

The group „Scientific Computing“ conducts research on high performance I/O optimizations, energy efficiency, and simulation of cluster infrastructure. We have expertise in parallel programming and environmental modeling.

Head of the group: Prof. Thomas Ludwig (ludwig@informatik.uni-hamburg.de)

### High Performance Input/Output

In high performance computing it is important to consider I/O capacity and bandwidth. A multitude of cluster file systems exist, each with different requirements, interfaces and behaviors. Benchmarks are used to evaluate their performance characteristics for specific use cases. However, due to the fact that the performance of file systems usually depends on the used access patterns it is difficult to compare them with each other. While storing large amounts of data is usually unproblematic, storing a large number of files poses another challenge, because of the associated management overhead. Some applications produce billions of files, pushing file systems to their limits. One important factor are file system semantics which can affect the overall performance heavily. The group’s focus lies on evaluating their effects and proposing new strategies with regards to these semantics.

Contact: Michael Kuhn (michael.kuhn@informatik.uni-hamburg.de)

### Simulation of Distributed Systems

Understanding supercomputers architecture is important to assess observed application performance and to identify bottlenecks. Simulating application behavior on various systems with individual components enables to project performance on future systems. Furthermore, the impact of replacing parts of the system like I/O subsystems with faster components can be analyzed in simulation. Simulation also allows to evaluate the impact of MPI-internal implementations without coding them into existing MPI libraries.

Contact: Julian Kunkel (kunkel@dkrz.de)



Figure 1: Analysis Workflow

- Execution of MPI applications is recorded in trace files.
- A cluster model describes the characteristics of the hardware environment and the interconnect topology. Application traces are mapped to the cluster model.
- A discrete-event simulation is performed. Internal events like server I/O operations and network activity is recorded as well.
- Results can be inspected visually in the same viewer as the original trace files.

