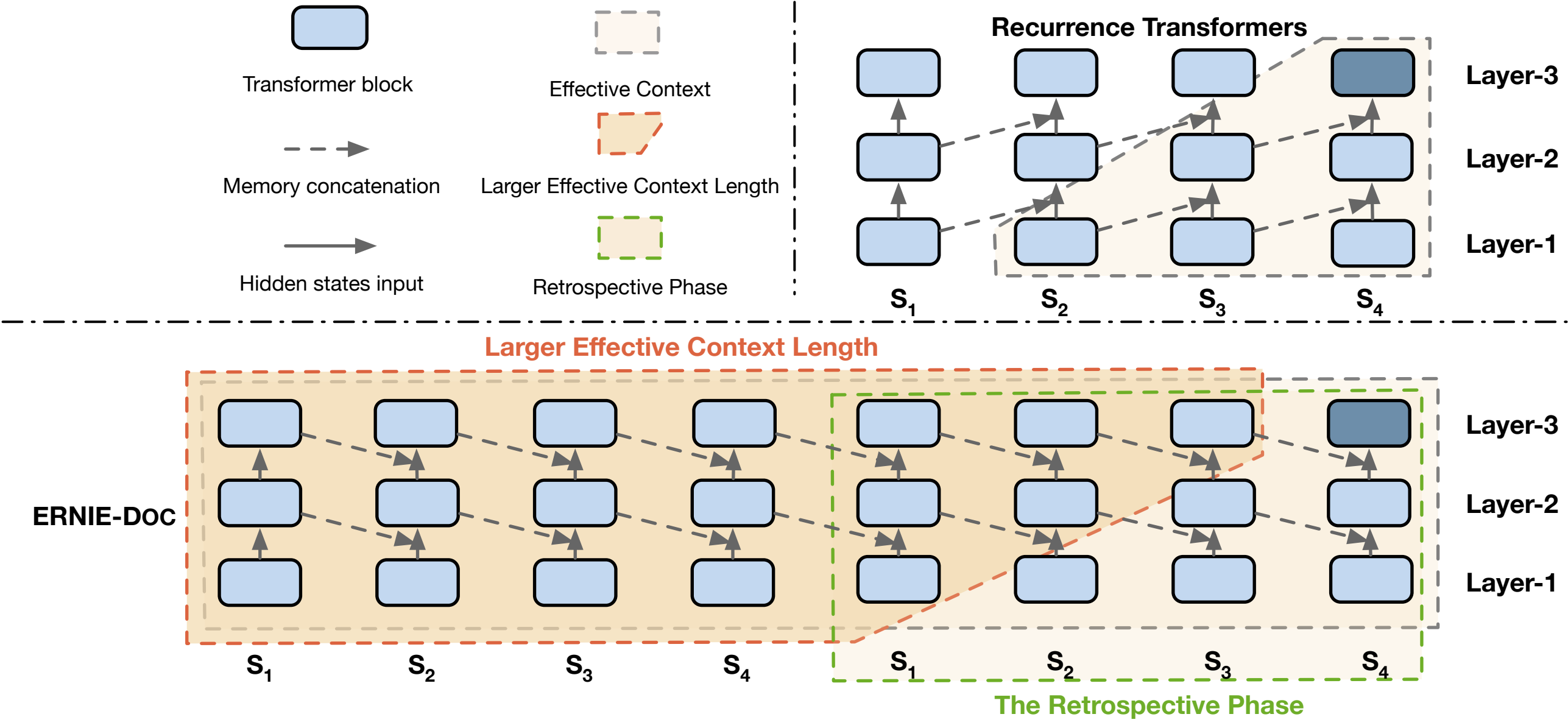
Transformer block
Effective Context
Memory concatenation
Larger Effective Context Length
Hidden states input
Retrospective Phase
Recurrence Transformers
Layer-3
Layer-2
Layer-1
S1
S2
S3
S4
Larger Effective Context Length
ERNIE-DOC
Layer-3
Layer-2
Layer-1
S1
S2
S3
S4
S1
S2
S3
S4
The Retrospective Phase