aim of this work is to contribute with a use case study from a running real service, i.e biobanking of digital data from clinical analyses, intended for a wide audience of practitioners in the field of medical and clinical research, as well for computer science experts supporting the development and evolution of the infrastructure herein described. The IT infrastructure dedicated to digital biobanking allows researchers and clinical practitioners to consult, annotate, segment, process and manage collections of medical images and clinical data [3–6]. In this study, we perform an experimental evaluation of an implementation of a digital Biobank, derived from XNAT [7] and operated by Bio Check Up Srl (BCU). In the analysis we adopt the viewpoint of the user, focusing on the perceived performance (in terms of transfer completion time) for the most frequent and most network-resources-demanding activities (upload and download of diagnostic images sets). Two different images sets sizes, representative of typical use cases, are considered, and network performance metrics are analyzed in detail. Three different network scenarios—local (virtualization in a single host), LAN, and remote VPN are considered, to provide references to the performance experienced over the (uncontrolled) Internet. Results highlight that the different Upload/Download activities, performed with different tools, exhibit significant different performance figures. The analyses confirm the usability of the system and the dimensioning of the network and server setups, also pointing at further analyses and future developments.

This paper is structured as follows. Section II introduces the state of the art and provides a brief discussion of digital images and information systems in clinical and research activities. Section III describes the System architecture, its framework and the implementation at BCU. Section IV includes the experimental evaluation of the platform with three different real case scenarios. Finally, Section V contains our conclusions and future work.

978-1-7281-8086-1/20/$31.00 ©2020 IEEE

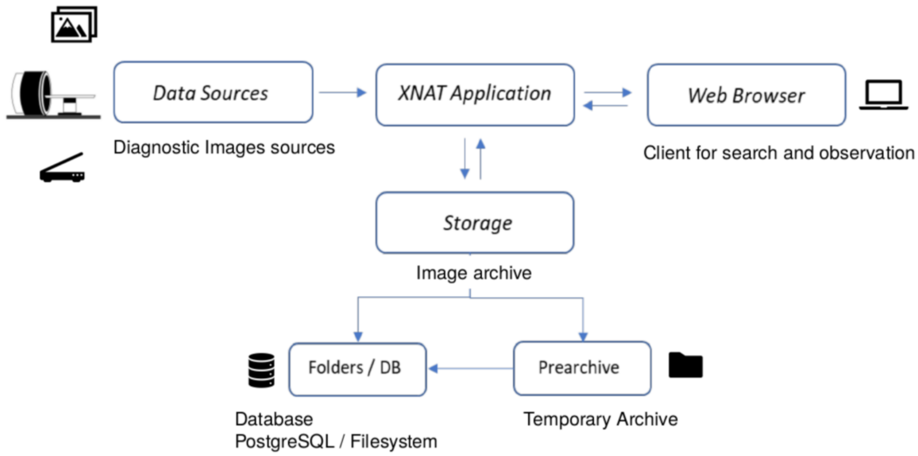


Fig. 1: Overall architecture of the system.

II. Background and Related Work

The impact of ICT in healthcare has rapidly evolved over time, as evidenced by the number and variety of systems and applications created to support and improve the activities of healthcare workers. An information system can be considered as a set of resources and tools with the aim of managing information within an organization. The medical field has seen the spread of the following information systems, defined hereafter [8–10]:

- HIS (Hospital Information System)—a complete and integrated set of IT tools, used to manage administrative and clinical information within a hospital. The purpose of HIS is to obtain the best support for health management and memorization of the patient's history, both in acquisition and in data processing. The main functions of HIS are those of archiving, storage, use and easy accessibility to the information of interest.
- RIS (Radiology Information System)—has the same purpose and function as HIS, but is used in radiological units only.
- PACS (Picture Archiving and Communication System)—an information system suitable for storing images acquired from various imaging modalities, managing the network of workstations connected to it and accessing images during reporting. PACS eliminates manual archiving and the use of radiographic films from X-ray flow processes

