

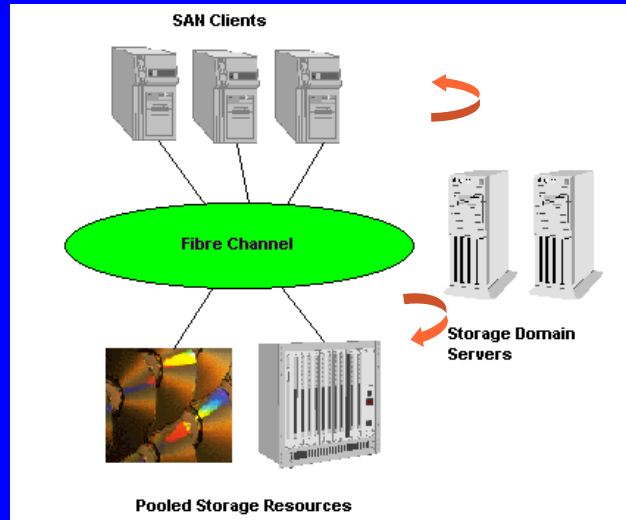
A model of SAN performance

Dr. Alexandre Brandwajn
PALLAS International Corporation
San Jose, CA

Plan

- ◆ Motivation
- ◆ SAN and model
 - logical I/O
 - Storage Domain Server
 - PCI bus model
- ◆ Results example

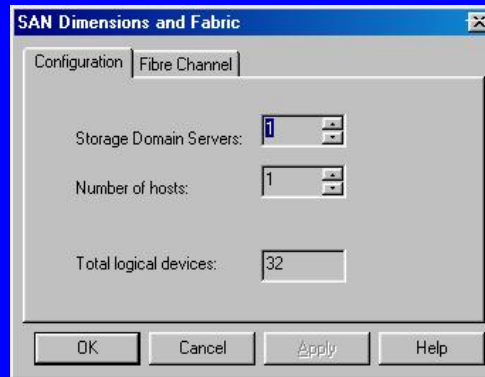
In-band SANsymphony



Model goals

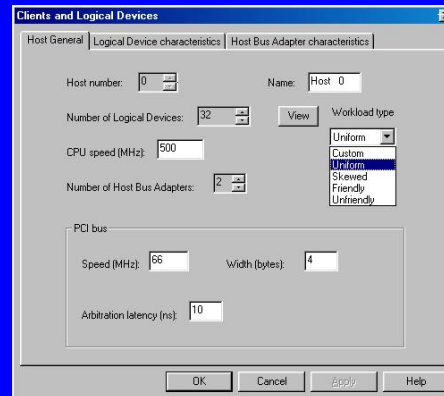
- ◆ Expected I/O performance for clients
- ◆ SAN dimensioning
 - number of servers
 - server power
 - server configuration
 - internal bottlenecks
- ◆ Operating range

SAN configuration



Client model

- ◆ I/O workload
 - logical devices
- ◆ “typical” workloads
- ◆ host computer
 - speed
 - SAN interface
 - future model



I/O workload model

- ◆ Extensive Lead Time experience
 - I/O burstiness
 - transfer and seek variability
 - limited queue
 - read hit behavior
- ◆ Mapping
 - mirroring

I/O workload model

Clients and Logical Devices

Host General | **Logical Device characteristics** | Host Bus Adapter characteristics

Host: Host 0 Logical device number: 0 Name: Ldv 0

All logical devices

Workload

<input checked="" type="checkbox"/> I/Os per second: 20	<input checked="" type="checkbox"/> I/O rate variability: 1
<input checked="" type="checkbox"/> Block length (Kbytes): 16	<input checked="" type="checkbox"/> Transfer variability: 0
<input checked="" type="checkbox"/> Fraction of zero seeks: 0.5	<input checked="" type="checkbox"/> Seek time variability: 1
<input checked="" type="checkbox"/> Fraction of reads: 0.75	<input checked="" type="checkbox"/> Hits on reads fraction: 0.85
<input checked="" type="checkbox"/> Queueing limit: 0	<input checked="" type="checkbox"/> Buffer hit probability: 0.5

