# NS-3模拟数据中心实验报告

李博杰 SA14011026

# 代码说明

.

├── NS-3-report.docx

├── pcap

│   ├── inter-cluster-10-1.pcap

│   ├── inter-cluster-1-1.pcap

│   ├── inter-cluster-4-1.pcap

│   ├── inter-cluster-7-1.pcap

│   ├── inter-cluster-merged.pcap

│   ├── many-to-one-1-1.pcap

│   ├── two-core-inter-cluster-10-1.pcap

│   ├── two-core-inter-cluster-1-1.pcap

│   ├── two-core-inter-cluster-4-1.pcap

│   ├── two-core-inter-cluster-7-1.pcap

│   ├── two-core-inter-cluster-merged.pcap

│   └── two-core-many-to-one-1-1.pcap

├── README

├── src

│   ├── inter-cluster.cc

│   ├── many-to-one.cc

│   ├── two-core-inter-cluster.cc

│   └── two-core-many-to-one.cc

pcap 目录下是抓包文件

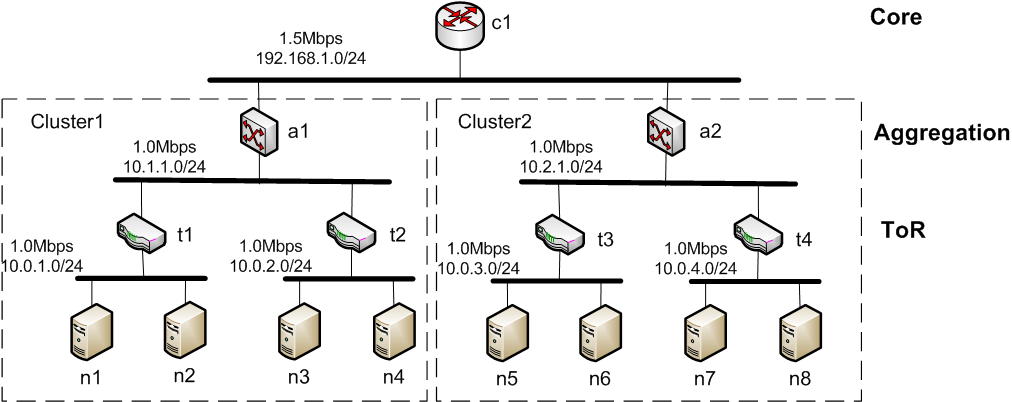
src 目录下是源代码文件

NS-3-report.docx 是实验报告

inter-cluster 和 many-to-one 是两种拓扑结构

two-core-\* 是增加 core switch 和采用 ECMP Routing 后的数据中心拓扑结构

实验拓扑结构：

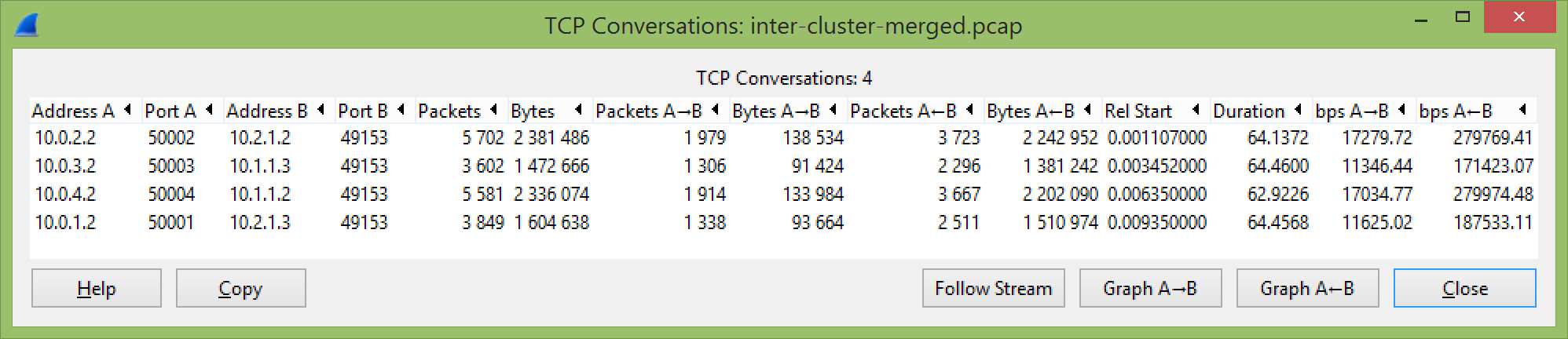


# Inter-Cluster

测试方式：跨aggregation switch的四条TCP流

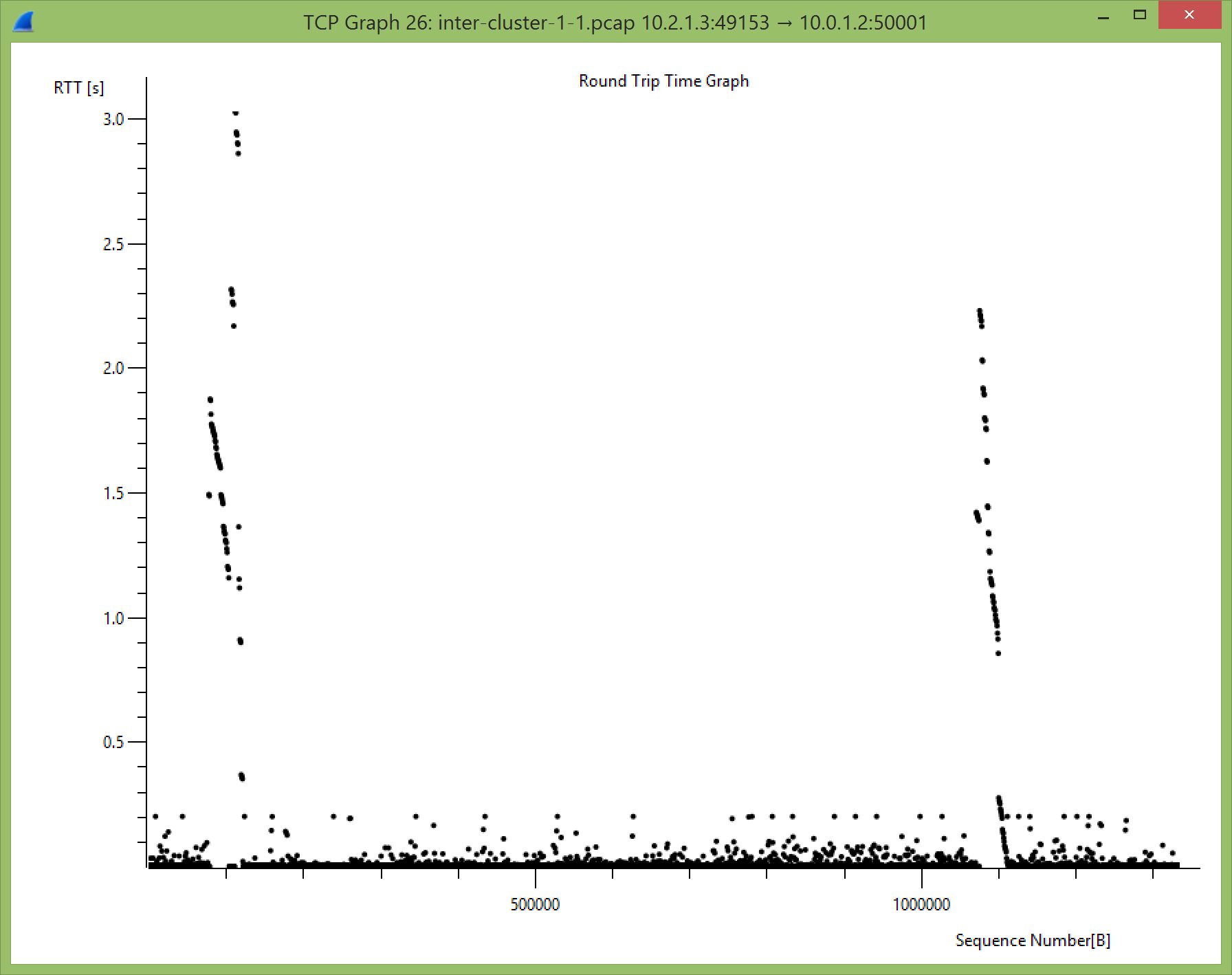
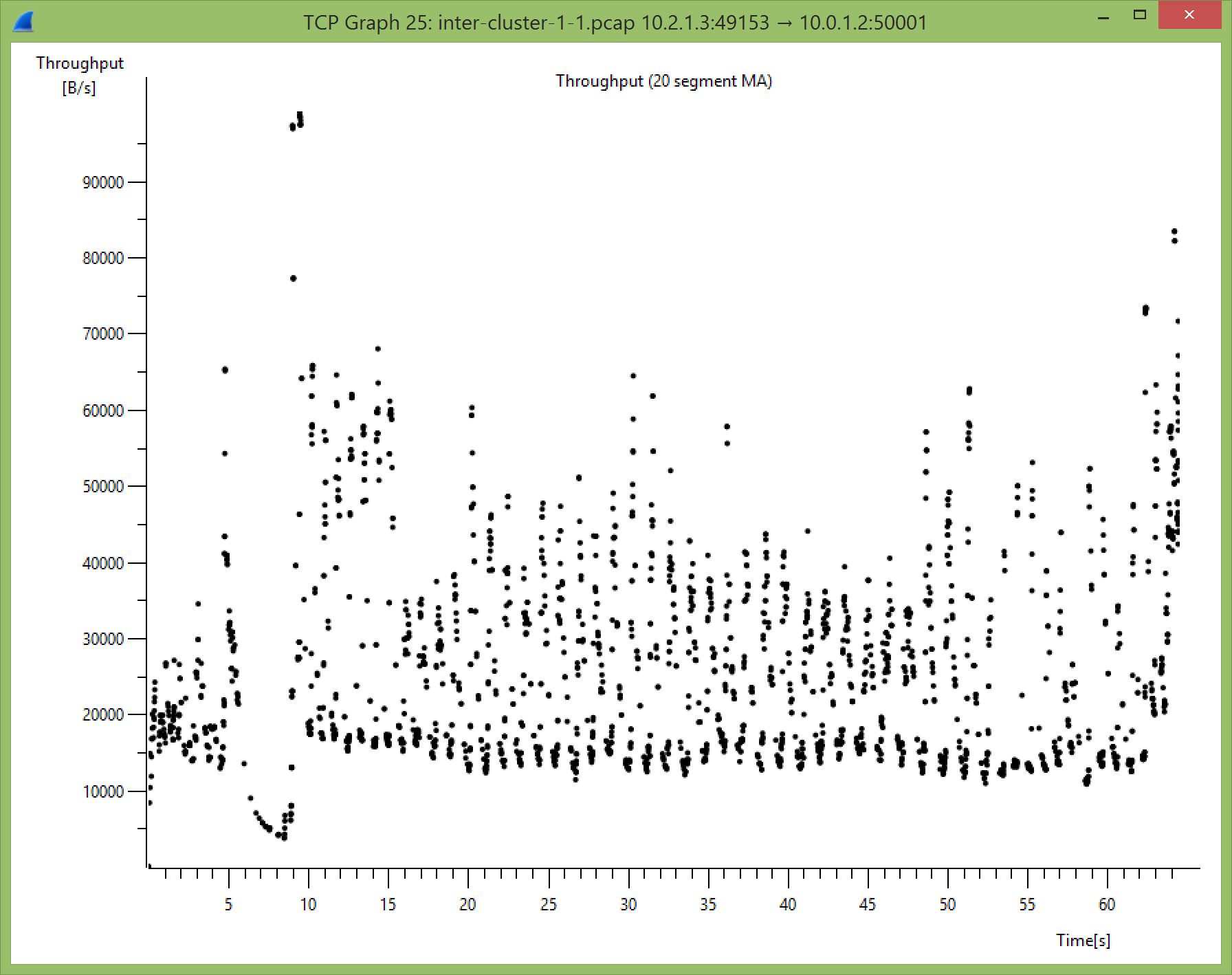
1. H7 => H2
2. H5 => H4
3. H3 => H6
4. H1 => H8

