 DOI: 10.15496/publikation-29062

Storage infrastructures to support advanced scientific workflows

Towards research data management aware storage infrastructures

Dirk von Suchodoletz* Ulrich Hahn† Bernd Wiebelt* Kolja Glogowski* Mark Seifert*

*eScience Department, Computer Center, University of Freiburg, Freiburg, Germany
†High Performance and Cloud Computing Group, University of Tübingen, Tübingen, Germany

The operators of the federated research infrastructures at the involved HPC computer centers face the challenge of how to provide storage services in an increasingly diverse landscape. Large data sets are often created on one system and computed or visualized on a different one. Therefore cooperation across institutional boundaries becomes a significant factor in modern research. Traditional HPC workflows assume certain preliminaries like POSIX file systems which cannot be changed on a whim. A modern research data management aware storage system needs to bridge from the existing landscape of network file systems into a world of flexible scientific workflows and data management. In addition to the integration of large scale object storage concepts, the long term identification of data sets, their owners, and the definition of necessary meta data becomes a challenge. No existing storage solution on the market meets all of the requirements, and thus the bwHPC-S5 project must implement these features. The joint procurement and later operation of the system will deepen the cooperation between the involved computer centers and communities. The transition to this new system will need to be organized together with the scientific communities being shareholders in the storage system. Finally, the created storage infrastructures have to fit well into the growing Research Data Repositories landscape.

 DOI: 10.15496/publikation-29058

1 Motivation

Modern research has become increasingly digital, leveraging a wide variety of hardware, data collection instruments, and software to gather, process, and visualize data in multitudes of ways. Such digital workflows are getting more complex and the volume of data processed or created is ever rising. From the perspective of a researcher, the typical workflow traditionally is executed on a local machine and, when the amount of data and computation exceeds local resources, handled by a larger system like an (external) HPC cluster. With the availability of new tools and options to process and view data, more systems will become involved in workflows.

Data Intensive Computing (DIC) involves big data or methods like deep learning to provide new perspectives on existing data (Schneider et al., 2019). This would require to bring data and compute resources to a common location efficiently. Depending on the type of data and workflow envisioned special resources like GPGPU are required which are not present at every site. Within Baden-Württemberg both the compute systems like the bwForClusters and the bwCloud as well as the LSDFs are distributed over different physical locations. The HPC systems on various tiers are complemented by e. g. bwCloud compute capacities to allow pre- or post-processing runs which would be a waste of resources on HPC systems. Remote visualization facilities – special systems sitting near the data – become relevant to render data and stream the results without the need to copy large data sets to the local machine of the user.[1] Additionally, requirements of reproducible science, the better understanding of the value of research data, the objective of open data publication all change the definition of data management.

Modern data management should extend beyond the traditional data-handling performed by a single scientist. Researchers often do not standardize metadata, making interoperability and sharing difficult. Data curation, the selection of data sets of relevance, and the removal of irrelevant data is often not a formalized step in the workflow. It often takes place when files must be copied across across systems, or whenever quotas were exceeded.[2] A storage system designed with research data management in mind should at least provide multiple ways to both automate workflows over the data lifecycle.

[1] The bandwidth consumed to stream visualization is usually much lower than to copy terabytes of data in reasonable time.

[2] This is not the proper trigger to achieve a high quality of data sets as valuable data might get thrown away because of size limitations.

The text is structured as follows: It gives an overview of the current state of data management and its challenges in HPC. The limitations of state-of-the-art of file systems used in HPC related scientific workflows will be explored. The requirements stemming from today's and future workflows of HPC and DIC user communities will be discussed. Further, it explores options to extend and optimize existing scientific workflows and local compute infrastructure setups. From this discussion it tries to provide a coherent conceptual framework for the design of a research data management aware large scale data facility. Storage-for-Science (bwSFS) is a DFG and state supported research infrastructure project to provide joint storage and research data management functionality for various research groups in Tübingen, Freiburg, Ulm and Stuttgart. The presented article extends upon the discussion provided in the paper published for the DFN-Forum 2017 (Meier et al., 2017) and outlines the framework for the intended design.

