The Analysis & Mining of Globally Distributed Data

Chapter 3.
Network & Data Protocols

Robert Grossman
Laboratory for Advanced Computing
University of Illinois at Chicago
&
Open Data Partners

Copyright 2003 Robert L. Grossman

3.1 Layers & Stacks

Stacks for web content,
internet applications and web services.

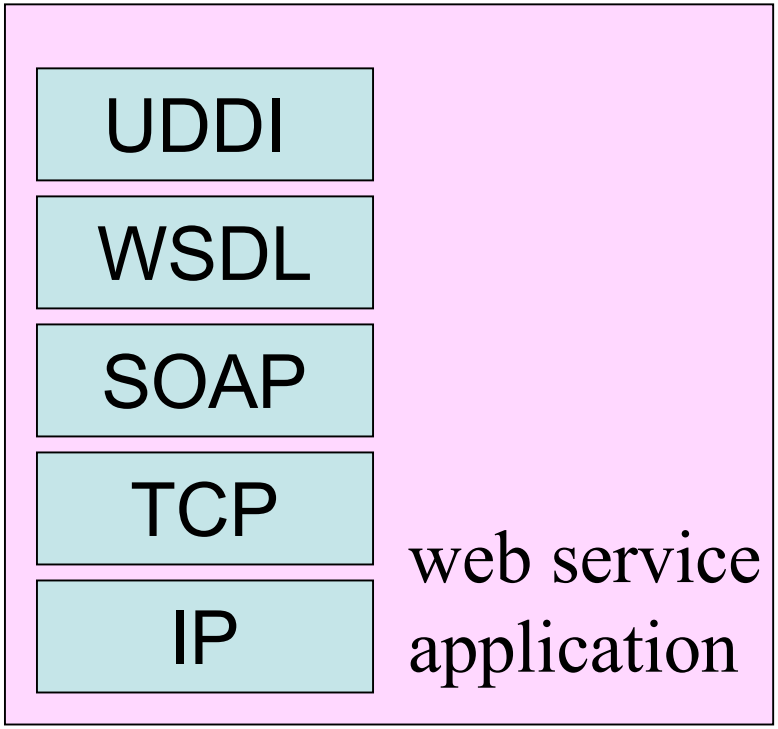
Copyright 2003 Robert L. Grossman

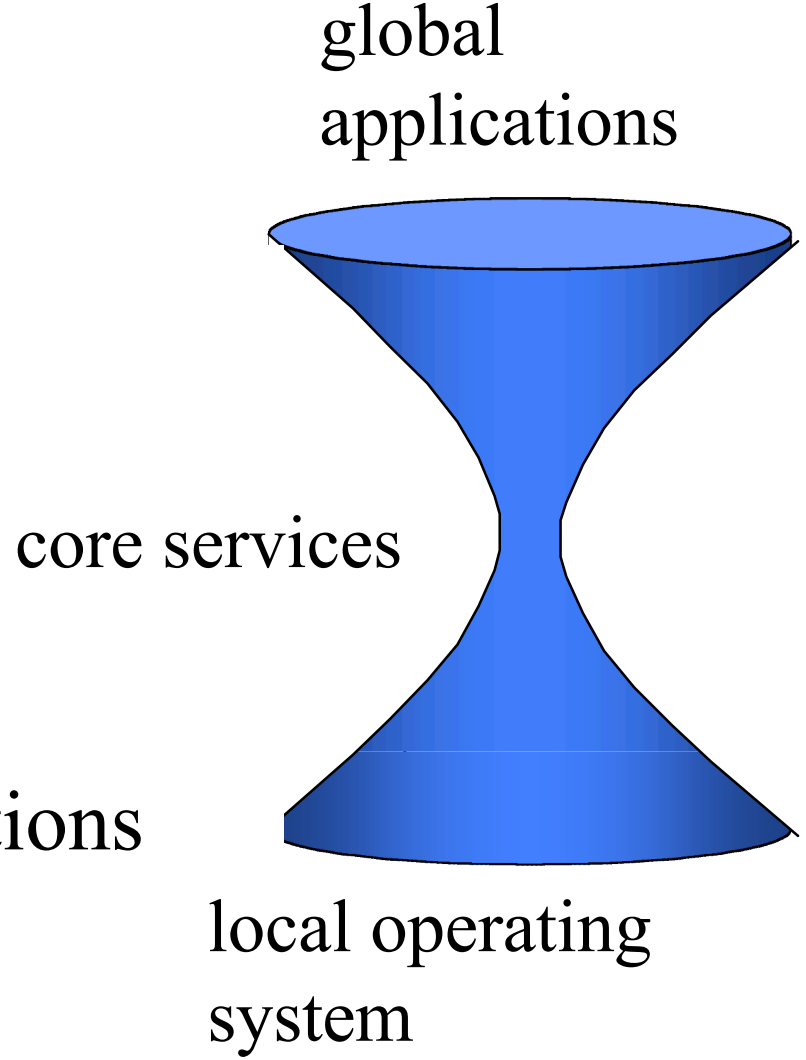
Four Standard Stacks

Stacks for data grids, data webs and distributed data mining.

