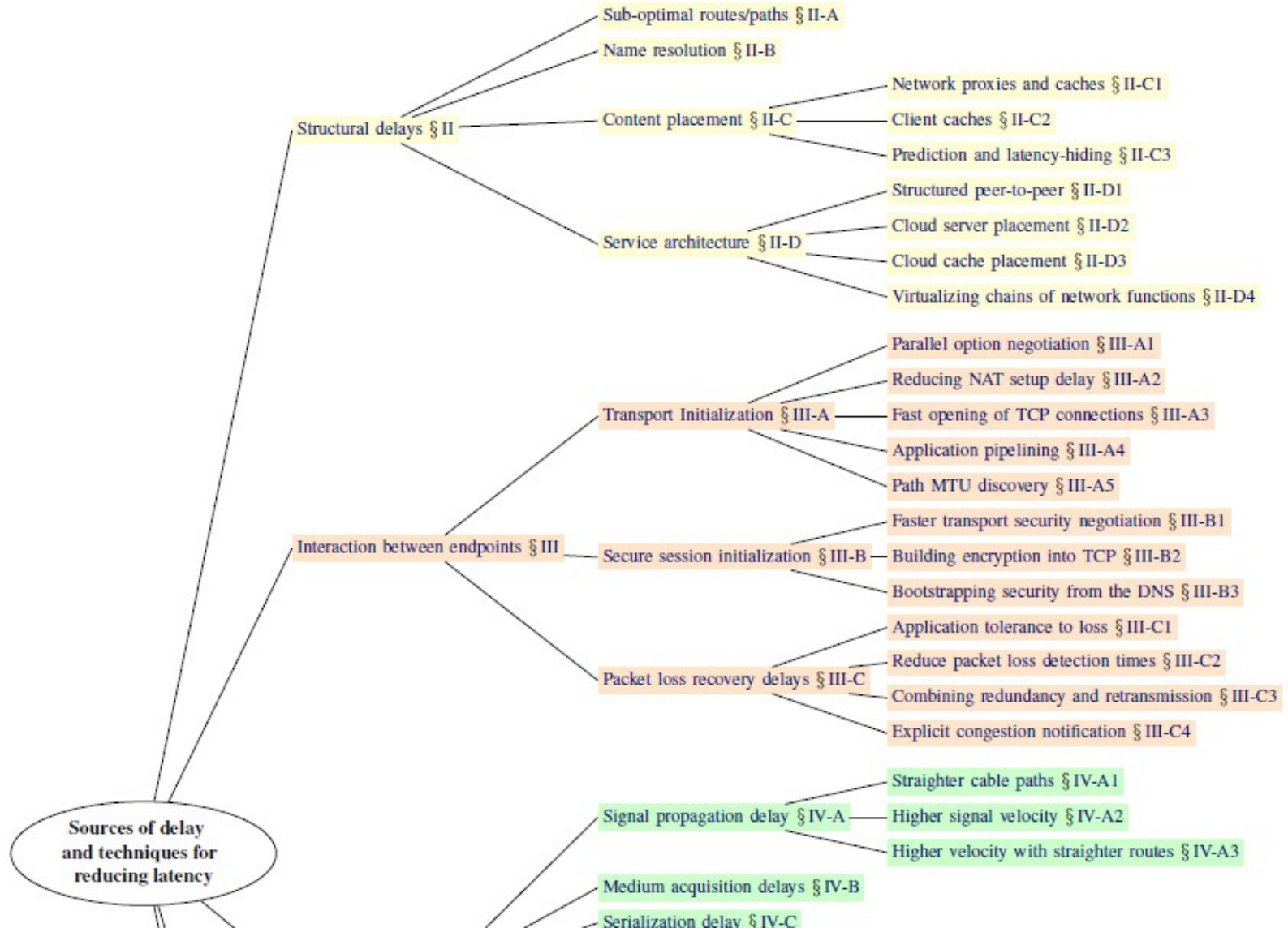# Reducing Internet Latency: a survey of techniques and their merits

Bob Briscoe, Anna Brunstrom, Andreas Petlund, David Hayes, David Ros, Ing-Jyh Tsang, Stein Gjessing, Gorry Fairhurst, Carsten Griwodz, Michael Welzl

summary

- industry roadmap of techniques
- gain vs pain
  - latency reduction against deployability

- "A Survey of Latency Reducing Techniques and their Merits"
  - 322 references
  - available via http://riteproject.eu/publications/

