

A Data Management Framework for Urgent Geoscience Workflows

Jason Cope and Henry M. Tufo

Department of Computer Science, University of Colorado at Boulder, 430 UCB,
Boulder, CO, 80309-0430, USA
{jason.cope, henry.tufo}@colorado.edu

Abstract. The emerging class of urgent geoscience workflows are capable of quickly allocating computational resources for time critical tasks. To date, no urgent computing capabilities for data services exists. Since urgent geoscience and Earth science workflows are typically data intensive, urgent data services are necessary so that these urgent workflows do not bottleneck on inappropriately managed or provisioned resources. In this paper we examine emerging urgent Earth and geoscience workflows, the data services used by these workflows, and our proposed urgent data management framework for managing urgent data services.

1 Introduction

The emergence of Grid computing as a viable high-performance computing (HPC) environment has provided several innovative technologies that enhance traditional scientific workflows. Dynamic data driven applications and workflows in particular benefit from improvements in data integration technology, distributed and dynamic computing resource integration, and wide-area network infrastructure. Recent research into urgent computing systems has further improved several of these workflows that perform emergency computations. Examples of these applications and workflows include Linked Environments for Atmospheric Discovery (LEAD) [1], the Southern California Earthquake Center's (SCEC) TeraShake [2], the Southeastern Universities Research Association (SURA) Coastal Ocean Observing and Prediction (SCOOP) project [3], and the Data Dynamic Simulation for Disaster Management project which is developing a Coupled Atmosphere-Fire (CAF) workflow for wildfire prediction [4].

The LEAD and SCOOP projects successfully use the Special PRIORITY and Urgent Computing Environment (SPRUCE) to obtain high-priority access to the shared computing resources available on the TeraGrid [5]. SPRUCE provides project users with elevated and automated access to TeraGrid computational resources so that high-priority applications run immediately or as soon as possible. SPRUCE currently provides urgent computational resource allocation capabilities but does not yet support urgent storage or data management capabilities. Urgent storage and data management capabilities provide prioritized usage of storage resources, such as file systems, data streams, and data catalogs.

Since a common definition for a scientific workflow is the flow of data between computational processes [6], providing urgent storage and data management is an essential and currently absent capability for urgent computing workflows. Supporting end-to-end urgent computing workflows requires support for common data capabilities, such as data storage, access, search, and manipulation capabilities, required by these workflows.

Our framework provides workflows and users with several urgent storage and data management capabilities, including the configuration of Service Level Agreements (SLAs) and Quality of Service (QoS) for data services, management of urgent and non-urgent data in shared computing environments, and autonomic management infrastructure that can adapt and tune data services without administrator intervention. These capabilities are designed as a series of shims that can integrate with existing data management infrastructure. In this paper, we present our proposed approach and framework in further detail. Section 2 describes the data requirements for urgent geoscience applications. Section 3 describes current urgent computing infrastructure. Section 4 describes common data services available to geoscience workflows. Section 5 describes our urgent data management framework. In the final sections, we present future work and conclusions.

2 Urgent Geoscience Applications and Grids

Advances in data and resource integration tools foster a computing environment capable of executing time critical workflows. These time-critical or urgent computing workflows harness distributed computational and data resources to quickly and reliably execute applications. The geoscience and Earth science community developed several applications with urgent computing use cases, such as earthquake, severe weather, flooding, and wildfire modeling applications. A typical characteristic of these applications is that they are I/O intensive. These applications often generate or ingest large amounts of data using various data resources, such as sensor networks, archival storage, and distributed storage systems. An example urgent application whose I/O requirements have been thoroughly analyzed in past work is SCEC TeraShake. The TeraShake simulations are constrained by data management resources because of the large amount of data produced [7]. A high-resolution SCEC simulation generated a total of 40TB of data on a 36TB storage resource. These limitations required developers to move the data as it was generated to other storage resources [2].

