

# Targeting Situation Awareness: The Challenges of Traffic Observation

Dr. Tanja Zseby  
Competence Center Network Research  
Fraunhofer Institute FOKUS  
Berlin



Fraunhofer

Institute for Open  
Communication Systems

SpoVNet Workshop  
July 17, 2008

# Network Measurements

---

- Basis for network research
  - But: Scientists from other disciplines are shocked:

*“[network research] had not managed to execute the usual elements of successful research... **measure**, model, make prototypes”*

*Source: Looking over the Fence at Networks, National Research Council, Washington DC, 2001*

- Essential for network operation
  - Network management
  - Network security
  - Support for service provisioning
- Required for Future Internet



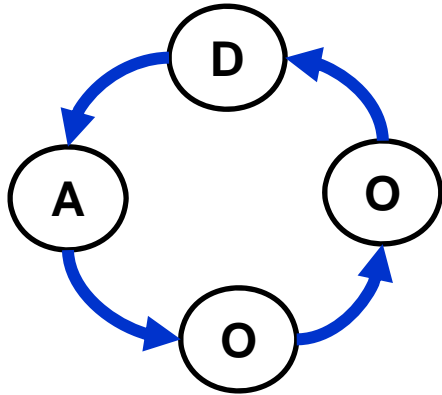
# Future Internet Promises

---

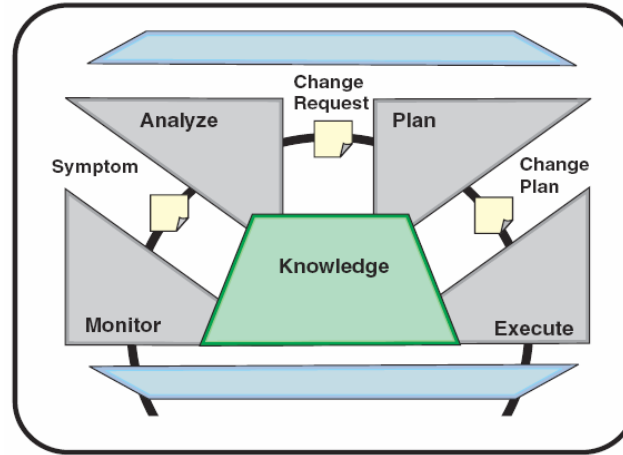
- Self-\* (self-management, self-protection)
  - Reducing human intervention
  - Providing mobility, security, QoS, GreenIT,...
- Evolutionary and revolutionary approaches
  - Overlays → Virtualization
  - In-network management
  - Functional composition
  - Business/Application-oriented networking
  - And many more

**Decision cycles within networks**

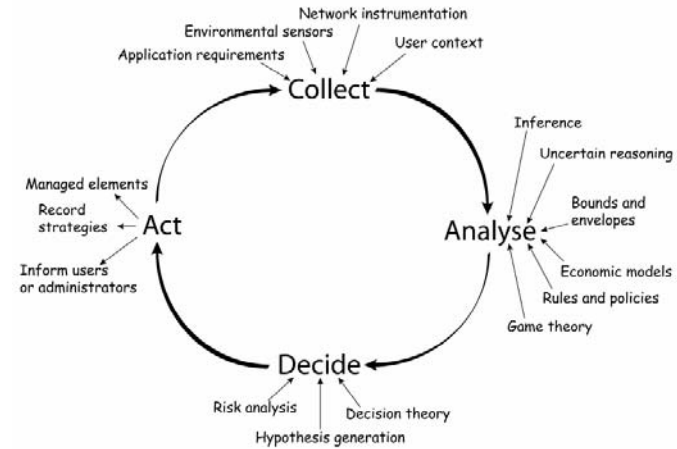
# Decision Cycles



Human [Boyd]: OODA  
Observe-Orient-Decide-Act



Autonomic Computing [IBM]: MAPE  
Monitor-Analyze-Plan-Execute



Autonomic Communication [Dobson]:  
Collect-Analyse-Decide-Act

**Establish Situation Awareness**

# Situation Awareness

---

- A state of knowledge of the situation
  - Constantly evolving picture of the state of the environment
  - 3 levels: Perception, Inference, Prediction