每条流均从0秒开始，60秒结束，测试时间持续60秒。

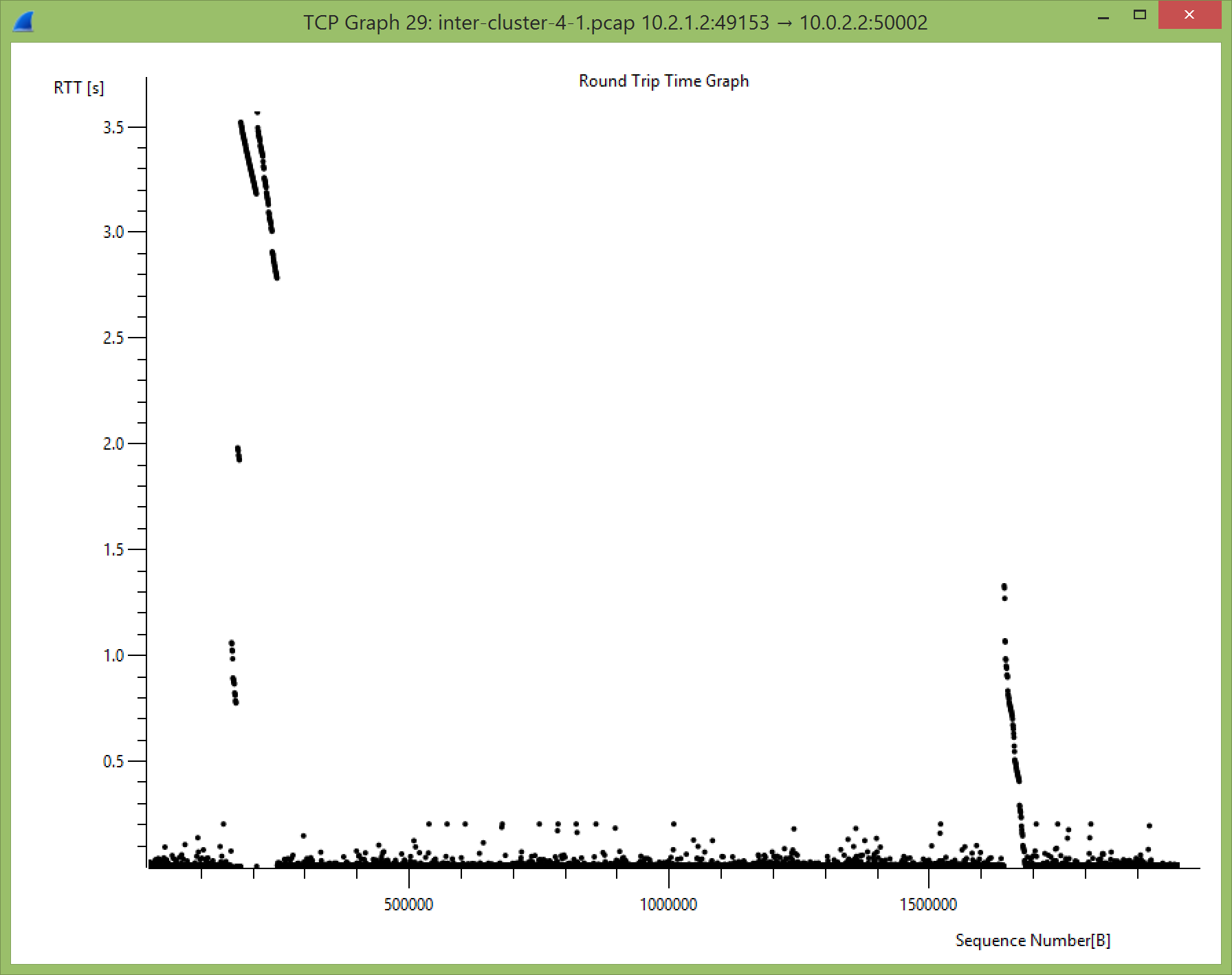
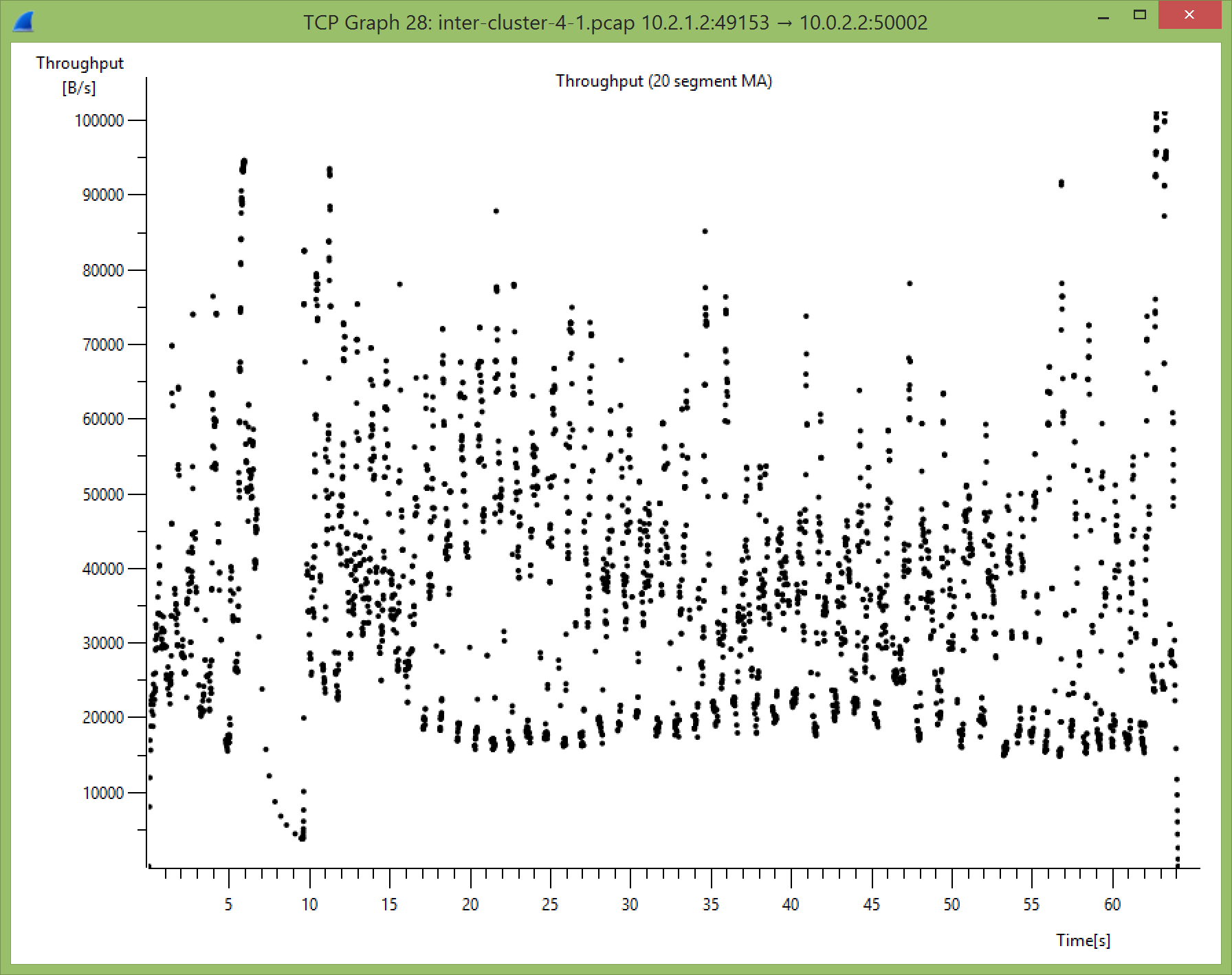


