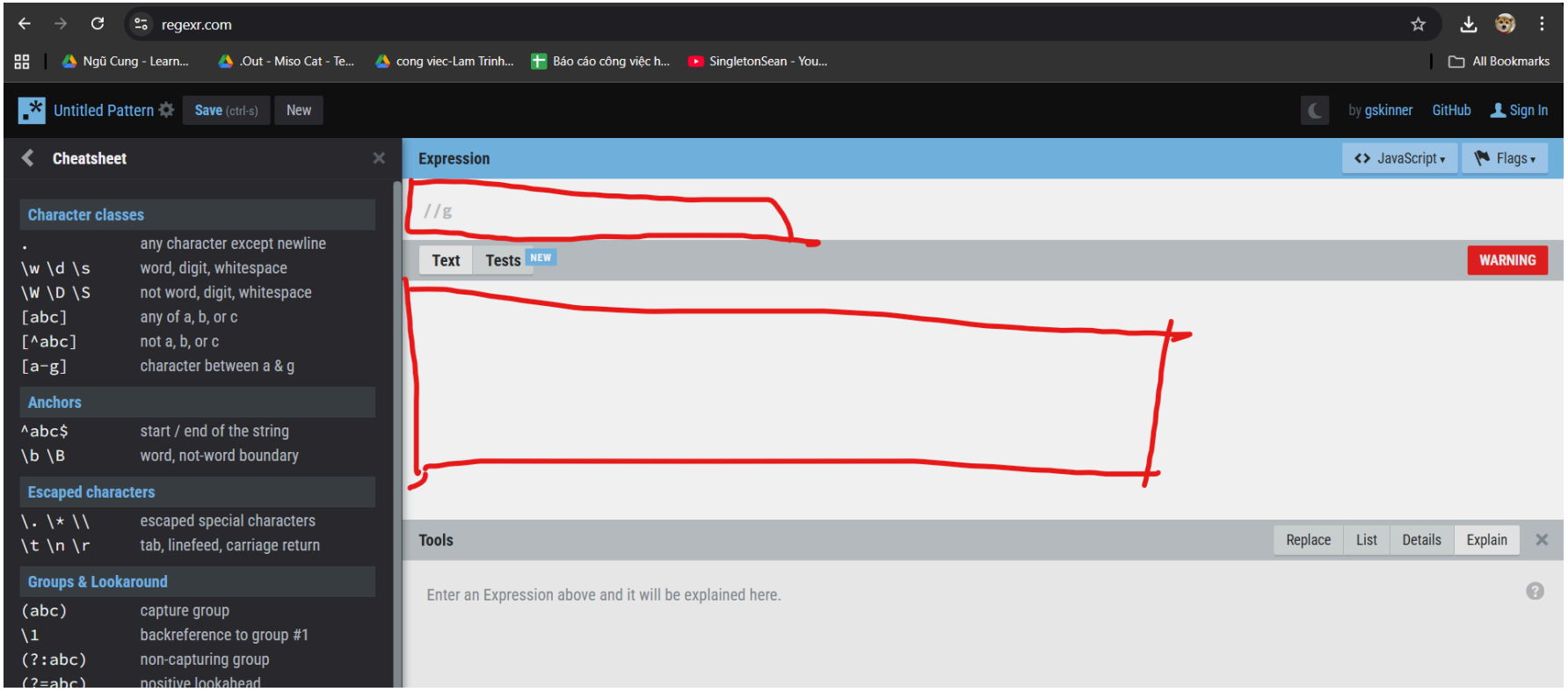
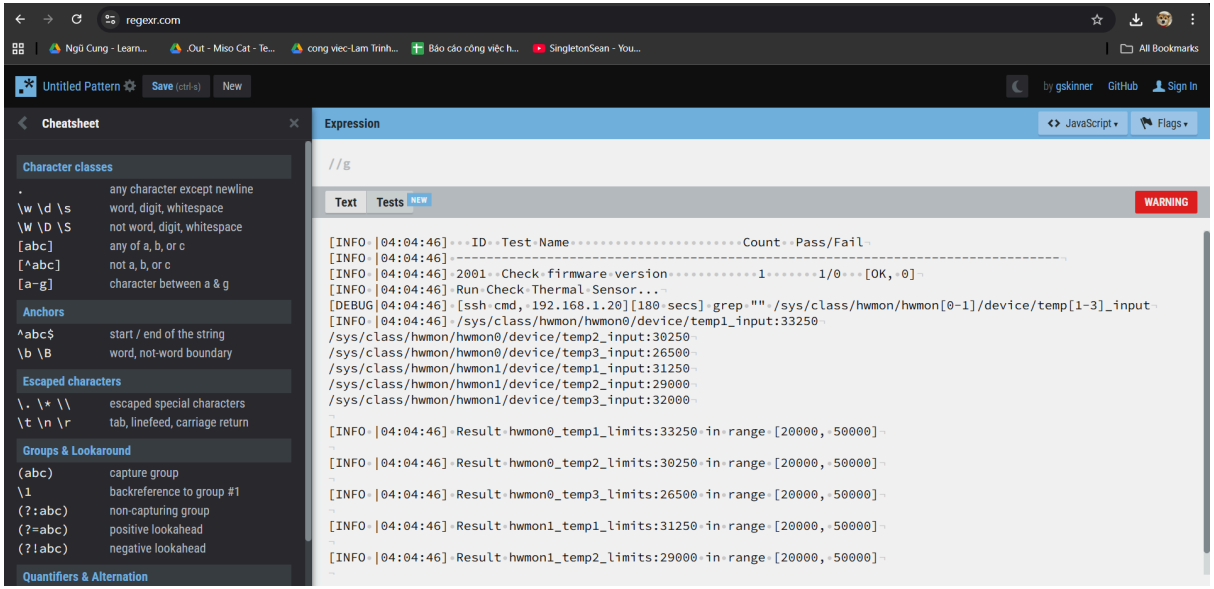
con cac
CÓ 17 TRANG NHƯNG MÀ TOÀN ẢNH THÔI KHÔNG CÓ CHỮ ĐÂU BẠN ĐỪNG SỢ NẾU BẠN SỢ THÌ BẠN NÊN UỐNG CHANH MUỐI. BẮT ĐẦU THÔI.
bạn tự viết regex tầm 2-3 lần là tự viết được mọi loại data cần get trong log. bạn có thể hỏi GPT nhưng mk không khuyến khích vì bạn chả hiểu vì sao lại viết như thế và cần sửa chỗ nào(nhiều khi nó bị ngu và không bắt được đâu). khi mà bạn hiểu thì có thể tự viết được mọi loại regex để bắt tất cả những loại data về sau mà bạn muốn.
1. để thử nghiệm xem có bắt được đoạn data không thì vào https://regexr.com/ để thử nghiệm trực tiếp. có 2 phần quan trọng là expression và Text



