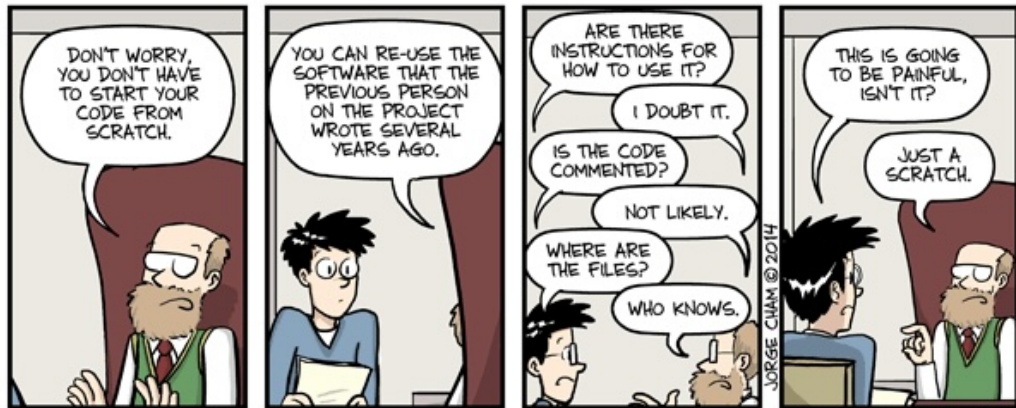


SIGCOMM and CoNEXT Artifact Evaluation and Infrastructure Needs

Sebastian Gallenmüller,
Damien Saucez, Georg Carle

Chair of Network Architectures and Services
School of Computation, Information and Technology
Technical University of Munich





WWW.PHDCOMICS.COM

1

¹ Jorge Cham, <https://phdcomics.com/comics/archive.php?comicid=1689>, Last accessed: Nov. 26, 2023

Goals

- Raising awareness for reproducibility in our community
- Rewarding papers and authors that make their research reproducible

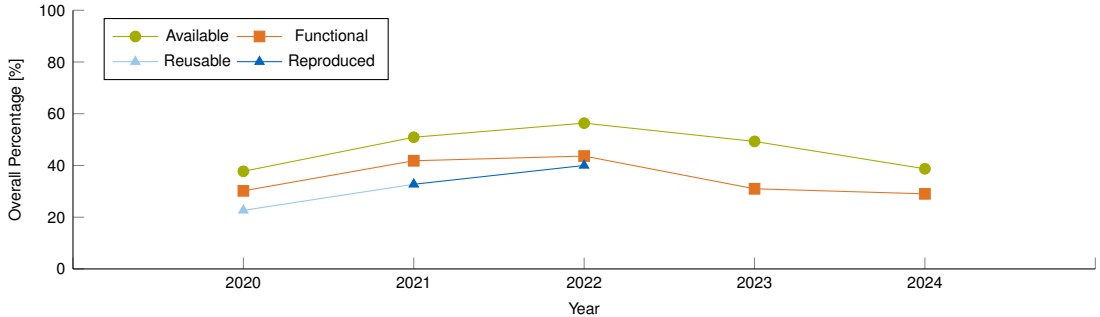
Means

- ACM definition for reproducible research:
 1. **Repeatability:** Same team executes experiment using same setup
 2. **Reproducibility:** Different team executes experiment using same setup
 3. **Replicability:** Different team executes experiment using different setup
- In 2015, ACM started its initiative to introduce badges to reflect different qualities of reproducibility

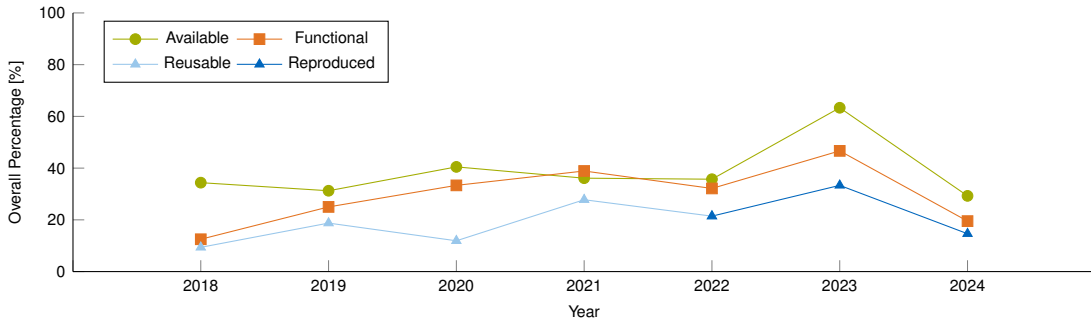


ACM reproducibility badge

Analysis of AE — Development of the SIGCOMM conference AE over Time



- Graph shows the percentage of papers that received one of the respective badges
- AE process for 2025 currently ongoing (results to be announced in October '25)
- SIGCOMM moved from [reusable](#) to [reproduced](#) badge in 2021
- Removed the evaluation of reproduced in 2022 (lower effort for authors and reviewers)
- Adoption stagnates over the years



- Graph shows the percentage of papers that received one of the respective badges
- CoNEXT moved from [reusable](#) to [reproduced](#) badge in 2022
- Adoption stagnates over the years

