creativecommons.org/licenses/by-nc-nd/4.0/legalcode

Bitmap Index on City (Uncompressed)

BY NC ND
Jens Dittrich

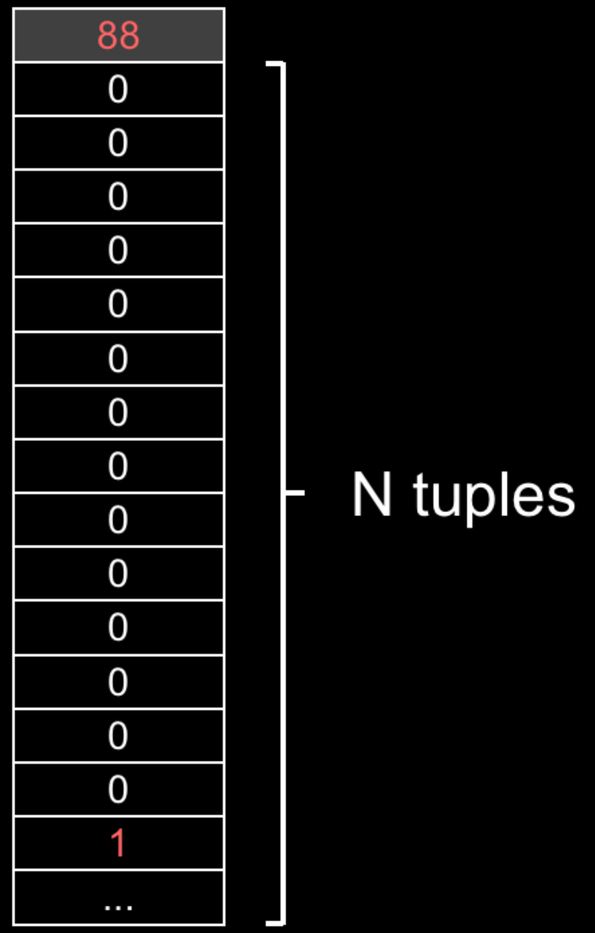
Colleagues		
name	**street**	**city**
peter	unistreet	653
steve	macstreet	42
mike	longstreet	7
tim	unistreet	7
hans	msstreet	653
jens	longstreet	42
frank	unistreet	653
olaf	macstreet	7
stefan	longstreet	7
alekh	unistreet	7
felix	macstreet	653
jorge	longstreet	7
michael	unistreet	9
albert	macstreet	323
volker	longstreet	88
...	...	...

city				
653	42	7	...	88
1	0	0		0
0	1	0		0
0	0	1		0
0	0	1		0
1	0	0		0
0	1	0		0
1	0	0		0
0	0	1		0
0	0	1		0
0	0	1		0
1	0	0		0
0	0	1		0
0	0	0		0
0	0	0		0
0	0	0		1
...	...	...		...

N tuples

D bit lists
= cardinality

Bitmap List for 88 (Uncompressed)

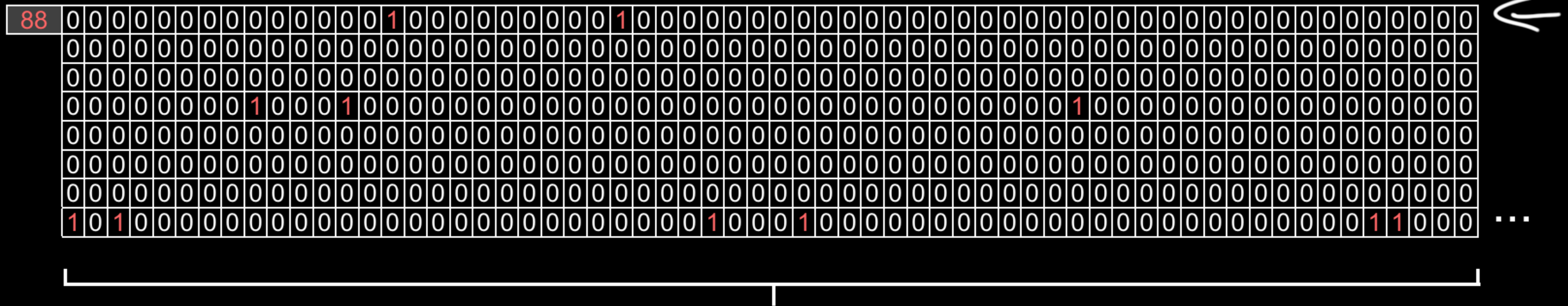


Bitmap List for 88 (Rotated)