The standard format for storing medical images is DICOM (Digital Imaging and COmmunications in Medicine) [11]. The DICOM standard includes protocols, rules and syntax to encode, transfer, archive and stamp a digital medical image. This standard promotes the interoperability between information systems. Imaging biobanks can be defined as "organised databases of medical images and associated imaging biomarkers (radiology and beyond) shared among multiple researchers, and linked to other biorepositories" [3]. They consist in a service unit specialized in collecting, archiving and retrieving medical images and data, organized in collections based on studies relating to a specific pathology, organ or disease. Unlike the other patient-based information systems that are usually installed within a single institution or pair between structures, an imaging biobank is a collection-based repository whose data could be shared among researchers. A survey realized by the working group of ESR has detected 27 imaging biobanks disease-oriented in Europe, of which the 80% are not open access [3]. Research efforts in this field should focus on the application of common standards and on validation and benchmarking of IT infrastructures used for the repository architecture.

III. System Description

Health information systems had a wide diffusion thanks to the increasing need of sharing and processing medical images and data in order to achieve better and faster diagnoses. Usually, a biobank contains biological samples only. In the last few years, imaging biobanks have been established and there are huge efforts on the standardization process. As for the other health information systems, an imaging biobank should be solidly developed/installed/implemented on suitable IT structures. Specifically, the minimum requirements for the realization of an imaging biobank are: (i) appropriate storage, in terms of security and dimension, (ii) internal and external connectivity, in order to upload, download files and manage the IT infrastructure, (iii) dedicated workstation or server, (iv) documentation, ethical principles and privacy statement. Starting from these considerations, having assessed the hardware infrastructure and the availability of the various components as well as the study and compliance with the rules / regulations, we moved on to choosing the platform that best suits our purposes. The Longitudinal Online Research and Imaging System (LORIS) [12] and the Extensible Neuroimaging Archive Toolkit (XNAT) [7] have been taken into consideration. LORIS, developed by the Montreal Neurological

(a) Local Virtual Machines (VM) (b) Local Area Network (LAN) (c) Virtual Private Network (VPN)

Fig. 2: Network setups considered in the experimental evaluations.

Institute, is a modular system that allows all aspects of a multi-site study. XNAT, developed by the Neuroinformatics Research Group at Washington University, is a web-based, open-source platform for neuroimaging research and processing.

As in [13], we defined a set of selection criteria and we based our choice on: availability of a database for managing and storing images and metadata, storage of measurements derived from images or non-imaging data (e.g. clinical data), support of different formats (e.g. DICOM, NIfTI, NRRD), possibility to customize and add extensions or plugins.

The final choice of XNAT was driven by the fulfillment of all criteria, and the presence of rich documentation and a wide community of active users. The installation of XNAT v1.7.5.6 was performed using Docker v19.03.3 [14] with the following specifications: Intel Core i3 - 8th Gen processor, 8GB RAM, 240GB SSD, OS Ubuntu 19.04. The main components of the system are shown in Fig. 1.

Afterwards, in order to evaluate the performance of the system, we conducted experimental evaluations as described in detail in the following section.

IV. Experimental Evaluation

An analysis of network performance has been designed and performed in order to both *(i)* assess the performance perceived by the operator in a realistic context, and *(ii)* evaluate the impact of the network technology, with the final goal of predicting the infrastructural requirements for providing the service. This benchmark was useful for assessing the performance of the network in the hypothesis that the platform is used at full capacity as an imaging biobank by different users both in downloading and in uploading images.

Usage scenarios

To perform measurements that are representative of realistic and common usage scenarios, instead of using synthetic but realistic traffic load [15, 16], in this paper we use traffic generated by clients supported by the application under analysis, actually engaging it. To this aim, expert operators and practitioners currently involved in clinical analysis data collection and usage have been interviewed, leading to the following choices. In the measurements, two sets of sample images of different sizes are considered: their sizes are 100MB and 1GB, representative of *medium size* and *big size* sets of diagnostic data, respectively. These choices are motivated by the fact that these files represent realistic cases of diagnostic images that will be treated within the biobank (averaged on actual sets of real diagnostic images, and rounded to a reference value for readability of results). A ten-fold enlargement has been considered to represent more rich diagnostic image sets. This is coherent with the usage levels expected for the service: an initial monthly contribution of ≈ 500 images sets has been estimated (≈ 100 transfers per week). In the first stage of deployment of the service a moderate influx of concurrent users has been planned, averaging to 10 concurrent users. Given the high specialization of the service and the *permission-based* process of authorization to its use, the maximum of concurrent users is known beforehand and can be controlled for. Therefore this aspect has been investigated at a basic level in current work, and further deeper investigations are deferred to future work.