Copyright 2003 Robert L. Grossman

Basic Idea - Layers

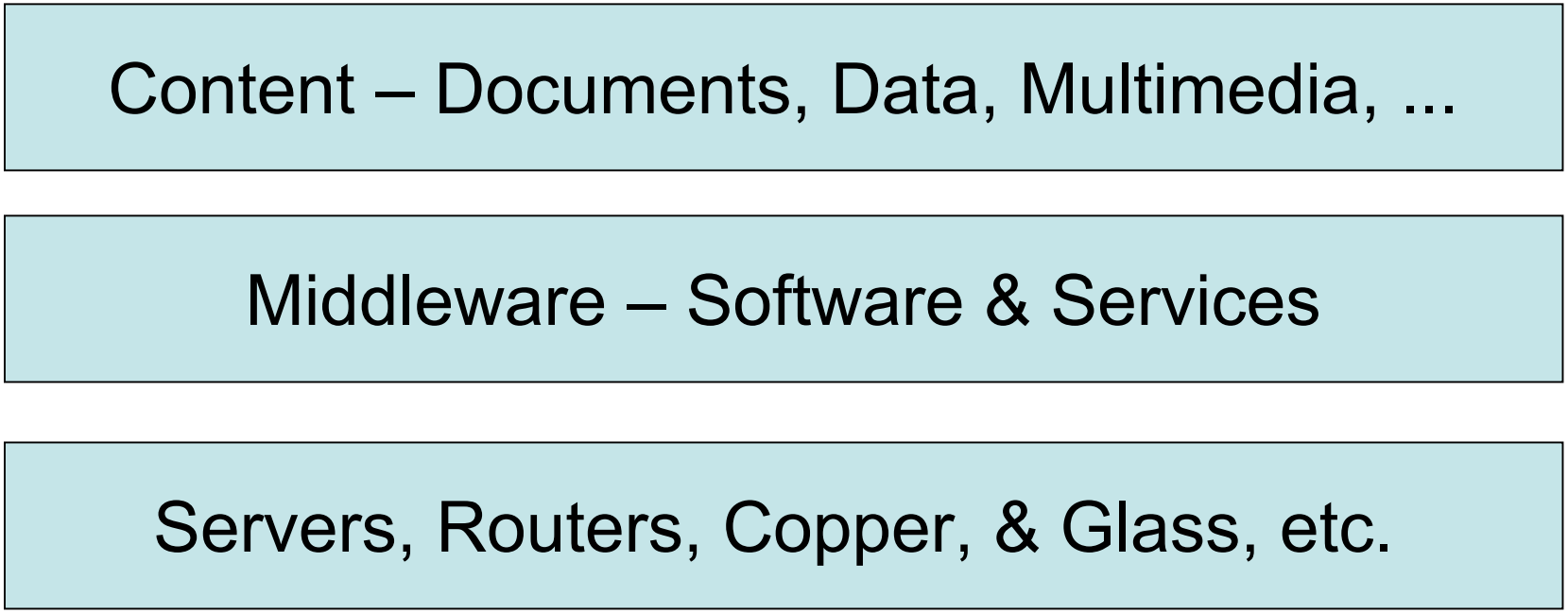




- diverse global applications
- few core services

Copyright 2003 Robert L. Grossman

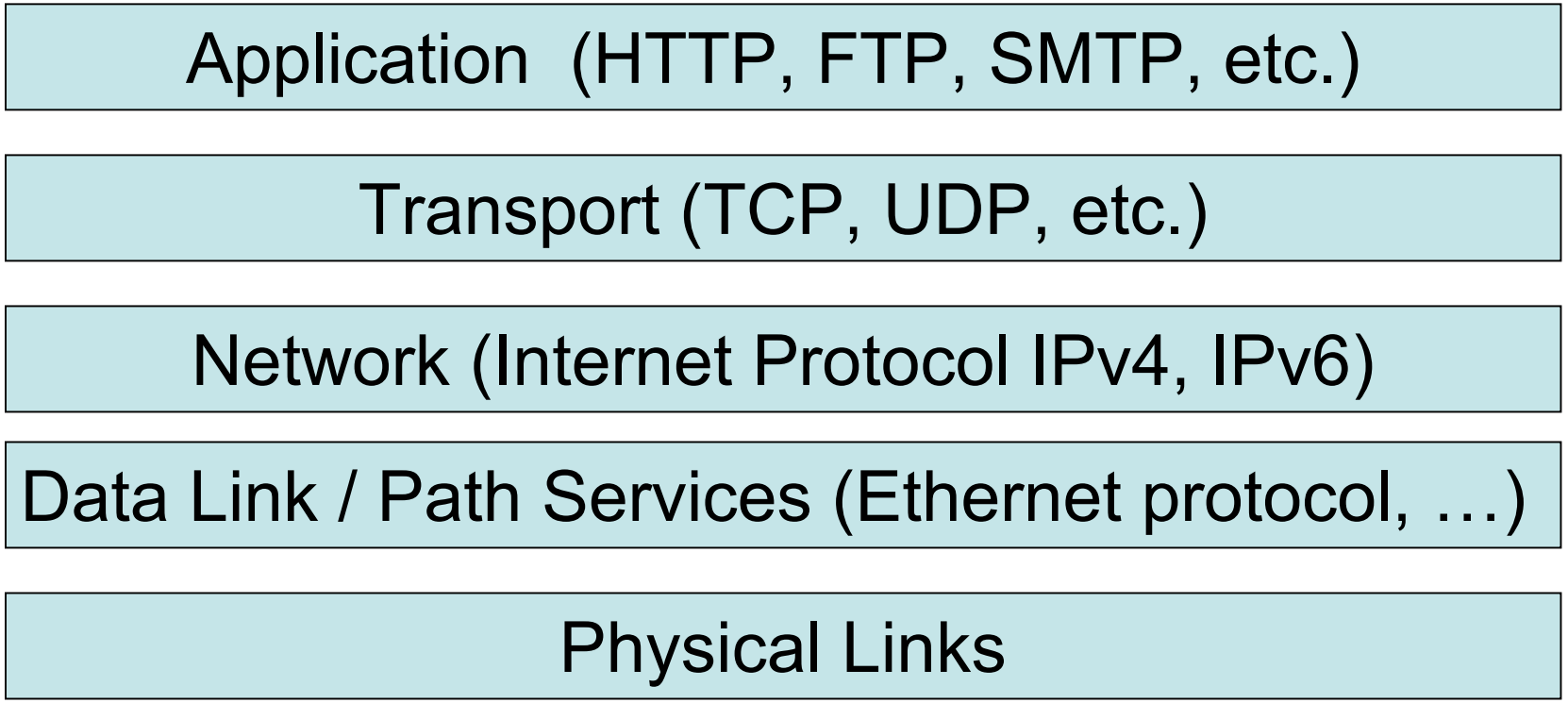
3 Layer Web Content Model



- External View – data grids, data webs, & semantic webs are the middle layer.
- Perspective of data content providers.

Copyright 2003 Robert L. Grossman

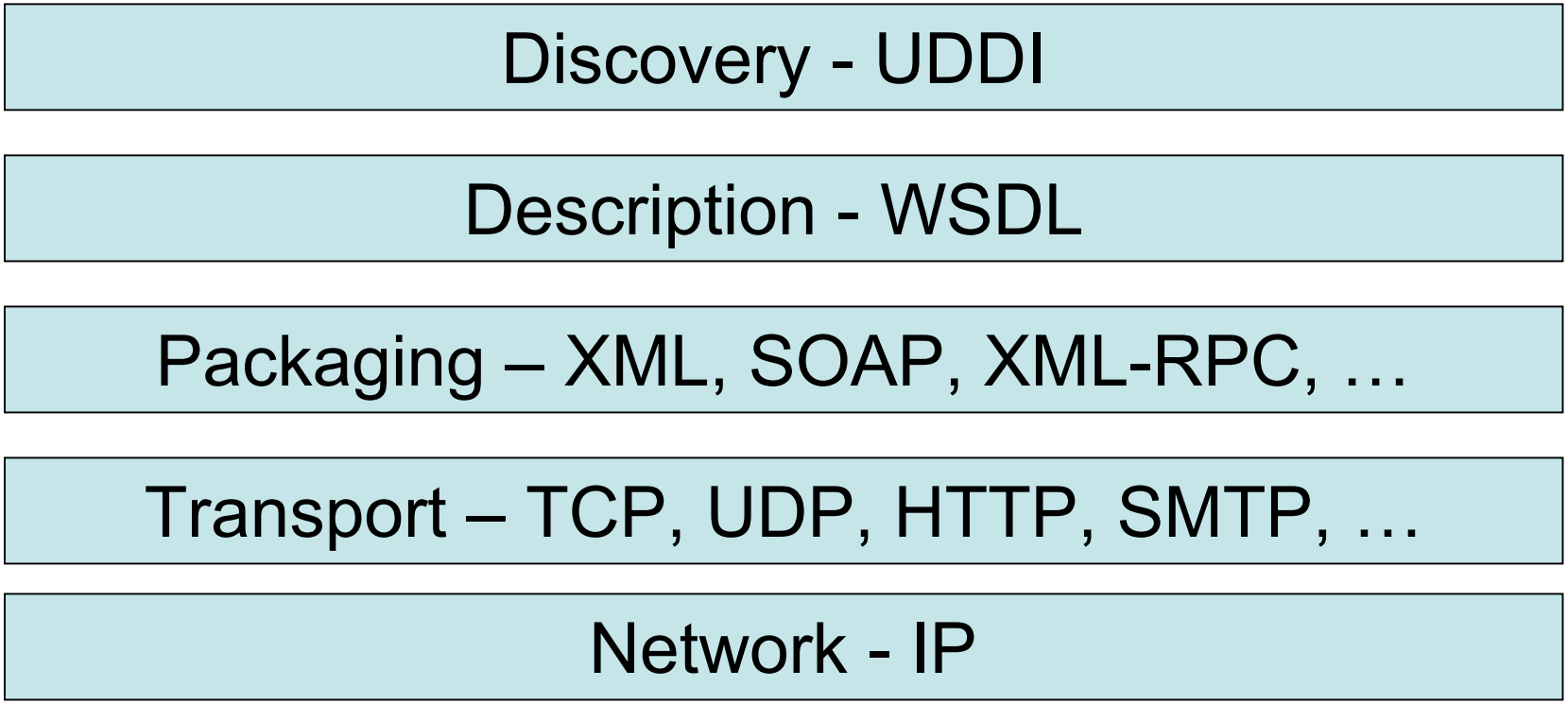
5 Layer Internet Model



- The action is below TCP: perspective is a wire stack.
- ISO 7 layer model a variant of this.

Copyright 2003 Robert L. Grossman

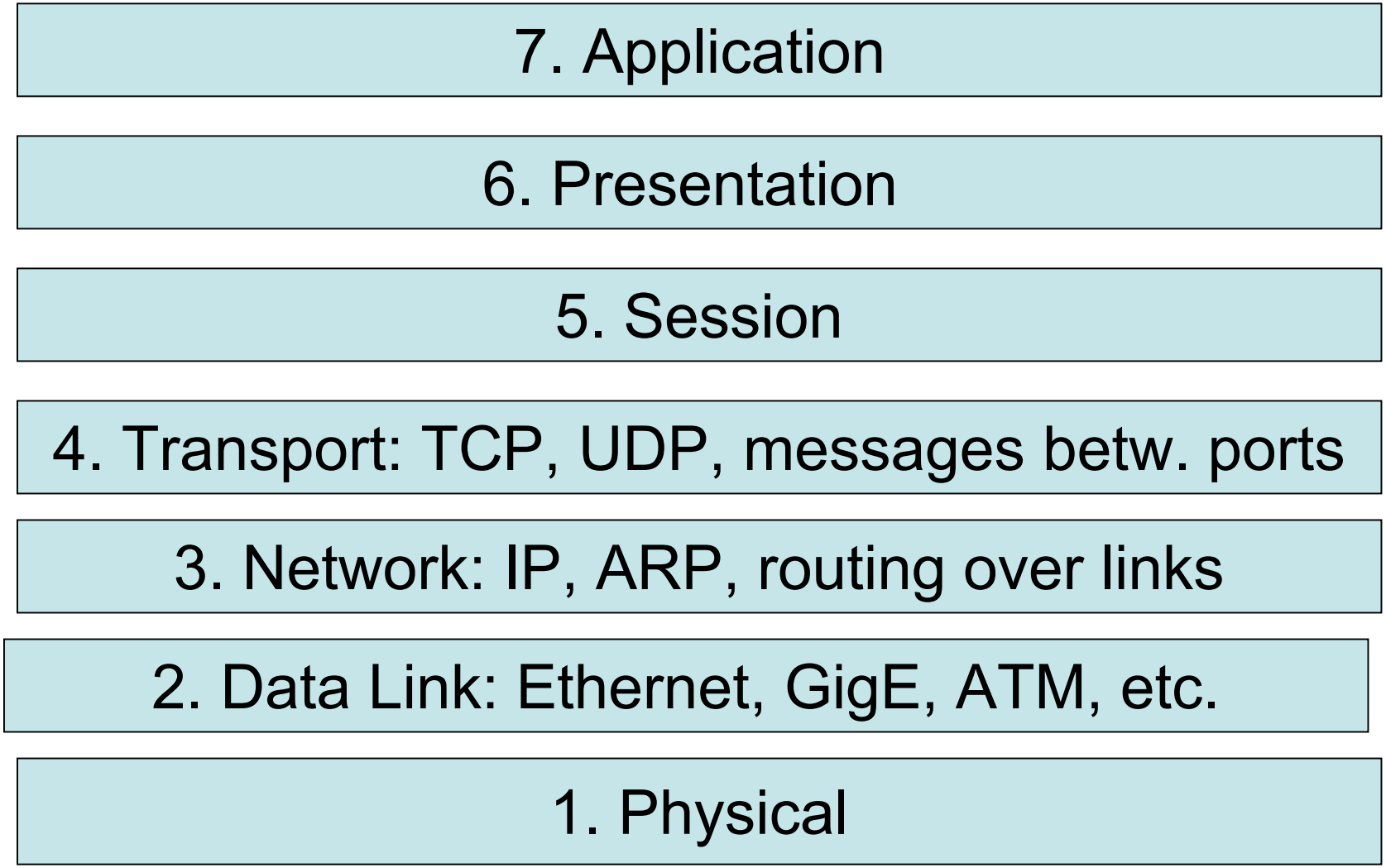
5-Layer Web Services Model



- The action above TCP: perspective discovery stack.
- Model for data webs over commodity networks.

