



Exploring the impact of Bit Rot

In an effort to understand the direct impact of “bit rot” on a file, and to inform potential research streams an image file was “attacked” by a script that sequentially flipped every bit, and tested the resulting image for (1) renderability and (2) similarity to the source image.

Jay Gattuso

Preservation, Research and Consultancy
National Library of New Zealand Te Puna Mātauranga o Aotearoa

Keywords:

Jpeg

Bit-Rot

RMSe

Render Test

Contents

Summary	2
Method	2
Example:	2
Comments.....	3
Original Image:	4
Specific examples of damage types:.....	4
Comb:.....	4
Detail:.....	4
Block:	4
Colour/ tone:.....	5
Catastrophic	5
Dimension Bytes:.....	6
Catalogue of Errored images:	7

Exploring the impact of Bit Rot

Summary

In an effort to understand the direct impact of “bit rot” on a file, and to inform potential research streams an image file was “attacked” by a script that sequentially flipped every bit, and tested the resulting image for (1) renderability and (2) similarity to the source image.

Method

A python script was written that loops over a bitstream, locates each bit in order, inverts the bit, and saves the bitstream as a new image.

Each image is presented to imagemagik, which attempts to open the file.

If the file successfully opens, a root mean squared error (RMSe) value is calculated as a measure of difference between the original image and the new (bit flipped) image.

Example:

The following bitstream is presented to the script:

```
1111 0000 1010 0000
```

This results in the following bitstreams being created:

```
0111 0000 1010 0000
1011 0000 1010 0000
1101 0000 1010 0000
1110 0000 1010 0000
1111 1000 1010 0000
1111 0100 1010 0000
1111 0010 1010 0000
1111 0001 1010 0000
1111 0000 0010 0000
1111 0000 1110 0000
1111 0000 1000 0000
1111 0000 1011 0000
1111 0000 1010 1000
1111 0000 1010 0100
1111 0000 1010 0010
1111 0000 1010 0001
```

Comments

A demonstration test image was selected that contains a number of image features that are of interest to image research (e.g. gradual colour change / luma change, high frequency detail, blocks of solid colour etc.) and would specifically highlight errors in a jpeg compressed image. This source image was reduced to 180 x 120 pixels in size, which results in a 17,514 byte image. This is equal to 140,112 bits of data per image, which results in 140,112 new images being created. These 140,112 consume some 2.3 GB of storage.

This process is extremely resource intensive - On an i5 64-bit Ubuntu 12.04LTS install, with 16 Gb of RAM, writing to a 120Gb SSD HDD, the demonstration image run through took approximately 36 hours with ~60% of the processor time being consumed by the RMSe calculation.

As this test was undertaken as a proof of concept, very little analysis has been undertaken on the resulting images – there are some notable groups of error type, and some examples have been provided in this report. The results are provided for the interest of the reader.

Raw test data is available on request (filename, rendered? [Y|N], RMSe) for this test and some other jpg images used to test the workflow. Some resultant jpg files are available.

Exploring the RMSe values in the main result data show that most significant bits (MSB) being flipped frequently produce larger RMSe errors than the least most significant bits (LSB) from the same byte being flipped.

It is expected that this process could be (1) more efficiently coded, and (2) rolled out with other image types and further to other file types (the main factor would be renderance metrics, however, anything that is visually or sonically rendered RMSe remains a viable base measure).

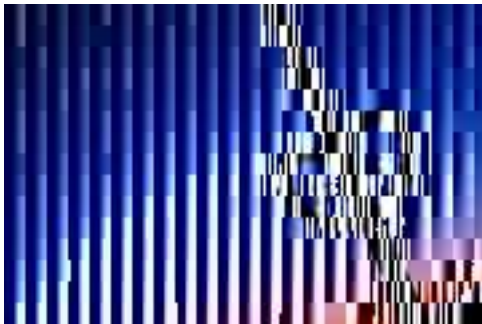
The next stage for the file type comparison would be to undertake the test on multiple files of the same type, and explore the resulting data looking for patterns in file structure that are particularly at risk. It is expected that different file types would have different bitstream locations that cause catastrophic file errors (for example, any bit flip attacks on the first or last 4 bytes of a PDF will result in a render fail as it damages the BOF or EOF marker the PDF renderer looks for).

