Precision Instrumentation of High Performance Trading Systems

Yong C. Shin, CEO

Ambrotos, Inc.

ycshin@ambrotos.co.kr

TS-Associates® Precision Instrumentation

AM·BROTOS

Performance

Precision

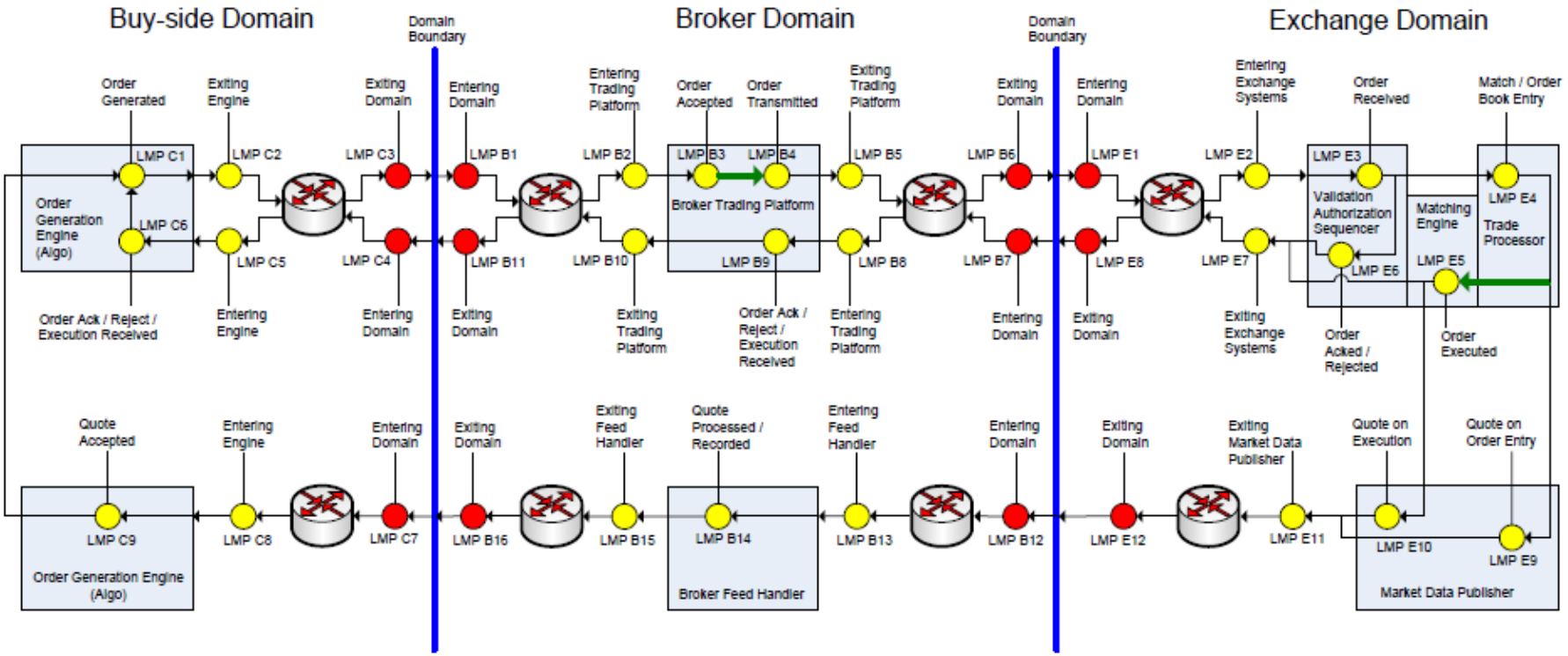
Parallelism

How did we get from here …

… to here …


REAR FAN 2
REAR FAN 1
intel
S/N
BQKT3120008B AA G74721-362
DIMM 3
DIMM 1
DIMM 4
DIMM 2
CPU

… end up here ?