2 Scientific data on the tier-3 HPC environment

When scientists use a storage system, their satisfaction largely depends on how easy it is to access and on whether it is available in all the various usage scenarios. Usually, the optimization of one characteristic might degrade others. E. g. a local home directory on the scientist's laptop offers the easiest form of access to the data but is usually more severely capacity limited compared to home directories or data shares provided e. g. via NFS or SMB from specialized storage appliances. Network file systems require a common authentication service and properly mapped identities and face limitations regarding performance in wide area network operation. Data can be shared among colleagues to a certain degree from a networked file system provided by a storage appliance but can not easily extend beyond institutional boundaries.

All file systems (as well as their POSIX completeness and their performance) rely on the operating system support of the machine they are deployed on. If different systems involved in a scientific workflow do not feature the same operating system additional challenges will arise. These include the availability features a file system provides.[3]

[3] E. g. some metadata and additional information is lost when copying data from a Mac OS X HFS to Linux or Windows or vice versa. Another problem is the availability of a certain network file system in an operating system or version of it (French et al., 2007).

In the typical HPC world specialized local or near local storage in the geographical sense is attached to each compute infrastructure. Most bw*Clusters share these storage characteristics.

Fast parallel distributed file systems like e. g. BeeGFS, Lustre or GPFS in HPC are meant to provide a short-term high-performance storage to cater to many concurrent, parallel jobs.[4] BeeGFS and Lustre have free-of-charge[5] or open-source(-like) licenses but generally should be operated with a support contract.[6] GPFS, now marketed under the name Spectrum Scale, is another option but proprietary and rather expensive.[7] All are POSIX compliant and require a client-side kernel module. Parallel file systems are designed for capacity plus speed and are typically configured with a low level of hardware redundancy. The HPC system in Freiburg e. g. is set up with with two metadata daemons configured with two 1200 GB NVMe mirrored disks each running on one physical machine. The data is stored on four rack mount storage containers with sixty 4 TB disks each controlled by four machines. Each machine provides six RAID 6 storage targets with roughly 32 TB of capacity adding up to 768 TB of total usable parallel storage capacity.[8] These file systems usually offer large volume and are available during the time of actual computing, but not necessarily provided for long term storage of valuable data sets. The storage space even if vast is limited as it is shared among many users. Parallel file systems in HPC are usually not backed up.

In an era dominated by spinning magnetic disks, parallel storage offered an attractive compromise between speed, capacity and costs. However, depending on built-in redundancy, additional SSD- or NVMe-caching, chosen filesystem, and solution provider it can be quite costly in some configurations. It usually offers the largest user and group quotas available on an HPC system. Often it is possible to allocate several tens of terabytes per group. To manage concurrent use the parallel file systems offer means to set quotas e. g. on the amount of space or inodes taken. Nevertheless, the huge quantities of storage provided produces new challenges for users, operators, and system tools.

[4] For discussion of various storage technologies, future development and parallel file systems refer to the survey (Lüttgau et al., 2018) and the report (Brinkmann et al., 2017).
[5] See the end-user license agreement at `http://www.beegfs.io/docs/BeeGFS_EULA.txt`.
[6] Longer outages of a crucial part of an HPC system are costly as the whole system is affected.
[7] Aside project prices a per-socket license is charged.
[8] The costs per terabyte of the BeeGFS in NEMO run below 450 €.

Traditional system tools for listing files or accessing file metadata become unusably slow when run with significant amounts of files, degrading performance for other users. To properly manage scarcity of resources and create a kind of fairshare mechanism for storage a workspace regime is deployed for the bw*Clusters.[9] Workspace tools allow for automatic cleaning of unused data after a predefined amount of time. On HPC cluster NEMO users can allocate a workspace for a maximum of 100 days. After that period the workspace gets deleted automatically if the duration is not extended manually by the creator of this workspace. Users can extend their workspaces a configurable number of times. They can have multiple workspaces that are only limited by their quota. Workspaces provide a mechanism to clean old and unused data automatically if no manual interaction is made. This is alleviates the common problem of old inactive accounts occupying significant resources which could be better allocated to active projects.