Copyright 2003 Robert L. Grossman

7-Layer OSI Model



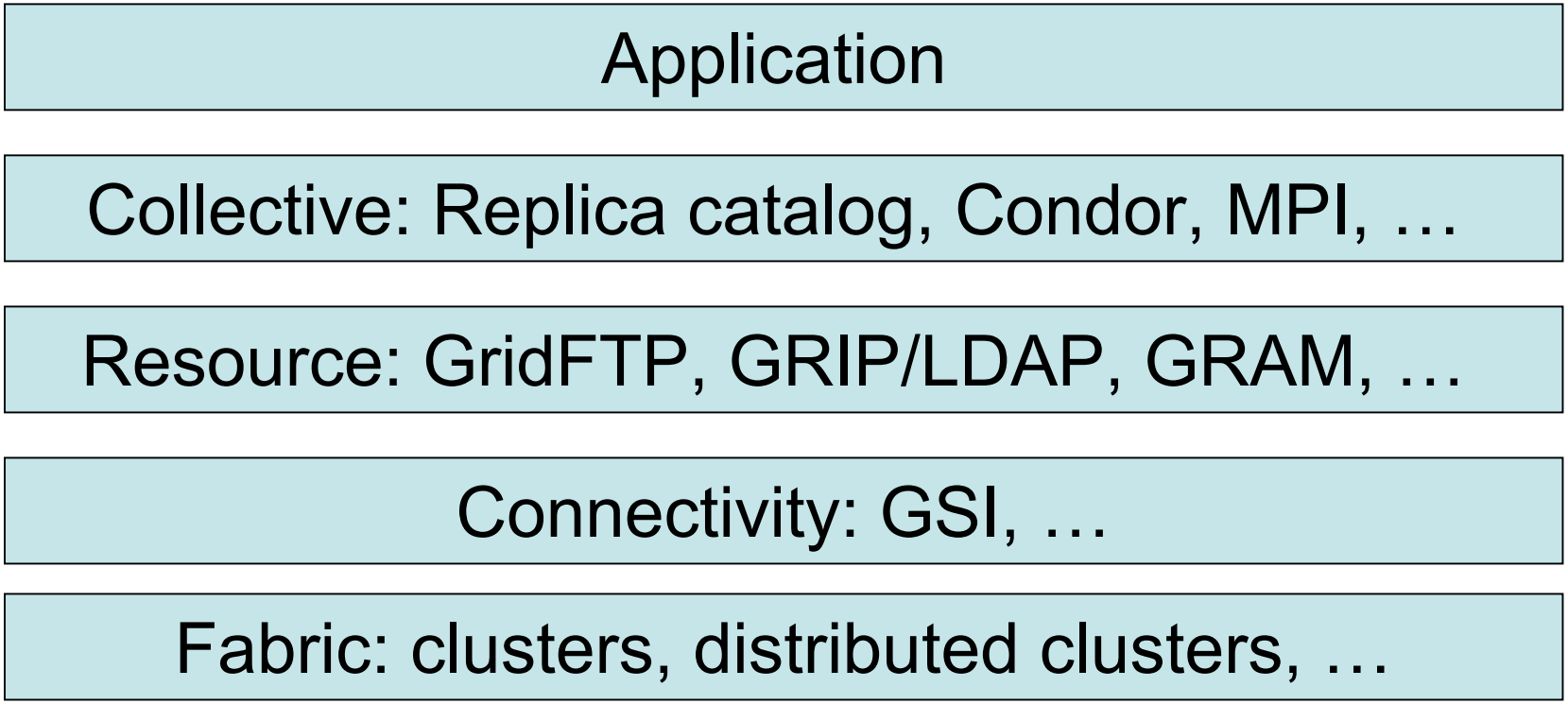
Copyright 2003 Robert L. Grossman

Three Data Stacks

Stacks for data grids, data webs and distributed data mining.

Copyright 2003 Robert L. Grossman

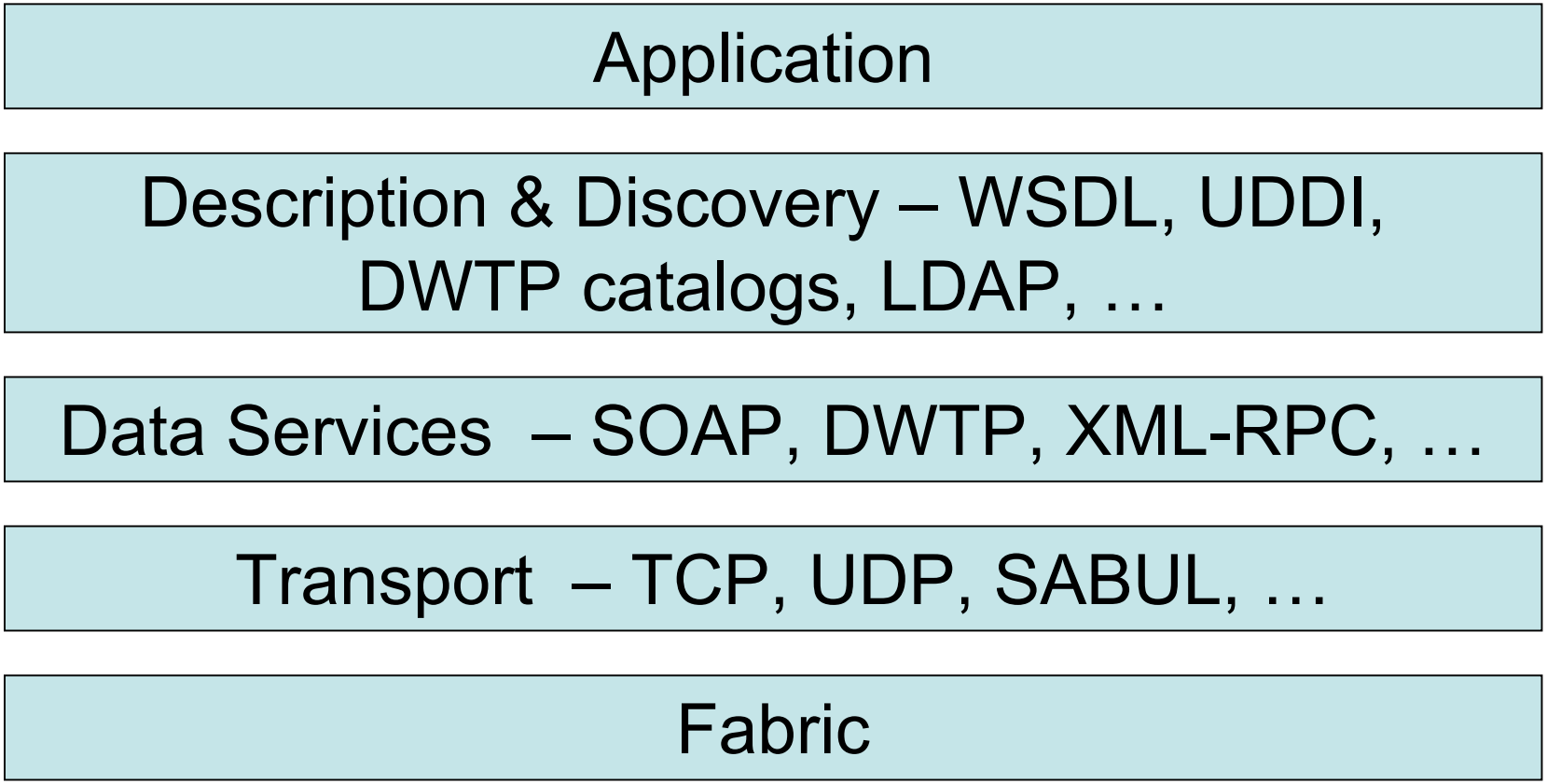
Data Grid Layers



- Much broader in scope than previous layers.
- Think of as virtual distributed computer.

