While the big data revolution in the consumer, business, and social networks domains is widely known, a similar revolution is taking place in the sciences and engineering driven by high-throughput instrumentation. This talk will feature big data challenges in genomics and systems biology, primarily due to advances in sequencing that resulted in several orders of magnitude throughput increases per unit cost during the last decade. These advances are democratizing big data generation capabilities and spawning new scientific inquiries that would not be feasible otherwise. High-throughput sequencers enable diverse applications, each requiring its own class of supporting algorithms. I will present an overview of my group’s research in addressing these challenges through the development of parallel string, graph, and machine learning algorithms, engineering high performance software, and community building efforts.

Srinivas Aluru is Executive Director of the Institute for Data Engineering and Science (IDEaS) and Professor in the School of Computational Science and Engineering at Georgia Institute of Technology. He co-leads the NSF South Big Data Regional Innovation Hub which nurtures big data partnerships between organizations in the 16 Southern States and Washington D.C., and the NSF Transdisciplinary Research Institute for Advancing Data Science. Aluru conducts research in high performance computing, large-scale data analysis, bioinformatics and systems biology, combinatorial scientific computing, and applied algorithms. He is a recipient of the NSF Career award, IBM faculty award, Swarnajayanti Fellowship from the Government of India, the John. V. Atanasoff Discovery Award from Iowa State University, and the Outstanding Senior Faculty Research Award, Dean’s award for faculty excellence, and the Outstanding Research Program Development Award at Georgia Tech. He is a Fellow of the AAAS, IEEE, and SIAM, and is a recipient of the IEEE Computer Society Golden Core and Meritorious Service awards.