Program Chair's Message

It has been a privilege to serve as the program chair for ASPLOS 2014. I am inspired and humbled by the breadth and quality of the work submitted and the commitment of the organizing team to an excellent conference.

The broad goal of a conference is to move the field forward. I was fortunate that recent PC chairs of ASPLOS led very successful conferences, and I mostly followed in their footsteps. Nevertheless I identified two narrower goals for additional efforts: (1) to continue to enhance the image of ASPLOS as a broad, multidisciplinary conference, and (2) to continue to raise the bar for quality and fairness in our review process. This document describes the processes I used, summarizes the program, and acknowledges those who made this program possible.

Scope of ASPLOS

The call for papers emphasized that ASPLOS is a broad multidisciplinary conference open to new and non-traditional systems related topics. I reached out to "non-traditional" SIGs and other organizations to advertise the call. One concrete outcome is that ASPLOS is "in-cooperation" with SIGBED this year. I hope our ties with SIGBED continue to strengthen in the coming years.

The call also said that papers should emphasize synergy of at least two ASPLOS disciplines (broadly defined) and the submission form explicitly asked authors to identify these areas.

We received 217 submissions, a 12% increase from last year and a new record. About 147 papers identified architecture as one of their areas, 89 identified PL or compilers, and 94 identified OS. Many papers identified other broad areas, including verification, graphics, big data, networks, cloud computing, mobile computing, embedded systems, software engineering, and more.

For specific research topics, again there was a large diversity. The following were listed by more than 20 submissions: power/energy/thermal management (34 papers), parallel architecture (31), heterogeneous architectures and accelerators (28), caches (27), high-performance computing (26), OS scheduling and resource management (26), compiler optimization (23), software reliability (23), virtualization (22), parallel programming languages (21), and programming models (21). The accepted papers constitute a similarly diverse set of topics.

In summary, ASPLOS continues to live up to its vision of a multidisciplinary conference, attracting a broad range of systems related researchers.

The Review Process: Setting the Stage

I was fortunate to have an excellent program committee (PC) of 35 members and external review committee (ERC) of 69 members. An additional 43 reviewers contributed to the process.

Review process overview: I used a two phase process similar to recent ASPLOS conferences. Unlike ASPLOS'13 but like other ASPLOS conferences, I obtained phase 2 reviews for only a subset of papers. The two review phases were followed by an author response period, an intense 2-week long online discussion period, and a 1-day in-person PC meeting. ERC reviewers were required to participate in all parts of this process except for the PC meeting.

Aids for reviewer assignments: I assigned all reviews. To help find the best reviewers, I used the following aids new to ASPLOS. (1) The submission instructions allowed unlimited pages for the bibliography and required all citations to include all authors (i.e., no et al.). This borrows from NSF and makes it easier to identify potential reviewers from related work. (2) I asked reviewers to

suggest other reviewers as part of their review, borrowed from ISCA'13. (3) I allowed authors to (optionally) suggest reviewers.

I required all reviewers to write their assigned reviews themselves. Consulting with others for small aspects was allowed if it brought clear added value (after checking with me for conflicts), but the assigned reviewers needed to write the bulk of the review and give their own scores.

System enhancements for conflicts: I used double-blind reviewing and handled conflicts of interest using community norms. Sandhya Dwarkadas handled the 20 papers with which I had a conflict using the same process described here. I requested Eddie Kohler, the author of the HotCRP reviewing software, to add a new mechanism to handle PC chair conflicts. This mechanism allows the PC chair to yield (most) chair privileges to a different "manager" for a given paper, and allows reviewers to email paper-specific issues to the (anonymous) paper manager. This system (1) significantly streamlined handling the review process for Sandhya, (2) was better at hiding sensitive information from me than alternatives I had previously used, and (3) was better at hiding the conflict from assigned reviewers, preventing inadvertent guessing of author identity. I am grateful to Eddie for implementing such an intrusive change at very short notice.

