Traffic Engineering with Forward Fault Correction

Harry Liu

Microsoft Research

06/02/2016

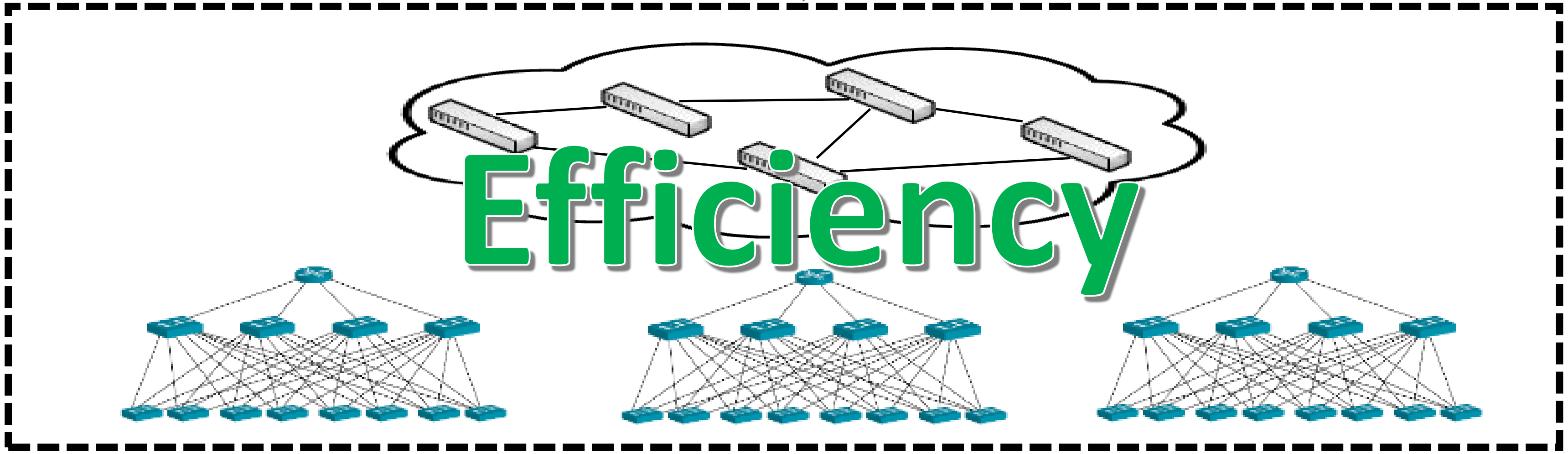
Joint work with Ratul Mahajan, Srikanth Kandula, Ming Zhang and David Gelernter

Cloud services require large network capacity

Cloud Applications

Growing traffic

Cloud Networks

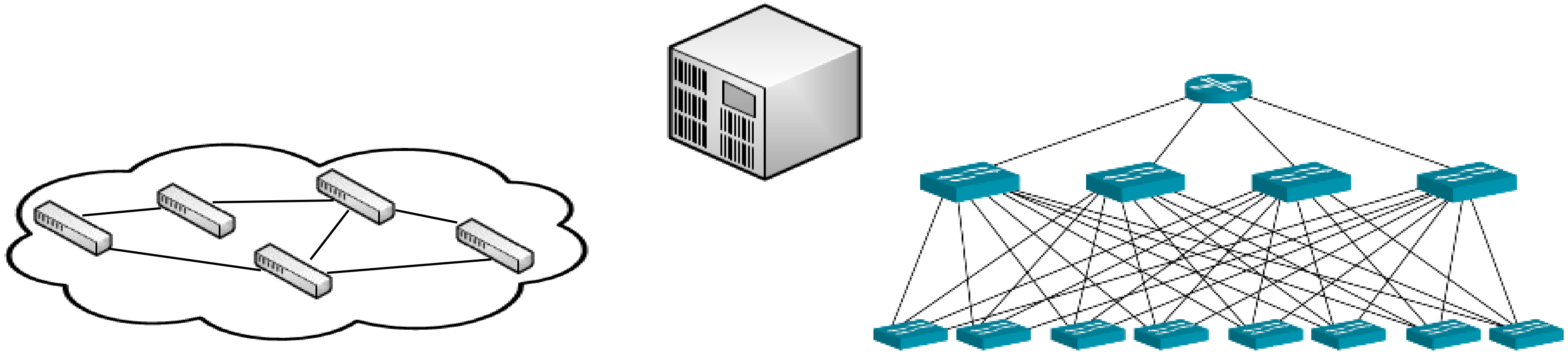


Expensive

(e.g. cost of WAN: $100M/year)

TE is critical to effectively utilizing networks

Traffic Engineering
(centralized & SDN-Based)



