

Project Rubric

The project score can be broken down into twelve components, each of which are scored independently. The total score is out of 100 points. The following are the components:

1. Implementation: 15 points
2. Design Document: 10 points
3. Evaluation Section: 10 points
4. Results Section: 10 points
5. Final Report - Abstract and Intro.: 5 points
6. Final Report - Approach/Design: 10 points
7. Final Report - Evaluation: 10 points
8. Final Report - Results: 10 points
9. Final Report - Conclusion: 5 points
10. Design Peer Review: 5 points
11. Evaluation Peer Review: 5 points
12. Results Peer Review: 5 points

We now describe each of the criteria and how they are graded.

Implementation: 15 points

The implementation points are primarily tied to whether the code properly implements each aspect of the system and how well it achieves its goals. The implementation serves as a pre-requisite to the results and conclusion sections. Without at least a partially functioning implementation, a team could score no higher than a 60% on the project. The points associated with the implementation are primarily focused on efficiency, clarity, effectiveness, and security.

Rubric:

- 15 points: The implementation is well done and represents the high-quality expected of a senior-year computer science student. The code is easy to read and is well documented. The code is designed to withstand attacks identified in the design document and runs reasonably efficiently, based on its language (i.e., interpreted languages are not penalized in comparison to compiled languages, but neither may be bloated). No functional errors are present.
- 13 points: The implementation generally works, but has one or two mild weaknesses. Weaknesses can include occasionally unclear code or mild inefficiencies. The code may not fully address an attack identified in the design document. A minor functional error or two may be present.
- 11 points: The implementation has obvious flaws, with a significant functional error or multiple minor ones. The code may be unclear in multiple places or has some significant instances of inefficiency.
- 9 points: The implementation mostly works, but has several missing or poorly implemented components. The code may be challenging to understand in multiple places or has significant inefficiencies.
- <9 points: The implementation has severe flaws that severely undermine its utility.

Design Document: 10 points

The design document will fully specify the system model, the components, the communication, and the implementation strategy that will be used to complete the implementation. The design document must include one or more figures of the system model and each of the communicating components. The design must specifically state the threat model of the adversary and all assumptions made by the approach. The design must be detailed enough to allow students to almost trivially reduce the implementation to code. A reasonably well-specified design is likely to be at least 4 pages of single-spaced text.

Rubric:

- Superior (10 points): The design document is extremely detailed and clear. A skilled sophomore-level CS student would be able to implement the code with little to no questions about interpreting the design. Each of the components is described in detail. The protocol between each of the parties is clear and the wire frame format of the messages is provided in complete technical detail, including order, size, and type of each field of communicated information (e.g., like an Internet RFC¹). The threat model is clear and comprehensively describes the attacker's abilities and limitations. The design should specify the number of computers and VMs that will be used, the programming language that will be used, and a general outline of the key functions (or classes) used along with the parameters that will go into each. A diagram of the system model and a diagram of the organization of each of these components is included. The diagrams are easy to interpret and are correct.
- Excellent (9 points): The design document is generally good, but has one or two mild weaknesses. Weaknesses can include occasional mild vagueness in the specification of the program design, components, or protocol. The threat model may be missing one or more considerations for the attackers. Some of the diagrams may have vagueness, inaccuracies, or unclear components. A skilled sophomore-level CS student would need to ask a few questions about interpreting the design in order to be able to implement the code according to the team's design.
- Good (8 points): The design document has a few obvious flaws, with omissions or vagueness surrounding key design points. Multiple weaknesses are present or at least one severe weakness is present. A skilled sophomore-level CS student would need to ask many questions about interpreting the design in order to be able to implement the code according to the team's design.
- Fair (7 points): The design document has many obvious flaws with at least two being severe. A skilled sophomore-level CS student would have significant difficulty using the design to implement the code.
- Poor (<7 points): The design document has severe flaws that severely undermine its utility.

