In [1]:

```python
#extracting lines for simplied verion
open('waddr.txt','w').writelines([ line for line in open("waddr-out.log") if "Enqueue" in
line])
print ("done")
```

done

In [2]:

```python
#extracting content from lines
csv_out = open('waddr_csv.txt','w')
with open ('waddr.txt', 'rt') as fft:
    csv_out.write("time,router,outport,inport,packet,flit_id,type,vnet,vc,src_ni,src_route
r,dst_ni,dst_router,enq_time\n")
    for line in fft:
        line_split = line.split()
        time = line_split[line_split.index("time:") + 1]
        router = line_split[line_split.index("SwitchAllocator") + 3]
        outport = line_split[line_split.index("outport") + 1]
        inport = line_split[line_split.index("inport") + 1]
        packet = line_split[line_split.index("addr") + 2][3:-1]
        flit_id = line_split[line_split.index("[flit::") + 1][3:]
        type = line_split[line_split.index("Id="+str(flit_id)) + 1][5:]
        vnet = line_split[line_split.index("Type="+str(type)) + 1][5:]
        vc = line_split[line_split.index("Vnet="+str(vnet)) + 1][3:]
        src_ni = line_split[line_split.index("VC="+str(vc)) + 2][3:]
        src_router = line_split[line_split.index("NI="+str(src_ni)) + 2][7:]
        dst_ni = line_split[line_split.index("Router="+str(src_router)) + 2][3:]
        dst_router = line_split[line_split.index("NI="+str(dst_ni)) + 2][7:]
        enq_time = str(line_split[line_split.index("Enqueue") + 1][5:])
        line_csv = time+","+router+","+outport+","+inport+","+packet+","+flit_id+","+type+
","+vnet+","+vc+","+src_ni+","+src_router+","+dst_ni+","+dst_router+","+enq_time+"\n"
        csv_out.write(line_csv)
print ("done")
```

done

In [3]:

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

df = pd.read_csv("waddr_csv.txt",delimiter=',')
df.to_csv('waddr.csv',index=False)
```

In [4]:

```python
#dataset
df = pd.read_csv('waddr.csv')

print(df.head())
```

```
    time   router  outport  inport  packet  flit_id  type  vnet  vc  src_ni  \
0      7        0     East   Local    1dc0        0     3     2   8       0
1     11        1     East    West    1dc0        0     3     2   8       0
2     15        2     East    West    1dc0        0     3     2   8       0
3     19        3    North    West    1dc0        0     3     2   8       0
4     23        7    Local   South    1dc0        0     3     2   8       0

    src_router  dst_ni  dst_router  enq_time
0            0      23           7         3
1            0      23           7         3
2            0      23           7         3
3            0      23           7         3
4            0      23           7         3
```

In [5]:

```python
df.dtypes
```

Out[5]:

```
time          int64
router        int64
outport       object
inport        object
packet        object
flit_id       int64
type          int64
vnet          int64
vc            int64
src_ni        int64
src_router    int64
dst_ni        int64
dst_router    int64
enq_time      int64
dtype: object
```

In [6]:

```python
df.shape
```

Out[6]:

```
(245551, 14)
```

In [7]:

```
df.describe()
```

Out[7]:

|	time	router	flit_id	type	vnet	v
count	2.455510e+05	245551.000000	245551.000000	245551.000000	245551.000000	245551.000000
mean	1.732599e+06	4.370021	1.666644	1.333332	2.935411	12.145505
std	1.059087e+06	4.384976	1.490712	0.942818	0.956338	3.887423
min	7.000000e+00	0.000000	0.000000	0.000000	2.000000	8.000000
25%	7.219425e+05	0.000000	0.000000	1.000000	2.000000	8.000000
50%	1.802783e+06	3.000000	1.000000	1.000000	3.000000	12.000000
75%	2.705146e+06	8.000000	3.000000	2.000000	4.000000	16.000000
max	3.152608e+06	15.000000	4.000000	3.000000	4.000000	19.000000

In [8]:

```
import matplotlib.pyplot as plt
import seaborn as sns
```

In [9]:

