SC-16 Birds of a Feather Session Submission Form

Submission Instructions

Use this form to submit a proposal to host a Birds of a Feather (BOF) session at SC16. Follow the directions on the form to make your submission, and be sure to click the "Submit button" at the bottom of the form. Submissions are being accepted from 12am (UTC-12) Feb 16, 2016 till 11:59pm (UTC-12) Jul 31, 2016. We plan to announce decisions on 11:59pm (UTC-12) Sep 1, 2016.

We remind you that BOFs are interactive, community building, non-commercial venues. Before submitting, consider whether your proposed session fits this description. Consider also whether there is another group of people who are likely to submit a similar BOF and whether you should merge your proposal with theirs. Upon review of your BOF proposal, the BOFs committee may suggest combining with other proposals.

Each BOF must produce a tangible outcome, such as a report summarizing decisions made by the attendees or results of a survey of attendees that will guide how the community should proceed). These tangible outcomes may be archived with the SC conference proceedings, or be made available via the SC public web site.

A proposal must be submitted for all BOFs and Meetings to be held at SC16. Even if your BOF has been held at SC since the first time the conference was held, you cannot assume that it will be held at SC16 - you must submit a proposal that will be reviewed and evaluated against the other submissions.

Selected BOF proposals will rank high in terms for their potential for community building, relevance to the expected audience, the level of audience participation, and the timeliness of the topic. Other criteria may be considered as well. BoFs are expected to have a limited amount of vendor-specific content.

Title

Provide the title of your proposed BOF. If your BOF is accepted, this is the title for the session that will be used in the program, so please use standard title capitalization (i.e., capitalizing the first and subsequent important words, not ALL CAPS and not all lower case).

An Open Failure Data Repository for Dependability Research and Practice: Data, Tools, and Analysis

{I believe we should use the same title as last year to show some continuity}

Title

Primary Session Leader Information

Provide information about the Primary Session Leader exactly as you would like it to appear in the program if your BOF submission is accepted. Please check your spelling, use standard capitalization (e.g., "Chris Smith" is acceptable but not "CHRIS SMITH" or "chris smith"), and confirm the email address.

People entered here will be given access to this submission like the original submitter.

Primary Session Leader: Saurabh Bagchi, Purdue University
Secondary Session Leaders: (Alphabetically)
1.   Nathan DeBardeleben, Los Alamos National Laboratory
2.   Ravishankar Iyer, University of Illinois at Urbana-Champaign
3.   Zbigniew Kalbarczyk, University of Illinois at Urbana-Champaign
4.   Carol Song, Purdue University

Insert your contact info from your "My Account" page?

☐

Title	First/Given Names	MI	Last/Family Name

Email   Company/Institution ⓘ

Primary Telephone Alternate Telephone Fax

You may enter a biography of this person, to appear in the program.

Biography (Maximum 150 words)   0 words

Upload a photograph of this person (JPG/GIF/TIFF/PNG) up to 1 MB.

Is this person on the Birds of a Feather reviewing committee? ☐ Yes

Secondary Session Leader Information

If desired, provide information for Secondary Session Leaders. Enter the information exactly as you would like it to appear in the program if your submission is accepted.

Please check spelling, use standard capitalization (e.g., "Chris Smith" but not "CHRIS SMITH" or "chris smith"), and confirm the email address for each.

People entered here will be given access to this submission like the original submitter.

Secondary Session Leader 1: ✗

Insert your contact info from your "My Account" page?

☐

Title	First/Given Names	MI	Last/Family Name
▼			

Email	Company/Institution ⓘ

Primary Telephone	Alternate Telephone	Fax

You may enter a biography of this person, to appear in the program.

Biography (Maximum 150 words)   0 words

Upload a photograph of this person (JPG/GIF/TIFF/PNG) up to 1 MB.

Is this person on the Birds of a Feather reviewing committee? ☐ Yes

BOF Topic Area

New this year: Please select the topic area that best fits your BOF submission. If none of the topic areas fit, please select "Other." New this year: Please select Meetings for society or SC meetings, and TOP 500.

