Writing a Successful Anton 2 Proposal

Marcela Madrid

Pittsburgh Supercomputing Center

anton-support@psc.edu

PSC
PITTSBURGH SUPERCOMPUTING CENTER

Carnegie Mellon University

University of Pittsburgh

Agenda

This webinar aims to explain:

- Anton 2 capabilities
- Who can apply for computer time
- How to write a successful proposal

Anton 2 at PSC

- Anton 2 is a special purpose supercomputer for molecular dynamics (MD) simulations of biomolecular systems with periodic boundary conditions.

- Designed and constructed by D. E. Shaw Research (DESRES).

- Made available without cost by DESRES for non-commercial research by US universities and other not-for-profit institutions.

- Hosted by the Pittsburgh Supercomputing Center (PSC) with support from the NIH National Institute of General Medical Sciences.

- Dramatically increases the speed of molecular dynamics (MD) simulations.

- By allowing MD simulations over time scales approaching milliseconds, it fundamentally changes the types of questions biomedical researchers can address.

- More information about Anton 2 can be found at:

  https://www.psc.edu/resources/anton/

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Anton 2 enables breakthrough science

- In just nine machine days on Anton 2 (the maximum yearly allocation per research group) a research group is able to produce results that would take years of continuous runtime on any other resource, reaching timescales of tens to hundreds of microseconds.

- Utilized for diverse biomolecular research: nucleic acids, lipids, proteins, designed nanostructures composed of nucleic acids or proteins, force field development, generation of long trajectories for analysis with AI methods

Carnegie Mellon University   **PSC**
PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Who and How

Who can benefit from running on Anton 2?

Researchers who need long-time scale (approaching millisecond-long) molecular dynamics simulations of biological systems (nucleic acids, proteins, lipids).

• How to get access to Anton 2?

Submit a competitive proposal. Request for Proposals are issued once a year.
Deadline: June 29, 2023.

Begin the submission process **a couple of days before the deadline**, because you will need to create an account at PSC.
https://www.psc.edu/resources/anton/anton-rfp/

Allocations are for one year.

Carnegie Mellon University   PSC   University of Pittsburgh

Tips for a successful proposal

1. Explain how Anton 2 will enable your group to achieve breakthrough scientific results that cannot be achieved in any other supercomputer.

2. Make sure your system will meet the technical requirements to run on Anton 2.

3. Provide all requested info in all required proposal sections.

Contact PSC in case of doubts: anton-support@psc.edu or grants@psc.edu

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

How can Anton 2 help you achieve scientific objectives

- Anton is designed primarily to accelerate classical MD simulations of biomolecular systems with periodic boundary conditions and explicit solvent.

- To maximize the benefit of Anton to the scientific community, proposed projects should focus exclusively on questions that will be greatly advanced by the unique Anton 2 capabilities.

- **Explain why achieving the scientific objectives requires access to Anton, and could not be efficiently achieved on any other high performance computing system, e.g.:**

  - Need for longer MD trajectories that are not feasible on conventional systems.

  - Applications that require long-time scale (approaching millisecond-long) MD to observe phenomena that cannot be accessible otherwise.

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Request for Proposals

June 29, 2023.

Who can apply:

- The principal investigator (PI) must be a faculty or staff member at a U.S. academic or non-profit research institution.

- A graduate student or postdoctoral researcher may not be a PI, but a qualified advisor may apply on their behalf.

- Each investigator can serve as a PI for only ONE application in a given allocation round.

- IMPORTANT: If a PI or co-PI is on another Anton proposal (as co-PI or PI, respectively), they must clearly demonstrate that the research efforts described in the two proposals are **distinct and independent** of each other, and do not constitute an effective doubling of the allocation. **You must demonstrate that the Anton work will happen primarily in the lab of the PI and will be directly supervised by the PI.**

- 25% of the total allocated time is reserved for PIs that have not previously had an allocation on an Anton system at PSC.