For each of the two types of transfer (medium / big) both *Upload* and *Download* activities have been considered. The *Upload* represents a contribution of a new set of images to the biobank, and is performed by means of the `DicomRemap` tool[1]. The *Download* represents the access to images sets that are already stored in the biobank and accessible to the user. Only the data transfer process is considered, without any other online interaction (browsing, selection, preview) that may be implied in this process. The data transfers are conducted towards and from the *prearchive* storage area of the biobank, and do not include any anonymization procedure. To perform Download measurements the HTTP XNAT API[2] has been accessed, by means of the `curl`[3] tool—a de-facto standard for HTTP(s) API access.

Network setups

The benchmark test has been carried out considering three different network setups, useful for performance comparison purposes. In all the considered setups, the installation of the XNAT platform was carried out using lightweight virtualization technologies (Docker containers), according to current best practices to ease scalability and maintainability of software systems [17].

Local (VM): This setup consists of a Client virtual machine used for sending images, and another virtual machine hosting the XNAT platform, both VMs are hosted on the same machine and are connected via an emulated private network. This setup from the usability point of view is representative

[1] https://wiki.xnat.org/xnat-tools/dicombrowser/using-dicombrowser-in-the-command-line
[2] https://wiki.xnat.org/display/XAPI/XNAT
[3] https://curl.haxx.se/

of a “legacy” configuration, similar to a monolitic application running locally, requiring direct access from the terminal.This case can also be considered as an ideal scenario from the network point of view, as the (local) virtualization environment emulates a very high performance network (Fig.2a).

Local Area Network: In this setup we have a client machine that communicates via the LAN (switched Gigabit Ethernet) with the XNAT platform (Fig.2b). This scenario is used to investigate the performance on the local network typical of access from inside the laboratory that houses the biobank. The setup also constitute another benchmarking reference for the (more complex) VPN scenario.

Virtual Private Network: This setup is composed of a client machine located on an external network in order to represent the case of a user who accesses the biobank of images from any location using the Internet connection, adopting a VPN authorized by the host of the platform to communicate in a secure way (Fig.2c). Server-side, the Internet access network is a *Fiber to the Home* installation, with nominal network layer capacity of $300Mbps$ in Upstream and $500Mbps$ in Downstream. Client-side, the Internet access network is a *Fiber to the Home* installation, with nominal network layer capacity of $100Mbps$both up- and down-stream directions. The VPN client software is OpenVPN[4], an open-source software implementing an SSL authenticated and encrypted sublayer in the application level. The Authentication and encryption algorithms are SHA-256 and AES (256-bit), respectively.

The experiments have been planned as measurement *sessions* each considering one network setup (*VM, LAN, VPN*), and one usage scenario (*medium/big upload/download*). The tests were fully automated, and the measurement procedure is described in details in the following section.