WAN Network

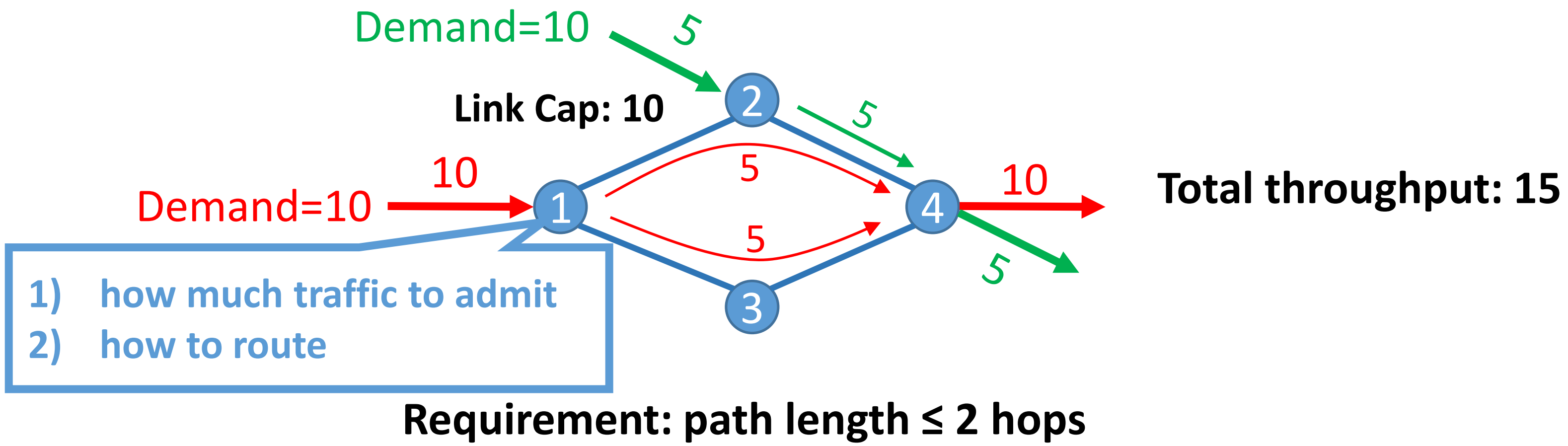
Datacenter Network

- Microsoft SWAN (SIGCOMM'13)
- Google B4 (SIGCOMM'13)
- ……

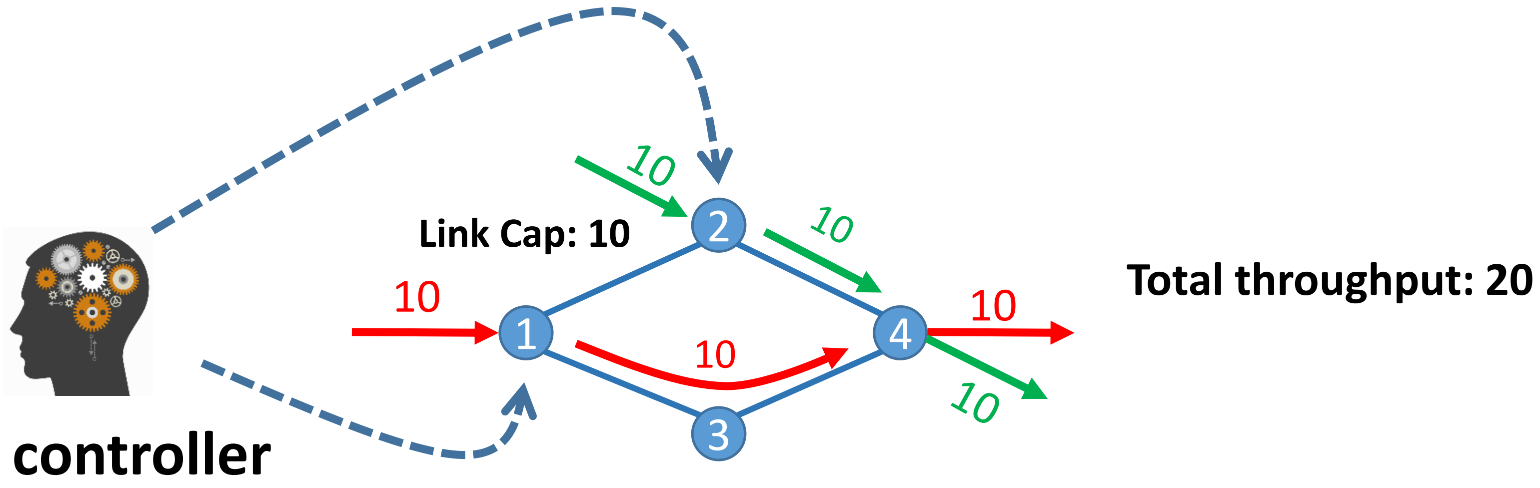
- Devoflow (SIGCOMM'11)
- MicroTE (CoNEXT'11)
- ……

