
Accelerating Oracle Databases with Solid-State File Caching

Michael Casey

**Solid Data Systems, Inc.
2945 Oakmead Village Court
Santa Clara, CA 95051**

Voice: 408.845.5802

Fax: 408.727.5496

mcasey@soliddata.com

Dimensions of Internet Performance

↑
Transaction
Performance

E-Transactions

- E-mail
- Messaging
- Databases !!

Information Highway

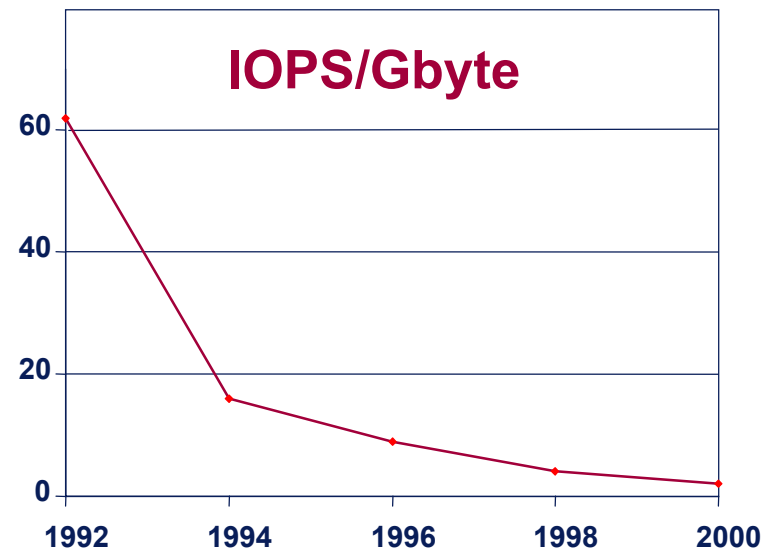
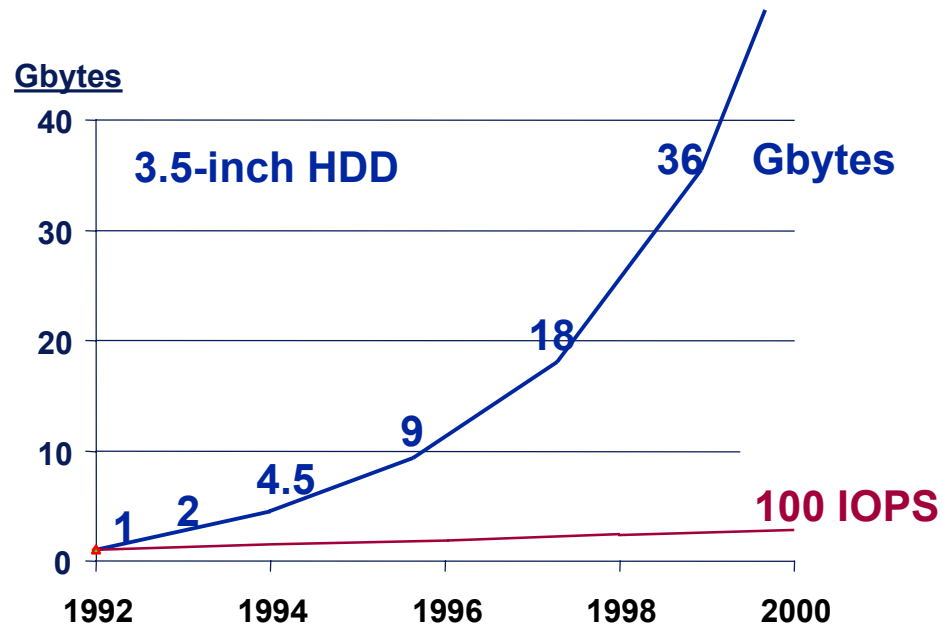
- Text
- Graphics
- Images

Bandwidth →

- Online stock trading
- Online order entry
- Real-time tracking databases
- CRM (Siebel, etc.)
- ERP (Peoplesoft, etc.)

Key Issue: I/O Scalability

Disk drive capacity scales rapidly ...