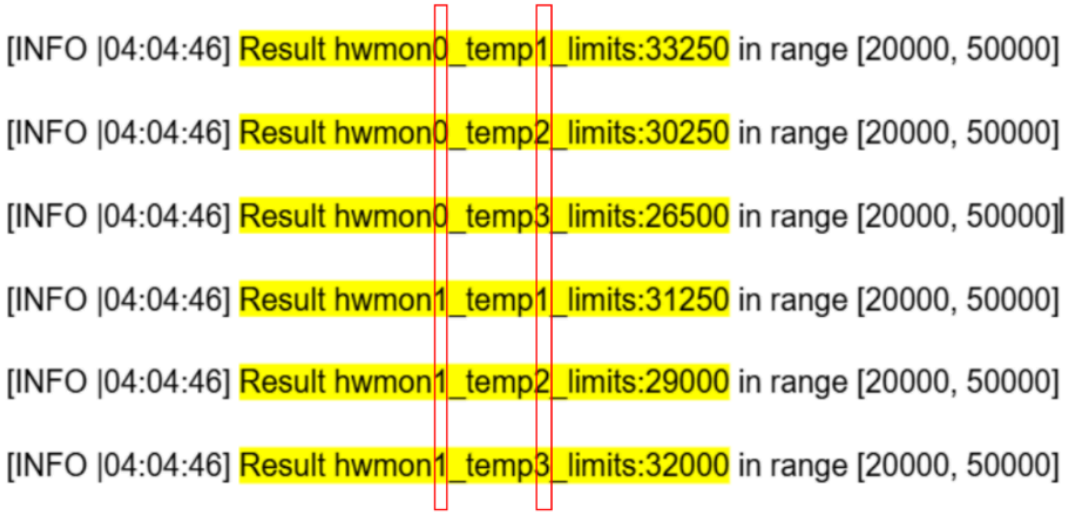
Giờ làm ví dụ với lấy sensor, ta có đoạn data sensor trong LOG như sau :

```
[INFO |04:04:46]   ID  Test Name                    Count  Pass/Fail
[INFO |04:04:46] -------------------------------------------------------------------------------
[INFO |04:04:46] 2001  Check firmware version          1       1/0   [OK, 0]
[INFO |04:04:46] Run Check Thermal Sensor...
[DEBUG|04:04:46] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/temp[1-3]_input
[INFO |04:04:46] /sys/class/hwmon/hwmon0/device/temp1_input:33250
/sys/class/hwmon/hwmon0/device/temp2_input:30250
/sys/class/hwmon/hwmon0/device/temp3_input:26500
/sys/class/hwmon/hwmon1/device/temp1_input:31250
/sys/class/hwmon/hwmon1/device/temp2_input:29000
/sys/class/hwmon/hwmon1/device/temp3_input:32000

[INFO |04:04:46] Result hwmon0_temp1_limits:33250 in range [20000, 50000]

[INFO |04:04:46] Result hwmon0_temp2_limits:30250 in range [20000, 50000]

[INFO |04:04:46] Result hwmon0_temp3_limits:26500 in range [20000, 50000]

[INFO |04:04:46] Result hwmon1_temp1_limits:31250 in range [20000, 50000]

[INFO |04:04:46] Result hwmon1_temp2_limits:29000 in range [20000, 50000]

[INFO |04:04:46] Result hwmon1_temp3_limits:32000 in range [20000, 50000]
```

Đoạn data màu vàng là đoạn chúng ta cần phải lấy ra và phải lấy ra toàn bộ không bỏ sót dòng nào.
ta paste đoạn log trên vào https://regexr.com/ ở mục text để tiến hành thử nghiệm.



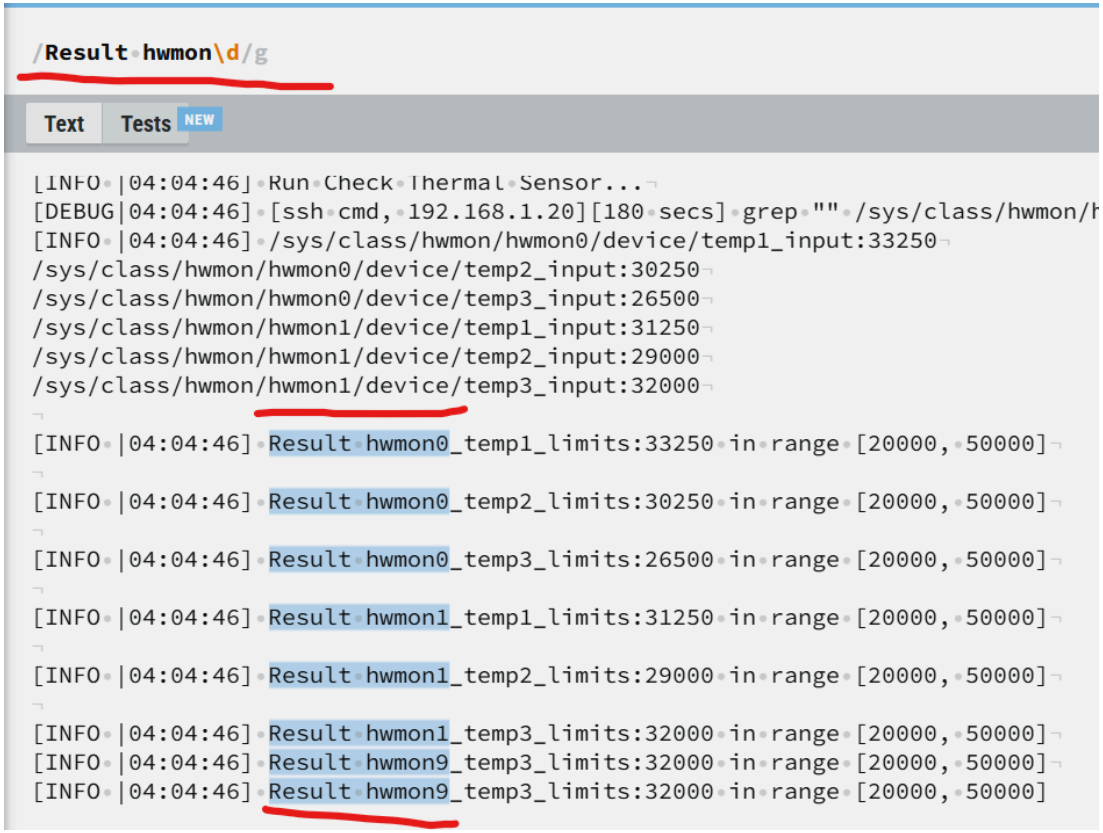
Ta cần xác định trong đoạn data muốn lấy ra có những thành phần nào cố định thành phần nào thay đổi. Xem ví dụ dưới đây để viết regex:



Chúng ta sẽ đi từ trái qua phải từng phần từng phần
Như hình chúng ta để ý phần Result hwmon thì dòng nào cũng giống nhau. chỉ khác mỗi số đằng sau đó :

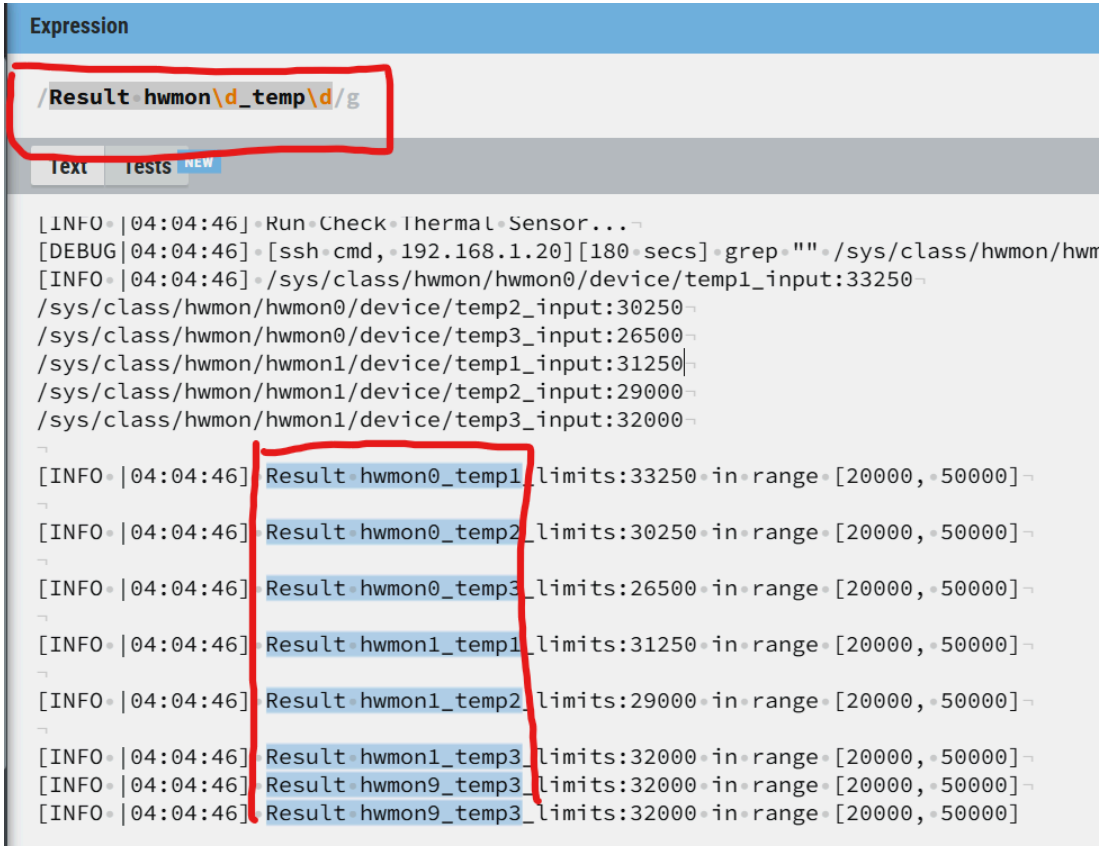
Result hwmon1
Result hwmon0

vậy thì regex sẽ là
Result hwmon\d
vì sao là \d vì \d là đại diện cho một số kiểu như là Result hwmon\d=Result hwmon+một số nào đó.
Giờ tôi thử paste vào expresstion đoạn regex kia xem kết quả:



đã bắt được bất kể là: Result hwmon9,Result hwmon1,Result hwmon3 hay Result hwmon0 thì đều bắt được.
cùng xử lí phần tiếp theo nào. Lại giống hệt phần trước _temp cố định và tiếp theo là một số, lại áp dụng công thức ta có :
_temp\d
Nối hai đoạn regex vào với nhau ta có :
Result hwmon\d và _temp\d ------> Result hwmon\d_temp\d

