Chair for Network Architectures and Services – Prof. Carle
Department of Computer Science
Technische Universität München

Dynamic Traffic Engineering in the Future(?) Internet

3rd GI-ITG/KuVS-Fachgespräch "Future Internet"

Nils Kammenhuber

kammenhuber@net.in.tum.de

DFG Schwerpunkt Nr. 1126
Algorithmik großer und komplexer Netzwerke

Structure

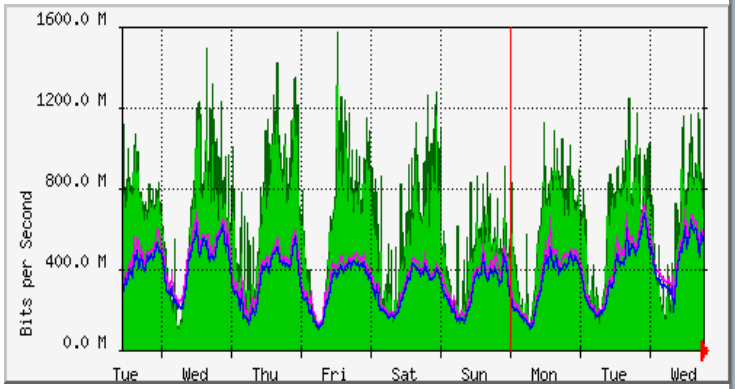
- Motivation: Traffic changes
- ReplEx algorithm for dynamic TE
- Evaluation in today's Internet
- ReplEx in the Future Internet
- Summary

- Spare slides:
  - Game theoretic background
  - How the algorithm works
  - Avoiding packet reordering

Motivation:
Traffic changes are inevitable

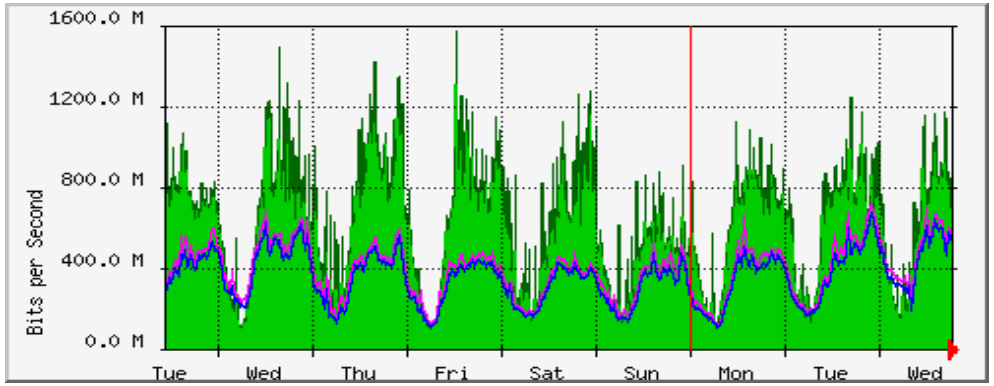
Traffic variability

- Network traffic is known to change
  - Circadian effects
  - Self-similar traffic properties
    ⇒ statistical spikes
  - Flash crowds
  - Worm/virus outbreaks
  - (D)DoS attacks
  - …



- Reason:
  Humans are non-linear, non-deterministic.
  - Future Internet: still made for humans
  - ⇒ **Traffic variability will not vanish!**

Variable traffic = bad



- *Some* traffic variability is regular/circadian
- *Most* traffic variability cannot be anticipated
- Traffic variability is the path to service degradation
  (… also in the Future Internet!)
  1. Traffic variability leads to traffic spikes
  2. Traffic spikes lead to congestion
  3. Congestion leads to service degradation
- Need to avoid this!

Handling variable traffic in today's Internet

- TCP congestion control
  - Responsible: End hosts
  - ✓ Good: Short timescale (sub-seconds)
  - × Bad: Only along one path; no alternative paths taken into account
- Traffic engineering
  - Responsible: Individual networks along path / network operators
  - ✓ Good: Entire(?) network topology taken into account
  - × Bad: Long timescale (hours)

Why not dynamic?

Dynamic routing: Mostly frowned upon!