**Basis for sound decisions**

# How to Achieve Situation Awareness?

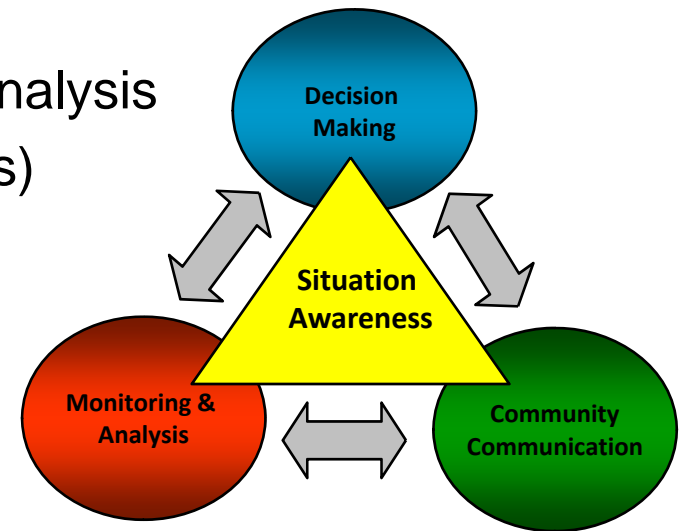
---

- **Monitoring & Analysis**

- Network measurements at multiple observation points
- Changing viewpoints (zoom in and out)
- Analysis (inference, prediction)

- **Community Communication**

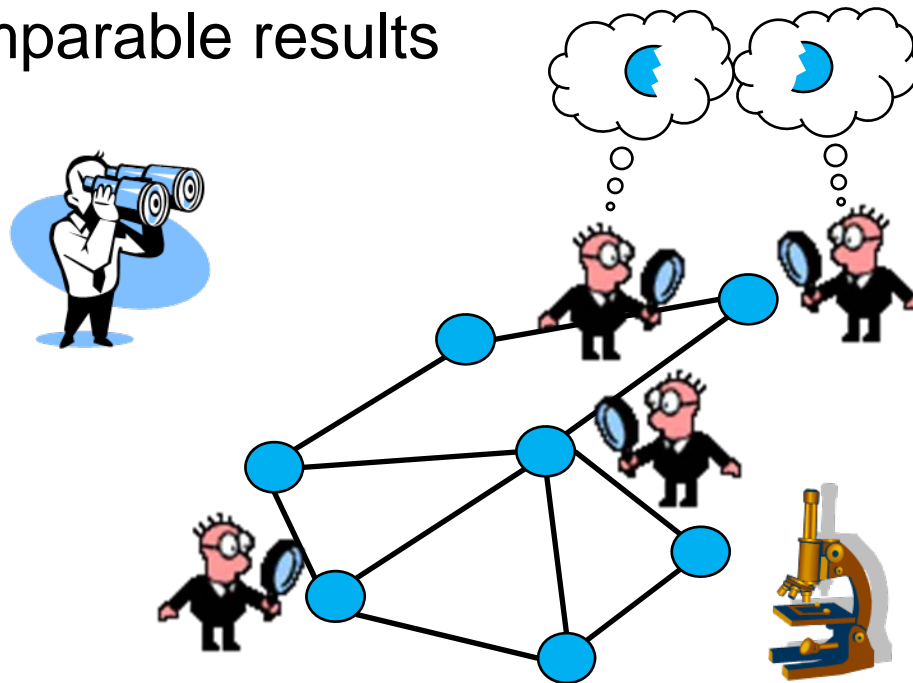
- Share viewpoints
- Share resources for monitoring and analysis
- Share local decisions (analysis results)



# Desired Features for Traffic Observation

---

- **Network-wide:** multiple observation points
- **Flexible:** change viewpoints
- **Shareable:** comparable results