Vào lại web kia paste thử đoạn regex xem sao

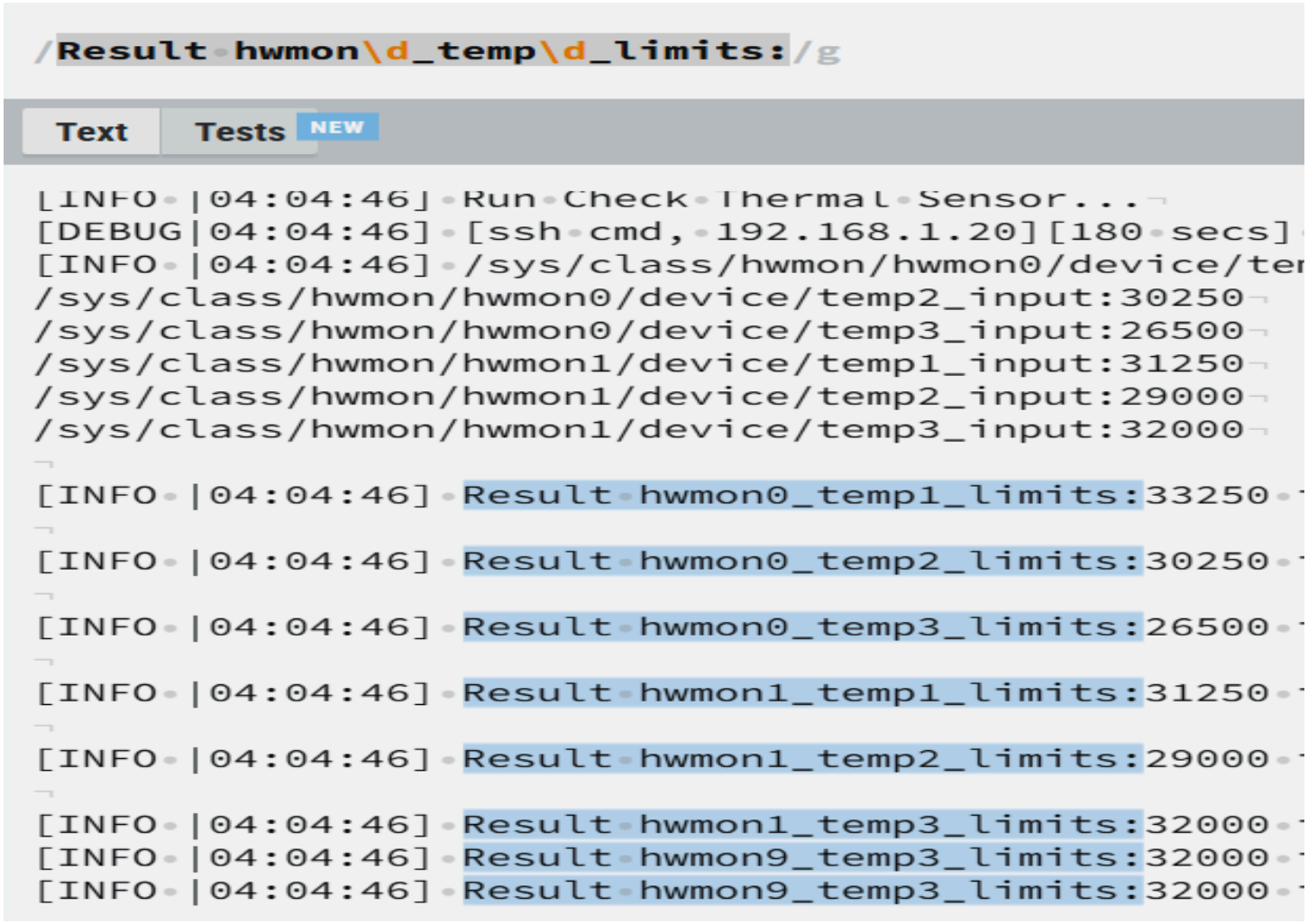


oke đã bắt được . Phần tiếp theo là _limits: phần này thì dòng nào cũng như nhau thế nên ghép thẳng vào regex ta có :

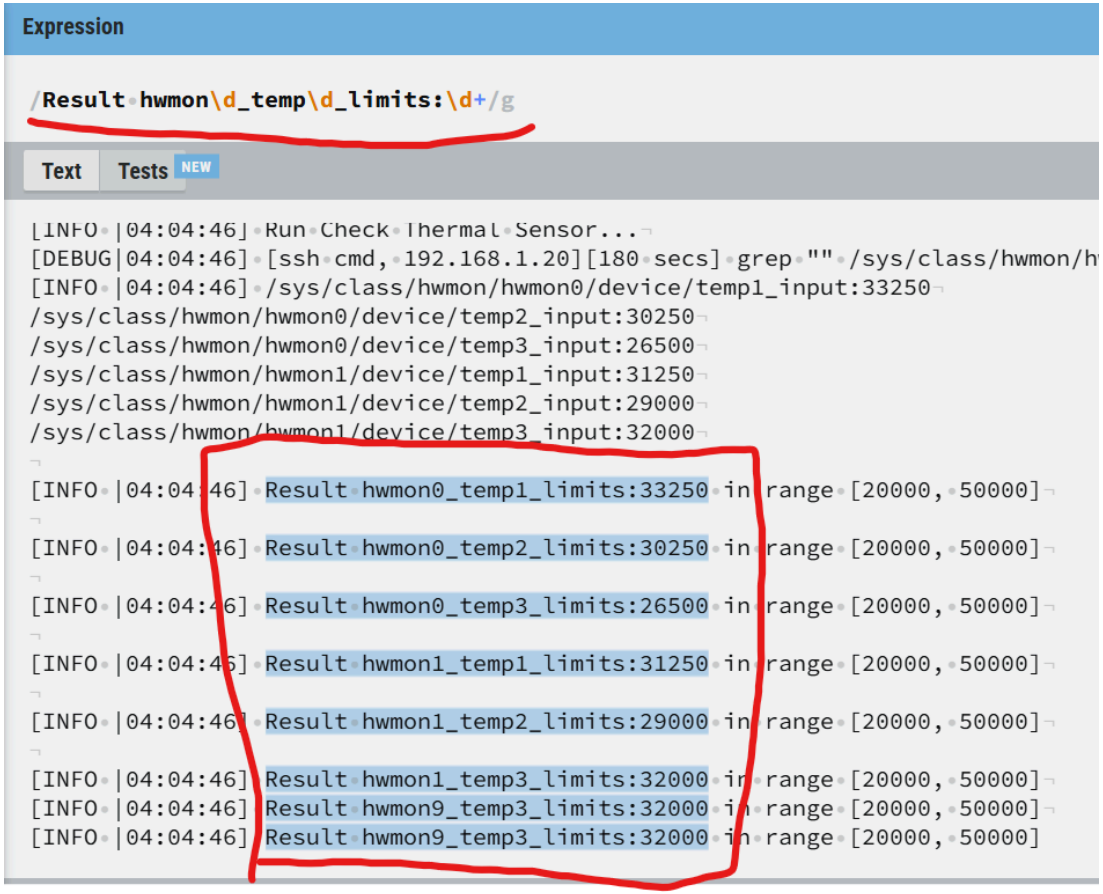
Result hwmon\d_temp\d_limits:

Cùng xem kết quả :

(ảnh to quá trôi sang trang dưới)



vậy là còn phần giá trị nữa là xong !
ta có 33250 là giá trị cần lấy, mà \d chỉ đại diện cho MỘT CHỮ SỐ mà số 33250 có nhiều hơn MỘT CHỮ SỐ thì ta dùng \d+ cùng xem kết quả :



Vậy là đã bắt được hết đoạn data muốn lấy. Nhưng còn 1 phần nữa trước khi paste regex vào tool đó là nhóm.