Buy-side Domain
Domain Boundary
Broker Domain
Domain Boundary
Exchange Domain
Order Generated
Exiting Engine
Exiting Domain
Entering Domain
Entering Trading Platform
Order Accepted
Order Transmitted
Exiting Trading Platform
Exiting Domain
Entering Domain
Entering Exchange Systems
Order Received
Match / Order Book Entry
LMP C1
LMP C2
LMP C3
LMP B1
LMP B2
LMP B3
LMP B4
LMP B5
LMP B6
LMP E1
LMP E2
LMP E3
LMP E4
Order Generation Engine (Algo)
LMP C6
Broker Trading Platform
Validation Authorization Sequencer
Matching Engine
Trade Processor
LMP C5
LMP C4
LMP B11
LMP B10
LMP B9
LMP B8
LMP B7
LMP E8
LMP E7
LMP E6
LMP E5
Order Ack / Reject / Execution Received
Entering Engine
Entering Domain
Exiting Domain
Exiting Trading Platform
Order Ack / Reject / Execution Received
Entering Trading Platform
Entering Domain
Exiting Domain
Exiting Exchange Systems
Order Acked / Rejected
Order Executed
Quote Accepted
Entering Engine
Entering Domain
Exiting Domain
Exiting Feed Handler
Quote Processed / Recorded
Entering Feed Handler
Entering Domain
Exiting Domain
Exiting Market Data Publisher
Quote on Execution
Quote on Order Entry
LMP C9
LMP C8
LMP C7
LMP B16
LMP B15
LMP B14
LMP B13
LMP B12
LMP E12
LMP E11
LMP E10
LMP E9
Order Generation Engine (Algo)
Broker Feed Handler
Market Data Publisher

Help - We're out of our depth

Billions of transistors

22 **nanometer** lithography

Gigabits of network bandwidth

Nanosecond precise timing

Our system of units is based at human scale

How can we make these things meaningful ?

What is a nanosecond ?

One billionth of a second : 1 ns = 10^{-9} s

Speed of light (in a vacuum) = 299,792,458 m/s

Light travels 30 cm in 1 ns (in a vacuum)

Copper transmission = light speed x NVP

CAT5e/6 NVP = 67%

1 ns = 20 cm

Fibre transmission = light speed / refractive index

850 nm multimode RI = 1.538

1 ns = 19.5 cm

What about a FIX message ?

Example New Order Single:

```
8=FIX.4.2|9=130|35=D|34=659|49=BROKER04|56=REUTERS|52=20070123-
19:09:43|38=1000|59=1|100=N|40=1|11=ORD10001|60=20070123-
19:01:17|55=HPQ|54=1|21=2|10=004|
```

Length = 153 Bytes = 1224 Bits

10GE Serialisation = 122.4 ns

Message Length = 24.5m

FIX has Performance Challenges

Verbose "tag = value" syntax

Variable length fields

ASCII wire encoding (text -> binary price conversion)

Encode/decode is CPU intensive

XML = 10 x worse

-> The "ASCII Backlash" -> Binary Protcols

Binary Order Entry Protocols

- Tag-less
- Fixed offset (for mandatory fields)
- Prices already in binary encoding
- Omission of redundant information
- Encode/decode is much less CPU intensive
- Messages are shorter -> less bandwidth / higher rates
- New Order in BATS BOE : Length = 41 bytes = 6.5m
- FIX Workgroup working towards higher performance

Principles of Refactoring

Refactoring : Restructuring a system, altering its internal structure without changing its external behaviour, undertaken to improve some of the non-functional attributes of the system.

Improving:

- Architecture
- Maintainability
- Performance

Principles of Refactoring

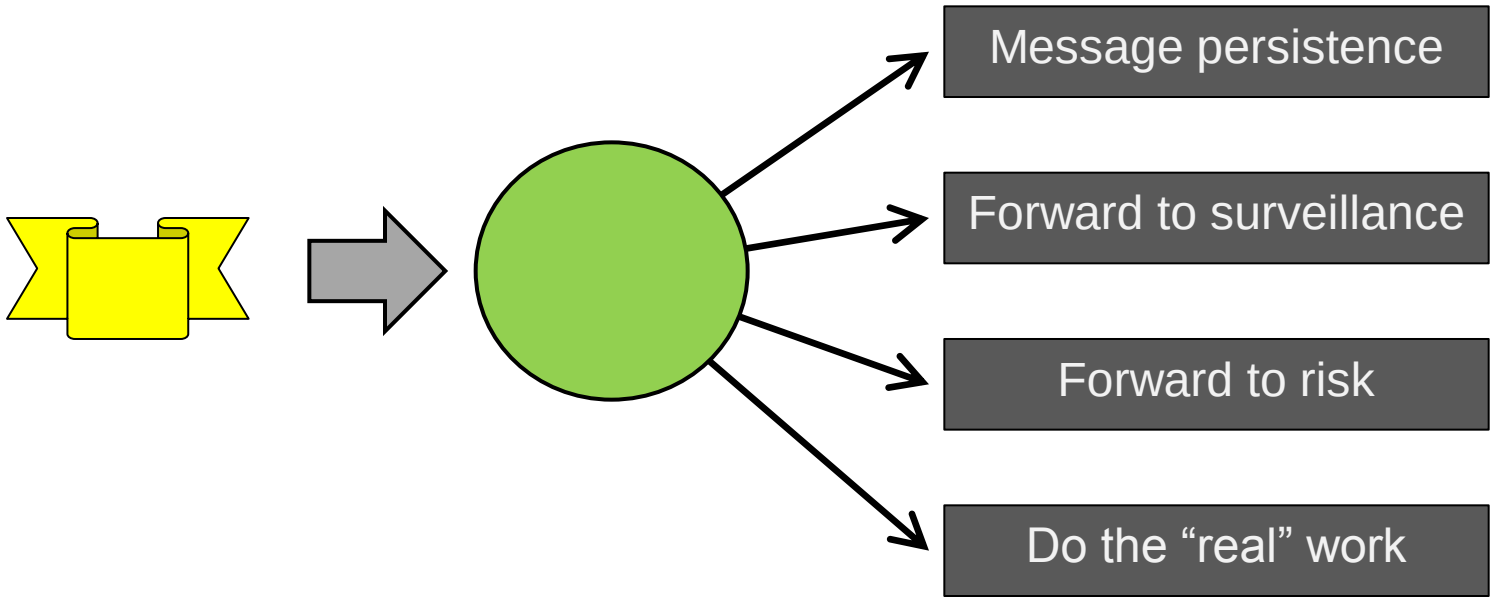
Refactoring : Restructuring a system, altering its internal structure without changing its external behaviour, undertaken to improve some of the non-functional attributes of the system.

"결과의 변경 없이 코드의 구조를 재조정함"

Improving:

- Architecture
- Maintainability
- Performance

Typical Message Processing Fan Out



Software fan out = Bottleneck

Creating Flow Parallelism

Layer 7 techniques are inefficient (software)

