

GridFTP-Lite

*Feature enhancements to improve GridFTP's utility in
non-production testbed uses.*

*An unsolicited proposal
To
Thomas Ndousse
DOE Program Manager for Network Research
14 May 2004*

1 Executive Summary / Introduction

The Globus Toolkit's GridFTP code base has become the de facto standard for data movement within "Grid" projects, and is in use in the vast majority of such projects in the US and abroad. These projects appreciate GridFTP's integration with the public key infrastructure (PKI)-based Grid Security Infrastructure (GSI), as well as its implementation of the fast, efficient, and robust GridFTP data transport protocol.

Another important user community for GridFTP comprises networking researchers and others interested in using GridFTP for a limited amount of time, in non-production, low risk scenarios—for example, as part of research projects exploring new network protocols. For this user community, GridFTP's reliance on GSI can represent a time investment that is not justified, due to the associated need to establish, configure, and manage an appropriate PKI. Thus, this community has expressed a strong interest in seeing extensions to GridFTP that would allow for alternative security solutions.

Motivated by these requests, we propose here a two-year program of work that will update both the GridFTP code and its supporting packaging and infrastructure to (a) facilitate this type of usage and (b) respond to other demands for new GridFTP features that have emerged from the research community. The effort level required for this work is 0.5 FTEs in Year 1 and 0.7 FTEs in Year 2, for a total of 1.2 FTEs.

2 Security Issues

Security is the most commonly cited obstacle to the use of GridFTP in research purposes. The key problem is the need to obtain PKI credentials, a process that can result in several days of delay while identities are confirmed and certificates are issued. For production systems that must participate in multi-institutional deployments and will need to withstand hackers, port scans, and heavy usage 24x7, there is no real alternative to the use of PKI mechanisms. However, in other settings—e.g., for a testbed that will exist only for a temporary set of tests—it *may* be feasible to consider alternative approaches.

While we wish to address ease of use issues relating to GridFTP security infrastructure, the answer to these ease of use issues is not simply to disable security. Instead, we need to provide alternative techniques that may have different security characteristics but still provide some support for access control. The researcher can then select the most appropriate

technique based on ease of use vs. security tradeoffs. Nevertheless, it must be stated explicitly that *anyone choosing to run GridFTP with a reduced level of security is accepting the fact that there is a higher probability that their system will be compromised.*

We propose the following staged approach to changes to GridFTP and its surrounding infrastructure to better support short term, non-production use with reduced security.

Re-enable user/password and anonymous authentication in the wuftp based GridFTP server: The wuftp FTP (not GridFTP) server has support for user/password authentication, as well as anonymous access. This feature was disabled in the wuftp based GridFTP server because of its poor security. However, it is relatively easy to re-enable and would provide a tool that people could use in a relatively short period of time. Note that we will likely disable authentication against the system password file, and simply provide for a separate wuftp password file. This approach prevents system passwords (though not wuftp passwords) from being sent in the clear. Note that if anonymous is chosen, there is *no* security at all: *anyone* who happens to stumble on the server can initiate a transfer.

Add user/password and anonymous support to the new GridFTP server: This feature was not planned for the new server, but can be added. The infrastructure is in place for adding new commands, and the semantics are well defined by appropriate RFCs. The wuftp server will not be supported in the future so we will need to provide equivalent functionality in the new server. Besides the specific USER and PASS commands, there is the associated configuration, anonymous access, password file maintenance, etc, that must be addressed.

Add access control by IP address to the new GridFTP server: Configuration options can be added to the server that will permit policy to be applied based on incoming IP address. Such policies can range from simple admission control (you can or cannot connect) to allowing a completely different configuration to be selected based on incoming IP address. This approach allows access from testbed hosts with no authentication, but prevents the general public or a normal port scan from using the server. However, it is still susceptible to IP spoofing.