Centralized TE is the key to network efficiency

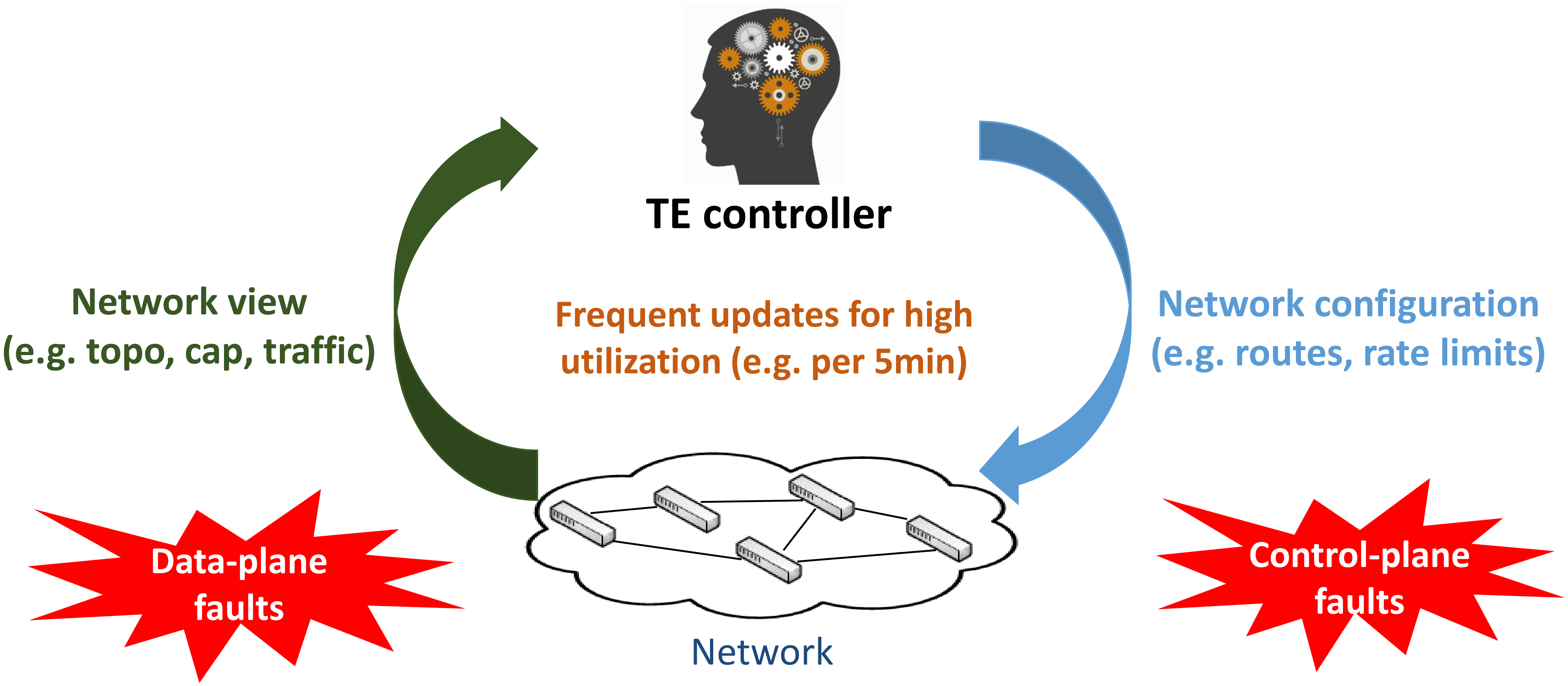
Sub-optimal resource allocation based on local view & control.



Optimal resource allocation based on global view & control.