**But: We can not measure everything everywhere!**

# Challenges of Traffic Observation

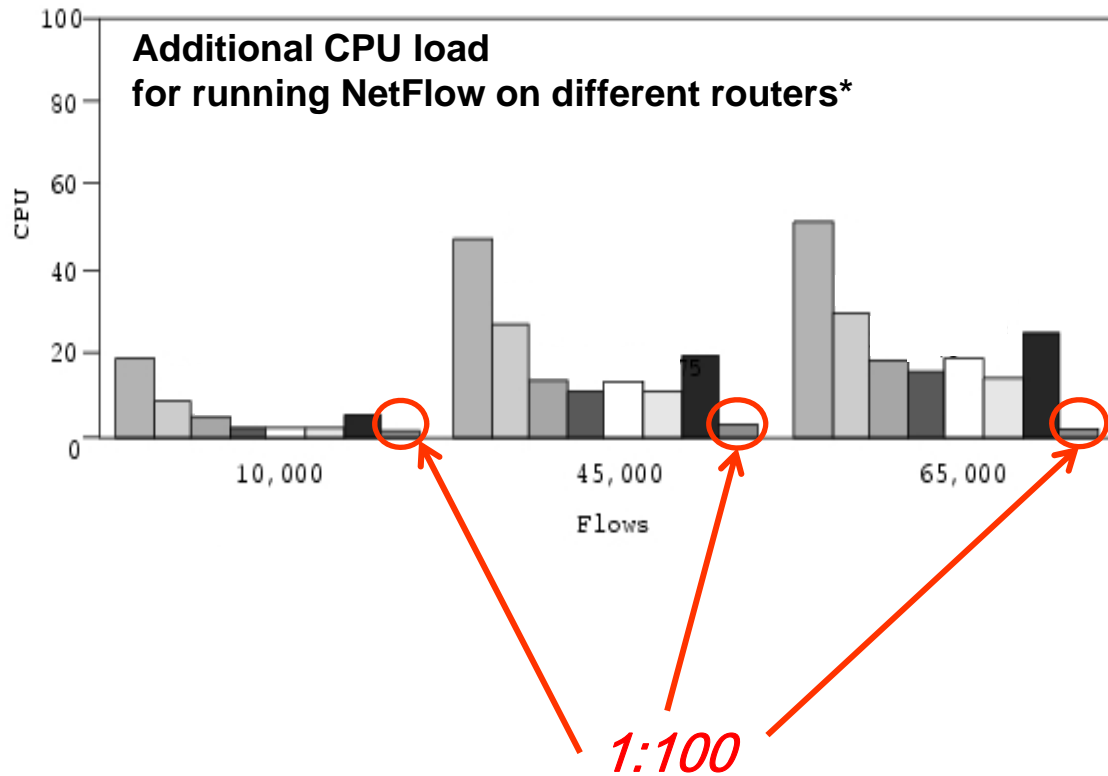
---

- Resource Limitations
  - Resource Consumption
  - Resource Control
  - Measurement adaptation
- Sharing Data
  - Multipoint Measurements
  - Comparability of measurement results
  - Standardized interfaces
  - Privacy
- Protection of measurement system
  - Against overload from traffic fluctuation
  - Against attacks



# Resource Consumption

- Resource Limitations
  - Processing power
  - Transmission
  - Storage
- Demand depends on
  - Data rates
  - Required granularity
- Solutions
  - Dedicated Hardware
  - Improved Algorithms
  - Data Selection