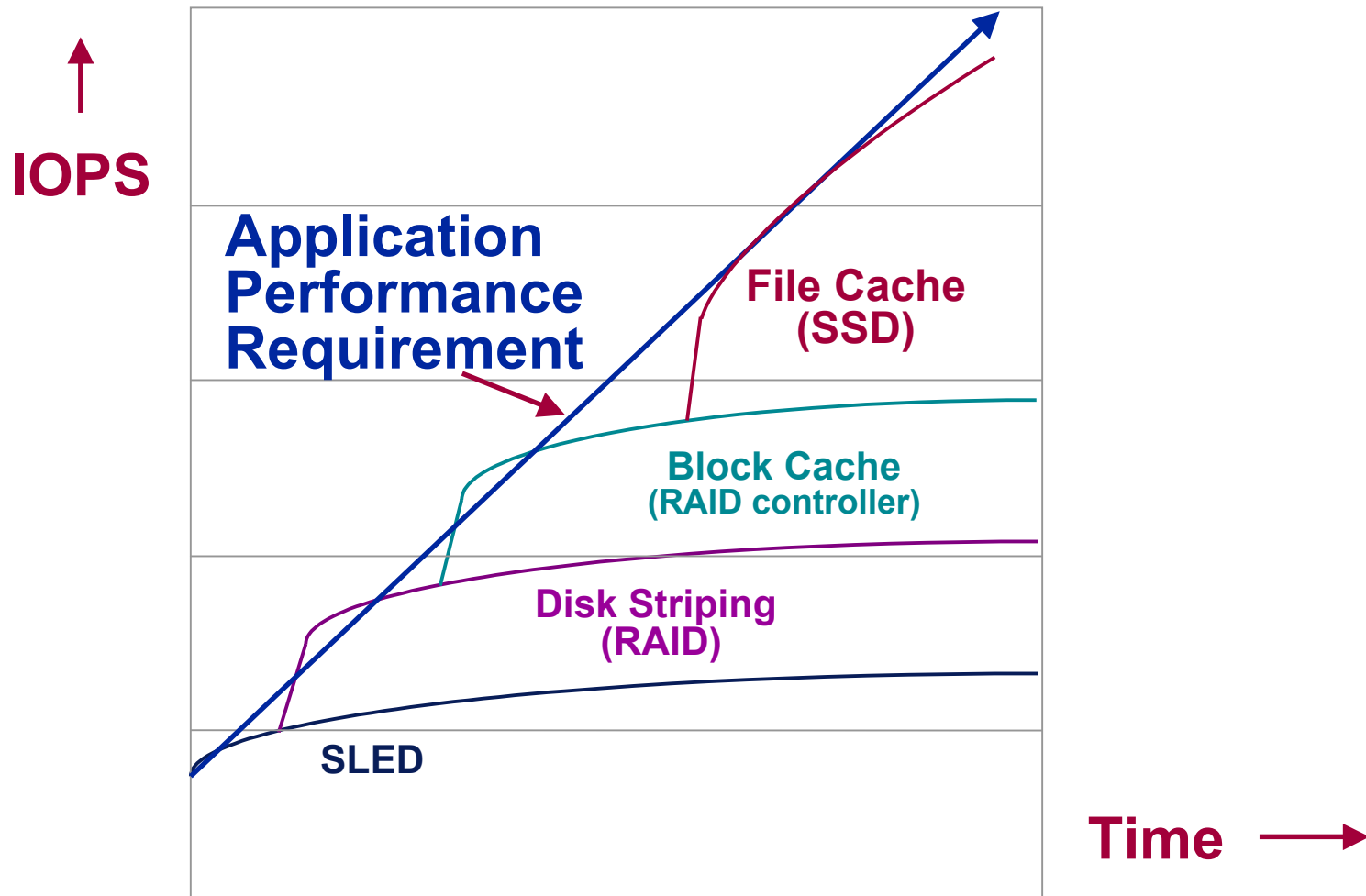


... but performance lags

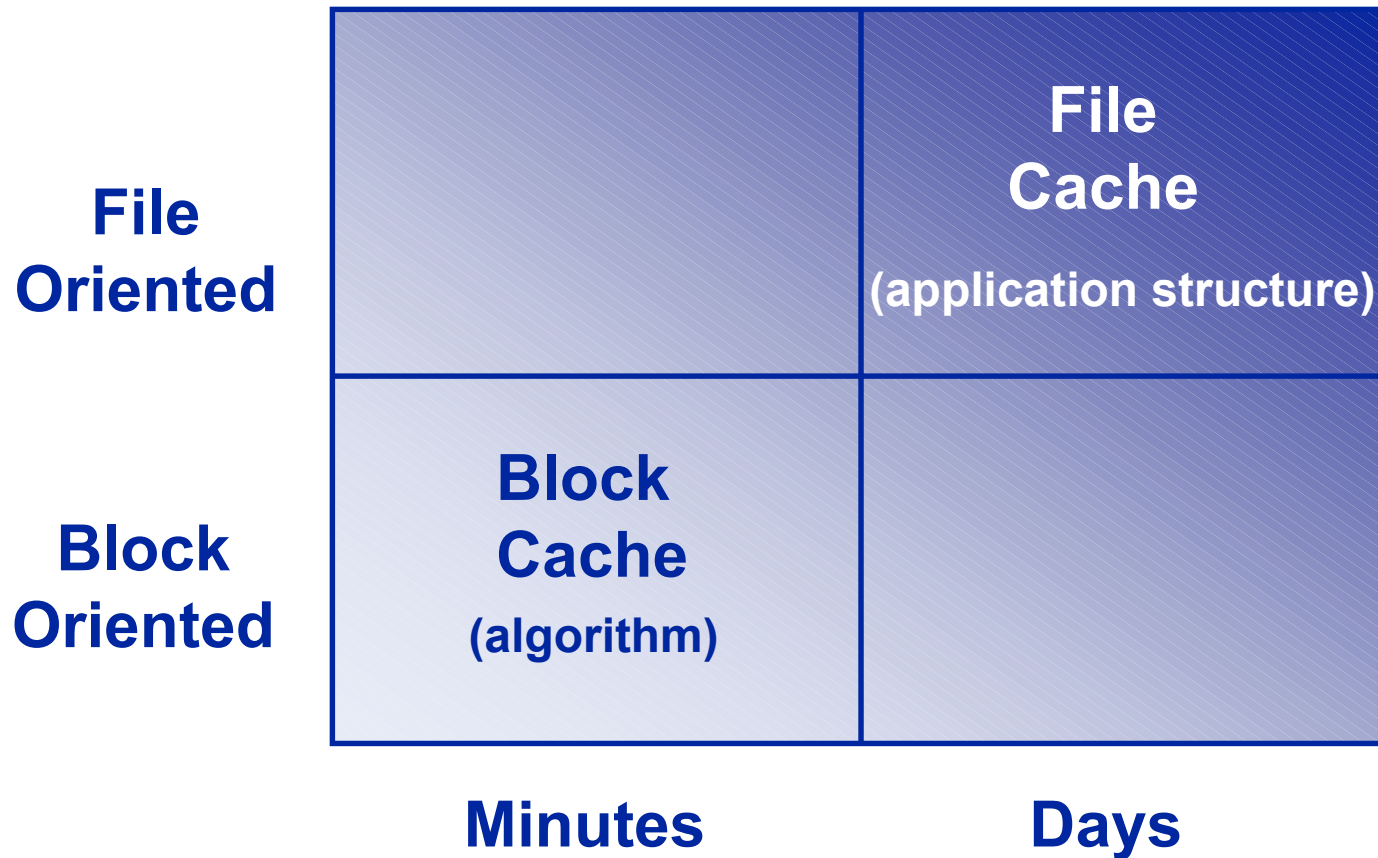
... so access density declines *

* "Access Density - Key to Disk Performance"
