endace

Designing a 100% Flow generator for high-speed networks from OC3 to 100GbE

Stuart Wilson

Spencer Greene

Chapters

Design discussion

Design results

Future & Conclusions

Design discussion

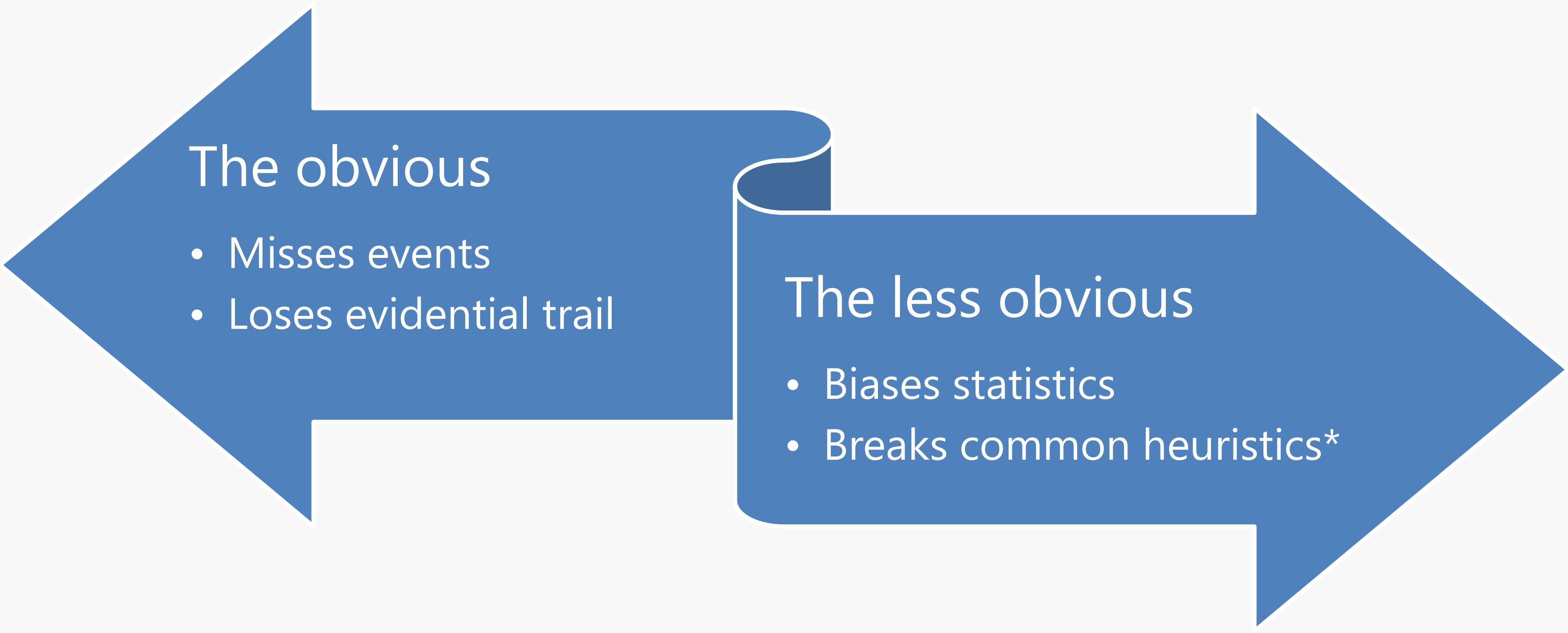
Borrowing from the über-talk…

	High value		**Packet**
	Low value	**Flow**	
		Low storage (1%-2%)	High storage (100%)

What about 10G/40G/100G?

High value			**Packet**
Medium value		**Flow**	
Lowest value	**Sampled Flow**		
	Lowest storage Low CPU	Low storage High CPU	High storage

Issues with sampling



The obvious

- Misses events
- Loses evidential trail

The less obvious

- Biases statistics
- Breaks common heuristics*

** Source: Mai et al, “Is Sampled Data Sufficient for Anomaly Detection?,” SIGCOMM ‘06*

Design objectives

Zero packet loss up to 100GbE @ small packets

Unsampled and sampled options at packet and NetFlow levels

Flexibility for additional metadata extraction

High Density / Low Space Weight & Power

Minimum number of unique SKUs

Small form factor for network-edge deployment

Monitor links via optical splitters (avoid in-line & span ports)

Parameters on a real network

Per interface	OC3	OC12	OC48	OC192	OC768	1GE	10GE	40GE	100GE
BW Gbps @ 75% load	0.10	0.40	1.6	7.2	29.	0.70	7.0	28.	70.
Mpps @ 100Byte	0.5	1.5	6.5	25	100	3.5	35	140	350
Active flows @ 30sec	35K	150K	600K	2.5M	10M	250K	2.5M	10M	25M
Flows/sec	1.5K	6K	25K	100K	400K	10K	100K	400K	1M

Architectural fundamentals

Router/Switch	**Server/appliance**
Range of high performance I/O	Typically poor I/O capability
Touches production traffic	Isolated from production traffic
Dedicated HW	Flexible software approach
Low CPU & Memory capacity	High CPU & Memory capacity
High SWaP➔Central	Low SWaP➔Distributed

=> I/O is key to unlocking high performance NetFlow !