BOF Topic Area
- ☐ Algorithms
- ☐ Applications
- ☐ Architecture and Networks
- ☐ Clouds & Distributed Computing
- ☉ Data Analytics, Visualization & Storage
- ☐ Education
- ☐ Meetings

- ☐ Performance Measurement, Modeling, and Tools
- ☐ Programming Systems
- ☐ State of the Practice
- ☐ System Software
- ☐ Other

Abstract

Provide a short description of your proposed BOF, clearly stating its goal(s), topic, format, and intended audience. If your BOF is accepted, this information will be used in the program. Indicate paragraph breaks by a blank line in the text field; all other text formatting will be lost.

This BoF builds on our last year's meeting and in it we plan to share our lessons learned while creating an open failure and system usage data repository to enable resiliency research for large-scale computing systems. We will share how we have taken the feedback from last year's BoF into account in designing our repository. We plan to extend the dialog with the HPC community by providing examples of actual analysis based on failure data logs collected from Blue Waters (peta-scale system at UIUC). We will stimulate discussion on issues in creating such a data repository and expected benefit in the context of future extreme-scale systems.

Abstract (Maximum 100 words)   0 words

Long Description

Provide a description of your proposed BOF, clearly stating its goal(s), topic, relevance to the expected HPC audience, whether the BOF has been held before and past attendance count(s), and the outcome (e.g., written report describing results of survey of BOF attendees for a specific question - be sure to indicate which survey questions will be asked). Indicate paragraph breaks by a blank line in the text field; all other text formatting will be lost. [Note: BOF Meetings - please enter "N/A"]

Building on the very positive experience with our BoF last year (https://engineering.purdue.edu/dcsl/news/2015/sc_bof.html) we wish to further draw on a community within SC that works on resiliency issues of large computing infrastructures. While we still plan to talk about the issues and experience in creating failure data repository we want to go beyond data collection and discuss how one would use such repository to assist development of more robust and scalable HPC systems.

We believe that analysis of failure data can provide an in-depth understanding of failure modes, fault/error propagation, detection/mitigation mechanisms, and failure

impact on applications. This understanding can form the basis for future instrumentation and tools to automate fault characterization.

Through an NSF and a DOE project, we are providing an annotated modern comprehensive dataset for extreme-scale systems to facilitate advances in resilience research. Through this BoF, we will seek to collect lessons to increase the value of such a repository:

- to characterize faults, in extreme-scale systems, based on root or most probable cause, likelihood of detection, frequency of occurrence, timescales for resultant system impact, and efficiency of error recovery.

- to determine new instrumentation, to support fault detection and recovery as well as means of improving quality of data collected by the system.

- to recommend at what levels of the system hierarchy (e.g., hardware infrastructure, operating system, runtime systems, and/or application software) one should place cost-effective error detection and/or recovery mechanisms.

The points made above will be used to involve (and perhaps provoke) the participants of the BoF session to share their experience and ideas as to how we, as a community, should collaborate to test those ideas in reality.

In addition, this time we plan to bring for discussion the needs for tools to automate failure data preprocessing and analysis. As an example we will discuss LogDiver, an open source tool developed at UIUC to facilitate analysis and measurement of system- and application-level resiliency in extreme-scale environments. The LogDiver approach is to create a unique dataset encapsulating events that are central in conducting resiliency and performability measurements. The tool allows precise identification of the reasons behind application termination, relates system errors and failures (e.g., network fabric errors, GPU errors, and file-system failures) to application failures, and provides a unified representation of the workload/error/failure logs, permitting workload-failure analysis and computation of a range of quantitative performance and dependability metrics.

While the tool was developed heaving in mind Cray systems (e.g., Blue Waters the supercomputer managed by NCSA), the LogDiver design and its capabilities are a good example of what one would need to improve efficiency of data analysis and to produce actionable information. We will involve participants to share their views, ideas, needs, and experience.