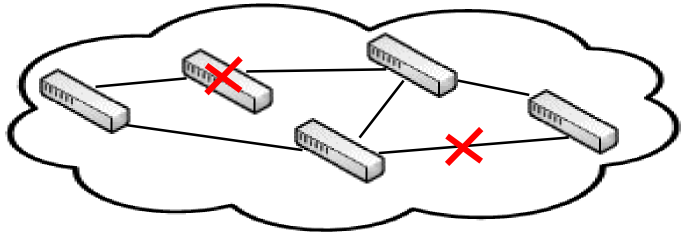


But, centralized TE is also vulnerable to faults

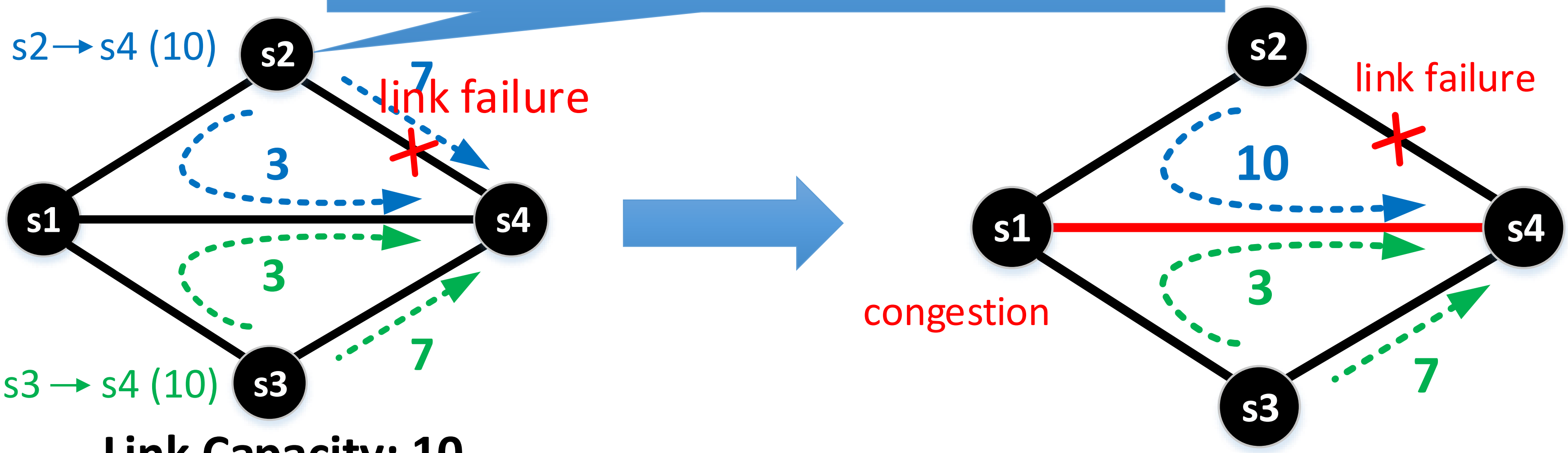


Data plane faults

Link and switch failures



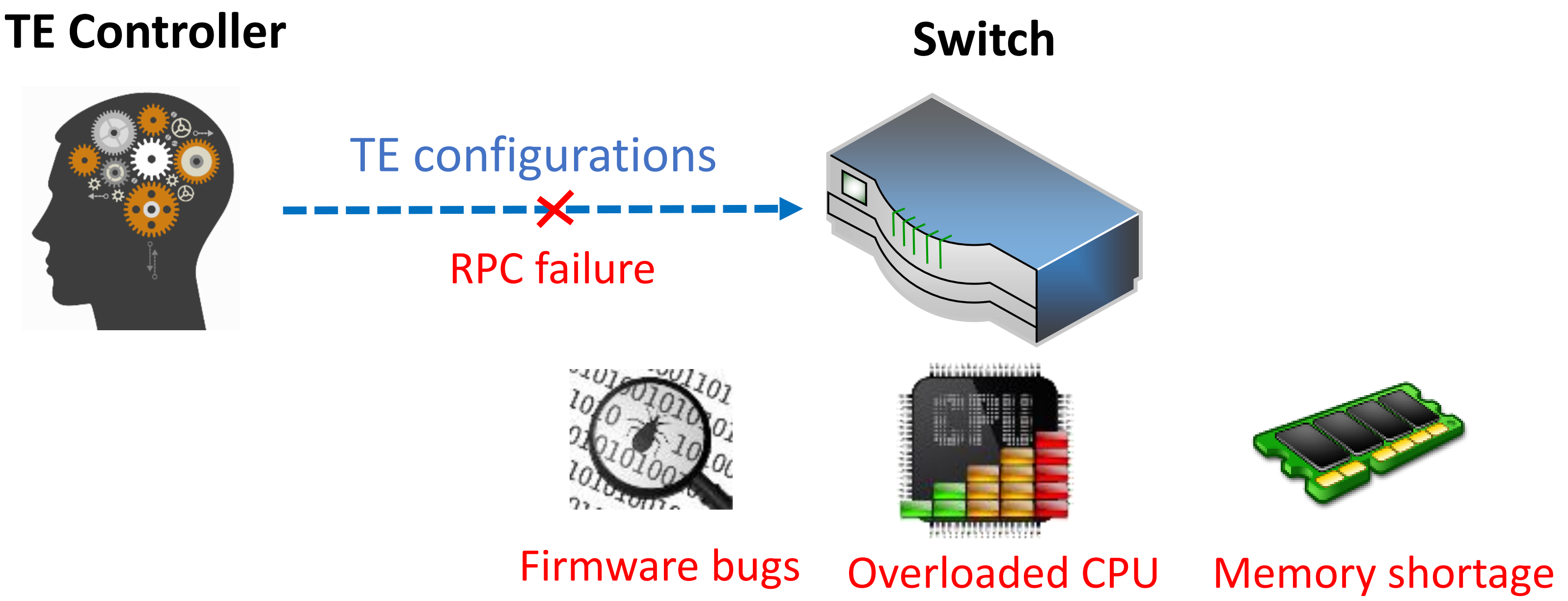
Rescaling: Sending traffic proportionally to residual paths



