THEORETICAL AND COMPUTATIONAL BIOPHYSICS GROUP

## NAMD 2005 SURVEY REPORT

## D. BRANDON, J. PHILLIPS, & M. PUNKE

Theoretical and Computational Biophysics Group Beckman Institute for Advanced Science and Technology University of Illinois at Urbana-Champaign www.ks.uiuc.edu

## EXECUTIVE SUMMARY

The NAMD 2005 User Survey was announced to 5,476 NAMD 2.5 users on April 11 of 2005 and ran through April 22 of that year. Survey questions examined user satisfaction, the impact of the program on work quality, and user ratings of existing and planned features.

- A total of 770 usable responses were returned for the survey, providing a final response rate of 14.1%.
- A large proportion, 39.2%, of survey respondents had downloaded more than one version of the program. The majority of NAMD users are affiliated with academic institutions (86.2%), and most use the program for research purposes (83.7%). Moderate to high levels of expertise in macromolecular modeling are reported by 71.7% of users, with 38.1% reporting similar levels of expertise in using NAMD. NIH funding supports the work of 18.0% of users. The most popular platform for running NAMD is a local Linux cluster (32.5%). Just over half (53.8%) of user sites had more than one NAMD user.
- Users are highly satisfied with NAMD, with 76.9% indicating they are satisfied or highly satisfied with NAMD. And, 64% agree or strongly agree that NAMD has improved the quality of their work.
- A majority of both repeat and non-repeat users of NAMD are satisfied with NAMD, though repeat users are significantly more satisfied than non-repeat users. A majority of both groups feel NAMD has had a positive impact on the quality of their work.
- Many users, 39%, report using NAMD for all or most of their molecular dynamics simulations, while about half, 47.8%, use NAMD for some of their simulation work.
- Users indicating moderate to high levels of expertise in macromolecular modeling or in using NAMD were more satisfied than those low in expertise.
- No significant differences in satisfaction were found by NIH funding status, or by academic/non-academic affiliation.



#### OVERVIEW

NAMD is a parallel, object-oriented molecular dynamics code designed for highperformance simulation of large biomolecular systems; more on NAMD is available via its webpage, <u>http://www.ks.uiuc.edu/Research/namd/</u>. The NAMD 2005 survey is part of an ongoing effort (similar survey were conducted in 2000 and 2003) to ensure that NAMD is up-to-date, relevant and of high quality. NAMD users were identified via registration records, and contacted via e-mail with a request that they complete an online survey about NAMD (see locations below for a copy of the survey) during April of 2005. The following report details the administration and results of the survey.

#### NAMD 2005 Survey (complete copy)

A link to the survey forms the users completed is available here. Note that for analysis, interpretation and review purposes that all references to the items within the report are based on the numbering of the items as was used in the original survey.

Web form: <u>http://www.ks.uiuc.edu/Research/namd/survey/survey2005.html</u>

#### Table of Contents

NAMD User Profile	4
Ratings of Satisfaction	6
Ratings of Impact on Work Quality	7
Ratings of NAMD Usage	8
Ratings of Support, Documentation, and Overall Usability	9
Ratings of Planned Features*	11
Ratings by Repeat/Non-repeat Users	13
Ratings by Level of Macromolecular Modeling Expertise	15
Ratings by Level of NAMD Expertise	17
Ratings by NIH Funding Status	19
Ratings by Academic/Non-academic Affiliation	21
• Downloading Source Code, Generating Input Files, and Citing NAMD	23
Appendix: Survey Method	24

\*Features planned for future versions of NAMD



#### NAMD USER PROFILE

User profile characteristics based on survey responses are illustrated below.

- 39.2% are repeat users
- 86.2% are affiliated with academic institutions
- 83.7% use NAMD primarily for research purposes
- 71.7% report moderate to high macromolecular modeling expertise
- 38.1% report moderate to very high levels of experience with NAMD
- 18.0% report at least partial funding by NIH
- 46.2% report one NAMD user at their site, and one-third run on a Linux cluster.