- Bad experiences in ARPANET (NCP)
  - Prone to overreactions
  - Prone to oscillations
- Additional signalling traffic
- Dynamic mechanism
  ⇒ harder to debug
  ⇒ harder to manage
- IGP metric is a parameter to BGP,
  and BGP is already unstable enough — no thanks

Solution: ReplEx

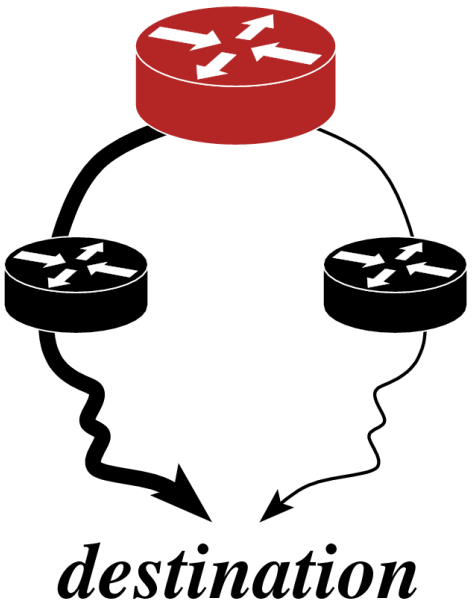
Solution idea: What we have (1/2)

ReplEx algorithm for dynamic traffic engineering

- Dynamic: reacts within seconds
- Distributed
- Oscillation-free; based on game-theory
- Only small signalling overhead (DV-like)
- Easy to implement
- Simple administration
- Evaluated with realistic self-similar TCP traffic in realistic IP backbone topologies

Solution idea: What we have (2/2)

- Not a routing protocol, but a traffic engineering protocol
- Sits on top of existing routing protocol(s)
- Balances traffic among equal-cost routes
- Traffic information:
  - Local measurements
  - Distance vector like protocol



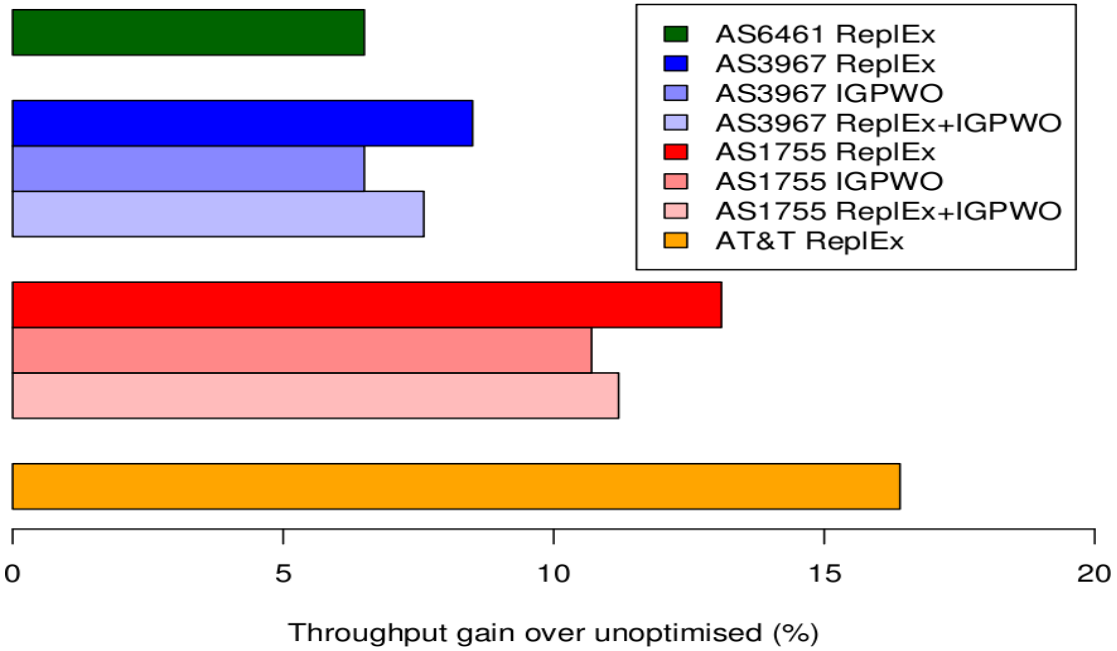
Evaluation results
(in today's Internet)

Evaluation results (1)

- Evaluation setup
  - SSFNet simulator
  - ISP topologies from Rocketfuel and TOTEM
  - OSPF routing provides multipath routes
  - Realistic traffic
    - Real TCP, no packet replays
    - Fractal flows (SURGE model)
- Comparison with standard OSPF link optimization (traditional static TE)

Evaluation results (2)