Home directories are usually not identical to traditional homes in a multi-user environment like at a typical research institution, especially in bwHPC as users from different universities share domain specialized clusters for their research.[10] In the NEMO cluster environment the user home is provided via NFS from a enterprise grade storage appliance.[11] The system offers several levels of redundancy and snapshots to allow the user to go back to earlier versions of files. The space provided is intended to host the relevant research related files and data. Nevertheless, it is not the best location to store large amount of research data, as the size of disk quota provided is rather low. The per gigabyte costs are driven up by the various levels of redundancy and number of snapshots taken. Data can be stored for the lifetime of an account active for the particular cluster. Because of performance implications both on the storage appliance and on the speed of remote file access in the cluster users are strongly discouraged to run (parallel) HPC jobs out of their homes.

Scratch space is typically a fast local file system in each node which is node-exclusive with no redundancy and not shared among other nodes within the cluster. Most installations in the bwHPC federation deploy local scratch space, which is

[9] Workspace tools on GitHub: `https://github.com/holgerBerger/hpc-workspace` (visited on 13. 02. 2019)

[10] The situation is slightly different at the other bwHPC sites as some operators mount additional shares for local users or deploy parts of the parallel file system to be used as home directories.

[11] The per terabyte costs run at about 1200 €.

usually set into the relation to the number of CPU cores in the compute node. It is meant to allow processes to write temporary results in a fast fashion or to take debugging and logging output. All modern systems have solid state disks installed, but with a rather moderate capacity of in between 200 and 500 gigabytes per node. To overcome these capacity-limitations, special block device extensions featuring a concatenation of the local disk with network block device part backed by the parallel file system have been developed by the operators of the JUSTUS cluster in Ulm (Neuer et al., 2016).

A method to create an on-demand parallel file system backed scratch space is BeeOND (Brinkmann et al., 2017). It uses local scratch space combined with high speed networks like InfiniBand and Omni-Path to create larger storage capacities for a certain parallel jobs. The costs of scratch space directly correlates with the costs of the individual disks in each node. The big advantage is that the local scratch space usually consists of SSDs and therefore file access and write and read performance can be better than on the central storage which can still have a traditional hard disk setup for the storage targets. With BeeOND the local SSDs is utilized as a central storage for parallel jobs, permitting use of storage which otherwise would be idle and unused. That way the load on the central storage is reduced and misbehaving jobs with high IO do not affect other jobs. BeeOND is not configured with any levels of (hardware) redundancy which applies to the worker nodes as well. Data can be stored for the duration of the job lifetime but is expected to be moved away to a workspace at the end of the job. The on-demand scratch space is destroyed at the end of the job and data which has not been saved at this point is lost.

2.1 Limitations of current setups

There are a couple of limitations for modern data management in today's solutions deployed in the various bwHPC clusters. They are driven by a combination of factors. Storage in HPC systems is not meant to safeguard data for longer periods. The publication of data sets is outside the scope of HPC clusters and only partially answered by projects like bwDATAarchive. Not only the quota in the dimension of time or capacity is limited but also the level of redundancy and data backup. Most

HPC systems are set up as »closed boxes«, the available storage is not exposed outside the cluster itself and data must be copied in and out.[12]

The amount of data regarding file sizes and number of files play a significant role. Copying of data is acceptable for a certain quantity of files and size but becomes unbearable if a single operation exceeds a couple of hours. Researchers face versioning and synchronization challenges as the same data sets may exist on different systems at the same time. They need to apply stringent data management to avoid conflicts or to run into quota limitations. Additionally, long preparation times might require proper data staging to avoid long startup delays of scheduled HPC jobs. To circumvent inefficient use of HPC systems, preprocessing, post processing or visualization of data is expected to take place somewhere else.

All mentioned storage options adhere to the POSIX standard which is convenient for application development but might restrict parallel performance (Lüttgau et al., 2018). For performance and simplicity reasons many systems apply IP based security which is acceptable within tightly controlled physical installations and networks. Parallel file systems usually skip certain security checks and do not record access time. The latter would be relevant e. g. for hierarchical storage management systems.

Further limitations were experienced when experimenting with Virtual Research Environments (VRE). VREs abstract scientific workflows – a packaged software stack with certain configuration – in a virtual machine or a container (Bauer, Suchodoletz et al., 2019). To allow simpler sharing of environments or to run in widely distributed setups VREs are made independent of the underlying physical hardware. When moved from one cluster to another or to a cloud resource (Heidecker et al., 2017; Bührer et al., 2018) the task of a VRE remains the same but the accessibility of data sources and sinks poses a challenge as they break out of the walled garden of simple IP based access control and security.