Evaluation Section: 10 points

The evaluation section describes the empirical setup, the evaluation methodology, and the metrics and criteria for success. The evaluation must be sufficiently detailed that another networking student or professional would be able to re-create the exact same setup. The evaluation methodology must cover all critical components in the system, both from a security and a performance standpoint. The criteria for success must be clear and sensible. The criteria must be appropriately rigorous: a good solution should be able to successfully meet the criteria while a poor solution should fail to meet the criteria.

Rubric:

- Superior (10 points): The evaluation of the project is comprehensive and clear. The document clearly demonstrates the students have considered all the critical components in the system and have devised a way to determine whether those components operate successfully. The document includes discussion of security and performance goals and has clear metrics for whether an evaluated implementation meets those goals or not. All success conditions are fail-able: a poor implementation or approach would reasonably be declared a failure by the relevant metrics.

¹These details are specified in the RFC for the IPv4 header at <https://tools.ietf.org/html/rfc791#page-11>

- Excellent (9 points): The evaluation is generally good, but has one or two mild weaknesses. Weaknesses can include missing one critical component or general vagueness in how one of the evaluation methodologies would be performed. One of the evaluation metrics may be too vague or unclear, such as a metric that would not be fail-able.
- Good (8 points): The evaluation has a few obvious flaws, with omissions or vagueness surrounding key evaluation methodologies or success metrics. Multiple weaknesses are present or at least one severe weakness is present. As a result, another networking student or professional would have difficulty implementing the testing methodology or evaluating whether a tested system was deemed successful or not. Two or more metrics may not be useful, excessively vague, or not fail-able.
- Fair (7 points): The evaluation has many obvious flaws with at least two being severe. The evaluation would not yield certainty in distinguishing between a successful implementation and one that is not.
- Poor (<7 points): The evaluation has severe flaws that severely undermine its utility.

Results Section: 10 points

The results section will present the findings of the empirical study across a sufficient number of trials to have confidence in the results. The results should be described in tabular or chart form where appropriate. In addition to the raw results, the students should provide explanation/interpretation of the results and provide a comparison to other systems to allow the reader to understand the results and independently reason about them. For example, a student might use a ICMP ping probe or the average DNS latency numbers as a baseline for the delay of having a query/response protocol that must traverse the Internet. The results should clearly indicate whether the implementation and approach met the previously stated evaluation criteria.

To be clear, there will not be a grading penalty if an otherwise good implementation fails to pass the success metrics specified if those criteria are simply too onerous for many implementations.

Pre-requisite: The implementation must be completed successfully for each of the components described in the results section to have value. If the team does not have a viable implementation, no points can be earned in the results section. See the implementation rubric for more details.

Rubric:

- Superior (10 points): The results of the experiments are clearly described and the number of trials removes doubt that the performance numbers are simply due to outliers. The results are depicted in figures or tables when appropriate. The students explained each result clearly and in sufficient detail for the reader to understand the significant. Where appropriate, the students compared the numbers to other systems with similar properties, allowing the reader to intuitively judge the approach independently.
- Excellent (9 points): The results of the experiment are described in an appropriate format, but there may not be quite enough trials to confirm the results are robust. The explanation of results may be vague or repetitive from earlier sections in the paper. A small number of results may lack appropriate context for the reader to reason about the results.
- Good (8 points): The results are depicted in an appropriate format, with only a subset having an explanation or context. The number of trials may lead to reasonable questions about whether the results are robust.
- Fair (7 points): The results are presented accurately, but without much elaboration or context.
- Poor (<7 points): Invalid results are presented or the results are difficult for the reader to understand.

Final Report - Abstract and Intro.: 5 points

The final report will describe the project as a whole, combining information from the previous sections along with an abstract, introduction, and conclusion. The abstract will provide a quick summary of the problem being solved, the significance of the problem, the approach to solve the problem, and the key results from evaluating the approach. The introduction will elaborate on each of these points, clearly motivating the reader on why solving the problem is valuable for network security and why the project is a reasonable strategy for doing so. The introduction must provide a high-level idea of what the approach will be and summarizes the key contributions and results of the approach.

