Listifyed (UDP) GRO

Steffen Klassert

secunet Security Networks AG

Dresden

Linux Netconf, Boston, June, 2019

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Generic Receive Offload (GRO)

- Software packet aggregation on Layer 2
- Merges packets of a flow into one big packet (up to 64 KB)
- Reduces the number of network stack traversals
- Good performance if a local socket can handle GRO
- Good performance on forwarding if the TX NIC can do LSO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Problematic GRO usecases

- Does not perform well if GSO is needed
- GRO and GSO add overhead
  - calculating checksums
  - modifies headers
  - touches packet data
- UDP needs needs software GSO if NIC can't do LSO
- IPsec needs software GSO if NIC can't offload ESP
- Is there a better way to do GRO if software GSO is needed?
- Yes: Listifyed GRO

Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

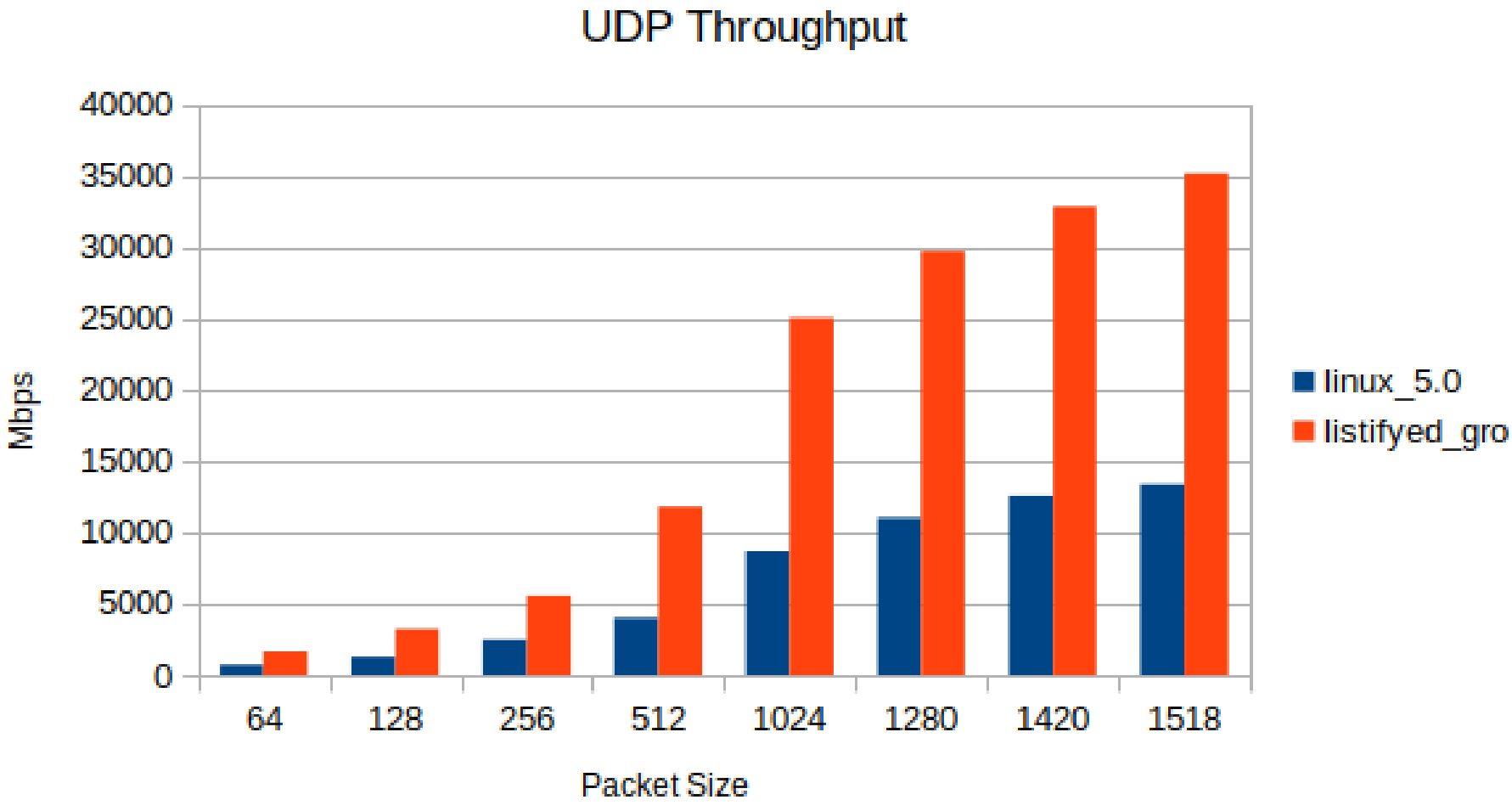
Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