```
sns.distplot(df['router'], kde = False, bins=30, color='blue')
```
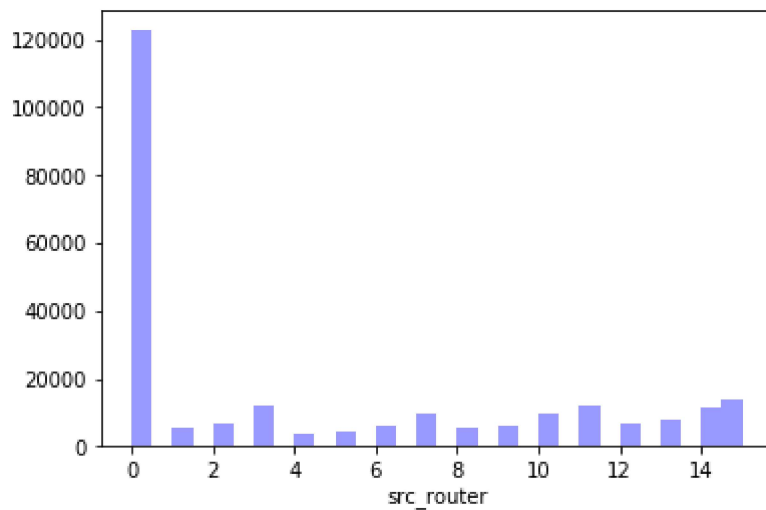
Out[9]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a0f60c8a88>
```
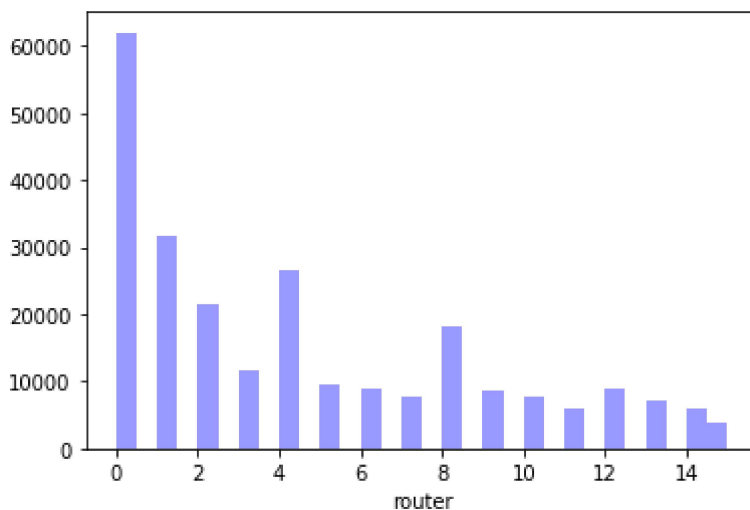
In [10]:

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

Out[10]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a0fb660948>
```


In [11]:

```
sns.distplot(df['dst_router'], kde = False, bins=30, color='red')
```
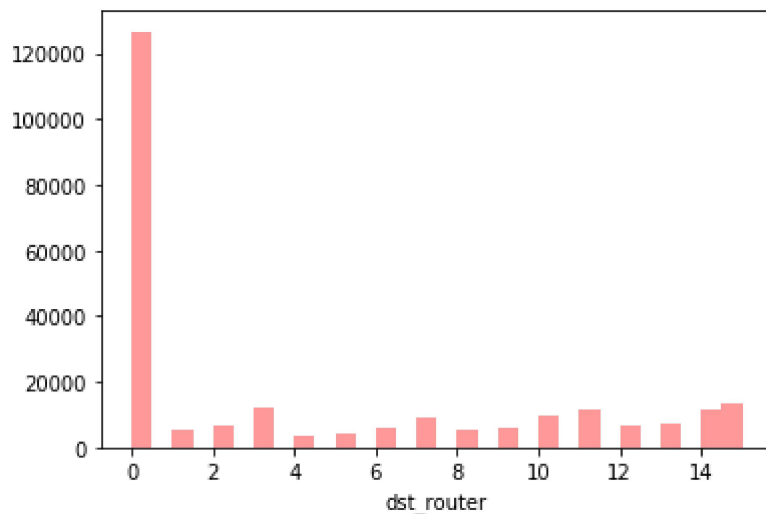
Out[11]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a0fb75c4c8>
```

In [12]:

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

Out[12]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a0fb81f9c8>
```
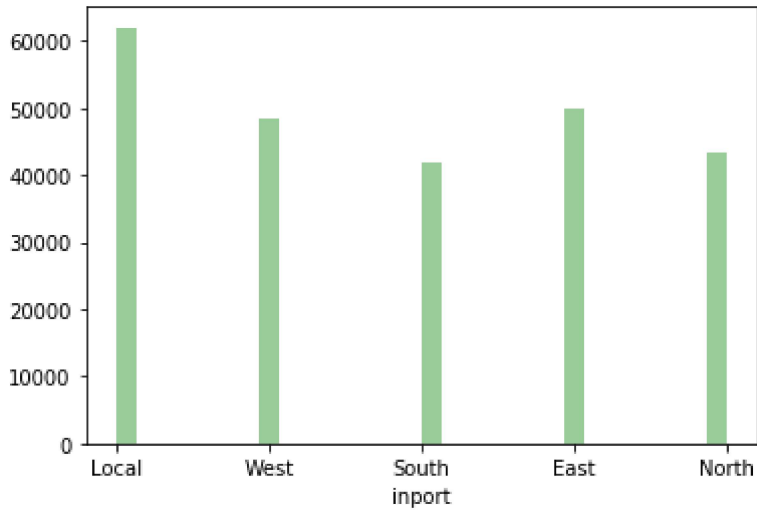

In [13]:

```python
sns.distplot(df['outport'], kde = False, bins=30, color='green')
```
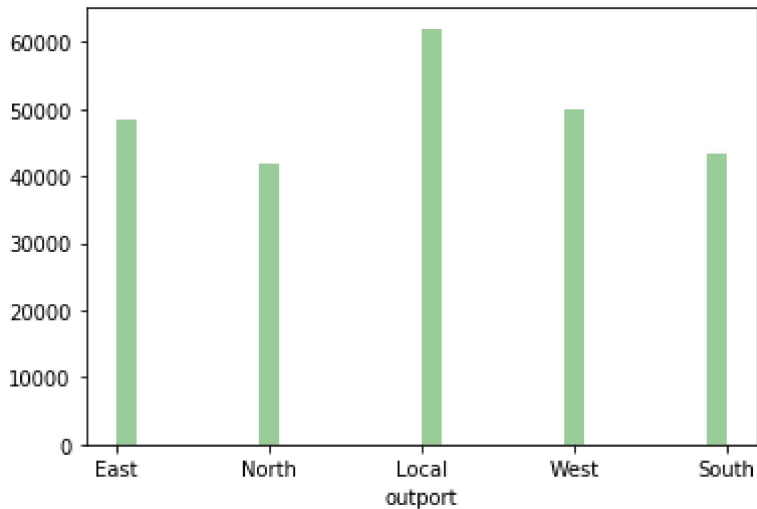
Out[13]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a0fb8c8c48>
```


In [41]:

```python
direction = {'Local': 0,'North': 1, 'East': 2, 'South':3,'West':4}
df = df.replace({'inport': direction, 'outport': direction})
```

In [42]:

```
df.dtypes
```

Out[42]:

```
time           int64
router         int64
outport        int64
inport         int64
packet        object
flit_id        int64
type           int64
vnet           int64
vc             int64
src_ni         int64
src_router     int64
dst_ni         int64
dst_router     int64
enq_time       int64
packet_id      int32
dtype: object
```

In [43]:

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

Out[43]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1a084b04848>
```
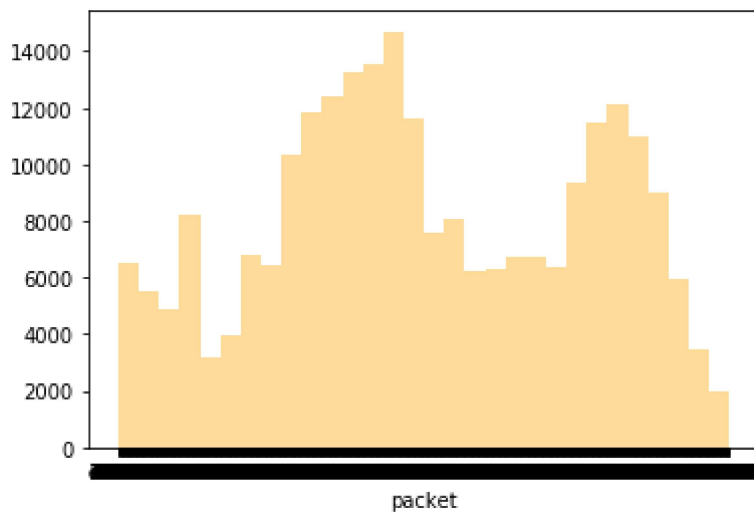

In [44]:

```
from sklearn.preprocessing import LabelEncoder
```

In [30]:

```
dff = pd.DataFrame(df['packet'], columns=['packet'])
```

In [33]:

```
le = LabelEncoder()
encode = dff[dff.columns[:]].apply(le.fit_transform)
```

In [34]:

```
df['packet_id'] = encode
```

In [35]:

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

Out[35]:

```
time            0
router          0
outport         0
inport          0
packet          0
flit_id         0
type            0
vnet            0
vc              0
src_ni          0
src_router      0
dst_ni          0
dst_router      0
enq_time        0
packet_id       0
dtype: int64
```

In [36]:

```
df.corr()
```

Out[36]:

	time	router	flit_id	type	vnet	vc	src_ni	src
time	1.000000e+00	0.035270	-0.000015	-7.662813e-07	-0.031665	-0.034442	-0.012504	0
router	3.526971e-02	1.000000	0.001808	-2.836210e-03	0.179307	0.169624	0.296786	0
flit_id	-1.520504e-05	0.001808	1.000000	7.905007e-02	0.104721	0.080540	0.009140	0
type	-7.662813e-07	-0.002836	0.079050	1.000000e+00	-0.165606	-0.127370	-0.014462	-0
vnet	-3.166479e-02	0.179307	0.104721	-1.656063e-01	1.000000	0.977442	0.930520	0
vc	-3.444174e-02	0.169624	0.080540	-1.273700e-01	0.977442	1.000000	0.902745	0
src_ni	-1.250373e-02	0.296786	0.009140	-1.446199e-02	0.930520	0.902745	1.000000	0
src_router	1.022479e-02	0.423714	0.007784	-1.229525e-02	0.771852	0.739236	0.927483	1
dst_ni	4.031869e-02	-0.111109	-0.009145	1.448329e-02	-0.931291	-0.914169	-0.936383	-0
dst_router	5.539027e-02	0.013340	-0.007813	1.237419e-02	-0.775392	-0.767836	-0.779631	-0
enq_time	1.000000e+00	0.035270	-0.000019	-5.171403e-06	-0.031659	-0.034435	-0.012498	0
packet_id	-2.649712e-01	-0.042977	-0.000004	-3.567544e-07	0.016125	0.018486	-0.002042	-0

In [45]:

```
df.dtypes
```

Out[45]:

```
time            int64
router          int64
outport         int64
inport          int64
packet         object
flit_id         int64
type            int64
vnet            int64
vc              int64
src_ni          int64
src_router      int64
dst_ni          int64
dst_router      int64
enq_time        int64
packet_id       int32
dtype: object
```

In [56]:

```python
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()
```

	time	router	outport	inport	flit_id	type	vnet	vc	src_ni	src_router	dst_ni	dst_router	enq_time	packet_id
time	1	0.035	-0.02	0.0073	-1.5e-05	-7.7e-07	-0.032	-0.034	-0.013	0.01	0.04	0.055	1	-0.26
router	0.035	1	0.31	0.045	0.0018	-0.0028	0.18	0.17	0.3	0.42	-0.11	0.013	0.035	-0.043
outport	-0.02	0.31	1	-0.36	0.0048	-0.0075	0.49	0.52	0.51	0.46	-0.49	-0.41	-0.02	0.016
inport	0.0073	0.045	-0.36	1	-0.0048	0.0076	-0.49	-0.58	-0.49	-0.41	0.52	0.47	0.0073	0.0024
flit_id	-1.5e-05	0.0018	0.0048	-0.0048	1	0.079	0.1	0.081	0.0091	0.0078	-0.0091	-0.0078	-1.9e-05	-3.8e-06
type	-7.7e-07	-0.0028	-0.0075	0.0076	0.079	1	-0.17	-0.13	-0.014	-0.012	0.014	0.012	-5.2e-06	-3.6e-07
vnet	-0.032	0.18	0.49	-0.49	0.1	-0.17	1	0.98	0.93	0.77	-0.93	-0.78	-0.032	0.016
vc	-0.034	0.17	0.52	-0.58	0.081	-0.13	0.98	1	0.9	0.74	-0.91	-0.77	-0.034	0.018
src_ni	-0.013	0.3	0.51	-0.49	0.0091	-0.014	0.93	0.9	1	0.93	-0.94	-0.78	-0.012	-0.002
src_router	0.01	0.42	0.46	-0.41	0.0078	-0.012	0.77	0.74	0.93	1	-0.78	-0.65	0.01	-0.025
dst_ni	0.04	-0.11	-0.49	0.52	-0.0091	0.014	-0.93	-0.91	-0.94	-0.78	1	0.93	0.04	-0.029
dst_router	0.055	0.013	-0.41	0.47	-0.0078	0.012	-0.78	-0.77	-0.78	-0.65	0.93	1	0.055	-0.049
enq_time	1	0.035	-0.02	0.0073	-1.9e-05	-5.2e-06	-0.032	-0.034	-0.012	0.01	0.04	0.055	1	-0.26
packet_id	-0.26	-0.043	0.016	0.0024	-3.8e-06	-3.6e-07	0.016	0.018	-0.002	-0.025	-0.029	-0.049	-0.26	1

In [62]:

```python
df_v2 = df[:]
```

In [64]:

```python
df_v2.to_csv('waddr_v2.csv',index=False)
```

In [65]:

```python
dff_500 = pd.read_csv('waddr_v2.csv',nrows=500)
plt.figure(figsize=(8,6))
sns.scatterplot(x='time',y='router',data=dff_500)
plt.show()
```