Original Image:



Specific examples of damage types:

Comb:



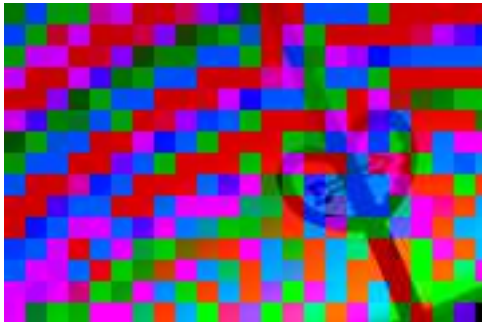
Detail:



Block:



Colour/ tone:



Catastrophic

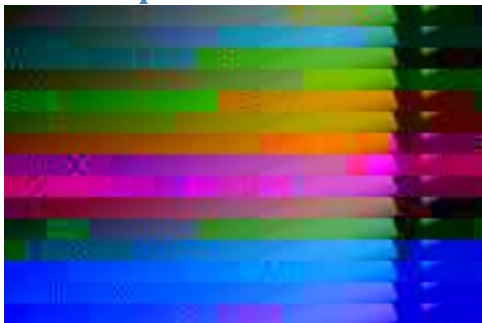


Image Dimensions:



Filename	W (pix)	H (pix)
163_0.jpg	180	32888
163_1.jpg	180	16504
163_2.jpg	180	8312
163_3.jpg	180	4216
163_4.jpg	180	2168
163_5.jpg	180	1144
163_6.jpg	180	632
163_7.jpg	180	376
164_0.jpg	180	248
164_1.jpg	180	56
164_2.jpg	180	88
164_3.jpg	180	104
164_4.jpg	180	112
164_5.jpg	180	124
164_6.jpg	180	122
164_7.jpg	180	121
163_0.jpg	32948	120
163_1.jpg	16564	120
163_2.jpg	8372	120
163_3.jpg	4276	120
163_4.jpg	2228	120
163_5.jpg	1204	120
163_6.jpg	692	120
163_7.jpg	436	120
164_0.jpg	52	120
164_1.jpg	244	120
164_2.jpg	148	120
164_3.jpg	164	120
164_4.jpg	188	120
164_5.jpg	176	120
164_6.jpg	182	120
164_7.jpg	181	120

This table shows the image dimensions for the 32 images that report a different image pixel size to the source image.

Catalogue of Errored images:

N.B The position of the flipped bit can be established from the file name:

XXX_N.jpg,

Where XXX is the byte location offset (decimal) in the bitstream

And N is the bit position in the stated byte of the flipped bit.

e.g. if the byte offset is 1,440

wellington.jpeg	0	1	2	3
1,408	11111000	11100011	10111001	01100000
1,424	10001110	11110011	00010011	11011001
1,440	00111011	11001100	01001011	10100000
1,456	00100001	11000101	01101110	10110001
1,472	11000011	00111100	10010100	01001101
1,488	01101111	00000001	00101100	00100100
1,504	10110110	11001111	10011001	01110101
1,520	11110101	11101011	11000001	10111100
1,536	10011001	10000001	11011110	11101111
1,552	00110111	10011000	01100111	11011111

And the 'bit to be flipped' is the first '1', the above byte would become:

wellington.jpeg*	0	1	2	3
1,408	11111000	11100011	10111001	01100000
1,424	10001110	11110011	00010011	11011001
1,440	10111011	11001100	01001011	10100000
1,456	00100001	11000101	01101110	10110001
1,472	11000011	00111100	10010100	01001101
1,488	01101111	00000001	00101100	00100100
1,504	10110110	11001111	10011001	01110101
1,520	11110101	11101011	11000001	10111100
1,536	10011001	10000001	11011110	11101111
1,552	00110111	10011000	01100111	11011111

