In this paper, we develop a decentralized probabilistic method for performance optimization of cloud services. We focus on Infrastructure-as-a-Service where the user is provided with the ability of configuring virtual resources on demand in order to satisfy specific computational requirements. This novel approach is strongly supported by a theoretical framework based on tail probabilities and sample complexity analysis. It allows not only the inclusion of performance metrics for the cloud but the incorporation of security metrics based on cryptographic algorithms for data storage. To the best of the authors’ knowledge this is the first unified approach to provision performance and security on demand subject to the Service Level Agreement between the client and the cloud service provider. The quality of the service is guaranteed given certain values of accuracy and confidence. We present some experimental results using the Amazon Web Services, Amazon Elastic Compute Cloud service to validate our probabilistic optimization method.