BẠN CẦN NHÓM LÀM 2 PHẦN

Phần 1 là tiền tố, phần 2 là giá trị. ta có đoạn regex đã bắt được toàn bộ đoạn data là

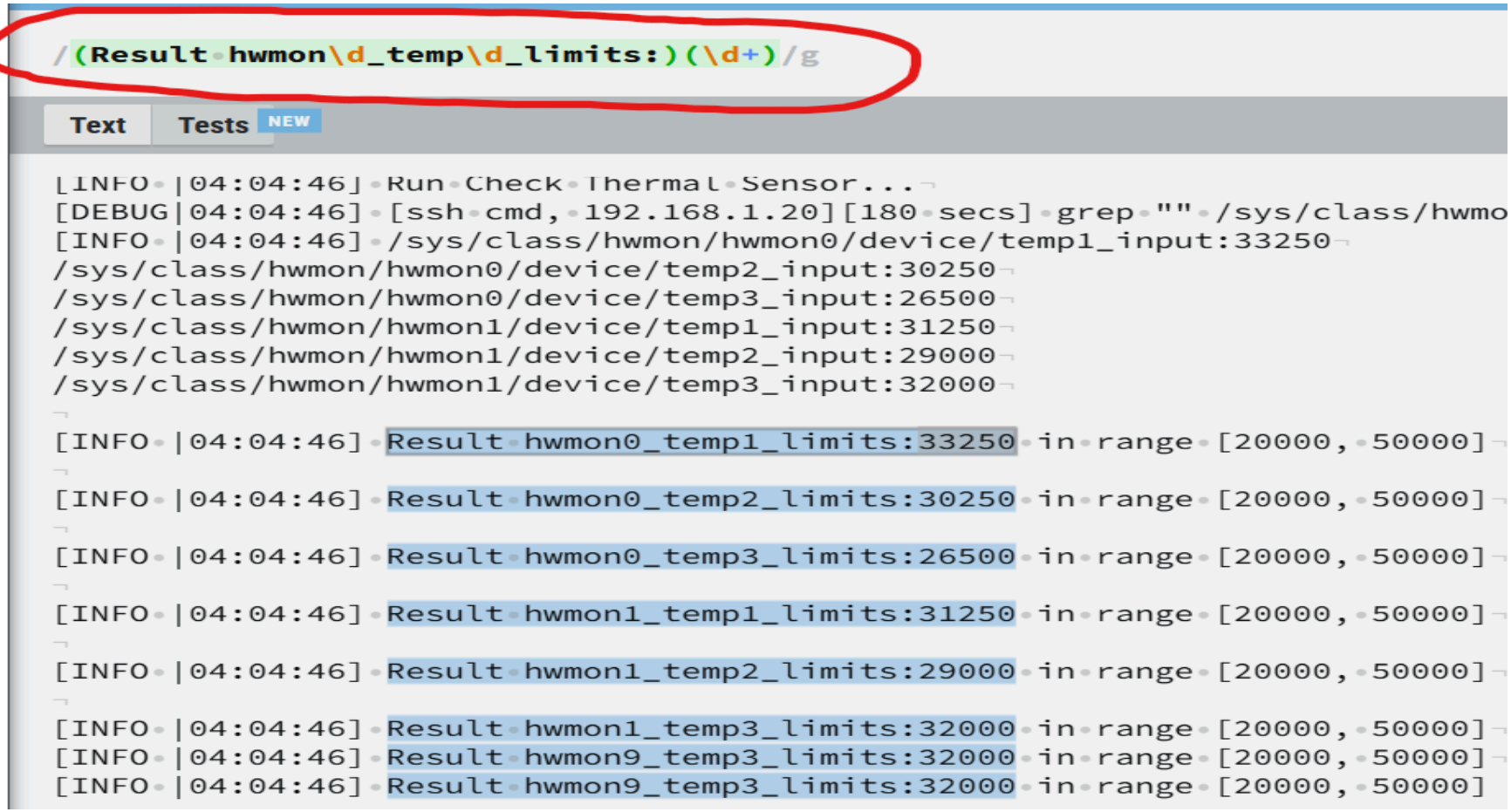
Result hwmon\d_temp\d_limits:\d+

phần tiền tố là Result hwmon0_temp1_limits:

phần giá trị là 33250

tương ứng trong regex là

phần tiền tố là Result hwmon\d_temp\d_limits:
Phần giá trị là \d+
Nhóm thì thêm ngoặc tròn vào là được giúp mình. regex cuối cùng sẽ là:
(Result hwmon\d_temp\d_limits:)(\d+)
tách nhóm này để tool chia data cho tiện chỉ là thêm ngoặc tròn thôi mà dễ hiểu mà đúng không bạn 🙂 .

