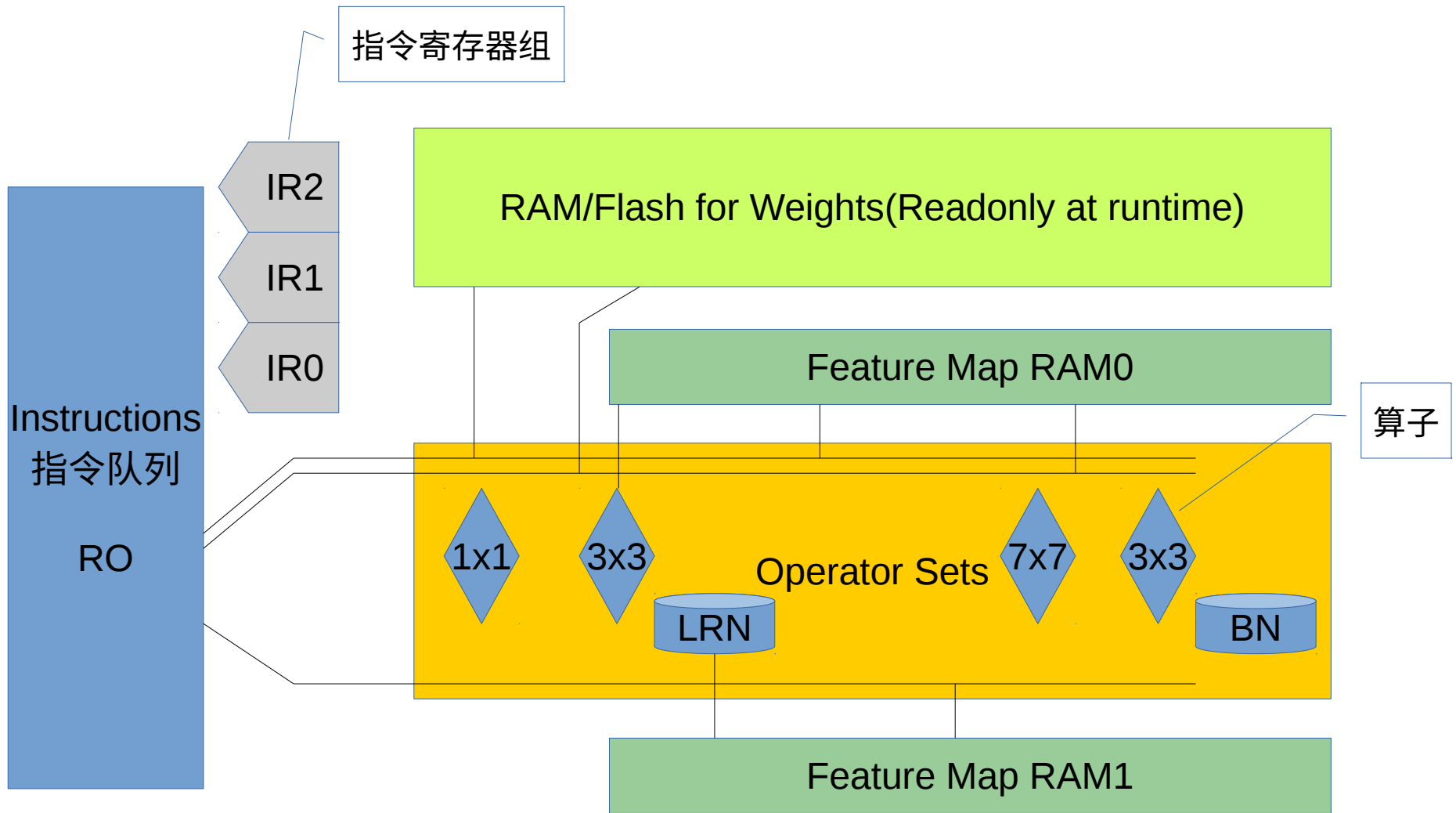
