

Distribution of query processing time with
3-gram based Transformer models