Sharing of datasets becomes more relevant as cooperation of (geographically distributed) research groups, as well as the demand for re-using existing data, grows (Tenopir et al., 2011). With the predominant POSIX file systems the limited export options for data sharing are typically confined within a single HPC cluster or work group. Further limitations apply to sophisticated ACL settings or different views on

[12] The HPC systems in Karlsruhe and Heidelberg provide an exemption as they offer the mounting of local file system from local user groups or provide direct access to the SDS@hd service running on the Large Scale Data Facility (LSDF) (Bauer, Suchodoletz et al., 2019; Baumann et al., 2017).

a collection of files. In a traditional file system everything is hierarchically organized in directories, different views in the form of an alternatively structured directory tree for different users are not available.[13] Assigning rights to some subdirectory or share exceeding those of the top level directory are possible. But, it becomes increasingly difficult to monitor if everyone requiring access has the proper rights and if the rights are completely updated, when a person changes its role.

3 Rethinking scientific storage – towards bwSFS

A next generation scientific storage system should include features to support research data management in the sense that data could stay within the same storage system over the complete data life cycle – spanning from data acquisition over the various stages of computation, visualization to long term archiving and publication (Tenopir et al., 2011; Demchenko et al., 2012; Meier et al., 2017). The system should provide tools and services to support researchers in their data management tasks and various workflow designs. With the new bwHPC-S5 project started mid 2018, data management in HPC and DIC will get added to the development and support activities (Barthel et al., 2019). Federated services like EUDAT can provide guidance on which services are to be provided an how the several challenges are tackled (Lecarpentier et al., 2013; Ardestani et al., 2015).

The HPC sites of BinAC in Tübingen and NEMO in Freiburg[14] plan to complement the compute infrastructure by a research data management aware storage system also focused on their scientific cluster communities. The system is planned to run in a cooperated, federated fashion spanning both locations. After multiple rounds of discussion with future users during the grant application process and after the requested sum was approved a couple of characteristics and abstract features of a future research data management system can be summarized:

- The system needs to scale well: regarding both the total capacity and the size of individual files.

[13] Links, possible in some file systems, are not an alternative as they are not updated, if a file gets relocated.

[14] More information on the individual clusters and bwHPC in general can be found at `http://www.bwhpc.de/ressourcen.php`.

- It should offer a good compromise of price per terabyte, performance, and capacity over the whole system. Currently a hierarchical storage solution is planned. IO metrics of individual workflows will help to determine the requirements for the according hierarchy levels.
- The system should provide various levels of redundancy including geographically distributed locations.
- For trusted, reproducible scientific workflows the archived data should be immutable.
- Definable service classes for different qualities of storage should be provided. A user should be able to declare, that a certain data set should be kept geographically redundant or that it may not be copied to a location outside a given campus.
- The system should provide (high-performance) interfaces to other research infrastructures like HPC clusters and cloud systems.
- It should allow the automation of workflows by providing appropriate interfaces like REST APIs allowing asynchronous operation so that users do not need to wait until a certain storage related process is finished.
- The system needs to implement an identity mapping for users and long term data owners and data objects abstracted from the local IDM to federated IDM systems.
- Copying of large data sets should be avoided whenever possible. Many storage systems solve this by using references to data sets and update these, when changes occur. The future planning of computational infrastructures will be more and more influenced by the physical location of data.
- Monitoring and accounting features are required to provide administrators with enough insight to detect and avoid bottlenecks both in performance and system usage.

Beside the technical specifications the non-technical characteristics should adhere to the following ideas:

- The implemented solution should support at least in some abstract form the FAIR principles (Wilkinson et al., 2016) and the data life cycle.
- It should support the implementation of workflows or should provide the relevant stubs to deal with the various stages in the data life cycle:

  - Registration of storage capacity (of a certain kind, for a project with time limits).
  - Curation of data sets (e. g. certain files are automatically removed after a certain period because of a tag, provide a UI for interactive use, API).
  - Enrichment of data sets with metadata in the various stages of the life cycle.
- Workflows should be automated as far as possible by the system. Manual intervention should nevertheless be possible.
- Avoid vendor and technology lock-ins as the system is intended to exceed the short living technological life cycles.
- Provide various (discipline defined) interfaces for data publication via external third party services to blend into the growing Research Data Repositories landscape (Pampel et al., 2013).
- An interesting feature of a research data management repository would be to allow authenticated third party user comments on the data sets.[15]
- A sustainable and quality assured research data management will require certification of the system and the involved workflows. They are various certification options available, see e. g (RLG-NARA Task Force, 2007; DINI, Working Group Electronic Publishing, 2011; CCSDS Secretariat, 2011; ESF & EUROHORCs (2011), 2011).