And the filename would become 1440_0.jpg





166_6.jpeg



169_6.jpeg



175_6.jpeg



198_7.jpeg



200_4.jpeg



200_7.jpeg



201_6.jpeg



202_6.jpeg



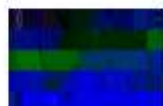
204_6.jpeg



205_6.jpeg



166_7.jpeg



172_6.jpeg



198_4.jpeg



199_4.jpeg



200_5.jpeg



201_4.jpeg



202_4.jpeg



204_4.jpeg



205_4.jpeg



206_4.jpeg



169_2.jpeg



175_2.jpeg



198_6.jpeg



199_5.jpeg



200_6.jpeg



201_5.jpeg



202_5.jpeg



204_5.jpeg



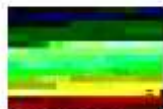
205_5.jpeg



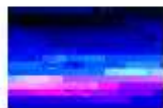
206_5.jpeg



206_6.jpeg



230_0.jpeg



230_3.jpeg



230_6.jpeg



231_1.jpeg



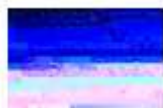
231_4.jpeg



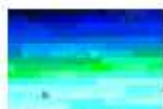
231_7.jpeg



232_2.jpeg



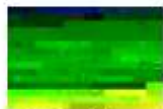
232_5.jpeg



233_0.jpeg



206_7.jpeg



230_1.jpeg



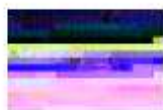
230_4.jpeg



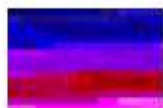
230_7.jpeg



231_2.jpeg



231_5.jpeg



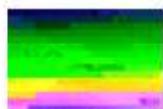
232_0.jpeg



232_3.jpeg



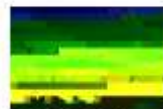
232_6.jpeg



233_1.jpeg



208_4.jpeg



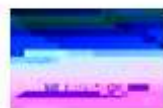
230_2.jpeg



230_5.jpeg



231_0.jpeg



231_3.jpeg



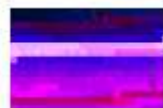
231_6.jpeg



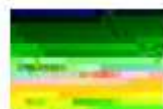
232_1.jpeg



232_4.jpeg



232_7.jpeg



233_2.jpeg



233_3.jpeg



233_6.jpeg



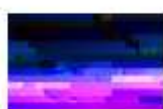
234_1.jpeg



234_4.jpeg



234_7.jpeg



235_2.jpeg



235_5.jpeg



236_0.jpeg



236_3.jpeg



236_6.jpeg



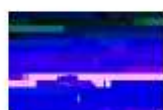
233_4.jpeg



233_7.jpeg



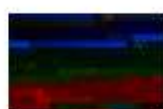
234_2.jpeg



234_5.jpeg



235_0.jpeg



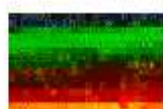
235_3.jpeg



235_6.jpeg



236_1.jpeg



236_4.jpeg



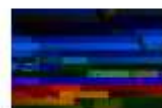
236_7.jpeg



233_5.jpeg



234_0.jpeg



234_3.jpeg



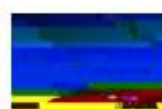
234_6.jpeg



235_1.jpeg



235_4.jpeg



235_7.jpeg



236_2.jpeg



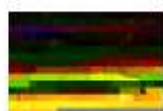
236_5.jpeg



237_4.jpeg



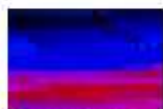
237_5.jpeg



237_6.jpeg



237_7.jpeg



238_0.jpeg



238_1.jpeg



238_2.jpeg



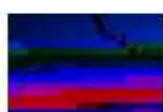
238_3.jpeg



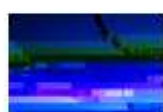
238_4.jpeg



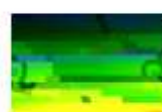
238_5.jpeg



238_6.jpeg



238_7.jpeg



239_0.jpeg



239_1.jpeg



239_4.jpeg



239_5.jpeg



239_6.jpeg



239_7.jpeg



240_0.jpeg



240_1.jpeg



240_2.jpeg



240_3.jpeg



240_4.jpeg



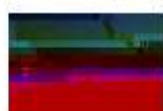
240_5.jpeg



240_6.jpeg



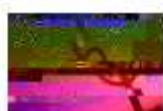
240_7.jpeg



241_0.jpeg



241_2.jpeg



241_4.jpeg



241_5.jpeg



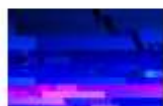
241_6.jpeg



241_7.jpeg



242_4.jpeg



242_7.jpeg



243_4.jpeg



243_7.jpeg



244_4.jpeg



245_3.jpeg



245_7.jpeg



249_1.jpeg



25_0.jpeg



242_2.jpeg



242_5.jpeg



243_0.jpeg



243_5.jpeg



244_0.jpeg



244_5.jpeg



245_5.jpeg



246_0.jpeg



249_4.jpeg



25_1.jpeg



242_3.jpeg



242_6.jpeg



243_2.jpeg



243_6.jpeg



244_1.jpeg



245_2.jpeg



245_6.jpeg



247_6.jpeg



249_6.jpeg



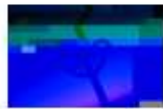
25_2.jpeg