下图中，左侧是吞吐量与时间的关系，右侧是RTT与时间的关系。

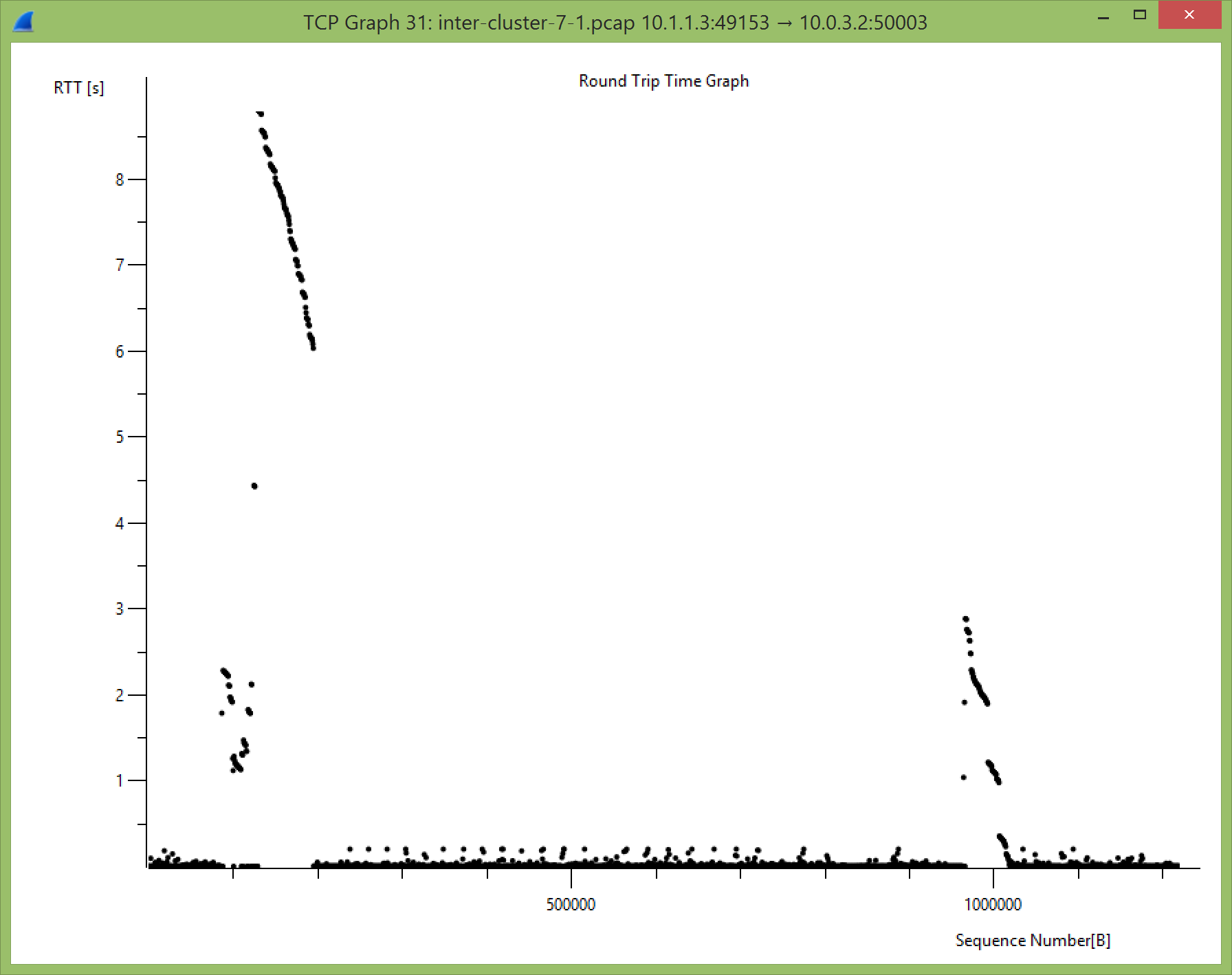
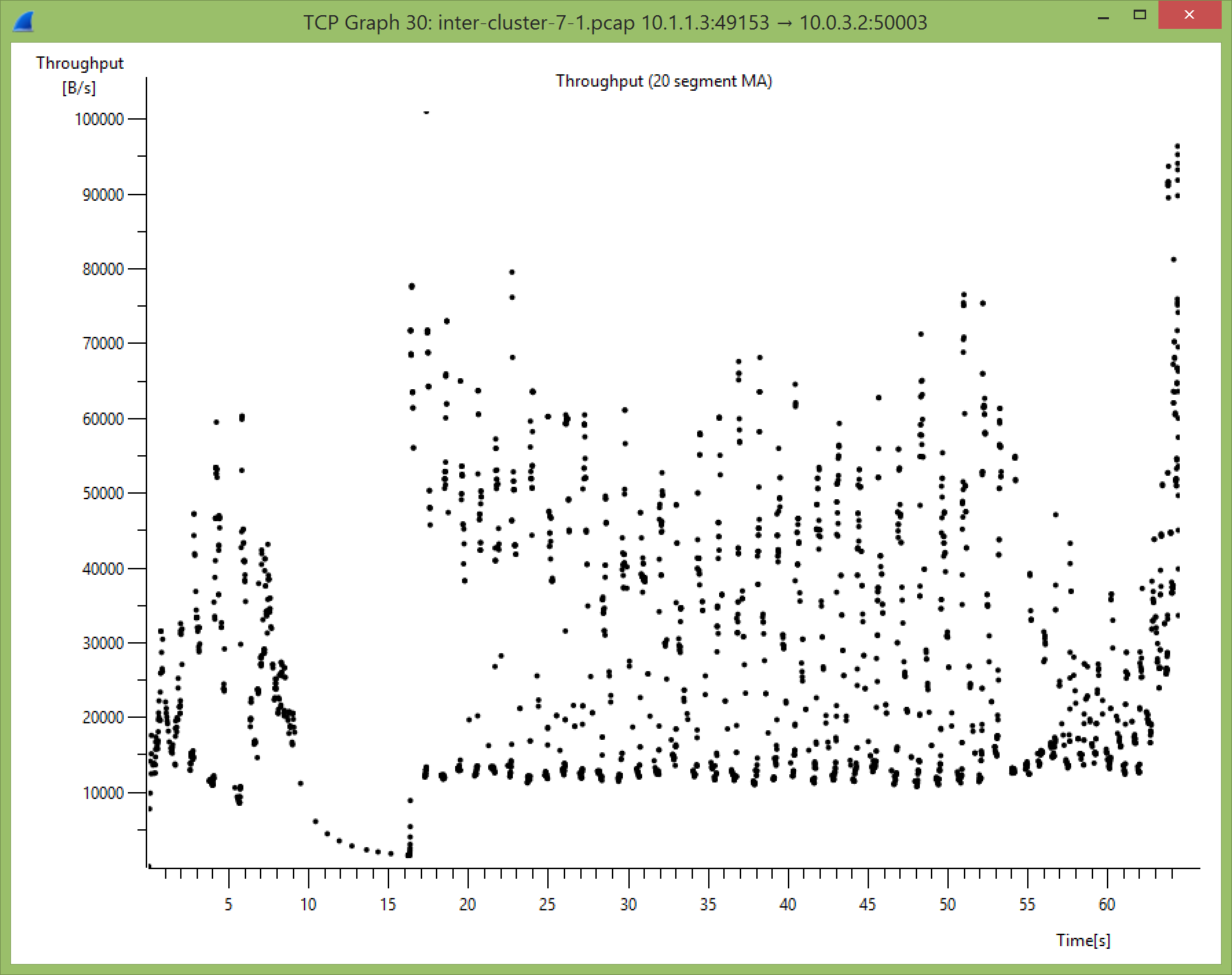
H7 => H2



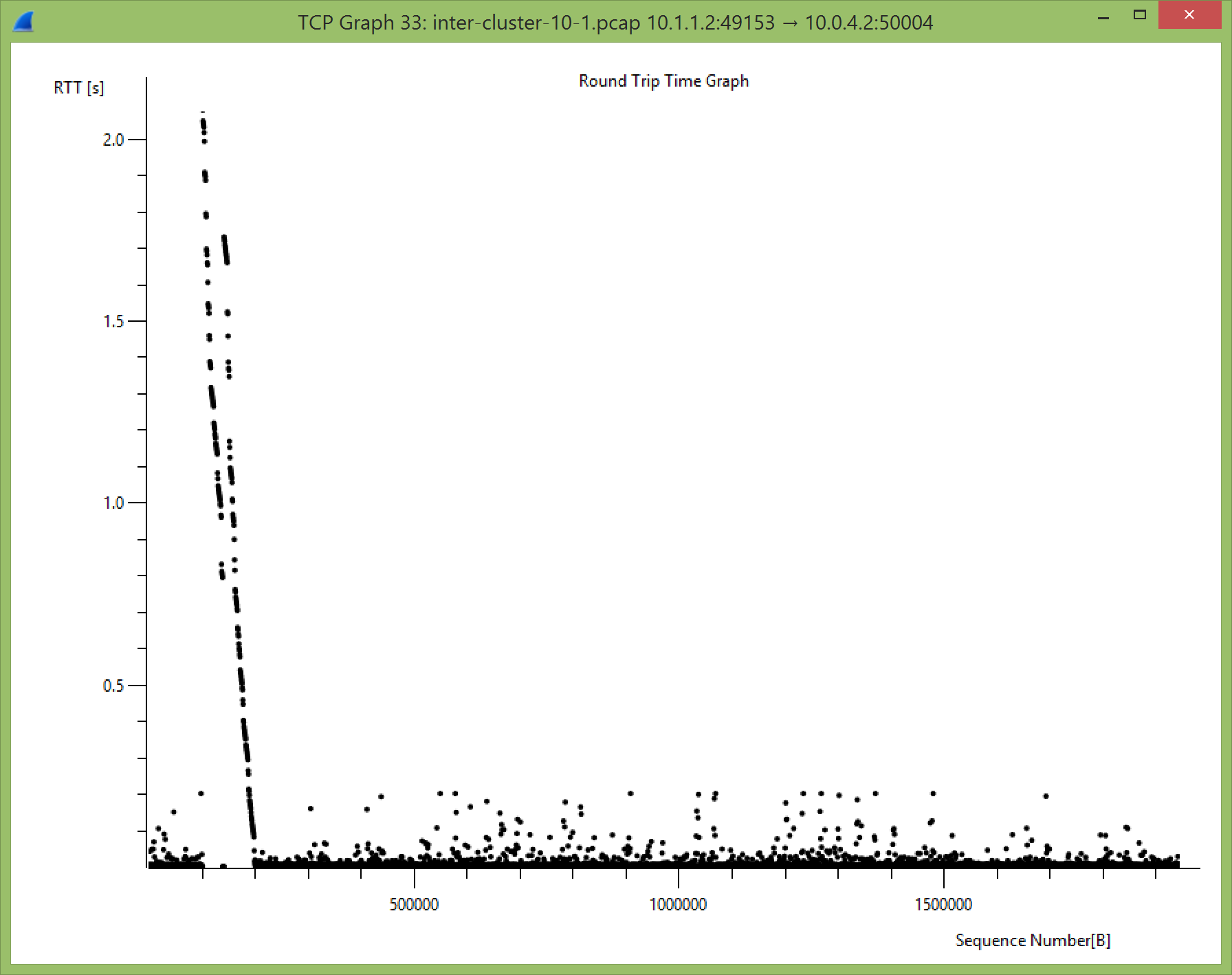
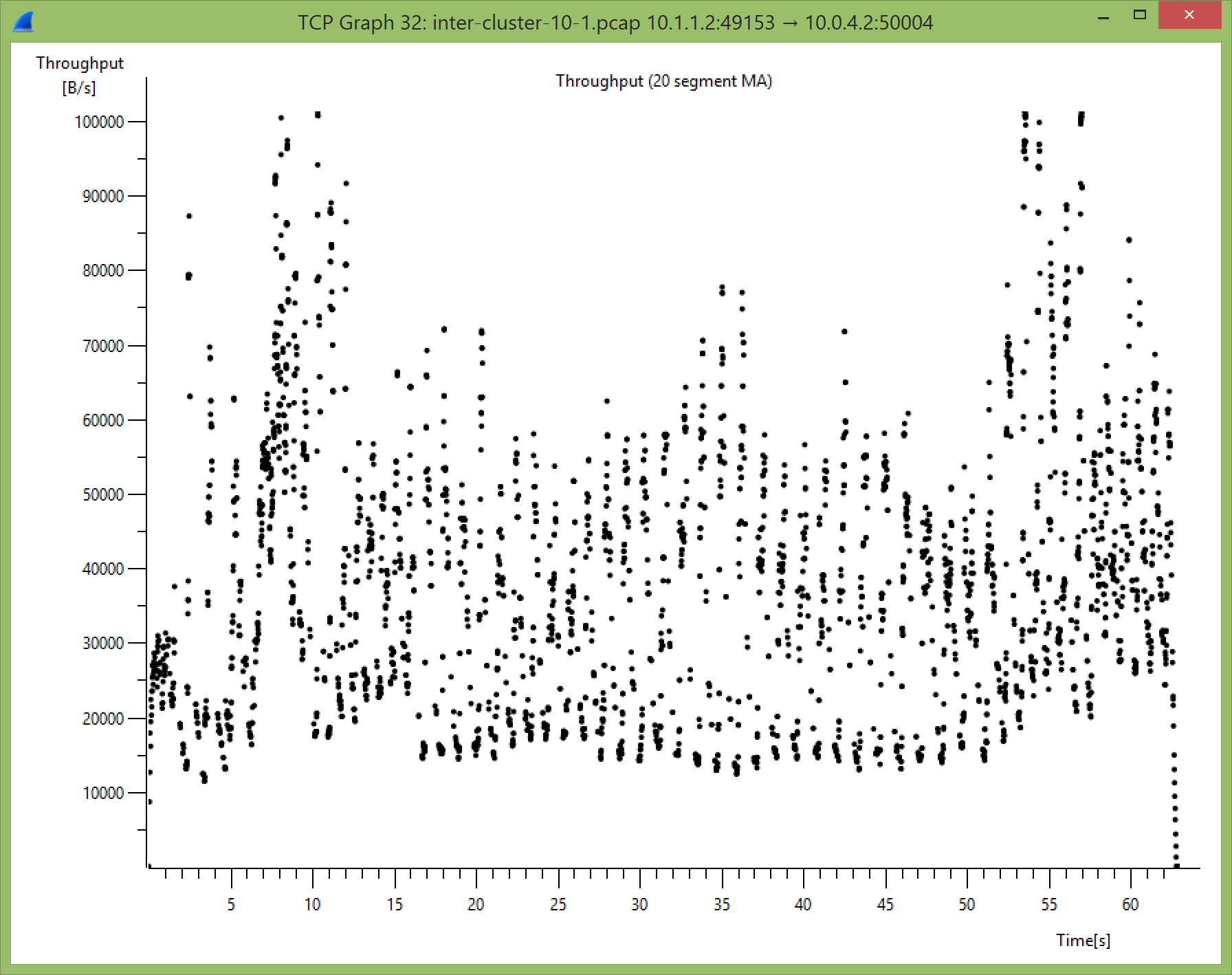
H5 => H4