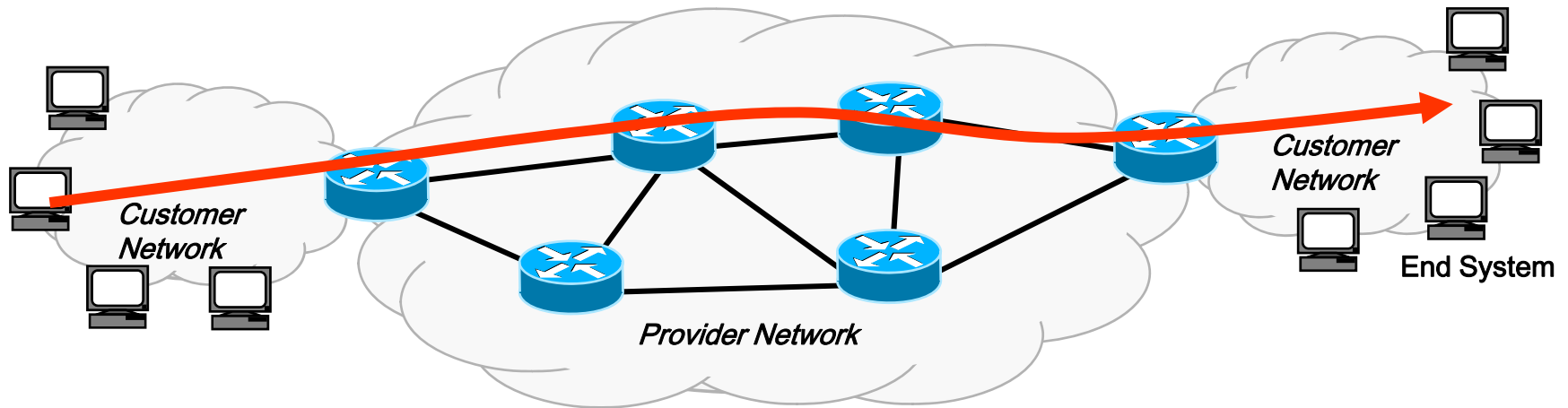
\*source: NetFlow Performance Analysis, Cisco white paper

# Example: Flow Measurements

**Flow:= packets with common properties**

## Examples

- all packets from host A to host B
- all small packets
- all TCP-SYN packets
- all HTTP packets with URL `www.whitehouse.gov`



# Example: Flow Measurements

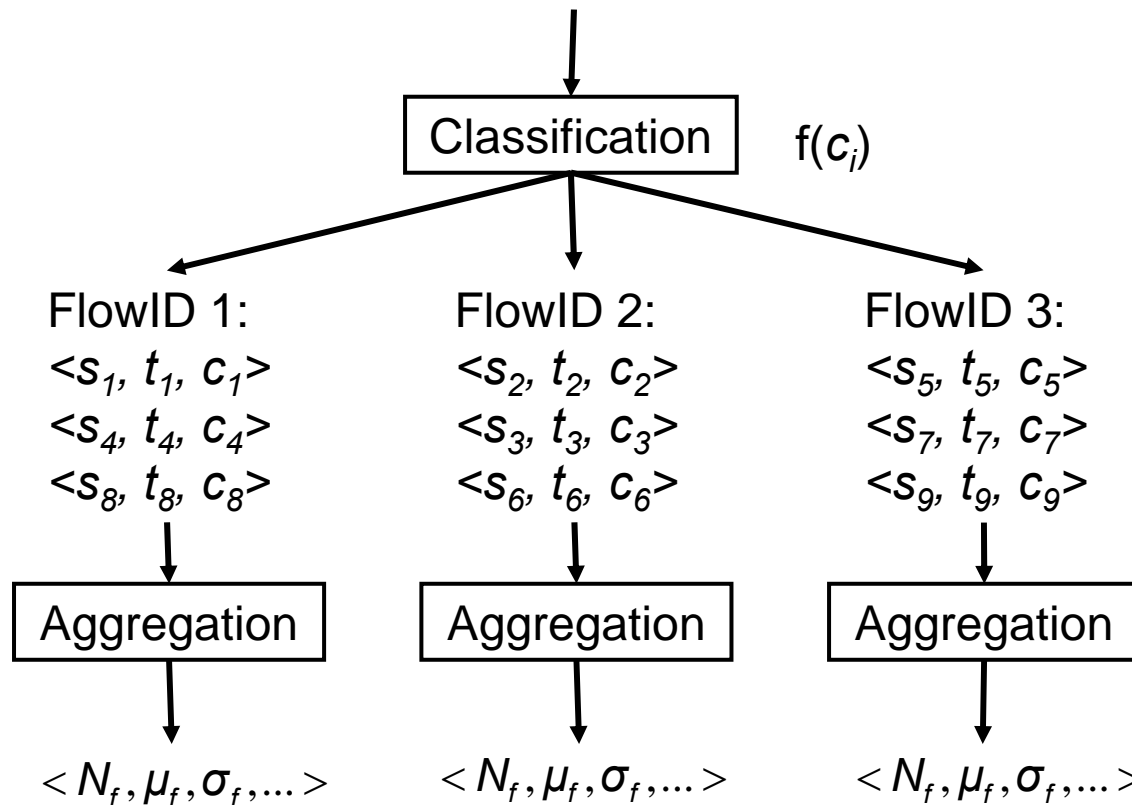
*Packet Attributes at Observation Point*  
 $s_i$  – sequence number  
 $t_i$  – arrival time  
 $c_i$  – content (header, payload)