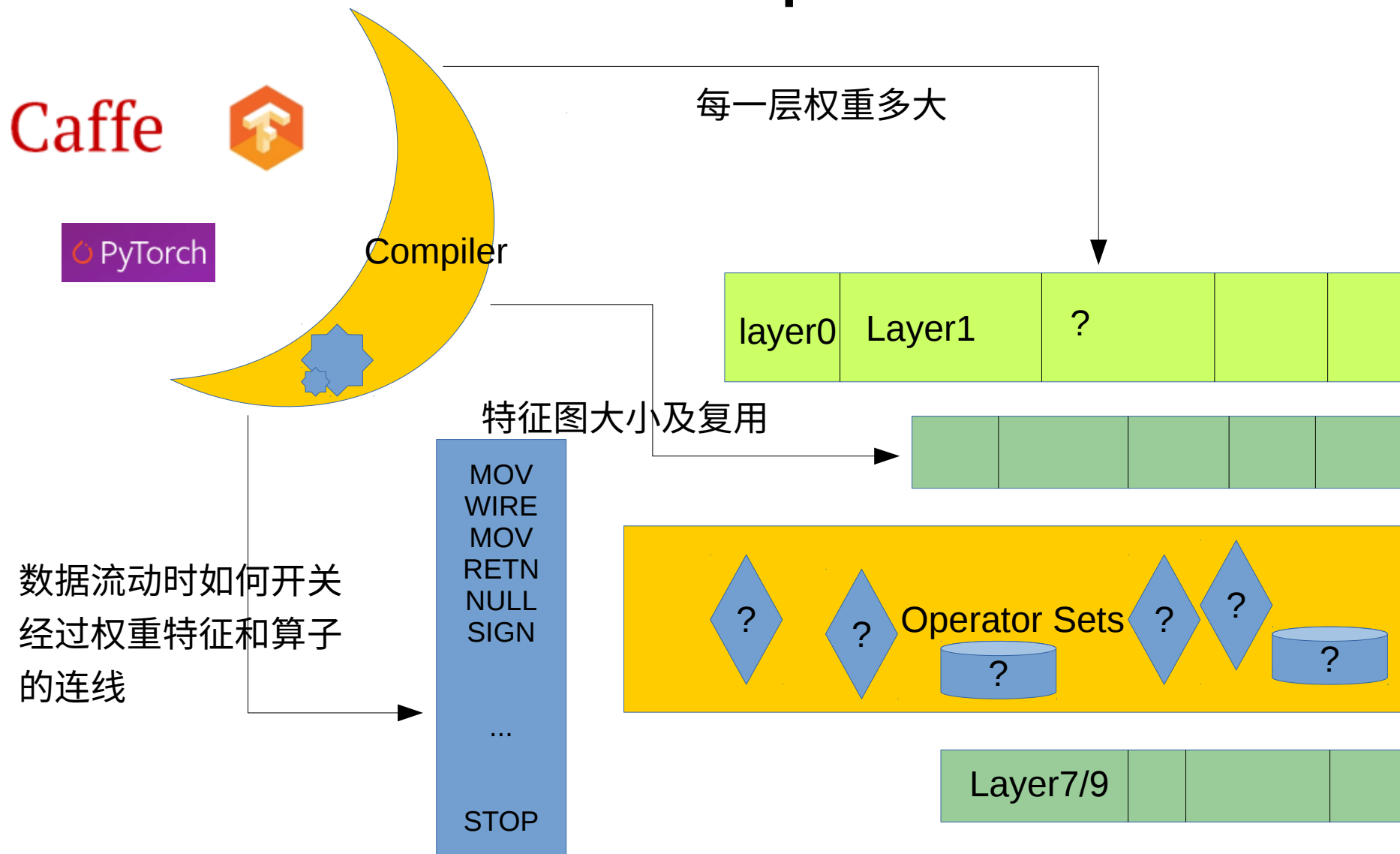