H3 => H6



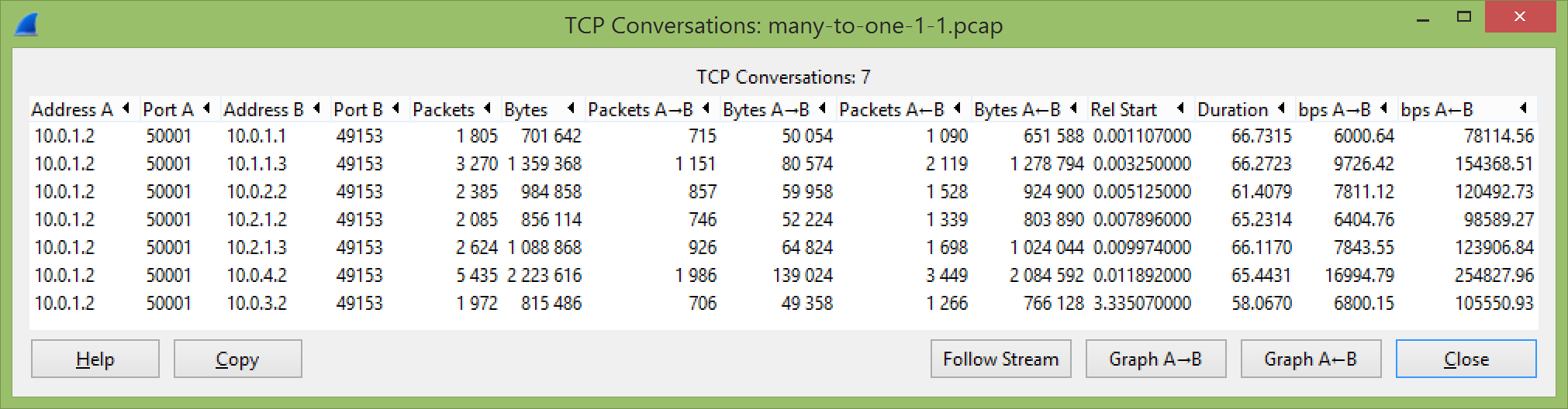
H1 => H8



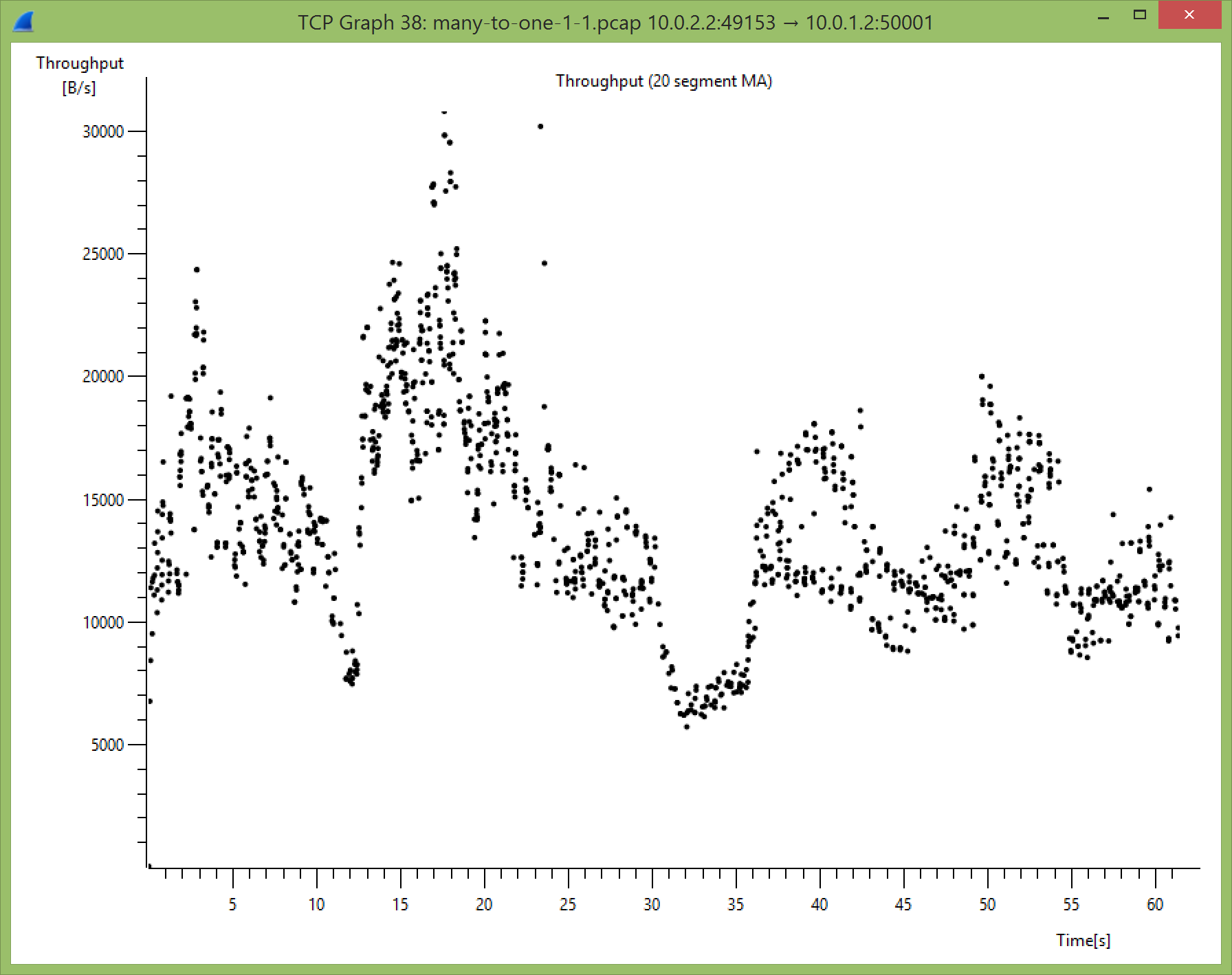
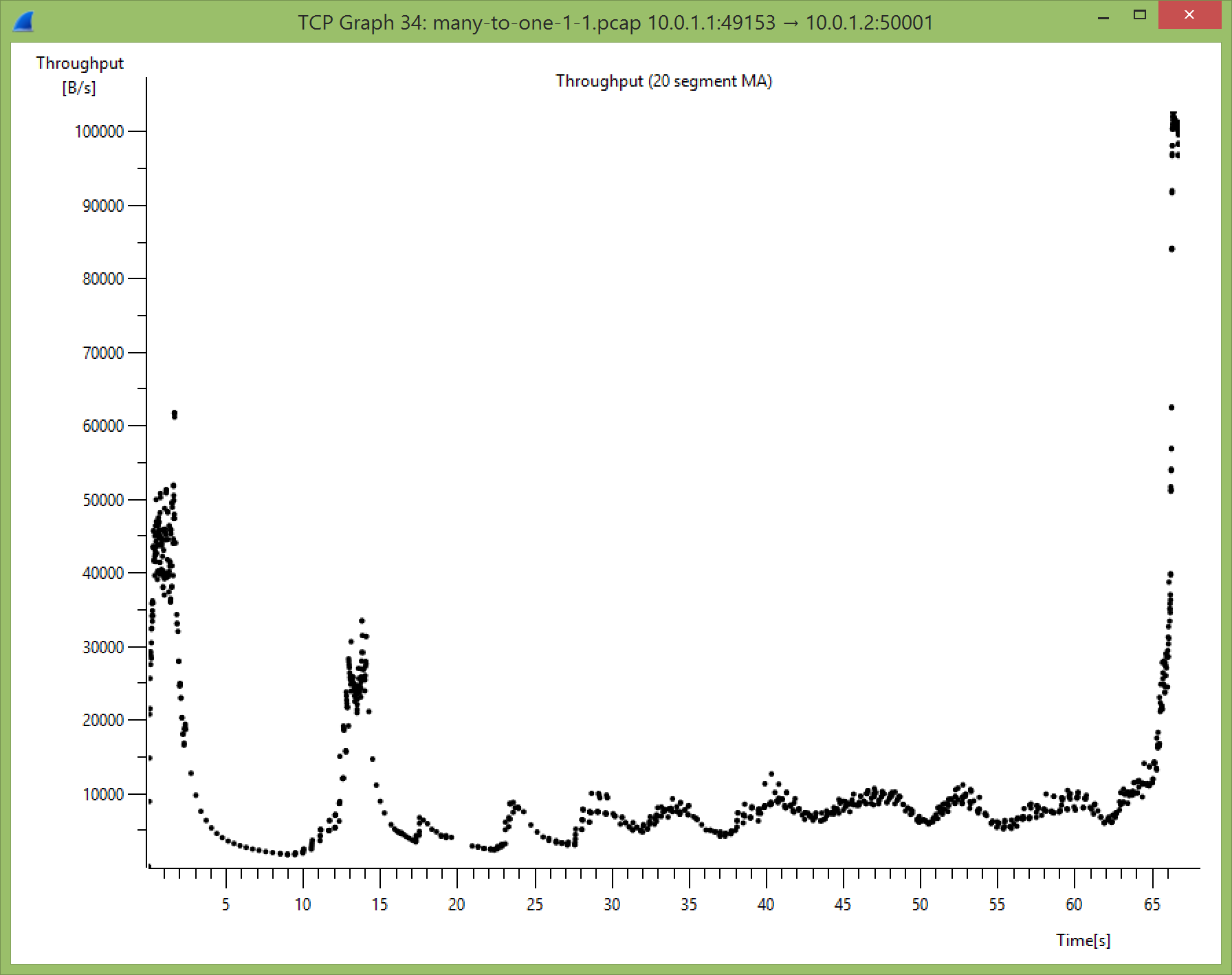
# Many to one