25_3.jpeg



25_6.jpeg



250_5.jpeg



251_0.jpeg



26_1.jpeg



26_4.jpeg



26_7.jpeg



268_7.jpeg



27_2.jpeg



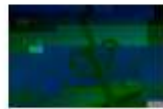
27_5.jpeg



25_4.jpeg



25_7.jpeg



250_6.jpeg



251_1.jpeg



26_2.jpeg



26_5.jpeg



266_5.jpeg



27_0.jpeg



27_3.jpeg



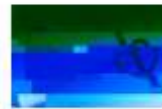
27_6.jpeg



25_5.jpeg



250_4.jpeg



250_7.jpeg



26_0.jpeg



26_3.jpeg



26_6.jpeg



268_0.jpeg



27_1.jpeg



27_4.jpeg



27_7.jpeg



270_0.jpeg



270_6.jpeg



273_5.jpeg



28_0.jpeg



28_1.jpeg



28_2.jpeg



28_3.jpeg



28_4.jpeg



28_5.jpeg



28_6.jpeg



28_7.jpeg



29_0.jpeg



29_1.jpeg



29_2.jpeg



29_3.jpeg



29_4.jpeg



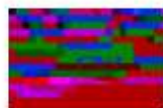
29_5.jpeg



29_6.jpeg



29_7.jpeg



295_4.jpeg



295_5.jpeg



295_6.jpeg



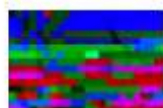
295_7.jpeg



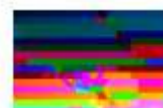
296_4.jpeg



296_5.jpeg



297_4.jpeg



298_4.jpeg



298_5.jpeg



298_6.jpeg



298_7.jpeg



299_5.jpeg



299_6.jpeg



299_7.jpeg



30_0.jpeg



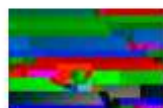
30_1.jpeg



30_2.jpeg



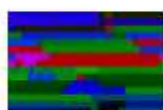
30_3.jpeg



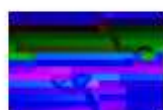
300_4.jpeg



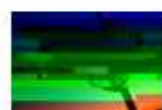
301_4.jpeg



301_5.jpeg



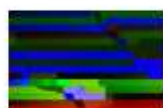
301_6.jpeg



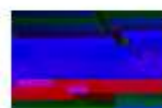
301_7.jpeg



302_4.jpeg



302_5.jpeg



302_6.jpeg



303_4.jpeg



303_5.jpeg



303_6.jpeg



303_7.jpeg



304_5.jpeg



304_7.jpeg



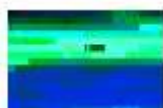
305_4.jpeg



305_5.jpeg



305_6.jpeg



327_0.jpeg



327_1.jpeg



327_2.jpeg



327_3.jpeg



327_4.jpeg



327_5.jpeg



327_6.jpeg



327_7.jpeg



328_0.jpeg



328_1.jpeg



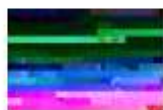
328_2.jpeg



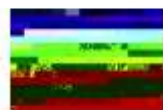
328_3.jpeg



328_4.jpeg



328_5.jpeg



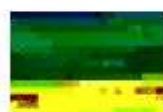
328_6.jpeg



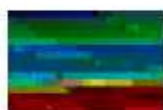
328_7.jpeg



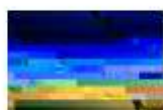
329_0.jpeg



329_1.jpeg



329_2.jpeg



329_3.jpeg



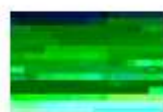
329_4.jpeg



329_5.jpeg



329_6.jpeg



329_7.jpeg



330_0.jpeg



330_1.jpeg



330_2.jpeg



330_3.jpeg



330_4.jpeg



330_5.jpeg



330_6.jpeg



330_7.jpeg



331_0.jpeg



331_1.jpeg



331_2.jpeg



331_3.jpeg



335_2.jpeg



335_5.jpeg



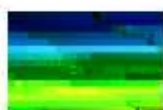
336_0.jpeg



336_3.jpeg



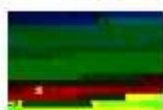
336_6.jpeg



337_1.jpeg



337_4.jpeg



337_7.jpeg



338_2.jpeg



338_5.jpeg



335_3.jpeg



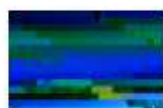
335_6.jpeg



336_1.jpeg



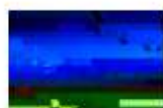
336_4.jpeg



336_7.jpeg



337_2.jpeg



337_5.jpeg



338_0.jpeg



338_3.jpeg



338_6.jpeg



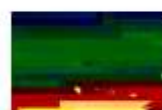
335_4.jpeg



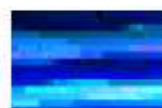
335_7.jpeg



336_2.jpeg



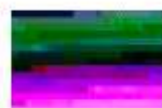
336_5.jpeg



337_0.jpeg



337_3.jpeg



337_6.jpeg



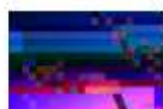
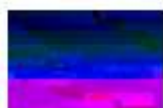
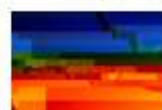
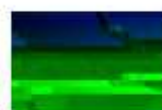
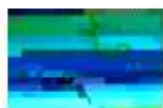
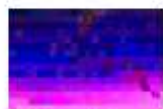
338_1.jpeg