Measurement procedure

As described in the previous sections, we recall that for measurement purposes both *Upload* and *Download* procedures have been implemented with command-line tools, to grant repeatability and exclude any human factor in the process: the whole procedure is completely automatic (also accounting for possible connection failures, repeating attempts up to a timeout). Moreover, all measurements are based on *raw packet traces* captured on the client machine: the utility `tcpdump` has been used to capture and filter traffic based on the biobank server IP (only network traffic related to the client-server biobank communication is captured); while the `tshark` utility has been used to extract measured metrics (described in detail in the following section). The measurement procedure is described as pseudocode in Alg. 1. It can be noted how different packet traces (lines 4 and 7) are captured for the Upload and Download activities respectively. The authentication procedures (*login* and *logout*, lines 1 and 19 respectively) are performed only once per measurement session, outside the repetitions loop and outside the packet captures. Moreover, each repetition begins with a `traceroute` such path is compared (line 12) with the one collected in the previous repetition, and if a change is detected the number of repetitions already performed is reset to 1. This way the desired number of repetitions are guaranteed to happen in the same network path configuration, preventing routing changes during the repetitions from impacting measured network performance. This is especially significant for the *VPN* network setup, where traffic between the client and the biobank server traverses the public Internet, on which there is no control possibility. Finally, to guarantee the same operating conditions for what regards the biobank server side, in each repetition the uploaded images set is deleted (line 11).

[4]https://openvpn.net

Measurement metrics

To characterize the performance experienced by the user in the considered usage scenarios and network setups, the following metrics have been considered.

Total transfer duration: This is the main metric, representative of the experience of the user of the digital biobank. Indeed any transmission issue causing delays and packet losses (in wired connections, as in the considered setups, mostly due to congestion) are surfaced to the user only as additional delay in completing the transfer. It is computed from the raw packet trace as the difference in time between the latest packet seen and the first one. This metric is strongly dependent on the volume of data to be transferred, therefore it is evaluated for the two considered sizes (medium/big) of figure sets.

Goodput: Another user-related metric is *goodput*, defined as the number of application bytes received per unit time in each direction. Differently from *duration*, this is not directly derived from the size of the transmitted payload. Notably, our definition excludes from the application data also the application-layer protocol overhead, while other uses of the same term includes such component [18]. Therefore this metric shows the overall effectiveness of the system (included application-level protocol) in transferring user provided files.

Average Byte Rate: The impact of the network conditions on the performance experienced by the user is described through the *byte rate*, calculated as the total amount of transmitted traffic (at network layer) divided by the duration of the transfer. This metric, differently from *goodput*, highlights the efficiency of the network, regardless of the application-layer protocol overhead. Therefore it can be more directly related to nominal network transfer rate, to check if the sending/receiving access connection is a bottleneck for the transfer. Notably, this metric (differently from *goodput*) conflates all network-layer-and-up protocols overhead with the user data, including possibly retransmissions due to lost packets and timeouts. For this reason we monitored retransmissions, as they can be inferred from the packet trace captured at the client side[5].

[5]Retransmissions are detected as a packet that: is not a keepalive packet; in the forward direction, has segment length is greater than zero or the SYN or FIN flag is set; has sequence number smaller than the expected sequence number.

Average Packet Rate: A measure of the impact that the transfer has on the network infrastructure is represented by the average *packet rate*. In the considered VPN setup a firewall and a VPN client and server middleboxes are on the path between the user client and the biobank server: each middlebox performance can be affected by high packet rates, therefore this metric is of interest for the considered usage scenarios.

Other metrics: While collecting the metrics mentioned above, other aspects have been monitored, namely *packet losses and retransmissions*, and *average packet size*, in order to have a comprehensive view of the communication conditions.

Algorithm 1: Measurement procedure pseudocode.

```
Input: ImageSet, Repetitions
Output: Measured network parameters
login on biobank;
repetitionCount ← 0;
repeat
    traceroute to biobank;
    start traffic capture;
    upload ImageSet;
    stop traffic capture;
    start traffic capture;
    download ImageSet;
    stop traffic capture;
    delete ImageSet;
    if same network path of previous measurement then
        repetitionCount ← repetitionCount + 1
    else
        repetitionCount ← 1;
        discard measurements before last one;
    end
until repetitionCount = Repetitions;
logout from biobank;
process packet traces;
return measured metrics;
```

A. Analysis of results

In this section we report the analyses of results from a measurement campaign conducted in March 2020. A synthetic summary of measured metrics is reported in Table I for all considered combinations of network setups and usage scenarios. For the most user-relevant metrics (*goodput* and *average byte rate*) results are reported also as box-plot showing quartiles (median is orange colored), minimum, maximum and possible outliers, calculated on immediate repetitions for each session. The small variability measured in all experimental campaigns has motivated the choice of stopping at 10 repetitions per session (as in some cases a single transfer can last more than 25 minutes, this number of repetitions already accounts for varying network conditions). In the following we describe and interpret the main results, on a per-metric basis.