Link Capacity: 10

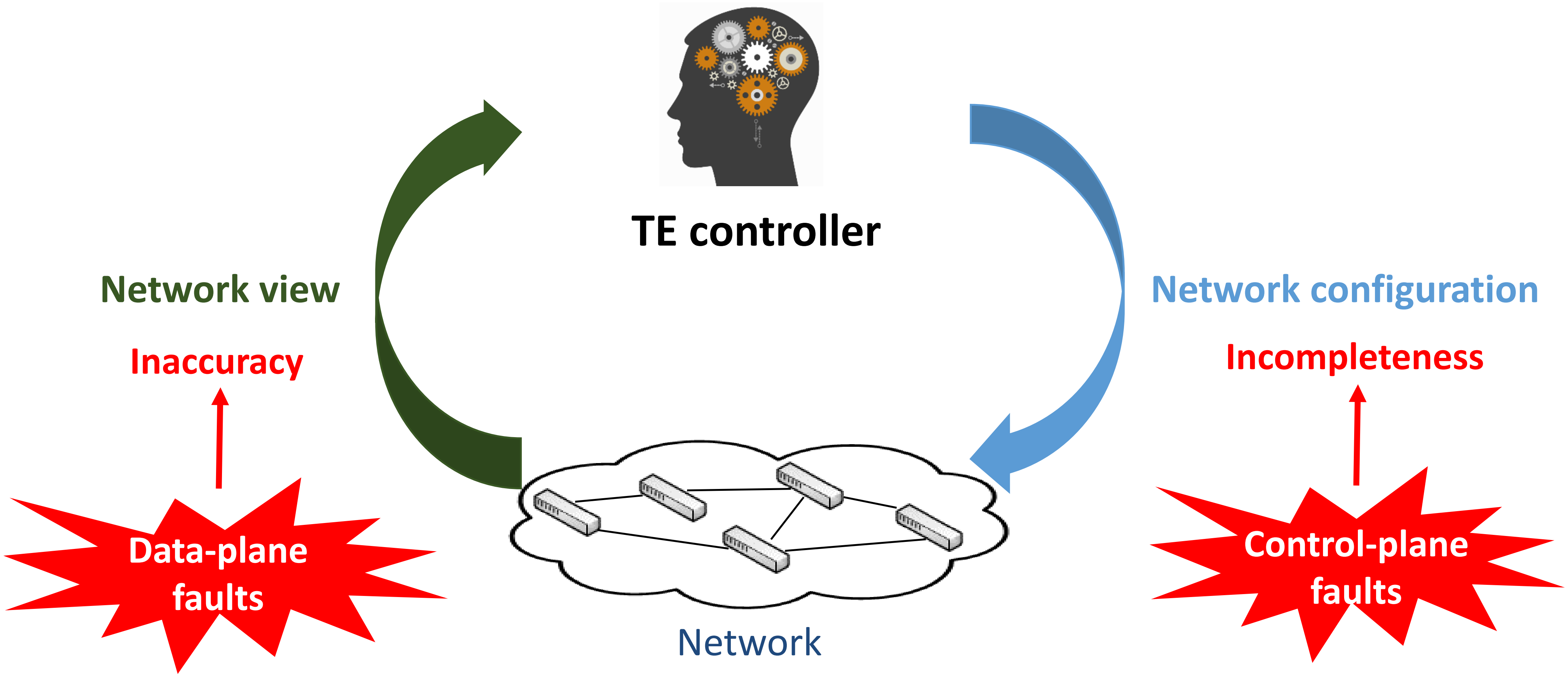
Control plane faults

Failures or long delays to configure a network device



Control plane faults can also result in congestion.

The TE controllability is undermined by faults



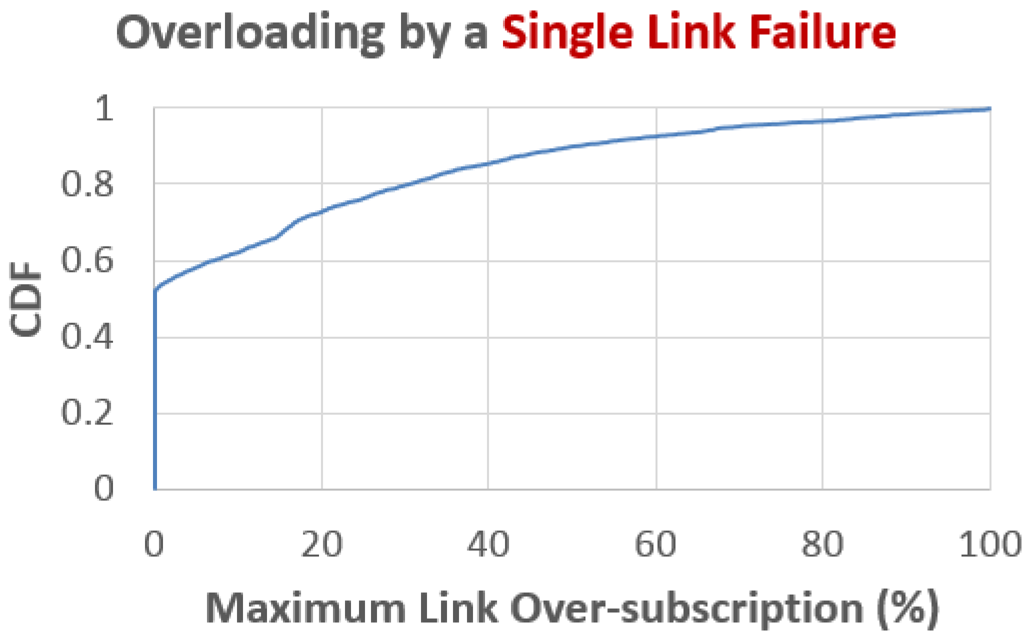
Control and data plane faults in practice