Add support to GridFTP for using ssh for authentication: ssh is a broadly deployed security mechanism. It does not provide delegation and other necessary functions for the Grid, but provides significantly more security than the methods listed above. One potential obstacle to this approach, which we have not yet checked out, is that we need to find open

source SSH protocol code that can be used under the Globus Toolkit License (code under the GPL can not). If we cannot, then we will have to opt for some form of scripting rather than actual code added to GridFTP—or we may have to drop this option all together.

Automate the use of Simple CA: The Globus Toolkit provides a component called the Simple CA, a utility that allows the user to establish their own CA, issue certificates, and so forth, and thus avoid the need to obtain PKI credentials from other sources. We propose to automate the security setup associated with using GSI. The install process would automatically create a testbed CA, issue certificates, and install them in the appropriate places. This process would be configured for non-privileged installation, i.e., the server would be run as a daemon under the user's account and certificate, rather than from inetd using a host certificate. This option has the advantage of retaining the full benefits of GSI security regarding access to the use of the service and it can still do data channel authentication. There would be, of course, the option for the use of an existing cert, should they already have one. Note that each individual would have to start a server within his account with his own certificate: either his existing personal certificate, or the testbed certificate created by Simple CA.

Summary of security improvements: With the changes listed above, we can provide a solution almost immediately with the existing wuftp, but with very little security and a high degree of risk. The other solutions will be added to the new GridFTP server and will provide a range of security options, allowing network researchers to balance the risks they take with the effort involved to ameliorate those risks.

3 Other GridFTP Extensions

Network researchers have requested several other new features in GridFTP, and we propose to incorporate those features into GridFTP as part of this work. We discuss these features in detail below.

Provide a GridFTP only Package: There are many people, network testers among them, whose only interest in the Globus Toolkit is GridFTP. It is confusing and inconvenient for them to have to download the entire data management package, or as is often done due to lack of knowledge, the entire toolkit. We will develop a package that contains only GridFTP and its dependencies and we will isolate the GridFTP specific installation and configuration documentation into one GridFTP only manual.

Provide a standard port for the data channel connection: Currently, GridFTP picks an ephemeral (random) port for the data channel. This makes

it nearly impossible for passive monitoring systems to detect GridFTP data channel traffic. We propose to modify GridFTP to use a default port on the listening server to allow passive monitoring systems to identify GridFTP data channel traffic. Initially, we will pick a port as the default and then will simply go to a random port if that port is in use. We will submit a request to IANA so that our chosen port can be registered as a “well known” port to prevent conflicts.

Allow memory-to-memory transfers: Currently, GridFTP can not do memory to memory transfers without modification. We propose to modify the server to allow access to special block devices (/dev/null, /dev/zero, /dev/random) and to add command line options to the client to run for a specified time, or a specified number of bytes. This approach will allow network researchers to take the disk subsystems out of the evaluation.

Extensive monitoring capabilities: When results are not as expected, it is useful to be able to “drill down” into the details of the execution of the tools and the transfer. We propose to add extensive monitoring capabilities to GridFTP. This work will include, at a minimum, adjustable monitoring levels, monitoring points around: all authorization tasks, external callouts (such as authz or CAS), channel setup, data transfer, transfer configuration (streams, TCP buffer size, etc), and for very detailed problems monitoring points around every callback, driver entry and exit in XIO and every disk and network read and write.

Firewall avoidance: It is a well-known limitation of the current GridFTP protocol that when using extended block mode (parallel streams), the sending side MUST perform the TCP connect. This can be a problem in many firewall situations. Resolution of this problem is a difficult one technically, but also socially and process wise. This requires standardization of a new protocol within the Global Grid Forum. Once standardized, we will need to implement and test the new protocol. The discussions for this new protocol are well underway within the GridFTP working group. A preliminary proposal will be reviewed at GGF11. However, it is likely that at least two, and probably three more GGFs will be required before consensus has been gained. We request a second years funding in order to implement this protocol, once it has been stabilized (though not necessarily accepted as a standard) within the GGF.

4 Cost and Timeline

The proposed work will require 0.5 FTE in the first year and 0.7 FTE in a second year. The first year will focus primarily on the security improvements proposed; the second year will implement the other improvements. Note that we are dependant on consensus being reached within GGF in order to

implement the firewall improvements. However, it seems unlikely that it will take longer than 18 months to accomplish this.

