

**NATIONAL SCIENCE FOUNDATION**  
4201 WILSON BOULEVARD  
ARLINGTON, VIRGINIA 22230

July 18, 2013

TO: Office of Information and Regulatory Affairs  
Office of Management and Budget  
Docket Library

FROM: Suzanne Plimpton  
Reports Clearance Officer  
National Science Foundation

SUBJECT: Amendment to OMB clearance 3145-0101, "Survey of Science and Engineering Research Facilities"

Attached for your records is the electronic submission of the amendment to the OMB clearance (expiration date: 10/31/14) for the Survey of Science and Engineering Research Facilities (3145-0101).

This amendment concerns the FY 2013 Survey of Science and Engineering Research Facilities questionnaire, which has two parts. The requested changes will reduce Part 2 of the questionnaire from 22 to 15 questions. This reduction in the length of Part 2 will be accomplished by eliminating questions and combining questions in situations where less data is required. Two new questions on cloud computing resources are introduced. Finally, questions on the time needed to complete the survey were added to both the Part 1 and Part 2 questionnaires.

Part 2 of the Survey of Science and Engineering Research Facilities focuses on Computing and Networking Capacity (cyberinfrastructure). NSF solicited feedback on Part 2 of the survey from field experts, NSF's Division of Advanced Cyberinfrastructure, data providers, and potential data users through interviews and a survey review panel. This feedback indicated that Part 2 of the survey does not adequately measure the rapidly changing field of academic research cyberinfrastructure. Several data points are no longer relevant, while other emerging resources need to be tracked. NSF has embarked on a plan to pare Part 2 of the Facilities Survey to core questions for FY 2013, add coverage of cloud computing questions for FY 2013 and investigate other metrics for the FY 2015 survey cycle.

NSF selected nine institutional survey participants to review a draft of Part 2 of the questionnaire, which was reduced from 22 to 15 questions. Phone debriefing interviews were then conducted with the participants to investigate their understanding of the questions and to

allow for further feedback. These institutions were selected based on criteria covering a range of networking, high performance computing and data storage resources. NSF believes the institutions were able to provide sufficient feedback because only two new questions were added. The other changes entailed eliminating questions or combining questions. Thus, the remaining thirteen questions have historical continuity with prior versions of the survey.

The nine survey respondents who reviewed the revised Part 2 questionnaire provided favorable comments on the elimination of questions, the focus on core cyberinfrastructure data points, and the addition and clarity of two new cloud computing questions. They noted that some of the questions from the past survey were no longer relevant due to changes in the field. All respondents indicated that the data were internally available for them to report and that the survey completion time was less than in FY 2011. NSF believes the survey burden will be reduced for participants by 30 minutes.

A list of the proposed changes is on page 1 of the attached questionnaire. Additional details are provided below.

Questions 1, 9, 10, and 11 have been modified to allow the precise reporting of bandwidth and data storage totals. In FY 2011, respondents selected a checkbox in a series of thresholds. This new method eliminates the need to periodically update those thresholds and provides greater accuracy. Discussions with respondents indicate that precise amounts are readily available and have been used when selecting checkbox ranges in the past.

Question 3 on connections to Internet2 and National LambdaRail combines questions 2 and 3 from the FY 2011 survey, while eliminating the requirement for providing specific bandwidth received from each source. The bandwidth provided by these sources has been (and continues to be) accounted for in question 1. The response options for question 3 were included in question 6 in the FYs 2005-2009 surveys.

The instructions for question 4 were modified to clarify that infeasible rights of use should be reported. Some state institutions do not technically own dark fiber though they might have infeasible rights of use that give them full control over this resource. All respondents understood the new guidelines.

The instructions for question 5 on architectures for centrally administered high-performance computing (HPC) were modified to drop the restriction that only systems that are “generally available to the campus community” should be considered. This phrase was added to the FY 2009 survey to ensure that computing systems reserved for single researchers were omitted. Subsequent respondent feedback indicated that this phrase was not necessary and that it may cause unintended confusion. The category for “Special purpose architectures” was dropped from question 5. This category is no longer relevant. Participants can provide responses in “Other architectures.”

Question 7 on centrally administered HPC systems asks respondents to provide the total amount of computing resources (for systems of 10 teraflops or faster), indicate the fastest individual system and identify its architecture. Computing speeds were collected for each architecture in questions 13-17 on the FY 2011 questionnaire. Restructuring this section to a single question

eliminates the need to provide more detailed responses. The respondents indicated that a breakdown by individual computing architectures was not necessary. Total computing capacity is the relevant point, which is now accounted for more succinctly.

Question 12 is new for the FY 2013 survey cycle. This item, which gathers information on archival storage from external cloud services, is a follow-up inquiry to question 11 on archival storage for centrally administered HPC. Cloud services are an increasing resource in academic research and development. NSF has been encouraged by field experts and potential data users to collect data on this area. All respondents were able to provide accurate responses to this question.

Question 13, which asks for information on research computing from external cloud services, is a new question for the FY 2013 survey cycle. It addresses the need to account for HPC access through the cloud. Respondents are asked whether their institution accessed these resources from cloud services in FY 2013 and whether they plan to access these resources in FY 2014. Question 13(c) clarifies whether researchers access these resources through a central IT department. Respondents found the wording and intent of this question to be clear. These data are available at the institutions.

Question 13 is being added at the request of data users who are interested in identifying the percentage of their peer institutions accessing cloud computing compared to those providing HPC resources internally. It will also help direct efforts to develop additional questions on this topic for the FY 2015 cycle.

Given the changes to the questionnaire, we would like to develop a more accurate burden estimate. Question 14 on survey completion time was added to account for the Part 2 burden. Since Part 1 (research space) and Part 2 (computing and networking capacity) are typically completed by different offices, we have added a survey completion question (question 19) to Part 1 as well. This will allow NSF to provide a more accurate burden estimate for the next OMB Facilities Survey clearance.

As noted above, the nine survey respondents indicated that the questions were clear and that the data were readily available. NSF will confirm this in the production setting through debriefing interviews. These data will be published on the national level. Data for Part 2 of the Facilities Survey are not imputed.

Based on estimates from past surveys and discussions with respondents regarding the changes, we expect a minor decrease to the burden. We estimate the survey burden will be reduced by approximately 30 minutes per respondent for the FY 2013 survey. Because this decrease in average burden is estimated to be less than one hour, we have retained the current burden estimate of 41 hours. Total burden hours were estimated to be 19,270 for FY 2011 based on 470 responses as confirmed by the OMB notice dated 10/19/2011. The FY 2011 response rate of 98% was greater than anticipated, as was the higher number of academic institutions surpassing the \$1 million threshold. Therefore, total burden hours for the FY 2013 Facilities Survey are estimated to be 22,140 based on 540 participants.

If you have any questions, please call Suzanne Plimpton at 703-292-7556.

Form 83-C

Attachment 1: FY 2013 Survey of Science and Engineering Research Facilities