Tone: I consistently set a tone that encouraged reviewers to not look for perfection but to focus on moving the field forward while being constructive and fair to all authors. I emphasized that there was no target acceptance rate and all worthy papers would be accepted. Accepting many papers implies multitrack sessions and potentially reduced interaction. To mitigate these downsides, the first day of the conference features lightning presentations and the second day features a poster session for all papers, borrowed from MICRO'12. Although I used numeric scores in the review forms, decisions were made based on the review texts.

Phase 1 Reviews

I assigned 2 PC and 1 ERC reviewers for each paper in phase 1. I received virtually all reviews within two days of the completion deadline (I learned to nag). Unlike ASPLOS'13 (but like other conferences), I did not move all papers to phase 2. An important benefit of two phase reviewing is to relieve reviewers of the burden of reviewing clearly weak papers. On the other hand, a process that rejects papers on the basis of just three reviews risks inadequately informed decisions and inadequate author feedback. I put significant effort to balance these opposing demands.

I classified papers with at least one accept score for overall merit to move to phase 2 by default. Papers with all reject scores with high confidence levels were candidates to not move to phase 2. I highlighted the remaining 38 papers (no accepts, but some undecided and/or low confidence scores) for online discussion among reviewers, using the HotCRP comments system which provides a persistent record. Many reviewers voluntarily discussed many other papers as well.

In parallel, I read most reviews and every review of every paper that was a candidate for terminating at phase 1 or under discussion. For each review that did not provide adequate justification for a low score (e.g., a subjective claim that the paper was incremental without adequate citations of prior work), I used a HotCRP comment to request the reviewer to update their review or risk the paper moving to phase 2. I ensured all score changes were accompanied with justifications in the review text and not just an easy reaction to adjust to the majority. I used the HotCRP color tags feature extensively to help reviewers and myself distinguish different paper categories and track online discussion results.

After this process, all papers with at least one accept or undecided score proceeded to phase 2. A few papers with all reject scores also moved to phase 2 if the review text did not justify the reject

scores or had low confidence scores (when in doubt, I took the authors' perspective and moved the paper to phase 2). 30% of the papers did not move to phase 2.

Phase 2 Reviews

The phase 1 discussions and reviewer suggestions were very helpful in assigning reviewers for phase 2. All phase 2 papers had at least 5 total reviews, with at least 3 from the PC. The maximum total number of reviews for a PC or ERC member was 19 or 7 respectively.

Mid-phase 2, I used a tool by Andrew Myers[1] to calibrate excessive negativity or exuberance among reviewers. I sent each reviewer their estimated bias score and an anonymized list of all reviewers' scores. A positive (negative) bias indicated a tendency to give scores that are higher (lower) than others who reviewed the same paper. I believe (hope) this calibration motivated some reflection as reviewers moved to the critical decision-making stages.

The phase 2 reviews were due three days before the author response period. I wanted to ensure high quality reviews to facilitate responses focused on substantive issues. Since there was limited time, I distributed the quality assurance task among all reviewers. The last reviewer that submitted the review on a paper was asked to do a *review sufficiency check (RSC)* to ensure that (1) the reviews provided sufficient information to make an informed decision for the paper and (2) the reviews provided sufficient feedback to the authors. If the paper passed the RSC, the reviewer colored the paper purple (using HotCRP color tags); otherwise, they noted the problems.

The RSC mechanism significantly enhanced my ability to find the problem papers and focus my attention on them. Many reviews were updated and we identified 18 papers for additional reviews. When the author response started, *all* reviews from phase 2 had already been received (I got good at nagging). All post-phase2 (post-RSC) reviews had been requested and some were already received! The authors were notified that all pending reviews were late requests and they wouldn't be penalized for not responding to them. At the end, we received a total of 902 reviews.

Online Discussion After Author Response

This was the most critical phase of the review process. The goal for the online discussion phase was to triage papers into preliminary accept (tagged green, to be presented quickly at the PC meeting), preliminary reject (tagged red, not to be presented at the PC meeting), and discuss (tagged yellow, and the focus of most of the PC meeting).