The integration of streaming data is demonstrated in two urgent computing workflows. LEAD uses Calder to integrate various data streams into the simulation environment. Calder can accommodate variable data sizes, data generation rates, and user access loads [8]. The CAF workflow has integrated sensor data and shown that prefetch of the data can improve application performance [9]. While the sensor data streams may not be high volume, their performance is limited by lack of network capacity and storage availability. In order to achieve urgent data management capabilities, the data requirements for the various ap-

plications must be accounted for. Other services, such as data processing tasks, utilize both computational and storage resources. Conflict free allocation of multiple resources that satisfy QoS requirements is necessary.

Several of these urgent workflows utilize Web services and service-oriented architectures. Both the LEAD and SCOOP projects utilize Web services to interface with a suite of data services. These services include access to archival storage, distributed storage systems, and metadata catalogs. To complicate matters, these workflows are dynamic and are not limited to urgent computing use. Many of these services, such as UCAR's Unidata services, are available for use by the general Earth science community. Access to these resources or services must be appropriately provisioned based on the need and urgency of the request.

3 Urgent Computing Infrastructure

The Special PRiority and Urgent Computing Environment (SPRUCE) [5] enables on-demand resource allocation, authorization, and selection for urgent computing applications. This environment provides on-demand access to shared Grid computing resources with a token-based authorization framework. SPRUCE allows Virtual Organizations (VOs) to utilize existing computing infrastructure for time critical tasks instead of procuring dedicated resources for these tasks. Users submitting SPRUCE jobs specify a color-coded urgency parameter with their job description. SPRUCE authorizes the urgent job by verifying that a user is permitted to execute tasks with the specified urgency on the target resource. Each VO defines policies for how the urgent tasks are handled on a per-resource basis. For example, a resource provider may choose to preempt non-urgent jobs for high-priority tasks or to give the urgent tasks next-to-run privileges. The infrastructure is currently deployed on several TeraGrid resources, including the NCAR's Frost Blue Gene/L, ANL's DTF TeraGrid cluster, and the SDSC's DataStar and DTF TeraGrid cluster. The LEAD and SCOOP projects use SPRUCE for urgent allocation of computational resources.

While SPRUCE currently provides access to computational resources, in the future it could also be adapted to manage other resources common in workflows, including storage and network resources. To completely support end-to-end urgent computing workflows in Grids, the usage and performance of storage and network resources must be accounted for in urgent computing management infrastructure. Therefore, we proposed the development of an urgent data management framework and services to support data-related tasks in urgent computing workflows. These capabilities will provide the appropriate SLAs and QoS for data services used in urgent computations. Several components are required to adapt current urgent computing capabilities to support these new resource types and several new capabilities are required to support the data requirements of urgent computing workflows. Tools are necessary to integrate existing urgent computing authorization infrastructure with common data services. Additional resource management tasks and processes are necessary to manage data products for Grid resources, resource users, and Grid workflows executing on these resources. In-

frastructure is also required to coordinate access to the multiple urgent resources and to ensure that conflicts in usage do not occur. Our proposed framework will provide these capabilities for common data services used by urgent applications.

4 Common Data Services

Across the spectrum of geoscience and Earth science Grids, there are several common data services available. These services can be classified as data storage, management, and processing services. The capabilities provided by these services are meant for general use and most provide little or no support for QoS, SLAs, or prioritized access to the data resources. The management of storage and data resources for urgent computing workflows is not addressed by any of these common data services. In this section, we describe the data services available in most Grids and how these service types can be augmented to support urgent storage and data management.