In a production WAN network
(200+ routers, 6000+ links):

- Faults are common.
- Faults cause severe congestion.

Data plane:
fault rate = **25%** per 5 minutes.

Control plane:
fault rate = **0.1% -- 1%** per TE update.



State of the art for handling faults

- Heavy over-provisioning:

Big loss in throughput

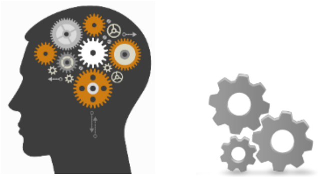
- Reactive handling of faults:
  - Control plane faults: retry
  - Data plane faults: re-compute TE and update networks

Cannot prevent congestion

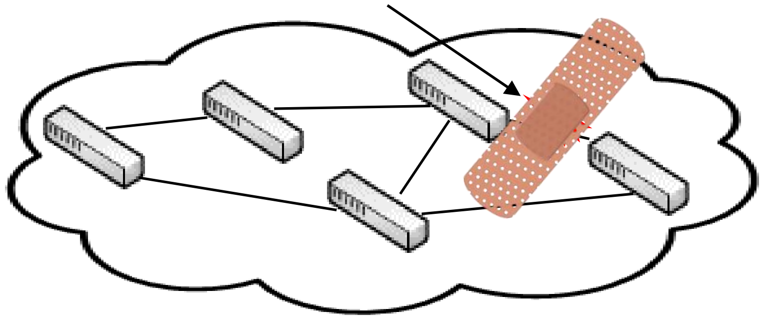
Slow
(seconds -- minutes)

Blocked by control plane faults

How about handling faults proactively?



TE Algorithm



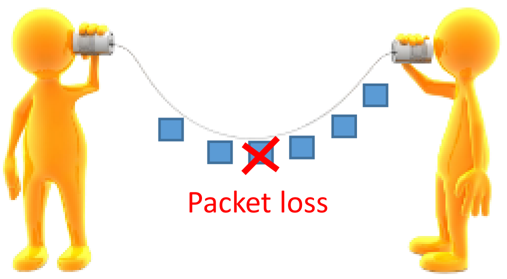
Network

making it robust

not robust enough

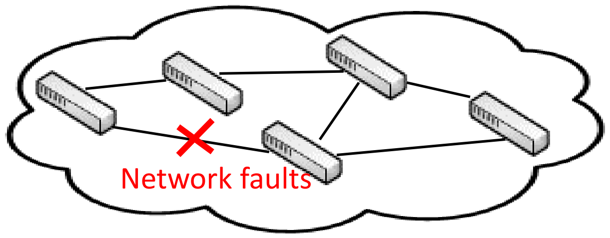
Forward fault correction (FFC) in TE

- [Bad News] Individual faults are **unpredictable**.
- [Good News] Simultaneous #faults is **small**.



FEC guarantees no information loss under up to k arbitrary packet drops.

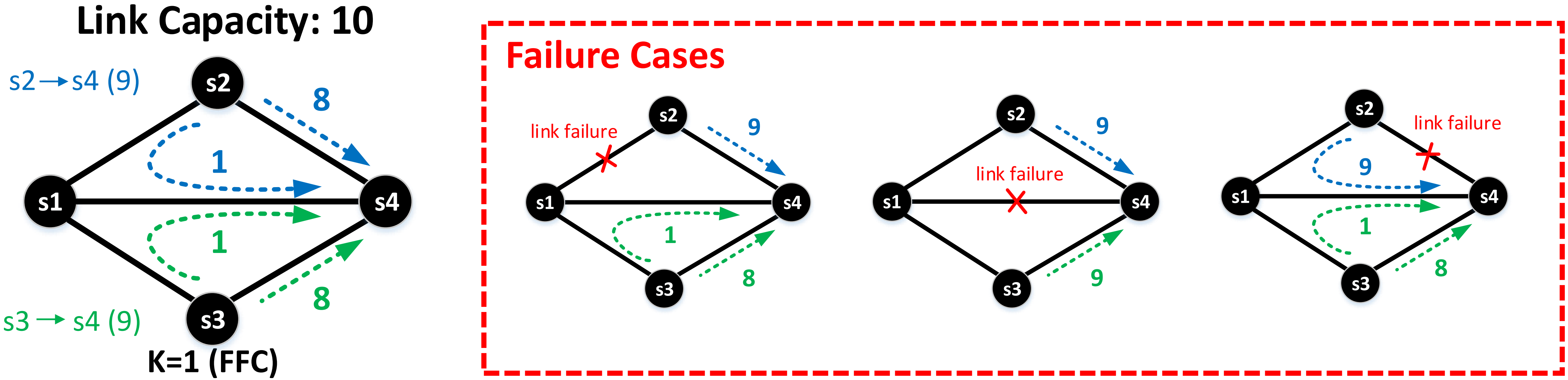
with careful **data encoding**



FFC guarantees no congestion under up to *k* arbitrary faults.