Each paper was assigned a lead from the PC or ERC to initiate discussion among its reviewers (tagging the paper purple) and lead it towards a **consensus** (unanimous agreement) for green (accept) or red (reject). If significant discussion did not lead to a unanimous agreement, the paper was marked yellow (discuss). For these, the discussion was expected to reconcile as many differences among the reviewers as possible, leaving only a few substantive differences for a focused PC meeting discussion. That is, tagging yellow was not simply a way to duck work during this phase. These decisions assumed liberal use of shepherding, but not for adding new results.

I monitored discussions for all papers, nagging reviewers, clarifying policies, reminding philosophies, and obtaining yet more reviews as needed, while paying extra attention to papers turning red and those leaning to yellow.

---

[1] http://www.cs.cornell.edu/andru/reviewAdj/

Although this phase required significant effort, it played a critical role in running a smooth PC meeting with well-informed decisions. No system is perfect, but I identify the following advantages of having an online discussion phase with mandatory participation from all reviewers.

(1) It ensured non-PC reviewers' opinions were adequately represented during the PC meeting. In a conference as broad as ASPLOS, often the non-PC reviewers provide key expertise. We accepted some papers where the only champions were ERC reviewers because they forcefully argued for the paper during the online discussion phase.

(2) It allowed reviewers to reflect on each others' opinions in ways not possible when the first discussion is at the PC meeting; e.g., by reading previous work or confirming an opinion from another expert. These reflections resolved many differences among reviewers that could have become unhelpful sticking points in the limited time of a PC meeting.

(3) It allowed me to provide clarifications on matters of policy and philosophy that sometimes needlessly take up too much PC meeting time (e.g., does a prior workshop submission constitute prior art or does the paper meet the multidisciplinary criteria of ASPLOS).

(4) Mandatory participation (and judicious nagging) reduced (hopefully eliminated) the number of PC members arriving at the meeting without reading author responses and other reviews.

 (5) It resulted in a more effective PC meeting with more engagement from all members. I reminded the PC that most of the meeting was focused on papers where ≥ 5 reviewers had failed to reach consensus after significant deliberation, elevating the importance of their input with clear issues to debate.

This phase identified 23 green and 65 yellow papers for the PC meeting.

PC Meeting

The PC meeting took place at the O'Hare Hilton on October 30th (a week day) from 8am to 6:15pm. One PC member could not attend due to an injury. I made all papers and reviews available to all PC members (sans conflicts), assigned a PC lead to each green and yellow paper, and sent all meeting procedures two days before the meeting. A PC member was allowed to turn any paper to yellow to discuss in the meeting.

The guiding philosophy for the PC discussion was to seek consensus rather than make decisions through narrow majority votes and to give the reviewers (who had read the papers) priority in making the final decision.

For discussion order, within each color, I mostly used the average overall merit score, but I interleaved sessions of green and yellow papers in the morning. We covered about 50 papers before lunch, with the green papers taking little time. For the yellow papers, if about 5 minutes of discussion did not reveal a path to consensus among the reviewers, then we took a full PC (without the reviewers) vote. This vote was advisory input for the reviewers and was recorded in HotCRP. In many cases, the PC vote changed some reviewers' minds and resulted in an immediate consensus decision among the reviewers. In 9 cases, the differences remained and the paper was tabled.

For tabled papers, the reviewers and any interested PC members got together during the ample breaks (with food!) to attempt a consensus. I nudged these small-group discussions along as needed; many were long and intense. A late afternoon session revisited these papers with the full PC. If the reviewers had reached the same agreement as the majority in the previously recorded PC vote, then this result was simply recorded as the final decision. If the new reviewer consensus

was different from the PC vote, then the reviewers provided their rationale, entertained questions from the PC, and the new consensus vote was recorded as the final decision. If the reviewers could not reach a consensus even after all the deliberation, then we used the majority vote among them as the final decision. A tie among reviewers used the earlier full PC majority vote for the decision.

During the meeting, a PC member conflicted with a paper left the room for that paper (Sandhya led the discussion for my conflicts). Papers of PC members were handled similar to the others.

We accepted 49 papers, 12 of which were shepherded. The acceptance rate of 23% was coincidentally the same as last year. 10 accepted papers were co-authored by a PC member out of 24 PC submissions. For most papers and all rejected papers that had an author response, the lead for the paper included a brief summary of the online and PC meeting discussions with the review.