- evolved from BT roadmap work, but repurposed
  - a company tries to prioritise the quick wins
  - an industry also needs to identify hard problems being avoided
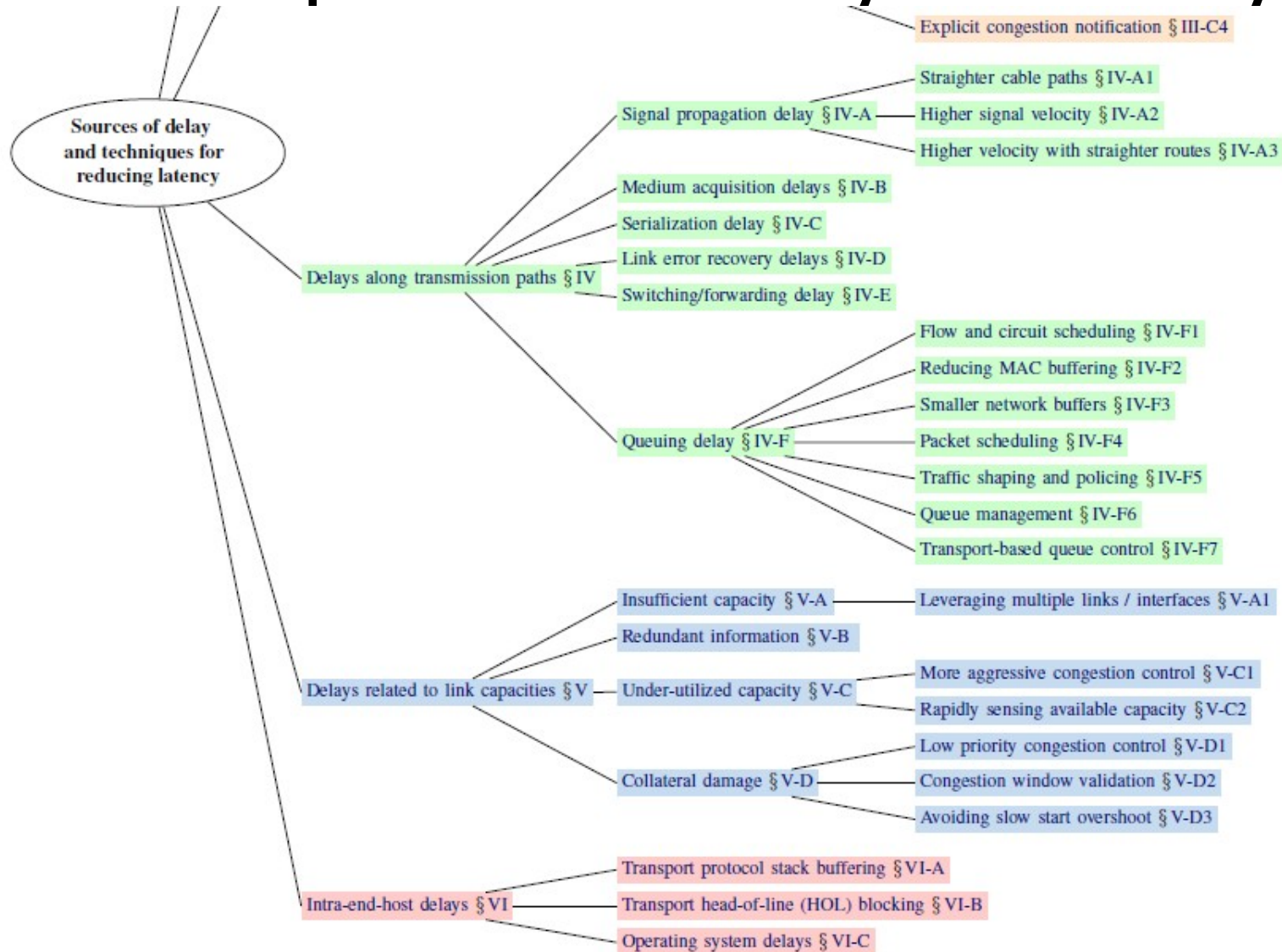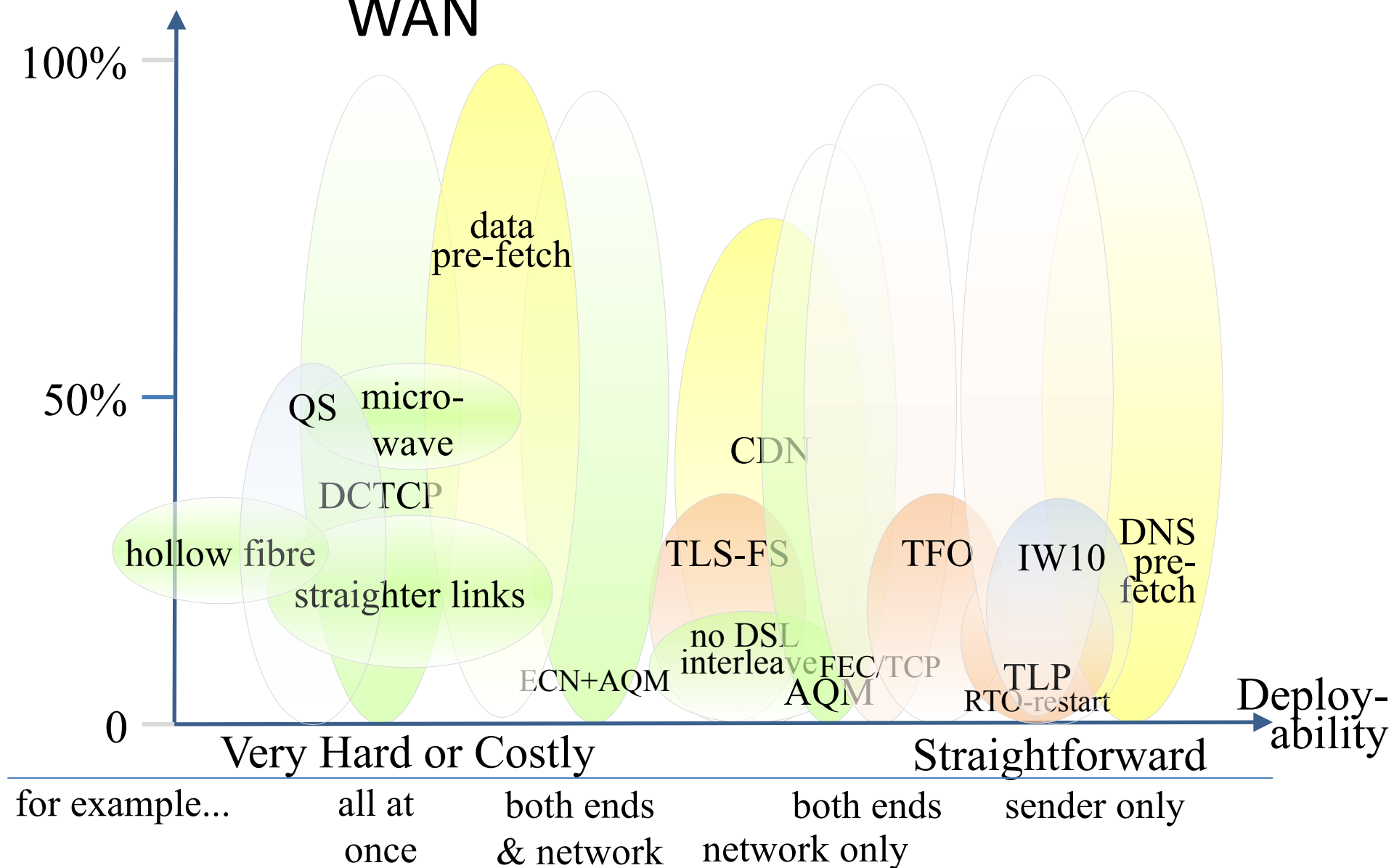
roadmap around body of survey I