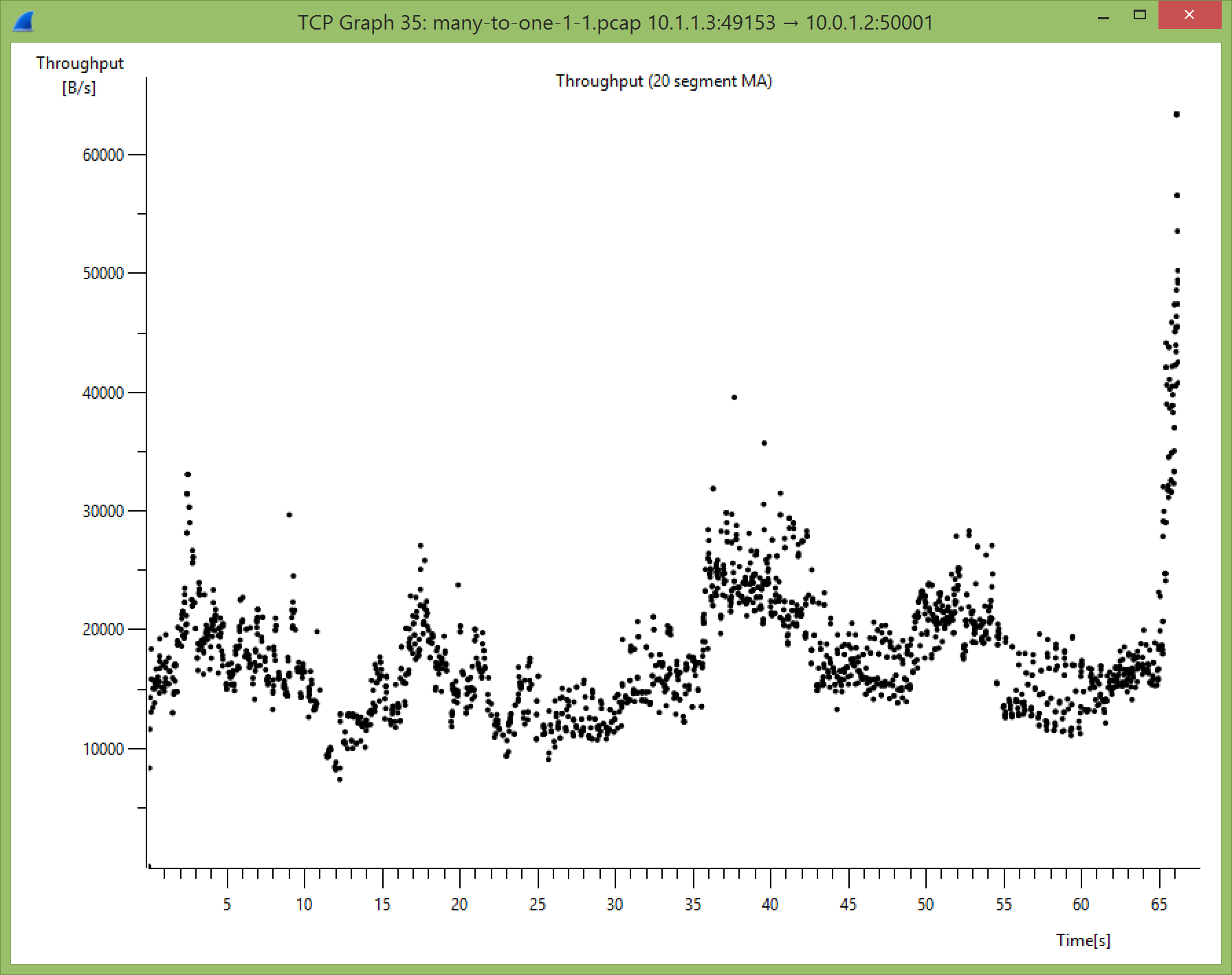
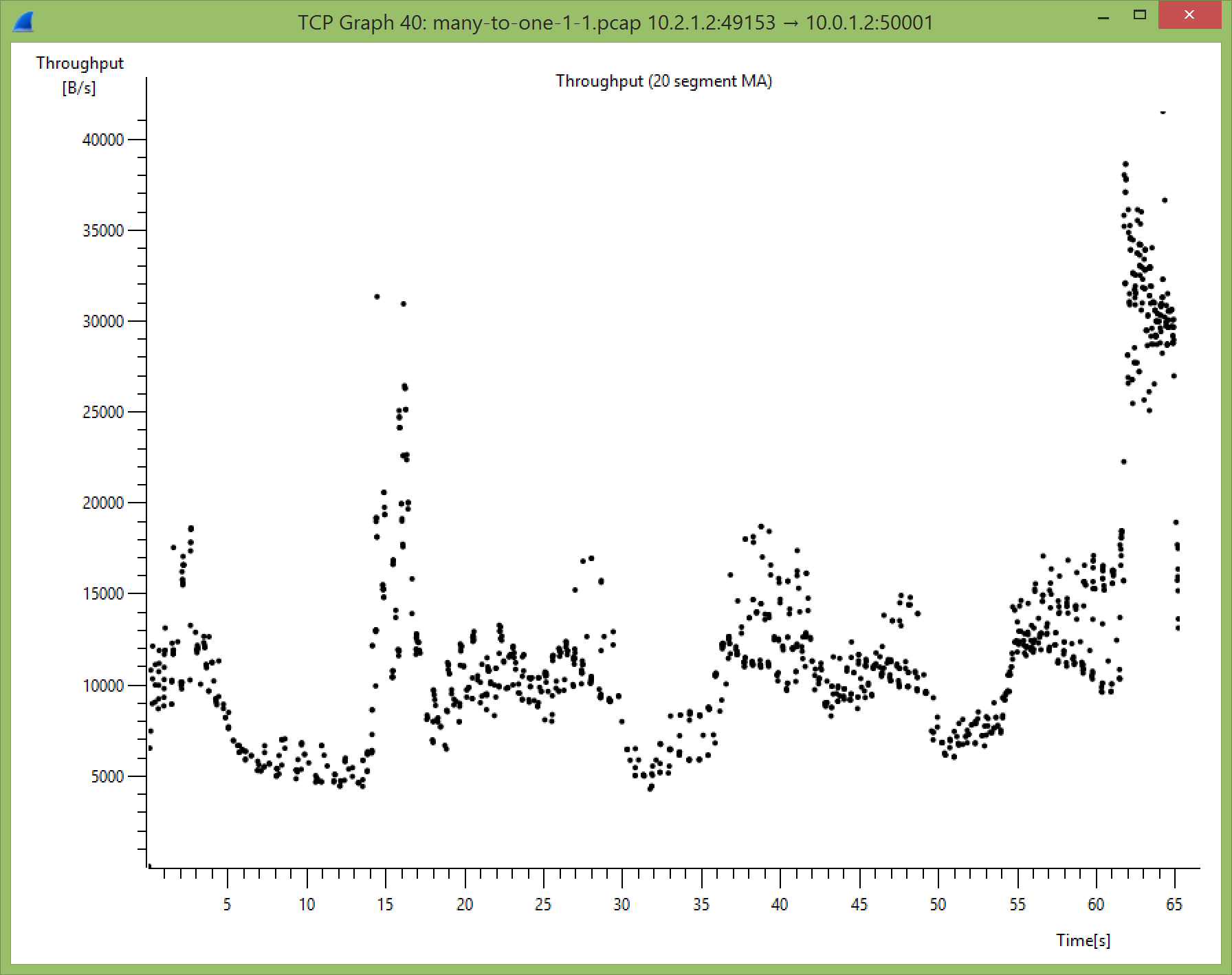
测试方式：H1, H3, H4, H5, H6, H7, H8 分别发起一条到H2的TCP流。