Note that the packages in the task list are not official Globus Toolkit releases. This proposal contains features that are not suitable for use in a production infrastructure. However, features that *are* suitable for production usage will be fully integrated with the toolkit and come out with the next release after its completion.

<i>Date</i>	<i>Task (Milestones in Bold)</i>
Q1, Y1	Re-enable User/Pass and Anonymous for wuftp Deliver patch and instructions for wuftp Add User/Pass and Anonymous to new server Protocol enhancement discussions at GGF-12
Q2, Y1	Add access control based on IP address to new server Deliver GridFTP only package that includes enhancements Protocol enhancement discussions at GGF-13
Q3, Y1	Provide an ssh based solution
Q4, Y1	Provide an automated GSI solution Deliver updated GridFTP package with full range of security solutions Protocol enhancement discussions at GGF-14
Q1, Y2	Make the data channel use a default port Add memory to memory capability Deliver updated GridFTP package Protocol enhancement discussions at GGF-15
Q2, Y2	Complete Instrumentation work Develop appropriate associated logging tools and services (netlogger based) Deliver updated GridFTP package Protocol enhancement discussions at GGF-16
Q3, Y2	New protocol implemented
Q4, Y2	New protocol tested and tuned Deliver updated GridFTP package Protocol enhancement discussions at GGF-17

Cover Page

Title of Proposed Project:

GridFTP - Lite

Office of Science Announcement Title/#:

Unsolicited

DOE/OSC Program Staff Contact:

Dr. Thomas Ndousse

Name of Laboratory:

Argonne National Laboratory

Principal Investigator(s):

William Allcock
Argonne National Laboratory
Mathematics and Computer Science Division
9700 So. Cass Avenue - Bld.221
Argonne, IL 60439
630-252-7573
allcock@mcs.anl.gov

Official signing for Laboratory:

Rick L. Stevens, Division Director
Argonne National Laboratory
Mathematics and Computer Science Division
9700 So. Cass Avenue - Bld.221
Argonne, IL 60439
630-252-3378
stevens@mcs.anl.gov

Requested funding for ANL for each year; total request

Year 1	\$	150,000
Year 2	\$	200,000
Year 3	\$	-
Year 4	\$	-
Year 5	\$	-
Total:	<u>\$</u>	<u>350,000</u>

Duration of Entire Project Period:

10/01/2004 to 09/30/2006

Use of human subjects in proposed project:

No

Use of vertebrate animals in proposed project:

No

Signature of PI, Date of Signature:

Signature on File- William Allcock

Signature of Official, Date of Signature:

Signature on File - Rick L. Stevens

Budget Page

(See reverse for Instructions)