Memory Wall Free Architecture for Neural Network Computing

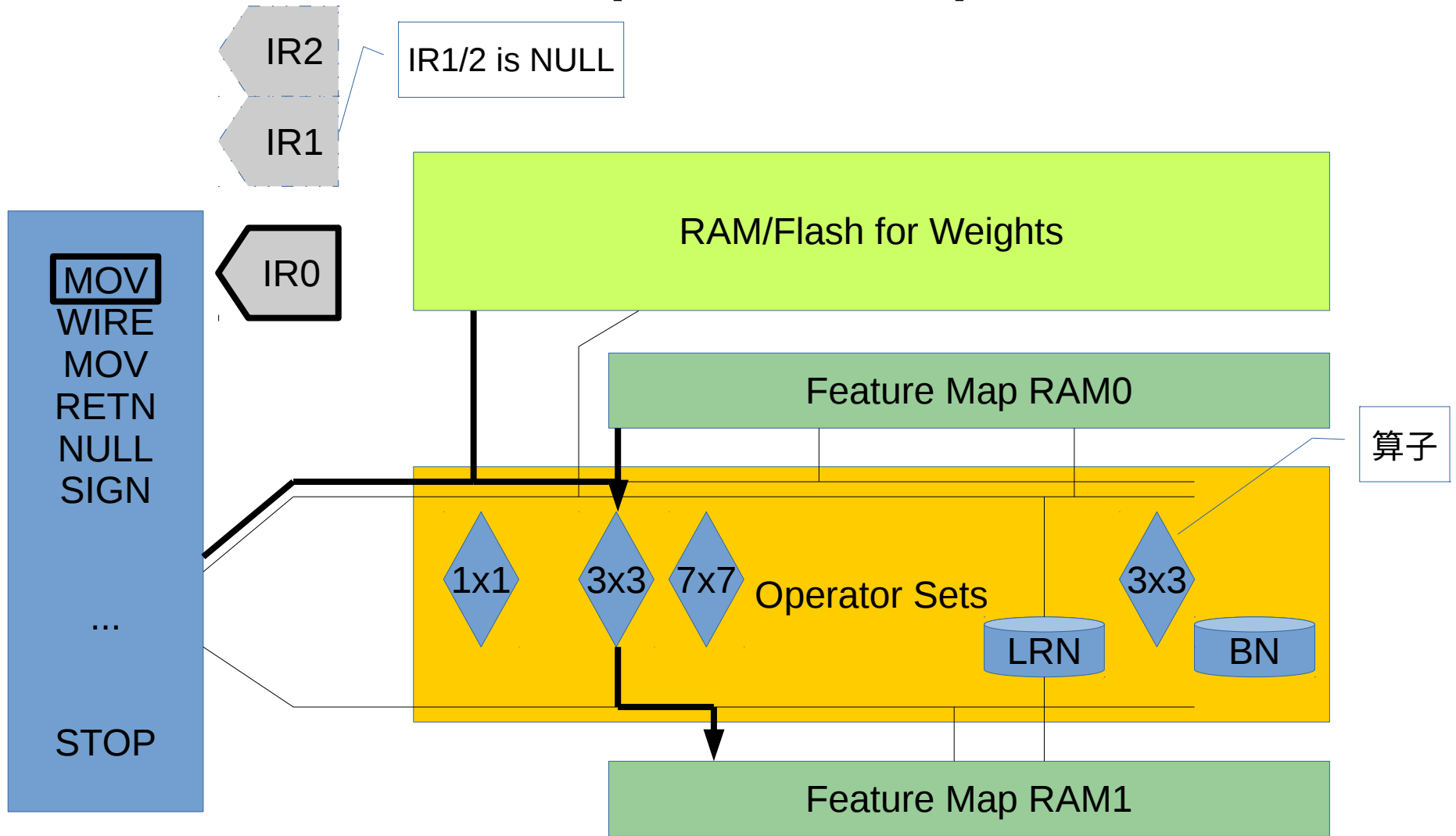


Graph Compute Architecture Compile



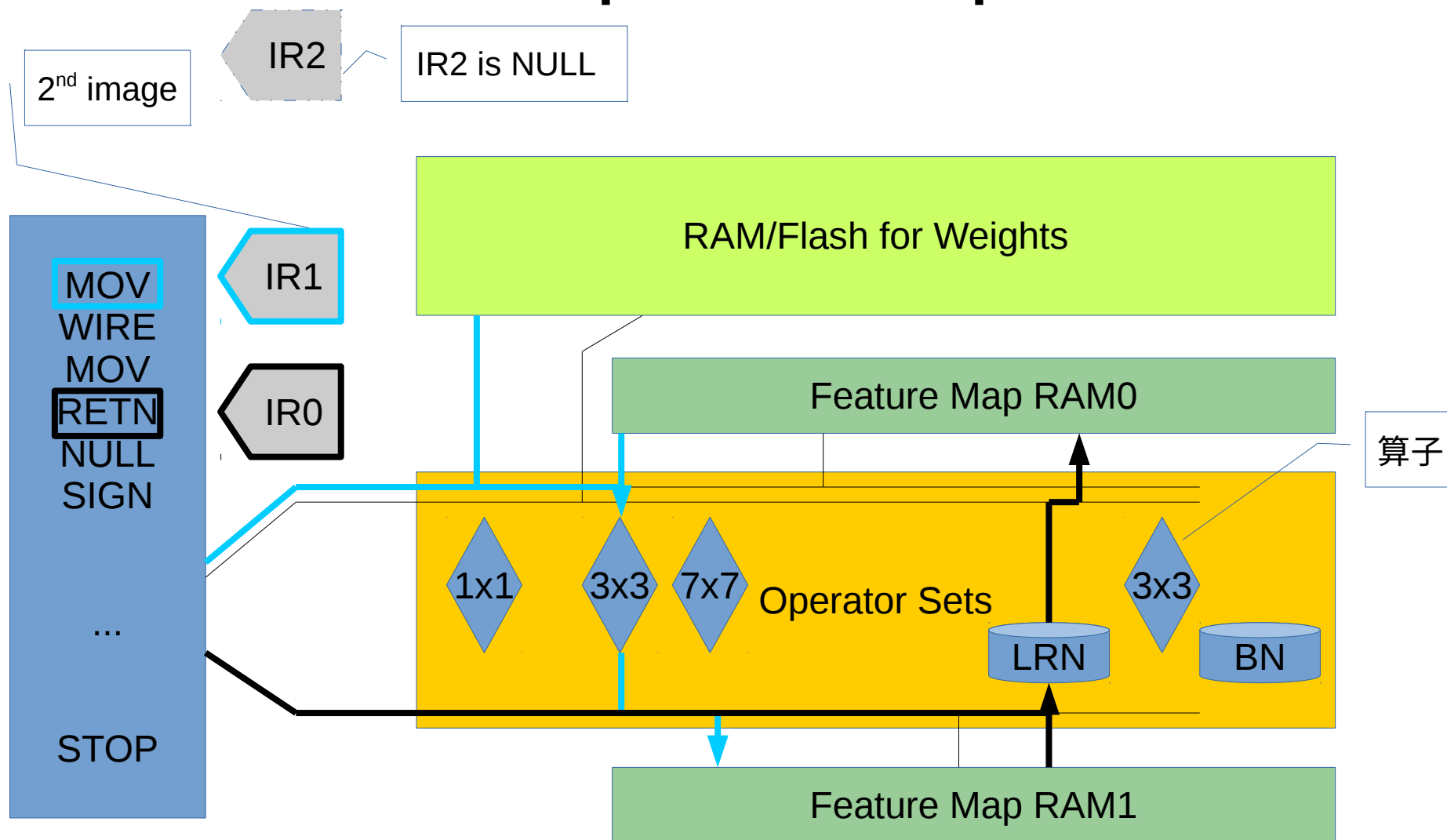
