In [1]:

```
#convert txt to csv
import pandas as pd
import matplotlib.pyplot as plt
```

In [2]:

```
dfg = pd.read_csv('wat_good.csv')
```

In [3]:

```
dfm = pd.read_csv('wats_mal.csv')
```

In [4]:

```
dfg.head()
```

Out[4]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
0	7	0	2	0	0x1dc0	0	0	3	2	8	0
1	11	1	2	4	0x1dc0	0	0	3	2	8	0
2	15	2	2	4	0x1dc0	0	0	3	2	8	0
3	19	3	1	4	0x1dc0	0	0	3	2	8	0
4	23	7	0	3	0x1dc0	0	0	3	2	8	0

In [5]:

```
dfm.head()
```

Out[5]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
0	7	0	2	0	0x1dc0	0	0	3	2	8	0
1	7	1	1	0	0xecf40	0	0	3	2	8	1
2	11	5	1	3	0xecf40	0	0	3	2	8	1
3	11	1	2	4	0x1dc0	0	0	3	2	8	0
4	15	2	2	4	0x1dc0	0	0	3	2	8	0

In [6]:

```
df = dfg.append(dfm, ignore_index=True)
```

In [7]:

```
df.head()
```

Out[7]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
0	7	0	2	0	0x1dc0	0	0	3	2	8	0
1	11	1	2	4	0x1dc0	0	0	3	2	8	0
2	15	2	2	4	0x1dc0	0	0	3	2	8	0
3	19	3	1	4	0x1dc0	0	0	3	2	8	0
4	23	7	0	3	0x1dc0	0	0	3	2	8	0

In [8]:

```
df = df.sort_values('time')
```

In [9]:

```
df.head(50)
```

Out[9]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	sr
0	7	0	2	0	0x1dc0	0	0	3	2	8	
245668	7	1	1	0	0xecf40	0	0	3	2	8	
245667	7	0	2	0	0x1dc0	0	0	3	2	8	
245669	11	5	1	3	0xecf40	0	0	3	2	8	
245670	11	1	2	4	0x1dc0	0	0	3	2	8	
1	11	1	2	4	0x1dc0	0	0	3	2	8	
245671	15	2	2	4	0x1dc0	0	0	3	2	8	
245672	15	9	1	3	0xecf40	0	0	3	2	8	
2	15	2	2	4	0x1dc0	0	0	3	2	8	
245673	19	13	0	3	0xecf40	0	0	3	2	8	
245674	19	3	1	4	0x1dc0	0	0	3	2	8	
3	19	3	1	4	0x1dc0	0	0	3	2	8	
4	23	7	0	3	0x1dc0	0	0	3	2	8	
245675	23	7	0	3	0x1dc0	0	0	3	2	8	
245676	58	13	3	0	0xecf40	1	0	0	4	16	
245677	59	13	3	0	0xecf40	1	1	1	4	16	
245678	60	13	3	0	0xecf40	1	2	1	4	16	
245679	61	13	3	0	0xecf40	1	3	1	4	16	
245680	62	9	3	1	0xecf40	1	0	0	4	16	
245681	62	7	4	0	0x1dc0	1	0	0	4	16	
5	62	7	4	0	0x1dc0	1	0	0	4	16	
6	63	7	4	0	0x1dc0	1	1	1	4	16	
245682	63	9	3	1	0xecf40	1	1	1	4	16	
245683	63	7	4	0	0x1dc0	1	1	1	4	16	
7	64	7	4	0	0x1dc0	1	2	1	4	16	
245684	64	7	4	0	0x1dc0	1	2	1	4	16	
245685	64	9	3	1	0xecf40	1	2	1	4	16	
245688	65	13	3	0	0xecf40	1	4	2	4	16	
245687	65	9	3	1	0xecf40	1	3	1	4	16	
245686	65	7	4	0	0x1dc0	1	3	1	4	16	
8	65	7	4	0	0x1dc0	1	3	1	4	16	
245689	66	6	4	2	0x1dc0	1	0	0	4	16	
245690	66	5	3	1	0xecf40	1	0	0	4	16	
9	66	6	4	2	0x1dc0	1	0	0	4	16	

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	sr
245691	67	6	4	2	0x1dc0	1	1	1	4	16	
245692	67	5	3	1	0xecf40	1	1	1	4	16	
10	67	6	4	2	0x1dc0	1	1	1	4	16	
245693	68	5	3	1	0xecf40	1	2	1	4	16	
245694	68	6	4	2	0x1dc0	1	2	1	4	16	
11	68	6	4	2	0x1dc0	1	2	1	4	16	
245695	69	5	3	1	0xecf40	1	3	1	4	16	
245696	69	6	4	2	0x1dc0	1	3	1	4	16	
245697	69	9	3	1	0xecf40	1	4	2	4	16	
245698	69	7	4	0	0x1dc0	1	4	2	4	16	
12	69	6	4	2	0x1dc0	1	3	1	4	16	
13	69	7	4	0	0x1dc0	1	4	2	4	16	
245699	70	5	4	2	0x1dc0	1	0	0	4	16	
14	70	5	4	2	0x1dc0	1	0	0	4	16	
245700	70	1	0	1	0xecf40	1	0	0	4	16	
245701	71	5	4	2	0x1dc0	1	1	1	4	16	

In [10]:

```
df.to_csv('wat-com.csv',index=False)
```

In [11]:

```
df = pd.read_csv('wat-com.csv')
```

In [12]:

```
df.head(50)
```

Out[12]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
0	7	0	2	0	0x1dc0	0	0	3	2	8	0
1	7	1	1	0	0xecf40	0	0	3	2	8	1
2	7	0	2	0	0x1dc0	0	0	3	2	8	0
3	11	5	1	3	0xecf40	0	0	3	2	8	1
4	11	1	2	4	0x1dc0	0	0	3	2	8	0
5	11	1	2	4	0x1dc0	0	0	3	2	8	0
6	15	2	2	4	0x1dc0	0	0	3	2	8	0
7	15	9	1	3	0xecf40	0	0	3	2	8	1
8	15	2	2	4	0x1dc0	0	0	3	2	8	0
9	19	13	0	3	0xecf40	0	0	3	2	8	1
10	19	3	1	4	0x1dc0	0	0	3	2	8	0
11	19	3	1	4	0x1dc0	0	0	3	2	8	0
12	23	7	0	3	0x1dc0	0	0	3	2	8	0
13	23	7	0	3	0x1dc0	0	0	3	2	8	0
14	58	13	3	0	0xecf40	1	0	0	4	16	29
15	59	13	3	0	0xecf40	1	1	1	4	16	29
16	60	13	3	0	0xecf40	1	2	1	4	16	29
17	61	13	3	0	0xecf40	1	3	1	4	16	29
18	62	9	3	1	0xecf40	1	0	0	4	16	29
19	62	7	4	0	0x1dc0	1	0	0	4	16	23
20	62	7	4	0	0x1dc0	1	0	0	4	16	23
21	63	7	4	0	0x1dc0	1	1	1	4	16	23
22	63	9	3	1	0xecf40	1	1	1	4	16	29
23	63	7	4	0	0x1dc0	1	1	1	4	16	23
24	64	7	4	0	0x1dc0	1	2	1	4	16	23
25	64	7	4	0	0x1dc0	1	2	1	4	16	23
26	64	9	3	1	0xecf40	1	2	1	4	16	29
27	65	13	3	0	0xecf40	1	4	2	4	16	29
28	65	9	3	1	0xecf40	1	3	1	4	16	29
29	65	7	4	0	0x1dc0	1	3	1	4	16	23
30	65	7	4	0	0x1dc0	1	3	1	4	16	23
31	66	6	4	2	0x1dc0	1	0	0	4	16	23
32	66	5	3	1	0xecf40	1	0	0	4	16	29
33	66	6	4	2	0x1dc0	1	0	0	4	16	23