ORGANIZATION The University of Chicago, Operator of Argonne National Laboratory				Budget Page No: <u>pg 1 of 3</u> Year 1 of 2		
PRINCIPAL INVESTIGATOR/PROJECT DIRECTOR William Allcock				Requested Duration: <u>12</u> (Months)		
A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title; A.6. show number in brackets)		DOE Funded Person-mos.		Funds Requested	Funds Granted	
		CAL	ACAD	SUMR	by Applicant	by DOE
1.	William Allcock, PI	1.20			\$14,351	
2.	Raj Kettimuthu	4.80			\$57,405	
3.						
4.						
5.						
6.	() OTHERS (LIST INDIVIDUALLY ON BUDGET EXPLANATION PAGE)					
7.	(2) TOTAL SENIOR PERSONNEL (1-6)	6.00			\$71,756	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1.	() POST DOCTORAL ASSOCIATES					
2.	() OTHER PROFESSIONAL (TECHNICIAN, PROGRAMMER, ETC.)					
3.	() GRADUATE STUDENTS					
4.	() UNDERGRADUATE STUDENTS					
5.	() SECRETARIAL - CLERICAL					
6.	() OTHER					
TOTAL SALARIES AND WAGES (A+B)					\$71,756	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)					\$17,036	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A+B+C)					\$88,792	
D. PERMANENT EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM.)						
TOTAL PERMANENT EQUIPMENT						
E. TRAVEL		1. DOMESTIC (INCL. CANADA AND U.S. POSSESSIONS)			\$2,400	
		2. FOREIGN			\$10,000	
TOTAL TRAVEL					\$12,400	
F. TRAINEE/PARTICIPANT COSTS						
1. STIPENDS (Itemize levels, types + totals on budget justification page)						
2. TUITION & FEES						
3. TRAINEE TRAVEL						
4. OTHER (fully explain on justification page)						
TOTAL PARTICIPANTS () TOTAL COST						
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES					\$12,503	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						
3. CONSULTANT SERVICES						
4. COMPUTER (ADPE) SERVICES						
5. SUBCONTRACTS						
6. OTHER						
TOTAL OTHER DIRECT COSTS					\$12,503	
H. TOTAL DIRECT COSTS (A THROUGH G)					\$113,696	
I. INDIRECT COSTS (SPECIFY RATE AND BASE)						
Section H. Direct cost X Aggregate rate of: 31.931%						
TOTAL INDIRECT COSTS					\$36,304	
J. TOTAL DIRECT AND INDIRECT COSTS (H+I)					\$150,000	
K. AMOUNT OF ANY REQUIRED COST SHARING FROM NON-FEDERAL SOURCES						
L. TOTAL COST OF PROJECT (J+K)					\$150,000	

Budget Page

(See reverse for Instructions)

ORGANIZATION The University of Chicago, Operator of Argonne National Laboratory				Budget Page No: <u>pg 2 of 3</u> Yr. 2 of 2		
PRINCIPAL INVESTIGATOR/PROJECT DIRECTOR William Allcock				Requested Duration: <u>12</u> (Months)		
A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title; A.6. show number in brackets)		DOE Funded Person-mos.		Funds Requested	Funds Granted	
		CAL	ACAD	SUMR	by Applicant	by DOE
1.	William Allcock, PI	1.2			\$15,004	
2.	Raj Kettimuthu	4.8			\$60,014	
3.	John Bresnahan	1.2			\$15,004	
4.	Joe Link	1.20			\$15,004	
5.						
6.	() OTHERS (LIST INDIVIDUALLY ON BUDGET EXPLANATION PAGE)					
7.	(4) TOTAL SENIOR PERSONNEL (1-6)	8.40			\$105,025	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1.	() POST DOCTORAL ASSOCIATES					
2.	() OTHER PROFESSIONAL (TECHNICIAN, PROGRAMMER, ETC.)					
3.	() GRADUATE STUDENTS					
4.	() UNDERGRADUATE STUDENTS					
5.	() SECRETARIAL - CLERICAL					
6.	() OTHER					
TOTAL SALARIES AND WAGES (A+B)					\$105,025	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)					\$24,972	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A+B+C)					\$129,997	
D. PERMANENT EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM.)						
TOTAL PERMANENT EQUIPMENT						
E. TRAVEL		1. DOMESTIC (INCL. CANADA AND U.S. POSSESSIONS)			\$2,500	
		2. FOREIGN			\$10,400	
TOTAL TRAVEL					\$12,900	
F. TRAINEE/PARTICIPANT COSTS						
1. STIPENDS (Itemize levels, types + totals on budget justification page)						
2. TUITION & FEES						
3. TRAINEE TRAVEL						
4. OTHER (fully explain on justification page)						
TOTAL PARTICIPANTS () TOTAL COST						
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES					\$8,698	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						
3. CONSULTANT SERVICES						
4. COMPUTER (ADPE) SERVICES						
5. SUBCONTRACTS						
6. OTHER						
TOTAL OTHER DIRECT COSTS					\$8,698	
H. TOTAL DIRECT COSTS (A THROUGH G)					\$151,595	
I. INDIRECT COSTS (SPECIFY RATE AND BASE)						
Section H. Direct cost X Aggregate rate of: 31.931%						
TOTAL INDIRECT COSTS					\$48,405	
J. TOTAL DIRECT AND INDIRECT COSTS (H+I)					\$200,000	
K. AMOUNT OF ANY REQUIRED COST SHARING FROM NON-FEDERAL SOURCES						
L. TOTAL COST OF PROJECT (J+K)					\$200,000	