338_4.jpeg



338_7.jpeg

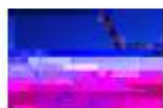




342_6.jpeg



343_1.jpeg



343_4.jpeg



343_7.jpeg



344_2.jpeg



344_5.jpeg



345_0.jpeg



345_3.jpeg



345_6.jpeg



346_1.jpeg



342_7.jpeg



343_2.jpeg



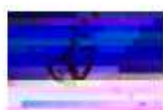
343_5.jpeg



344_0.jpeg



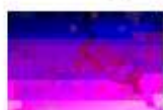
344_3.jpeg



344_6.jpeg



345_1.jpeg



345_4.jpeg



345_7.jpeg



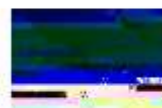
346_2.jpeg



343_0.jpeg



343_3.jpeg



343_6.jpeg



344_1.jpeg



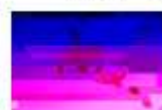
344_4.jpeg



344_7.jpeg



345_2.jpeg



345_5.jpeg



346_0.jpeg



346_3.jpeg



346_4.jpeg



346_5.jpeg



346_6.jpeg



346_7.jpeg



347_0.jpeg



347_1.jpeg



347_2.jpeg



347_3.jpeg



347_4.jpeg



347_5.jpeg



347_6.jpeg



347_7.jpeg



348_0.jpeg



348_1.jpeg



348_2.jpeg



348_3.jpeg



348_4.jpeg



348_5.jpeg



348_6.jpeg



348_7.jpeg



349_0.jpeg



349_1.jpeg



349_2.jpeg



349_3.jpeg



349_4.jpeg



349_5.jpeg



349_6.jpeg



349_7.jpeg



350_0.jpeg



350_1.jpeg



350_2.jpeg



350_3.jpeg



350_4.jpeg



350_5.jpeg



350_6.jpeg



350_7.jpeg



351_0.jpeg



351_1.jpeg



351_2.jpeg



351_3.jpeg



351_4.jpeg



351_5.jpeg



351_6.jpeg



351_7.jpeg



352_0.jpeg



352_1.jpeg



352_2.jpeg



352_3.jpeg



352_4.jpeg



352_5.jpeg



352_6.jpeg



352_7.jpeg



353_0.jpeg



353_1.jpeg



353_2.jpeg



353_3.jpeg



353_4.jpeg



353_5.jpeg



353_6.jpeg



353_7.jpeg



354_0.jpeg



354_3.jpeg



354_6.jpeg



355_1.jpeg



355_4.jpeg



355_7.jpeg



356_2.jpeg



356_5.jpeg



357_0.jpeg



357_3.jpeg



354_1.jpeg



354_4.jpeg



354_7.jpeg



355_2.jpeg



355_5.jpeg



356_0.jpeg



356_3.jpeg



356_6.jpeg



357_1.jpeg



357_4.jpeg



354_2.jpeg



354_5.jpeg



355_0.jpeg



355_3.jpeg



355_6.jpeg



356_1.jpeg



356_4.jpeg



356_7.jpeg



357_2.jpeg



357_5.jpeg



357_6.jpeg



357_7.jpeg



358_0.jpeg



358_1.jpeg



358_2.jpeg



358_3.jpeg



358_4.jpeg



358_5.jpeg



358_6.jpeg



358_7.jpeg



359_0.jpeg



359_1.jpeg



359_2.jpeg



359_3.jpeg



359_4.jpeg