#### RATINGS OF SATISFACTION

- A majority of users are satisfied with NAMD: 76.9% agreed or strongly agreed with the statement "I am satisfied with NAMD" (Q15).
- Mean satisfaction was 3.98 on a 5-point scale (1=strongly disagree, 5=strongly agree).
- Satisfaction with NAMD is slightly higher than in the NAMD 2003 and 2000 User Surveys, where mean satisfaction rated 3.93 and 3.74 respectively.



#### RATINGS OF IMPACT ON WORK QUALITY

- NAMD does positively impact work quality: 64% agreed or strongly agreed with the statement "NAMD has improved the quality of my work" (Q16).
- The mean response was 3.80 on a 5-point scale (1=strongly disagree, 5=strongly agree).





#### RATINGS OF NAMD USAGE

- Many users, 39%, report using NAMD for all or most of their molecular dynamics simulations, while about half, 47.8%, use NAMD for some of their simulation work (Q9).
- Excluding from the dataset those who don't use molecular dynamics, the proportion
  of those using NAMD for all or most of their molecular dynamics efforts increases to
  41.5%, with just over half, 50.8%, using NAMD for some of their molecular dynamics
  efforts.



NAMD Scalable Molecular Dynamics

#### RATINGS OFSUPPORT, DOCUMENTATION AND OVERALL USABILITY

- Responses to usability (Q10), and support and documentation items (Q14) indicated why respondents use NAMD, and their agreement with statements about specific aspects of the program.
- The three highest rated qualities are: NAMD is free (M=4.56), NAMD is a well-written program (M=4.07), and NAMD meets my needs (M=3.83). Amongst these qualities, regression analysis indicates the four strongest predictors of satisfaction with NAMD (Q15) are NAMD is a well-written program (Q14a), NAMD meets my needs (Q10a), and NAMD documentation is clear (Q14d).



Figure 5B: Support, Documentation and Overall Usability		
Question Stem	Mean <sup>†</sup>	Std Deviation <sup>†</sup>
Q10 I use NAMD because it:		
Q10b Is free	4.56	.84
Q10a Meets my needs	3.83	.97
Q10c Includes source code	3.74	1.25
Q10e Is better than other MD programs*	3.44	.83
Q10d Is user friendly	3.56	1.02
Q14 Rate your agreement with these statements describing NAMD:		
Q14a NAMD is a well written program	4.07	.79
Q14c NAMD support meets my needs	3.65	.87
Q14b NAMD developers respond to my requests	3.60	.88
Q14d NAMD documentation is clear		.98
Q14e NAMD documentation is complete	3.43	.97

<sup>†</sup>Responses from a 5-point scale: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree \*MD – Molecular Dynamics.



#### RATINGS OF PLANNED FEATURES

- Planned features are functionalities being considered for future versions of NAMD, e.g. making it easier to extend source code. In Q13 of the survey, users were asked to rate the value of 14 planned features to their work, using a 5-point importance scale (1-very unimportant, 5-very important).
- Mean results indicate that the most desirable feature is adding a trajectory analysis suite (M=4.11), while the least desired is SSE/Altivec acceleration (M=3.23). See Figs. 6A, 6B.





Figure 6B: Ratings for Planned NAMD Features, continued		
Question Stem	Mean <sup>†</sup>	Std Deviation <sup>†</sup>
Q13 Rate the importance of these PLANNED features to your work:		
13n Trajectory analysis suite	4.11	1.00
13i Improved user interface	3.97	1.01
13m Quantum/classical simulations	3.85	1.10
13k Implicit solvent models	3.81	1.07
13h Automated simulation setup	3.81	1.03
13I Polarizable force fields	3.80	1.04
13g Fault tolerance & recovery	3.74	1.00
13f Repeatable parallel runs	3.60	1.12
13a Improved serial performance	3.58	1.07
13d Scaling on 100's of CPUs	3.40	1.18
13j Easier to extend source code	3.37	1.16
13c Add-in hardware acceleration	3.36	1.10

<sup>†</sup>Responses on a 5-point scale: 1-Very unimportant, 2-Unimportant, 3-Unsure, 4-Important, 5-Very important



#### RATINGS BY REPEAT/NONREPEAT USERS

- Repeat users are defined as any user of NAMD 2.5 who has downloaded any other version of NAMD.
- A majority of repeat users, 79.8%, agreed or strongly agreed with the statement "I am satisfied with NAMD" (Q15) compared to 75.0% for non-repeat users. Mean ratings show greater satisfaction for repeat users (M=4.04) than non-repeat users (M=3.95), though these are not statistically significant differences. See Fig. 7A.
- A majority of both repeat (67.5%) and non-repeat (61.7%) users of NAMD agree with the statement "NAMD has improved the quality of my work" (Q16). See Fig. 7B.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: Repeater, 302; Non-repeater, 468.