Graph Compute Architecture

Compute Step 1



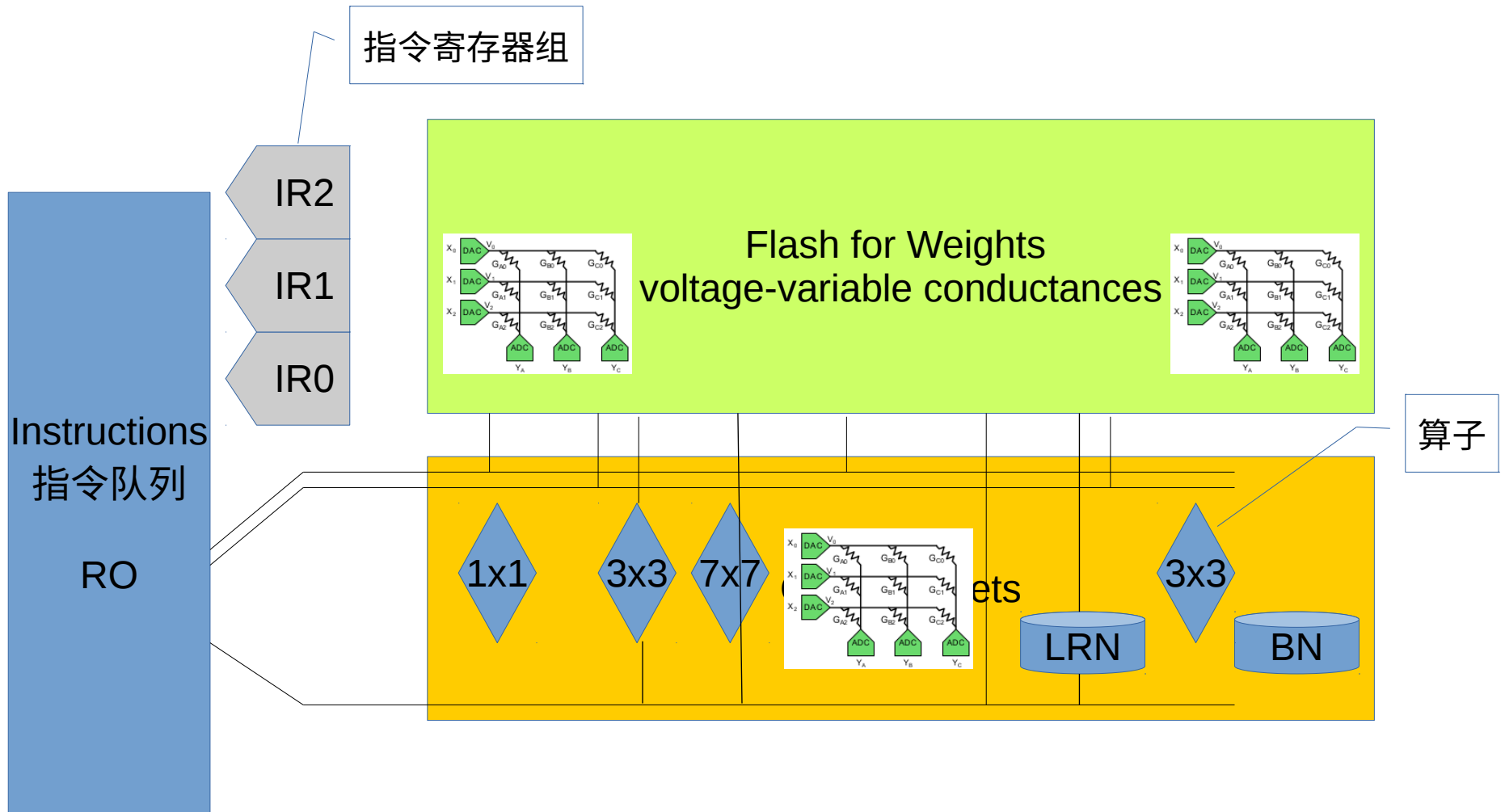
Graph Compute Architecture

Computer Step 2



Graph Compute Architecture

Low Power Consumption