Sub-optimal routes/paths §II-A

Name resolution §II-B

Structural delays §II

Content placement §II-C
- Network proxies and caches §II-C1
- Client caches §II-C2
- Prediction and latency-hiding §II-C3

Service architecture §II-D
- Structured peer-to-peer §II-D1
- Cloud server placement §II-D2
- Cloud cache placement §II-D3
- Virtualizing chains of network functions §II-D4

Interaction between endpoints §III

Transport Initialization §III-A
- Parallel option negotiation §III-A1
- Reducing NAT setup delay §III-A2
- Fast opening of TCP connections §III-A3
- Application pipelining §III-A4
- Path MTU discovery §III-A5

Secure session initialization §III-B
- Faster transport security negotiation §III-B1
- Building encryption into TCP §III-B2
- Bootstrapping security from the DNS §III-B3

Packet loss recovery delays §III-C
- Application tolerance to loss §III-C1
- Reduce packet loss detection times §III-C2
- Combining redundancy and retransmission §III-C3
- Explicit congestion notification §III-C4

Signal propagation delay §IV-A
- Straighter cable paths §IV-A1
- Higher signal velocity §IV-A2
- Higher velocity with straighter routes §IV-A3

Medium acquisition delays §IV-B

Serialization delay §IV-C

Sources of delay and techniques for reducing latency

roadmap around body of survey II

Fig. 1. Techniques for reducing latency organized by sources of delay.