Layer 2/3 techniques are non-invasive

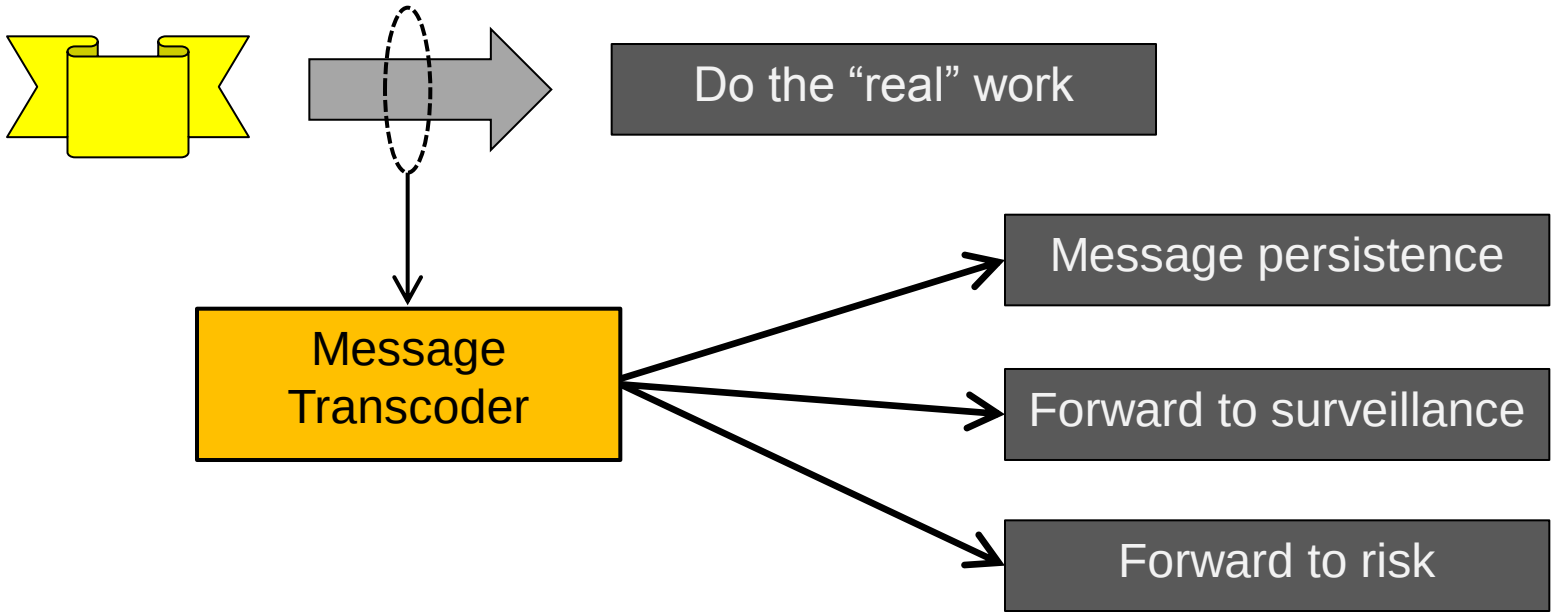
Layer 2

- Network Tap

Layer 3

- Switch SPAN / Mirror port

Refactored Message Processing Fan Out



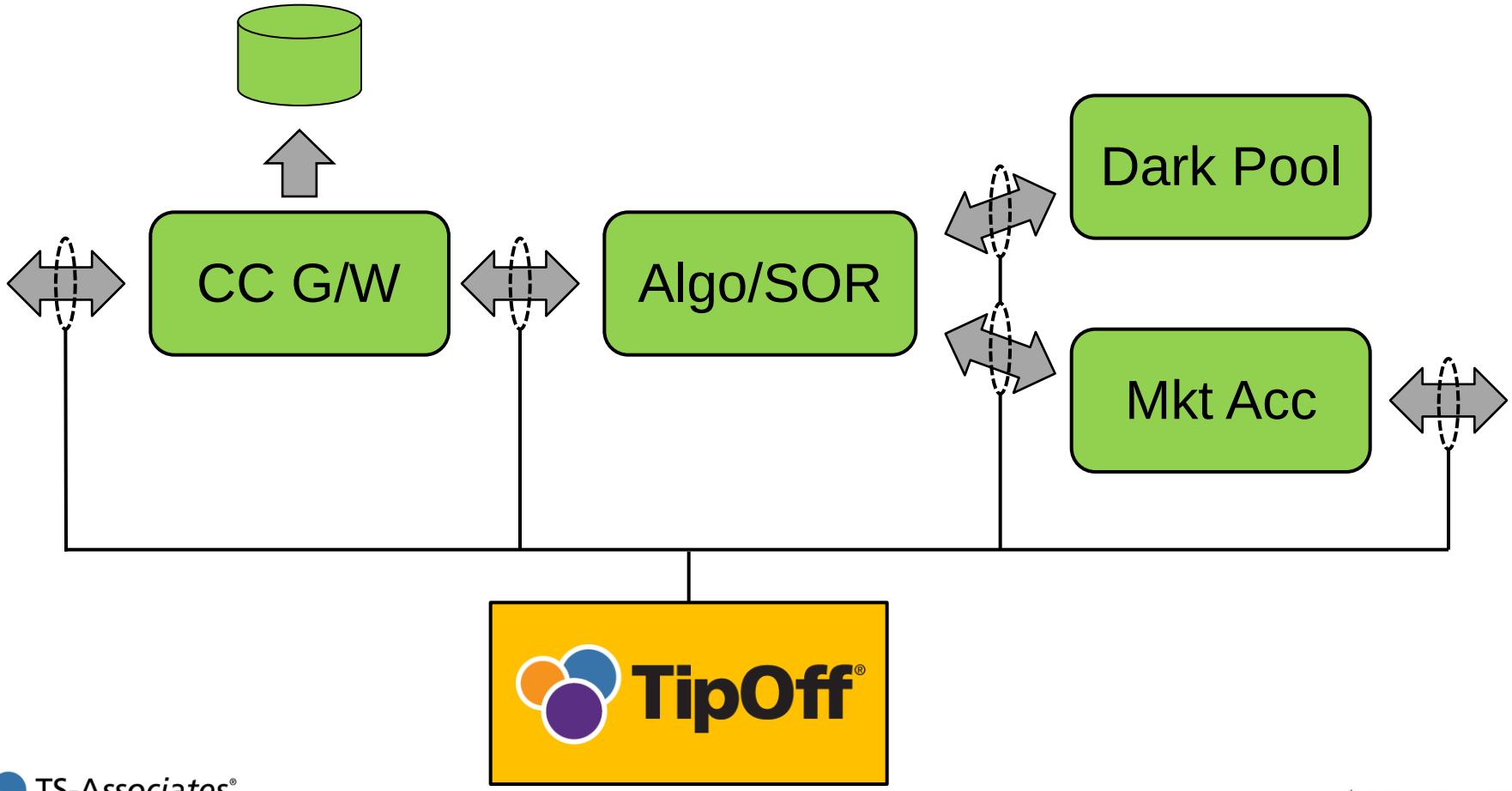
Message Transcoder