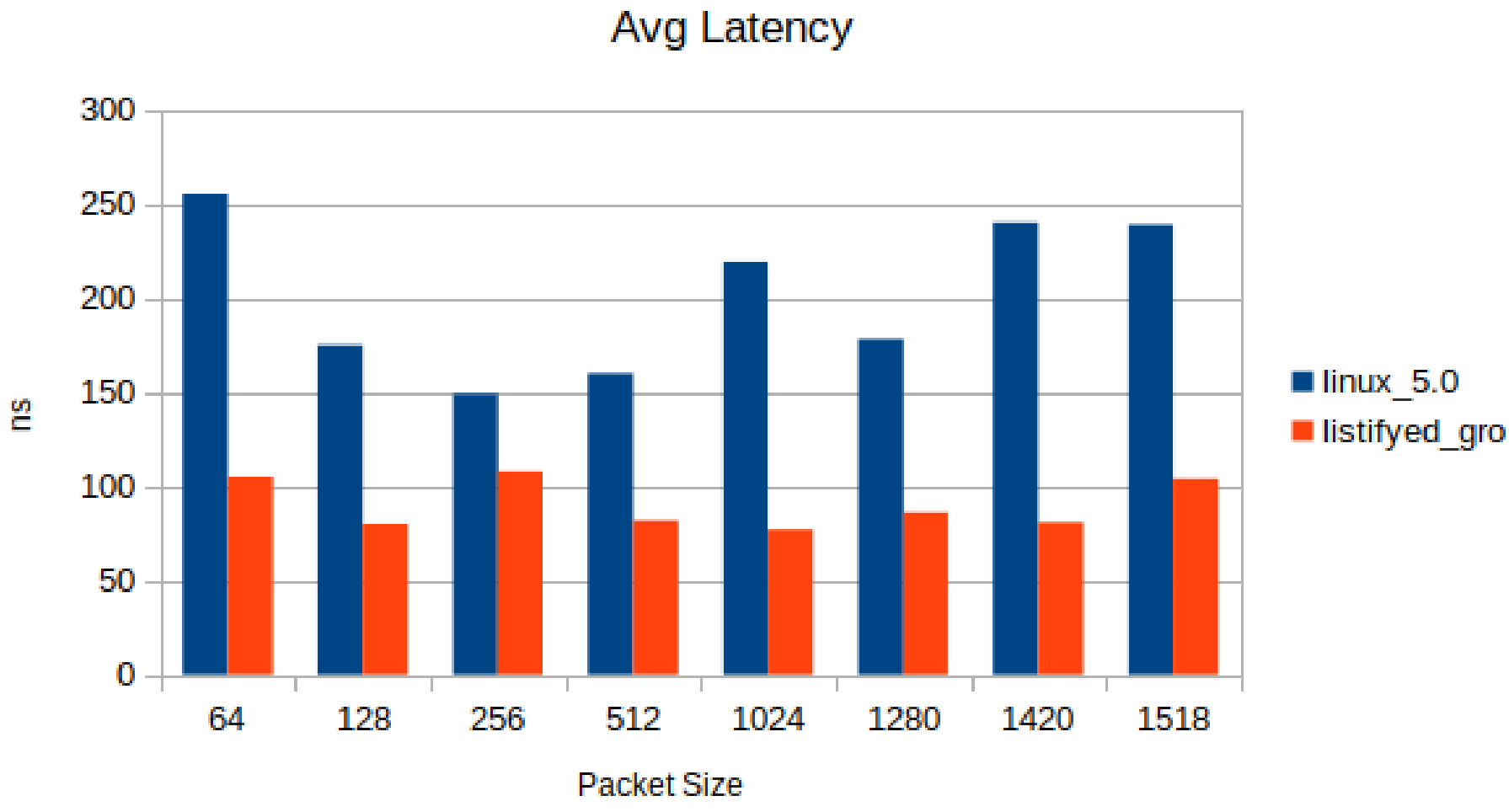
Listifyed GRO

- Does not merge packets into a big packet
- Builds a linked list of packets
- Linked packets travel together though the stack
- GSO just needs to unlink the packets
- Listifyed GRO/GSO does not need to touch packet data
- Listifyed GRO/GSO performs better than standard GRO/GSO