Configuration

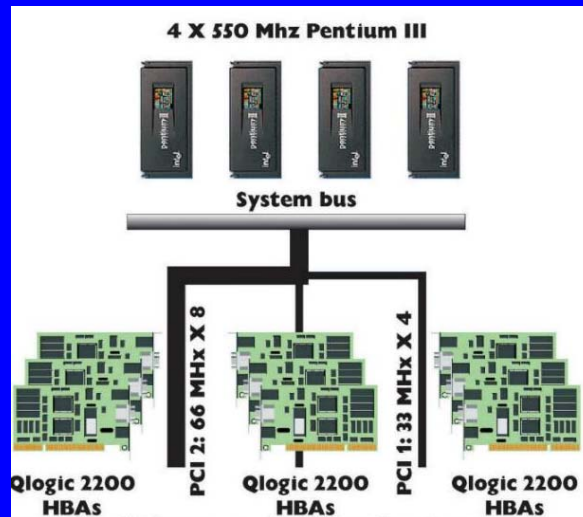
<input checked="" type="checkbox"/> SDS number: 0	<input checked="" type="checkbox"/> Physical device: 0
<input checked="" type="checkbox"/> Mirroring type: None	
<input checked="" type="checkbox"/> Index of mirror SDS: 0	<input checked="" type="checkbox"/> Mirror device: 0

OK Cancel Apply Help

Storage Domain Servers

- ◆ Set of physical devices
- ◆ Server CPUs
 - polling
- ◆ PCI buses
- ◆ Host Bus Adapters
 - PCI bus \Leftrightarrow Fibre Channel

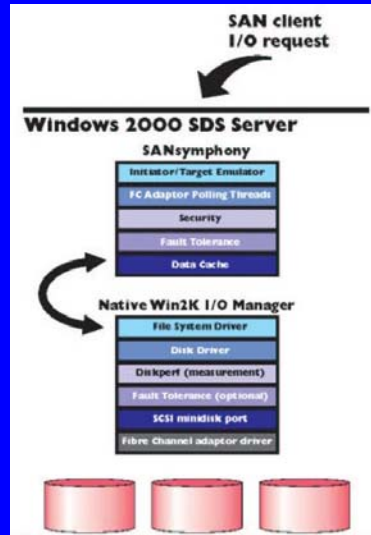
Example of SDS



I/O operations

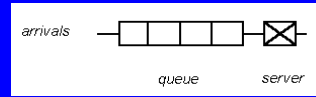
- ◆ Read hit
 - served from SDS cache
- ◆ Read miss
 - I/O Request Packet ==> Windows I/O Manager
 - data first copied into SDS cache
- ◆ Write
 - synchronous and asynchronous mirroring
 - mirror in another SDS
 - asynchronous write back

I/O operations



SDS model

- ◆ Physical disk
 - G/G/1/N approximation
 - queueing and service time
- ◆ CPU and polling
 - overheads and threads
- ◆ PCI buses and HBAs
 - effective bandwidth
 - future: main bus, memory access



SDS model - general

The screenshot shows a configuration window titled "Storage Domain Server" with several tabs. The "SDS General" tab is selected. The window contains the following fields and values:

Field	Value
SDS number:	0
Number of CPUs:	4
Number of PCI buses:	3
Number of Host Bus Adapters:	9
Number of physical disks:	16
Number of device types:	1
CPU speed (MHz):	550

At the bottom of the window are four buttons: "OK", "Cancel", "Apply", and "Help".

SDS model - CPU overheads

The screenshot shows the 'Storage Domain Server' configuration window with the 'CPU Overheads & Polling' tab selected. The window contains several input fields for configuring CPU overheads and polling parameters.

