In this notebook, we will process data on network usage obtained from monitoring the nodes.

In [1]:
```
import pandas as pd
import numpy as np
import plotly.express as px
import plotly.graph_objects as go
pd.set_option('max_columns', None)
```

In [2]:
```
def getDataframeFromNethogsOutput(df):
    df['timestamp'] = pd.to_datetime(df[0].str.strip('['))
    df[['sent','received']] = df[1].str.split(' ', expand=True)[[2,3]].astype(float)
    df = df.drop([0,1],axis=1)
    return df
```

In [3]:
```
def getAverageByMinute(df):
    df1 = df.copy().drop(['timestamp'],axis=1)
    # convert to datetime
    df1['datetime'] = pd.to_datetime(df['timestamp'])
    df1 = df1.set_index('datetime')
    # get average in every minute
    df1 = df1.resample('1T').mean()
    df1 = df1.reset_index()
    df1 = df1.rename(columns={"datetime": "timestamp"})
    return df1
```

1205_nethogs_ipfs.txt:

```
[2022-12-05 00:00:02] ipfs/17286/1002 0.141797 0.130273
[2022-12-05 00:00:12] ipfs/17286/1002 0.113672 0.0820312
[2022-12-05 00:00:22] ipfs/17286/1002 0.0632813 0.0375
[2022-12-05 00:00:32] ipfs/17286/1002 0.0832031 0.0824219
[2022-12-05 00:00:42] ipfs/17286/1002 0.608594 1.49453
...
```

In [4]:
```
df1 = pd.read_csv('1205_nethogs_ipfs.txt', sep=']', header=None)

df1 = getDataframeFromNethogsOutput(df1)
df1 = getAverageByMinute(df1)
df1['node'] = 'ipfs'

df1.shape
```

Out[4]: (1440, 4)

In [5]:
```
df1.describe()
```

Out[5]:

	sent	received
count	1287.000000	1287.000000
mean	0.566173	1.103379
std	0.828065	1.801272
min	0.000000	0.000000
25%	0.044043	0.041016
50%	0.165234	0.339727
75%	0.768099	1.426545
max	6.863254	18.315500

1205_nethogs_swarm.txt:

```
[2022-12-05 00:00:02] /usr/bin/bee/15233/998 2.22402 1.86152
[2022-12-05 00:00:12] /usr/bin/bee/15233/998 2.25742 1.92031
[2022-12-05 00:00:22] /usr/bin/bee/15233/998 2.29082 1.9791
[2022-12-05 00:00:32] /usr/bin/bee/15233/998 1.96348 2.18848
[2022-12-05 00:00:42] /usr/bin/bee/15233/998 2.05859 1.77598
...
```

In [6]:
```
df2 = pd.read_csv('1205_nethogs_swarm.txt', sep=']', header=None)

df1 = getDataframeFromNethogsOutput(df2)
df2 = getAverageByMinute(df2)
df2['node'] = 'swarm'

df2.shape
```

Out[6]: (1440, 4)

In [7]:
```
df2.describe()
```

Out[7]:

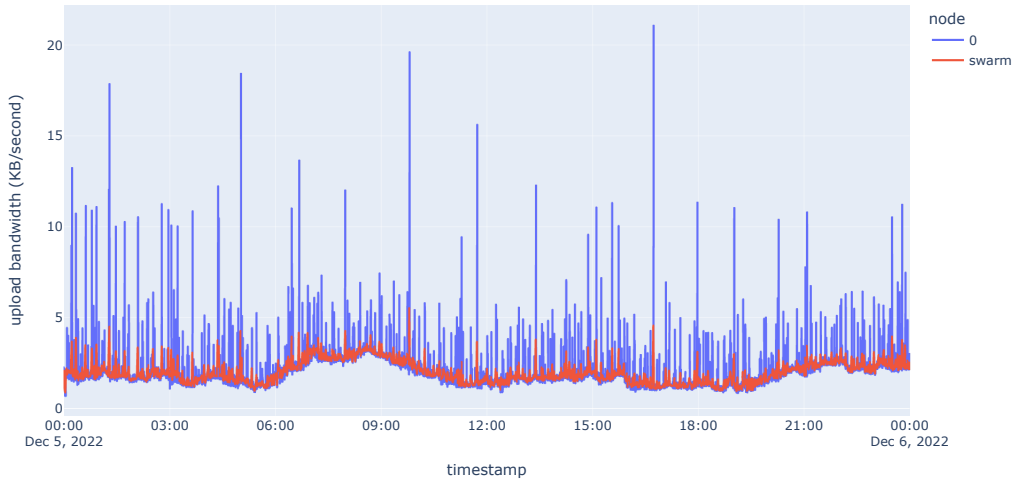
	sent	received
count	1440.000000	1440.000000
mean	2.029880	1.986188
std	0.624513	0.769133
min	0.920346	0.884375
25%	1.596063	1.466309
50%	1.887224	1.776709
75%	2.390829	2.391341
max	5.546067	10.026823

In [8]:
```
df = pd.concat([df1, df2])
df = df.fillna(0)
df.head()
```

Out[8]:

	timestamp	sent	received	node
0	2022-12-05 00:00:02	2.22402	1.86152	0
1	2022-12-05 00:00:12	2.25742	1.92031	0
2	2022-12-05 00:00:22	2.29082	1.97910	0
3	2022-12-05 00:00:32	1.96348	2.18848	0
4	2022-12-05 00:00:42	2.05859	1.77598	0

In [9]:
```
fig = px.line(df, x='timestamp', y='sent', color='node')
fig.update_yaxes(title="upload bandwidth (KB/second)")
fig.show()
```



In [10]:
```
fig = px.line(df, x='timestamp', y='received', color='node')
fig.update_yaxes(title="download bandwidth (KB/second)")
fig.show()
```