<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: Repeater, 302; Non-repeater, 468.



#### RATINGS BY LEVEL OF MACROMOLECULAR MODELING EXPERTISE

- A majority of users across the three macromolecular modeling expertise (MME) levels\* agreed or strongly agreed (low-68.4%; moderate-76.7%; high-83.5%) with the statement "I am satisfied with NAMD" (Q15). High and moderate MME users are significantly more satisfied with NAMD than low MME users. See Fig. 8A.
- A majority of high (72.0%), moderate (65.4%), and low (51.8%) MME users agreed or strongly agreed with the statement "NAMD has improved the quality of my work" (Q16), while just over half of low MME users (51.8%) indicated similar ratings. High MME users indicate a significantly greater impact on work quality than low MME users. See Fig. 8B.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: Low expertise, 218; Moderate expertise, 266; High expertise, 286.





<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree Total N: Low expertise, 218; Moderate expertise, 266; High expertise, 286.

- High MME users rated NAMD significantly higher than low MME users on nearly all aspects of support, documentation, and usability; exceptions were NAMD being free, critical for work, and having complete documentation.
- High and moderate MME users rated these planned items higher than low MME users: implicit solvent models, polarizable force fields, quantum/classical simulations, and trajectory analysis suite. High MME users also rated improved serial performance more than low MME users.

\*Level of expertise categories were derived from the survey question "My level of expertise with molecular modeling is . . ." (Q5) that users answered on a 5-point scale (1-very low, 5-very high). For ease of interpretation, the two lowest expertise values were collapsed together, as were the two highest expertise categories, to produce the low, moderate, and high expertise categories used above.



#### RATINGS BY LEVEL OF NAMD EXPERTISE

- A majority of all NAMD expertise groups\* agreed or strongly agreed with the statement "I am satisfied with NAMD" (Q15), with 94.1% of high, 87.3% of moderate, and 66.7% of low NAMD expertise groups indicating agreement with this statement. Statistical comparisons indicate high NAMD expertise users are significantly more satisfied with NAMD than low NAMD expertise users. See Fig. 9A.
- Most high (85.9%) and moderate (79.9%) NAMD expertise users agreed or strongly agreed with the statement "NAMD has improved the quality of my work" (Q16), while nearly half (49.4%) of low expertise users agreed with this statement. High NAMD expertise users indicate a statistically stronger impact on work quality than low expertise users. See Fig. 9B.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree.

Total N: Low expertise, 417; Moderate expertise, 268; High expertise, 85.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree Total N: Low expertise, 417; Moderate expertise, 268; High expertise, 85.

- High and in most cases moderate NAMD expertise users rated NAMD significantly higher on all aspects of support, documentation, and overall usability. No mean differences were found in ratings of NAMD being free, including source code, or having complete documentation.
- High NAMD expertise users are significantly more interested in scaling on 100's of CPUs, repeatable parallel runs, easy to extend source code, polarizable force fields, and a trajectory analysis suite than users with low NAMD expertise.

\*Level of expertise categories were derived from the survey question "My level of expertise with NAMD is . . ." (Q5) that users answered on a 5-point scale (1-very low, 5-very high). For ease of interpretation, the two lowest expertise values were collapsed together, as were the two highest expertise categories, to produce the low, moderate, and high expertise categories used above.