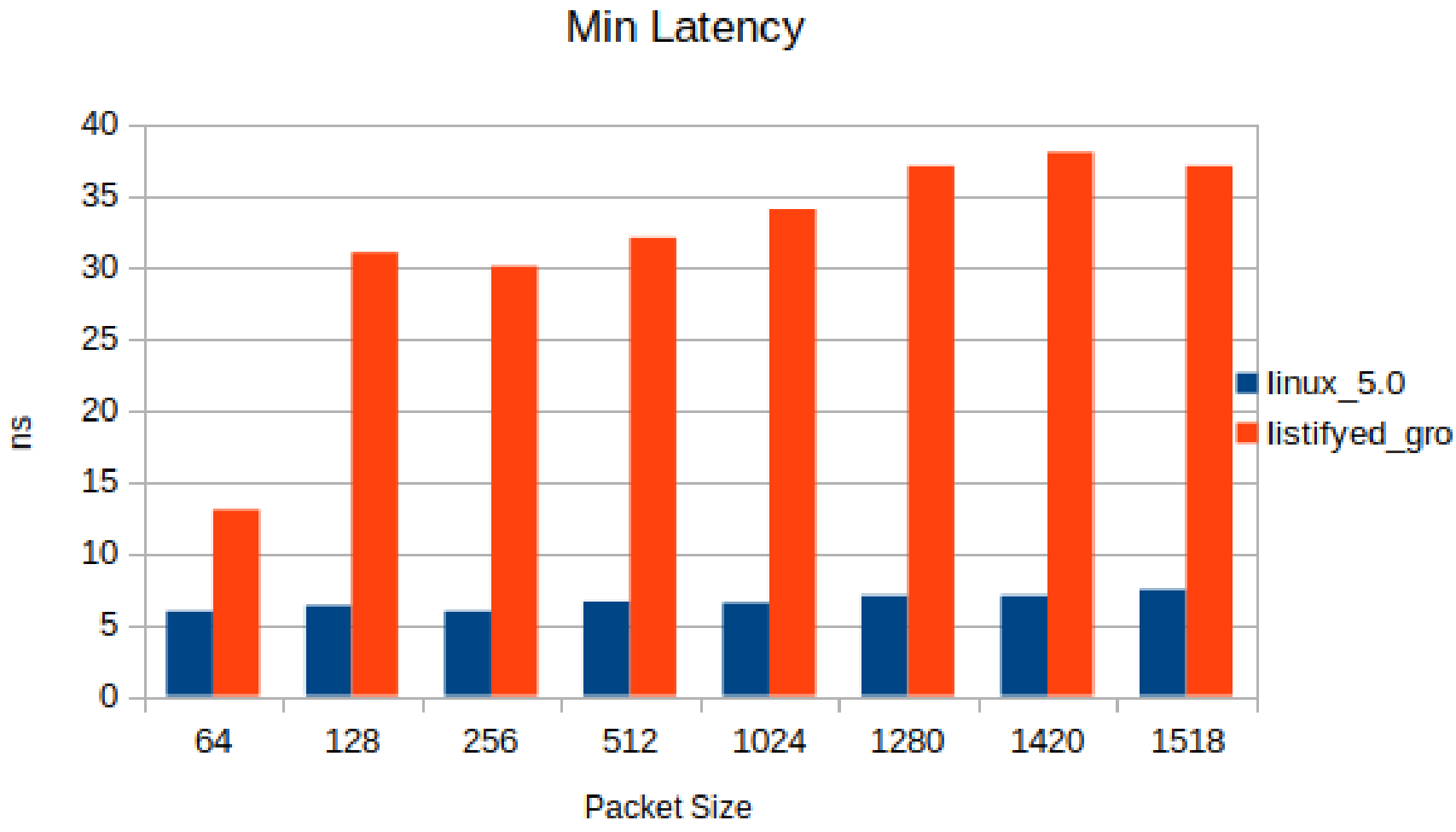
Max UDP Throughput in Mbps (single flow / 30 sec)