The ongoing consultation with the stake- and shareholder of the federated research management system will further complete the matrix of the characteristics of the data sets like ACLs, size, value, share-ability. Alternatives to the up to now prevalent network file systems, like the S3 protocol or alike for object store systems, will be evaluated to gradually change workflows to better suit them to a sustainable research data management.

[15] Challenges remain how to distinguish between (il-)legitimate posts, though (Golbeck, 2018).

3.1 Security and privacy implications

Beside technical considerations legal implications will drive the requirements of a shared, multi-user storage systems. Funding agencies require retention of data for defined periods, commercial partners might want to restrict access to data sets and sensitive data like from clinical trials, surveys should not be disclosed to third-parties, or some projects require embargos for defined timespans. Certain datasets should not get copied outside the home institution even if redundancy would require it. Secure storage is not only achieved through local encryption of disks and filesystems but also through access management and secure transportation. Such challenges are discussed in the connection with cloud storage (Li et al., 2013; Wang et al., 2012). Security and privacy matters could be part of certification processes.

3.2 Mapping of identifiers

The provisioning of storage infrastructures in a broader sense regarding potential users and storage periods requires proper handling of user identities and identifiers for the objects. The identifiers for users will differ from traditional user accounts provided by the IDM of the local institution. This is already the case for the existing cluster and cloud logins. Unified credentials might be desirable for the network file systems which could in turn create challenges for the storage systems and proper user mapping. Switching to object store oriented operation access tokens could provide a more versatile solution than complex ACLs. The management of the tokens can be shifted more easily to the user side than traditional role management within a file system.

The need to identify data sets is completely different for the system and the users. Identifiers of data sets need to be mapped between the system and the interfaces for data management e. g. for indexing, citation, or access by external infrastructures. Metadata of various types[16] is crucial here as implicit information stored within a POSIX file system like access time or ownerships might get lost when files are moved.

[16]Ranging from common structures to describe projects, runtimes and owners to community specific information structures.

4 Managing transition

In Freiburg, the existing local compute services of both the HPC clusters of NEMO and the ATLAS experiment, as well as the bwCloud SCOPE and de.NBI cloud become more tightly integrated. For efficiency and manageability reasons they share the relevant base infrastructures like server racks, high speed Ethernet connectivity and hardware monitoring. Workflows which require the processing of data in different facilities like preprocessing in the cloud and parallel computing in NEMO should not require the copying of data sets as the relevant storage systems get more easily accessible within the conglomerate.

Learning from the results of the bwVisu project (Schridde et al., 2017) visualization servers could be added to an HPC cluster or a visualization cloud or cluster could be set up close to a large storage system for scientific data using cloud operation or container models to share the hardware. Services like CVMFS get deployed in a way to be available in all kinds of infrastructures ranging from the HPC clusters to clouds and VREs. Furthermore the provisioning of LSDF resources like SDS@hd (Baumann et al., 2017) from cloud, HPC and VREs got explored and enabled. Local storage of the various research groups should be available as well. This results in a matrix of accessible shares.

As large scale compute infrastructures become commonplace for most researchers and workflows get more diverse future data services should keep these developments in mind. In Freiburg a step by step approach is envisioned to provide truly flexible workflows. It works on different stages. As a compromise to the preexisting security and access control models of network file systems VLANs are used to virtually separate the installed infrastructures. As a flexible boot model defines into which operation mode a server is deployed the IP networks are assigned dynamically (Bauer, Messner et al., 2019). Thus, each infrastructure can mount resources depending on the operation model and workflow in use. Resources from the other infrastructures are not visible to the nodes because of the logical network separation.