Copyright 2003 Robert L. Grossman

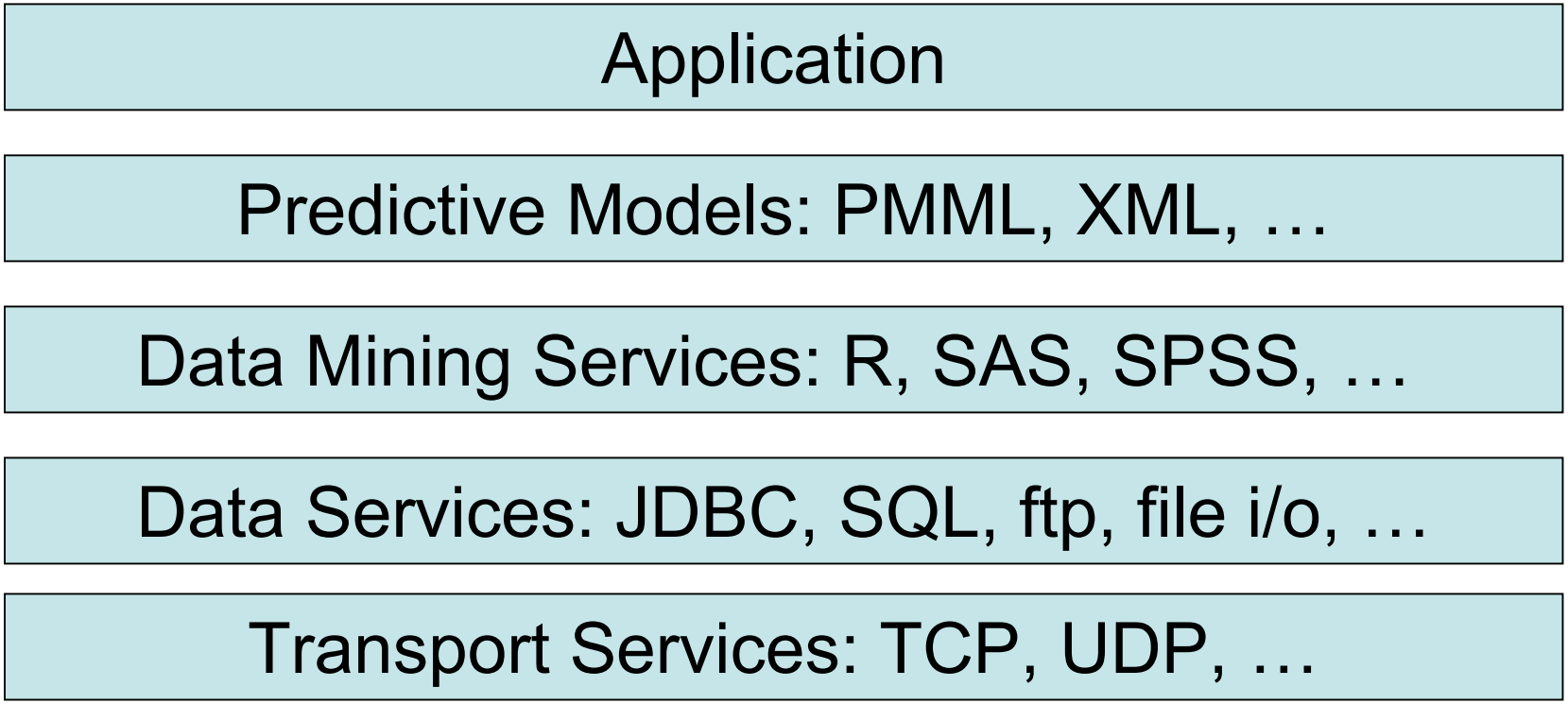
Data Web Layers



- data & metadata separated
- transport has specialized services for large data

Copyright 2003 Robert L. Grossman

Data Mining Layers



- Emergence of a predictive model layer.
- Standardization on JDBC, ODBC, SQL, …

Copyright 2003 Robert L. Grossman

3.2 Network Protocols

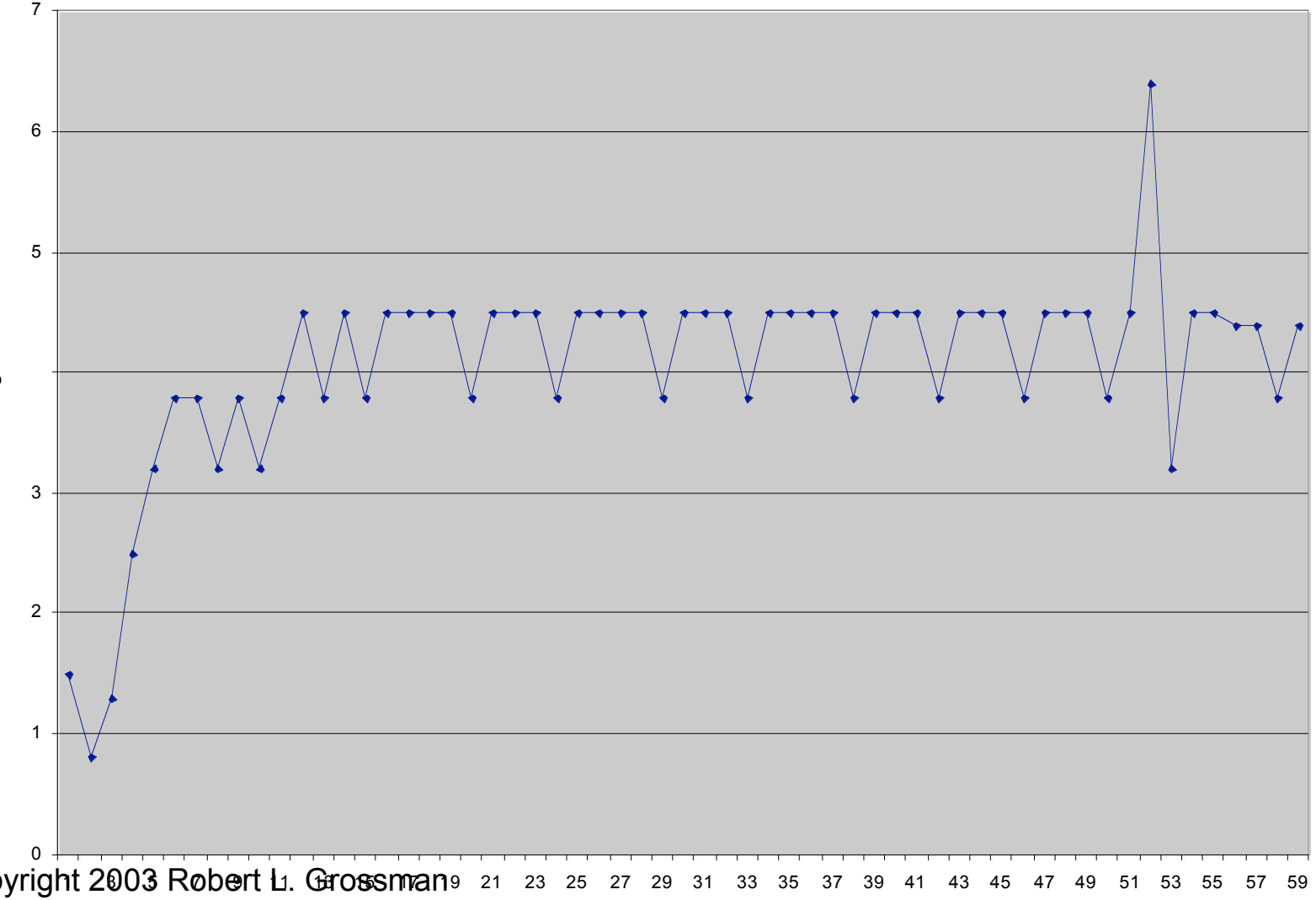
Protocols for moving bits.

Adjusting Window Size, Network Striping, Separating Data and Control, and Other Tricks

Copyright 2003 Robert L. Grossman

TCP Data Transport Chicago to Amsterdam over 622 Mb/s Link

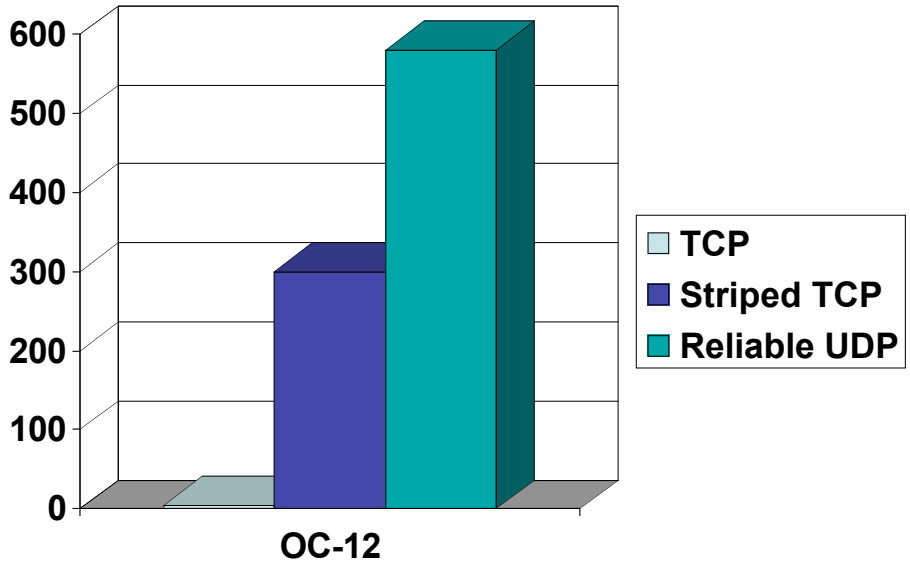
4.5 Mb/s



Copyright 2003 Robert L. Grossman

What is the Problem?