Parameter	Value
SDS number	0
Reference speed (MHz)	550
Average CPU times (ms) for:	
Unproductive poll	0.00266
Read hit	0.073
Read miss	0.073
Write direct	0.087
Write back	0
Mirror primary	0
Mirror secondary	0
Productive poll threshold	0.5

Buttons at the bottom: OK, Cancel, Apply, Help.

SDS model - polling

The screenshot shows the 'Storage Domain Server' configuration window. It has several tabs: 'PCI bus characteristics', 'Host Bus Adapter characteristics', 'More Overheads', 'Physical Disks', 'SDS General', 'CPU Overheads & Polling', and 'More Polling and Overheads'. The 'CPU Overheads & Polling' tab is selected. The 'SDS number' is set to 0. There are two sections for 'Numbers of polls for' and 'Fibre Channel Interface card overheads (ms) for'. The 'Numbers of polls for' section has input fields for Read hit (2), Read miss (5), Write direct (3), Write back (3), Mirror primary (3), and Secondary (3). The 'Fibre Channel Interface card overheads (ms) for' section has input fields for Read hit (0.094), Read miss (0.094), Write direct (0.11), Write back (0), Mirror primary (0), and Secondary (0). There is also a 'PCI bus access' field set to 0.004. At the bottom are buttons for 'OK', 'Cancel', 'Apply', and 'Help'.

Category	Read hit	Read miss	Write direct	Write back	Mirror primary	Secondary
Numbers of polls for	2	5	3	3	3	3
Fibre Channel Interface card overheads (ms) for	0.094	0.094	0.11	0	0	0

PCI bus access: 0.004

SDS model - Host Bus Adapters

The screenshot shows the 'Storage Domain Server' configuration window, specifically the 'Host Bus Adapter characteristics' tab. The window is titled 'Storage Domain Server' and has a standard Windows-style title bar with a close button. The main content area is divided into two sections: 'PCI bus characteristics' and 'Host Bus Adapter characteristics'. The 'Host Bus Adapter characteristics' section contains the following fields:

- SDS number: 0
- HBA number: 0
- Model: iQLA2200/66
- PCI bus number: 0
- Simplex bandwidth (MB/s): 100
- Width (bytes): 8
- Operation on PCI bus (latency in bus cycles):
 - Acquisition latency: 2
 - Initial Target Latency: 8
 - Subsequent transfers: 1
 - End of transfer: 1
 - Burst size (data phases): 32

At the bottom of the window, there are four buttons: 'OK', 'Cancel', 'Apply', and 'Help'.

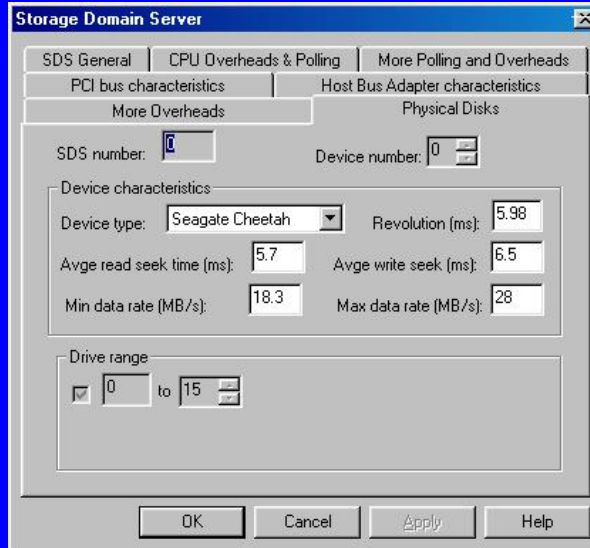
SDS model - PCI bus

The screenshot shows the 'Storage Domain Server' configuration window with the 'PCI bus characteristics' tab selected. The window contains several input fields and buttons for configuring the PCI bus parameters.