We will deploy NLP CTC-RNN and Attention LSTM

For an input,
like speech



Predict a
sequence of
tokens

h e e ϵ l ϵ l l o o !

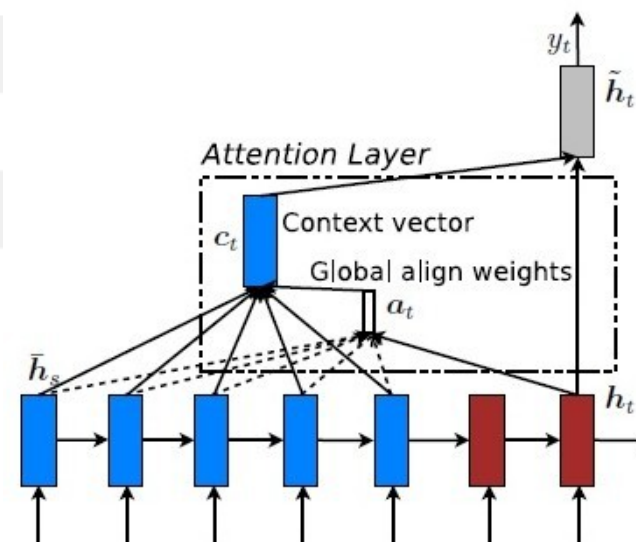
Merge repeats,
drop ϵ

h e l l o !