**Packets:**

$\langle s_1, t_1, c_1 \rangle, \langle s_2, t_2, c_2 \rangle, \dots \langle s_N, t_N, c_N \rangle$

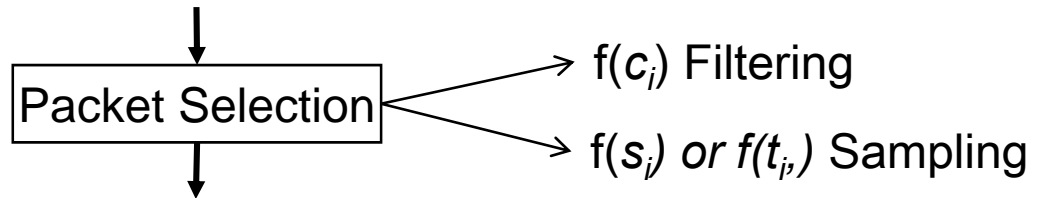
**Flows:**



# Packet Selection

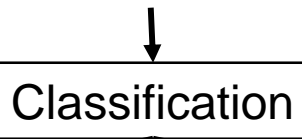
**Packets:**

$\langle s_1, t_1, c_1 \rangle, \langle s_2, t_2, c_2 \rangle, \dots \langle s_N, t_N, c_N \rangle$



**Selected Packets:**

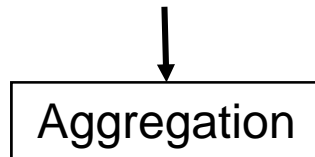
$\langle s_2, t_2, c_2 \rangle, \langle s_6, t_6, c_6 \rangle \dots \langle s_n, t_n, c_n \rangle$



**Flows:**

FlowID 1:

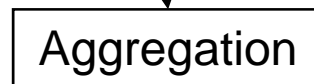
$\langle s_8, t_8, c_8 \rangle$



FlowID 2:

$\langle s_2, t_2, c_2 \rangle$

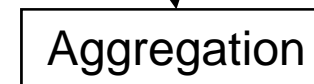
$\langle s_6, t_6, c_6 \rangle$



FlowID 3:

$\langle s_5, t_5, c_5 \rangle$

$\langle s_9, t_9, c_9 \rangle$



**Flow Records:**

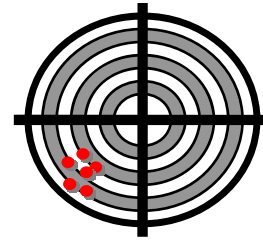
$\langle \hat{N}_f, \hat{\mu}_f, \hat{\sigma}_f, \dots \rangle$

$\langle \hat{N}_f, \hat{\mu}_f, \hat{\sigma}_f, \dots \rangle$

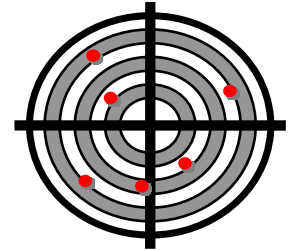
$\langle \hat{N}_f, \hat{\mu}_f, \hat{\sigma}_f, \dots \rangle$