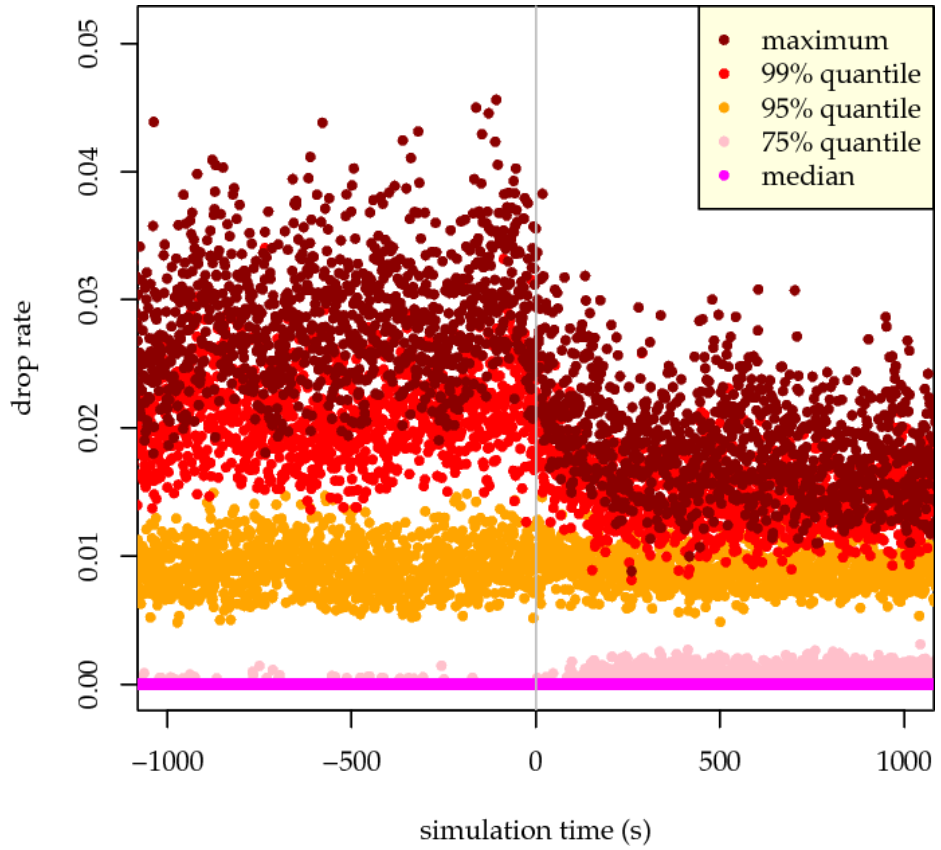
Throughput gains:



Topology	Method	Gain
1755	IGPWO	10.7%
1755	**RepIEx**	**13.1%**
3967	IGPWO	6.5%
3967	**RepIEx**	**8.5%**
AT&T	**RepIEx**	**16.4%**

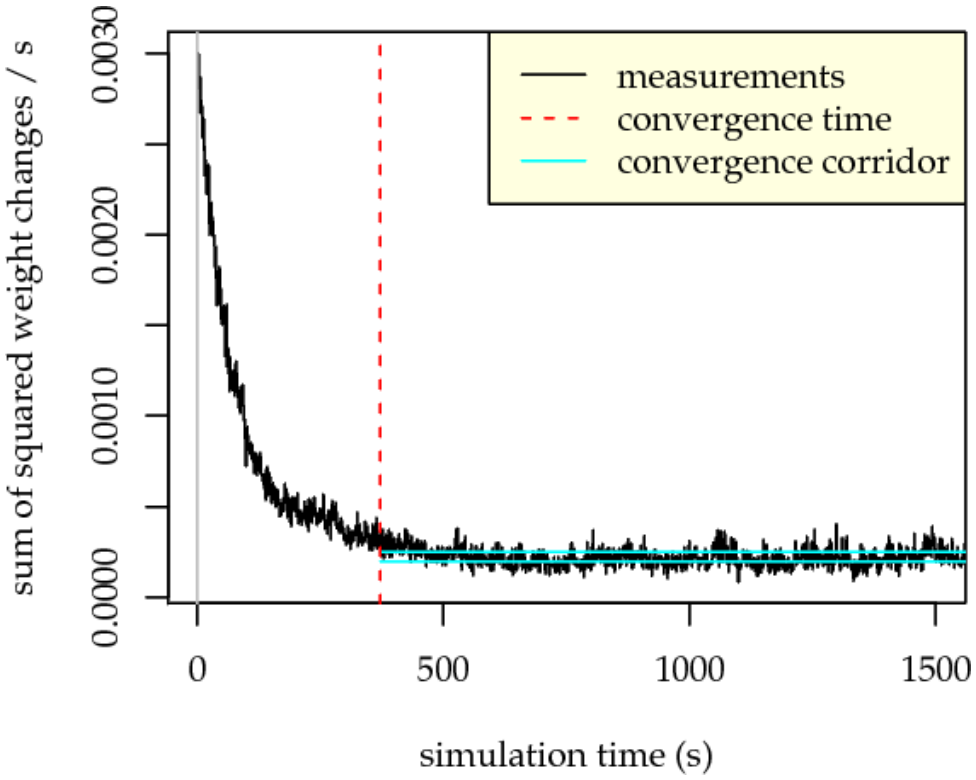
Evaluation results (3)