- TCP requires acknowledgement of packets – 1/RTT limit & is slow to recover from congestion
- Over 622 Mb/s OC-12 between Chicago and Amsterdam
  - TCP – 5 Mb/s
  - Striped TCP 300 Mb/s (PSockets)
  - Reliable UDP 580 Mb/s (SABUL)



Copyright 2003 Robert L. Grossman

Related Work

Approach	Implementations	Challenges
Improve TCP	Multiple	Will it scale?
Striped TCP	GridFTP, PSockets	Improve Performance
Reliable, Friendly UDP	SABUL, FAST, TSUNAMI	Make friendly
Striped UDP	P-SABUL	Interface to parallel I/O
Alternative Prot.	XCP	Requires new router code

Copyright 2003 Robert L. Grossman

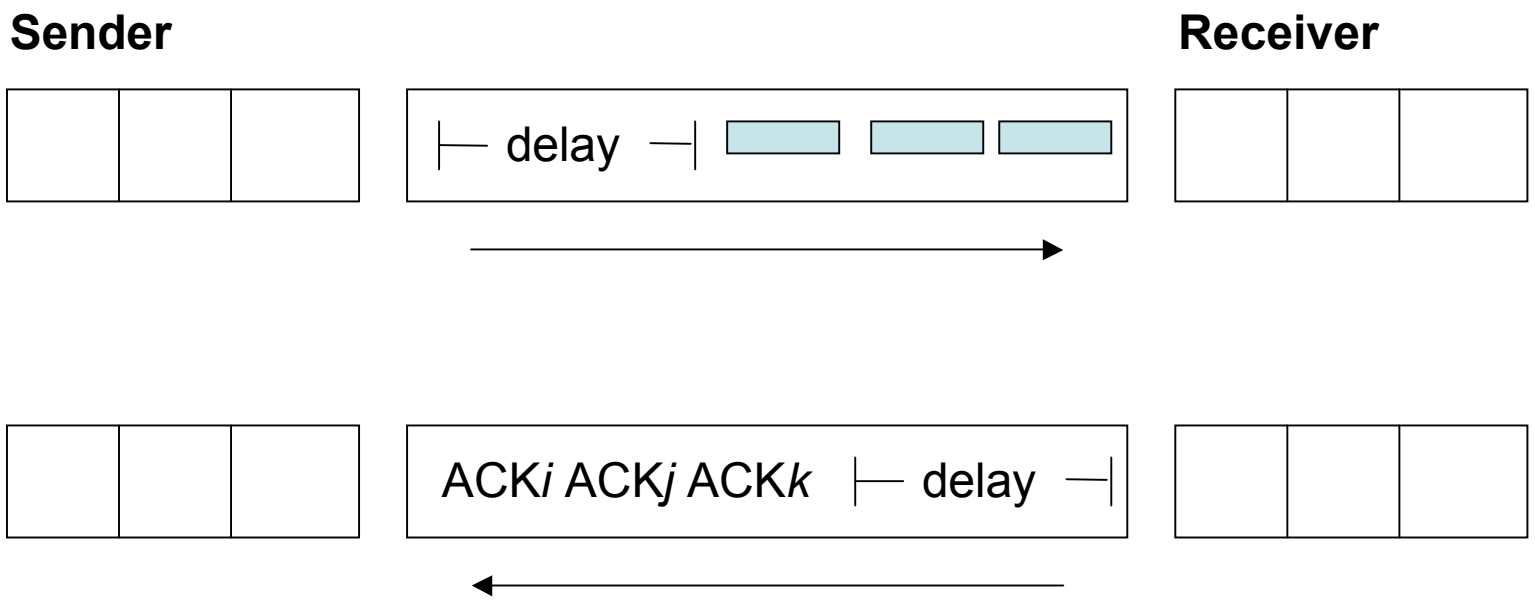
Approach 1: Change TCP Window Size

- Advantages:
  - improves bandwidth
  - does not require new protocol

- Disadvantages
  - requires careful tuning
  - requires kernel modifications along route
  - performance levels out

Copyright 2003 Robert L. Grossman

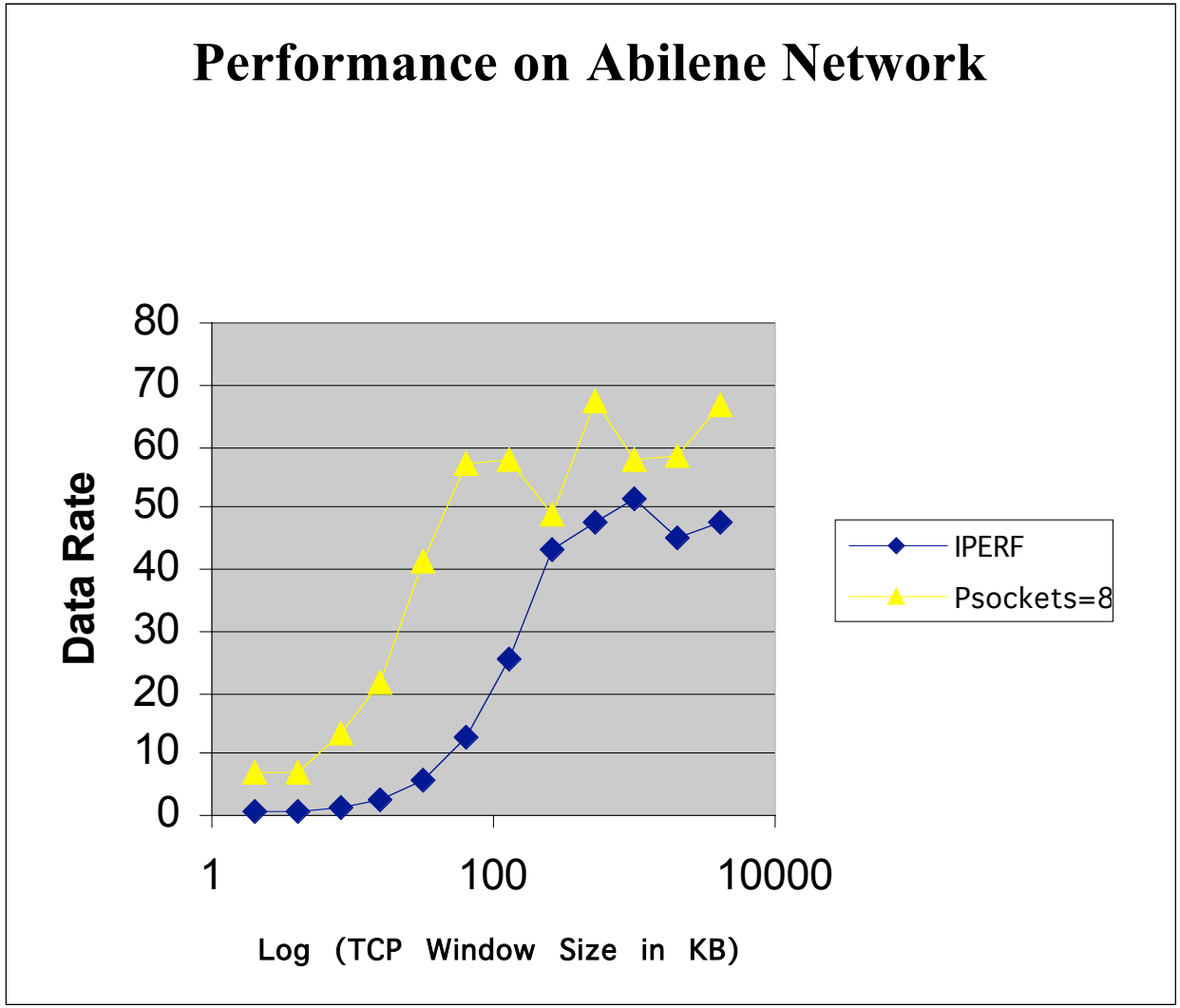
TCP *Bandwidth*Delay* Problem



Default TCP window size, RFC 1323 not enabled.