Parameter	Value
SDS number	0
PCI bus index	0
Speed (MHz)	66
Width (bytes)	8
Arbitration latency (ns)	10
HBA's	3

Buttons: OK, Cancel, Apply, Help

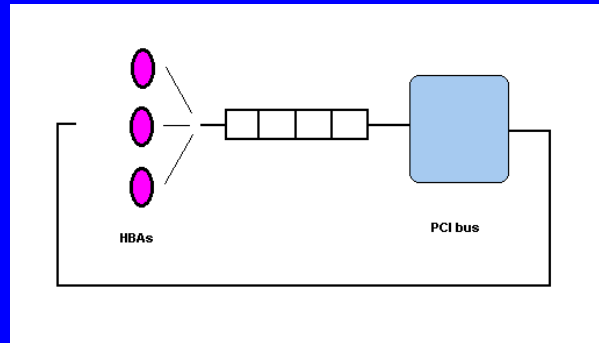
SDS model - physical disks



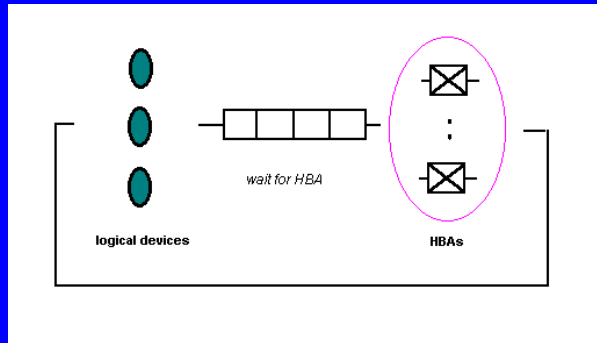
PCI bus model

- ◆ PCI bus model
 - effective data rate, wait for bus
- ◆ Host Bus Adapter
 - traffic fraction, wait for HBA
- ◆ fixed-point iteration between models

PCI bus model - bus



PCI bus model - adapter traffic



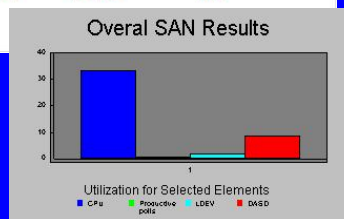
Results

Overall SAN Analysis Results

Summary Totals			
Device	Number	I/Os/sec	I/O time
Logical	32	640.0	0.86
Physical	16	232.0	6.38

Storage Domain Servers		
CPU utiliz.	Productive polls	Poll ratio
33.3	2136	0.59

Averages for Logical and Physical Devices				
Device	Queueing	Service	I/Os/sec	Utiliz.
Logical	0.05	0.81	20.00	1.6
Physical	0.54	5.84	14.50	8.5



Results - client

Client Results for Host 0

Device	Number	I/Os/sec	I/O time
Logical	32	640.0	0.86
Physical	16	232.0	6.38



Client	Queueing	Service	I/Os/sec	Utiliz.
Host 0	0.05	0.81	20.00	1.6

Device	Queueing	Service	I/Os/sec	Utiliz.
Logical	0.05	0.81	20.00	1.6
Physical	0.54	5.84	14.50	8.5



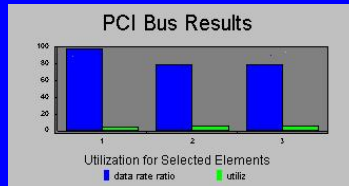
Results - PCI bus

PCI Bus Results for Multiple devices

Device	Summary Totals	I/Os/sec	I/O time
Logical	Number	32	640.0
Physical		16	232.0

Device	PCI Bus	MB/s ms per I/O	I/Os/sec	Utiliz.
PCI0 of A		97.1	0.16	291.41
PCI1 of A		79.3	0.20	290.28
PCI2 of A		79.3	0.20	290.28

Device	Averages for SDS PCI buses	MB/s ms per I/O	I/Os/sec	Utiliz.
PCI of SDS A		85.3	0.19	290.65



Conclusions

- ◆ Model of in-band SAN
- ◆ Logical I/O for clients
- ◆ Storage Domain Server model
 - CPU overheads and polling
 - PCI buses and Host Bus Adapters
 - physical devices
- ◆ Future: leverage Lead Time
 - controllers attached to SDS