Final output

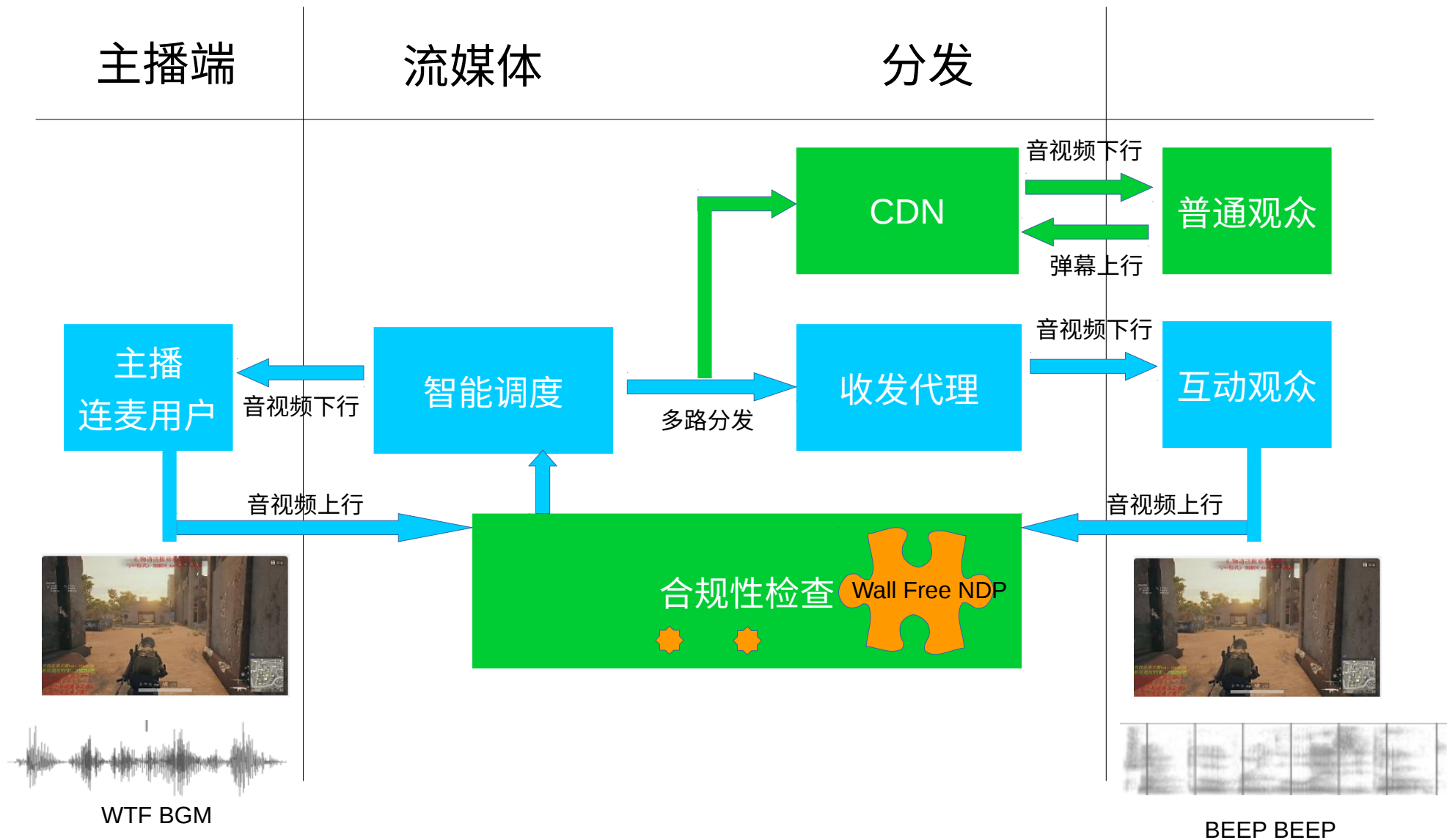
h e l l o !

