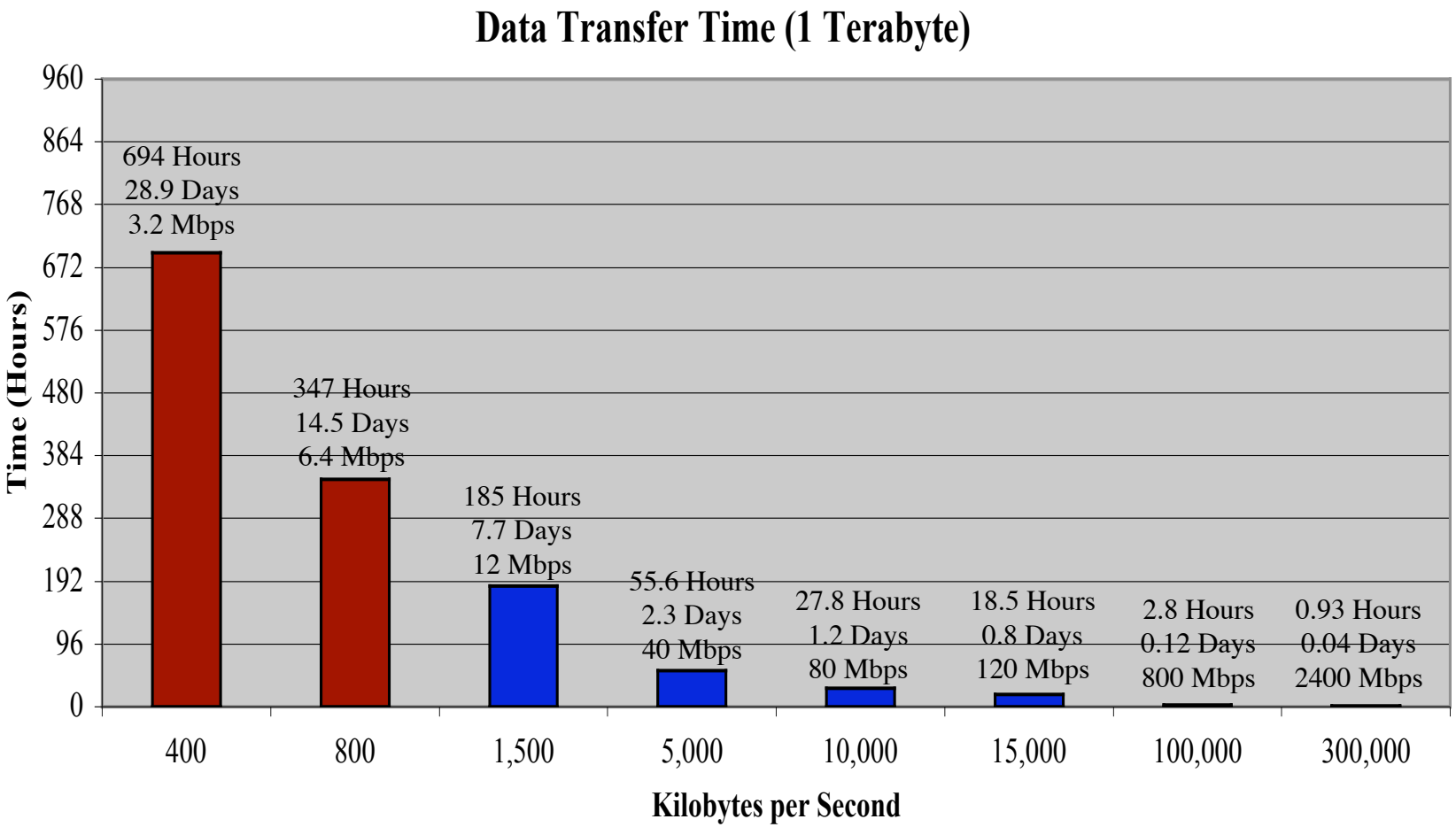
Improving Network Data Transfer Rates



September 2006

Agenda

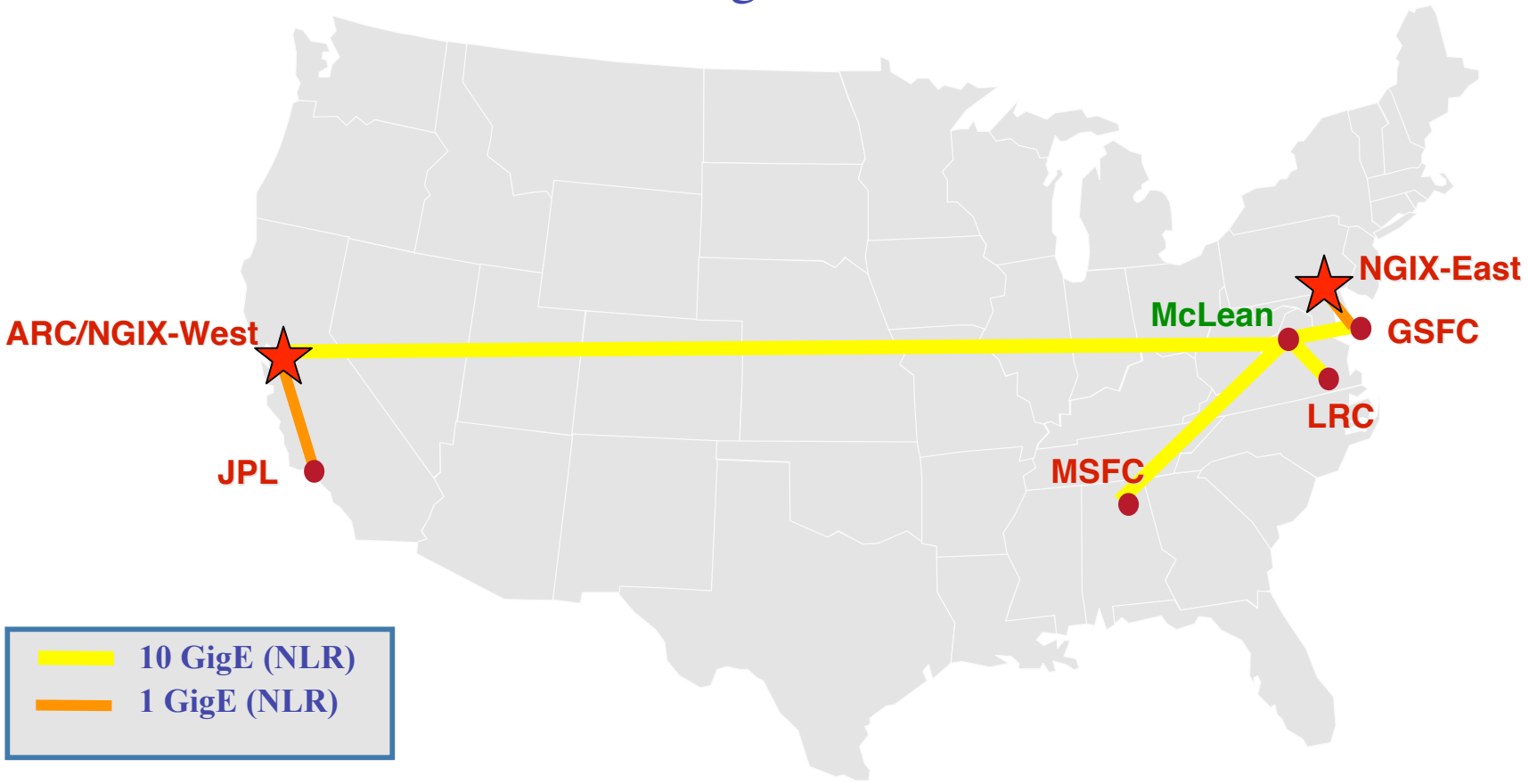
- NREN Overview
- Goddard to Ames Network Overview
- TCP Windows Sizes
- TCP Performance Tuning on End-Systems
- BBFTP
- Impact of Firewalls
- NPAD

NASA Research and Engineering Network (NREN) Overview

- NREN is a research network, that is currently focused on supporting NASA's High-End Computing (HEC) networking requirements
- Specifically, the Columbia Supercomputer is located at NASA Ames Research Center
- NASA Advanced Supercomputing (NAS) Division manages both NREN and the Columbia Supercomputer

