XTR/ScaNN (document top k=100, token top k=1000), BEIR SciFact (iterations=300, total=107.6ms) Query Encoding search batched enumerate scores **Estimate Missing Similarity** get did2scores add ems get final score sort scores

400

Latency (ms)

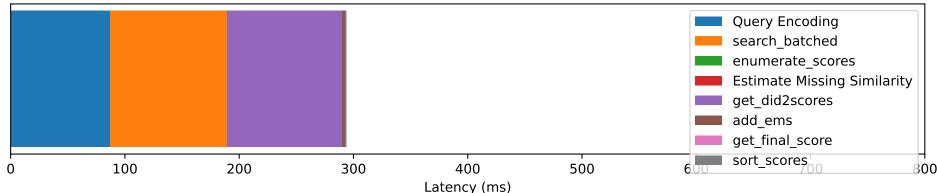
500

100

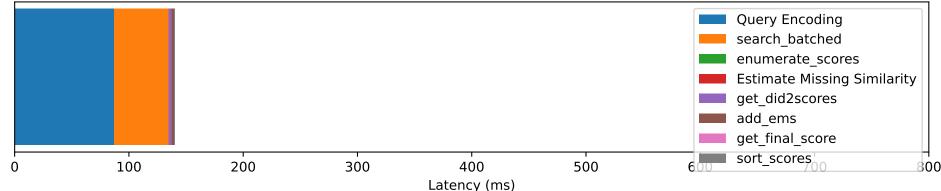
200

300

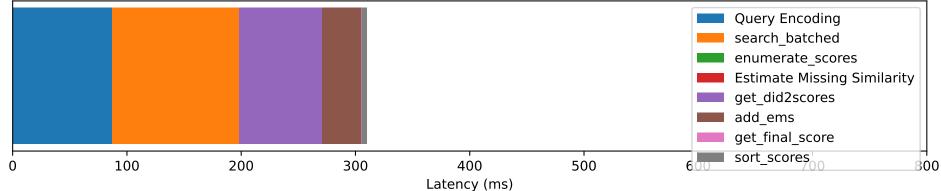
## XTR/ScaNN (document\_top\_k=100, token\_top\_k=40000), BEIR SciFact (iterations=300, total=293.7ms)



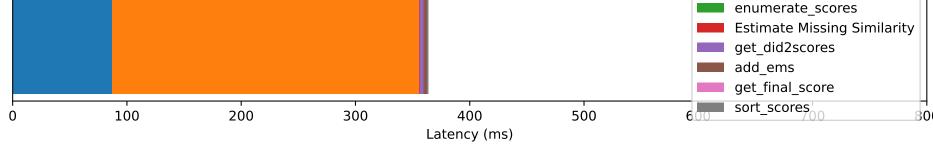
XTR/ScaNN (document\_top\_k=100, token\_top\_k=1000), LoTTE Lifestyle (iterations=661, total=140.4ms)



XTR/ScaNN (document\_top\_k=100, token\_top\_k=40000), LoTTE Lifestyle (iterations=661, total=309.9ms)



XTR/ScaNN (document top k=100, token top k=1000), LoTTE Technology (iterations=596, total=363.9ms) Query Encoding search batched enumerate scores **Estimate Missing Similarity** get did2scores add ems get final score



XTR/ScaNN (document\_top\_k=100, token\_top\_k=40000), LoTTE Technology (iterations=596, total=710.3ms)