- **Researchers with no previous experience with Anton are encouraged to apply.**

Carnegie Mellon University   PSC   University of Pittsburgh

Requirements: Chemical Systems and Force Fields

- **The system that you propose to simulate must fulfill these conditions:**

- Chemical systems must have periodic boundary conditions and explicit solvent.
- Simulation cell can have only right angles (i.e., cubic or orthorhombic box), with a minimum of 45 Angstroms on each side.
- Between 25,000 and 700,000 atoms (including solvent atoms), Between 50,000 and 600,000 atoms are recommended for maximum efficiency.
- Consist of protein, DNA, RNA, lipids, water, and standard ions.
- **Proposed simulations must use recent variants of the following standard, non-polarizable biomolecular force fields:**
- CHARMM (e.g., CHARMM22, CHARMM27 CHARMM36– including CMAP corrections, AMBER (e.g., AMBER99, AMBER99SB, AMBER03, and AMBER19SB).
- Modified versions of the CHARMM and AMBER force fields, based on published research by DESRES.
- Water should be modeled with the SPC, TIP3P, or TIP4P models, or their variants.

Carnegie Mellon University   PSC   University of Pittsburgh
PITTSBURGH SUPERCOMPUTING CENTER

Enhanced sampling

- Can apply a uniform constant electric field.
- Position restraints on a per atom basis.
- Enhanced sampling is also available in four forms as follows:
- (i) simulated tempering, including adaptive weighting
- (ii) application of restraints between the centers of mass of groups of atoms
- (iii) application of conformational restraints, each based on the calculation of RMSD (root mean squared deviation) with respect to atomic positions of a given reference structure, and
- (iv) tempered binding.

- For restraints in both (ii) and (iii), equilibria and spring constants can be varied during a simulation according to a schedule or adaptively to implement a form of umbrella sampling

Contact us

- Applicants with systems:

- With custom parameters or molecules that are not included in the standard distribution of the supported force fields.

- Have dozens of restraints and/or restraints involving thousands of atoms.

- Shaped such that one dimension of the simulation cell is much larger than the others

- Please contact anton-support@psc.edu or grants@psc.edu

Estimating the amount of simulation time to ask for:

Chemical system (PDB ID)	Number of atoms	~ Performance (microseconds/machine-day)
DHFR (5DFR)	23,558	61.3
aSFP (1SFP)	48,423	53.0
FtsZ (1FSZ)	98,236	26.0
T7Lig (1A01)	116,650	21.9
bILAP (1BPM)	132,362	18.7
f1atpase	327,506	7.9
Tiled FDH-H	700,184	3.6

- 2.5-femtosecond time steps with long-range interactions evaluated at every other time step and Nose-Hoover thermostat applied every 100 time steps.

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Example of estimating MD simulation units

- Example:
- T7Lig (1A01) 116,650 atoms   21.9 microseconds in one machine day.

- Definition: An "MD simulation unit" is the amount of machine-time required to simulate 1 nanosecond for 50,000 atoms with Anton production parameters.
- There are 50,550 MD simulation units per machine-day

- A 100 microseconds-long simulation of T7Lig would require:
- (100 microseconds / 21.9 microseconds per machine-day) x 50,550 MD simulation units per machine-day = 230,822 MD simulation units.
- Maximum that can be requested: 460,000 MD simulation units (9.1 machine-days).
- Show these calculations in your proposal so it is easy to see how you made your estimates!

Carnegie Mellon University   PSC
PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Requesting MD simulation units

- Applicants are encouraged to target their requested resources at one of these two levels:

- 10 to 20 allocations will be made at or near 460,000 MD simulation units.

- 30 to 40 allocations will be made at or near 230,000 MD simulation units.

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

The proposal

- Proposals should be two to six pages in length, not including references.

- PIs who previously received an allocation on Anton at PSC must also provide a separate progress report (2 pages maximum) demonstrating their successful use of Anton to produce high-impact scientific results.

- May submit up to two additional supporting documents (e.g., published papers) in PDF format on the submission page.

- Proposal must be self-contained. Reviewers are not be required to refer to supporting documents or any other external documents when reviewing proposals.

- Grantsmanship is important

Carnegie Mellon University   PSC   University of Pittsburgh
PITTSBURGH SUPERCOMPUTING CENTER

Main Proposal Document Sections

- **The main proposal document must have these 6 sections:**

1. Summary of the project, including descriptive title of proposed research (400 words maximum for summary & title).

2. Name, address, email, and telephone number of the Principal Investigator and co-PIs.

3. Background information (1 page maximum): Include sufficient background information on the research field to allow reviewers to judge the scientific merit of the proposed research.