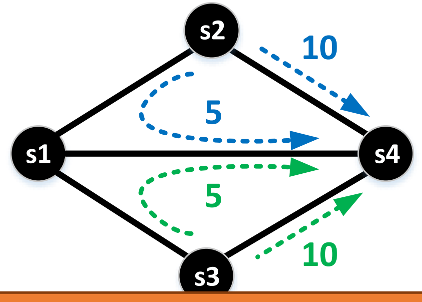
with careful **traffic distribution**

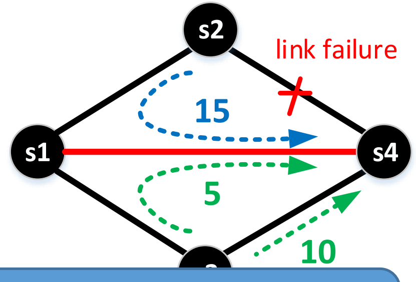
Example: FFC for link failures



Trade-off: network efficiency v.s. robustness

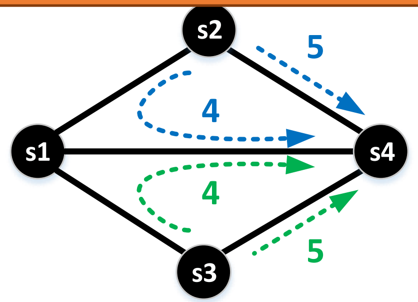
Non-FFC
(Throughput: 30)

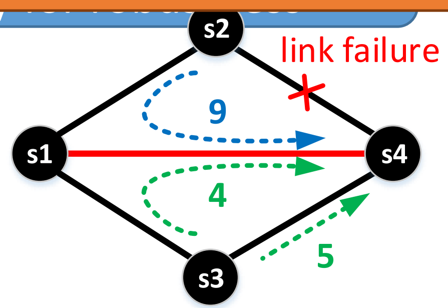




Achieving the **optimal throughput** with FFC guarantee

Non-FFC
(Throughput: 18)





Systematically realizing FFC in TE

Formulation:
How to merge FFC into existing TE framework?

Computation:
How to find FFC-TE efficiently?

Basic TE linear programming formulations

LP formulations

TE decisions:
- Sizes of flows — b_f
- Traffic on paths — $l_{f,t}$

TE objective: Maximizing throughput — max. $\sum_{\forall f} b_f$

Basic TE constraints:
- Deliver all granted flows — s.t. $\forall f$: $\sum_{\forall t} l_{f,t} \geq b_f$
- No overloaded link — $\forall e$: $\sum_{\forall f} \sum_{\forall t \ni e} l_{f,t} \leq c_e$
- … — …

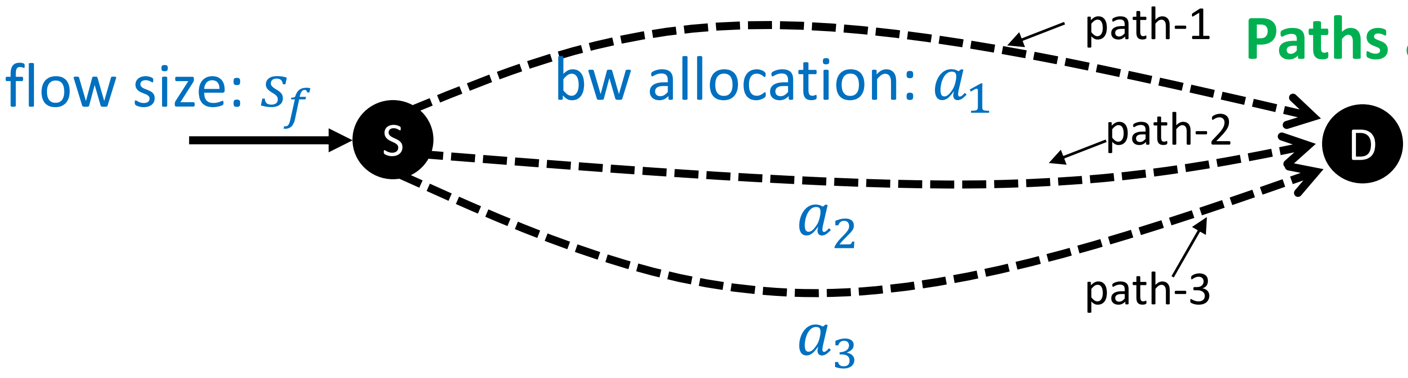
+

FFC constraints:

No overloaded link up to
- k_c control plane faults
- k_e link failures
- k_v switch failures

+

Formulating data-plane FFC



Paths are link-disjoint.