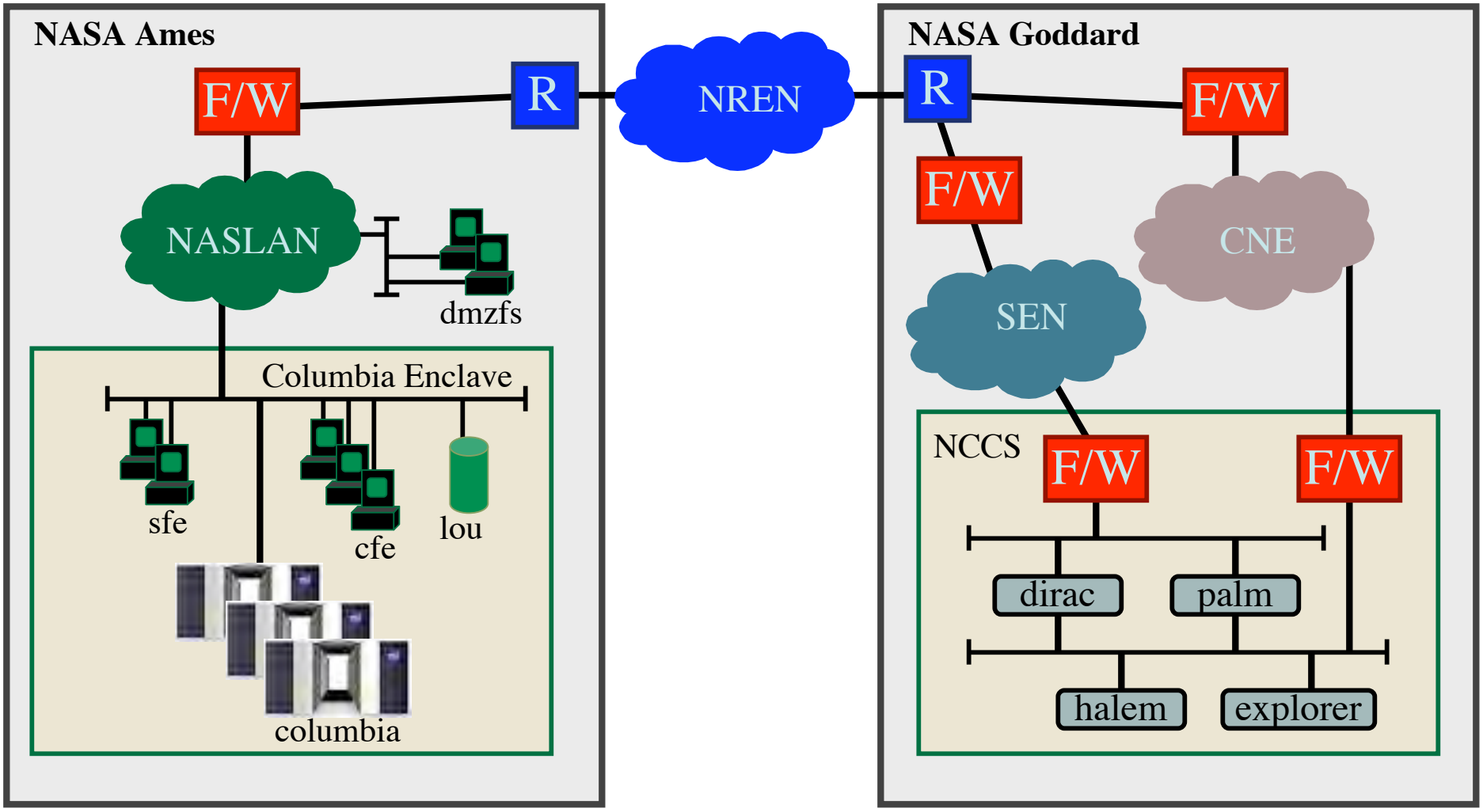
NREN Overview

Target CY06



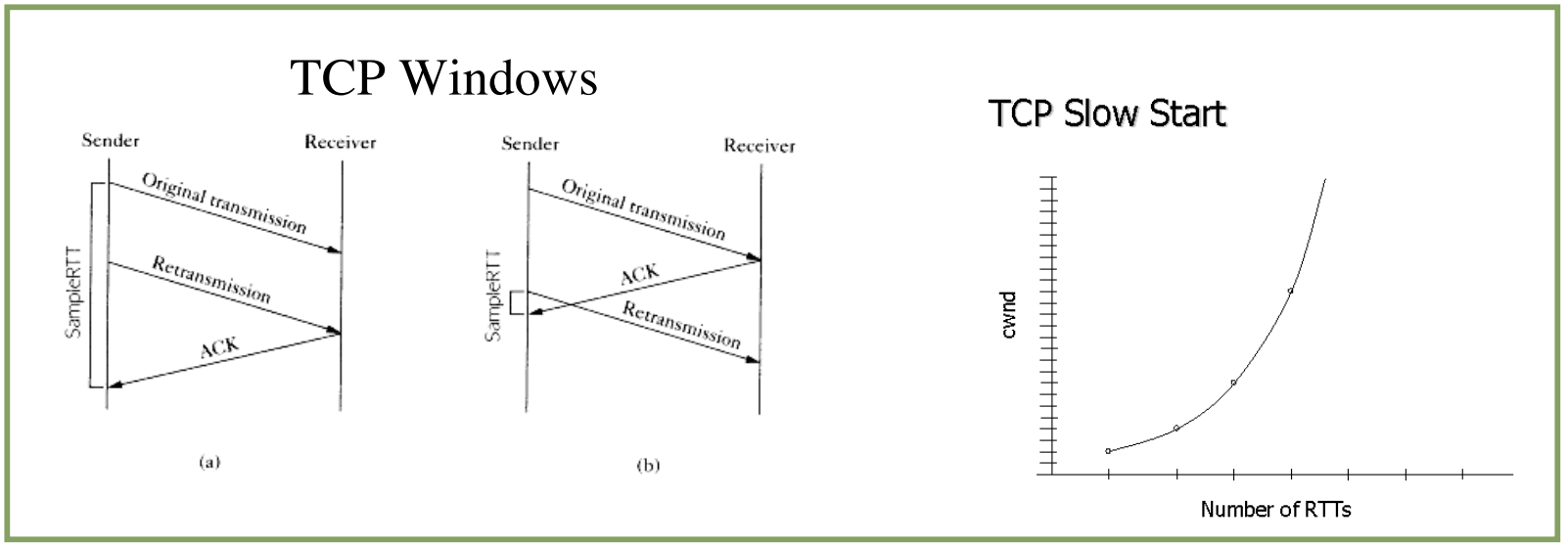
Maximum Disk-to-Disk Data Transfer Rate
GSFC (dirac.nccs.nasa.gov) - ARC (lou.nas.nasa.gov)
450,000 kilobytes per second (3.6 Gbps) - September 2006

Goddard to Ames Network Overview



TCP Windows Sizes

- IP Packet round-trip times between ARC and GSFC over NREN is nominally 88 ms.
- Given this RTT and the network capacity, a large amount of data must be "in flight" to efficiently utilize the network.



- TCP/IP hosts (by default) are not configured for these large "in flight" data rates.
- TCP Send and Receiver windows (aka buffers) should be set to AT LEAST 5 MBytes, to achieve transfer rates greater than 300 Mbps (375,000 kilobytes per second) with 88 ms RTT.

TCP Performance Tuning on End-Systems

- Most OS vendors provide parameters to adjust window sizes.
  - UNIXes typically have two to four adjustable kernel parameters.

Linux parameters can be set with ***sysctl*** command:

```
# Set maximum TCP window sizes to 100 megabytes
 net.core.rmem_max = 104857600
 net.core.wmem_max = 104857600
 # Set minimum, default, and maximum TCP buffer limits
 net.ipv4.tcp_rmem = 4096 524288 104857600
 net.ipv4.tcp_wmem = 4096 524288 104857600
 # Set maximum network input buffer queue length
 net.core.netdev_max_backlog = 30000
 # Disable caching of TCP congestion state (2.6 only)  *Fixes a
  bug in some Linux stacks.
 net.ipv4.tcp_no_metrics_save = 1
```

TCP Performance Tuning on End-Systems

Mac OS X parameters can be set with *sysctl* command:

 # Set maximum TCP window sizes to 16 megabytes

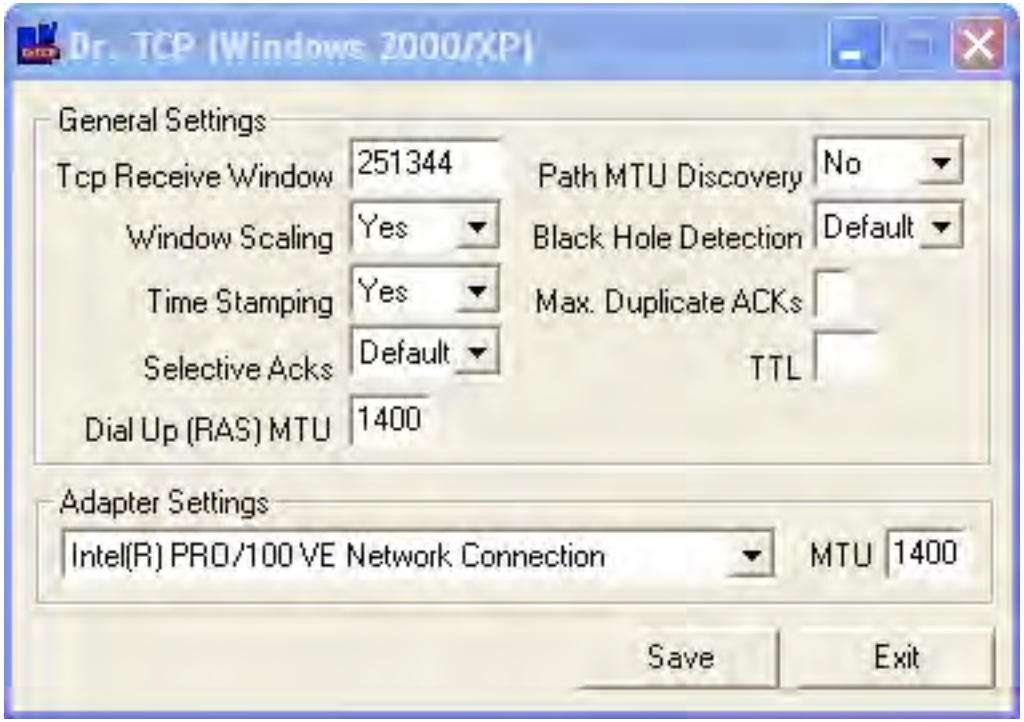
- net.inet.tcp.sendspace= 16772216
- net.inet.tcp.recvspace= 16772216

 # Set maximum Socket Buffer sizes to 128 megabytes

- kern.ipc.maxsockbuf= 134217728

TCP Performance Tuning on End-Systems

- For an 88ms round-trip-time link with 1 GbE interfaces, you'll probably want to use 8-10MB for the maximum TCP window size.
- Suggested tool: Dr TCP
  - http://www.dslreports.com/drtcp
  - Screenshot of DR TCP settings for WinXP on a 100 Mbps-attached host (fast ethernet MTU is limited to 1500):