The most prevalent data services are data storage services. These services usually tightly couple to a computational resource, provide a staging area for transferring data between distributed resources, and provide a scratch work space for applications to store temporary data. The most common storage in Grids are file and archival systems accessible to one or more resource within a single VO. Transferring data between VOs requires the use of data transfer tools, such as GridFTP [10] and the Reliable File Transfer (RFT) service [11]. Recent developments and research in Grid storage systems have adapted cluster file systems to wide-area computing environments. Examples of wide-area file systems on the TeraGrid include the deployment of IBM's GPFS file system and Sun Microsystem's Lustre file system [12, 13]. Recent work with Grid data transfers has begun to address providing quality of service in Grids with variable throughput links and resource availability using tools such TeraPaths and autonomic computing [14, 15]. To date, none of these capabilities have addressed urgent computing data storage services. The capabilities of recent services, such as the GridFTP QoS provisioning [16], the Managed Object Placement Service (MOPS) [17], and the Data Placement Service (DPS)[18], provide the means for obtaining and sustaining QoS for a storage resource but cannot manage urgent data requirements without additional support. To adequately support urgent computations and workflows, an additional management layer is required to obtain the QoS best suited for the I/O footprint of an urgent computation, adjust the QoS of other concurrent workflows so that urgent workflows are not starved for resources, and negotiate end-to-end workflow scheduling for all urgent computing resources, such as storage, compute, or network resources through the utilization of available capabilities.

The immense amount of data produced by some emerging computations has frequently been cited as a hurdle to scaling applications to larger systems. Numerous tools have been developed and integrated into geoscience workflows to support data management tasks. Example data management tools include the Storage Resource Broker (SRB) [19], the Globus Replica Location Service (RLS)

[20], and GriPhyn [21]. These services allow users to store and retrieve data from a variety of Grid storage systems or to build catalogs of data products. Several geoscience-specific data services also provide data access and management capabilities for users. Examples include the myLEAD metadata catalog [22], Unidata's Thematic Realtime Environmental Distributed Data Services (THREDDS) [23], and data stream services such as Calder [8]. There is no explicit support in these data management services for urgent computing applications. There is little or no support for indicating QoS or SLA requirements to these services. Data management features for urgent computing workflows are necessary. These required capabilities include data replica and lifetime management of urgent and non-urgent data.

The last set of common data services are data processing services. These services provide various data integration or manipulation tasks. In this set of services, we include data discovery, assimilation, validation, and visualization services. Geoscience workflows, such as those in LEAD, SCOOP, and Grid-BGC [24], use several of these services. These services are generally computationally intensive and urgent computing capabilities built from these services would benefit from coupling urgent data provider and computation. The interplay of provisioning multiple resource required to support these services increases the management complexity of these data services. An additional resource management component is likely required to coordinate these provisioning tasks.

5 Urgent Data Management Capabilities and Framework

To support common data services described in urgent workflows, we proposed an additional data management layer. This Urgent Data Management Framework (UDMF) will leverage existing QoS, SLA, and resource provisioning infrastructures to allocate data resources for urgent computing workflows. These capabilities will integrate into a data and resource management layer responsible for allocating the appropriate SLAs and QoS for data resources and provide appropriate access levels and services. The UDMF will provide application and Web service interfaces so that workflows can invoke the urgent data management capabilities.

Several challenges exist that must be addressed by our framework so that it can provide urgent data services. First, our framework must interoperate with a variety of heterogeneous Grid resources managed by different VOs, such as the data services mentioned in the previous section. Since Grids are heterogeneous computing environments, UDMF must cope with differences in data service types, management policies, and characteristics. Another challenge that must be addressed is how to configure and manage these urgent data services. The urgent data services should require minimal human interaction for configuration, but should be intelligent enough to adapt to environmental changes. Another non-trivial integration issue for our framework involves how much effort is required for software developers and users to interact with our framework.

While configuring QoS for a data service will be required, the amount of effort to use our framework in existing tools should be minimized.