Rubric:

- 5 points: The abstract and introduction are of high quality and are comparable to such sections in the papers read in this course. The problem and approach are clearly described and motivated. The key results and contributions are clear.
- 4 points: The abstract and introduction are reasonable, but may not fully motivate or explain the problem or approach. Some key results and contributions may be missing.
- 3 points: The abstract and introduction vaguely or poorly describe the problem and approach.
- <3 points: The abstract and introduction fail to motivate or describe the problem and approach.

Final Report - Approach/Design: 10 points

The criteria and rubric for the final report's Approach or Design section are equivalent to those described in the design document. However, the approach must be merged into a final report with appropriate transitions rather than being treated as a separate document.

Teams are expected to have improved their approach/design section from its original version based on feedback and insights gained from the peer review process. A "change log" of all the alterations from the original version should be included in a separate "Improvements" section following the conclusion section of the paper. Given the benefit of the peer review, the expectations for the final version of this section are higher than the initial submission. If students do not make appropriate improvements based on the peer review, they should anticipate a lower score for this component.

Final Report - Evaluation: 10 points

The criteria and rubric for the final report's Evaluation section are equivalent to those described in the original Evaluation document. However, the evaluation must be merged into a final report with appropriate transitions rather than being treated as a separate document.

Teams are expected to have improved their evaluation section from its original version based on feedback and insights gained from the peer review process. A "change log" of all the alterations from the original version should be included in a separate "Improvements" section following the conclusion section of the paper. Given the benefit of the peer review, the expectations for the final version of this section are higher than the initial submission. If students do not make appropriate improvements based on the peer review, they should anticipate a lower score for this component.

Final Report - Results: 10 points

The criteria and rubric for the final report's results section are equivalent to those described in the Results document. However, the results must be merged into a final report with appropriate transitions rather than being treated as a separate document.

Teams are expected to have improved their results section from its original version based on feedback and insights gained from the peer review process. A "change log" of all the alterations from the original

version should be included in a separate “Improvements” section following the conclusion section of the paper. Given the benefit of the peer review, the expectations for the final version of this section are higher than the initial submission. If students do not make appropriate improvements based on the peer review, they should anticipate a lower score for this component.

Final Report - Conclusion: 5 points

The conclusion section should provide a paragraph summary of the work and the key findings. The conclusion should also highlight any remaining components that others could do as future work, such as scale testing.

Further, in a separate section after the conclusion, called “Improvements,” the team should provide a “change log” of all major alterations from the initial versions of the design, evaluation, and results sections when placed in the final document. These changes should include feedback from the peer review process, both direct feedback and insights gained from reviewing other teams’ work, along with any new ideas the students had along the way.

Peer Reviews for Design, Evaluation, and Results Sections: 5 points each

The peer reviews are graded individually for each student rather than as a team score. Each team member must perform a peer review on two other team’s projects, as assigned by the InstructAssist system. The reviews must be of high quality, providing actionable information about how to improve the work along with a critical assessment of the strengths and weaknesses associated with the document. The peer review score must be consistent with the written reviews. Quality reviews are likely to be at least half a page of single-spaced text.

Rubric:

- Superior (5 points): The review evaluates the document thoroughly, providing a specific, clear, and accurate list of the strengths and weaknesses of the document. The review provides actionable next steps for the authors to improve the work. The review is neither unduly harsh nor does it sugarcoat weaknesses.
- Good (4 points): The reviews is generally good, but has at least one instance that is mildly vague or lacks clear action items for the the authors. The review may miss one or two strengths or weaknesses. The review score may mildly deviate from the staff’s score.
- Fair (3 points): The review has multiple instances of vagueness or unclear clear actions. The review may miss multiple strengths or weaknesses, indicating the reviewer does not have a clear understanding of the submission. The review score may moderately deviate from the staff’s score.
- Poor (<3 points): The review provides little value to the authors and may be vague or inaccurate. The strengths and weaknesses may be incomplete. The review score may differ significantly from the assessment of the teaching staff.