BBFTP

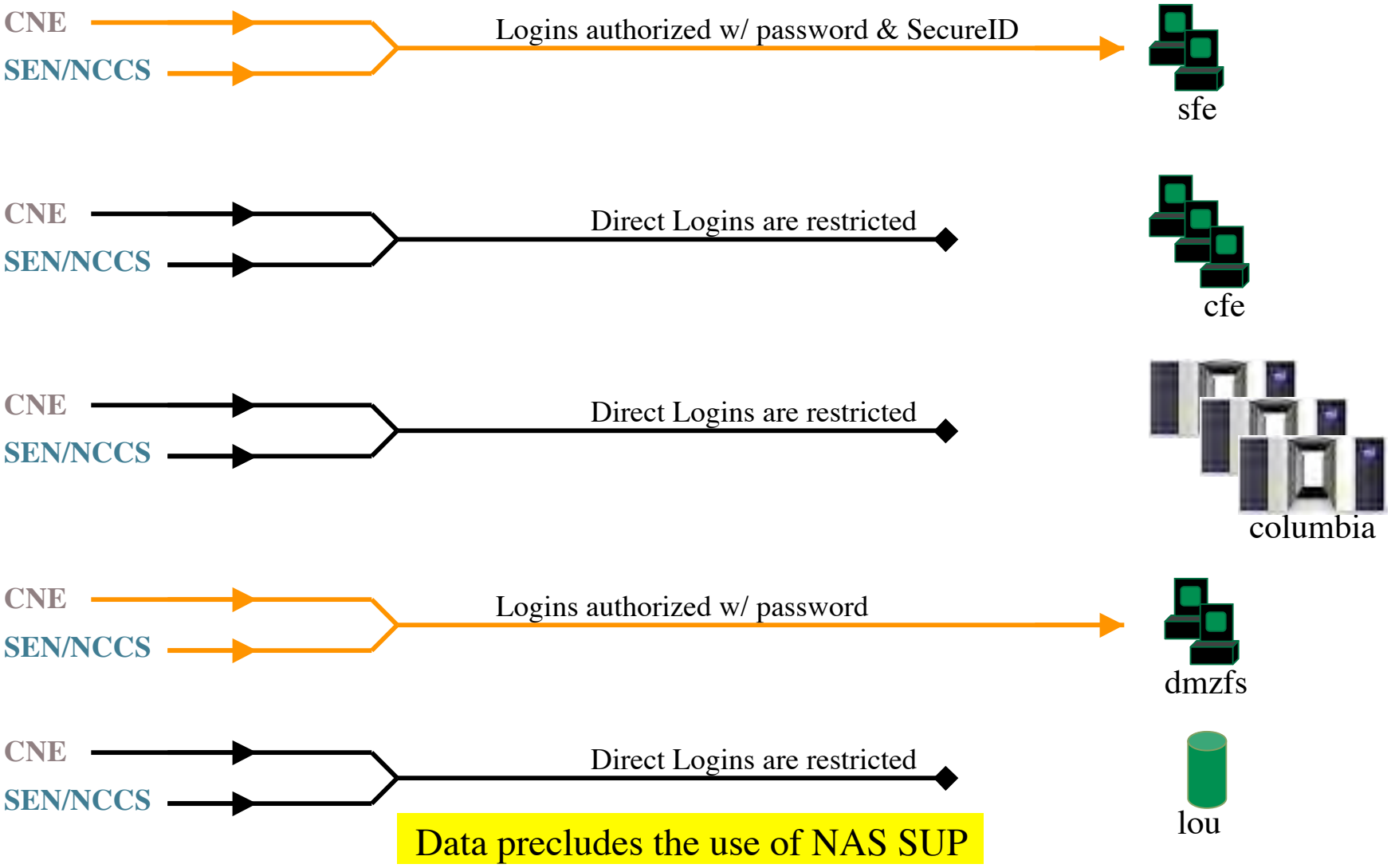
- BBFTP is a non-interactive FTP-like system that supports parallel tcp streams for data transfers, allowing it to achieve bandwidths that are greater than normal FTP. Because of these characteristics, it is the preferred method for transferring large data files
- BBFTP only **encrypts** usernames and passwords, it does **NOT** encrypt the data being transferred.
- In order to use BBFTP, you will need to obtain a BBFTP client. The BBFTP source code can be obtained from the BBFTP web site.
  - http://doc.in2p3.fr/bbftp/

Impact of Firewalls

- In general firewalls will limit single stream data transfer rates.
- A significant amount of this limitation can be reduced by utilizing multi-stream file transfers (including multiple simultaneous SCP file transfers