- Packet capture
- Decode – transport, session, message payload
- Content normalisation (optional)
- Message encode (optional)
- Message persistence / transmission

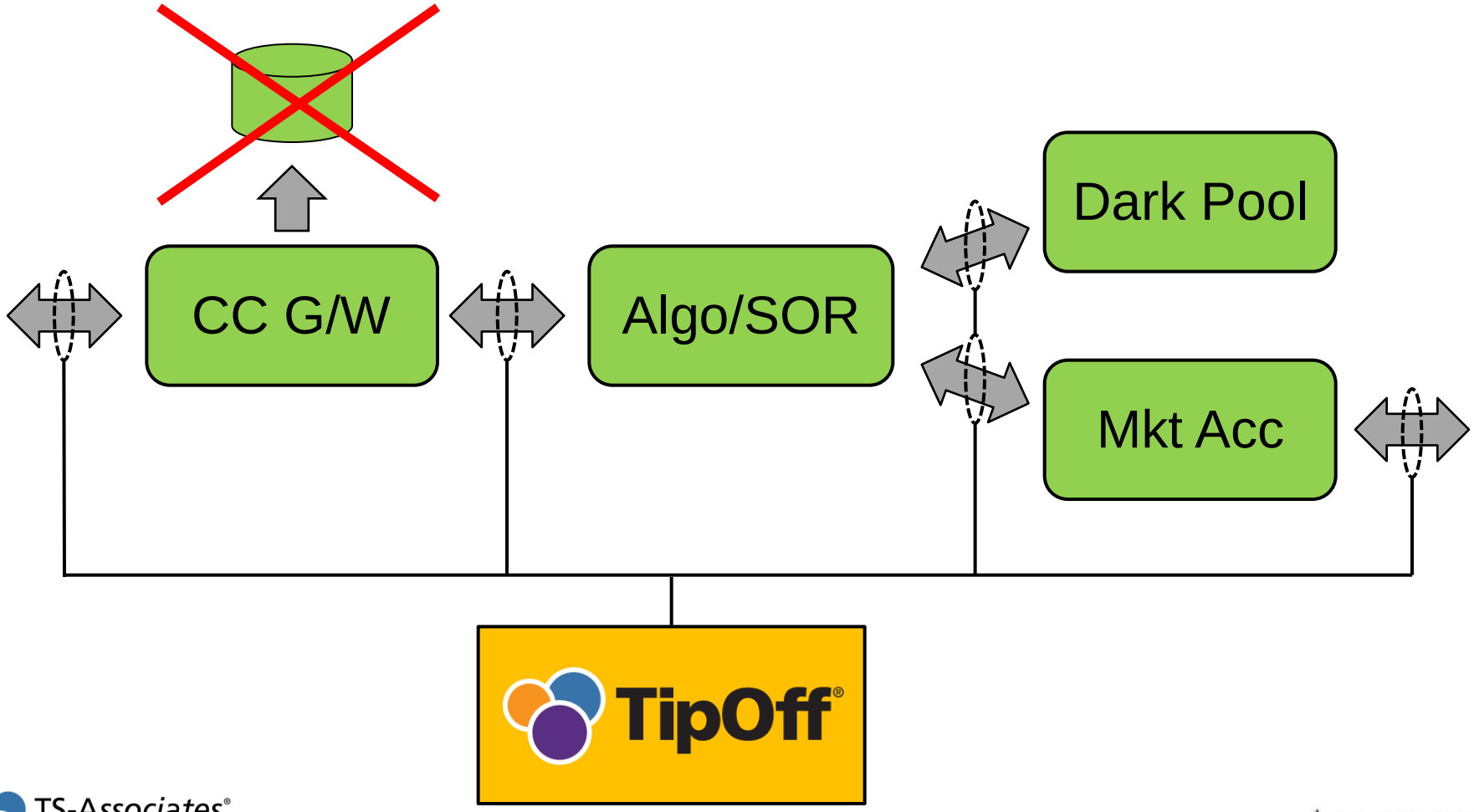
Example of a Message Transcoder:

