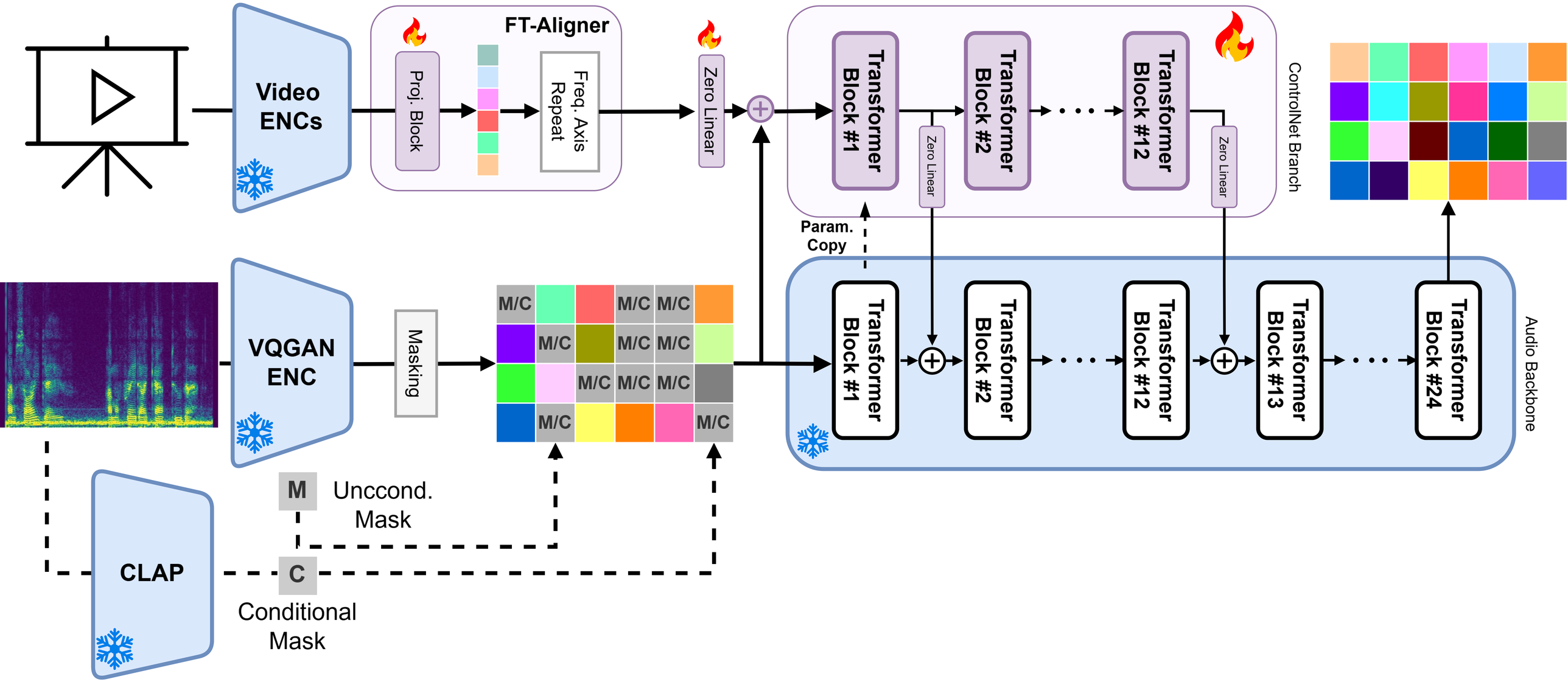
Video ENCs
FT-Aligner
Proj. Block
Freq. Axis Repeat
Zero Linear
Transformer Block #1
Zero Linear
Transformer Block #2
Transformer Block #12
Zero Linear
ControlNet Branch
Param. Copy
VQGAN ENC
Masking
M/C
Transformer Block #1
Transformer Block #2
Transformer Block #12
Transformer Block #13
Transformer Block #24
Audio Backbone
M
Unccond. Mask
CLAP
C
Conditional Mask