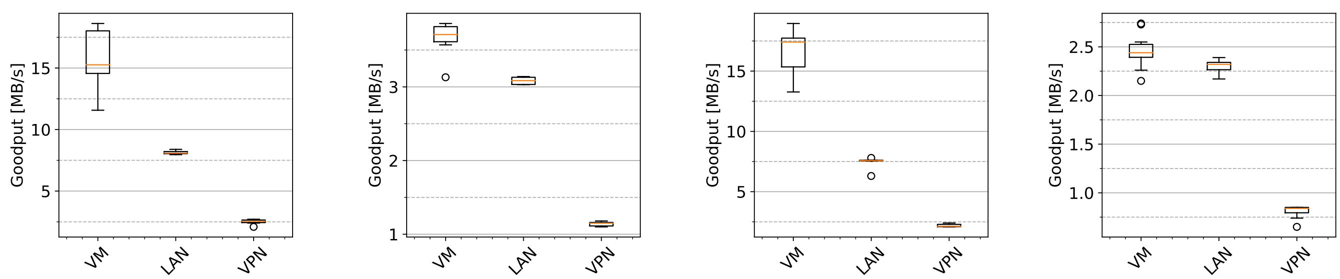
We start the analysis from results regarding transfer *duration*. It can be noted that for the activity most frequently envisioned (Upload of a medium-sized images set) the average time to complete the transfer is less than $40s$ in the VPN setup. This value is compatible with a presence of an operator waiting for the completion of the transfer to continue other activities. The case of a 10-fold increase of size of the image set for the Upload usage case roughly causes a proportional increase of the waiting time, for all three network setups (with a small additional increase of $\approx 10\%$ for VPN). This implies that for *Upload big size* use case the user waits ≈ 8 minutes on average in the VPN setup, while for the LAN and VM ones the waiting is less than 3 and 1 minute(s), respectively. A significantly different situation is found for Download activities: in this case the duration is increased to more than twice the time of the Upload for the *medium size* transfer, and almost three times for the *big size*. This phenomenon happens for the VM and LAN setups as well, therefore can not be attributed to asymmetry in the network path (both VM and LAN setups are symmetrical) and should be instead attributed to the two different methods that have been adopted to transfer the Images Sets to and from the biobank.

Regarding the overall performance of the system in transferring the user data, in Fig. 3 the measurements of *goodput* are reported. Coherently with the *duration* measured values, and as expected given the involved the network technology, the best results are obtained for the VM setup. Differently from the *duration*, in this case the impact of the total transfer size is of minor significance: besides the expected relative differences among the three technologies, the main variation is not between the medium–big Images Set sizes, but between the Upload–Download usage scenarios: the goodput of Upload is several times (from more than twice, for VPN setup, up to almost 7 times, for VM setup) higher than the one of Donwload. This phenomenon is more evident for big size scenario. As noted for the *duration*, this can not be ascribed to network path asymmetry, as it is present also for VM and LAN setups, where paths are guaranteed to be symmetrical.

To more directly relate these perceived performance results to the network infrastructure, the *average byte rate* is reported in Fig. 4. Differently from the previous graphs, in this case for each network setup both directions upstream, and downstream, characterized by the suffix *up* and *down*, respectively. As expected, for the usage case *Upload* the almost-entirety of traffic is found in the upstream direction, or client-to-server. Analogous condition happens conversely, downstream for *Download* case, in the server-to-client direction. The difference between the forward (client-to-server for Upload, server-to-client for Download) and the backward direction (the remaining two cases) is of about two orders of magnitude, across all the Download use cases. The most extreme difference is found for *Upload, medium size* usage (Fig. 4a), VM setup (client-to-server accounting for $1.2 \cdot 10^3$ times the byte rate of the opposite direction). This difference is much less prominent for the VPN setup (a ratio of $6.5 \cdot 10^1$) for the *Upload, big size* usage scenario (Fig. 4c). These variations could be in principle due to retransmissions, but we have monitored this aspect, finding very limited occurrences of retransmissions (less than 20 in a single transfer session). Therefore other factors (available bandwidth, latency, and protocols overhead) are to be investigated. An evident contribution to this phe-

TABLE I: Measured metrics. Values are in form *average* (± *standard deviation*).