Quick reduction of packet losses:



Evaluation results (3)

Quick reduction of packet losses:



ReplEx for a Future Internet

Why ReplEx is ready for the Future Internet

- ReplEx = extremely flexible and generic
- Just needs multipaths to choose from; works with any underlying routing protocol
  - Tested with intradomain OSPF
  - Can work with BGP, MPLS, wireless routing, layer-2 switching, …
    or **arbitrary future routing architectures!**
- Metric to optimise = fairly arbitrary
  - So far: link load, packet drop probability
  - Why not: Reliability? Trust? Battery power? Combination of metrics?
    … whatever is needed by a future Internet!

RepIEx in a Future Internet: Popular F.I. aspects (1)

- Locator-ID split: ✓
  - ✓ Dumb forwarding ⇒ no ID functionality needed
- Sensor networks: ✓
  - ✓ Simple calculations, linear with # route destinations
  - ✓ Low signalling overhead
- Mobile networks: ✓
  - Communication with neighbourhing RepIEx routers must not be disturbed to great extent
  - Drastic topology changes not too frequent
  - ✓ Can include mobility metrics! (e.g., battery power)
- Overlays and virtual networks: ✓
  - Just need multipath alternatives

RepIEx in a Future Internet: Popular F.I. aspects (2)

- Security: ✓
  - × Most metrics require co-operation between nodes
  - ✓ Can include “trust”, “security” into optimization metric
- Resilience: ✓
  - ★ Being investigated in EU FP7 project *ResumeNet*
  - ✓ Can help mitigating traffic bursts (attacks, worms,...)
  - ✓ Can include “reliability” into metric
  - ● Reaction timescale = seconds, but not much less
  - × No immediate 100% avoidance of congested path!

RepIEx in a Future Internet: Popular F.I. aspects (3)

- Flow routing/switching: ✓
  - ✓ Information on flow properties can be useful
  - Need enough flows (or splittable flows) for traffic ratios
- Quality of service: (✓)
  - ✓ Call admission (i.e., flow) information can be useful
  - × RepIEx makes no performance guarantees!
  - ... unless metric is constructed to do so
- Cross-layer design, disappearance of layering: ✓
  - Need source with multipath alternatives
  - Need ability to arbitrarily split traffic
  - ✓ Additional “functionality brick” for dynamic TE

RepIEx in a Future Internet: Popular F.I. aspects (4)

- Business considerations, “tussle space”: ✓
  - ✓ Information on network conditions only in aggregated form ⇒ hiding information to outsiders
  - ✓ Flexible metrics ⇒ can include business metrics
- Self-*, Auto-* properties: ✓
  - ✓ Set some global parameters for all nodes—that’s it!

Almost done

Summary and Outlook

- ReplEx = simple yet effective algorithm for dynamic traffic engineering
- Distributed, low signalling overhead, low administrative overhead
- No oscillations

- Can be applied in any network that features multipaths — no matter what the Future Internet will look like
  (Your design still lacks dynamic routing/traffic engineering? Simply plug in ReplEx!)

Nils Kammenhuber:
Traffic-Adaptive Routing.
Dissertation, Technische Universität München 2009.

Simon Fischer, Nils Kammenhuber, Anja Feldmann:
REPLEX—Dynamic Traffic Engineering Based on Wardrop Routing Policies.
Proceedings of ACM CoNext, Lisboa, 2006.

ReplEx home page (with more references/papers):
`http://www.net.in.tum.de/~hirvi/replex/`

Ongoing work in EU FP7 project *ResumeNet.*

Questions?