4. Scientific Objectives (2 pages maximum) explaining:
   - Why Anton is necessary for the planned project
   - Why the project could not be efficiently completed on conventional supercomputers
   - Scientific impact of their proposed project

5. **Project Feasibility and Team Qualifications** (2 pages maximum): Address these five points (the Simulation Requirements outlined in the RFP):

1. Simulations must be standard MD in constant NVE, NPT or NVT.
2. Simulation cell must have only right angles.
3. Supported Force Field
4. Number of atoms (between 25,000 and 700,000 atoms, including solvent)
5. Simulations will not finish in less than 1 hour on Anton

- State whether the proposed system has already been built and equilibrated. If not, provide evidence that the proposed system can be successfully built and simulated.
- Describe the expertise and experience of the research, to ensure that the proposed simulations can be successfully completed on Anton.
- If there is a co-PI on a proposal, provide information here demonstrating that the work described in the proposal will be conducted primarily in the PI's lab and will be directly supervised by the PI.

Carnegie Mellon University   PSC   University of Pittsburgh

6. Requested Resources (1 page maximum):

- Clearly state and provide justification for the number of Anton MD simulation units requested.

- The justification should provide strong scientific arguments as to why the **length** and **number** of proposed simulation runs will be both sufficient and necessary to achieve the stated scientific objectives.

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Progress Report for PIs with previous allocations

- Progress Report: Separate document. Required only for groups with previous allocations – 2 pages maximum

- Describe how the PI's previous allocation was used. Summarize findings or results.

- Results can be discussed in the context of publications, presentations, but **all information required to assess progress should be clearly communicated in the text of the progress report itself** and should not depend on reviewers reading supporting documents or any other external documents. (Progress Report should be self-contained.)

- Explain any significant digressions from the originally proposed resource usage plan for the prior period, if applicable.

- Describe reasons for underutilization of previous allocation, if applicable.

Summary of Common Proposal Pitfalls

- Poor grantsmanship (probably #1 issue), including:
    - *lack of clear hypothesis ("fishing expeditions")*
    - *lack of explanation about why the system you are studying is important or what contribution the results will have on the field (looking for breakthrough science)*
- Insufficient explanation of the differences between 2 proposals submitted by a PI/Co-PI pair and demonstration that the work will happen in PI's lab.
- Insufficient justification of the number of requested simulation units
- Insufficient justification of the need for Anton over standard supercomputing cluster
- Weak explanation of progress since last received allocation
- Insufficient justification for unused prior Anton 2 allocation

Carnegie Mellon University   **PSC**   PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

Submit the proposal

- Submit through the Portal as described in
- https://www.psc.edu/resources/anton/anton-rfp/

- Proposals are first reviewed for feasibility, missing documents, etc. at the PSC. Then sent to a review panel convened by the National Academy of Sciences (NAS).

- Rigorous and transparent process: The review committee is made public and there is a community comment period prior to the review. The entire committee must agree on the recommendation report.

- If notified of acceptance, you will receive info about login in to Anton and how to get started.

Carnegie Mellon University   **PSC** PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

After the Review

- Expect a notice of proposal acceptance by mid-November. The full NAS report is also published on their website.

- Allocation is for one year. Access to files for an extra year.

- A "Getting Started Workshop" is offered in early December:
an intensive 1-day Anton training workshops for new users
to help them prepare and start running their own biomolecules on Anton.

- Acknowledge Anton 2 in all publications resulting from this work.

- We collect publications and Ph.D. Theses resulting from Anton 2 work,
so keep them coming!

Trajectories

- Trajectories generated on Anton and all their associated data will be kept private until publication, or up to one year after the end of an Anton allocation, whichever comes first.

- After this period, the trajectories and data will become part of a collection of hundreds of long-timescale trajectories that will be made widely available to the research and education communities.

Carnegie Mellon University   PSC
PITTSBURGH SUPERCOMPUTING CENTER   University of Pittsburgh

- Slides and RFP are at
https://www.psc.edu/resources/anton/

- Please fill out Survey
https://forms.gle/kXbZ2jsR9Y2bfUz7A

- Best of Luck
- Questions?
- Thank you!