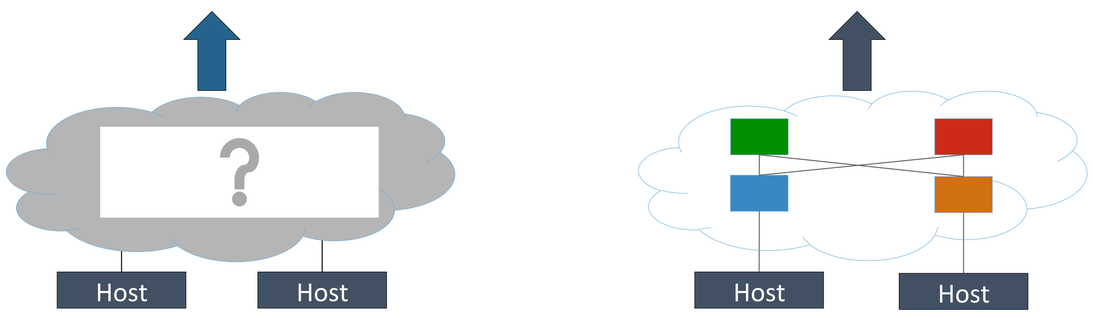
INT: In-band Network Telemetry in P4

Embedding device-internal metadata directly into the data plane

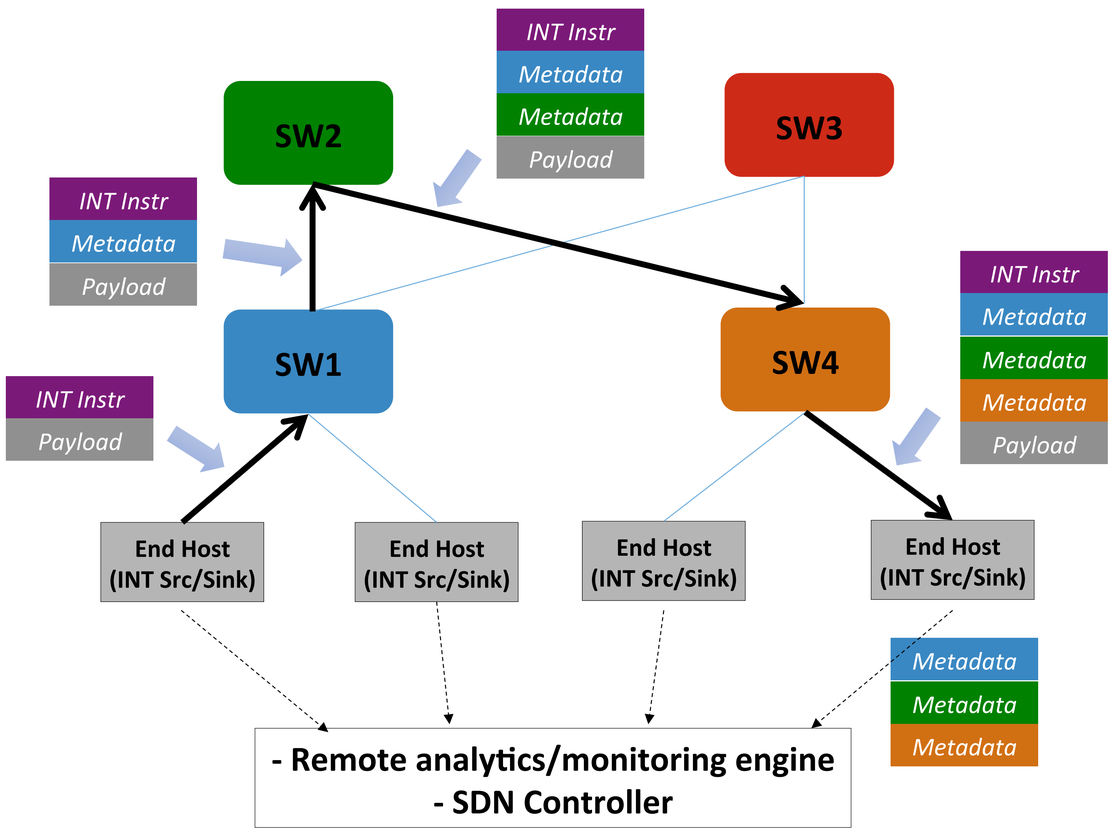
VMware and Barefoot Networks

(1) Why INT?