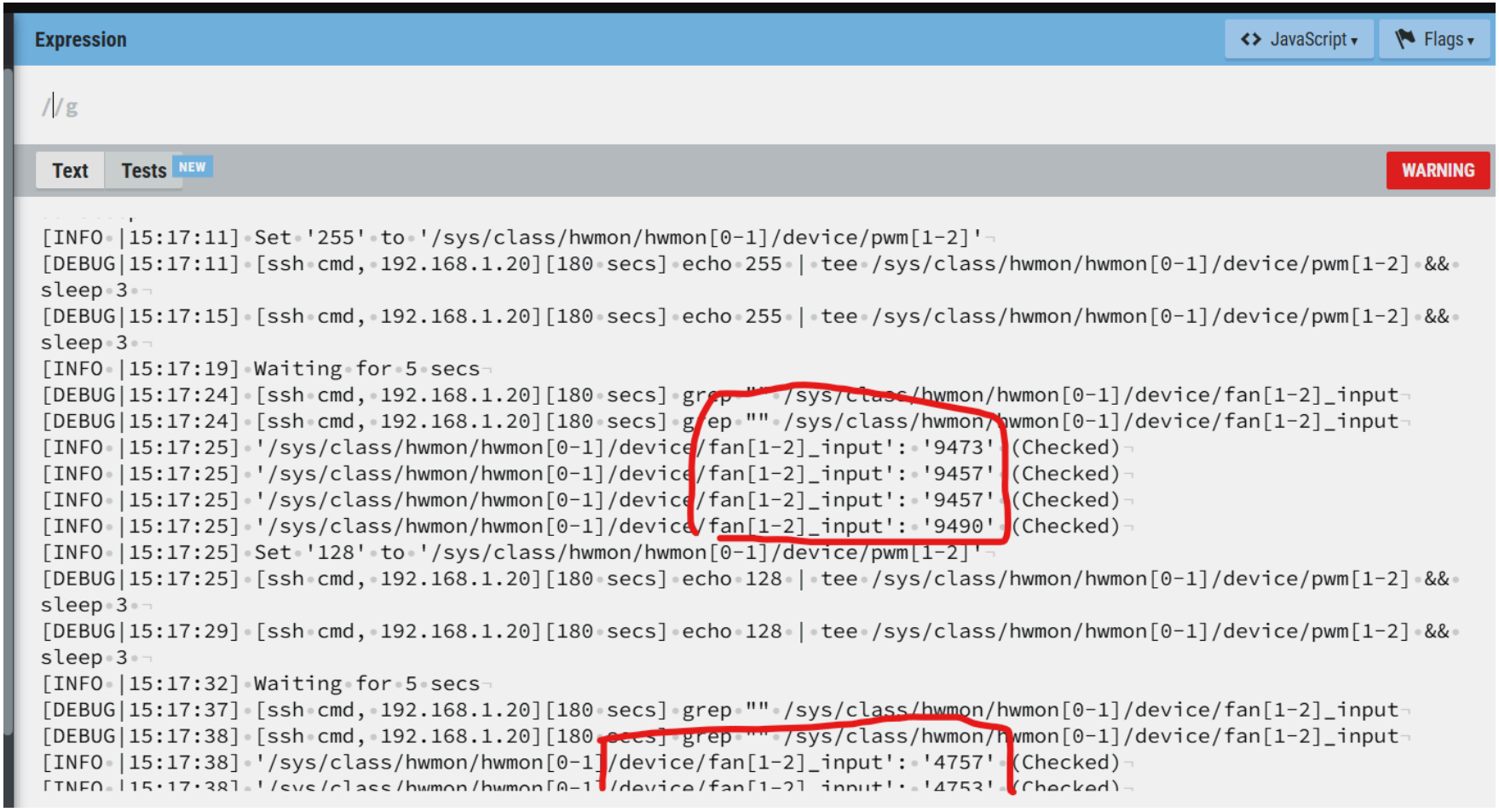


LÀM TIẾP VÍ DỤ NỮA NHé !
bạn tự viết regex tầm 2-3 lần là tự viết được mọi loại data cần get trong log
lần này lấy fan data (thằng này mỗi con hàng 1 khác nhau nhưng mà đừng lo dễ như ăn bánh vì parttern vẫn như cũ mà thôi)
ta lại có log:

[INFO |15:17:09] Disabled PID auto fan
[DEBUG|15:17:09] [ssh cmd, 192.168.1.20][180 secs] sed -i "/fanctrl_wrapper.sh/ s/^/#/" /etc/inittab; init -q
[INFO |15:17:10] Set '/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]_enable' to 'manual' mode
[DEBUG|15:17:10] [ssh cmd, 192.168.1.20][180 secs] echo 1 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]_enable && sleep 1
[INFO |15:17:11] Set '255' to '/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]'
[DEBUG|15:17:11] [ssh cmd, 192.168.1.20][180 secs] echo 255 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[DEBUG|15:17:15] [ssh cmd, 192.168.1.20][180 secs] echo 255 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[INFO |15:17:19] Waiting for 5 secs
[DEBUG|15:17:24] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[DEBUG|15:17:24] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[INFO |15:17:25] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '9473' (Checked)
[INFO |15:17:25] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '9457' (Checked)
[INFO |15:17:25] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '9457' (Checked)
[INFO |15:17:25] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '9490' (Checked)
[INFO |15:17:25] Set '128' to '/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]'
[DEBUG|15:17:25] [ssh cmd, 192.168.1.20][180 secs] echo 128 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[DEBUG|15:17:29] [ssh cmd, 192.168.1.20][180 secs] echo 128 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[INFO |15:17:32] Waiting for 5 secs
[DEBUG|15:17:37] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[DEBUG|15:17:38] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[INFO |15:17:38] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '4757' (Checked)
[INFO |15:17:38] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '4753' (Checked)
[INFO |15:17:38] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '4745' (Checked)
[INFO |15:17:38] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '4741' (Checked)