Bitmap List for 88 (more data)

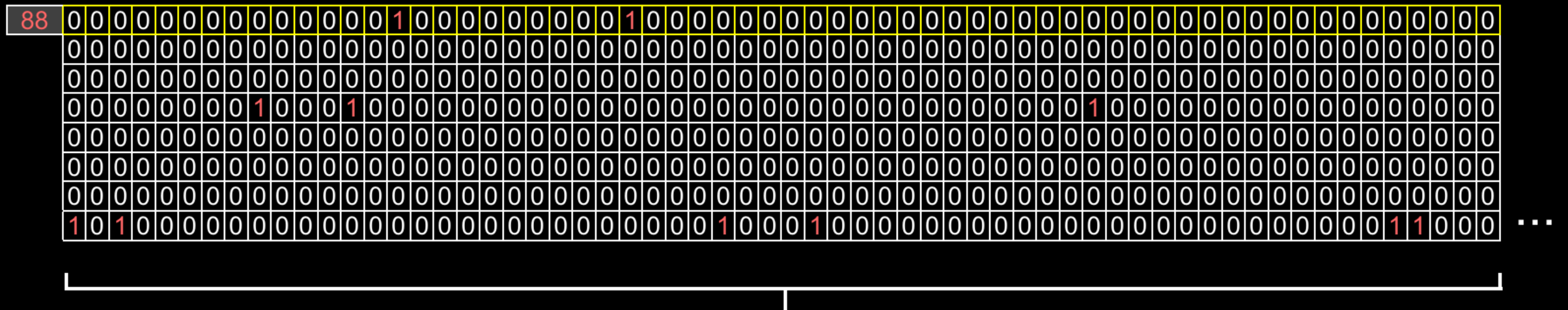
WAH = ~~X~~-Bit encoding + RLE encoding

7 → 31

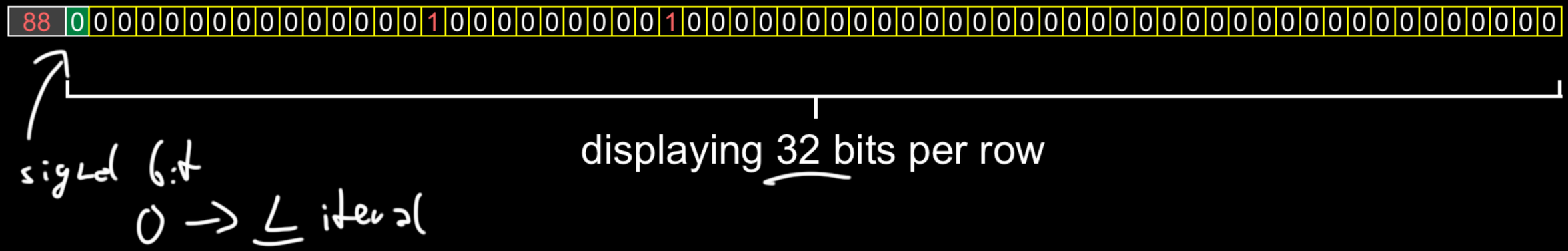


Inspect first 31 Bits

Uncompressed:

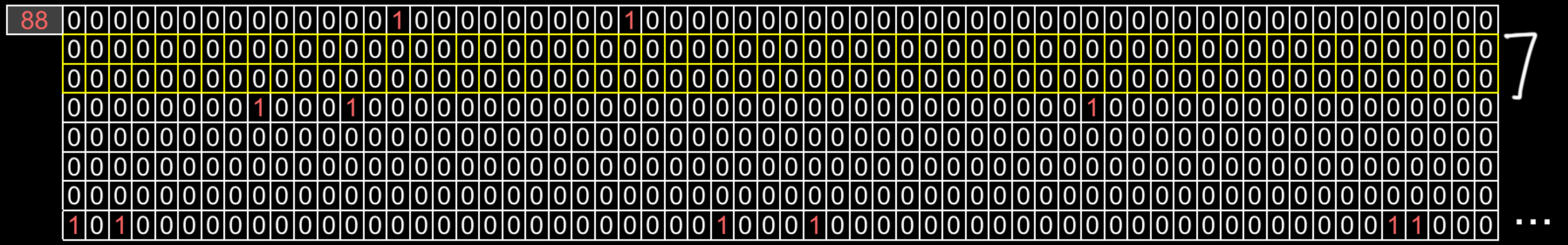


Compressed:

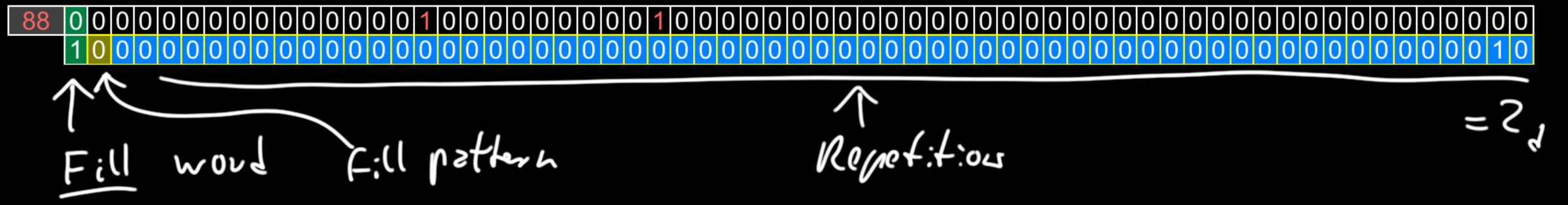


Inspect second and third 31 Bits

Uncompressed:



Compressed:



Inspect fourth 31 Bits

Uncompressed:

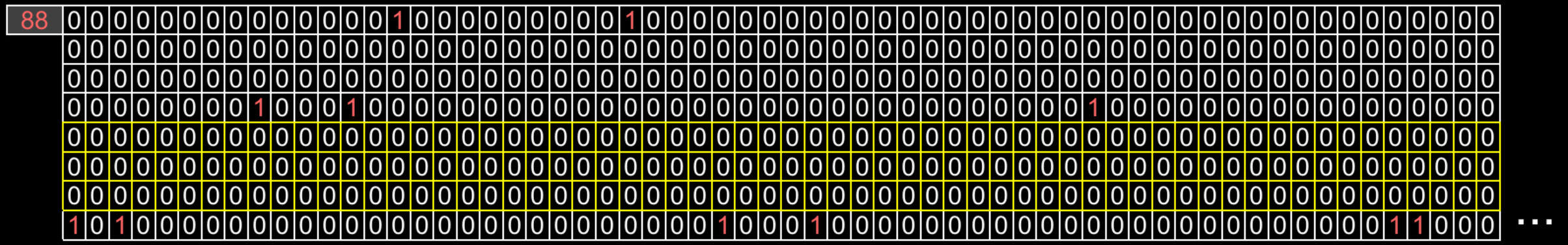
```
88 00000000000000100000000010000000000000000000000000000000000000
   00000000000000000000000000000000000000000000000000000000000000
   00000000000000000000000000000000000000000000000000000000000000
   00000000100010000000000000000000000000000000100000000000000000   <-
   00000000000000000000000000000000000000000000000000000000000000
   00000000000000000000000000000000000000000000000000000000000000
   00000000000000000000000000000000000000000000000000000000000000
   10100000000000000000000000001000100000000000000000000000110000   ...
```

Compressed:

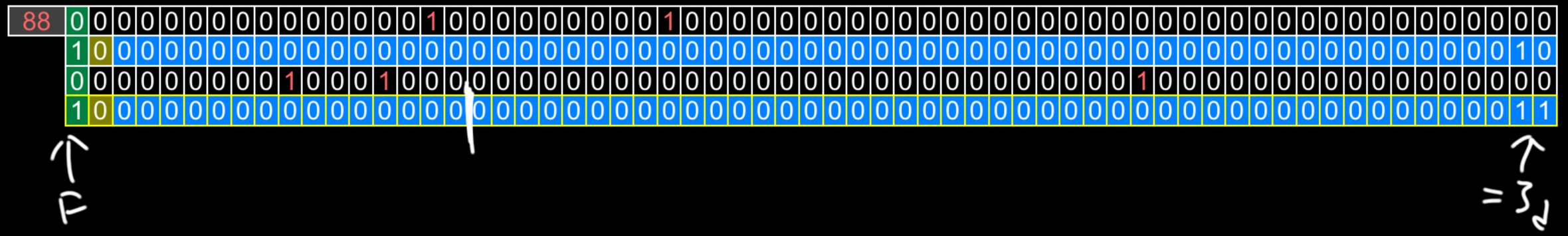
```
88 0 00000000000000100000000010000000000000000000000000000000000000
   1 0 00000000000000000000000000000000000000000000000000000000000010
   0 00000000100010000000000000000000000000000000100000000000000000
   ↑
   =L
```

Inspect fifth, sixth, and seventh 31 Bits

Uncompressed:



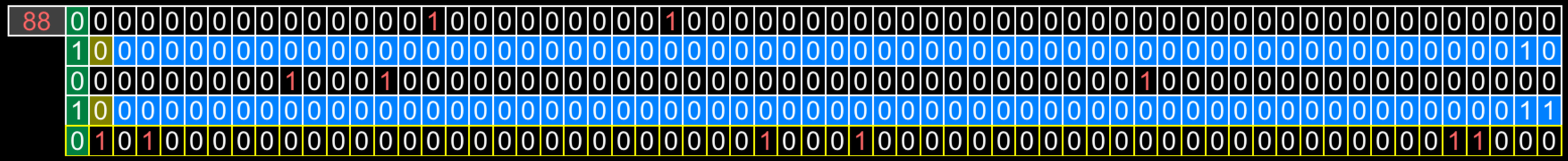
Compressed:



Inspect eighth 31 Bits

Uncompressed:

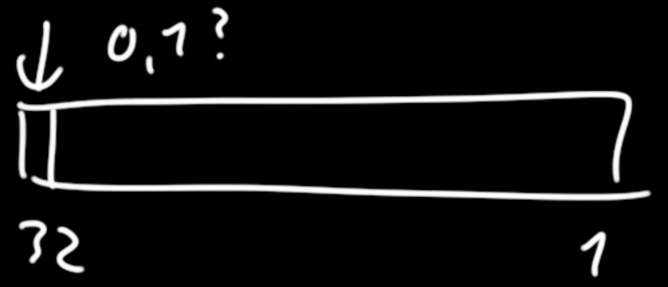
Compressed:



Decompression

32 bit - chunks

Decompression



first bit of a word (i.e. bit 32):

0 ⇒ **literal word L**

take remaining 31 bits as is

1 ⇒ **fill word**

Decompression

first bit of a word (i.e. bit 32):

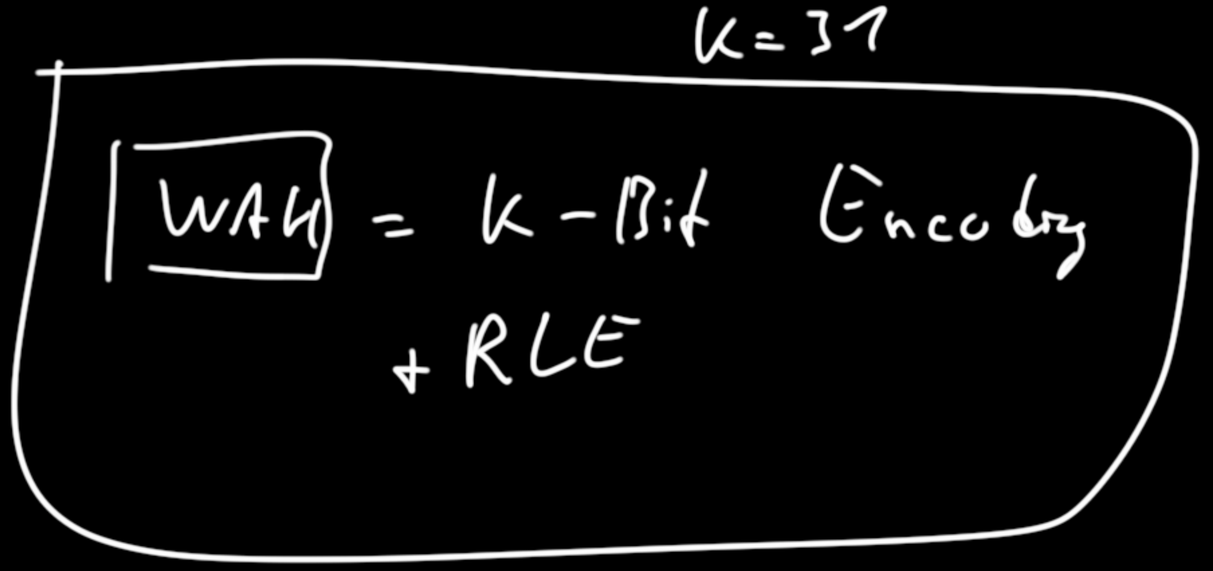
0 ⇒ **literal word L**

take remaining 31 bits as is

1 ⇒ **fill word**

fill pattern := second bit (i.e. bit 31)

Decompression



first bit of a word (i.e. bit 32):

0 ⇒ **literal word L**

take remaining 31 bits as is

1 ⇒ **fill word**

fill pattern := second bit (i.e. bit 31)

#repetitions := remaining 30 bits

31-bit pattern