WarpTCP에서 WarpRATE로

윈도우 대신 레이트를 택한 이유

Dennis Vadura, BADU Networks

2026년 11월

WarpTCP는 처음부터 WarpEngine에서 동작해 왔습니다. WarpTCP는 정교하게 다듬은 CUBIC으로, HyStart를 튜닝하고 TCP 친화(TCP-friendly) 하한을 두었으며 표준 CUBIC에 없던 수정 사항을 더했습니다. 이제 저희는 WarpTCP를 은퇴시키고 WarpRATE로 대체했습니다.

1 네트워크 패브릭은 점점 빨라지고 있습니다

1500바이트 패킷 하나를 직렬화하는 데 1 Gbit/s에서는 12 μs, 10 Gbit/s에서는 1.2 μs, 100 Gbit/s에서는 불과 120 ns가 걸립니다. 그러나 호스트와 네트워크 링크의 배칭(batching)은 같은 비율로 줄어들지 않습니다. 64 KB TSO 트레인은 여전히 하나의 버스트로 나가고, 인터럽트 병합(interrupt coalescing)은 ACK를 수십 마이크로초 동안 붙잡아 두며, 무선 스케줄러는 밀리초 단위로 동작합니다. 회선 속도가 한 단계 올라갈 때마다 같은 배칭이 더 많은 패킷에 걸치게 되고, 그 결과 ACK 클럭에 따른 모든 결정에 잡음이 더 많아집니다. 이러한 잡음 증가는 송신자가 다뤄야 할 대역폭-지연 곱(BDP)이 커지는 시기와 맞물려 있어, 잘못된 제어 결정의 대가도 함께 커졌습니다. 요컨대 네트워크는 얼마나 많은 데이터를 보낼지 올바르게 판단하는 WarpTCP의 능력을 앞질러 버렸고, 더 나은 해법이 필요했습니다.

2 ACK 잡음 앞에서 윈도우 제어는 신뢰할 수 없습니다

WarpTCP에서 혼잡 윈도우는 수신된 ACK에 의해 클럭됩니다. 도착하는 각 ACK는 여러 역할을 하지만, 가장 중요한 역할은 다음 데이터 블록의 전송을 허용하는 것입니다. 문제는 각 ACK의 타이밍이 혼잡 윈도우의 페이싱과 증가 결정에 반영된다는 점이며, ACK에 잡음이 많으면 윈도우의 증가와 감소가 더 혼란스러운 양상을 띠게 됩니다. 혼란은 처리량에 해롭습니다. 오늘날의 경로에서 ACK 스트림이 매끄러운 경우는 드뭅니다. 지연 ACK(delayed ACK), 수신 측의 GRO, Wi-Fi 프레임 집적(aggregation), LTE 그랜트 스케줄링, Starlink의 재구성 주기는 모두 ACK를 붙잡아 두었다가 한꺼번에 전달합니다. 윈도우는 이러한 묶음을 회선 위의 버스트로 바꾸고, 제어기는 압축된 ACK 트레인을 버퍼가 가득 찼다는 신호로 해석해 전송 속도를 낮춥니다. WarpRATE는 이러한 결합을 피하도록 설계되었으며, 데이터 전달 품질에 따라 제어되는 페이싱 레이트로 데이터를 보냅니다. 오류가 없으면 계속 보내고, 큐가 쌓이거나 오류가 생겼다는 징후가 보이면 속도를 늦춥니다. 또한 WarpRATE는 탄력적인 시간 버퍼를 사용해, 도착 시점이 들쭉날쭉한 지연 ACK가 전달 성공을 알려 올 때까지 충분한 여유를 둡니다.

3 퀀텀보다 긴 구간의 전달량 측정이 핵심입니다

모든 NIC(네트워크 인터페이스 카드)와 수신 측은 트래픽을 묶어서 처리하므로, ACK는 그룹 단위로 도착합니다. 그룹 사이에는 최소 간격이 있으며, 이 간격은 RTT와 수신 측에서 일어나는 다른 모든 일에 따라 정해집니다. 저희는 이 간격을 측정 퀀텀(measurement quantum)이라고 부릅니다. 퀀텀보다 짧은 시간 척도의 타이밍 신호는 배칭을 측정할 뿐, 경로에 대해서는 거의 알려 주지 않습니다. WarpRATE는 레이트를 V, 즉 최소 8개 에포크(측정 주기, 약 1 RTT) 동안 전달된 데이터 양으로 추출합니다. 이 구간은 경로가 만드는 배칭을 충분히 적분할 만큼 길기 때문에, 측정은 구조적으로 퀀텀의 영향을 받지 않습니다.

4 ALLOW, LOSST, RATE를 분리하면 제어 법칙이 더 안정됩니다

혼잡 윈도우 기반 제어기 설계에서는 숫자 하나가 세 가지 일을 합니다. 윈도우는 전송 중인(in flight) 데이터의 상한을 정하고, 전송 속도를 결정하며, 손실 허용 범위를 제어하는 데에도 쓰입니다. 그래서 한 가지 역할에 닿는 신호가 세 가지 모두에 영향을 줍니다. 순서 뒤바뀜을 손실로 읽으면 전송 속도가 깎이고, 손실 검출기에 지터 여유를 두면 전송 중인 데이터가 부풀어 오릅니다. WarpRATE는 역할마다 하나의 제어 변수를 둡니다. ALLOW는 전송 중인 바이트의 상한을, RATE는 패킷 간 지연을 정하고, LOSST는 확인 응답을 받지 못한 세그먼트를 언제 손실로 판단할지 결정합니다. ACK 타이밍은 LOSST에만 영향을 주어 그 판단 구간을 넓힙니다. 페이스는 전달된 데이터 양과 이상 신호, 즉 손실 표시, CE 에코, 자체적으로 유발한 지연에 의해서만 바뀝니다. 지터 스파이크로 LOSST(손실 판정 대기 구간)가 넓어져도 페이스와 전송 중인 바이트는 그대로 유지됩니다. 이러한 분리는 앞서 설명한 원인들로 ACK 타이밍이 흔들리는 동안 처리량을 지키는 데 매우 효과적이었습니다.

WarpRATE는 현대적인 혼잡 제어 알고리즘으로, 저희는 PFQD와 함께 사용할 때 장거리 인터넷 경로와 데이터센터 내부의 짧은 경로 모두에서 이점이 있다고 믿습니다. 초기 실제 배포 결과는 실험실 테스트와 검증 결과를 뒷받침하고 있어 고무적입니다.