Traditional networks		INT-enabled networks
- Low visibility into network state - Polling-based - Aggregated counters - Limited by control-plane speed	vs.	- Complete network visibility - Real-time monitoring - Per-packet metadata - Full line rate, zero switch CPU involvement



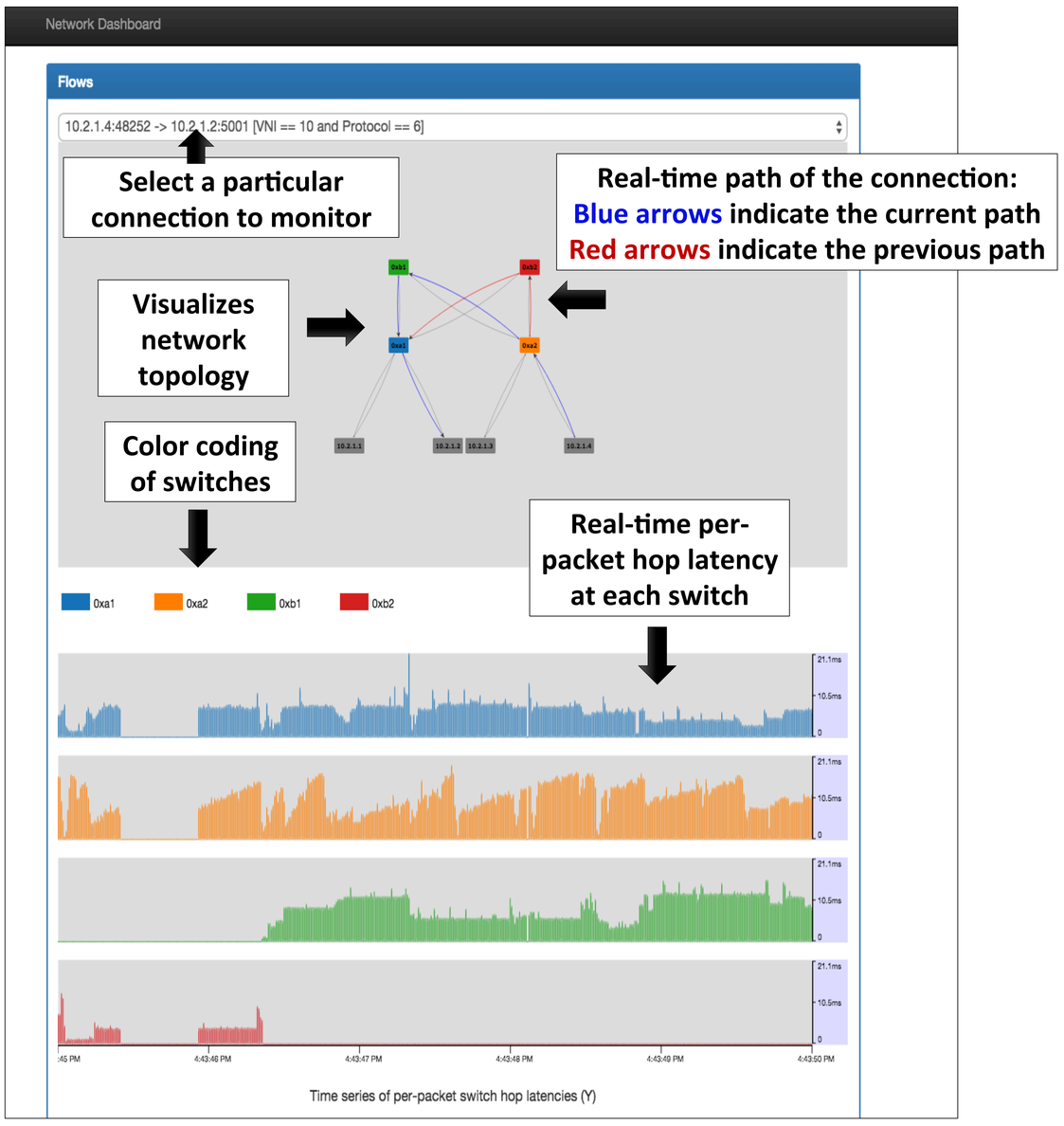
(2) How does INT work?



