**About the job**

**Job Summary**As an SDE at NetApp India’s R&D division, you will be responsible for the design, development and validation of software for Big Data Engineering across both cloud and on-premises environments. You will be part of a highly skilled technical team named NetApp Active IQ.  
  
The Active IQ DataHub platform processes over 10 trillion data points per month that feeds a multi-Petabyte DataLake. The platform is built using Kafka, a serverless platform running on Kubernetes, Spark and various NoSQL databases. This platform enables the use of advanced AI and ML techniques to uncover opportunities to proactively protect and optimize NetApp storage, and then provides the insights and actions to make it happen. We call this “actionable intelligence”  
  
You will be working closely with a team of senior software developers and a technical director. You will be responsible for contributing to the design, and development and testing of code. The software applications you build will be used by our internal product teams, partners, and customers.  
  
We are looking for a hands-on developer who is familiar with Spark and Scala, Java and/or Python. Any cloud experience is a plus. You should be passionate about learning, be creative and have the ability to work with and mentor junior engineers.  
  
**Job Requirements  
  
Your Responsibility**

* Design and build our Big Data Platform, and understand scale, performance and fault-tolerance
* Interact with Active IQ engineering teams across geographies to leverage expertise and contribute to the tech community.
* Identify the right tools to deliver product features by performing research, POCs and interacting with various open-source forums
* Build and deploy products both on-premises and in the cloud
* Work on technologies related to NoSQL, SQL and in-memory databases
* Develop and implement best-in-class monitoring processes to enable data applications meet SLAs
* Should be able to mentor junior engineers technically.
* Conduct code reviews to ensure code quality, consistency and best practices adherence.

**Our Ideal Candidate**

* You have a deep interest and passion for technology
* You love to code. An ideal candidate has a github repo that demonstrates coding proficiency
* You have strong problem solving, and excellent communication skills
* You are self-driven and motivated with the desire to work in a fast-paced, results-driven agile environment with varied responsibilities

**Education**

* 5+ years of Big Data hands-on development experience
* Demonstrate up-to-date expertise in Data Engineering, complex data pipeline development.
* Design, develop, implement and tune distributed data processing pipelines that process large volumes of data; focusing on scalability, low -latency, and fault-tolerance in every system built
* Awareness of Data Governance (Data Quality, Metadata Management, Security, etc.)
* Experience with one or more of Python/Java/Scala
* Proven, working expertise with Big Data Technologies Hadoop, HDFS, Hive, Spark Scala/Spark, and SQL
* Knowledge and experience with Kafka, Storm, Druid, Cassandra or Presto is an added advantage