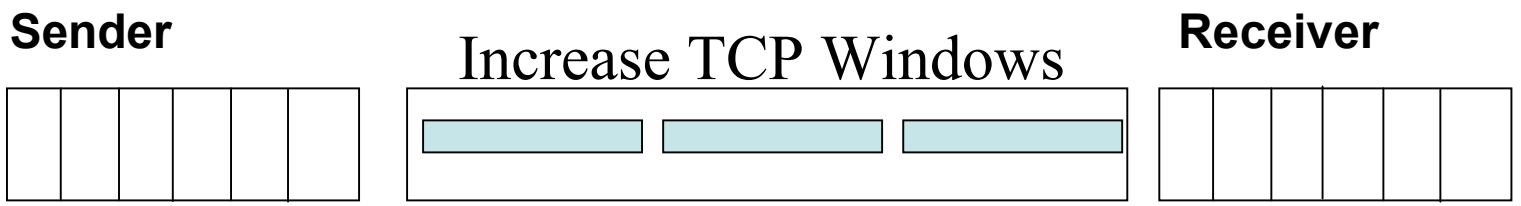
Copyright 2003 Robert L. Grossman

Varying TCP Window Size - OC-3

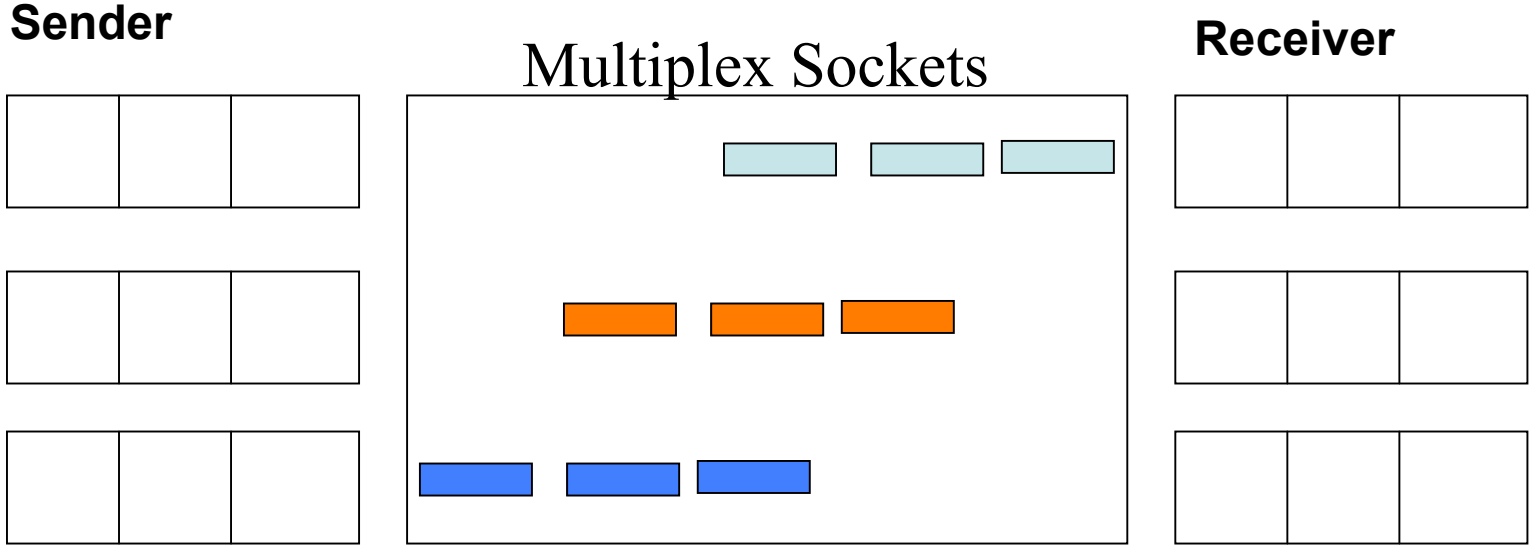


Copyright 2003 Robert L. Grossman

Approach 2: Network Striping



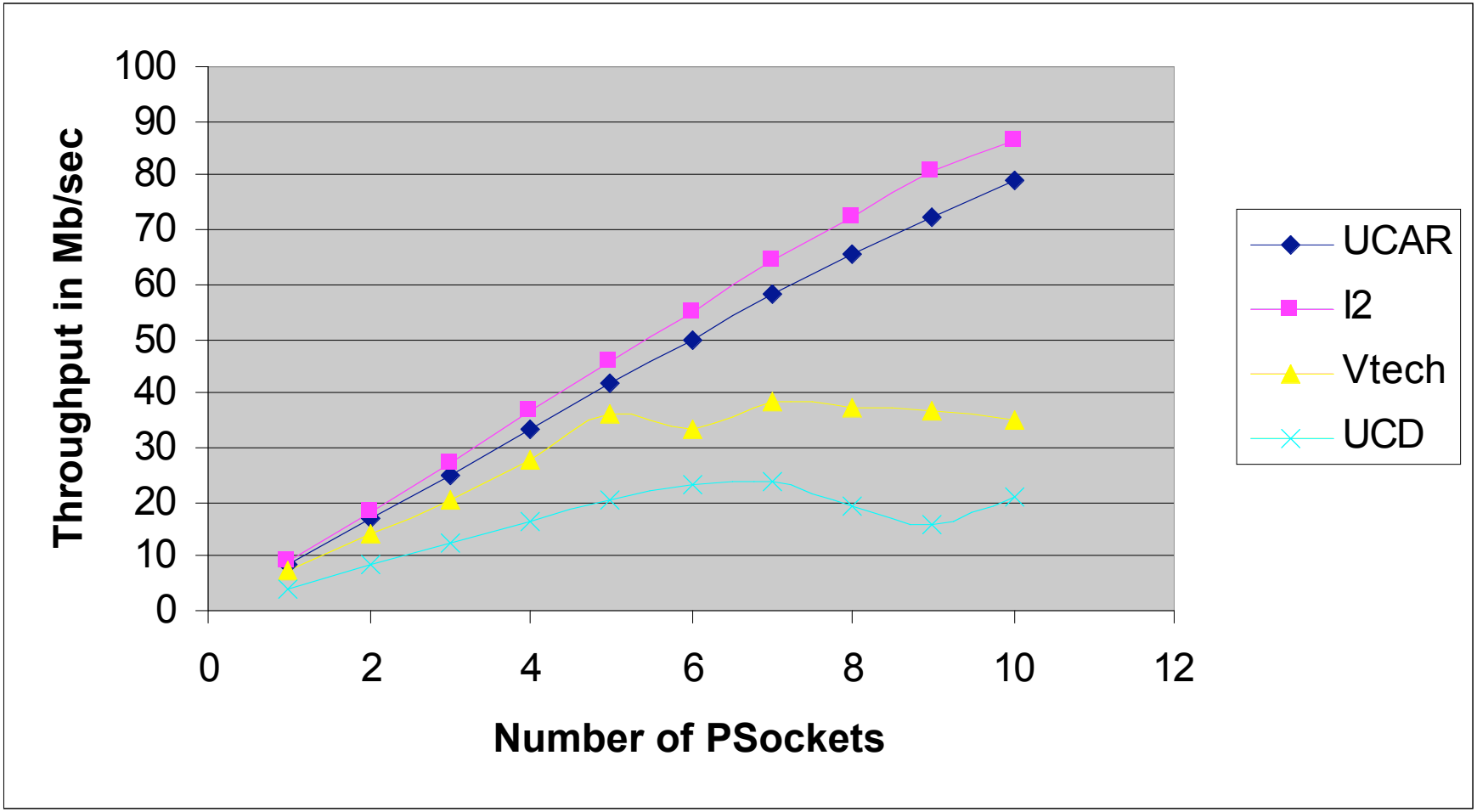
Set TCP window size to bandwidth*delay product, RFC 1323 enabled.



Default TCP window size, applications using PSockets with value 3

Copyright 2003 Robert L. Grossman

Network Striping with TCP Sockets (1998)



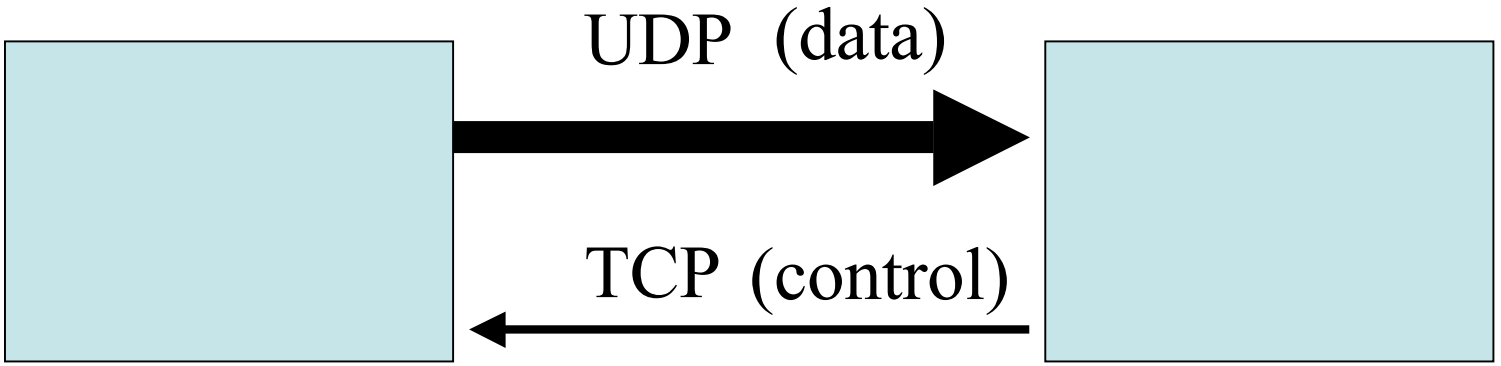
Copyright 2003 Robert L. Grossman

Approach 2: Network Striping Summary

- Advantages:
  - improves bandwidth
  - application level library; no changes to os
  - many implementations: PSockets, GridFTP, …
- Disadvantages
  - levels out after 10-20 sockets
  - better techniques for OC-12 and higher links

Copyright 2003 Robert L. Grossman

Approach 3: UDP Data and TCP Control



Advantages

- Available bandwidth utilization
- TCP friendly behavior
- Performance independent of latency