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
34	67	6	4	2	0x1dc0	1	1	1	4	16	23
35	67	5	3	1	0xecf40	1	1	1	4	16	29
36	67	6	4	2	0x1dc0	1	1	1	4	16	23
37	68	5	3	1	0xecf40	1	2	1	4	16	29
38	68	6	4	2	0x1dc0	1	2	1	4	16	23
39	68	6	4	2	0x1dc0	1	2	1	4	16	23
40	69	5	3	1	0xecf40	1	3	1	4	16	29
41	69	6	4	2	0x1dc0	1	3	1	4	16	23
42	69	9	3	1	0xecf40	1	4	2	4	16	29
43	69	7	4	0	0x1dc0	1	4	2	4	16	23
44	69	6	4	2	0x1dc0	1	3	1	4	16	23
45	69	7	4	0	0x1dc0	1	4	2	4	16	23
46	70	5	4	2	0x1dc0	1	0	0	4	16	23
47	70	5	4	2	0x1dc0	1	0	0	4	16	23
48	70	1	0	1	0xecf40	1	0	0	4	16	29
49	71	5	4	2	0x1dc0	1	1	1	4	16	23

In [13]:

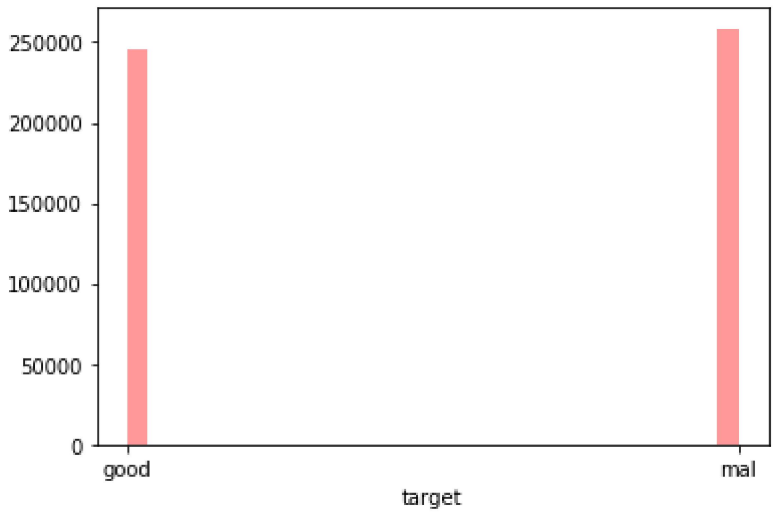
```
import seaborn as sns
```

In [14]:

```
sns.distplot(df['target'], kde = False, bins=30, color='red')
```

Out[14]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21085f6db88>
```



In [15]:

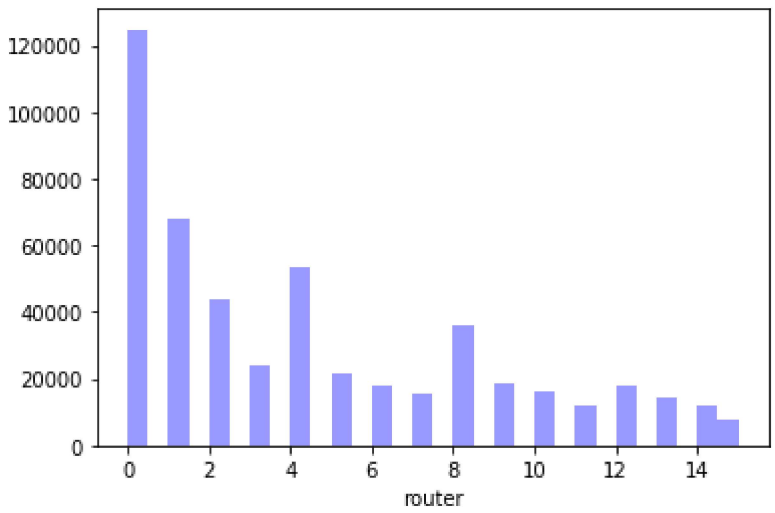
```
tar = {'good': 1,'mal': 0}
df = df.replace({'target': tar})
```

In [16]:

```
sns.distplot(df['router'], kde = False, bins=30, color='blue')
```

Out[16]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21086093c48>
```

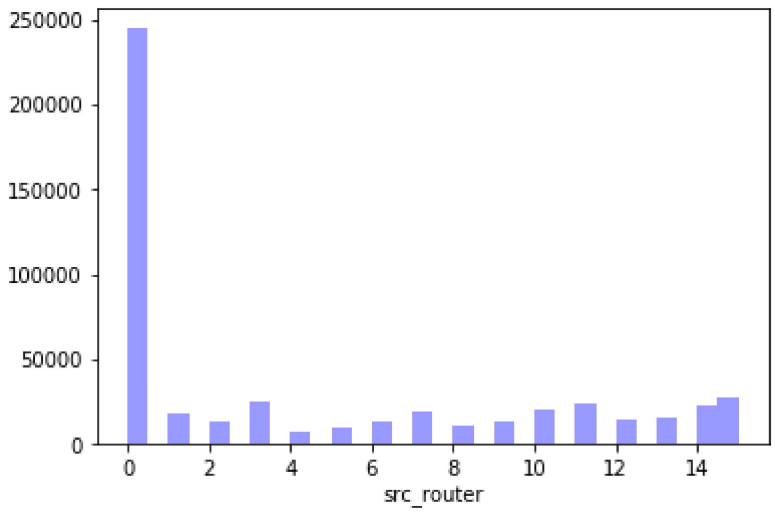


In [17]:

```
sns.distplot(df['src_router'], kde = False, bins=30, color='blue')
```