(4) Example: Real-time path and latency monitoring

- Real-time path information for each connection in the network

- Real-time per-packet hop latency at each switch



(5) Open-source INT implementation in P4

INT Header Definition

```
                                                        /* headers.p4 */
header_type int_header_t {
  fields {
    ver                     : 2;
    rep                     : 2;
    c                       : 1;
    e                       : 1;
    rsvd1                   : 5;
    ins_cnt                 : 5;
    max_hop_cnt             : 8;
    total_hop_cnt           : 8;
    instruction_mask_0003   : 4; /* split the bits for lookup */
    instruction_mask_0407   : 4;
    instruction_mask_0811   : 4;
    instruction_mask_1215   : 4;

header_type int_switch_id_header_t {
  fields {
    bos         : 1;
    switch_id   : 31;
```

INT Header Parsing

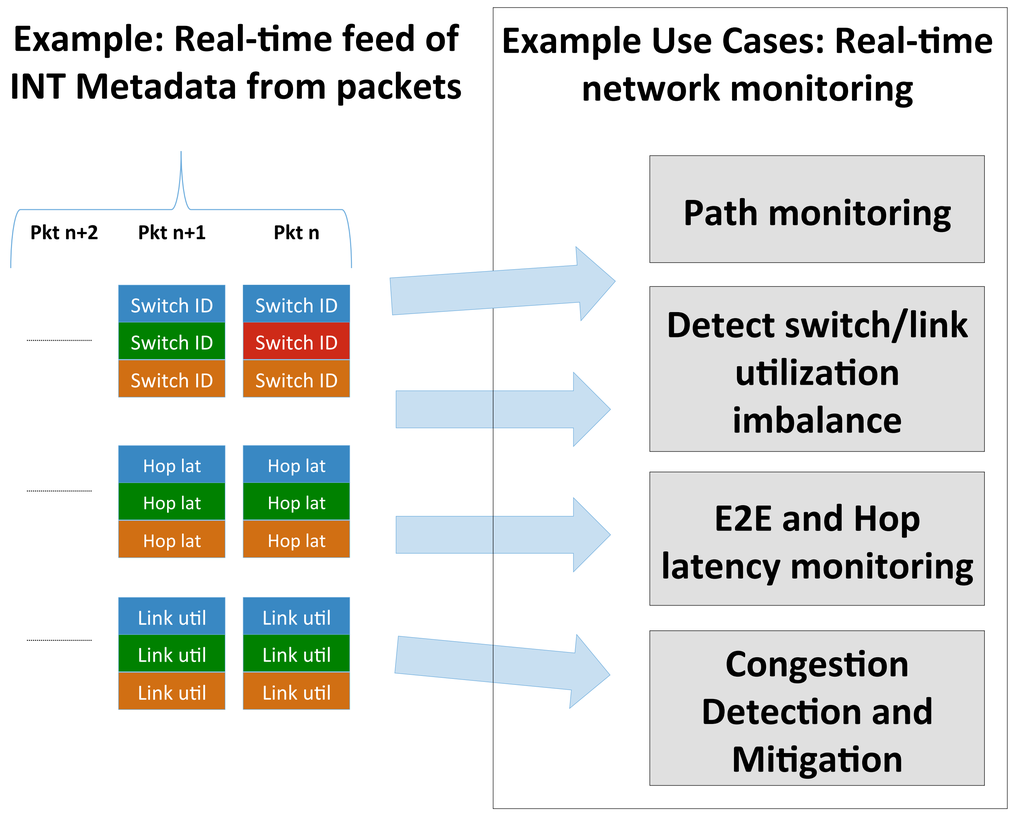
```
                                                        /* parser.p4 */
parser parse_int_header {
    extract(int_header);
    set_metadata(int_metadata.instruction_cnt, latest.ins_cnt)
    return select (latest.rsvd1, latest.total_hop_cnt) {
       0x000: ingress;
```

Control flow to append INT metadata as per the instructions in the INT header

```
                                                        /* int_transit.p4 */
control process_int_insertion {
  apply(int_insert) {
    int_transit {
      if (int_metadata.insert_cnt != 0) {
        apply(int_inst_0003);
        apply(int_inst_0407);
        apply(int_inst_0811);
        apply(int_inst_1215);
        apply(int_bos);
      }

      /* update E-bit or total_hop_cnt in the INT header */
      apply(int_meta_header_update);
```

(3) Example types of INT metadata



- Switch ID
- Hop latency
- Ingress port id
- Ingress timestamp
- Ingress port RX pkt count
- Ingress port RX drop count
- Ingress port RX utilization
- Queue id
- Instantaneous queue length
- Egress port id
- Egress timestamp
- Egress port TX pkt count
- Egress port TX drop count
- Egress port TX utilization

(6) Full Specification and Code - Available

Spec, Video & Blog:
http://p4.org/p4/inband-network-telemetry/

Code:
https://github.com/p4lang/p4factory/tree/master/apps/int