	Duration [s]		Goodput [MB/s]		Byte rate [kB/s]		Packet rate [pkt/s]		Average Packet Size [Byte]	
Upload, medium size (100MB)										
VM_up	6.55	(±1.12)	15.65	(±2.49)	16564.49	(±2635.3)	973.37	(±151.19)	17010.87	(±316.39)
VM_down					13.78	(±1.98)	133.5	(±18.31)	103.15	(±1.93)
LAN_up	12.32	(±0.2)	8.12	(±0.13)	8588.43	(±145.62)	407.75	(±13.94)	21074.76	(±405.27)
LAN_down					136.94	(±3.5)	2036.37	(±52.71)	67.26	(±0.02)
VPN_up	39.9	(±3.24)	2.52	(±0.19)	2766.53	(±203.93)	2095.88	(±154.54)	1320.02	(±0.02)
VPN_down					46.23	(±3.91)	871.33	(±73.66)	53.06	(±0.01)
Download, medium size (100MB)										
VM_up	27.36	(±1.78)	3.67	(±0.22)	15.69	(±0.76)	185.55	(±8.72)	84.54	(±0.21)
VM_down					802.62	(±42.96)	156.7	(±7.18)	5120.63	(±65.83)
LAN_up	32.42	(±0.56)	3.08	(±0.05)	16.39	(±0.37)	203.96	(±4.61)	80.37	(±0.05)
LAN_down					684.09	(±15.25)	179.48	(±4.07)	3811.6	(±10.85)
VPN_up	87.54	(±2.23)	1.14	(±0.03)	12.76	(±0.31)	224.51	(±5.42)	56.85	(±0.01)
VPN_down					260.6	(±6.3)	215.46	(±5.21)	1209.48	(±0.07)
Upload, big size (1GB)										
VM_up	60.52	(±7.05)	16.71	(±1.78)	19394.68	(±2067.97)	1176.21	(±106.5)	16468.68	(±398.34)
VM_down					15.14	(±1.55)	139.37	(±14.07)	108.6	(±1.2)
LAN_up	134.26	(±8.72)	7.47	(±0.42)	8662.5	(±486.24)	365.84	(±15.13)	23669.53	(±686.91)
LAN_down					133.81	(±8.19)	1987.18	(±121.88)	67.34	(±0.02)
VPN_up	456.07	(±25.32)	2.2	(±0.13)	2647	(±151.73)	2004.14	(±114.88)	1320.77	(±0.02)
VPN_down					44.05	(±2.72)	829.78	(±51.15)	53.09	(±0.02)
Download, big size (1GB)										
VM_up	409.29	(±30.62)	2.45	(±0.19)	14.46	(±0.69)	181.46	(±8.8)	79.69	(±0.17)
VM_down					1081.45	(±51.24)	160.37	(±7.86)	6744.8	(±99.7)
LAN_up	434.97	(±12.39)	2.3	(±0.07)	19.4	(±0.57)	257.66	(±7.64)	75.31	(±0.02)
LAN_down					1054.1	(±29.52)	237.41	(±7.08)	4440.27	(±11.94)
VPN_up	1242.98	(±117.92)	0.81	(±0.07)	15.28	(±1.25)	277.56	(±22.67)	55.04	(±0)
VPN_down					381.37	(±31.42)	301.68	(±24.84)	1264.14	(±0.09)