case (1a): small (20kB) flow over WAN

reduction in completion time
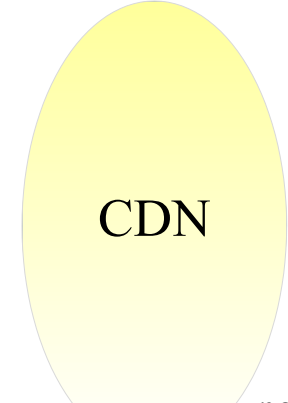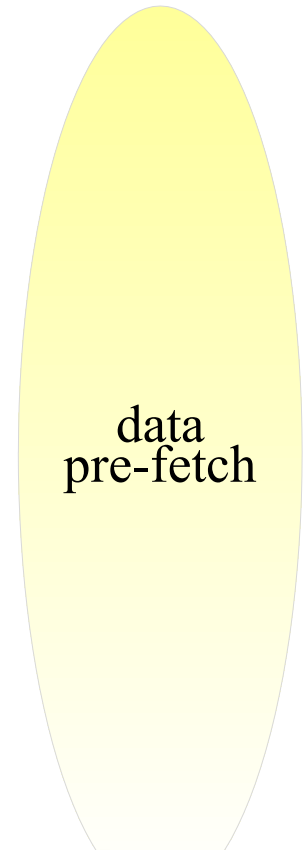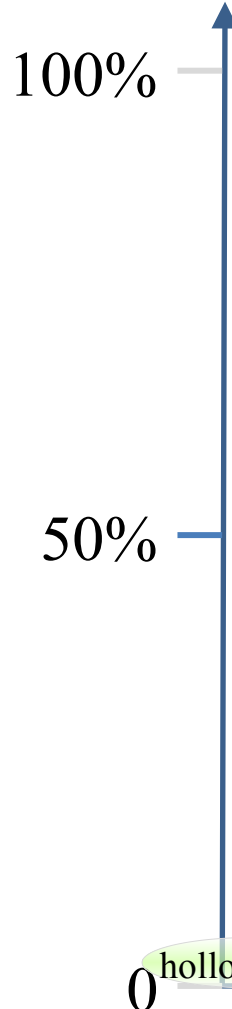
100%

50%

0

Deploy-ability

Very Hard or Costly · · · · · · · Straightforward

for example...   all at once   both ends & network   both ends network only   sender only

data pre-fetch

QS   micro-wave

DCTCP

hollow fibre

straighter links

CDN

TLS-FS   TFO   IW10   DNS pre-fetch

no DSL interleave   FEC/TCP

ECN+AQM   AQM   TLP   RTO-restart

reduction in completion time

case (2a): large flow over WAN

100%

data pre-fetch

CDN

50%

0

hollow fibre QS DCTCP microwave ECN TLS-FS no DSL interleave AQM TFO IW10 DNS pre-fetch

Deploy-ability

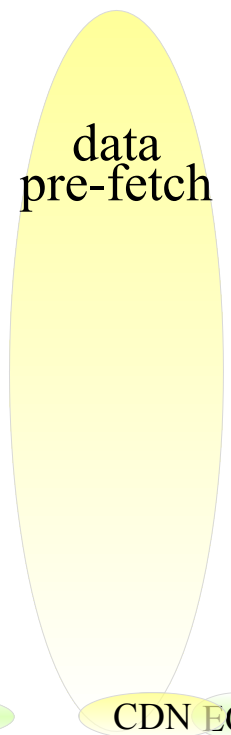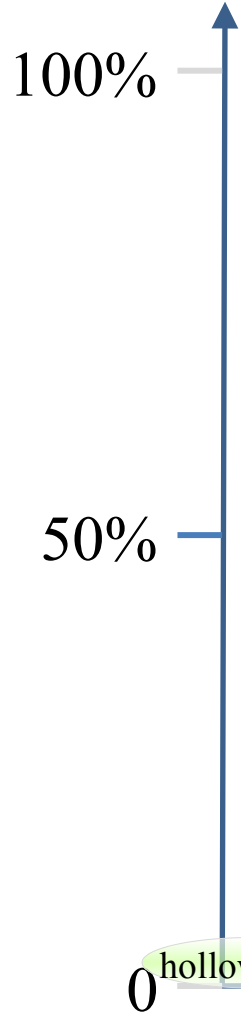Very Hard or Costly   Straightforward

for example...   all at once   both ends & network   both ends network only   sender only

case (2b): large flow over LAN

reduction in
completion
time

100%

50%

0

data
pre-fetch

hollow fibre QS DCTCP

CDN ECN

TLS-FS

AQM

TFO

IW10 DNS pre-fetch

Deploy-
ability

Very Hard or Costly

Straightforward

for example...

all at
once

both ends
& network

both ends
network only

sender only