I/O design

PCI-E (II) capable of 25Gbps per slot
⇒ Max 2 ports x 10Gbps per card
⇒ Front-end required for 40G/100G

1RU server includes 2 slots
⇒ Max 4 ports x 10Gbps per RU

New FPGA silicon enables “universal” receiver 155Mbps-10Gbps
⇒ 10GE/1GE and OC192/48/12/3

Appliance design basics

Intel class devices offer huge CPU & Memory performance

Combined with dedicated high speed front end

OC3 through 10GE

Single SKU 1U Server

Multiple cores load balanced

Fulfills performance objectives

1U server gives 2xPCI-E(II)

4 multi-rate interfaces in 1U

OC768, 40GE, 100GE

Saturates PCI-E(II)

=> ! Server design

Use dedicated head unit

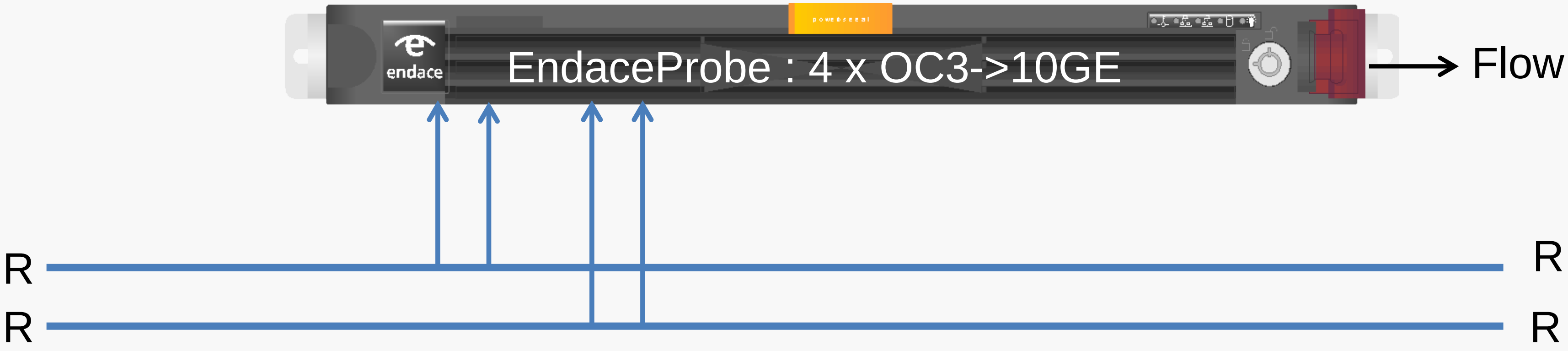
2 x 40GE / 1 x 100GE

Design results

4 x 10Gbps system

2 x bi-directional links, 10GE down to OC3

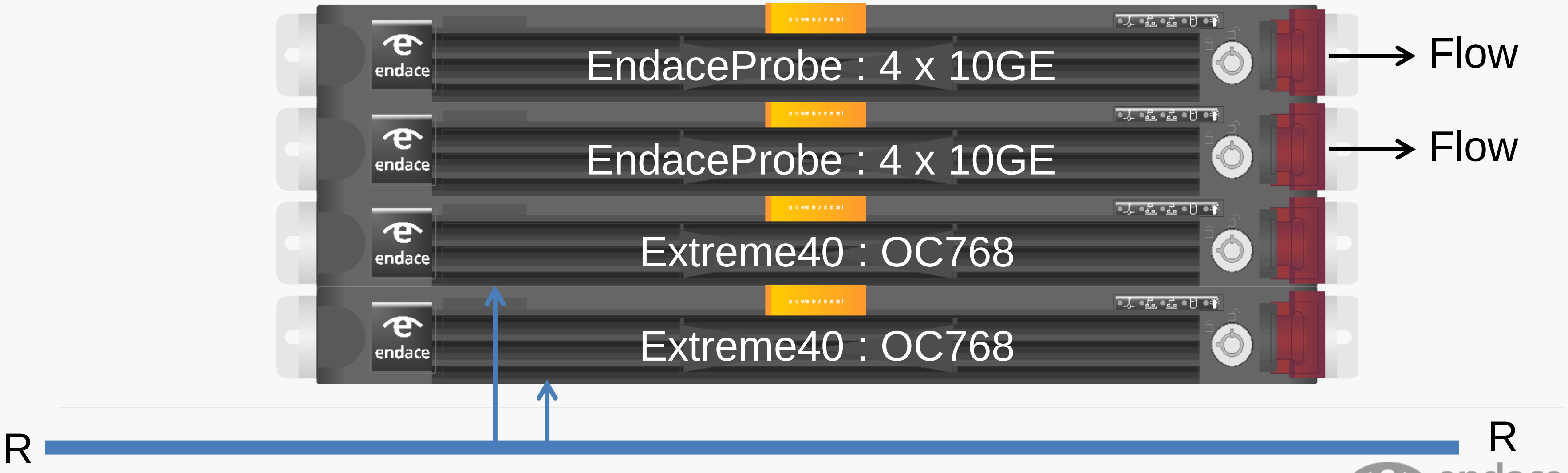
Per system	Gbps @ 100B	Mpps	Flows	Flows/Sec	Rack
Performance	30	100	10M	400K	1U (300W)