#### RATINGS BY NIH FUNDING STATUS

- Users were asked to answer 'yes' or 'no' to the question "The work I do with VMD is funded (at least partially) by NIH" (Q5).
- The majority of both NIH-funded (77.5%) and non-funded users (77%) agreed or strongly agreed with the statement "I am satisfied with NAMD" (Q15). Statistical comparisons found no significant differences between NIH funded users and users who are not NIH funded in ratings of satisfaction with NAMD. See Fig. 10A.
- Majorities of both funded (70.3%) and non-funded groups (62.8%) agree or strongly with the statement "NAMD has improved the quality of my work" (Q16). Mean comparisons found no significant differences between NIH funded users and users who are not NIH funded in ratings of impact of NAMD on work quality. See Fig. 10B.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: NIH funded, 138; no NIH funds, 629.





<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: NIH funded 138; no NIH funds, 629.



#### RATINGS BY ACADEMIC/NON-ACADEMIC AFFILIATION

- Survey respondents were asked to identify their affiliation as academic, government, industry, or other (Q2); below are the findings using academic/non-academic categories.
- A majority of academic users (78.0%) agreed or strongly agreed with the statement "I am satisfied with NAMD" (Q15), as did a majority of non-academic users (69.8%). The mean ratings of the academic and non-academic user ratings of satisfaction are not significantly different from each other. See Fig. 11A.
- At 65.2% and 56.6% respectively, majorities of both academic and non-academic users agreed or strongly agreed with the statement "NAMD has improved the quality of my work" (Q16). No statistically significant differences were found between academic and non-academic mean ratings of work quality. See Fig. 11B.



<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: Academic, 664; Non-academic, 106.





<sup>†</sup>Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree. Total N: Academic, 664; Non-academic, 106.



#### DOWNLOADING SOURCE CODE, GENERATING INPUT FILES, AND CITING NAMD

• Results for why users download source code, applications used to generate input files for NAMD, and willingness to cite use of NAMD are below.

### Figure 12: Downloading Source Code, Generating Input Files, and Citing NAMD





#### APPENDIX: SURVEY METHODOLOGY

Following are details about the administration of the survey, including survey method, target population, survey schedule and response rates, sample validity, and questions used on the survey.

#### Survey Method

Population members received an e-mail solicitation asking them to complete an on-line survey, with the link to the survey containing information about the user. Participants were asked to complete all items on the survey form and submit their responses; upon submission, participants were to complete any items they had skipped, with an option to submit without doing so. After submission, users were thanked for their participation.

#### Target Population

Users of NAMD 2.5, as identified via registration records, constituted the target population of the survey.

#### Survey Schedule and Response Rates

	Dates/Activities		
	Initial		Closing
	Solicitation	Reminder	Totals
Date survey notice sent	April 11	April 18	4/22 End
Total Population (5,476)			
Number of persons receiving by date*	4,761	4,379	-
Responses up to next date	376	446	822
Response rate for total population (all 5,476)	6.9%	8.1%	15.0%
Cumulative response rate	6.9%	15.0%	15.0%
Repeat Users (1,335)			
Responses up to date of next notice	177	143	317
Response rate for total repeat user population	13.3%	10.7%	23.7%
Cumulative response rate	13.3%	24.0%	24.0%

\*Due to bounces.

#### Data Editing

Those responses that were considered incomplete were deleted from our dataset. The deletions fell into two categories: Unresponsive and duplicates.

 Unresponsive records were those instances in which respondents did not answer most of the questions in the survey, specifically those cases in which more than 28% of the questions were not answered.



 Duplicates were those instances in which there was more than one response for a person, based on their e-mail address.

Deletions left 770 valid records for analyses, as shown in the table below.

Deleted Survey Responses			
Deletions category	Unresponsive	Duplicates	Total
	50	1	51
Number of records in dataset after removing deletions 770			

• The response rate, after accounting for deleted records, for the initial population of 5,476 is 14.1%. Excluding those addresses that bounced back, i.e. a total population of 4,761, the response rate is 16.2%.

#### Statistical Significance Tests

Mean differences were checked via one-way ANOVA tests, with Scheffe tests used for post-hoc multiple comparisons. Regression analyses used a stepwise entry of variables as a means of identifying the strongest predictors of a given dependent variable. A .05 alpha was used as the significance criterion for all tests. All statistics were computed using the SPSS 13.0 statistical package.