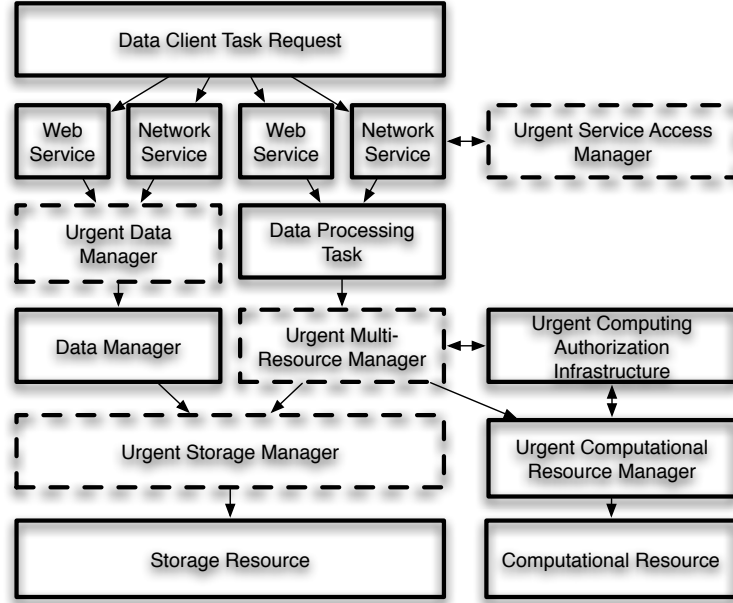


Fig. 1. The architecture and component interactions within UDMF.

Fig. 1 illustrates the proposed UDMF components and interactions. The components in boxes with solid lines represent existing infrastructure while the components in boxes with dashed lines are the proposed UDMF components. A variety of data and computational resources interact within our proposed architecture. Data resources, computational resources, service providers, service consumers, and urgent resource managers interact to provide urgent data services and data management capabilities to users. UDMF consists of several shims that fit between existing layers of data, service, and computing infrastructure. These shims will intercept normal data operations through hooks into the existing resources and adapt these resources for urgent computing demands or requirements. For example, we have begun development of the urgent computing access and provisioning tools for Grid services. This urgent computing infrastructure hooks into the authorization framework of Grid services to negotiate consumer access to the service based on the applicable urgent computing policy for the consumer. Hooks into other existing tools are available and include the Data Storage Interface (DSI) of GridFTP or intercepting I/O requests with overloaded operations similar to Trickle [25].

UDMF addresses the requirements and challenges for supporting urgent data services that we have defined throughout this paper. To fulfill these urgent data requests at the user-interface level, we will provide several Grid or Web services that allow users to define urgent computing requirements for a specific data service. These services will be accessible for use in existing workflow managers, such as Kepler [6]. A specific urgent data manager will be available for each data service type and these managers will coordinate through a general urgent resource manager. Resource administrators will define policies that these managers will follow. Autonomic computing tools will use the policies to adapt, tune, and manage these resources during urgent computations. The use of autonomic computing will relieve the need for human intervention during urgent computing events. The capabilities provided by each resource manager will vary based on the service. The urgent storage manager will negotiate access to storage and data resources, such as data transfer tools and scratch file systems. The urgent data manager will provide data replication, migration, and removal tasks to adapt existing data and resources for urgent computations. The multi-resource storage manager will negotiate access between coupled data and computational resources used by an urgent data task. This manager will leverage real-time computing and deadline scheduling techniques to schedule resource allocations and to identify potential scheduling conflicts. The UDMF architecture is intentionally small and designed to be as unobtrusive as possible. The managers will be easy to deploy and manage through modular integration or instrumentation of existing services. Workflows will only require minimal changes to properly allocate or access urgent data services by invoking the urgent data service APIs or Web service operations.

6 Conclusions

In this paper, we described our proposed approach to provisioning geoscience and Earth science data services for urgent computing applications and workflows. Our proposed approach is based on analyses of existing geoscience and Earth science data services and how these services are used in urgent computing workflows. Based on our studies, we devised an urgent data management framework that consists of additional layers and shims that can augment existing data management systems with urgent computing capabilities. This infrastructure is lightweight and is designed to not interfere with existing workflow execution, but to provide additional capabilities to urgent computing workflows. We have begun development of the autonomic management infrastructure, urgent data manager, and urgent service access components of our proposed framework. We expect to demonstrate the operation of these tools during the summer of 2008.