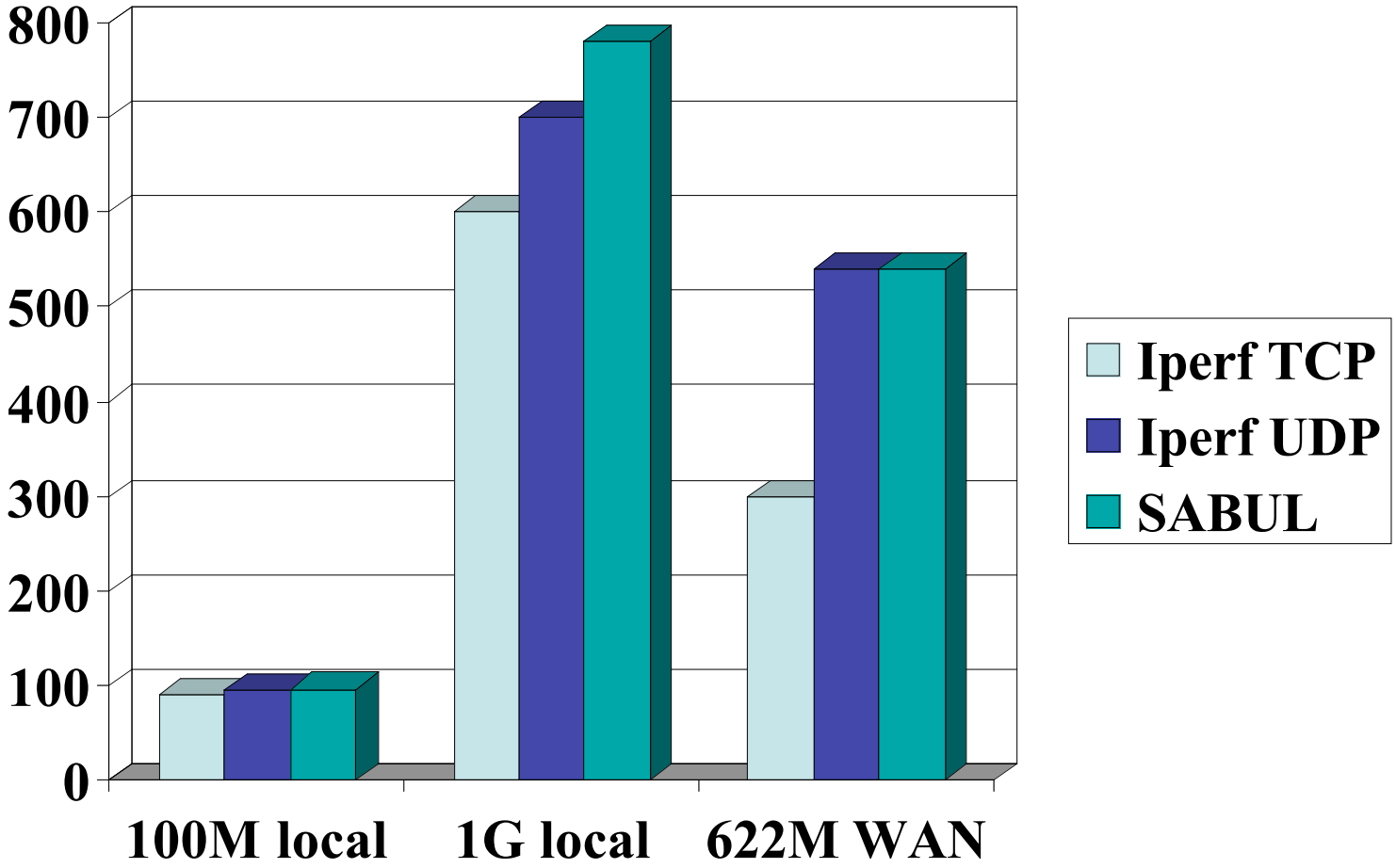
Copyright 2003 Robert L. Grossman

Approach 3: Reliable, Friendly UDP

- High performance reliable protocol
- Take advantage of UDP's fast transmission and TCP's reliability.
- Send packets with sequence number by UDP
- Exchange control information by TCP to revise packet error/loss and achieve traffic control
- SABUL is an implementation
  - currently version 2.2

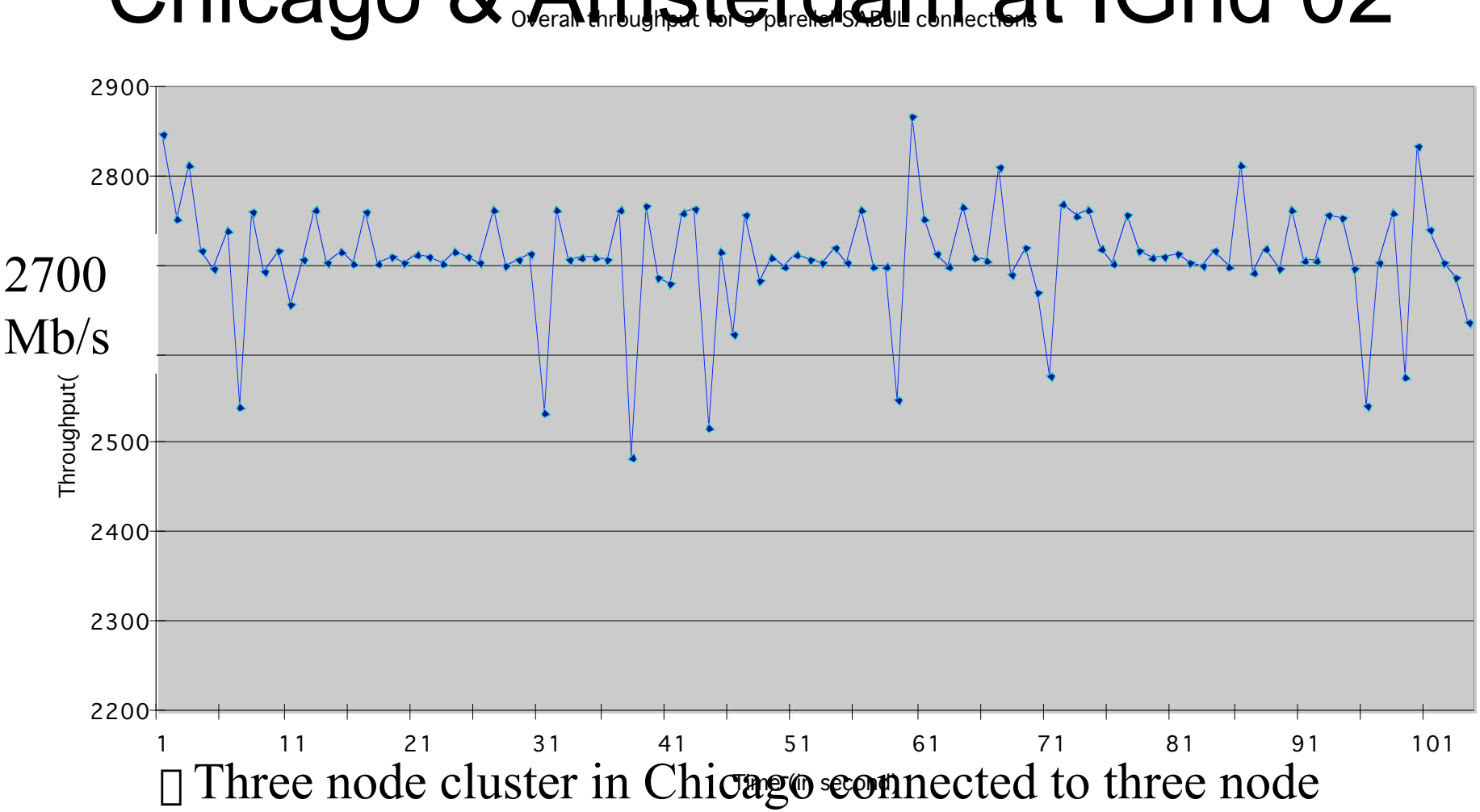
Copyright 2003 Robert L. Grossman

SABUL Release 1.0 (2001) Performance



Copyright 2003 Robert L. Grossman

PDS Data Transport between Chicago & Amsterdam at IGrid 02



- Three node cluster in Chicago connected to three node cluster in Amsterdam connected with 10 Gb/s link

Copyright 2003 Robert L. Grossman

P-SABUL (Striped SABUL) Chicago - Amsterdam

TCP Stream	SABUL Stream 1	SABUL Stream 2	SABUL Stream 3	Striped SABUL Stream
4.36 Mb/s	902.8 Mb/s	902.9 Mb/s	907.1 Mb/s	2712.8 Mb/s

- Three node cluster in Chicago connected to three node cluster in Amsterdam connected with 10 Gb/s link

Copyright 2003 Robert L. Grossman

Comparing Reliable UDP & Striped TCP

Data Set (MBs)	GridFTP (Mb/s)	SABUL (Mb/s)
100	94.9	527
500	246	476
1000	324	506
2000	315	506

- Experiments between Chicago and Amsterdam over OC-12

Copyright 2003 Robert L. Grossman

Approach 3: Reliable UPD Summary

- Advantages:
  - application level library
  - friendly protocol
  - several implementations: SABUL, BlastUDP, …
  - tested above 2.5 Gb/s
- Disadvantages
  - not as standard as TCP
  - requires that applications interface to new library

Copyright 2003 Robert L. Grossman

3.3 Data Protocols

Protocols for Moving Data

or why data repositories are different than bit repositories.

Copyright 2003 Robert L. Grossman

What is the Problem?

- Data is different than bits
- Data has metadata, data has keys, data can be merged, etc.

- By using specialized protocols for moving data, data intensive applications can be built more easily
- With these protocols, distributed data mining applications are also easier to build

Copyright 2003 Robert L. Grossman

What are the Choices?