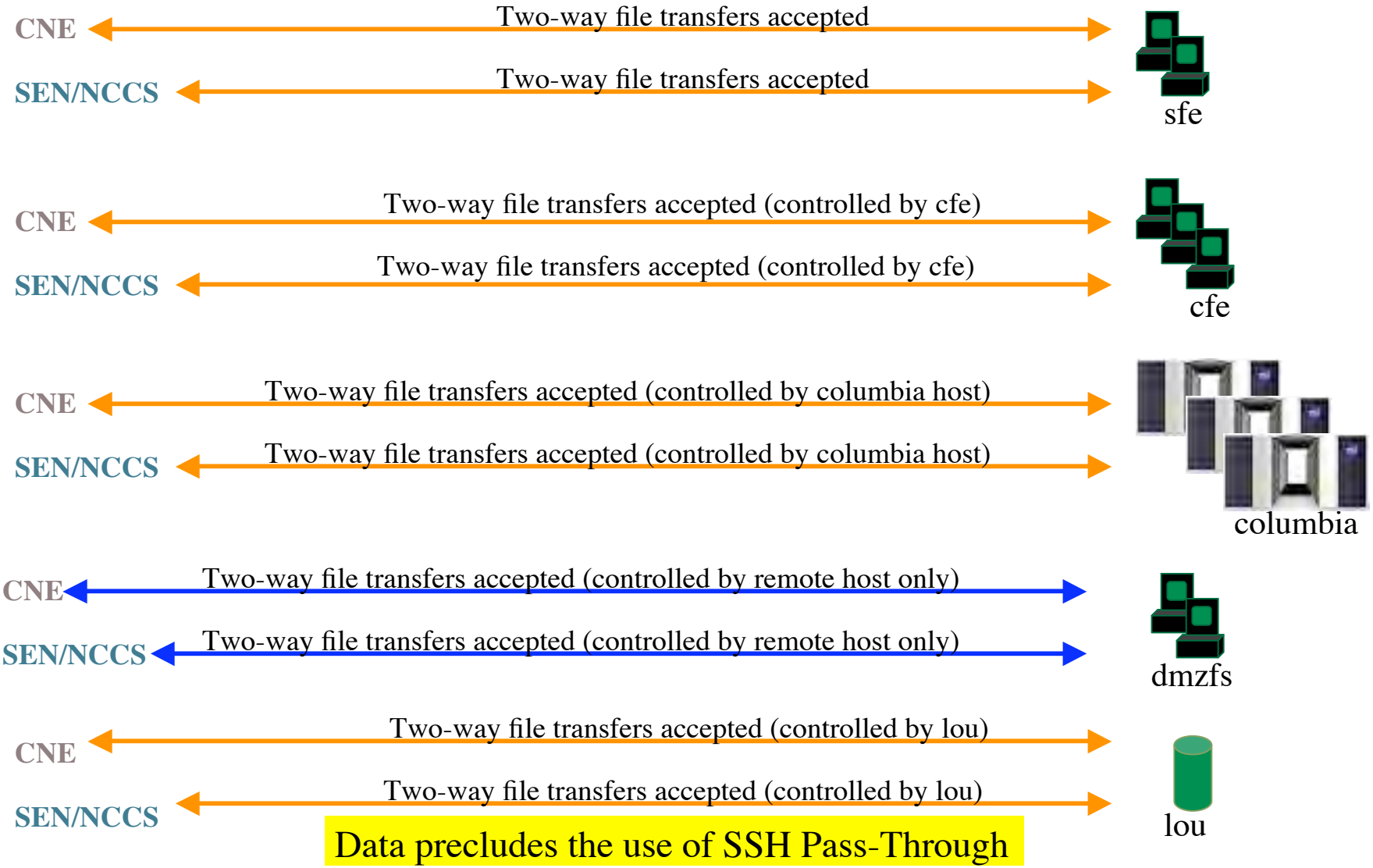
Impact of Firewalls (Bastions)