Acknowledgments. University of Colorado computer time was provided by equipment purchased under DOE SciDAC Grant #DE-FG02-04ER63870, NSF ARI Grant #CDA-9601817, NSF MRI Grant #CNS-0421498, NSF sponsorship of the National Center for Atmospheric Research, and a grant from the IBM

Shared University Research (SUR) program. We would like to thank the members of the SPRUCE project, including Pete Beckman, Suman Nadella, and Nick Trebon, for their guidance and support of this research.

References

1. Droegemeier, K., Chandrasekar, V., Clark, R., Gannon, D., Graves, S., Joesph, E., Ramamurthy, M., Wilhelmson, R., Brewster, K., Domenico, B., Leyton, T., Morris, V., Murray, D., Pale, B., Ramachandran, R., Reed, D., Rushing, J., Weber, D., Wilson, A., Xue, M., Yalda, S.: Linked Environments for atmospheric discovery (LEAD): A Cyberinfrastructure for Mesoscale Meteorology Research and Education. In: Proceedings of the 20th Conference on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology, Seattle, WA, American Meteorological Society (January 2004)
2. Cui, Y., Moore, R., Olsen, K., Chourasia, A., Maechling, P., Minster, B., Day, S., Hu, Y., Zhu, J., Majumdar, A., Jordan, T.: Enabling Very-Large Scale Earthquake Simulations on Parallel Machines. In: Proceedings of the International Conference on Computational Science (ICCS) 2007, Beijing, China, Springer (May 2007)
3. Bogden, P., Gale, T., Allen, G., MacLaren, J., Almes, G., Creager, G., Bintz, J., Wright, L., Graber, H., Williams, N., Graves, S., Conover, H., Galluppi, K., Luettich, R., Perrie, W., Toulany, B., Sheng, Y., Davis, J., Wang, H., Forrest, D.: Architecture of a Community Infrastructure for Predicting and Analyzing Coastal Inundation. *Marine Technology Society Journal* **41**(1) (2007) 53–71
4. Mandel, J., Beezley, J., Bennethum, L., Chakraborty, S., Coen, J., Douglas, C., Hatcher, J., Kim, M., Vodacek, A.: A Dynamic Data Driven Wildland Fire Model. In: Proceedings of the International Conference on Computational Science (ICCS) 2007, Beijing, China (May 2007) 1024–1049
5. Beckman, P., Beschatnikh, I., Nadella, S., Trebon, N.: Building an Infrastructure for Urgent Computing. *High Performance Computing and Grids in Action* (to appear) (2008)
6. Ludascher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger, E., Jones, M., Lee, E., Tao, J., Zhao, Y.: Scientific Workflow Management and the Kepler System. *IEEE Internet Computing* **18**(10) (December 2005) 1039–1065
7. Faerman, M., Moore, R., Cui, Y., Hu, Y., Zhu, J., Minster, B., Maechling, P.: Managing Large Scale Data for Earthquake Simulations. *Journal of Grid Computing* **5**(3) (September 2007) 295–302
8. Liu, Y., Vijayakumar, N., Plale, B.: Stream Processing in Data-driven Computational Science. In: 7th IEEE/ACM International Conference on Grid Computing (Grid'06). (September 2006)
9. Douglas, C., Beezley, J., Coen, J., Li, D., Li, W., Mandel, A., Mandel, J., Qin, G., Vodacek, A.: Demonstrating the Validity of Wildfire DDDAS. In: Proceedings of the 6th International Conference on Computational Science (ICCS) 2006, Reading, United Kingdom (May 2006)
10. Allcock, B., Bester, J., Bresnahan, J., Chervenak, A., Foster, I., Kesselman, C., Meder, S., Nefedova, V., Quesnal, D., Tuecke, S.: Data Management and Transfer in High Performance Computational Grid Environments. *Parallel Computing Journal* **28**(5) (May 2002) 749 – 771
11. Allcock, W., Foster, I., Madduri, R.: Reliable Data Transport: A Critical Service for the Grid. In: Global Grid Forum 11. (June 2004)