DFG
Schwerpunkt Nr. 1126
Algorithmik
großer und komplexer Netzwerke

Game theory

From theory to practice (1)

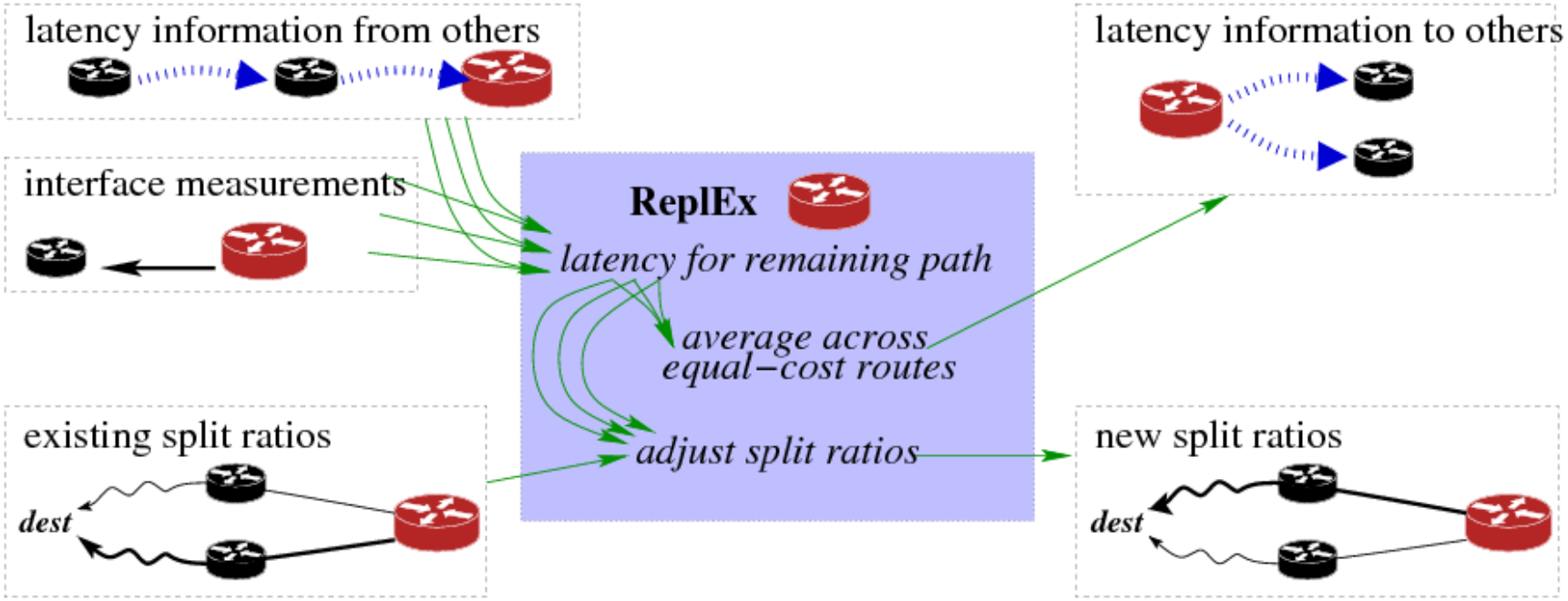
- Wardrop model
  - Infinite number of agents at network edges
  - Each controls infinitesimally small traffic share
  - Each selfishly optimizes its traffic performance
  - How? → agent's *strategy*
- Exploration-replication strategy
  - Agents continuously watch other agents
  - Imitate behaviour of more successful agents with certain probability
  - Provably converges without oscillations
- Nice theory, but not applicable in Internet:
  - Destination-based routing—not only at edges!
  - No information sharing (competing flows!)

From theory to practice (2)

- Aggregate behaviour of individual agents as probability
  - Probabilistic routing at one node
  - I.e., changing multipath ratios at a router
  - Many agents → one router
- Distribute decisions of Wardrop agents at edges along network path
  - One agent → multiple routers

➔ 1 Wardrop agent ≠ 1 ReplEx router!

From theory to practice (3)



Avoiding packet reordering

Avoiding packet reordering

- Forward packets to one of the multipath routes according to ratio (i.e., probability)
- Must not reorder packets within TCP stream
- Standard method: calculate hash of packet header, assume to be random number
- Not truly random?
  No problem: RepIEx fixes resulting overload!