Budget Page

(See reverse for Instructions)

ORGANIZATION The University of Chicago, Operator of Argonne National Laboratory				Budget Page No: <u>pg. 3 of 3</u>		
PRINCIPAL INVESTIGATOR/PROJECT DIRECTOR William Allcock				Total Project <u>2</u> years		
Requested Duration: <u>24</u> (Months)						
A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title; A.6. show number in brackets)		DOE Funded Person-mos.		Funds Requested	Funds Granted	
		CAL	ACAD	SUMR	by Applicant	by DOE
1.	William Allcock, PI	2.40			\$29,355	
2.	Raj Kettimuthu	9.60			\$117,420	
3.	John Bresnahan	1.20			\$15,004	
4.	Joe Link	1.20			\$15,004	
5.						
6.	() OTHERS (LIST INDIVIDUALLY ON BUDGET EXPLANATION PAGE)					
7.	(4) TOTAL SENIOR PERSONNEL (1-6)	14.40			\$176,782	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)						
1.	() POST DOCTORAL ASSOCIATES					
2.	() OTHER PROFESSIONAL (TECHNICIAN, PROGRAMMER, ETC.)					
3.	() GRADUATE STUDENTS					
4.	() UNDERGRADUATE STUDENTS					
5.	() SECRETARIAL - CLERICAL					
6.	() OTHER					
TOTAL SALARIES AND WAGES (A+B)					\$176,782	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)					\$42,007	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A+B+C)					\$218,789	
D. PERMANENT EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM.)						
TOTAL PERMANENT EQUIPMENT						
E. TRAVEL		1. DOMESTIC (INCL. CANADA AND U.S. POSSESSIONS)			\$4,900	
		2. FOREIGN			\$20,400	
TOTAL TRAVEL					\$25,300	
F. TRAINEE/PARTICIPANT COSTS						
1. STIPENDS (Itemize levels, types + totals on budget justification page)						
2. TUITION & FEES						
3. TRAINEE TRAVEL						
4. OTHER (fully explain on justification page)						
TOTAL PARTICIPANTS () TOTAL COST						
G. OTHER DIRECT COSTS						
1. MATERIALS AND SUPPLIES					\$21,201	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION						
3. CONSULTANT SERVICES						
4. COMPUTER (ADPE) SERVICES						
5. SUBCONTRACTS						
6. OTHER						
TOTAL OTHER DIRECT COSTS					\$21,201	
H. TOTAL DIRECT COSTS (A THROUGH G)					\$265,291	
I. INDIRECT COSTS (SPECIFY RATE AND BASE)						
TOTAL INDIRECT COSTS					\$84,709	
J. TOTAL DIRECT AND INDIRECT COSTS (H+I)					\$350,000	
K. AMOUNT OF ANY REQUIRED COST SHARING FROM NON-FEDERAL SOURCES						
L. TOTAL COST OF PROJECT (J+K)					\$350,000	

Budget Justification

Funding Mechanism:

Argonne National Laboratory is to receive funds through the financial plan for only the dollars identified on the ANL budget pages.

Sections A, B, and C.- Senior Personnel/Other Personnel/Fringe Rates:

Key Staff, Total Person Months and Primary Role/Responsibility

Senior Personnel:

William Allcock, PI	2.40
Raj Kettimuthu	9.60
John Bresnahan	1.20
Joe Link	1.20
Post Doctoral Associates	0.00

Other Professionals: 0.00

Undergraduate Students \$0

See attached explanation of costing procedures that relate to effort rates and fringe benefits.