Out[17]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21086929a88>
```

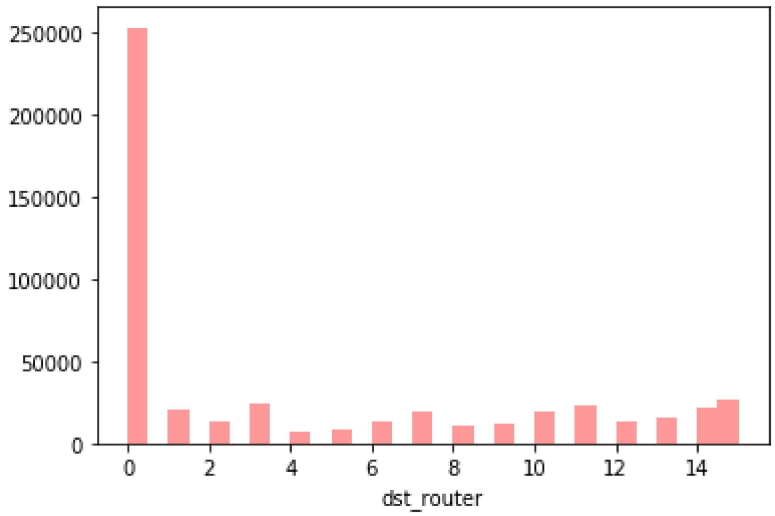


In [18]:

```
sns.distplot(df['dst_router'], kde = False, bins=30, color='red')
```

Out[18]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x210869e0608>
```

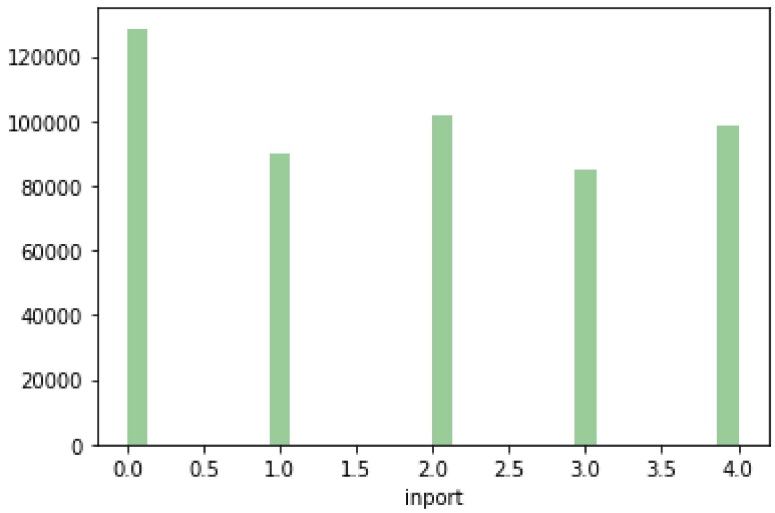


In [19]:

```
sns.distplot(df['inport'], kde = False, bins=30, color='green')
```

Out[19]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21086a8bd08>
```

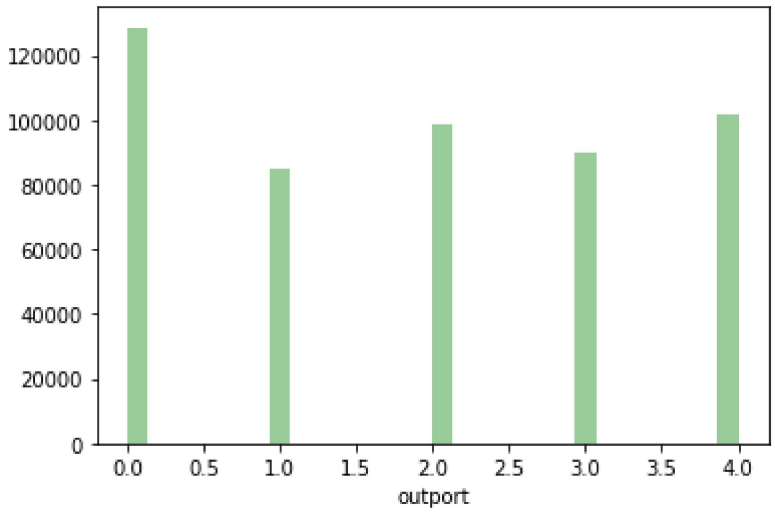


In [20]:

```
sns.distplot(df['outport'], kde = False, bins=30, color='green')
```

Out[20]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21086a9a6c8>
```

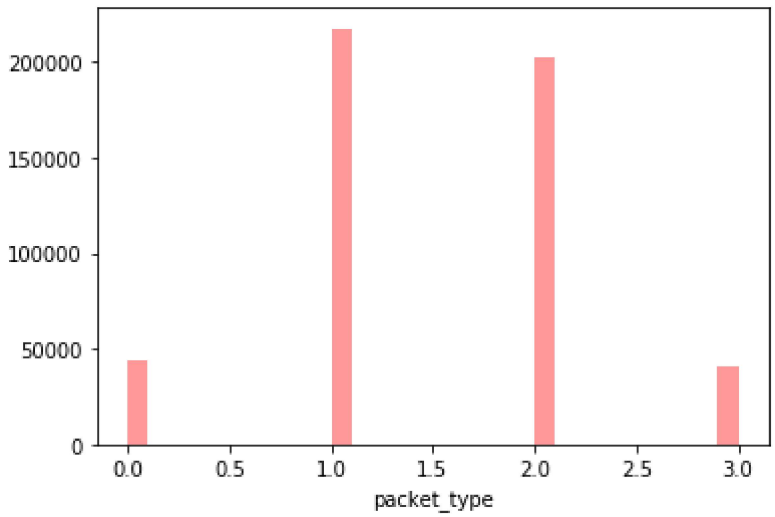


In [21]:

```
sns.distplot(df['packet_type'], kde = False, bins=30, color='red')
```