每条流均从0秒开始，60秒结束，测试时间持续60秒。

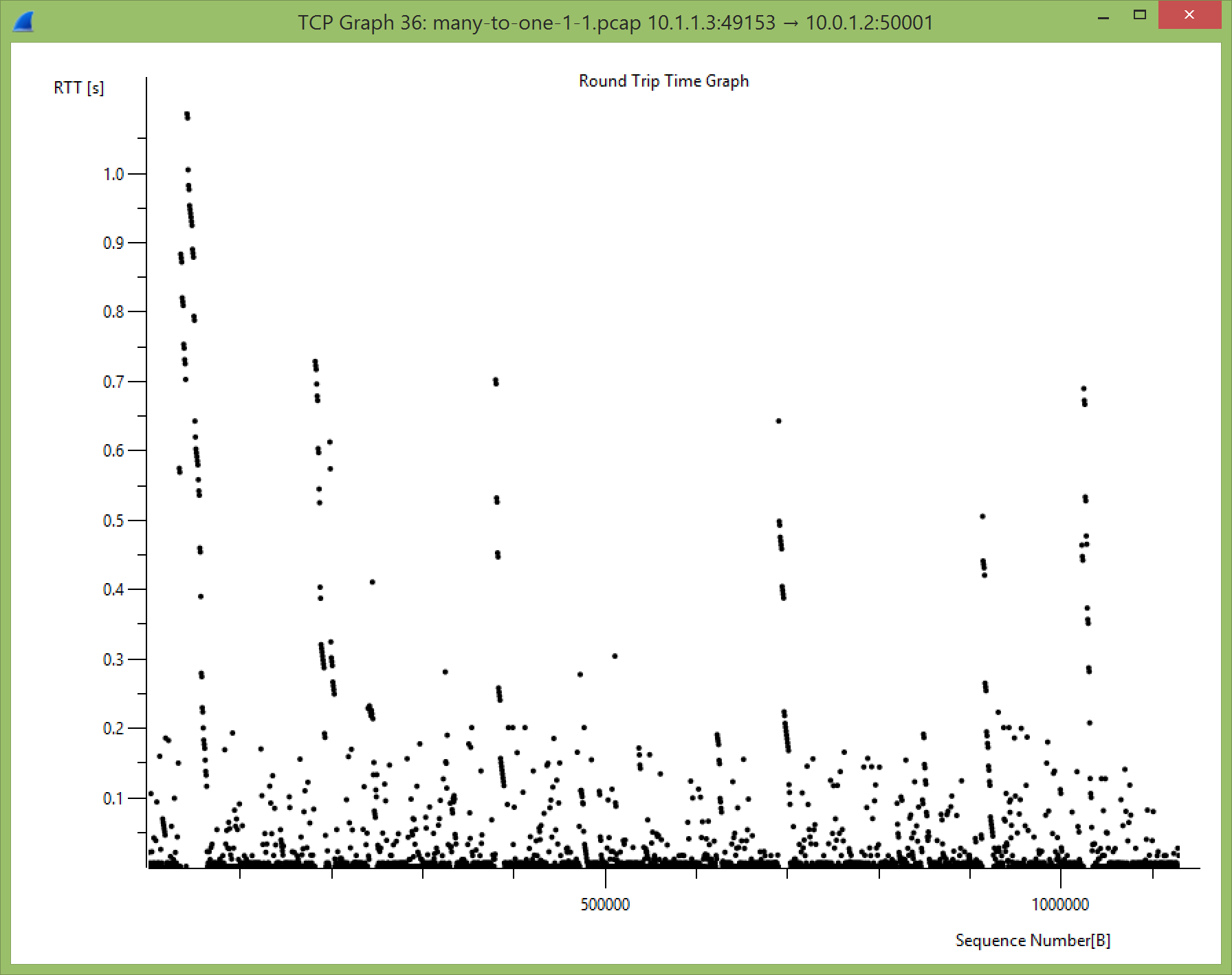
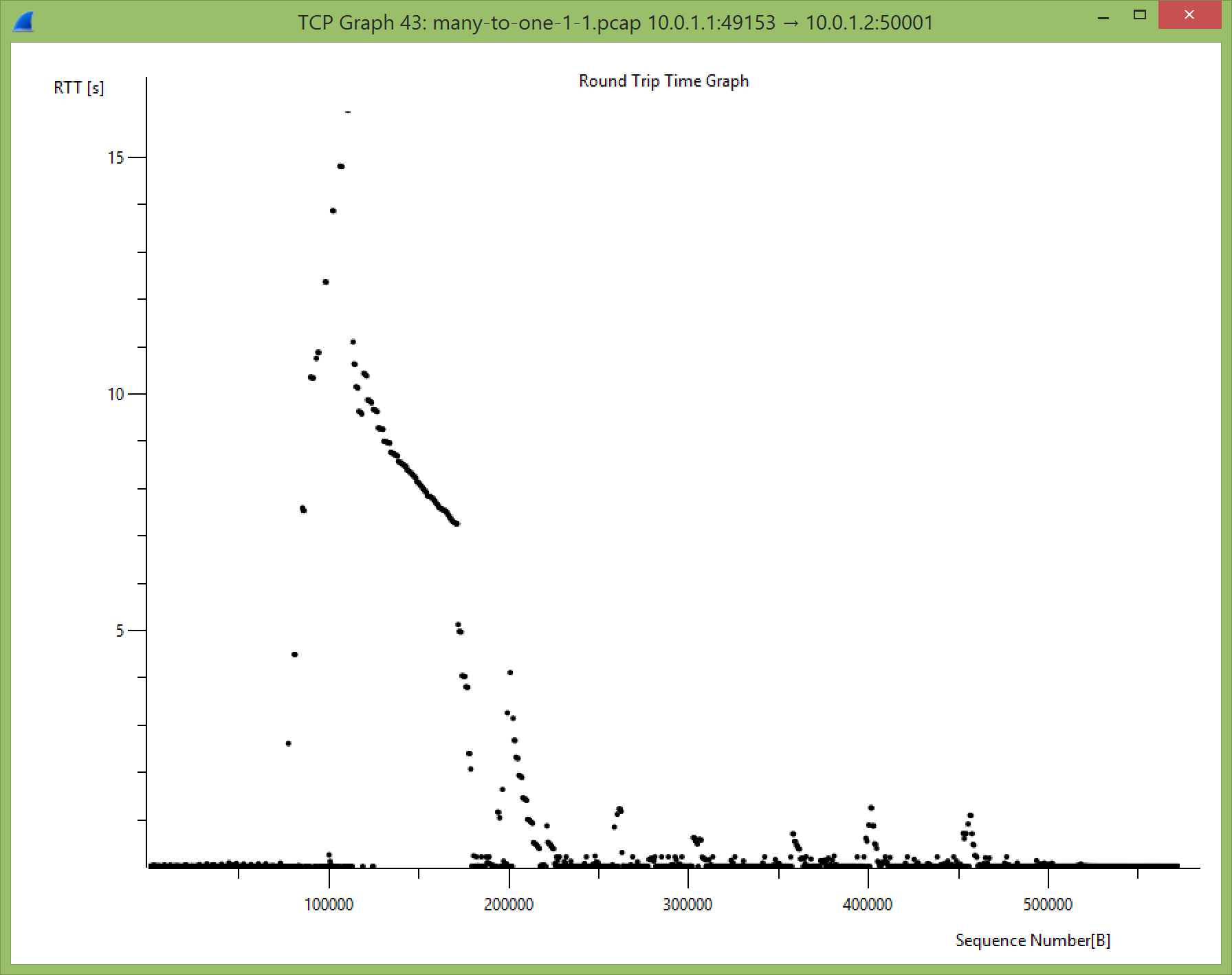


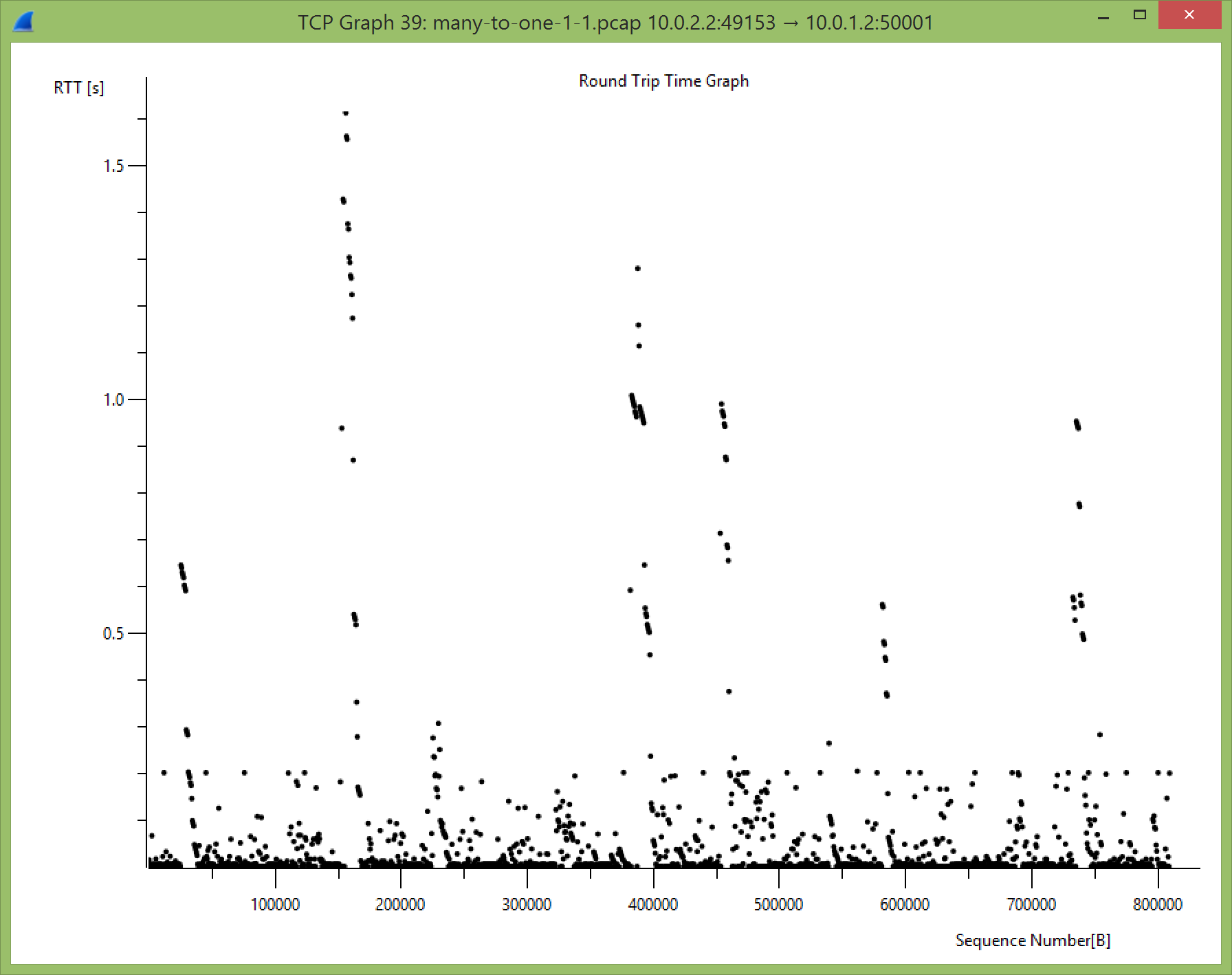
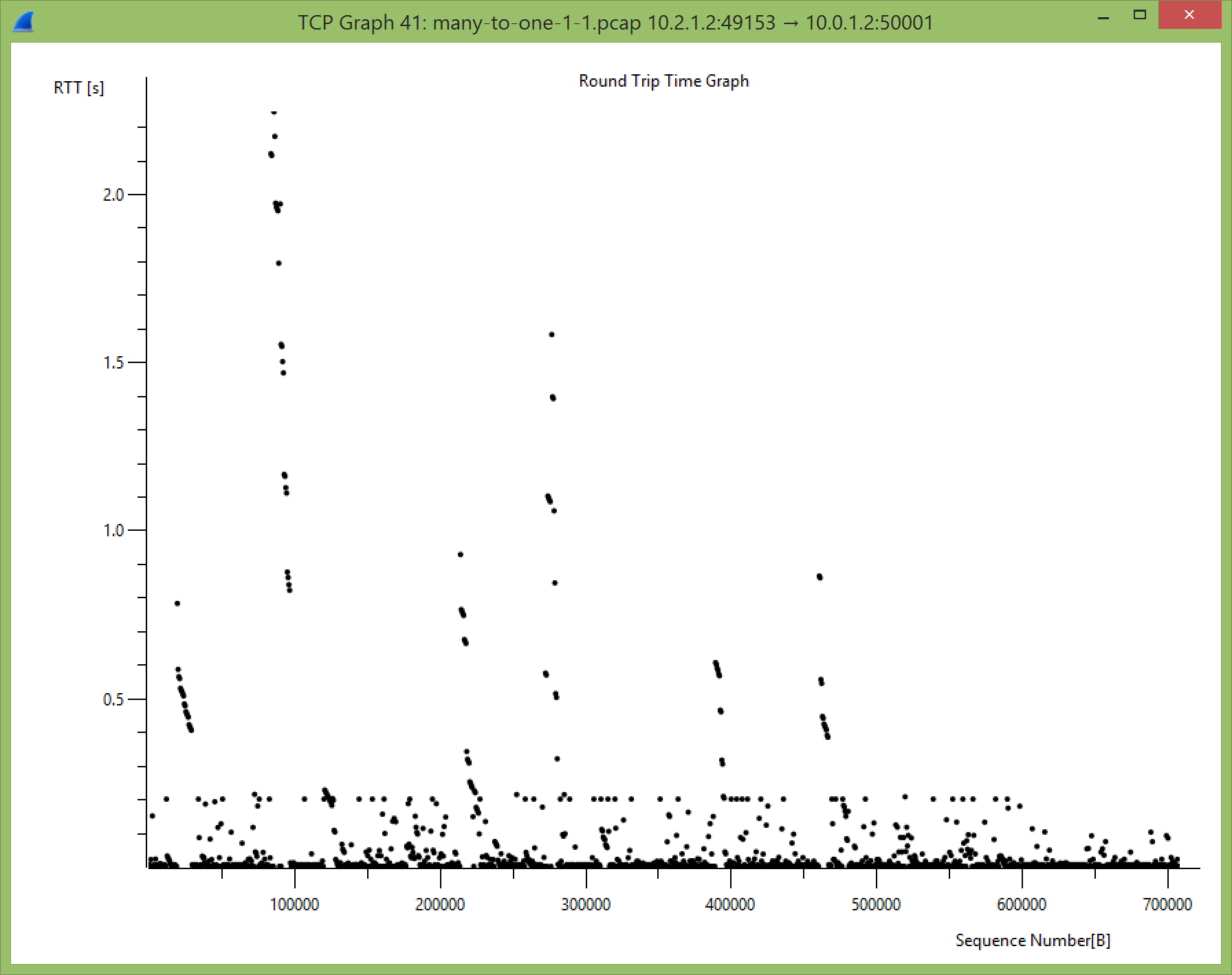
共有7个流，下面列出4个流的吞吐量。