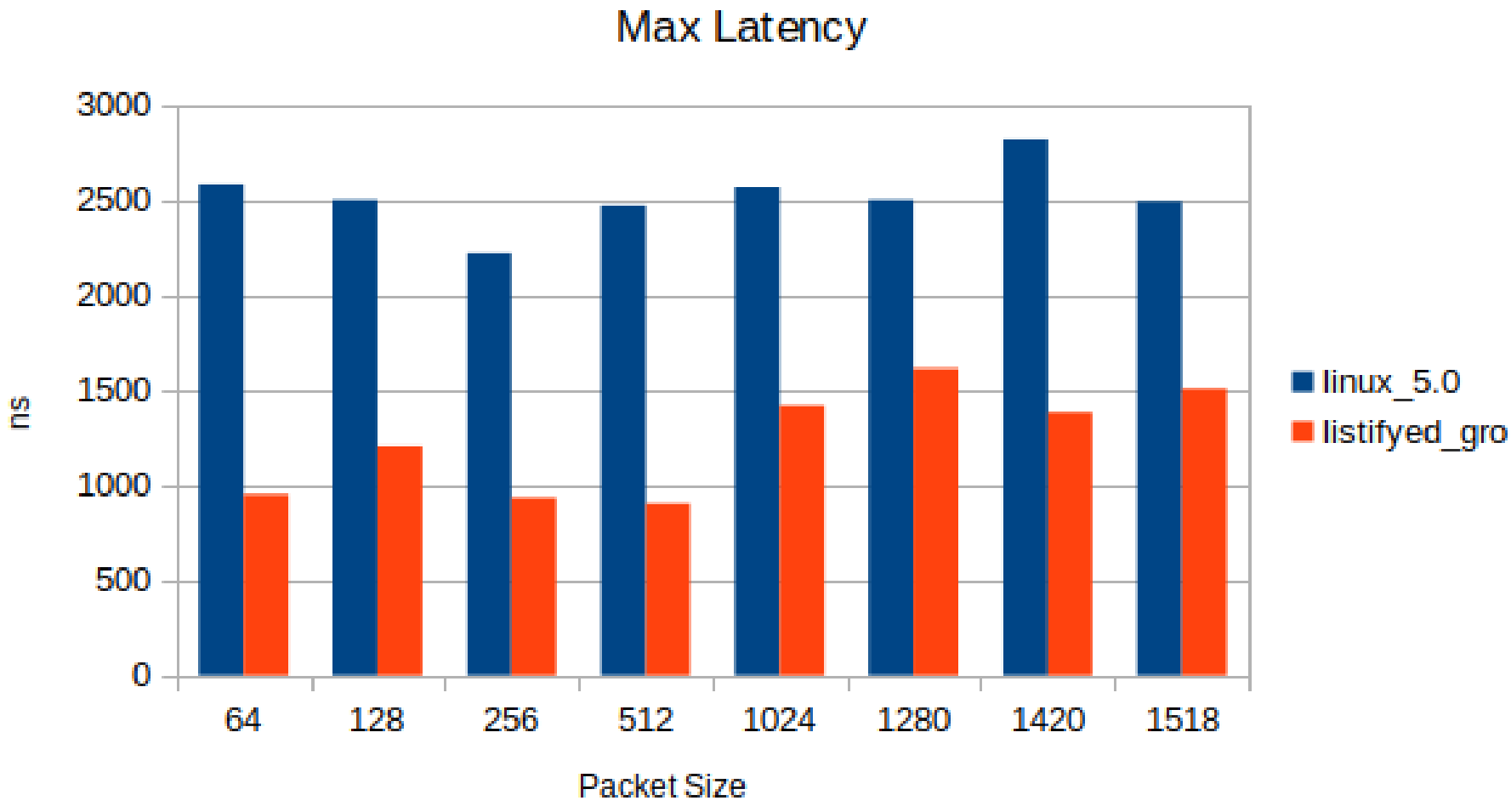
Avg Latency at Max Throughput



Min Latency at Max Throughput



Max Latency at Max Throughput



Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?

Status of listifyed GRO

- Implemented for UDP (RFC state)
- TCP forwarding could benefit too if software GSO is needed
  - Should listifyed GRO support TCP too?
- Problem: How to decide between standard and listifyed GRO?
  - Global knob to enable/disable
  - Do a route lookup at GRO layer
    - Gives a good guess on output NIC and xfrm
    - But: Could be rerouted by netfilter etc.
  - Add some TC/NF hook before GRO that can add route infos
    - NF flowtable?
    - eBPF programm?