1. HTTP
2. XML Messaging/SOAP
3. Data Web Transfer Protocol (DWTP)

- Note well: ftp and its variants are for moving bits…don't even think about it…
- Other specialized data protocols being developed…

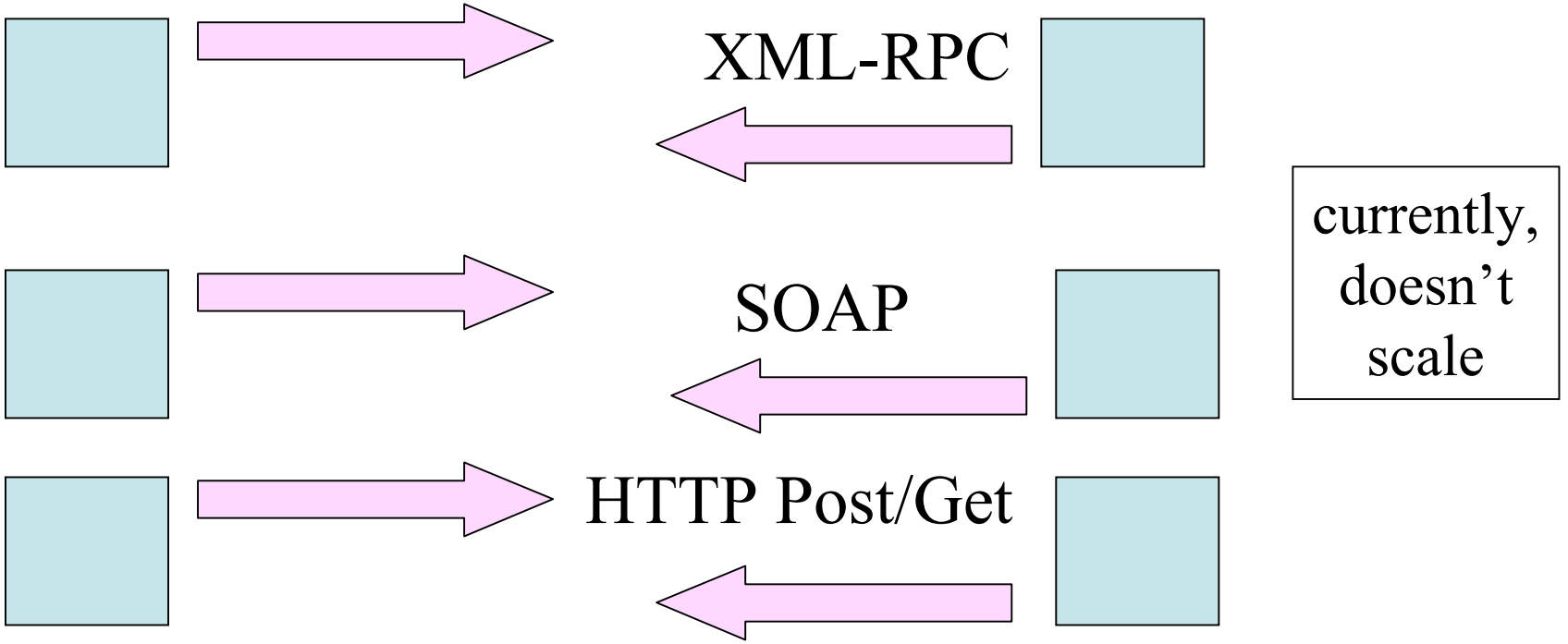
Copyright 2003 Robert L. Grossman

1. HTTP – Common in Practice

- In practice, data is often sent over HTTP
- Advantages:
  - very widely available
  - good for presenting data, not working with data
  - but SOAP, XML, etc. can be sent over HTTP
- Disadvantages:
  - not specific to data
  - HTML not well suited for data
  - XML is better for data, especially small data

Copyright 2003 Robert L. Grossman

2. XML Messaging



- XML good for metadata
- XML good for small data

Copyright 2003 Robert L. Grossman

SOAP

- XML messaging provides an application and platform independent means of sharing data
- SOAP is a good mechanism for sending XML messages
- Focus to date is on using SOAP for sending XML-RPCs over HTTP
- SOAP messages consists of
  - SOAP envelope
  - SOAP header
  - SOAP body

Copyright 2003 Robert L. Grossman

Example: SOAP Google Interface

```
<SOAP-ENV:Envelope
   xmlns:SOAPENV="http://schemas.xmlsoap.org/... >
   <SOAP-ENV:Body>
        <ns1:doGoogleSearch xmlns:ns1="urn:GoogleSearch"  … ">
        <key xsi:type="xsd:string">XXXXXXXXXXXX</key>
        <q xsi:type="xsd:string">data </q>
        <start xsi:type="xsd:int">0</start>
        <maxResults xsi:type="xsd:int">10</maxResults>
        <filter xsi:type="xsd:boolean">true</filter>
        …
        </ns1:doGoogleSearch>
   </SOAP-ENV:Body>
</SOAP-ENV:Envelope>
```

Copyright 2003 Robert L. Grossman

3. Data Web Transfer Protocol (DWTP)

- protocol designed for data
- supports data, metadata, and keys
- separates control from data channels
- can subset data by rows or columns
- mechanisms for sampling, merging data by key, working with missing values
- interoperates with web services

Copyright 2003 Robert L. Grossman

Example: DWTP Session

- Discover DWTP server containing appropriate data using web services
- DWTP client connects to DWTP server
- retrieve data set metadata using TCP
- set data set
- retrieve attribute metadata using TCP
- retrieve 25 columns of data using 20% subset of rows using SABUL

Copyright 2003 Robert L. Grossman

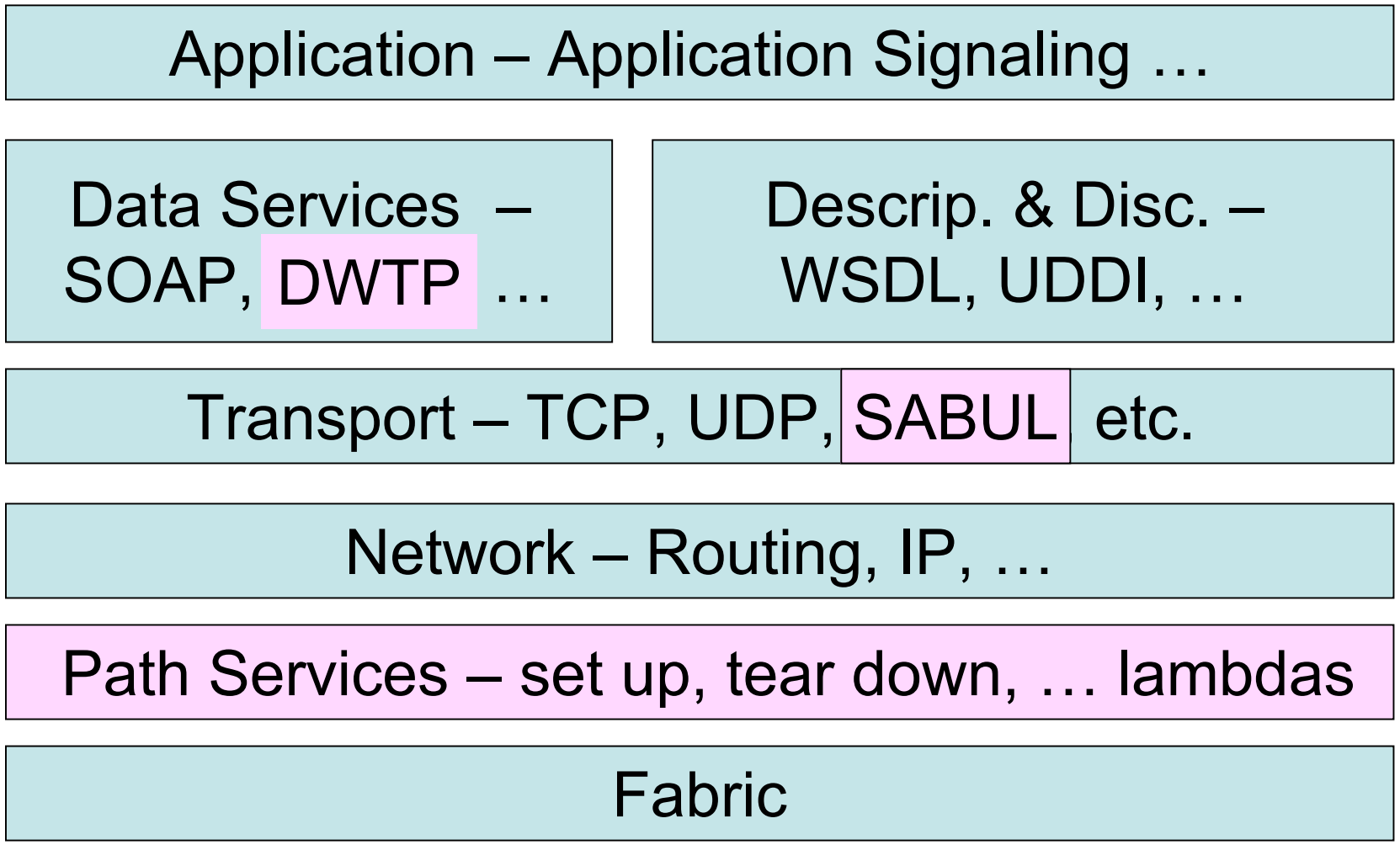
3.4 Photonic Data Services

Protocols for Moving Data over Lambdas

or what to do if you have your own lambda.

Copyright 2003 Robert L. Grossman

Photonic Data Services



Copyright 2003 Robert L. Grossman