Out[21]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x21086c2f388>
```



In [22]:

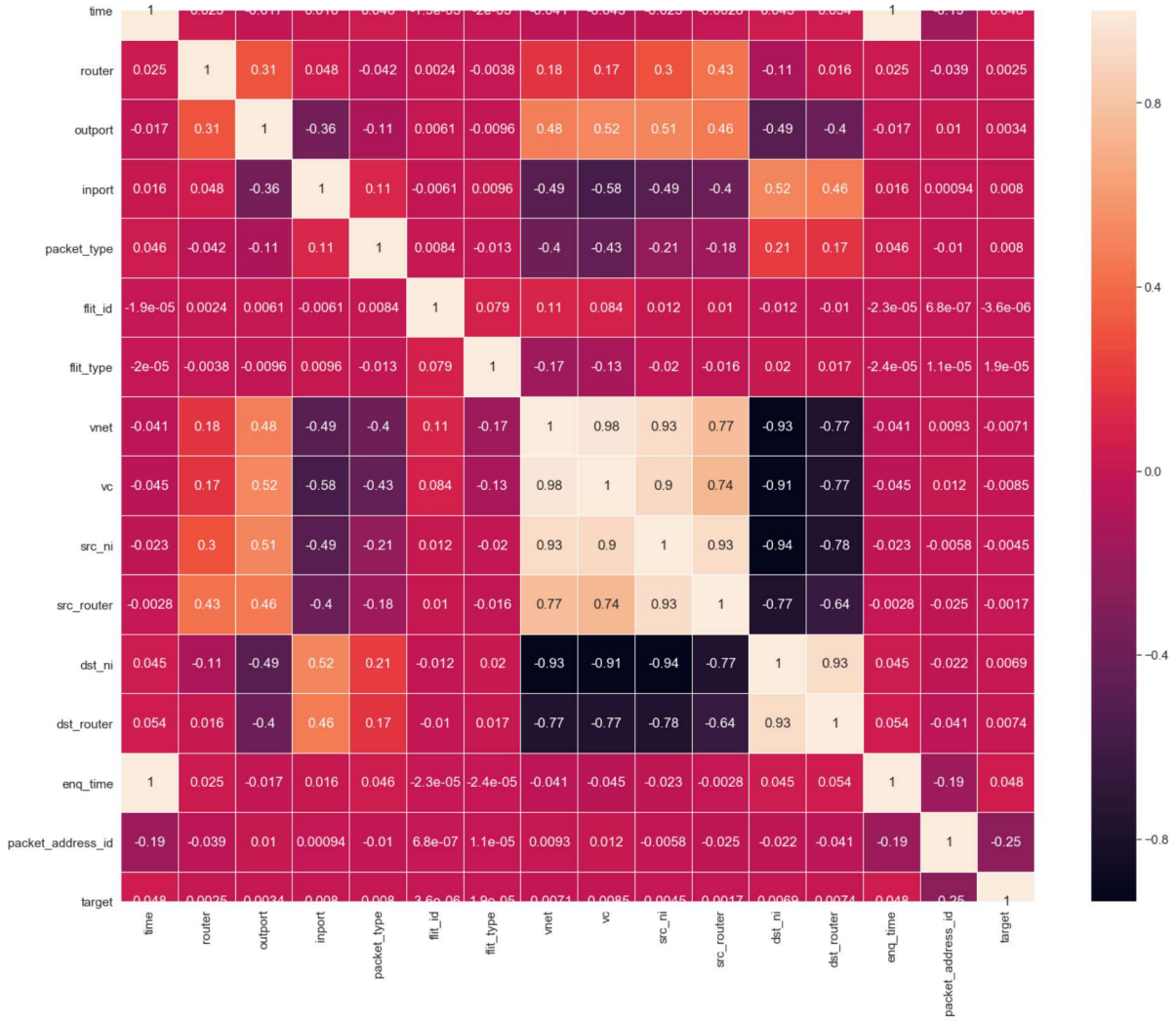
```
df.isnull().sum()
```

Out[22]:

```
time                0
router              0
outport             0
inport              0
packet_address      0
packet_type         0
flit_id             0
flit_type           0
vnet                0
vc                  0
src_ni              0
src_router          0
dst_ni              0
dst_router          0
enq_time            0
packet_address_id   0
target              0
dtype: int64
```

In [23]:

```
plt.figure(figsize=(25,20))
plt.rcParams["axes.labelsize"] = 20
sns.set(font_scale=1.4)
sns.heatmap(df.corr(), annot = True ,linewidths=.1)
plt.show()
```



In [25]:

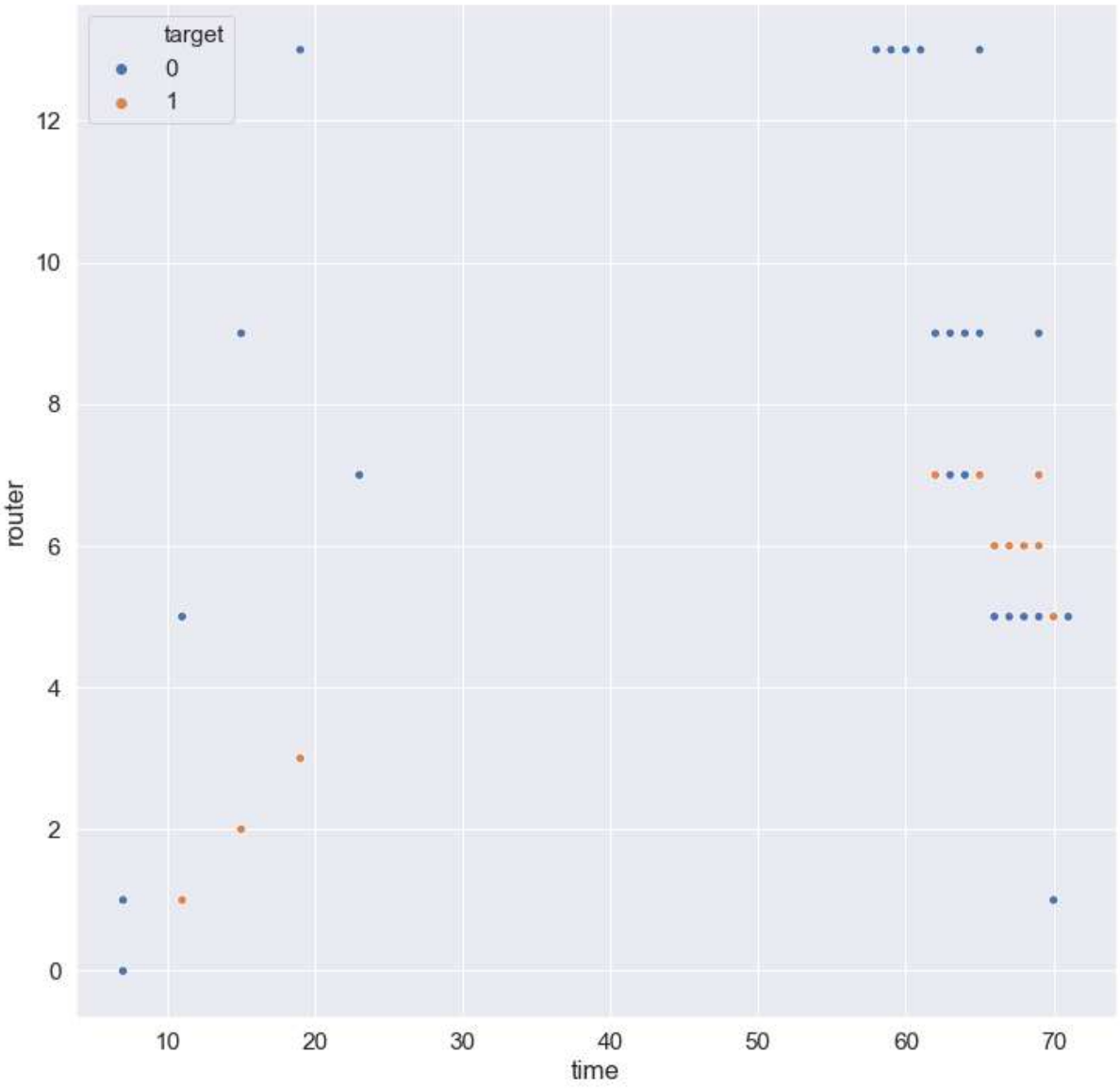
```
df.to_csv('wat-model.csv',index=False)
```

In [36]:

```
dff = pd.read_csv('wat-model.csv',nrows=50)
```

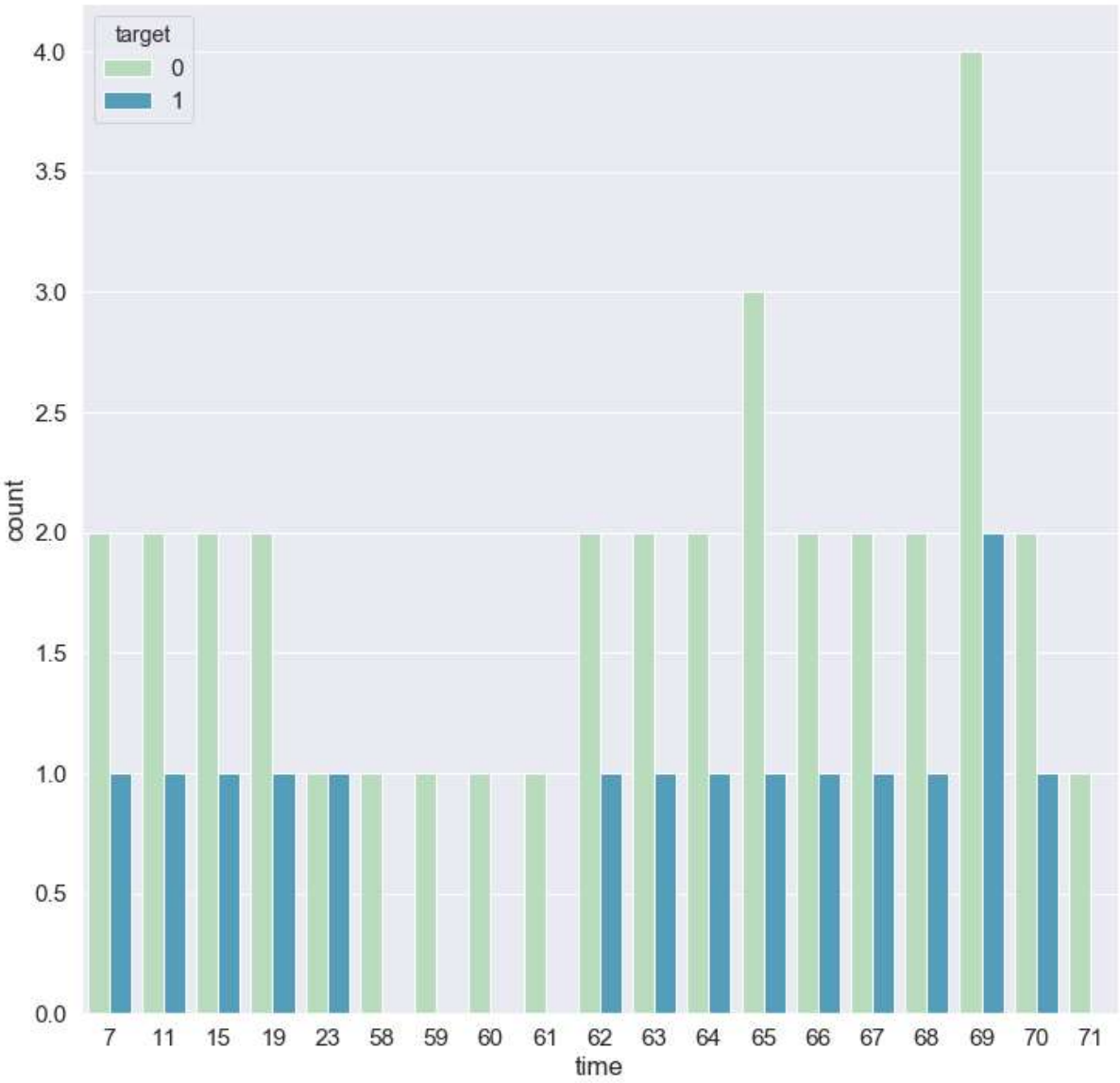
In [40]:

```
plt.figure(figsize=(12,12))
sns.scatterplot(x='time',y='router',data=dff, hue='target')
plt.show()
```



In [41]:

```
plt.figure(figsize=(12,12))
sns.countplot(x='time',data = dff, hue = 'target',palette='GnBu')
plt.show()
```



In [32]:

```
grouped = df.groupby(['time','router','target']).count().reset_index()
grouped
```

Out[32]:

	time	router	target	outport	inport	packet_address	packet_type	flit_id	flit_type	v
0	7	0	0	1	1	1	1	1	1	
1	7	0	1	1	1	1	1	1	1	
2	7	1	0	1	1	1	1	1	1	
3	11	1	0	1	1	1	1	1	1	
4	11	1	1	1	1	1	1	1	1	
...	...	...	...	...	...	...	...	...	...	
503819	3152966	8	1	1	1	1	1	1	1	
503820	3152967	8	1	1	1	1	1	1	1	
503821	3152967	9	1	1	1	1	1	1	1	
503822	3152968	4	1	1	1	1	1	1	1	
503823	3152969	4	1	1	1	1	1	1	1	

503824 rows × 17 columns

In [44]:

```
dff
```

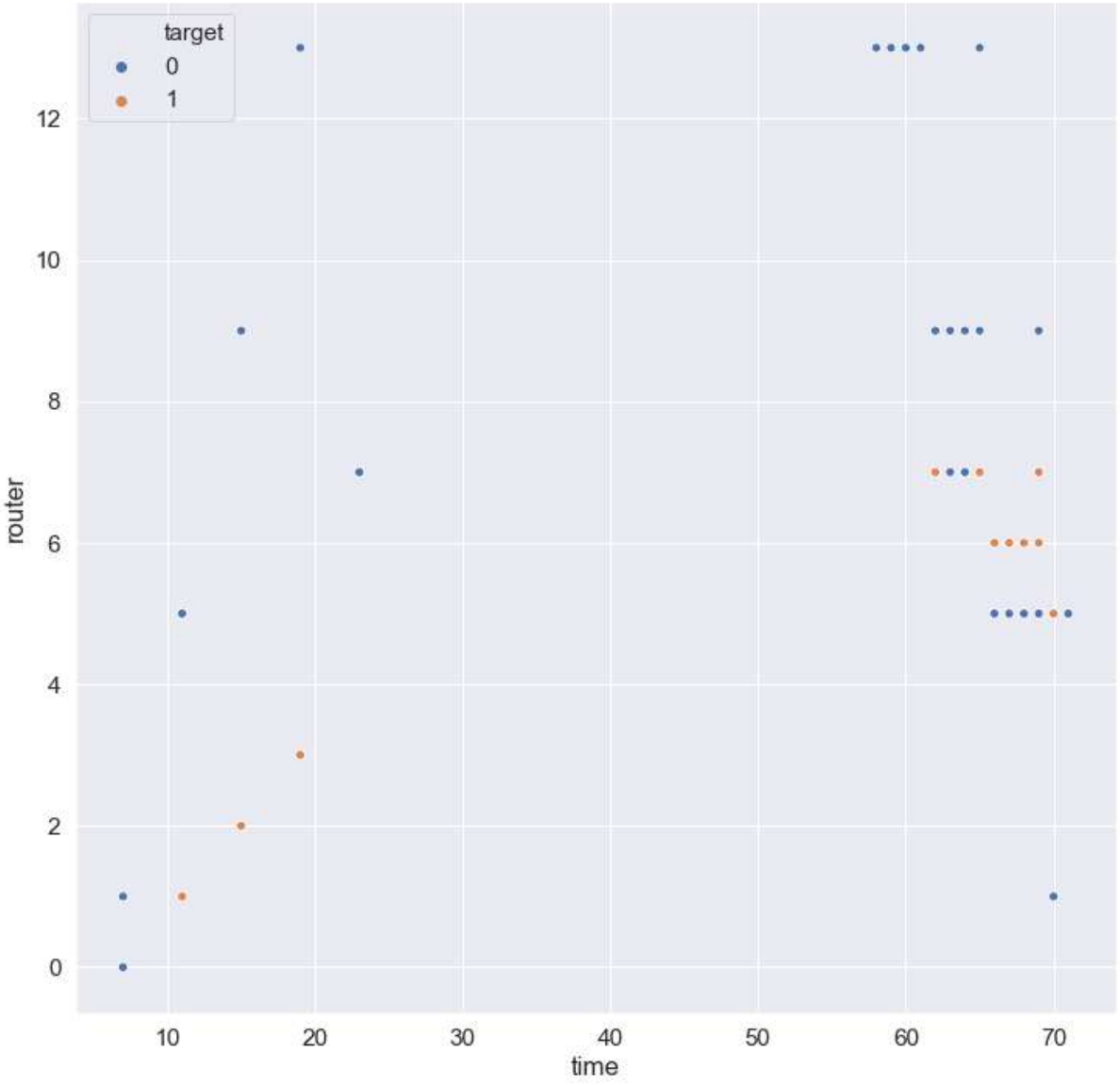
Out[44]:

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
0	7	0	2	0	0x1dc0	0	0	3	2	8	0
1	7	1	1	0	0xecf40	0	0	3	2	8	1
2	7	0	2	0	0x1dc0	0	0	3	2	8	0
3	11	5	1	3	0xecf40	0	0	3	2	8	1
4	11	1	2	4	0x1dc0	0	0	3	2	8	0
5	11	1	2	4	0x1dc0	0	0	3	2	8	0
6	15	2	2	4	0x1dc0	0	0	3	2	8	0
7	15	9	1	3	0xecf40	0	0	3	2	8	1
8	15	2	2	4	0x1dc0	0	0	3	2	8	0
9	19	13	0	3	0xecf40	0	0	3	2	8	1
10	19	3	1	4	0x1dc0	0	0	3	2	8	0
11	19	3	1	4	0x1dc0	0	0	3	2	8	0
12	23	7	0	3	0x1dc0	0	0	3	2	8	0
13	23	7	0	3	0x1dc0	0	0	3	2	8	0
14	58	13	3	0	0xecf40	1	0	0	4	16	29
15	59	13	3	0	0xecf40	1	1	1	4	16	29
16	60	13	3	0	0xecf40	1	2	1	4	16	29
17	61	13	3	0	0xecf40	1	3	1	4	16	29
18	62	9	3	1	0xecf40	1	0	0	4	16	29
19	62	7	4	0	0x1dc0	1	0	0	4	16	23
20	62	7	4	0	0x1dc0	1	0	0	4	16	23
21	63	7	4	0	0x1dc0	1	1	1	4	16	23
22	63	9	3	1	0xecf40	1	1	1	4	16	29
23	63	7	4	0	0x1dc0	1	1	1	4	16	23
24	64	7	4	0	0x1dc0	1	2	1	4	16	23
25	64	7	4	0	0x1dc0	1	2	1	4	16	23
26	64	9	3	1	0xecf40	1	2	1	4	16	29
27	65	13	3	0	0xecf40	1	4	2	4	16	29
28	65	9	3	1	0xecf40	1	3	1	4	16	29
29	65	7	4	0	0x1dc0	1	3	1	4	16	23
30	65	7	4	0	0x1dc0	1	3	1	4	16	23
31	66	6	4	2	0x1dc0	1	0	0	4	16	23
32	66	5	3	1	0xecf40	1	0	0	4	16	29
33	66	6	4	2	0x1dc0	1	0	0	4	16	23

	time	router	outport	inport	packet_address	packet_type	flit_id	flit_type	vnet	vc	src_ni
34	67	6	4	2	0x1dc0	1	1	1	4	16	23
35	67	5	3	1	0xecf40	1	1	1	4	16	29
36	67	6	4	2	0x1dc0	1	1	1	4	16	23
37	68	5	3	1	0xecf40	1	2	1	4	16	29
38	68	6	4	2	0x1dc0	1	2	1	4	16	23
39	68	6	4	2	0x1dc0	1	2	1	4	16	23
40	69	5	3	1	0xecf40	1	3	1	4	16	29
41	69	6	4	2	0x1dc0	1	3	1	4	16	23
42	69	9	3	1	0xecf40	1	4	2	4	16	29
43	69	7	4	0	0x1dc0	1	4	2	4	16	23
44	69	6	4	2	0x1dc0	1	3	1	4	16	23
45	69	7	4	0	0x1dc0	1	4	2	4	16	23
46	70	5	4	2	0x1dc0	1	0	0	4	16	23
47	70	5	4	2	0x1dc0	1	0	0	4	16	23
48	70	1	0	1	0xecf40	1	0	0	4	16	29
49	71	5	4	2	0x1dc0	1	1	1	4	16	23

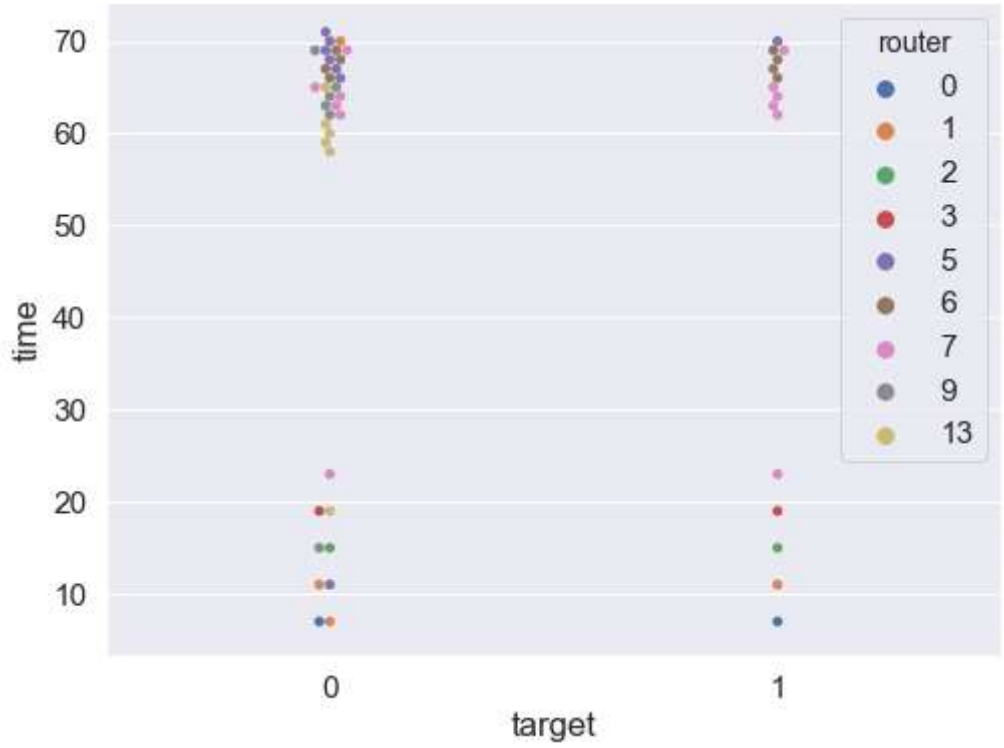
In [43]:

```
plt.figure(figsize=(12,12))
sns.scatterplot(x='time',y='router',data=dff, hue='target')
plt.show()
```


target
0
1
router
time
0
2
4
6
8
10
12
10
20
30
40
50
60
70

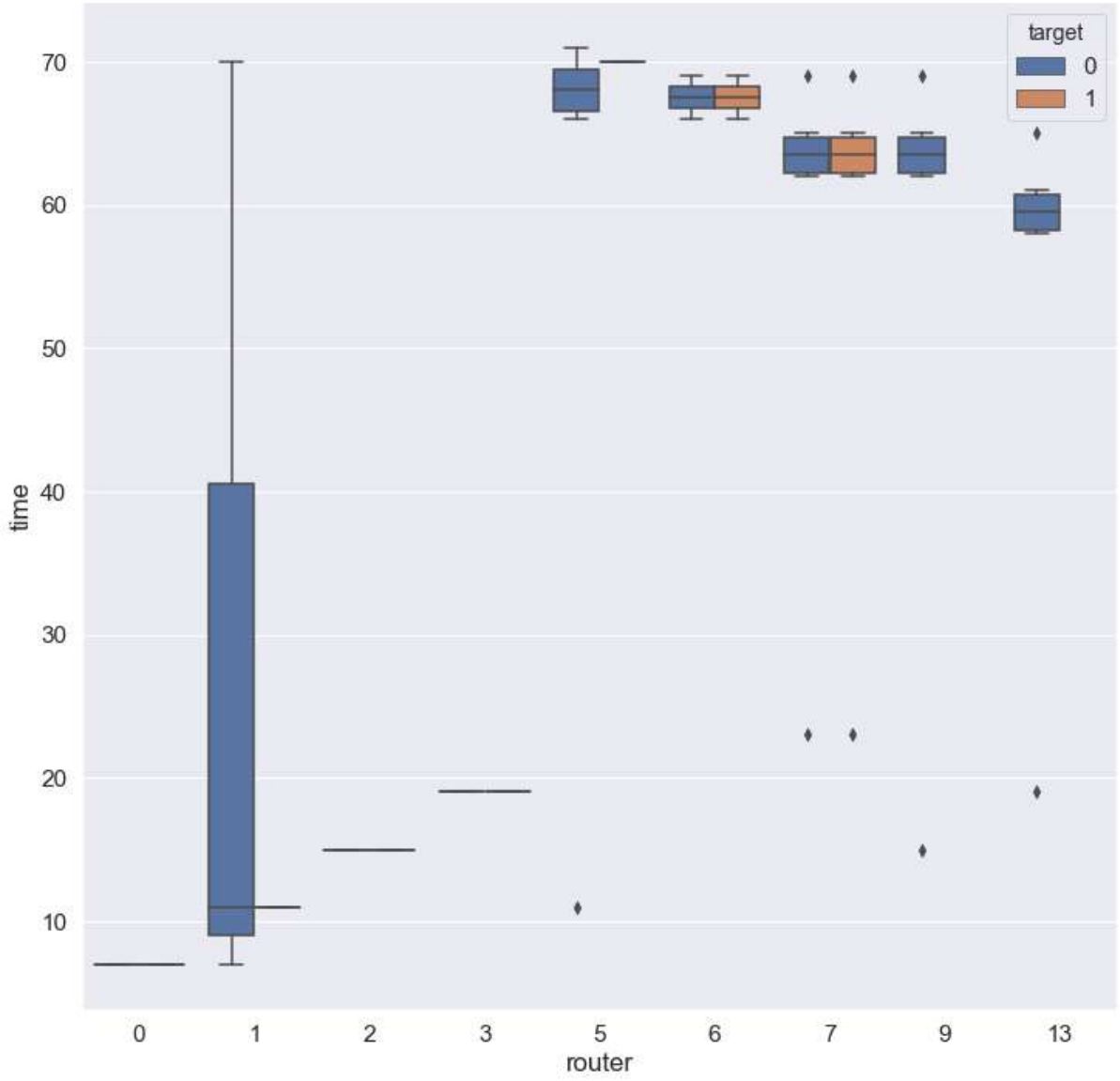
In [46]:

```
plt.figure(figsize=(8,6))
sns.swarmplot(x='target',y='time',data=dff, hue='router')
plt.show()
```



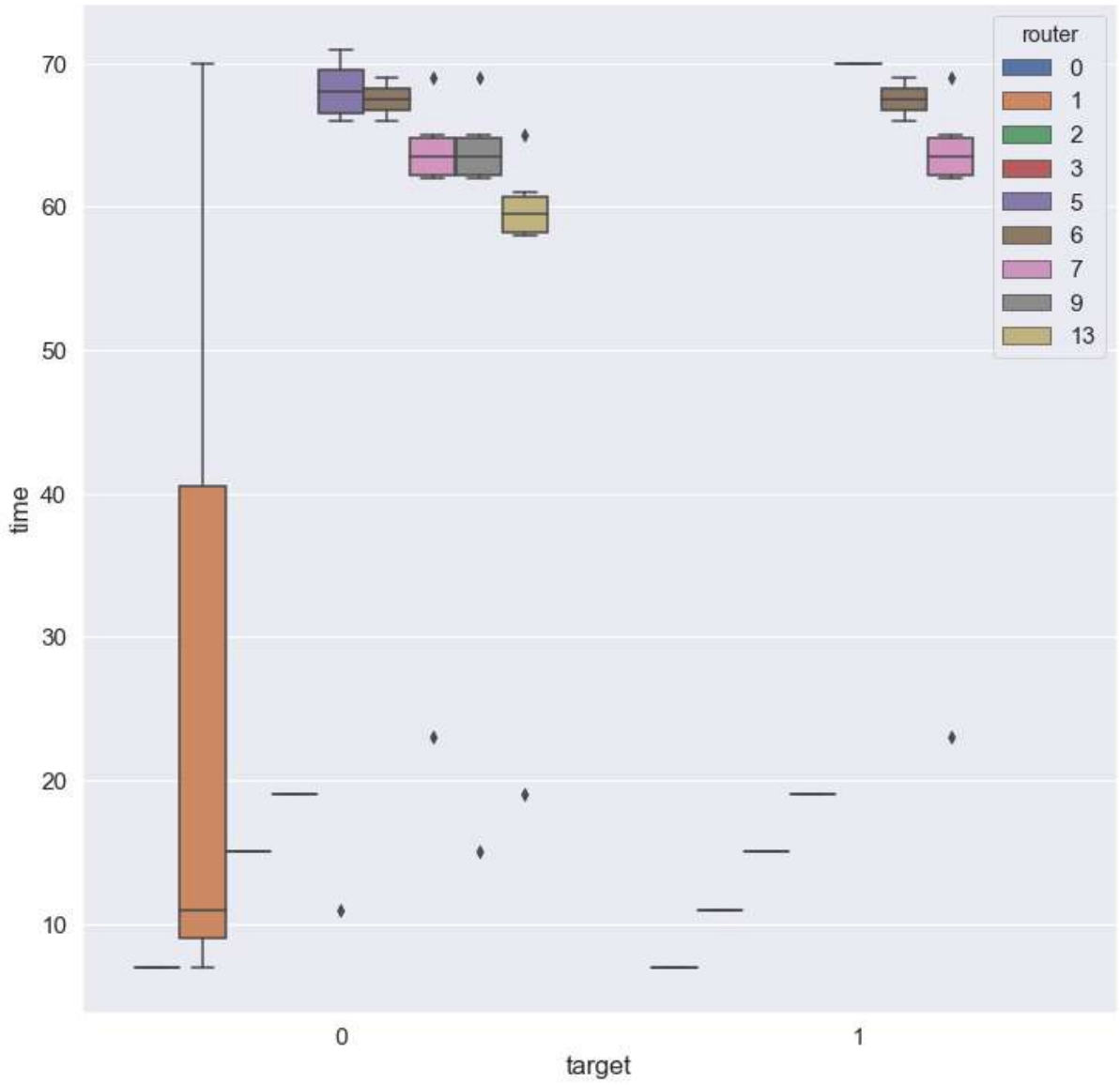
In [50]:

```
plt.figure(figsize=(12,12))
sns.boxplot(x='router',y='time',data=dff, hue='target')
plt.show()
```


target
0
1
70
60
50
40
30
20
10
time
0
1
2
3
5
6
7
9
13
router

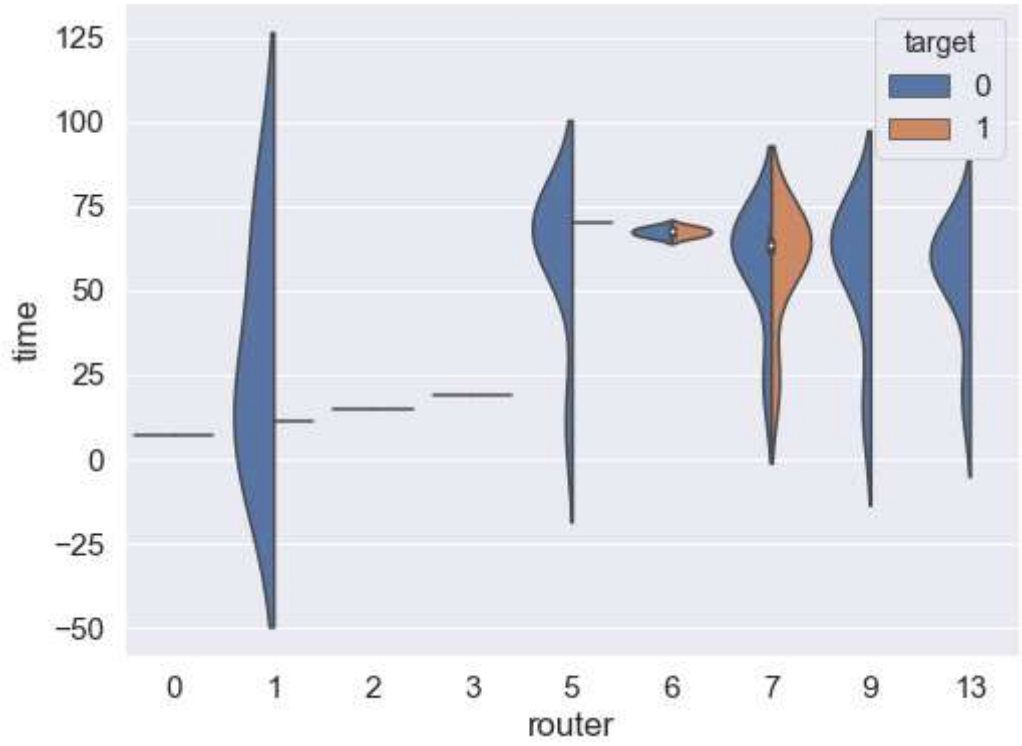
In [51]:

```
plt.figure(figsize=(12,12))
sns.boxplot(x='target',y='time',data=dff, hue='router')
plt.show()
```



In [52]:

```
plt.figure(figsize=(8,6))
sns.violinplot(x='router',y='time',data=dff, hue='target', split=True)
plt.show()
```



In [53]:

```
df.query('target == "1"').target.count()
```

Out[53]:

245667

In [54]:

```
df.query('target == "0"').target.count()
```

Out[54]:

258373