共有7个流，下面列出4个流的RTT。

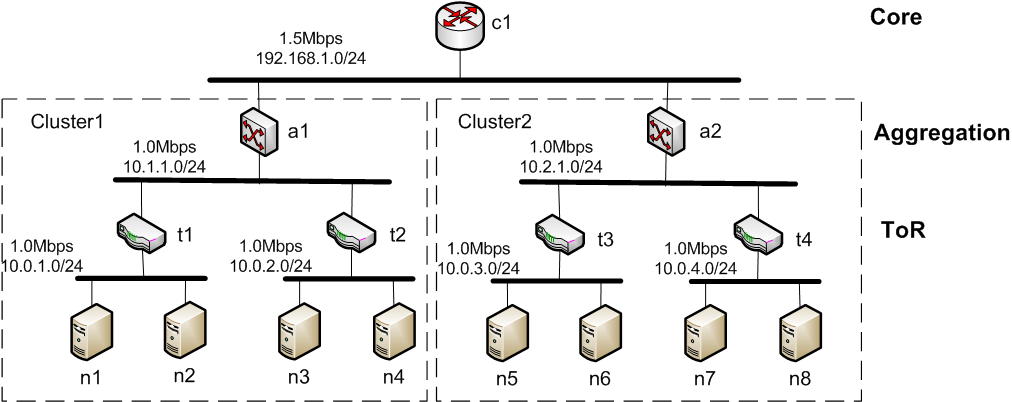


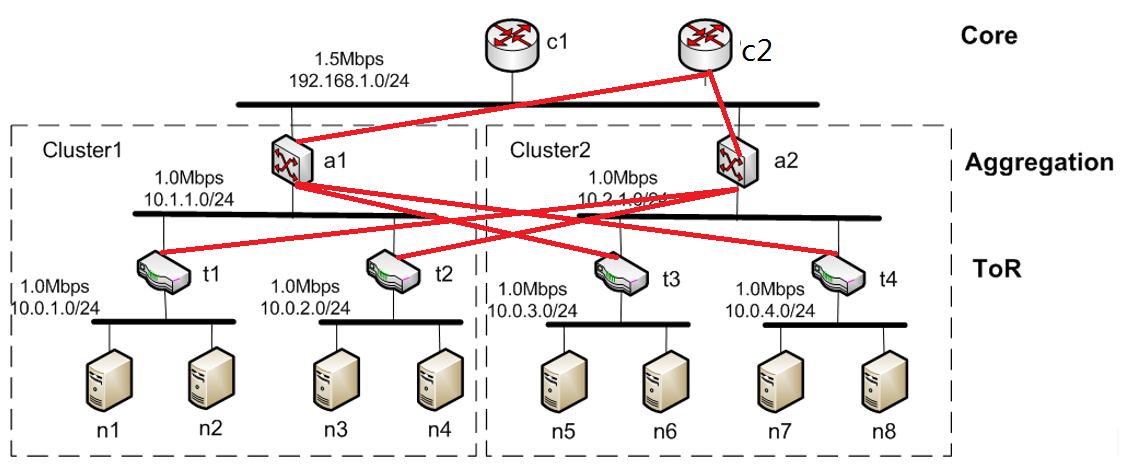
# 优化数据中心拓扑结构

在Inter-cluster的流量模式下，系统的瓶颈位于Core Switch和Aggregation Switch。左侧Cluster1向右侧Cluster2全速发送数据时，叶子节点的总出站带宽可达4 Mbps，但Aggregation Switch 只有1.0 Mbps，核心交换机只有 1.5 Mbps，也就是达不到full bisection，会成为系统的瓶颈。

因此，优化方法应该是增加Core和Aggregation层之间，以及Aggregation与ToR层之间的互连度。



如下图所示，添加一个核心交换机c2，并在Aggregation层和ToR 层之间增加连接，使得每个Aggregation Switch向上连接到两个核心交换机，每个ToR Switch向上连接到两个Aggregation Switch。红线表示新增的连接。



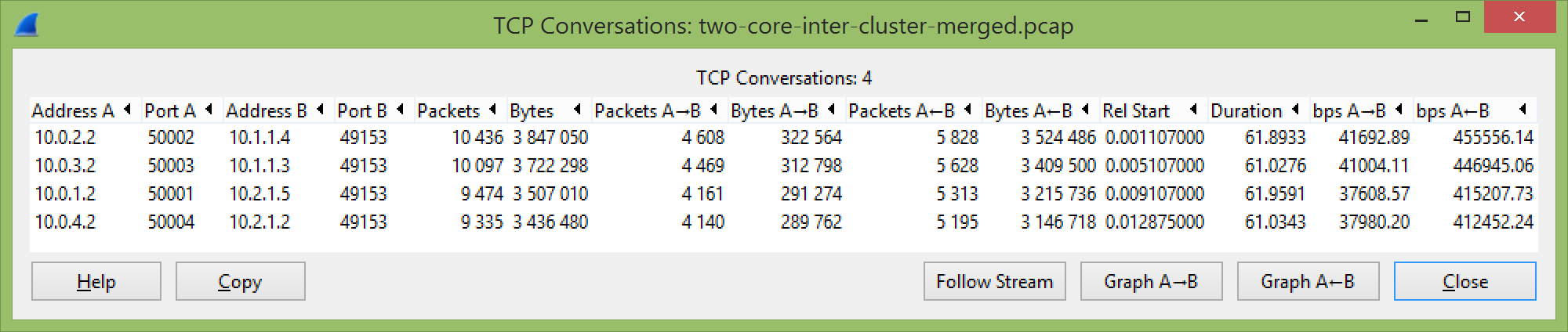
采用优化后的拓扑结构，辅以ECMP路由，不仅可以增加Bisection Bandwidth（具体优化效果见下），还可以实现failover，即任何一个core switch或者aggregation switch宕机都不会影响网络的连通性。在NS-3中开启ECMP Routing的代码为：

Config::SetDefault("ns3::Ipv4GlobalRouting::RandomEcmpRouting", BooleanValue(true));