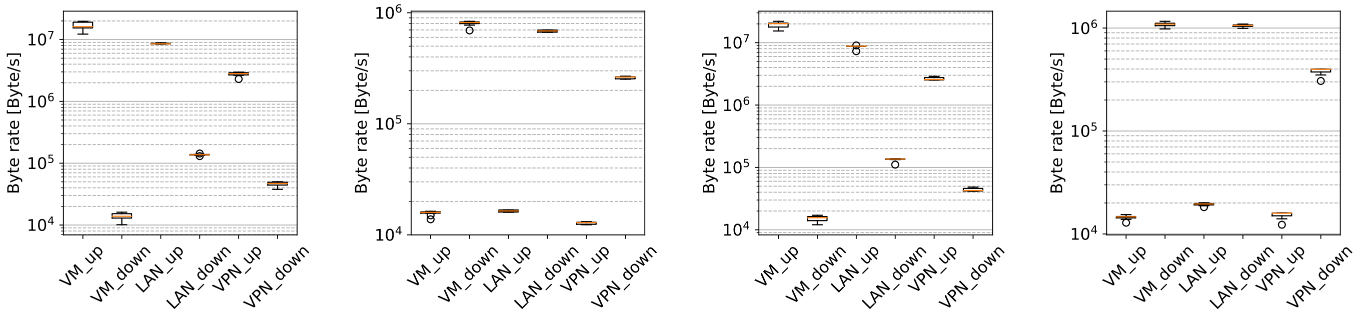


(a) Upload, medium size (b) Download, medium size (c) Upload, big size (d) Download, big size

Fig. 3: Goodput (ratio of Images Set total size over the total duration of transfer). Medium and Big sizes are for 100MB and 1GB, respectively.

nomenon can be related to protocol overhead associated to the network technology, as surfaced by the average packet size (reported in Table I). Indeed it can be noted that both the virtualized private network of the VM setup and the Gigabit Ethernet of the LAN setup use (Super-) Jumbo Frames, while the VPN setup is constrained by the 1500 bytes size of Internet Maximum Transfer Unit, and the additional overhead of the SSL sublayer.

Given the constraint posed by the nominal network bandwidth of the client and server Internet access connections (100Mbps among all direction combinations), the experimental setup of the network does not constitute a bottleneck. From the measurement results, an estimation is possible for the suitability of the current setup to fulfill the planned service level for the expected number of concurrent users. Under the hypothesis that each single user experiences the network performance as in the measurement campaign, current setup can accommodate more than 20 concurrent users for the Upload usage case: for the planned intake of concurrent users (10 in the current stage of the service deployment), this amounts to a $2\times$ overprovisioning factor. Regarding the Download scenario, the ratio between server-side upstream bandwidth and client downstream byte rate is greater than 90, leading to a $9\times$ overprovisioning factor. Hence, the current operating system is validated for suitability for the planned number of concurrent users, also allowing for servicing exceptional cases, and for ample time for planning an upscaling if the number of users admitted to the service is increased.



(a) Upload, medium size (b) Download, medium size (c) Upload, big size (d) Download, big size

Fig. 4: Average byte rate over the total duration of transfer. Medium and Big sizes are for 100MB and 1GB, respectively.

V. Conclusion

In this work we have described the biobanking system for diagnostic images implemented and and currently operated by *Bio Check Up Srl*. The design criteria that led to the choice of XNAT platform as a basis for the system have been described, and a measurement campaign for the assessment of network performance of the operating service has been conducted, for the envisioned usage case (remote access through VPN). Additional experiments have been conducted on a single-host reference setup, and a LAN setup, to better infer the impact of the (uncontrolled) Internet path in the VPN case. Results highlight that the different Upload/Download activities, performed with different tools provided by the XNAT platform, exhibit significantly different performance figures, and deserve further investigation. The analysis confirms that the current setup is able to satisfy the planned number of concurrent users with a wide margin ($2\times$ and $9\times$ the required level for Upload and Download use cases, respectively). Further analyses, suggested by the presented experimental campaign, regard the impact on the software and hardware middleboxes traversed by the service traffic, and the performance evaluation of alternative tools (and related protocols) to enact the same Upload/Download use cases. A qualitative analysis of Quality of Experience, based on scores provided by practitioners in medical and clinical research, will be investigated as well.