Reflections

In the last several years, our reviewing processes have come a long way, but there appears to still be scope for improvement. My review checks and the distributed review sufficiency checks (RSCs) resulted in many reviews being updated. The mandatory online discussions, adopted by some but not all recent conferences, clarified many misunderstandings and changed opinions. I believe the high quality of these pre-PC meeting deliberations and the consensus-driven approach led to more informed (and efficient) decisions during the PC meeting. For the papers where consensus was hardest, the face-to-face, small group deliberations during the PC meeting provided opportunity for more reflection before a final decision.

I would like to credit all past PC chairs from whom I liberally borrowed many ideas. Specifically, Ras Bodik (ASPLOS'13) first exposed me to a review process that employed intensive online discussion and was consensus-driven. The ASPLOS'14 process was mostly based on his framework. I also benefited from discussions with Vikram Adve (ASPLOS'10), Christos Kozyrakis (MICRO'13), Margaret Martonosi (ISCA'13), and Onur Mutlu (MICRO'12).

The Program

The 49 accepted technical papers are organized in two tracks. Monday morning will feature two lightning sessions with overviews of all papers to help you choose the sessions to attend. For closer interaction with authors, Tuesday afternoon features a poster session with all papers.

Several plenary sessions complement the technical papers. Monday morning kicks off with a keynote by Brad Calder (Microsoft) describing his experiences with a modern commercial cloud operating system, a topic that directly resonates with many ASPLOS submissions. Tuesday morning starts with a keynote by Jeff Gehlhaar (Qualcomm) describing a brain-inspired processor that overcomes the obstacles of post-Dennard scaling but creates new challenges that could inspire many future ASPLOS submissions. Tuesday ends with the ASPLOS tradition of a "Wild and Crazy Ideas (WACI)" session. The WACI team, ably led by Luis Ceze and Karin Strauss, has done a remarkable job attracting frontier ideas that are sure to inspire, with a cornerstone invited WACInote by Josh Smith (U Washington) on RF-powered computing and communication. Wednesday will start with a debate on whether specialized systems will supplant general-purpose alternatives, moderated by David Wood, again a hot topic touching many ASPLOS submissions.

This year, ASPLOS will have a business meeting and an awards lunch for the first time. Also new is that the technical papers will be openly available in the ACM digital library about a week before the conference. In spite of the packed program, we have included many breaks and a one-of-a-kind excursion – we hope these enable serendipitous interactions leading to technical inspirations and collaborations that make some of the most memorable conference experiences.

Acknowledgements

The program is the result of the hard work of many people, several of whom went above and beyond the call of duty. The general co-chairs, Rajeev Balasubramonian and Al Davis, were an absolute pleasure to work with. I started to imagine them with a magic wand in their hands – no matter what I asked of them, they said yes, even though it meant more fund raising and attention to more countless details.

I benefitted throughout from the wisdom and quick response times of the steering committee, and earlier mentioned my gratitude to past PC chairs. The submission chairs, my students Rakesh Komuravelli and Hyojin Sung, tirelessly and cheerfully took care of the submission web site. They also attended the PC meeting, recording decisions, moving conflicts in and out, and handling numerous logistics required for a smooth meeting. The review process I described would not be possible without HotCRP and the generosity of its author, Eddie Kohler. Sandhya Dwarkadas handled my conflicts, diligently following all the details of this year's review process. I was originally undecided on whether to have a WACI session, but the infectious enthusiasm of Luis Ceze and Karin Strauss and their volunteering to lead the charge enabled me to make the right decision. David Wood agreed to organize the debate even though I interrupted his vacation. Manju Shevgoor diligently kept the conference web site up to date and Niti Madan contacted many places new to ASPLOS as publicity chair. Michelle Osborne, my assistant, helped with logistics for the PC meeting.

Finally, none of this would be possible without the tireless efforts of the PC, ERC, and other reviewers and the work of the authors. Many of these were truly inspiring.

Sarita Adve
ASPLOS'14 Program Chair
University of Illinois at Urbana-Champaign