

新流量模型下的拥塞控制算法评估

网络创新训练汇报

刘般若

指导助教：刘志文

2022 年 6 月 9 日



① 实验背景

② 实验设置

③ 实验结果

④ 结论与心得

① 实验背景

② 实验设置

③ 实验结果

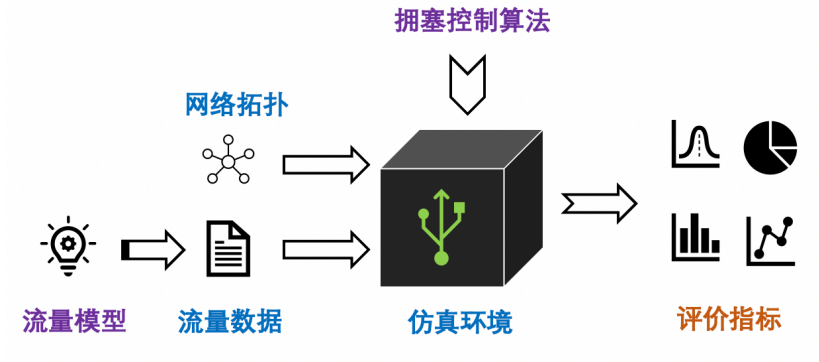
④ 结论与心得

数据中心下的网络需求

- 数据中心网络比普通网络追求更高的吞吐和更低的时延。
- 其中拥塞控制 (CC, congestion control) 是提升性能的重要部分。研究人员不断努力设计更好的算法。
 - DCTCP, 2010
 - TIMELY, 2015
 - DCQCN, 2015
 - HPCC, 2019

实验动机

现有的论文中的评估基于均匀的流量数据，不能完全反映真实业务下有偏的流量数据。需要进行对比实验。



① 实验背景

② 实验设置

③ 实验结果

④ 结论与心得

流量模型

一个流可以抽象为 $f = (src, dst, time, size)$ 四元组。

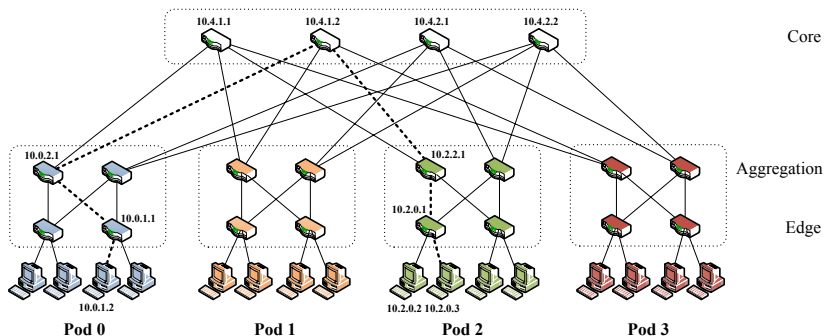
本实验中对比两种流量模型

- 均匀模型
 - 源-目的对完全随机抽取。
 - 时间间隔和流大小也从某个经验分布中抽取
 - 四个特征完全独立
- 有偏模型
 - 考虑了四元组的联合分布。
 - 在空间上有偏

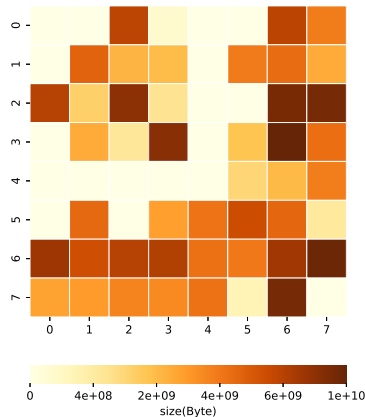
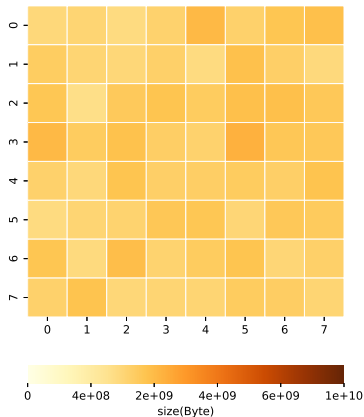
在测试时控制变量保证两种模型生成的流在统计意义下的平均负载相同。

网络拓扑

使用 Fat Tree 拓扑



流量输入的空间分布



① 实验背景

② 实验设置

③ 实验结果

④ 结论与心得

测试指标

流完成时间 (FCT, flow completion time):

一个流从源主机发出到全部被目的主机接收所经过的时间。

理论最小 FCT:

流大小 / 链路带宽

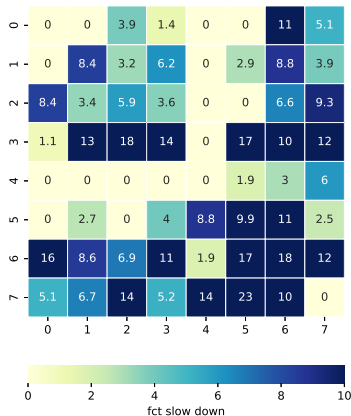
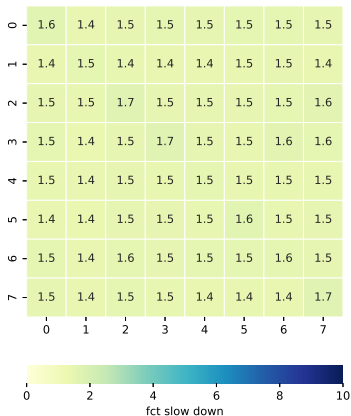
FCT slow down:

实测 FCT / 理论最小 FCT

体现了拥塞程度，并且排除了流大小的影响。

FCT 的空间分布

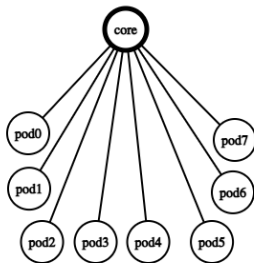
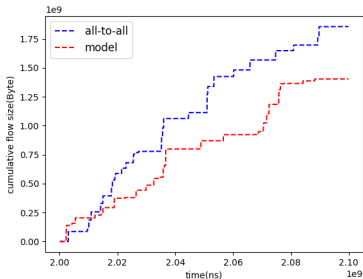
以 HPCC 为例,95% 分位点



实例分析

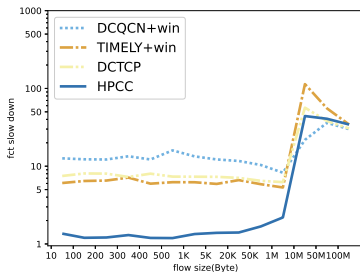
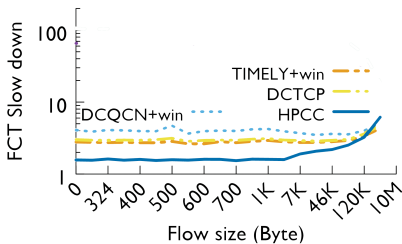
考虑 Pod3 到 Pod5 的流。

Pod3-Pod5	均匀输入	有偏输入
流量	1.77GB	1.34GB
50%FCTSD	1.0	2.2
95%FCTSD	1.5	16.9
出口带宽占比	9.8%	3.4%



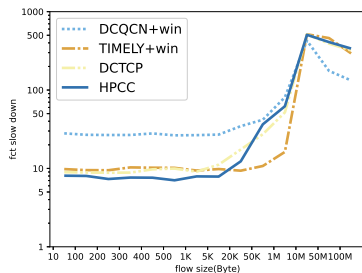
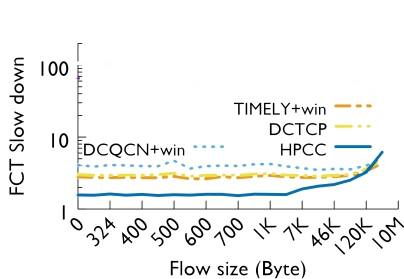
多个 CC 对比: 均匀流量输入

95% 分位点



多个 CC 对比: 有偏流量输入

95% 分位点



① 实验背景

② 实验设置

③ 实验结果

④ 结论与心得

流量特征影响 CC 表现

- 在不同的流量输入下，CC 的表现不同。
- 现实业务中，流量特征一般是有偏的。
- 不难推测，对于不同业务的流量特征，这些算法的表现又会有所不同。
- 更进一步的，我认为一致最优的算法很难找到，需要根据流量特征调整算法中的参数，甚至设计新的算法。

拥塞控制算法的选择

不同的算法部署的代价不同，需要与性能提升进行权衡。

- DCTCP: 2 个主要参数，需要 ECN
- TIMELY: 4 个主要参数，需要准确 RTT 测量
- DCQCN: 约 15 个参数，需要 ECN, RED
- HPCC: 3 个参数，需要 INT

收获和感受

- 学习 NS3 仿真。
- 学习 CC 算法及其实现。
- 学习数据的分析方法。
- 修改数据生成，拓扑生成代码等。
- 约 2.5k 行 python+shell+Cpp 代码
- Debug 要比分析数据难的多！

Thanks!