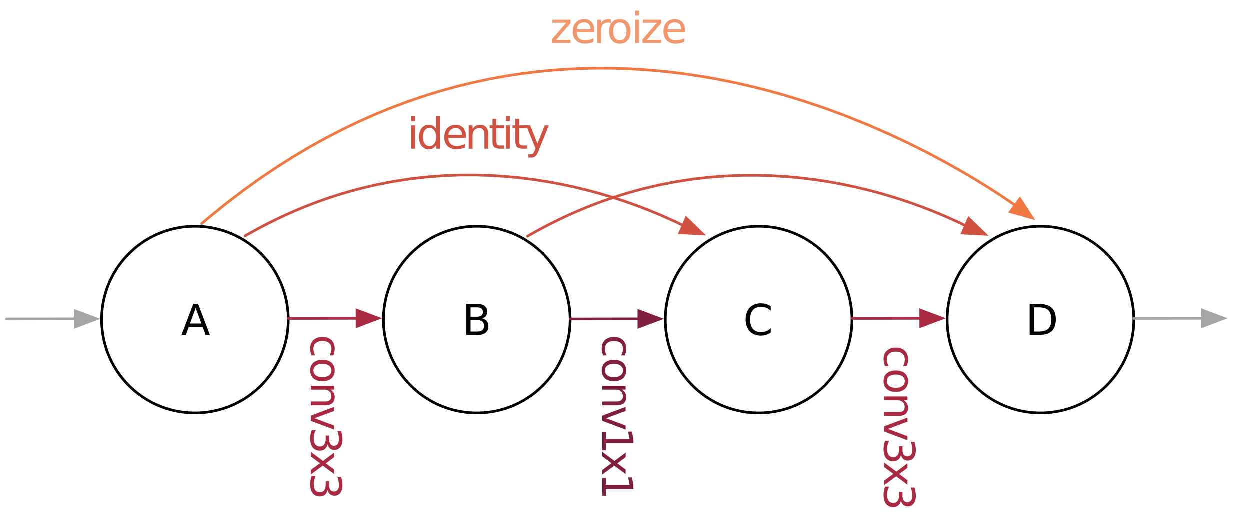
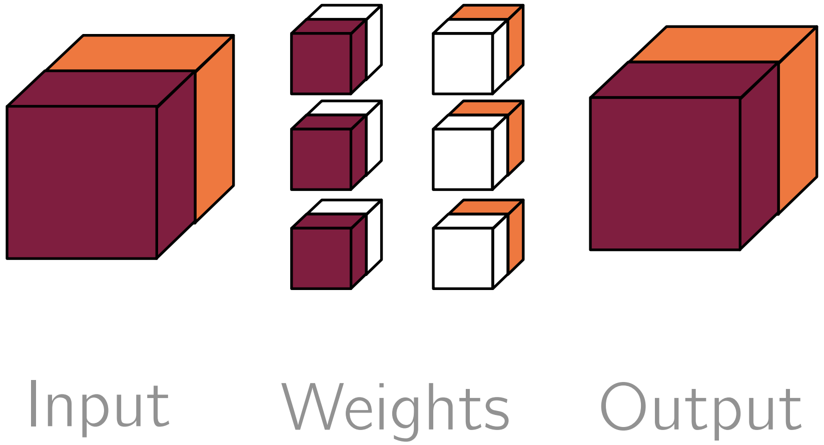
NAS is too coarse-grained



conv3x3
conv1x1
identity
zeroize
pool

Grouped convolutions



conv3x3
gconv(2)
gconv(4)
gconv(N)

Convolution

```
for co in range(CO):               # out_channels
 for ci in range(CI):              # in_channels
  for oh in range(OH):             # out_height
   for ow in range(OW):            # out_width
    for kh in range(KH):           # kernel_height
     for kw in range(KW):          # kernel width
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```

Convolution

```
for co in range(CO):
 for ci in range(CI):
  for oh in range(OH):
   for ow in range(OW):
    for kh in range(KH):
     for kw in range(KW):
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```

Group convolution

```
for g in range(G):
 for co in range(CO//G*g, CO//G*(g+1)):
  for ci in range(CI//G*g, CI//G*(g+1)):
   for oh in range(OH):
    for ow in range(OW):
     for kh in range(KH):
      for kw in range(KW):
        O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```

Generalisation

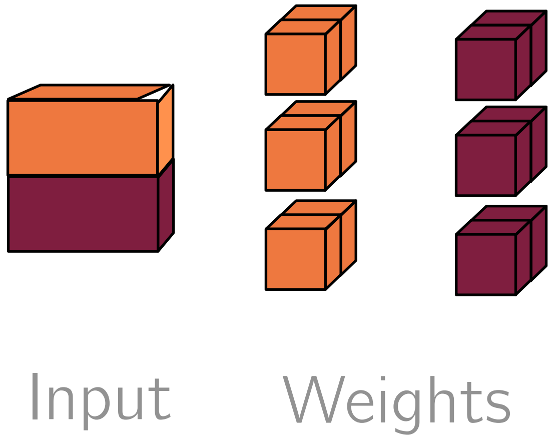
```
for i in range(I):
 for j in range(J):
   # do thing
```

$\xrightarrow{\text{group}(G)}$

```
for g in range(G):
 for i in range(I//G*g, I//G*g+1):
  for j in range(J//G*g, J//G*g+1):
    # do thing
```

Example 1

```
for co in range(CO):     # out_channels
 for oh in range(OH):    # out_height
  for ci in range(CI):   # in_channels
   for ow in range(OW):  # out_width
    for kh in range(KH): # kernel_height
     for kw in range(KW):
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```



Example 2

```
O[0][0][0] += I[0][0+0][0+0] *
                 W[0][0][0][0]
...
O[U][OH][OW]+= I[CI][OH+KH][OW+KW] *
                 W[CO][CI][KH][KW]
```

```
for co in range(U, CO):
 for ci in range(CI):
  for oh in range(OH):
   for ow in range(OW):
    for kh in range(KH):
     for kw in range(KW):
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```

Example 2

```
O[0][0][0] += I[0][0+0][0+0] *
                 W[0][0][0][0]
...
O[U][OH][OW]+= I[CI][OH+KH][OW+KW] *
                 W[CO][CI][KH][KW]

for co in range(U, CO):
 for ci in range(CI):
  for oh in range(OH):
   for ow in range(OW):
    for kh in range(KH):
     for kw in range(KW):
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```

```
O[0][0][0] += I[0][0+0][0+0] *
                 W[0][0][0][0]
...
O[U][OH][OW]+= I[CI][OH+KH][OW+KW] *
                 W[CO][CI][KH][KW]

for g in range(G):
 for co in range(U+CO//G*g, U+CO//G*(g+1)):
  for ci in range(CI//G*g, CI//G*(g+1)):
   for oh in range(OH):
    for ow in range(OW):
     for kh in range(KH):
      for kw in range(KW):
       O[co][oh][ow] +=
              I[ci][oh+kh][ow+kw] *
              W[co][ci][kh][kw]
```