To include cloud resources into a workflow, e. g. for pre and/or post processing a new rather flexible storage option becomes available. It provides a completely independent storage environment: Ceph block storage which can get configured and attached to VMs in various (user defined) ways for a wide range of different capacities. It adds another cost effective solution for temporary data. In further steps new approaches like the inclusion of S3 object storage will be evaluated. Suitable workflows are evaluated together with various research groups. A couple of *tiger teams* – an ad hoc group formed of scientists and practitioners to work on a particular issue – around data management were created within the context of bwHPC-S5 (Barthel et al., 2019).

To better support future data management data sets are to be tagged and/or described when created or processed, so that information becomes available and get added to metadata automatically. Optimally individual scientific workflows attach metadata already available or generated by them.

5 Outlook

The definition and design of a research data management system is an ongoing process and only tentative results can be given here. After securing the initial funding the next steps include the definition of required features and the decision on their realization. Not all requirements are met in systems available on the market yet. Hierarchical storage management systems offer a good yield on capacity, performance and prices per terabyte. Nevertheless, they need to be complemented by workflows helping both researchers and system providers to implement sustainable research data management. Here, the extended bwHPC-S5 project can provide the necessary resources to implement missing bits and pieces in-house and support the various communities in their data management needs. The joint procurement of the system both in Tübingen and Freiburg helps to gain a significant scaling-up and features like geographical redundancy. The sizing of the system allows to include more communities beside the core HPC cluster users of NEMO and BinAC. Upon a generic system core of standard technical features community specific options like specialized metadata and data services can be added. A similar »growth and upgrade path« as established with the HPC clusters should be implemented for the storage system as well: It should be possible that research groups, institutes etc. can pool-in financial resources and get equivalents in storage capacities of the requested quality level.

The transition from ad-hoc data handling today to sustainable data management is to be organized together with the scientific communities and institutions being shareholders in the storage system. On the technical layer a couple of steps were already implemented by the dynamic provisioning of HPC and cloud nodes via remote boot stateless provisioning (Bauer, Messner et al., 2019) or (Schmelzer et al., 2014) including the dynamically configured access to the relevant storage systems. Governance like already implemented globally and locally for the HPC systems (Wesner et al., 2017; Suchodoletz et al., 2017) needs to be in place as well. Monitoring and accounting need to provide the relevant input for decision making and the sustainable re-financing of the system in the long run. Finally, the automation of workflows from the definition of project requirements in a Data Management Plan to the deployment in the storage system would be an attractive feature (Bakos et al., 2018).

Acknowledgement

The research infrastructures described in this publications are part of the bwHPC project sponsored by the Ministry of Science, Research and the Arts, Baden-Württemberg (MWK) and the German Science Foundation. The same applies to the work presented from the ViCE project which got supported by the MWK as part of it's eScience initiative. The support is gratefully acknowledged.

Corresponding author

Dirk von Suchodoletz: dirk.von.suchodoletz@rz.uni-freiburg.de
eScience Department, Computer Center, University of Freiburg
Hermann-Herder-Str. 10, 79104 Freiburg, Germany

ORCID

Dirk von Suchodoletz https://orcid.org/0000-0002-4382-5104
Ulrich Hahn https://orcid.org/0000-0003-4471-9263
Bernd Wiebelt https://orcid.org/0000-0003-2771-4524
Kolja Glogowski https://orcid.org/0000-0002-1361-5712
Mark Seifert https://orcid.org/0000-0002-1042-6107

References

Ardestani, S. B. et al. (2015). »B2share: An open escience data sharing platform«. In: *2015 IEEE 11th International Conference on e-Science (e-Science)*. IEEE, pp. 448–453.

Bakos, A., T. Miksa and A. Rauber (2018). »Research Data Preservation Using Process Engines and Machine-Actionable Data Management Plans«. In: *International Conference on Theory and Practice of Digital Libraries*. Springer, pp. 69–80.

Barthel, R. and J. Salk (2019). »bwHPC-S5: Scientific Simulation and Storage Support Services. Unterstützung von Wissenschaft und Forschung beim leistungsstarken und datenintensiven Rechnen sowie großskaligem Forschungsdatenmanagement«. In: *Proceedings of the 5th bwHPC Symposium. HPC Activities in Baden-Württemberg*. Freiburg, September 2018. 5th bwHPC Symposium. Ed. by M. Janczyk, D. von Suchodoletz and B. Wiebelt. TLP, Tübingen, pp. 17–28. DOI: 10.15496/publikation-29039.