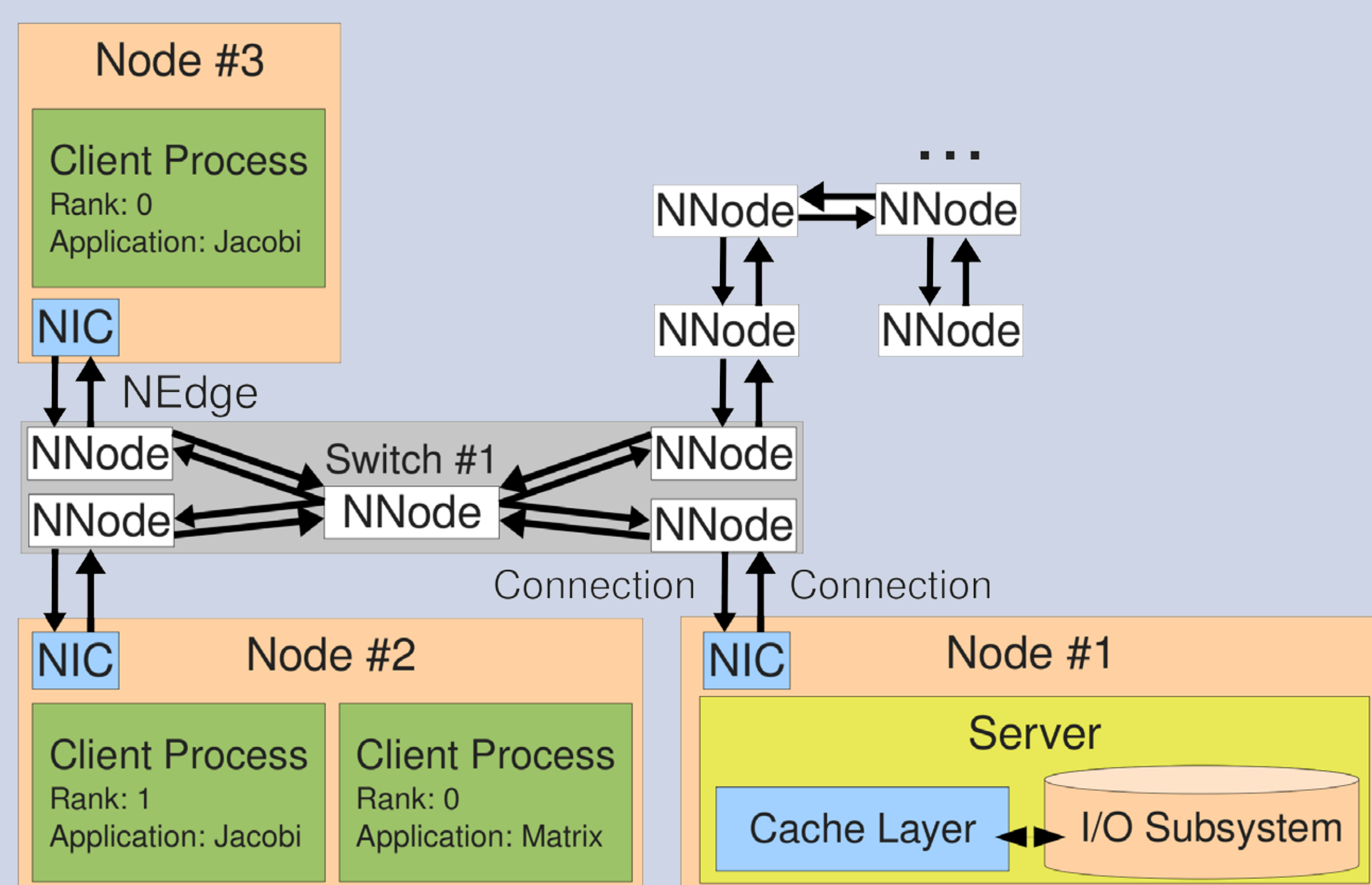


Figure 2: Cluster Model

### Environmental Modeling

For the use of HPC environmental modelling plays an important role. Climate models are well known as typical users of HPC infrastructure. Nevertheless, a number of other environmental modeling aspects are also reliable on the access to both, high computational power and large storage facilities for the simulation results. At our group models representing the ecosystem of the North Sea are in the focus of environmental modelling activities. For example, based on the hydrodynamical model HAMSOM (Hamburg Shelf Ocean Model) the effects of offshore wind farm installations are looked at. The changes in the marine environment is analyzed in relation to the wake effect, which results from the rotation of the propellers. The calculation of the nearly one million wet grid points from the North Sea topography needs up to date computational power and the possibility to store large volumes of simulation results.

Contact: Dr. Hermann Lenhart (hermann.lenhart@zmaw.de)