2 x OC768 system

1 x bi-directional OC768 link

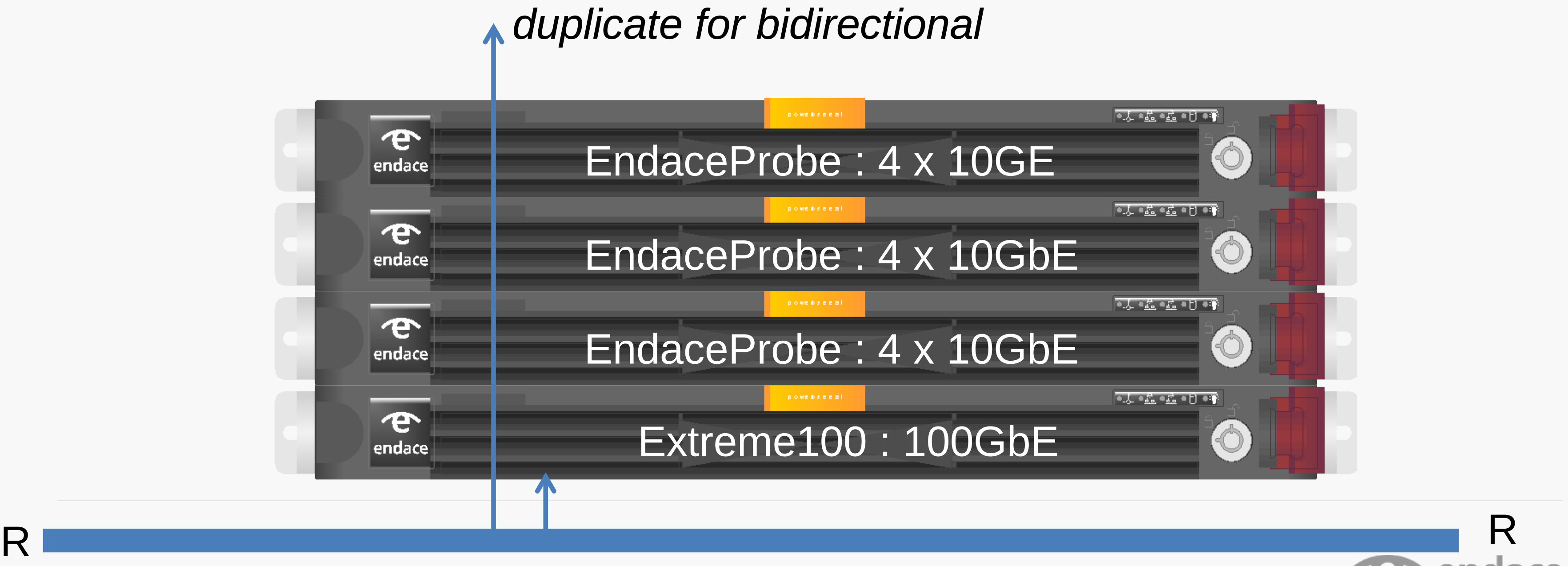
Per system	Gbps @ 100B	Mpps	Flows	Flows/Sec	Rack
Performance	60	200	20M	800K	4U (1kW)



1 x 100GE system

1 x uni-directional 100GE link

Per system	Gbps @ 100B	Mpps	Flows	Flows/Sec	Rack
Performance	90	300	30M	1.2M	4U (1.3kW)



Future & Conclusions

Uncharted territory

Channelized SONET/SDH

Eliminate rack(s) of SONET gear

Application awareness

DPI generated application type added to IPFIX

Identity – beyond IP address

NAT binding?

IMSI, IMEI, M-ISDN ?

Server/Software approach allows flexible derivatives

Conclusions

Big Iron is not required for production NetFlow generation

Server based designs with proper I/O

Modern FPGAs can do an awful lot

↓

Very high performance

Good flexibility

High I/O density – low space/weight/power endace

power to see all