Acknowledgment

This work was partially funded by the Italian Ministry for Education, University and Research - Project "MOLIM ONCOBRAIN LAB Metodi innovativi di imaging molecolare per lo studio di malattie oncologiche e neurodegenerative" ARS01_00144 Prot. U.0013785 31-08-2018.

References

[1] C. Peterson, C. Hamilton, and P. Hasvold, *From innovation to implementation – eHealth in the WHO European Region*, 2016.

[2] J. O'Donoghue and J. Herbert, "Data management within mhealth environments: Patient sensors, mobile devices, and databases," *J. Data and Information Quality*, vol. 4, no. 1, 2012.

[3] E. S. of Radiology (ESR *et al.*, "Esr position paper on imaging biobanks," *Insights into imaging*, vol. 6, no. 4, pp. 403–410, 2015.

[4] J. Kinkorová, "Biobanks in the era of personalized medicine: objectives, challenges, and innovation," *EPMA Journal*, vol. 7, no. 1, p. 4, 2016.

[5] E. Neri and D. Regge, "Imaging biobanks in oncology: European perspective," *Future Oncology*, vol. 13, no. 5, pp. 433–441, 2017.

[6] L. Coppola, A. Cianflone, A. M. Grimaldi, M. Incoronato, P. Bevilacqua, F. Messina, S. Baselice, A. Soricelli, P. Mirabelli, and M. Salvatore, "Biobanking in health care: evolution and future directions," *Journal of translational medicine*, vol. 17, no. 1, p. 172, 2019.

[7] D. S. Marcus, T. R. Olsen, M. Ramaratnam, and R. L. Buckner, "The extensible neuroimaging archive toolkit," *Neuroinformatics*, vol. 5, no. 1, pp. 11–33, 2007.

[8] Y. Mehdipour and H. Zerehkafi, "Hospital information system (his): At a glance," *Asian Journal of Computer and Information Systems (ISSN: 2321–5658)*, vol. 1, no. 02, 2013.

[9] "Reference Guide in Information Technology for the Practicing Radiologist," https://www.acr.org/Practice-Management-Quality-Informatics/Informatics/Reference-Guide, accessed: 2020-04-15.

[10] R. H. Choplin, J. Boehme 2nd, and C. Maynard, "Picture archiving and communication systems: an overview." *Radiographics*, vol. 12, no. 1, pp. 127–129, 1992.

[11] "DICOM Standard," https://www.dicomstandard.org/, accessed: 2020-04-15.

[12] S. Das, A. P. Zijdenbos, D. Vins, J. Harlap, and A. C. Evans, "Loris: a web-based data management system for multi-center studies," *Frontiers in neuroinformatics*, vol. 5, p. 37, 2012.

[13] P. Kidambi, M. Sherwood, J. G. Parker, and R. DeVelvis, "Benchmarking medical image databases." in *IFAC HMS*, 2013, pp. 240–243.

[14] "Docker," https://www.docker.org/, acc.: 2020-04-15.

[15] S. Avallone, A. Pescapé, and G. Ventre, "Distributed internet traffic generator (d-itg): analysis and experimentation over heterogeneous networks," in *ICNP 2003 poster Proceedings, International Conference on Network Protocols, Atlanta, Georgia*, 2003.

[16] S. Avallone, D. Emma, A. Pescapé, and G. Ventre, "A distributed multiplatform architecture for traffic generation," in *International Symposium on Performance Evaluation of Computer and Telecommunication Systems, San Jose, CA, USA*, 2004.

[17] "Container Service Plugin," https://marketplace.xnat.org/plugin/container-services-plugin/, accessed: 2020-04-15.

[18] D. P. Pezaros, D. Hutchison, F. J. Garcia, R. D. Gardner, and J. S. Sventek, "Service quality measurements for ipv6 inter-networks," in *Twelfth IEEE International Workshop on Quality of Service, 2004. IWQOS 2004.* IEEE, 2004, pp. 129–137.