Login-In (Authorization) Restrictions



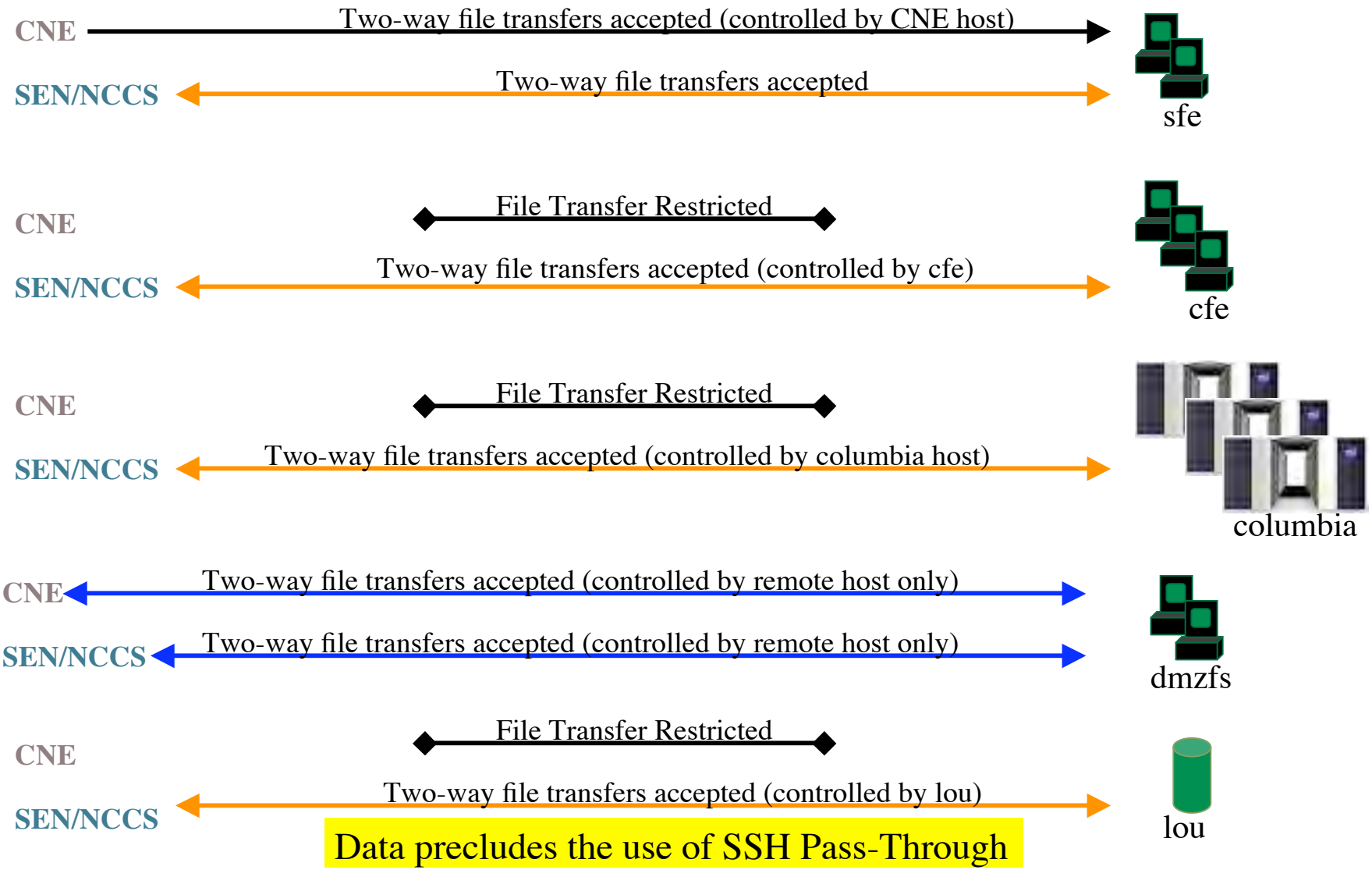
Impact of Firewalls (Bastions)

File Transfers (SCP)



Impact of Firewalls (Bastions)

File Transfers (BBFTP)



Impact of Firewalls (Bastions)

CNE Firewall Waiver Request

FOR CFRB USE ONLY

Date: ☐ Approved ☐ Denied ☐ Held ☐ CIO Waiver, *duration:*

Comments:

Member	DCSEs									ITSM	CNE Sec	GSFC NSM	CFRB Chair	CNE Engr'g		Total
	100	150	200	300	400	500	600	700	800					Gmbh	WFF	
Approve																
Deny																
Abstain																

CNE Firewall Waiver Request No. ____-____
(Number Assigned by CFRB)

1. Date: 7/11/2006
2. ☒ New ☐ Renewal ☐ Modification
3. Project or Organization Name: NASA Ames - High End Computing Columbia
4. Code: Ames/TN
5. Description of application (include version number where applicable) or Network Service :
6. Justification, alternatives, and impacts:
7. Classification of transmitted data: ☒ Public; ☐ SBU; ☐ Export Controlled (e.g. ITAR, EAR);
   When transmitted, the data is: ☐ Encrypted; ☒ Not encrypted; ☐ Other
8. ☐ Valid IPAMS record

9. Source IP Address	10. Destination IP Address	11. Destination Port/Protocol
		TCP/5020 - 5050
		TCP/5020 - 5050
		TCP/5020 - 5050
		TCP/5020 - 5050

Special requirements, comments:

NPAD (Network Path and Application Diagnosis)

- Intended to help users (and network folks) troubleshoot their network connections.
- NPAD is a server accessible via HTTP and/or the CLI.
- Users connect to the server and initiate the test.
- The test will perform several diagnostic routines and suggest a possible solution.
- The results are recorded (on server) for network folks, who can then examine the results for the user and suggest a possible solution.

- Available at ARC (for troubleshooting NREN/Columbia issues):
  - http://binkley.nren.nasa.gov:8000/

- Demo

Contact Information

- Ken Freeman
  - Kenneth.Freeman-1@nasa.gov
- Dave Hartzell
  - dhartzell@mail.arc.nasa.gov
- Websites
  - www.nren.nasa.gov/customer.html
- Help Desk
  - NAS Control Room (650) 604-4444