# Problem1: Accuracy Assessment

Accuracy Assessment required



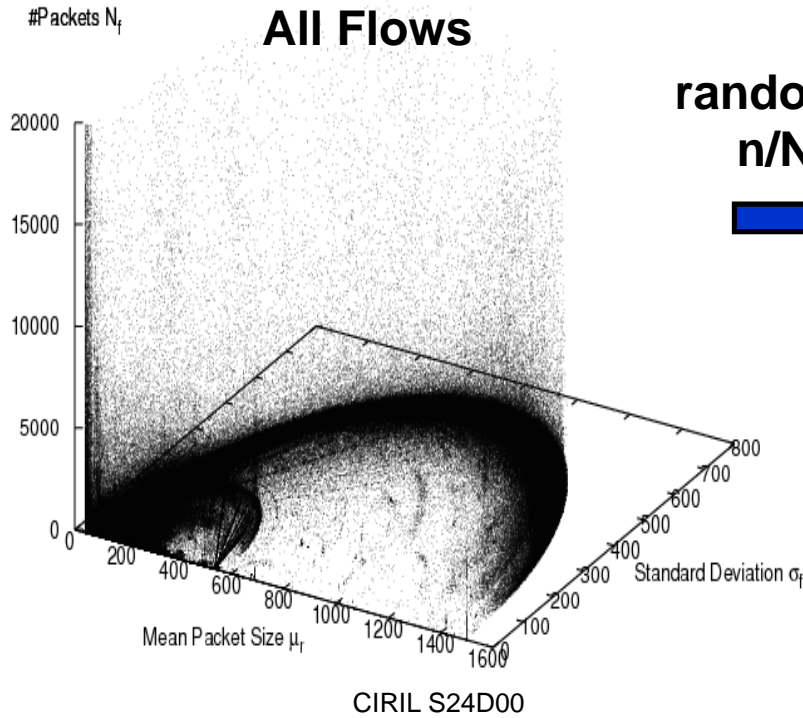
Bias



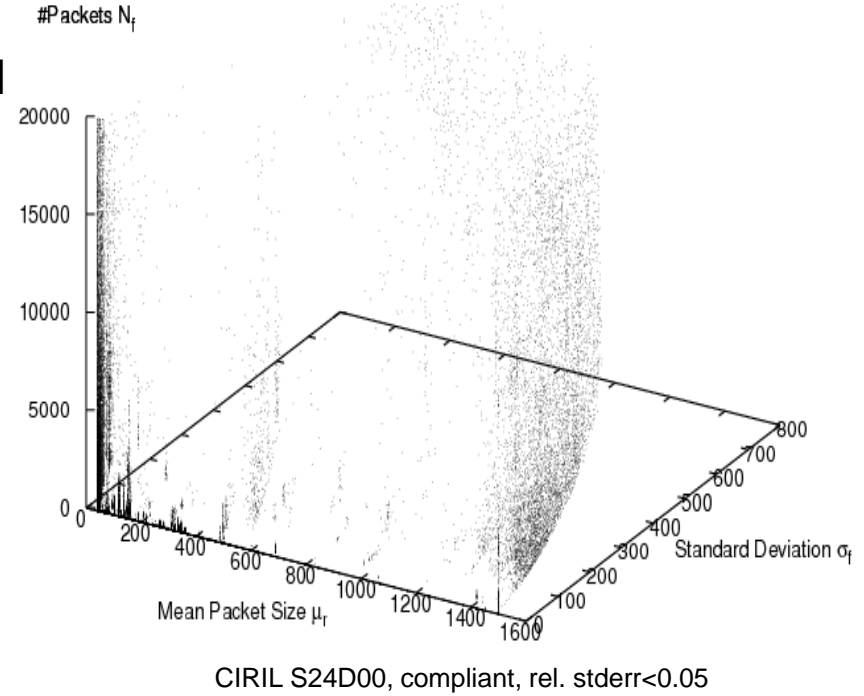
Precision

- Achievable accuracy depends on
  - Sampling and estimation method
  - Sampling parameters
  - Population characteristics ← **unknown and highly dynamic**
- Accuracy assessment **during** measurement
  - For each measurement interval
  - For each flow
  - Based on sampled data