#### Sample Validity

The validity of a sample size for representing an entire population is always a concern in survey research. Sample size calculators can provide measures of confidence intervals (+/- figures, i.e. 'margin of error') and confidence level measures (how certain you can be that an answer falls within a confidence interval). For a sample of 770 and a population of 5,476, using a standard test percentage of 50%, sample size calculations indicate that it can be said with 95% confidence that a given result for a question falls within a +/-3.3% confidence interval. (Figures were generated using Survey System sample size calculator: http://www.surveysystem.com/sscalc.htm).



#### Question Sets

# To aid in interpreting survey results, it is useful to view the question stems viewed by survey participants. Below are the survey questions, grouped by purpose: DEMOGRAPHIC ITEMS

Q. #	Торіс	Question Stem	
2.	Affiliation	Affiliation	
Respo	nses: Academic, gover	nment, industry, other (fill in blank provided)	
3.	Macromolecular modeling expertise	My level of expertise with macromolecular modeling is	
Respo	nses: Very low, low, mo	oderate, high, very high	
4.	NAMD expertise	My level of expertise with NAMD is	
Respo	nses: Very low, low, mo	oderate, high, very high	
5.	NIH funding	The work I do with NAMD is funded (at least partially) by NIH	
Respo	nses: Yes, no		
6.	Platform	I primarily use NAMD on	
Responses: Windows, Linux, Mac OS X, other UNIX; Local Linux cluster, large supercomputer, other (fill in blank provided)			
7.	Primary use of NAMD	I use NAMD primarily for	
Responses: Research, teaching, business, personal			
8.	Users at site	The number of people using NAMD at my site is	
Responses: 1, 2-4, 5-10, 11-20, 21 or more			
18.	Use of other programs	I consider myself a user of	
18a.	VMD		
18b.	BioCoRE		
Responses: Yes or No.			



QUEST	TIONS ABOUT SUPPORT, DOCUMENTATION, AND OVERALL USABILITY	
Q. 10.	Question stem: I use NAMD because it	
10a.	Meets my needs	
10b.	Is free	
10c.	Includes source code	
10d.	Is user friendly	
10e.	Is better than other molecular dynamics programs	
Q. 14	Question stem: Rate your agreement with these statements describing NAMD	
14a.	NAMD is a well written program	
14b.	NAMD developers respond to my requests	
14c.	NAMD support meets my needs	
14d.	NAMD documentation is clear	
14e.	NAMD documentation is complete	
Q. 15.	Question stem: I am satisfied with NAMD	
Q. 16.	Question stem: NAMD has improved the quality of my work.	
Responses: 1-Strongly disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly agree		



QUESTIONS ABOUT PLANNED ITEMS		
Q. #	Question Stem	
13.	Rate the importance of these PLANNED features to your work.	
a.	Improved serial performance	
b.	SSE/Altivec acceleration	
C.	Add-in hardware acceleration	
d.	Scaling on 100's of CPUs	
e.	Scaling for small molecules	
f.	Repeatable parallel runs	
g.	Fault tolerance & recovery	
h.	Automated simulation setup	
i.	Improved user interface	
j.	Easier to extend source code	
k.	Implicit solvent models	
I.	Polarizable force fields	
m.	Quantum/classical simulations	
n.	Trajectory analysis suite	
Responses: 1-Very unimportant, 2-Unimportant, 3-Unsure, 4-Important, 5-Very important		



QUESTIONS ON OTHER TOPICS			
9.	Quantity of simulations	I use NAMD for of my molecular dynamics simulations	
Responses: All, Most, Some, None, I don't use molecular dynamics			
11.		I have downloaded the NAMD source code to	
11a.	Use of Source Code	Examine algorithms	
11b.		Compile executables	
11c.		Locate bugs	
11d.		Add new features	
11e.		Reuse in my own programs	
Reponses: Yes or No.			
12.	Input file generation	I primarily generate input files for NAMD with	
Responses: VMD/psfgen, X-PLOR, CHARMM, AMBER, GROMACS, other (fill in blank provided)			
17.	Citation Likelihood	I would cite my use of NAMD in resulting publications.	
Responses: 1-Strongly Disagree, 2-Disagree, 3-Unsure, 4-Agree, 5-Strongly Agree			
19.	Suggestions for NAMD	What suggestions do you have for improving NAMD and NAMD support?	
Responses: Text box.			