Examples of hardware requirements for reviews

- 3 × artifacts require Nvidia GPUs
- 3 × artifacts require Intel Tofino switch(es)
- 1 × artifact requires Intel SGX-capable CPUs
- RAM requirements:
 - Most demanding artifact required 512 GB in one machine
 - Another artifact requires several machines with at least 64 GB
- **SIGCOMM'23**: Large AWS instance (>1000 USD costs for reviewing)

	NVIDIA H200 NVL, 141GB HBM3 (900-21010-0040-000)	ab € 31550,29 4 Angebote
	PNY A100, 80GB HBM2e (TCSA100M-80GB-PB)	€ 27965,00 1 Angebot
	NVIDIA H100 NVL, 94GB HBM3 (900-21010-0020-000)	ab € 28853,99 12 Angebote
	NVIDIA GH100, 80GB HBM2e (900-21010-0000-000)	€ 38675,00 1 Angebot

Prices for Nvidia GPUs (June 2025)



Evaluating the artifacts of SIGCOMM papers

Damien Saucez
Université Côte d'Azur, Inria, France
damien.saucez@inria.fr

Luigi Iannone
Telecom ParisTech, France
luigi.iannone@telecom-paristech.fr

Olivier Bonaventure
UCLouvain, Belgium
Olivier.Bonaventure@uclouvain.be

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.
The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

Observations

- AE participation stagnates
- Hardware requirements may prevent effective reproduction

Survey² among authors and artifact evaluators

- Main message:
 - + AE is useful and interesting
 - AE is time consuming for authors and evaluators

ABSTRACT

A growing fraction of the papers published by CCR and at SIGCOMM-sponsored conferences include artifacts such as software or datasets. Besides CCR, these artifacts were rarely evaluated. During the last months of 2018, we organized two different Artifacts Evaluation Committees to which authors could submit the artifacts of their papers for evaluation. The first one evaluated the papers accepted by Conext'18 shortly after the TPC decision. It assigned ACM reproducibility badges to 12 different papers. The second one evaluated papers accepted by CCR and any SIGCOMM-sponsored conference. 20 papers received ACM reproducibility badges. We report on the results of a short survey among artifacts authors and reviewers and provide some suggestions for future artifacts evaluations.

CCS CONCEPTS

• General and reference → Evaluation;

KEYWORDS

Artifacts, Reproducibility

1 INTRODUCTION

Latest years have witnessed a steadily growing number of the papers, accepted by Computer Communication Review and the SIGCOMM-sponsored conferences, including artifacts such as simulation models, measurement datasets, software implementation, etc. These artifacts are an essential part of many of these papers, and artifacts' availability encourages other researchers to build upon and reproduce and extend previous results.

The ACM has proposed guidelines for assessing the quality of artifacts in publications¹.

These two evaluations focused on assessing if artifacts were available, functional, or reusable, which definitions are given by the ACM as follows.

- **Artifacts Available** author-created artifacts relevant to this paper have been placed on a publicly accessible archival repository.
- **Artifacts Evaluated - Functional** the artifacts associated with the research are found to be documented, consistent, complete, executable, and include appropriate evidence of verification and validation.
- **Artifacts Evaluated - Reusable** the artifacts associated with the paper are of a quality that significantly exceeds minimal functionality.

¹<https://www.acm.org/publications/policies/artifact-evaluation-badging>



Figure 1: Artifacts badges used for SIGCOMM evaluation.

The ACM proposes two additional definitions for results validation. **Results Replicated** and **Results Reproduced**. In an ideal world the evaluation committee should also have validated results. However, validating results is time consuming and the committees were not having enough resource to accomplish this mission. Meanwhile, it is worth to mention that most of the time when artifacts were evaluated as functional or reusable in our two evaluations, the paper results were also replicated. However, as we didn't define strict guidelines for results validation we could not conclude on the actual validity of results. Hence the choice of focusing on the artifacts only.

The ACM associates a badging system to these definitions. These badges can be used to visually indicate the conclusions of the artifacts evaluation committee. Badges used in our evaluations are presented in Fig. 1.

As the objective was to promote reproducibility and open science, the evaluation process was incremental with interactions with the authors to improve the quality of artifacts when possible. For that reason, the artifacts study was optional and authors had to expressly apply in order to have their artifact evaluated. Therefore, the absence of badge on a 2018 SIGCOMM-sponsored venue paper doesn't indicate a lack of reproducibility of a paper.

2 CONEXT'18 ARTIFACTS EVALUATION RESULTS

The evaluation of CONEXT'18 papers' artifacts was carried out shortly after the acceptance notification. Out of 14 accepted papers proposing an artifact, 12 have been awarded a badge. Seven of them received the Artifacts Available badge.

- DemoVLC: A C++-Free Massive MEMO System with Distributed LETs [5]

²Damien Saucez, Luigi Iannone, Olivier Bonaventure: Evaluating the artifacts of SIGCOMM papers. Comput. Commun. Rev. 49(2): 44-47 (2019)

- Pre-approval process for reproduction of papers (IMC)
- Testbeds can provide access to a diverse set of hardware
- Limiting the effort for authors and reviewers through a shared platform
- A framework enabling reproducibility by design
- Conferences suggest and incentivize the use of testbeds:
 - SIGCOMM '25 recommended CloudLab for reviewers (but not for authors)
 - Testbeds will provide long-term availability of environment to run artifacts

