• Project status updates

• Review of topics covered
  – Overall point of systems evaluation
    * Providing data necessary to make a decision
    * Presenting it on a convincing manner
    * Performance evaluation is not the only relevant data, but it’s important
  – Understand what you are measuring
    * Understand what make convincing evidence
    * Understand what you need to measure and how to measure it
    * Do some baseline tests of your workload generator and monitors to make sure they exhibit expected results
  – Understand the system under test
    * Where are the boundaries?
    * What is inside those boundaries that you are actually measuring?
    * How do the SUT boundaries relate to a deployment environment?
    * How do the SUT boundaries relate to the claims of the paper?
  – Evaluation should be a part of the research and development process
    * Convince yourself with data, not just bias
    * Often need preliminary evaluations
    * Understand the scope and limitations of your work
    * The more evaluation you do along the way, the less biased you are likely to be
  – Recognize the strengths and weaknesses in evaluations that you read
    * There is a huge difference between active and passive reading
    * Think actively about what you need to see to be convinced
    * Look for biases or basic mistakes
  – Common mistakes in systems evaluation
    * No goals or biased goals
    * Ignoring significant factors
    * Analysis without understanding the problem—a reason to have a concrete problem statement
    * No sensitivity analysis
    * Ignoring variability
  – Use the statistical tools available to you
* For selecting the number of runs
* For understanding the confidence in your results
* For showing difference or sameness in results
* The question is “have you found something real?”
* Don’t just eyeball graphs
  · Compute CIs
  · Do linear regressions
  · Compares means, percentiles, etc.

- Work done to make work repeatable helps everyone
  * Not just for others, but for yourself too
  * Lets you re-run things when something changes
  * Makes it easier for others to build on your work
  * These things bitrot incredibly fast
  * The closer you can get to “experiment as function call”, the happier you will be
  * Assuming your environment is fragile forces you to build more repeatable research
  * Keep track of everything

- You are running experiments all the time, the only question is whether you are learning anything from them

• For next time

  – Weekly 3 due Friday midnight
  – Papers 4 due next Tuesday
  – Talks next Thursday (Anil Kumar, Jithu, George, Jonathon)
    * Schedule a time to meet with me
    * Any questions about what’s expected?