12. TeraGrid GPFS WAN, www.teragrid.org/userinfo/data/gpfswan.php (2008)
13. Simms, S., Pike, G., Balog, D.: Wide Area Filesystem Performance using Lustre on the TeraGrid. In: Proceedings of the TeraGrid 2007 Conference, Madison, WI (June 2007)
14. Katramatos, D., Yu, D., Gibbard, B., McKee, S.: The TeraPaths Testbed: Exploring End-to-End Network QoS. In: 2007 3rd International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities (TridentCom 2007). (May 2007)
15. Bhat, V., Parashar, M., Khandekar, M., Kandasamy, N., Klasky, S.: A Self-Managing Wide-Area Data Streaming Service using Model-based Online Control. In: Proceedings of the IEEE Conference on Grid Computing 2006 (Grid'06), Barcelona, Spain (September 2006)
16. Bresnahan, J., Link, M., Khanna, G., Imani, Z., Kettimuthu, R., Foster, I.: Globus GridFTP: What's New in 2007. In: Proceedings of the First International Conference on Networks for Grid Applications (GridNets 2007). (October 2007)
17. Baranovski, A., Bharathi, S., Bresnahan, J., Chervenak, A., Foster, I., Fraser, D., Freeman, T., Gunter, D., Jackson, K., Keahey, K., Kesselman, C., Konerding, D., Leroy, N., Link, M., Livny, M., Miller, N., Miller, R., Oleynik, G., Pearlman, L., Schopf, J., Schuler, R., Tierney, B.: Enabling Distributed Petascale Science. In: Proceedings of SciDAC 2007, Boston, MA (June 2007)
18. Chervenak, A., Schuler, R.: A Data Placement Service for Petascale Applications. In: Petascale Data Storage Workshop, Supercomputing 2007, Reno, NV (Nov. 2007)
19. Rajasekar, A., Wan, M., Moore, R., Schroeder, W., Kremenek, G., Jagatheesan, A., Cowart, C., Zhu, B., Chen, S., Olschanowsky, R.: Storage Resource Broker - Managing Distributed Data in a Grid. *Computer Society of India Journal* **33**(4) (October 2003) 42-54
20. Chervenak, A., Palavalli, N., Bharathi, S., Kesselman, C., Schwartzkopf, R.: Performance and Scalability of a Replica Location Service. In: International IEEE Symposium on High Performance Distributed Computing (HPDC-13), Honolulu, HI (June 2004)
21. Nefedova, V., Jacob, R., Foster, I., Liu, Z., Liu, Y., Deelman, E., Mehta, G., Su, M., Vahi, K.: Automating Climate Science: Large Ensemble Simulations on the Teragrid with the GriPhyN Virtual Data System. In: 2nd International IEEE Conference on e-Science and Grid Computing. (December 2006)
22. Plale, B., Gannon, D., Alameda, J., Wilhelmson, B., Hampton, S., Rossi, A., Droegeleier, K.: Active Management of Scientific Data. *IEEE Internet Computing* **9**(1) (January / February 2005) 27-34
23. Domenico, B., Caron, J., Davis, E., Kambic, R., Nativi, S.: Thematic Real-time Environmental Distributed Data Services (thredds): Incorporating Interactive Analysis Tools into NSDL. *Journal of Digital Information* **2**(4) (May 2002)
24. Cope, J., Hartsough, C., Thornton, P., Tufo, H.M., Wilhelmi, N., Woitaszek, M.: Grid-BGC: A Grid-Enabled Terrestrial Carbon Cycle Modeling System. In: Proceedings of the 11th European Conference on Parallel Processing (Euro-Par). (August 2005)
25. Eriksen, M.A.: Trickle: A Userland Bandwidth Shaper for Unix-like Systems. In: Proceedings of USENIX'05, Anaheim, CA (April 2005)