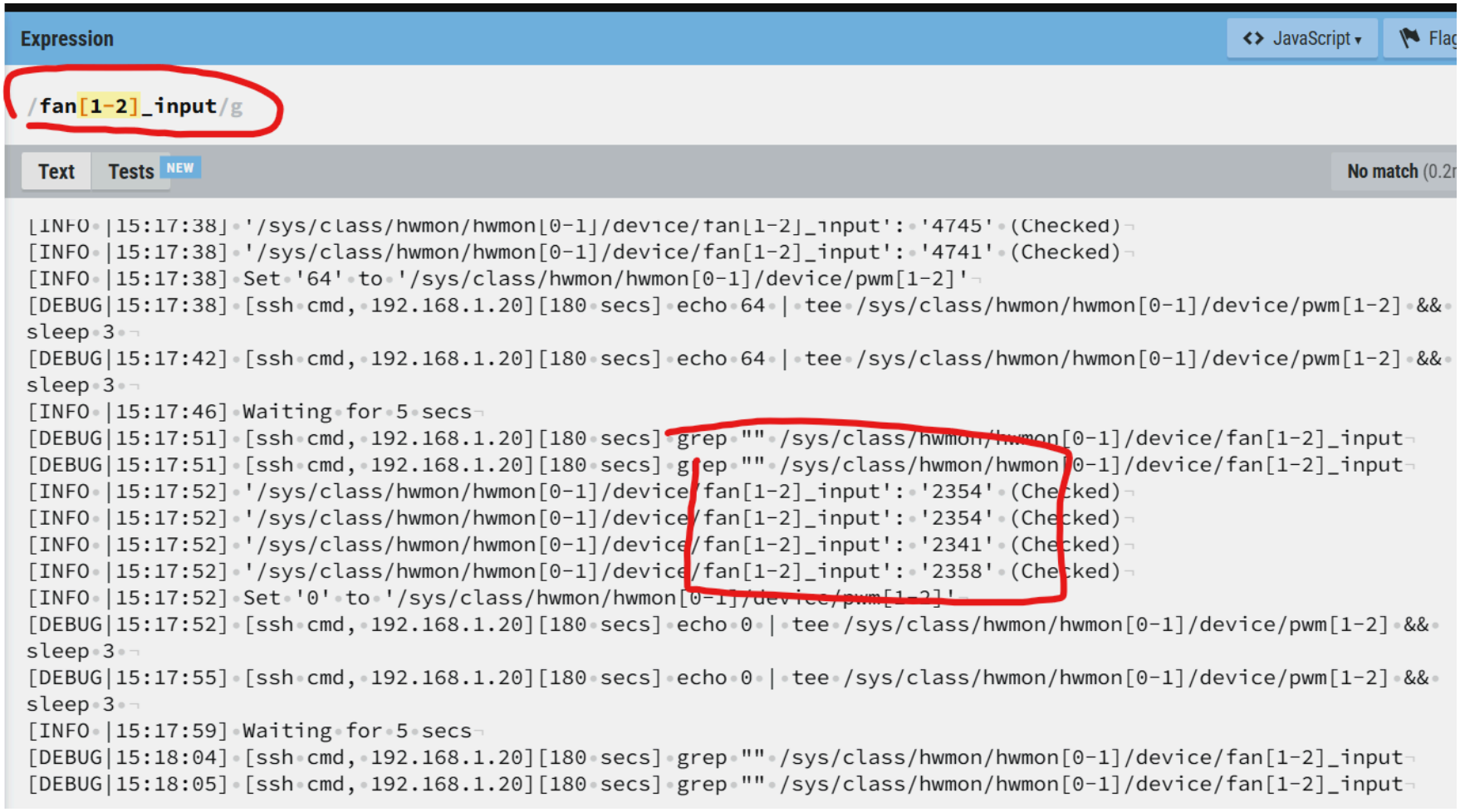
[INFO |15:17:38] Set '64' to '/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]'
[DEBUG|15:17:38] [ssh cmd, 192.168.1.20][180 secs] echo 64 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[DEBUG|15:17:42] [ssh cmd, 192.168.1.20][180 secs] echo 64 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[INFO |15:17:46] Waiting for 5 secs
[DEBUG|15:17:51] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[DEBUG|15:17:51] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[INFO |15:17:52] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '2354' (Checked)
[INFO |15:17:52] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '2354' (Checked)
[INFO |15:17:52] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '2341' (Checked)
[INFO |15:17:52] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '2358' (Checked)
[INFO |15:17:52] Set '0' to '/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]'
[DEBUG|15:17:52] [ssh cmd, 192.168.1.20][180 secs] echo 0 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[DEBUG|15:17:55] [ssh cmd, 192.168.1.20][180 secs] echo 0 | tee /sys/class/hwmon/hwmon[0-1]/device/pwm[1-2] && sleep 3
[INFO |15:17:59] Waiting for 5 secs
[DEBUG|15:18:04] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[DEBUG|15:18:05] [ssh cmd, 192.168.1.20][180 secs] grep "" /sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input
[INFO |15:18:05] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '0' (Checked)
[INFO |15:18:05] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '0' (Checked)
[INFO |15:18:05] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '0' (Checked)
[INFO |15:18:05] '/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input': '0' (Checked)
[INFO |15:18:05] Enabled PID auto fan

Bước 1 vẫn như cũ, xác định mục tiêu chính là đoạn màu vàng cần lấy ra. lại paste vào https://regexr.com/ để thử nghiệp tiếp thôi



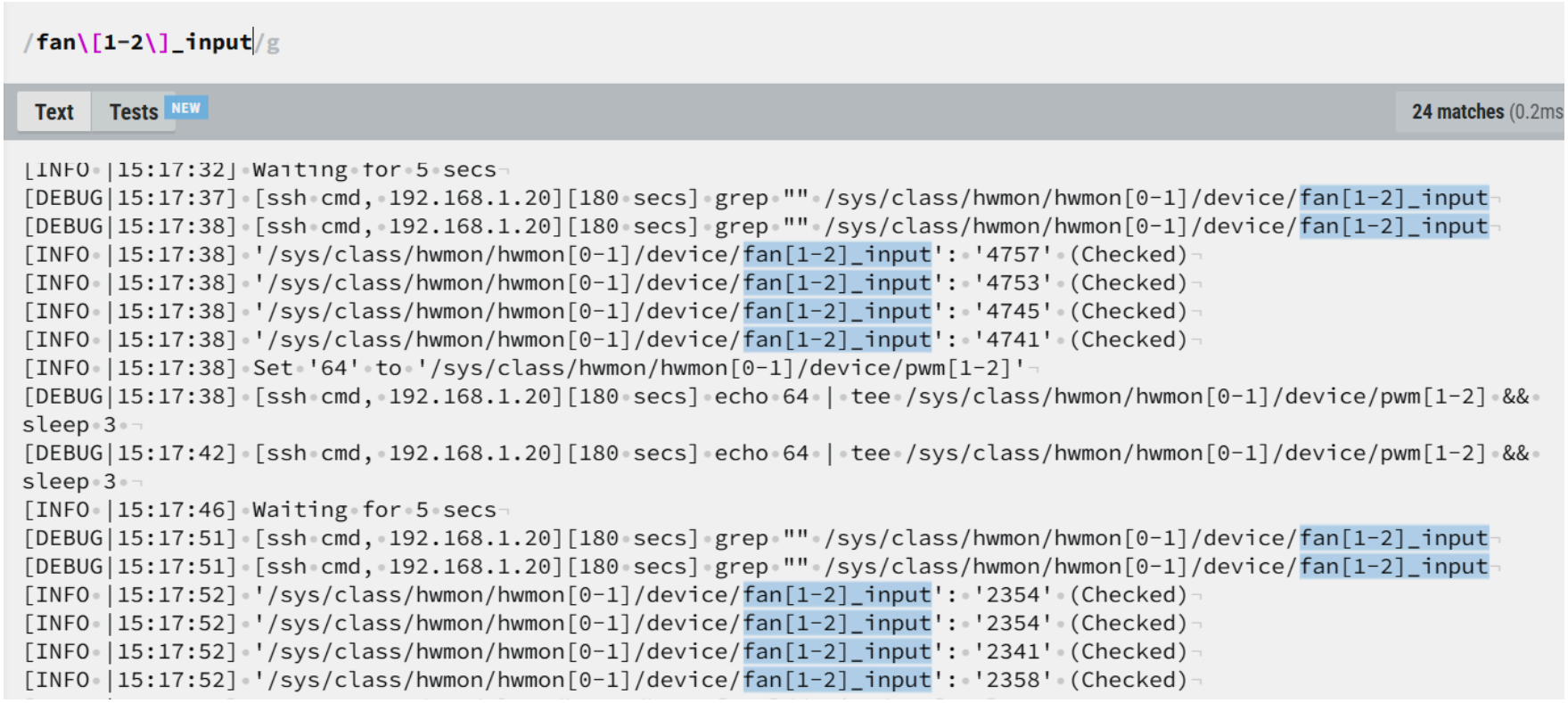
Giờ vẫn phân tích như cũ xem 16 dòng bôi vàng kia giống nhau ở những chữ nào. ở này chúng ta sẽ được học thêm về ký tự đặc biệt trong regex.