# Estimation Accuracy



random nofN  
 $n/N=5\%$



$$StdErr_{rel} = \frac{1}{N_f \cdot \mu_f} \cdot \sqrt{\frac{N \cdot N_f \cdot (\sigma_f^2 + \mu_f^2)}{n} - \frac{N_f^2 \cdot \mu_f^2}{N}}$$

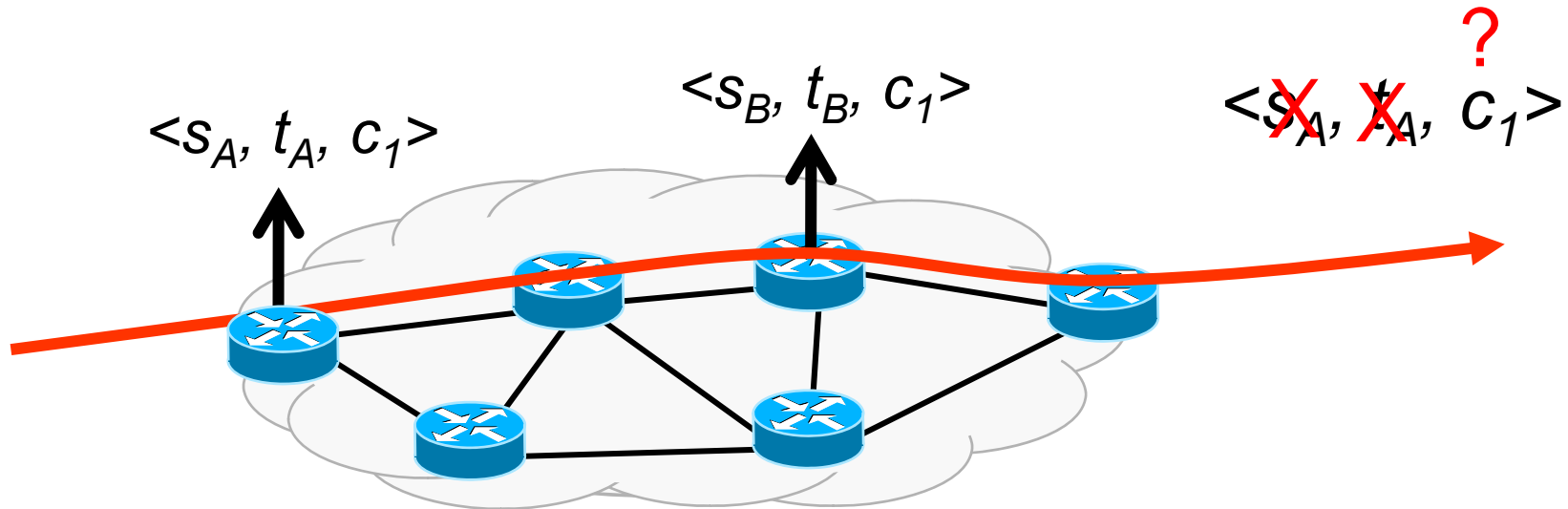
**NetFlow Records**  $N_f, \sum x_f$   $\sum x_f^2$

