Bioc Technical Advisory Board Minutes

7 March 2024

Present: Vince Carey, Laurent Gatto, Charlotte Soneson, Robert Shear, Helena Crowell, Henrik Bengtsson, Brian Schilder, Lori Kern, Jen Wokaty, Hervé Pagès, Alexandru Mahmoud, Ludwig Geistlinger, Stephanie Hicks, Marcel Ramos, Kasper Hansen, Rafael Irizarry
Apologies: Mike Love, Wolfgang Huber, Levi Waldron, Sean Davis, Davide Risso

:04 - :05 Previous [minutes](#) approved

:05 - :13 Issues with mac builds on kjohnson3
- Could we use a 'continuous run' rather than cron jobs (which seem less reliable on macOS Silicon). Interaction with Simon Urbanek.
- kjohnson2 (also mac Silicon) does not have these issues - could we rent an additional machine for this release?
- Would be good to increase stability before the release.

:13 - :14 Conference-related business
- Nominations for Bioc Awards - can't have diversity without lots of nominations!
    - https://bioconductor.org/about/awards/
- Long Workshops at BioC2024 are invite only. Does the TAB have any insight/opinions on who or what topics would be good to cover?
    - Suggestions from Marcel Ramos, based on previous workshops:
      📄 Intro course materials for BioC conferences
    - 4 intro-type workshops, annotation, public data, GRanges ('Fundamentals of Bioconductor')
- Potential joint Bioconductor/Galaxy conference in 2025 is being discussed.

:14 - :21 Upcoming release
- Contact R core for date, brief discussion of goals for the build system and core team.
- "Ephemeral infrastructure", build/check/distribute processes can be moved, decomposed, distributed.
- Initial test case: GPU-enabled builder kakopo1. Use strictly within the BBS container.
- Do TAB members have GPU-aware code that should be tested directly on GPU? What should we do to facilitate this, and what kinds of reports are useful? Traces of resource consumption via nvidia-smi should be feasible.
    - Kasper working on GPU-backed package for spatial transcriptomics. Currently in python, plan to use basilisk to wrap into an R package. Also runs out of the box on mac GPU.
- Quality enhancements to invisible infrastructure and to core software components for genomic analysis.
- Career development in information science and applications to genome biology.

- - Continuous learning.
  - Yearly conferences are good sources of information.
  - Bioc blog, R Journal.

:21 - :32 CRANhaven - Henrik
- Repository where recently archived CRAN packages can find shelter and time for maintenance (for a short time - not indefinitely).
- Packages on CRAN get archived with no notice to the public (only affected maintainers get a heads-up) -> potential for confusion and frustration.
- Many packages are eventually unarchived (36% - data from Lluis Revilla).
- CRANhaven hosts recently archived CRAN packages for 4 weeks.
- Gives end users some extra leeway.
- Can append CRANhaven to repos to use for installation.
- Built on top of R-universe - https://cranhaven.r-universe.dev/builds
- Currently contains 85 packages.
- Idea: Bioconductor can fall back on CRANhaven by default (could append https://cranhaven.r-universe.dev to the repos option). Perhaps only for release (to protect the user), not in devel (to make sure we notice the issue on time and give time for downstream dependencies to fix it)

:32 - :36 WASM/WebR package repo/demo - Alex

:36 - :47 GitHub Actions
- What guidelines should we add to contributions.bioconductor.org to ensure effective use in various contexts?
  - New package, not yet in ecosystem, help predict outcome of SPB/BBS evaluation, so include BiocCheck - the BBS container is relevant for Linux platform.
  - In-ecosystem package with associated GitHub repo.
  - Legacy package without associated GitHub repo.
- Comments:
  - biocthis is useful (including for beginners).
  - r-hub2 is coming out (GitHub self-hosted).
  - At the end of the day, the Bioc servers have the last word on whether things 'work'.
  - rworkflows was designed to provide a single action - minimal workflow script. No need to (find how to) update the workflow file when it breaks.
  - r-universe could be leveraged to regularly build packages - https://r-releases.r-universe.dev/builds
  - Considering options to 'build on demand' / not build packages that have not changed every day.

:47 - :52 Quick introduction to revised scRNAseq, gypsum, artifactdb - language-agnostic "ExperimentHub" analog

- ArtifactDb service currently hosted by Genentech (could be cloned by Bioc).
  - Discovery of missing values in PaulHSCData.
  - [Discovery](#) of missing columns in rowData of SegerstolpePancreasData.

:52 - :60 Open discussion
- New linked book on [python](#) and Bioconductor by Jayaram.
- (Tentative) release schedule has been posted:
  https://bioconductor.org/developers/release-schedule/
- Thought experiment: what if the Bioconductor release would be hosted on CRAN?
  - Bioconductor is less stringent in terms of issues, deprecation, …
  - Many packages on Bioc have a complex dependency structure.
  - A Bioc package for a given release only has to be compatible with a single R version.
  - Freedom to update packages at any time may make things harder for the user.
  - Continued discussion in #tech-advisory-board slack channel.