We will ask the participants specific questions:

1. What examples can you bring that demonstrate use of failure data analysis to guide design of future systems?

2. What insights would you expect to get from analyzing field failure data?

3. What kinds of data analysis tools would you find beneficial for analyzing failure and system usage data?

Long Description (Maximum 500 words)   0 words

Session format

How much of the session will be used for interaction between audience and session leaders/presenters?

- ○ 100%
- ○ 75%
- ◉ 50%
- ○ 25%
- ○ 10%

What is the primary format for content that does not directly involve audience discussion?

- ◉ Sequence of presentations
- ○ Panel of experts
- ○
- Other: [   ]

Does the BOF topic deal with commercial technology?

- ◉ Not commercial
- ○ Vendor-neutral
- ○ Vendor-specific
- ○
- Other: [   ]

Description of the session format

Provide a short description on how you plan to organize the BOF session, to justify, clarify and expand on your selections in the "Session format" part of the submission. In your response, keep in mind the guidelines from the Call for BOFs about what is appropriate for an SC BOF. [Note: BOF Meetings, please enter "N/A"]

1. We formulate research questions that could be answered using data on failures of large-scale computing systems/infrastructures [30%]

- Identify who conducts such research and why

- Discuss extra information and technical support needed in addition to raw failure and system usage data?

2. We present data we have been collecting at Purdue and UIUC [40%]

- Demonstrate an example analysis of error propagation in the network fabric of Blue Waters

- Discuss data-driven experimental methods (e.g., fault injection) that take advantage of failure data to drive methods for characterization of resilience

Description of the session format (Maximum 150 words)   0 words

BOFs at recent SCs

Provide information whether your proposed BOF has been held at any or all of the four most recent SC conferences. If it *was* held at **any** of those conferences, answer all questions. If it *was not* held at any of the past four SC conferences, answer the first question and skip the remaining questions in this section.

Has your BOF been held at recent SC conferences? ☐ SC12
☐ SC13
☐ SC14
☑ SC15

If so, approximately how many attendees did your BOF attract the most recent year it was held?
◉ <10
◉ 10-24
◉ 25-49
◉ 50-74
◉ 75-149
◉ >149

Scheduling Information

Please provide scheduling preferences for your BOF. If desired, provide the preferred date and time for your BOF. Note that we may not be able to satisfy every accepted BOF's preference, and we may not schedule BOFs in each of the indicated time slots. Please also provide one to three keywords or phrases that describe the topic of your BOF.

Preferred date and time: [▼]
Amount of time requested: [▼] 1h

Expected Attendance: [ ] 50

Keyword/Phrase 1: [ ] failure data repository

Keyword/Phrase 2: [ ] error and failure analysis

Keyword/Phrase 3: [ ] dependability

Conference Presentations

SC is gathering all presentations and other 'live' material (PowerPoint presentations, PDF slides, etc.) to archive at the ACM/IEEE archive. This material, collected at the conference, will be a permanent record of the conference and available for download after the conference. You will be requested to submit it around the time of the conference.

Can SC archive and distribute your conference presentation?

◉ Yes

◯ No

Acknowledgement

BOF leaders are required to produce some artifact from the BOF, such as a report describing audience discussion or the results of an audience survey, as described in the SC16 call for BOFs.

If your BOF is accepted, you will be asked to confirm your participation.

Also, changes to the information you submit here will not be allowed after July 31, even if your submission is accepted. This means you will not be able to change the title, abstract, session leaders or their affiliations, or any other information. If your submission is accepted, the information you submit here will go directly into the conference's final program. (You can, however, change this information at any time before the submission deadline of July 31.)

By checking this box, you acknowledge and agree to abide by the requirements listed above. ☐

Submissions close in 6 days, 17 hours, 7 minutes

Important Notes

When you submit the form, wait to see if any errors are reported. If errors are not fixed, it will not be counted as submitted.

A confirmation email will be sent to you and each primary session leader and secondary session leader entered above when this submission is received; keep this email as a receipt. If you do not get this email, it is likely that your submission has not been received. Verify that you can see it on your "My Submissions" page, and if not, resubmit. It is your responsibility to confirm that the submission has been received.

You and each primary session leader and secondary session leader entered above can update this submission until submissions close.