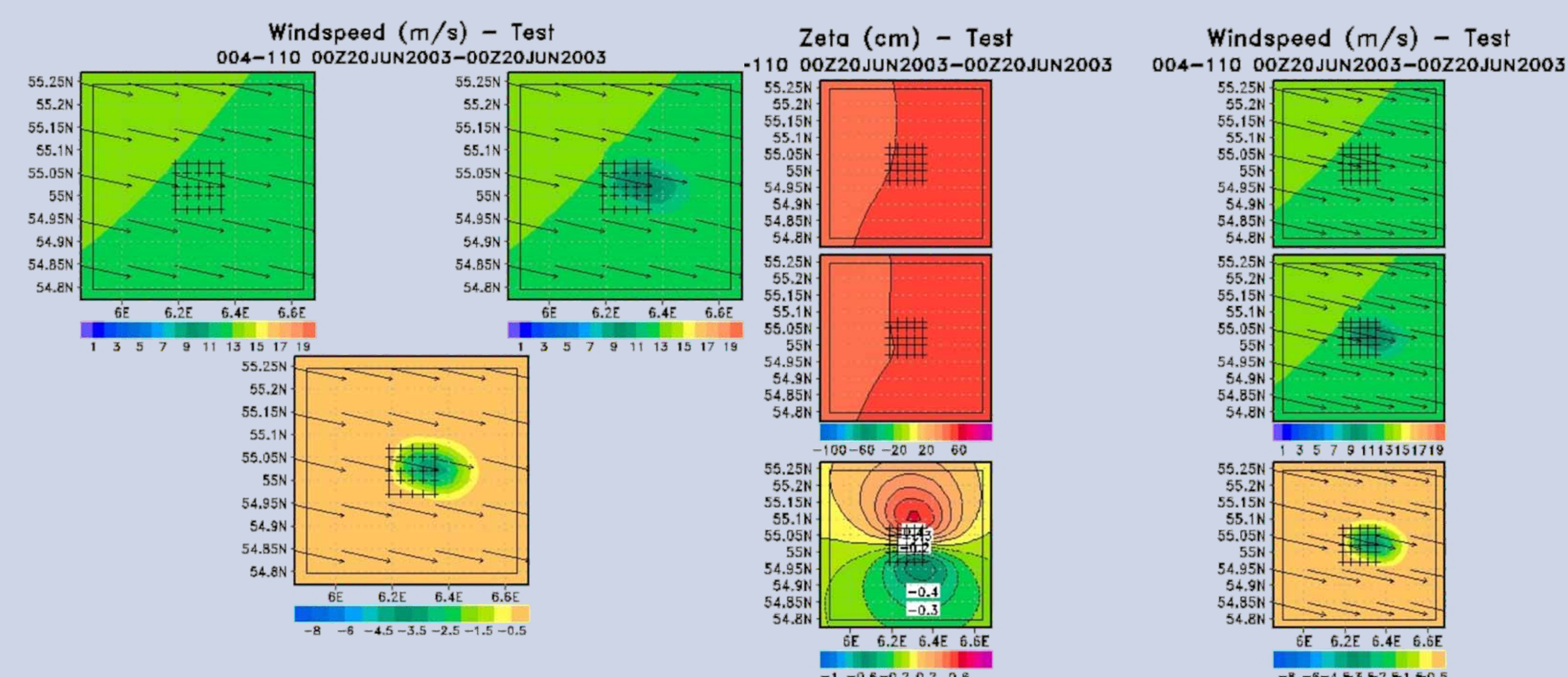


Figure 3: Application of the wake effect for the upper layer of the North Sea on the 20th June 2003. Left: Realistic wind conditions (m/s) including the wind reduction downstream of the offshore wind farm (indicated by the rectangle in the center). Right: Resulting difference in sea level (cm) between the undisturbed and the wake simulation.

### Energy Efficiency of HPC Installations

During the last years we see a strong increase in energy consumption and related costs in high performance computing. These costs are already in the range of acquisition costs of the whole computer. Our project – called Energy-Efficient Cluster Computing – aims at making high performance computing more efficient with respect to economic and ecological aspects. The basic idea is to determine relationships between the behavior of parallel programs and their impact on the energy consumption of the underlying compute cluster. Further, strategies will be developed to reduce the energy consumption with as little impact as possible on program performance.

Contact: Timo Minartz (minartz@informatik.uni-hamburg.de)

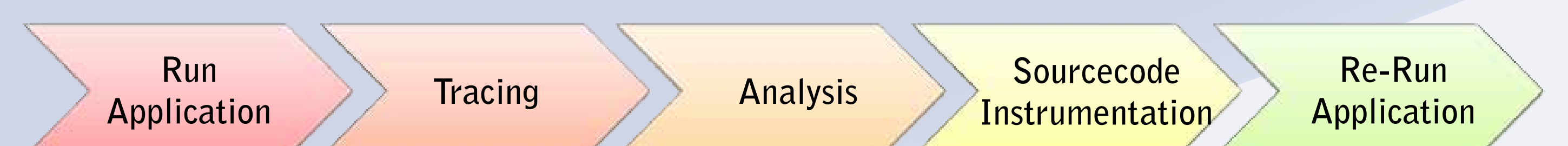


Figure 4: Project Workflow