Trông có vẻ chữ fan[1-2]_input ở dòng nào cũng có paste thử vào web kia xem sao



Ồ điều gì đang xảy ra sao không bắt được thằng nào ??????????
đó chính là tại ký tự ngoặc vuông chính là đoạn [1-2] giờ sao ??? vậy ký tự ngoặc vuông trong regex là gì ?
quy tắc chung nhé nếu bạn nghĩ ký tự nào là đặc biệt thì bạn thêm dấu \ trước ký tự đó ví dụ ký tự [trong regex sẽ là \[
ký tự] sẽ là \]

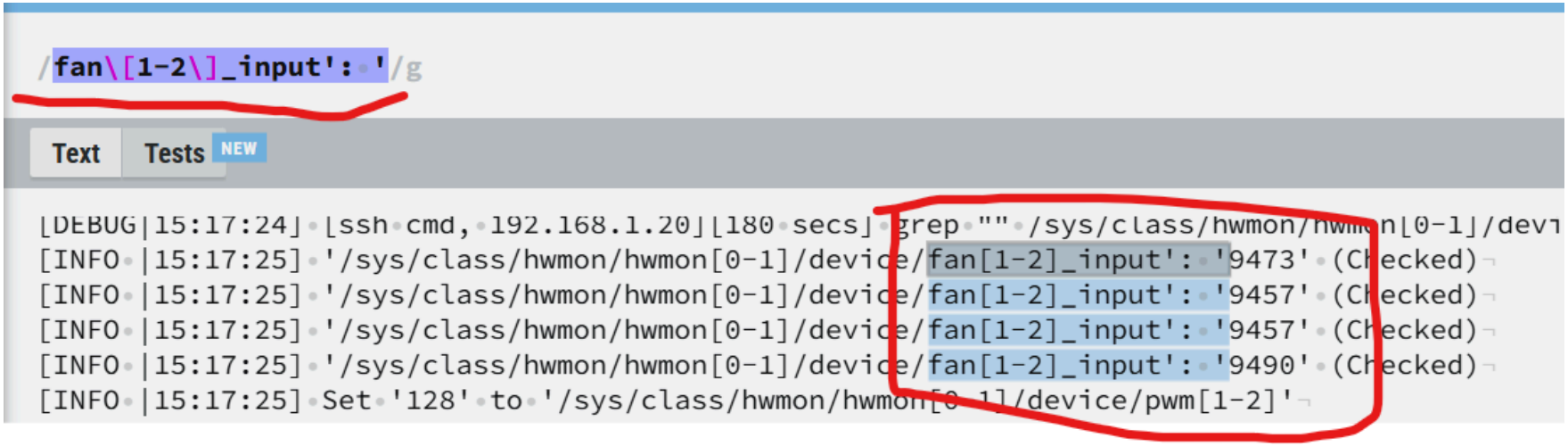
vậy sửa đoạn fan[1-2]_input thành fan\[1-2\]_input
à sẽ có bạn thắc măc sao không có \d kiểu như fan[\d-\d]_input vì dòng nào cũng là fan[1-2]_input vì số 1-2 có thay đổi đâu nếu có 1-3 hay 1-4 chúng ta mới phải dùng \d như đã làm ở ví dụ 1.
cùng xem kết quả nào :



ngon luôn ! đã bắt được
bạn sẽ thấy nó bắt thêm cả mấy cái chúng ta không muốn bắt nhưng không sao còn phần sau nữa mà !
giờ xử lí phần còn lại. lấy 1 mẫu ra bóc tách .

fan[1-2]_input': '4757'

còn đoạn nhì nhằng đằng sau fan[1-2]_input thì sao thì chả sao cả copy paste thẳng vào regex :



bắt ngon luôn !
giờ lấy phần giá trị nữa thôi là xong
vẫn như cũ giá trị là số tự nhiên vẫn dùng \d+

```
/fan\[1-2\]_input':·'\d+/g
```

Text | Tests NEW

```
[INFO·|15:17:19]·Waiting·for·5·secs¬
[DEBUG|15:17:24]·[ssh·cmd,·192.168.1.20][180·secs]·grep·""·/sys/class/hwmon/h
[DEBUG|15:17:24]·[ssh·cmd,·192.168.1.20][180·secs]·grep·""·/sys/class/hwmon/h
[INFO·|15:17:25]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'9473'·
[INFO·|15:17:25]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'9457'·
[INFO·|15:17:25]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'9457'·
[INFO·|15:17:25]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'9490'·
[INFO·|15:17:25]·Set·'128'·to·'/sys/class/hwmon/hwmon[0-1]/device/pwm[1-2]'¬
[DEBUG|15:17:25]·[ssh·cmd,·192.168.1.20][180·secs]·echo·128·|·tee·/sys/class/
sleep·3·¬
[DEBUG|15:17:29]·[ssh·cmd,·192.168.1.20][180·secs]·echo·128·|·tee·/sys/class/
sleep·3·¬
[INFO·|15:17:32]·Waiting·for·5·secs¬
[DEBUG|15:17:37]·[ssh·cmd,·192.168.1.20][180·secs]·grep·""·/sys/class/hwmon/h
[DEBUG|15:17:38]·[ssh·cmd,·192.168.1.20][180·secs]·grep·""·/sys/class/hwmon/h
[INFO·|15:17:38]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'4757'·
[INFO·|15:17:38]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'4753'·
[INFO·|15:17:38]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'4745'·
[INFO·|15:17:38]·'/sys/class/hwmon/hwmon[0-1]/device/fan[1-2]_input':·'4741'·
```

vậy là xong và dĩ nhiên không lấy dấu phẩy đằng sau số kia vì nó tác dụng gì đâu
Đừng quên nhóm thành 2 nhóm nhé tiền tố và giá trị

fan[1-2]_input': '4757

(fan\[1-2\]_input': ')(\d+)

đó vậy là xong dễ vl ta có regex cuối cùng là : (fan\[1-2\]_input': ')(\d+)

vậy là bạn đã tốt nghiệp khóa regex cơ bản chúc mừng bạn !
PASTE TIỀN TỐ VÀ GIÁ TRỊ VÀO TOOL LÀ OKE !