Bauer, J., M. Messner et al. (2019). »A Sorting Hat For Clusters. Dynamic Provisioning of Compute Nodes for Colocated Large Scale Computational Research Infrastructures«. In: *Proceedings of the 5th bwHPC Symposium. HPC Activities in Baden-Württemberg*. Freiburg, September 2018. 5th bwHPC Symposium. Ed. by M. Janczyk, D. von Suchodoletz and B. Wiebelt. TLP, Tübingen, pp. 217–229. DOI: 10.15496/publikation-29055.

Bauer, J., D. von Suchodoletz, J. Vollmer and H. Rasche (2019). »Game of Templates. Deploying and (re-)using Virtualized Research Environments in High-Performance and High-Throughput Computing«. In: *Proceedings of the 5th bwHPC Symposium. HPC Activities in Baden-Württemberg*. Freiburg, September 2018. 5th bwHPC Symposium. Ed. by M. Janczyk, D. von Suchodoletz and B. Wiebelt. TLP, Tübingen, pp. 245–262. DOI: 10.15496/publikation-29057.

Baumann, M., V. Heuveline, O. Mattes, S. Richling and S. Siebler (2017). »SDS@hd–Scientific Data Storage«. In: *Proceedings of the 4th bwHPC Symposium October 4th, 2017, Alte Aula Eberhard Karls Universität Tübingen*. Ed. by J. Krüger and T. Walter, pp. 32–36. DOI: 10.15496/publikation-25204.

Brinkmann, A., K. Mohror and W. Yu (2017). *Challenges and Opportunities of User-Level File Systemsfor HPC*. Tech. rep. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States). DOI: 10.2172/1424647. URL: https://www.osti.gov/servlets/purl/1424647.

Bührer, F. et al. (2018). »Dynamic Virtualized Deployment of Particle Physics Environments on a High Performance Computing Cluster«. In: *Computing and Software for Big Science*. arXiv: 1812.11044 [physics.comp-ph].

CCSDS Secretariat (2011). *Audit and certification of trustworthy digital repositories. Recommended Practice*. Recommendation for Space Data System Practices. Version CCSDS652.0-M-1. Space Communications and Navigation Office: Council of the Consultative Committee for Space Data Systems. URL: https://public.ccsds.org/pubs/652x0m1.pdf.

Demchenko, Y., Z. Zhao, P. Grosso, A. Wibisono and C. de Laat (2012). »Addressing Big Data challenges for Scientific Data Infrastructure«. In: *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pp. 614–617. DOI: 10.1109/CloudCom.2012.6427494.

DINI, Working Group Electronic Publishing (2011). *DINI-Certificate Document and Publication Services 2010 [March 2011]*. Deutsche Initiative für Netzwerkinformation (DINI). DOI: 10.18452/1494.

ESF & EUROHORCs (2011) (2011). *Basic Requirements for Research Infrastructures in Europe*. URL: http://www.dfg.de/download/pdf/foerderung/programme/wgi/basic_requirements_research_infrastructures.pdf.

French, S. M. and S. Team (2007). »A New Network File System is Born: Comparison of SMB2, CIFS and NFS«. In: *Linux Symposium*. sn, p. 131.

Golbeck, J. (2018). »Data We Trust—But What Data?« In: *Reference & User Services Quarterly* 57.3, pp. 196–199.

Heidecker, C., M. Giffels, G. Quast, K. Rabbertz and M. Schnepf (2017). »High precision calculations of particle physics at the NEMO cluster in Freiburg«. In: *Proceedings of the 4th bwHPC Symposium October 4th, 2017, Alte Aula Eberhard Karls Universität Tübingen*. Ed. by J. Krüger and T. Walter, pp. 28–31. DOI: 10.15496/publikation-25203.

Lecarpentier, D. et al. (2013). »EUDAT: a new cross-disciplinary data infrastructure for science«. In: *International Journal of Digital Curation* 8.1, pp. 279–287. DOI: 10.2218/ijdc.v8i1.260.

Li, M., S. Yu, Y. Zheng, K. Ren and W. Lou (2013). »Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption«. In: *IEEE transactions on parallel and distributed systems* 24.1, pp. 131–143.