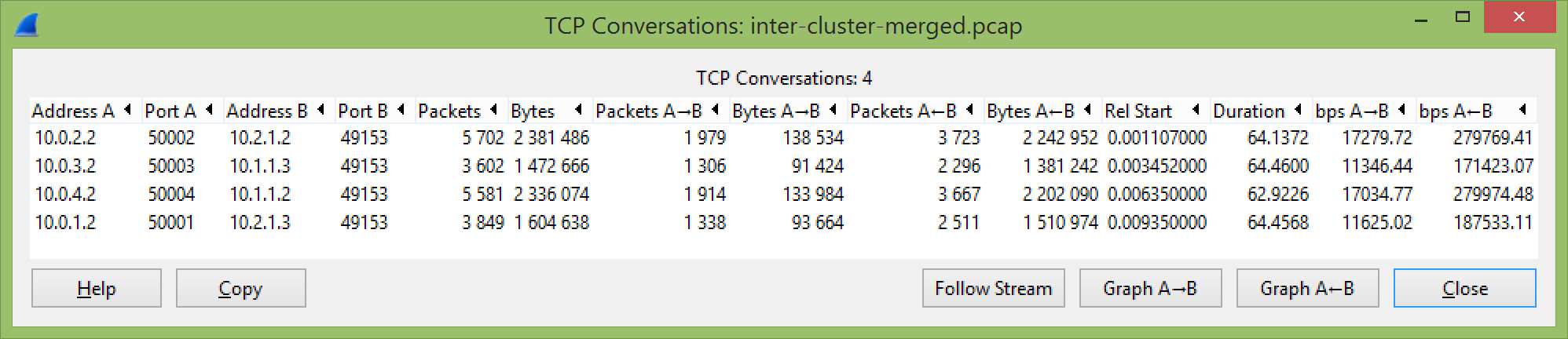
优化效果：

## Inter-cluster

优化后

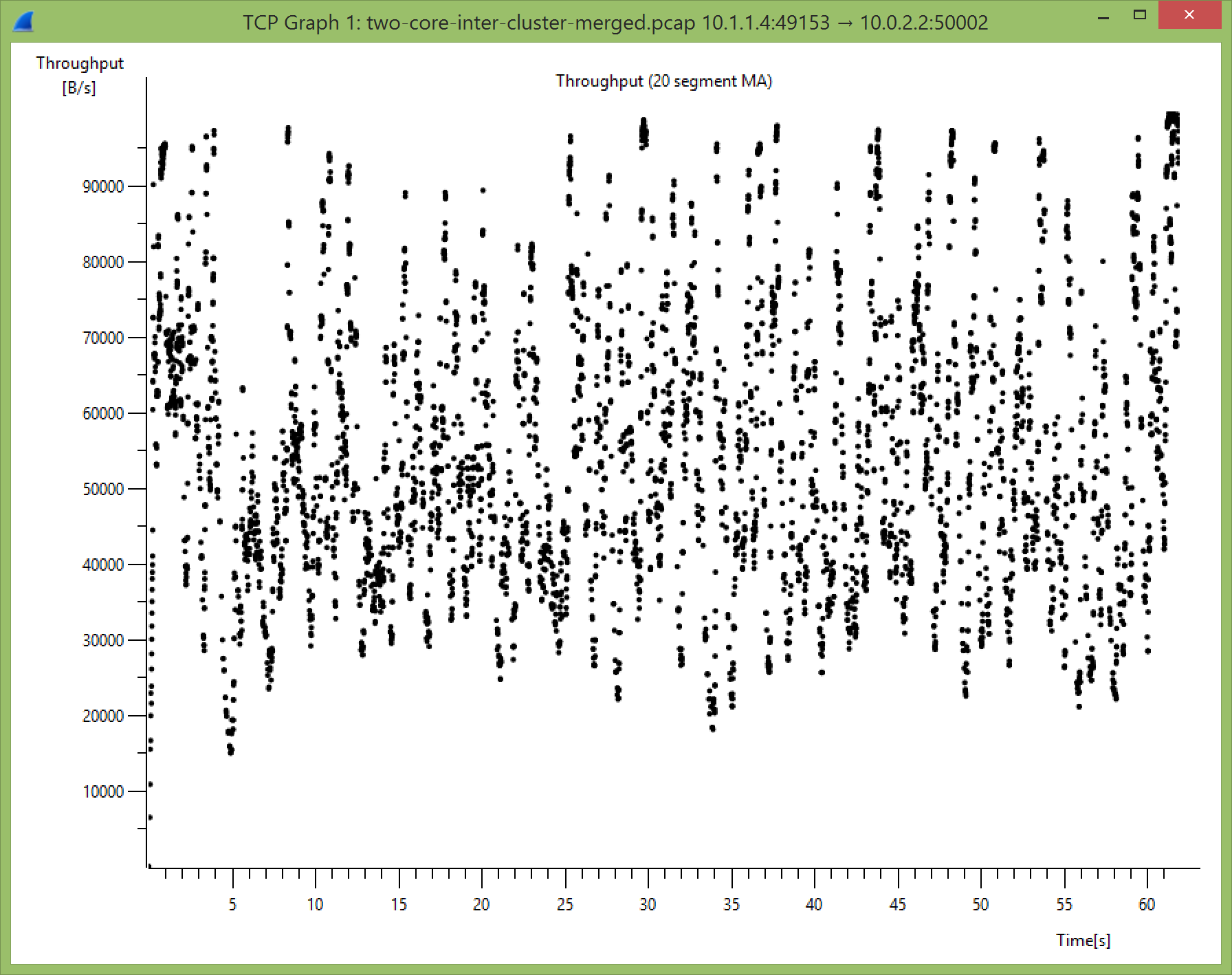
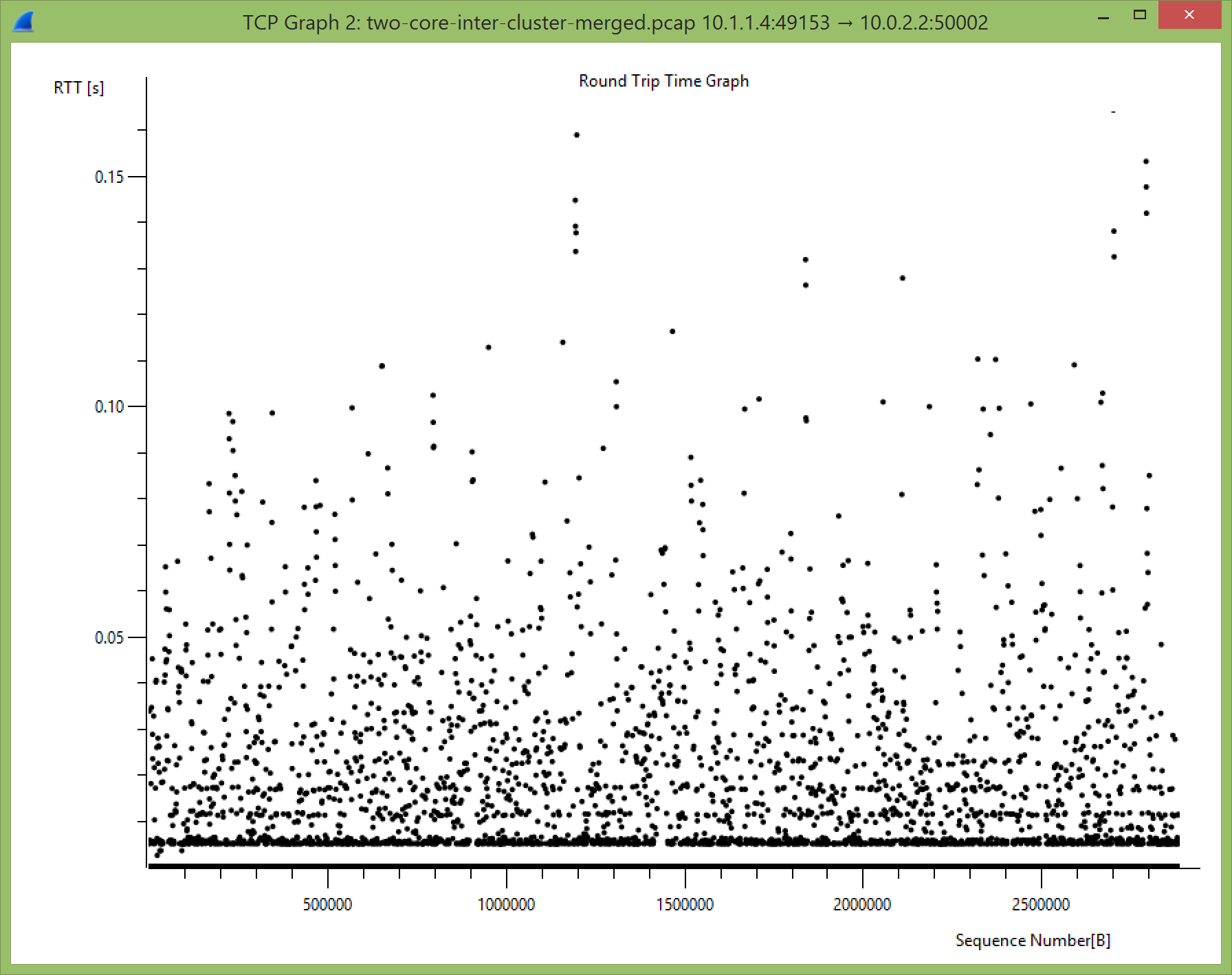


优化前



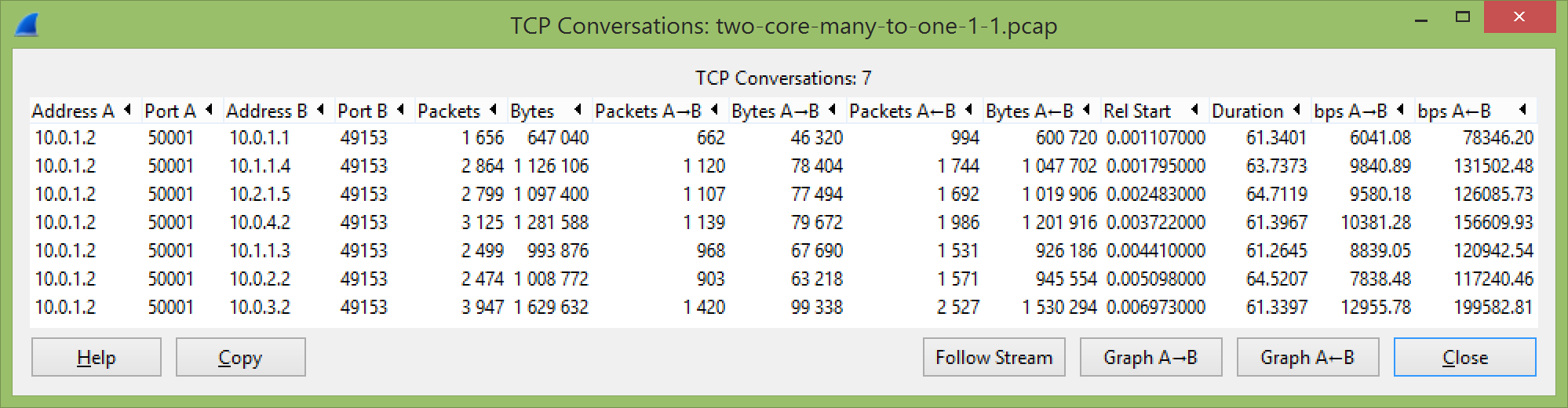
优化后的吞吐量明显增加。这是消除了核心层和汇聚层瓶颈的效果。

下图是优化后一条TCP流的吞吐量和RTT。

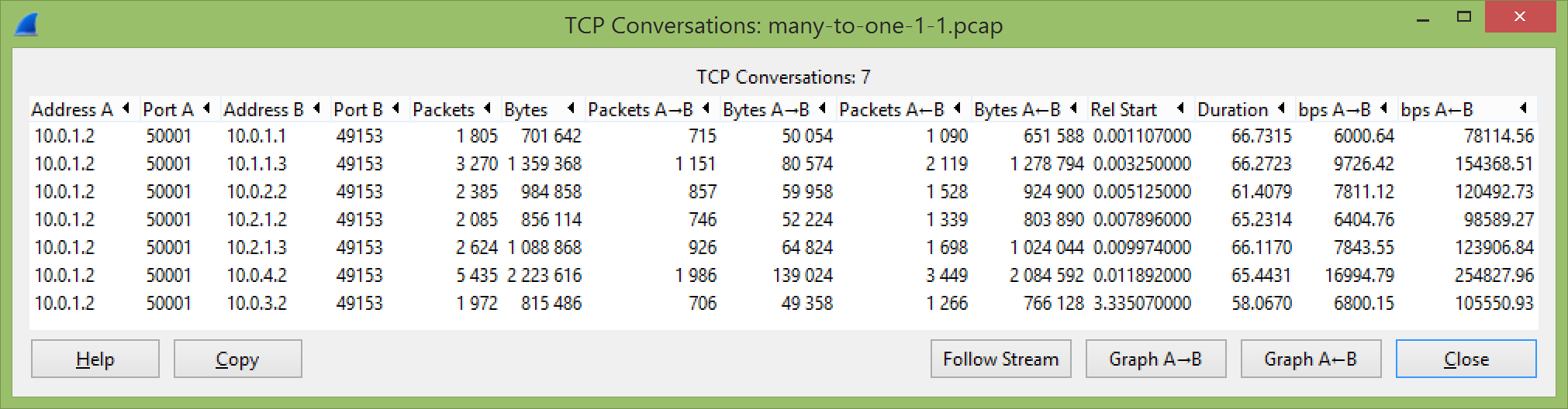
 

## Many-to-one

优化后



优化前



优化后的总吞吐量尽管没有增加（瓶颈是1号节点的总带宽1.0 Mbps），但

1. 各条流之间的带宽分配平均了很多（如上所示）
2. 同一条流的吞吐量随时间的变化平滑了很多，大起大落比较少。

也就是说，采用ECMP之后的数据中心网络在多对一的情况下吞吐量更可预测。