**Store sum of squares !**

VEGAS project, funded by Cisco Systems

# Problem2: Multipoint Data Selection

Goal: Select same packet at different observation points



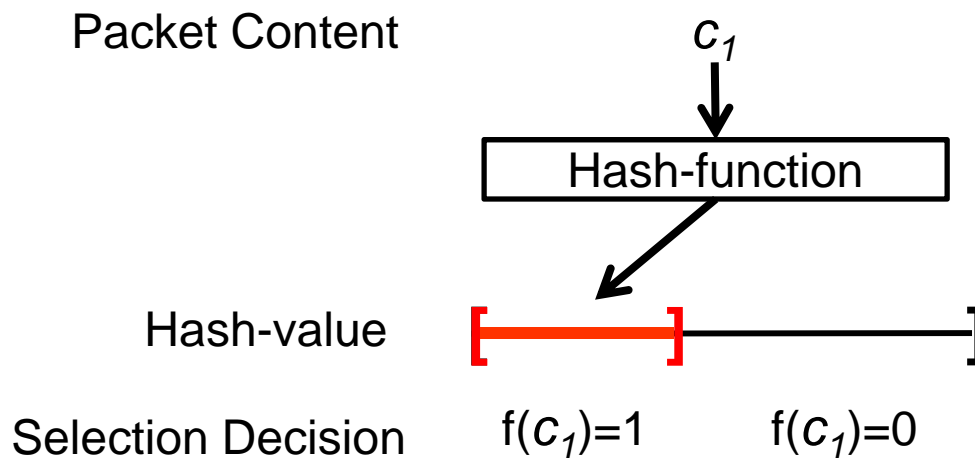
Sampling:  $f(s_i)$  or  $f(t_i)$  → ☹️

Filtering:  $f(c_i)$  → 😊

*Packet Attributes at Observation Point*  
 $s_i$  – sequence number  
 $t_i$  – arrival time  
 $c_i$  – content (header, payload)

# Coordinated Packet Selection

Trajectory Sampling [Duffield&Grossglauser, 2001]



**Goal: Emulate random selection**

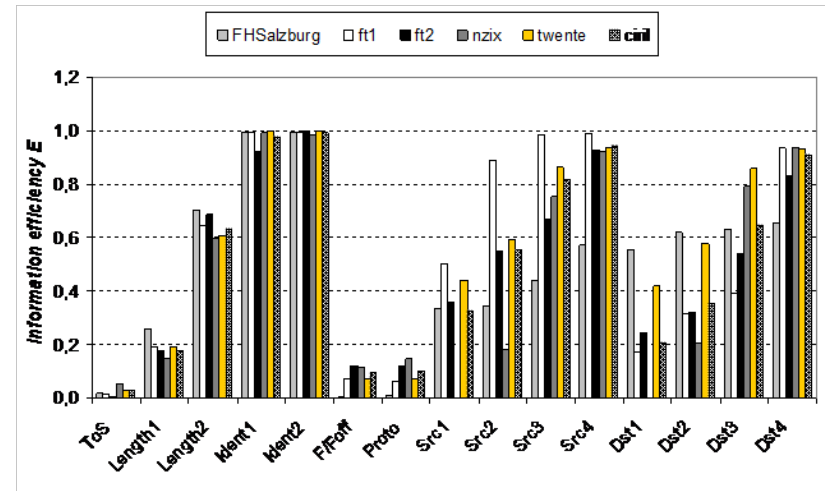


# Coordinated Packet Selection

- Problem1:** Content Selection
- Problem2:** Detection Avoidance
- Problem3:** Random Emulation
- Problem4:** Sample size variation

## FOKUS Contribution

- Empirical investigations on suitable content
  - invariant on the path, variable among packets
  - Initial investigations on IPv6
- Empirical investigations on bias
  - different traces and functions
- Standardization of suitable hash functions



# IETF Standardization

---

- IP Flow Information Export (IPFIX) → RFC5101
  - Export of flow information from routers and probes
  - Flexible flow definition
- Packet Sampling (PSAMP)
  - Export of packet information over IPFIX
  - Specification of packet selection techniques
  - Configuration of packet selection techniques
- FOKUS contribution
  - IPFIX requirements (RFC3917), implementation guidelines, applicability statement, file format, testing, extensions
  - PSAMP selection techniques
  - FOKUS Open Source IPFIX library

# Conclusion

---

- Traffic observation is essential
  - For network research and operation
  - For future Internet concepts
- Some solutions, but still many challenges
  - Example: Resource consumption
  - Example: Multipoint data selection
- Ongoing activities towards future Internet
  - ANA: Functional Composition
  - 4WARD: In-Network management based on situation awareness
  - PRISM: privacy-preserving methods to achieve situation awareness
  - Onelab2: support for experimental research



# Thank you!

*[tanja.zseby@fokus.fraunhofer.de](mailto:tanja.zseby@fokus.fraunhofer.de)*

FOKUS Open Source IPFIX Library:

*<http://net.fokus.fraunhofer.de/libipfix/>*

Measurement data always welcome at:

*<http://www.ist-mome.org/>*



Fraunhofer

Institute for Open  
Communication Systems