Both are state-of-the-art algorithm
with sequence data



Cloud User Scenarion

Beep illegal word for live video

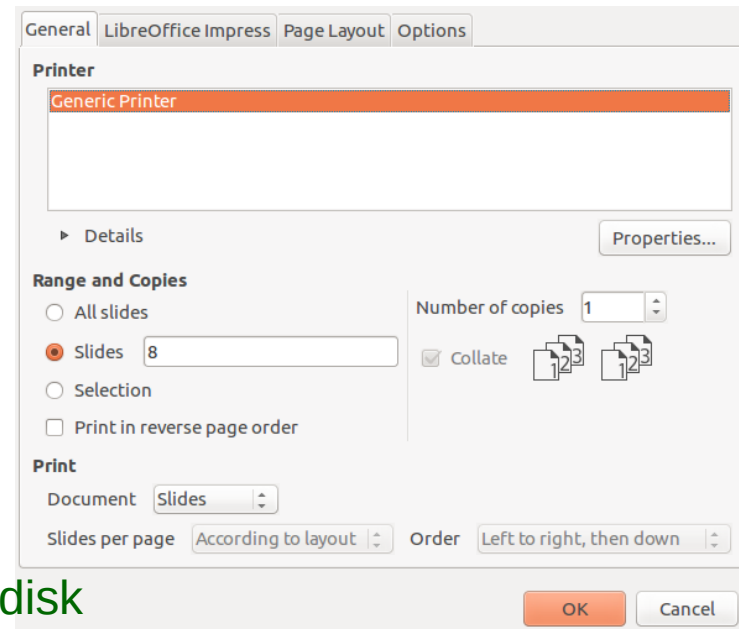


Edge User Scenarion 1

NLP for Office Device



The printer in the office has so many functions:
Print Copy Scan Fax ...
Have you tired with the printer dialog
And the front panel with keyboard & LCD?



How about this:

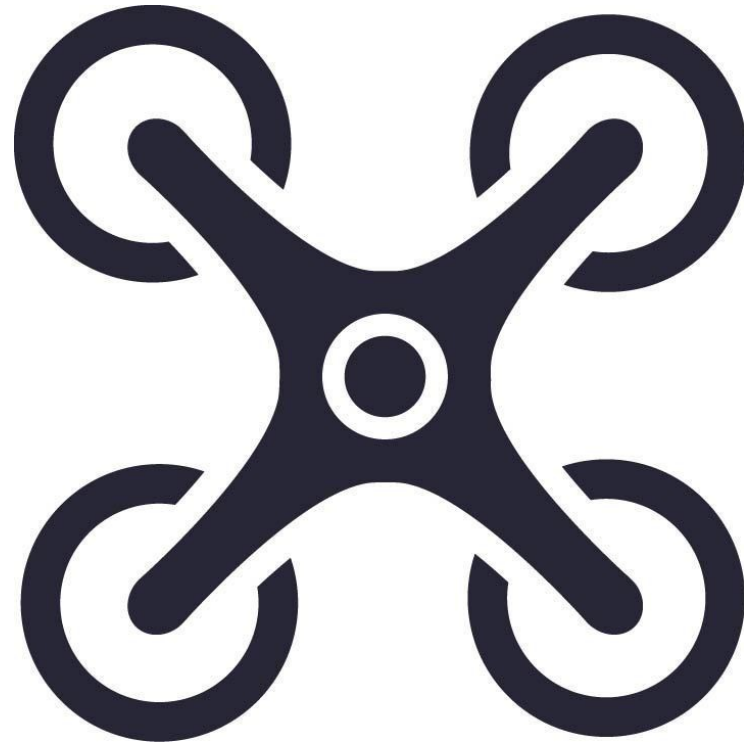
- 1.Scan paper to Udisk
- 2.Fax paper to +852 9327 0934

No need to jump finger with button and you can do other things with you eyes and hands if you can.

Edge User Scenarion 2

NLP and Low Power for Drones

Drone can only fly for 20 minutes
We can deploy NLP with
Low Power Consuming



Edge User Scenarion 3

NLP for Room Device



No need to touch



Because you are busing
Just say: 煮饭

We want do more

Word for cmd and print for Priority



This is the future Robot can help US