359_5.jpeg



359_6.jpeg



359_7.jpeg



360_0.jpeg



360_1.jpeg



360_2.jpeg



360_3.jpeg



360_4.jpeg



360_5.jpeg



360_6.jpeg



360_7.jpeg



361_0.jpeg



361_1.jpeg



361_2.jpeg



361_3.jpeg



361_4.jpeg



361_5.jpeg



361_6.jpeg



361_7.jpeg



362_0.jpeg



362_1.jpeg



362_2.jpeg



362_3.jpeg



362_4.jpeg



362_5.jpeg



362_6.jpeg



362_7.jpeg



363_0.jpeg



363_1.jpeg



363_2.jpeg



363_3.jpeg



363_4.jpeg



363_5.jpeg



363_6.jpeg



363_7.jpeg



364_0.jpeg



364_1.jpeg



364_2.jpeg



364_3.jpeg



364_4.jpeg



364_5.jpeg



364_6.jpeg



364_7.jpeg



365_0.jpeg



365_1.jpeg



365_2.jpeg



365_3.jpeg



365_4.jpeg



365_5.jpeg



365_6.jpeg



365_7.jpeg



366_0.jpeg



366_1.jpeg



366_2.jpeg



366_3.jpeg



366_4.jpeg



366_5.jpeg



366_6.jpeg



366_7.jpeg



367_0.jpeg



367_1.jpeg



367_2.jpeg



367_3.jpeg



367_4.jpeg



367_5.jpeg



367_6.jpeg



367_7.jpeg



379_6.jpeg



380_3.jpeg



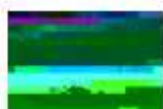
380_7.jpeg



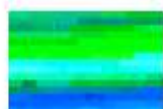
381_7.jpeg



382_3.jpeg



382_7.jpeg



383_6.jpeg



383_7.jpeg



384_3.jpeg



384_7.jpeg



387_7.jpeg



388_0.jpeg



388_1.jpeg



388_2.jpeg



388_3.jpeg



388_5.jpeg



397_2.jpeg



397_3.jpeg



397_4.jpeg



398_0.jpeg



399_5.jpeg



399_6.jpeg



409_1.jpeg



423_6.jpeg



425_1.jpeg



425_6.jpeg



426_6.jpeg



463_7.jpeg



63_0.jpeg



63_1.jpeg



63_2.jpeg



63_3.jpeg



63_4.jpeg



63_5.jpeg



63_6.jpeg



63_7.jpeg



64_0.jpeg



64_1.jpeg



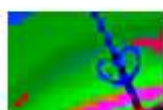
64_2.jpeg



64_5.jpeg



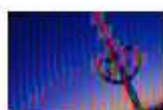
94_0.jpeg



94_3.jpeg



94_6.jpeg



95_1.jpeg



95_4.jpeg



95_7.jpeg



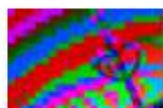
96_2.jpeg



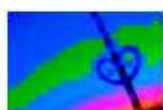
64_3.jpeg



64_6.jpeg



94_1.jpeg



94_4.jpeg



94_7.jpeg



95_2.jpeg



95_5.jpeg



96_0.jpeg



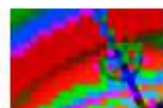
96_3.jpeg



64_4.jpeg



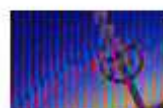
64_7.jpeg



94_2.jpeg



94_5.jpeg



95_0.jpeg



95_3.jpeg



95_6.jpeg



96_1.jpeg