Lüttgau, J. et al. (2018). »Survey of Storage Systems for High-Performance Computing«. In: *Supercomputing Frontiers and Innovations* 5.1, pp. 31–58. DOI: 10.14529/jsfi180103.

Meier, K., B. Grüning, C. Blank, M. Janczyk and D. von Suchodoletz (2017). »Virtualisierte wissenschaftliche Forschungsumgebungen und die zukünftige Rolle der Rechenzentren«. In: *10. DFN-Forum Kommunikationstechnologien, 30.-31. Mai 2017, Berlin, Gesellschaft für Informatik eV (GI)*, pp. 145–154.

Neuer, M. et al. (2016). »Motivation and Implementation of a Dynamic Remote Storage System for I/O Demanding HPC Applications«. In: *International Conference on High Performance Computing*. Springer, pp. 616–626.

Pampel, H. et al. (2013). »Making Research Data Repositories Visible: The re3data.org Registry.« In: *PLOS ONE* 8.11. DOI: 10.1371/journal.pone.0078080.

RLG-NARA Task Force (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist*. Report. Version 1.0. RLG- National Archives and Records Administration Digital Repository Certification Task Force. URL: http://bibpurl.oclc.org/web/16713.

Schmelzer, S., D. von Suchodoletz, M. Janczyk and G. Schneider (2014). »Flexible Cluster Node Provisioning in a Distributed Environment«. German. In: *Hochleistungsrechnen in Baden-Württemberg. Ausgewählte Aktivitäten im bwGRiD 2012*. Beiträge zu Anwenderprojekten und Infrastruktur im bwGRiD im Jahr 2012. Ed. by J. C. Schulz and S. Hermann. KIT Scientific Publishing, Karlsruhe, pp. 203–219. ISBN: 978-3-7315-0196-1. DOI: 10.5445/KSP/1000039516. URN: urn:nbn:de:0072-395167.

Schneider, G. et al. (2019). »Umsetzungskonzept der Universitäten des Landes Baden-Württemberg für das High Performance Computing (HPC), Data Intensive Computing (DIC) und Large Scale Scientific Data Management (LS^2DM)«. Gekürzte Fassung. In: *Proceedings of the 5th bwHPC Symposium. HPC Activities in Baden-Württemberg.* Freiburg, September 2018. 5th bwHPC Symposium. Ed. by M. Janczyk, D. von Suchodoletz and B. Wiebelt. TLP, Tübingen, pp. 3–16. DOI: 10.15496/publikation-29040.

Schridde, D., M. Baumann and V. Heuveline (2017). »Skalierbare und flexible Arbeitsumgebungen für Data-Driven Sciences«. In: *E-Science-Tage 2017: Forschungsdaten managen*. Ed. by J. Kratzke and V. Heuveline. Heidelberg: heiBOOKS, pp. 153–166. DOI: 10.11588/heibooks.285.377.

Suchodoletz, D. von, B. Wiebelt and M. Janczyk (2017). »bwHPC Governance of the ENM community«. In: *Proceedings of the 3rd bwHPC-Symposium.* (2016). Ed. by S. Richling, M. Baumann and V. Heuveline. Heidelberg: heiBOOKS. DOI: 10.11588/heibooks.308.418.

Tenopir, C. et al. (2011). »Data Sharing by Scientists: Practices and Perceptions«. In: *PLOS ONE* 6.6, pp. 1–21. DOI: 10.1371/journal.pone.0021101.

Wang, C., Q. Wang, K. Ren, N. Cao and W. Lou (2012). »Toward secure and dependable storage services in cloud computing«. In: *IEEE transactions on Services Computing* 5.2, pp. 220–232.

Wesner, S., D. von Suchodoletz, B. Wiebelt, G. Schneider and T. Walter (2017). »Overview on governance structures in bwHPC«. In: *Proceedings of the 3rd bwHPC-Symposium: Heidelberg 2016.* (2016). Ed. by S. Richling, M. Baumann and V. Heuveline. Heidelberg: heiBOOKS. DOI: 10.11588/heibooks.308.418.

Wilkinson, M. D. et al. (2016). »The FAIR Guiding Principles for scientific data management and stewardship«. In: *Scientific data* 3. DOI: 10.1038/sdata.2016.18.