Sections D. Permanent Equipment

\$0

Sections E. Travel

Domestic: \$1.2K per trip/escalate~ 4.0% per yr. 3-year estimate.	\$4,900
Projecting 1-trips per staff member per year to present project status to Sponsor-Washington D.C.	
Projecting 1-trips per staff member per year to present results/status to collaborative institutions-TBD.	
Projecting 1-trips per staff member per year to present results to Supercomputing Conference.-TBD	
Foreign: \$2.5k per trip/escalate 4.0% per yr. 2-year estimate.	\$20,400

Sections G. Other Direct Costs

1. Materials and Supplies: 3-year estimate.	\$21,201
Software, low-end computers (<\$5k), computer supplies/ peripherals, and misc supplies.	
2. Publications Costs/ Documentation 3-year estimate.	\$0
Books, subscriptions, publishing related costs to research.	
4.. Computer Services/Support 3-year estimate.	\$0
General Infrastructure Software/Hardware maintenance	
6. Other 3-year estimate.	\$0
Relocation Expenses related to hiring of new hires	

Sections I. Indirect Costs

See attached explanation of costing procedures that relate to indirect rates.

Costing Procedures

Argonne National Laboratory is a government-owned facility, operated by the University of Chicago. As a contractor for the Department of Energy (DOE), Argonne National Laboratory must comply with DOE general policies and procedures on budgeting and accounting. The costing procedures are based on the assumption that all costs incurred will be recovered. When there is a demonstrated direct programmatic benefit to the Department of Energy, work may be performed for other federal agencies, and the activity is charged on the same basis as work supported by the Department of Energy.

The costing procedures use standard rates, which are utilized throughout the Laboratory on a consistent basis and uniformly applied to all work supported by the Department of Energy and other federal agencies. Standard rates established at the beginning of the fiscal year for each Division, are monitored, and revised as necessary. All labor costs are distributed at standard rates, which are developed by the Laboratory Budget Office for each major payroll classification within the Division. The standard rates are an average of the base wages, fringe benefits, paid absence, and divisional overhead accounts (Division Management and Direct Allocations). The Division Management account includes only those costs associated with the operation of the Division and directly related administrative activities such as management, personnel administration, procurement and budget administration, and cost for materials, which cannot be directly associated with any specific program activity. The Direct Allocation account includes costs for custodial services, building maintenance, utilities and related services. The fringe benefits include payroll-related items such as annuities, social security, and hospital and medical payments. The following rates have been estimated for fiscal year 2004 through FY2009

Mathematics and Computer Science Division:

	FY2004	FY2005	FY2006	FY2007	FY2008	FY2009
Exempt, Regular (Sr. Personnel, Other Professionals)						
Salary (FTE Month)	\$7,962	\$8,351	\$8,744	\$9,154	\$9,584	\$10,035
Fringe (34.0 %)	\$2,707	\$2,839	\$2,973	\$3,112	\$3,259	\$3,412
Paid Absence/Div.Mgt/Direct Allocations.	\$3,464	\$3,608	\$3,759	\$3,917	\$4,080	\$4,251
Total FTE Monthly Rate	\$14,132	\$14,799	\$15,476	\$16,183	\$16,924	\$17,698
Exempt, Temporary (Postdoctoral and Visiting Scientists)						
Salary (FTE Month)	\$5,292	\$5,530	\$5,790	\$6,063	\$6,348	\$6,646
Fringe (11.0 %)	\$582	\$608	\$637	\$667	\$698	\$731
Paid Absence	\$1,174	\$1,226	\$1,280	\$1,336	\$1,395	\$1,456
Total FTE Monthly Rate	\$7,049	\$7,364	\$7,707	\$8,066	\$8,441	\$8,833

Standard rates are also developed for Laboratory General and Administrative (G&A) expense. The procedures for distribute Laboratory G&A and program expense is applied on the basis of the total cost of the work performed. The following indirect rates have been estimated for FY2004-2009:

PBCS Program Expenses @2.3%
Laboratory G&A:
Common Support @ 25.70%
Equipment/Subcontracts@ 6.8%
G&A Burden @ 3.83%