TS-Associates TipOff® appliance

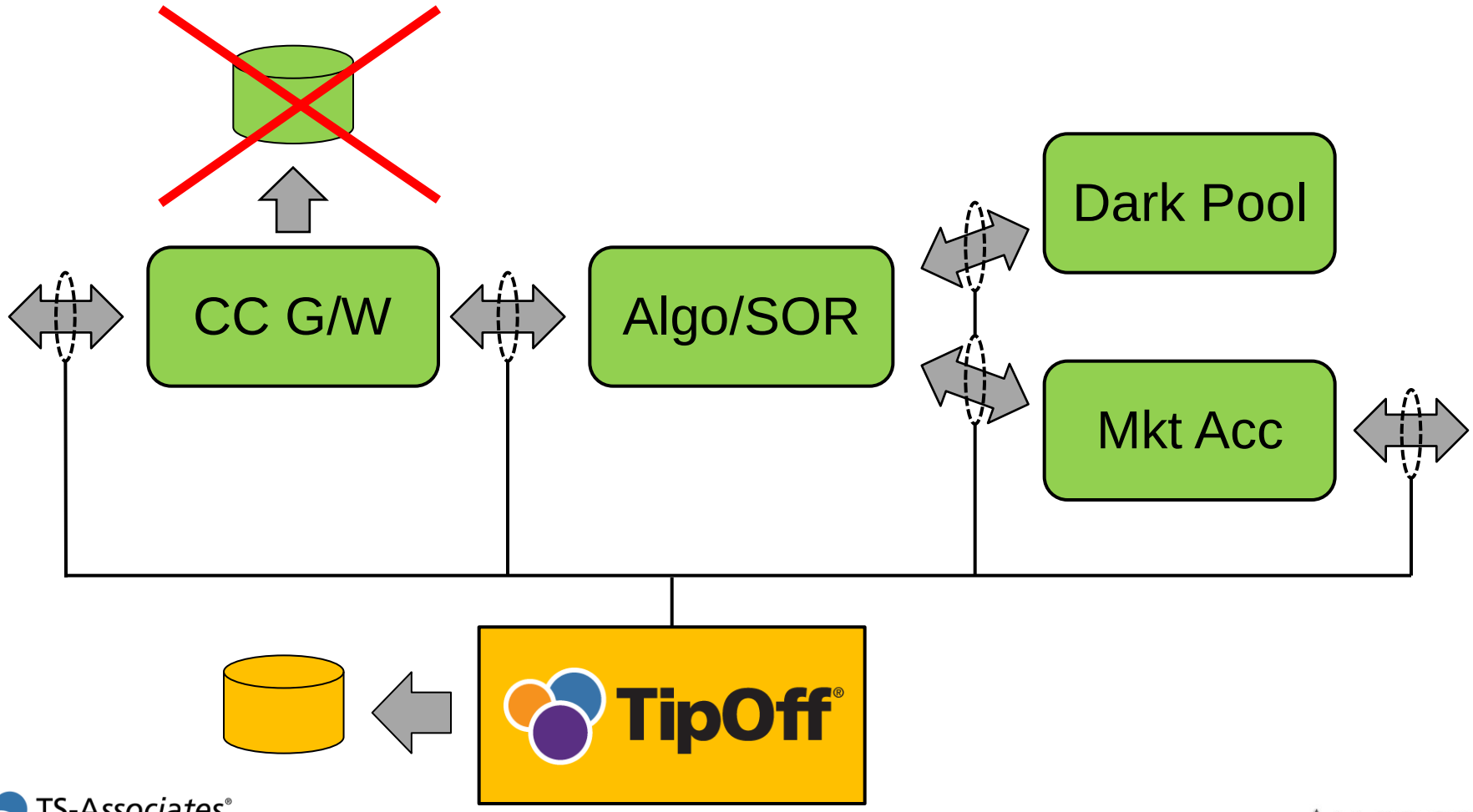
Case Study – Equity Brokerage



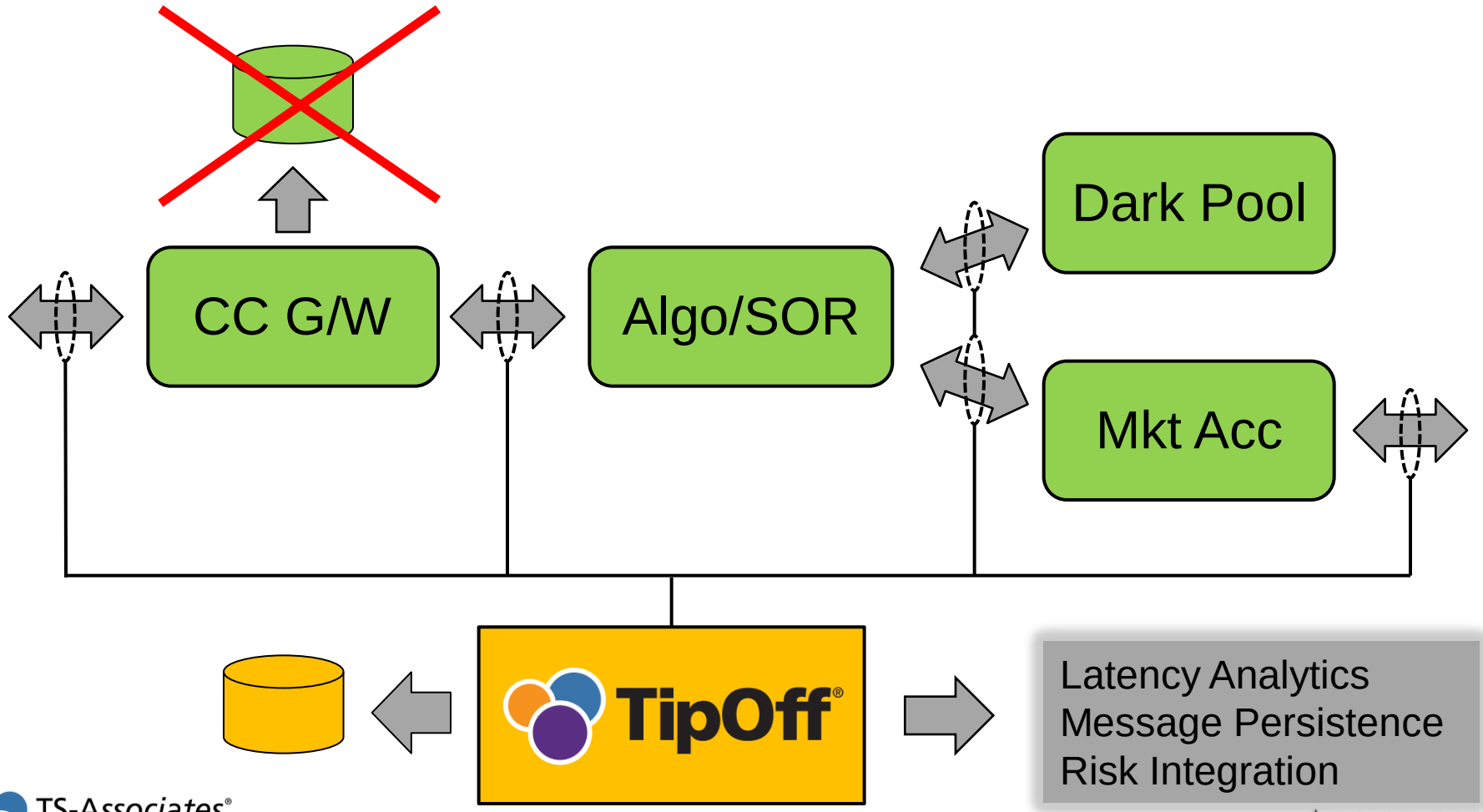
Case Study – Equity Brokerage



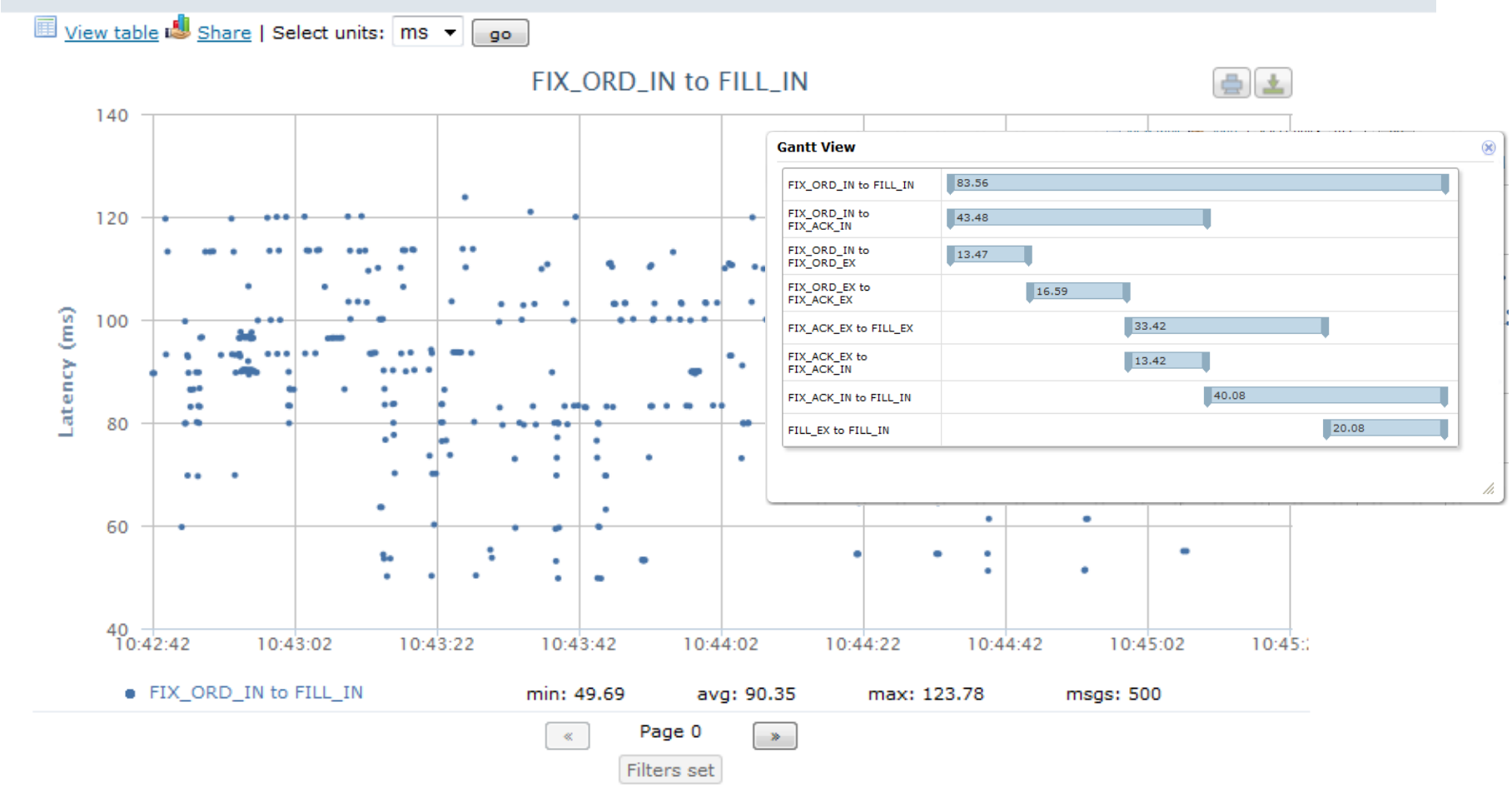
Case Study – Equity Brokerage



Case Study – Equity Brokerage



Example Screen Shot - TipOff



Precision Instrumentation

Standard instrumentation techniques no longer adequate:

- Commodity servers use low stability clocks – XO
- Managing time in software is inaccurate and invasive
- NTP time sync accuracy in range 1ms – 200us

Precision Instrumentation = real world time + ns accuracy

- High stability clocks – TCXO, OCXO, Rb, CSAC
- Time stamping in hardware (10ns resolution)
- Precision time sync – PPS or PTP (100ns accuracy)
- GPS time reference for inter-site instrumentation

Drilling Deeper – Software Component Latency

- ETS consolidation : distributed -> multi-core
- Traditional monitoring solutions lose visibility
- TS-Associates invented Application Tap
- Software instrumentation API
- Supported by deployment ecosystem
  - FPGA based Application Tap card
  - Integrated with Solarflare 10GE NICs
  - Software emulation, remote daemon, dev stub, etc

Thank you for listening

Yong C. Shin, Ambrotos, Inc.

www.ambrotos.co.kr
ycshin@ambrotos.co.kr

TS-Associates® Precision Instrumentation

AM·BROTOS