FFC k=1

Fault on path-1: $s_f \leq a_2+a_3$

Fault on path-2: $s_f \leq a_1+a_3$

Fault on path-3: $s_f \leq a_1+a_2$

$\binom{3}{2}$

Lemma: FFC is achieved when path-i's weight is $a_i/a_1+a_2+a_3$

An efficient and precise solution to FFC

k-sum linear constraint group (k-sum group):

Given ***n*** paths and $A = \{a_1, a_2, \dots, a_n\}$, FFC requires that
the sum of ***arbitrary n-k*** elements in A is $\geq$ flow size

$O(\binom{n}{k})$

FFC-TE LP-formulation:

TE Objective

Basic TE Constraints

k-sum group-1

⋮

k-sum group-N

FFC Constraints (too many)

Lossless compression of a k-sum group:

$O(\binom{n}{k})$

- O(kn) bubble sorting network (SIGCOMM 2014)
- O(n) strong duality (MSR TR 2016)

http://www.hongqiangliu.com/publications.html

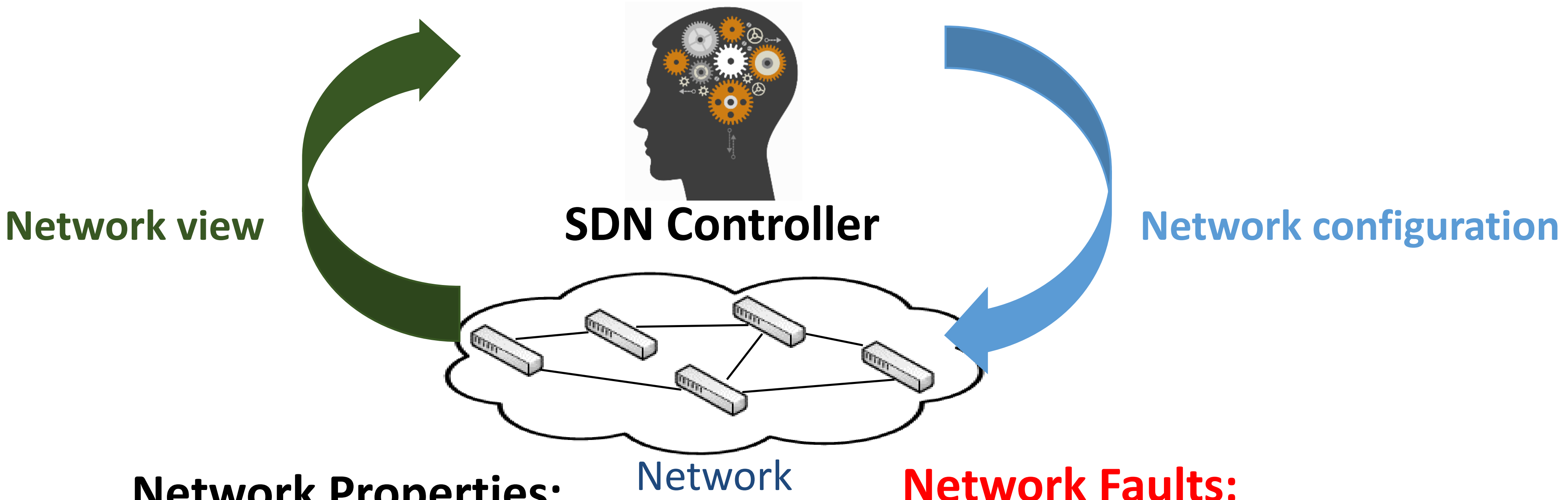
FFC extensions

- Differential protection for different traffic priorities
- Minimizing congestion risks without rate limiters
- Control plane faults on rate limiters
- Uncertainty in current TE
- Different TE objectives (e.g. max-min fairness)
- …

Implementation & evaluation highlights

- Testbed experiment (8 switches & 30 servers)
  - FFC can be implemented in commodity switches
  - FFC has no data loss due to congestion under faults
- Large-scale simulation
  - A WAN network with O(100) switches and O(1000) links
  - One-week traffic trace
  - Fault injection according to real failure trace
  - Results: with negligible throughput loss, FFC can reduce
    - data loss by a factor of 7-130 in well-provisioned networks
    - data loss of high priority traffic to almost zero in well-utilized networks

Conclusion and future work



Network Properties:

- High throughput
- No congestion
- Security
- Availability
- Connectivity
- ……

Network Faults:

- Data-plane
- Control-plane
- Misconfigurations
- Attacks
- Traffic spikes
- ……

FFC

Q&A