Computer Technology Review / Storage Inc., 1999 Q2

Architectural Solutions

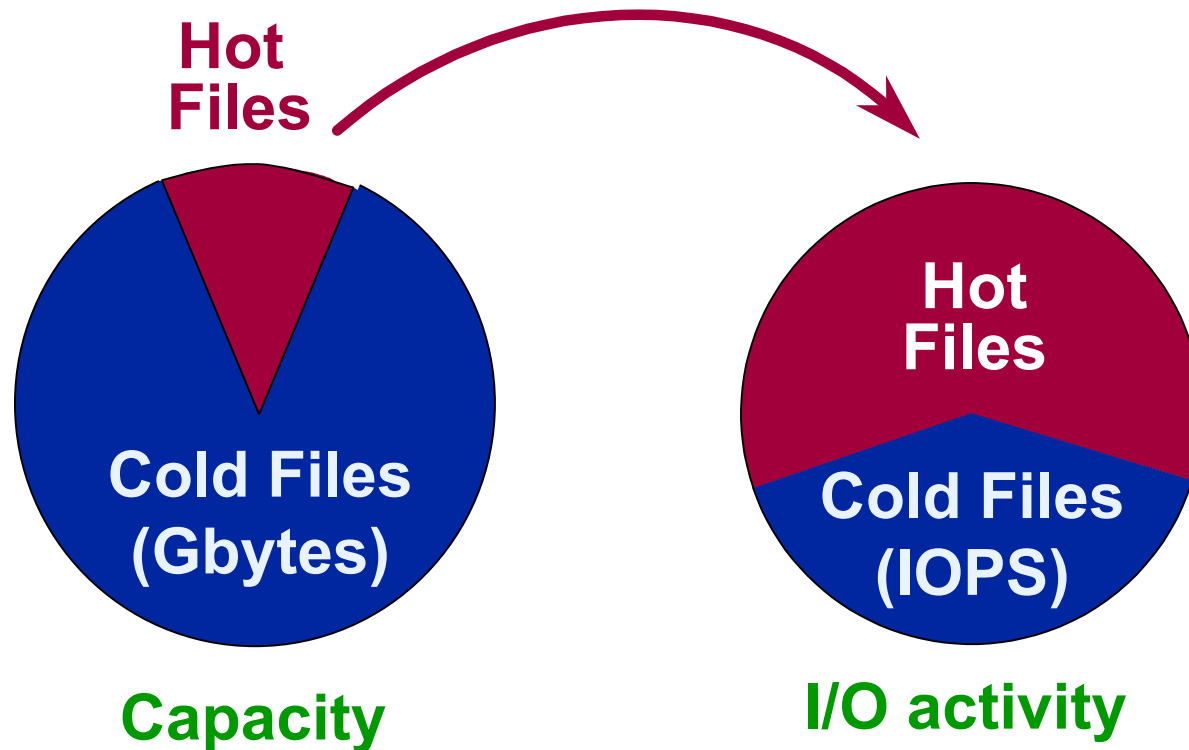


Block Cache and File Cache



Application-Specific I/O Skew

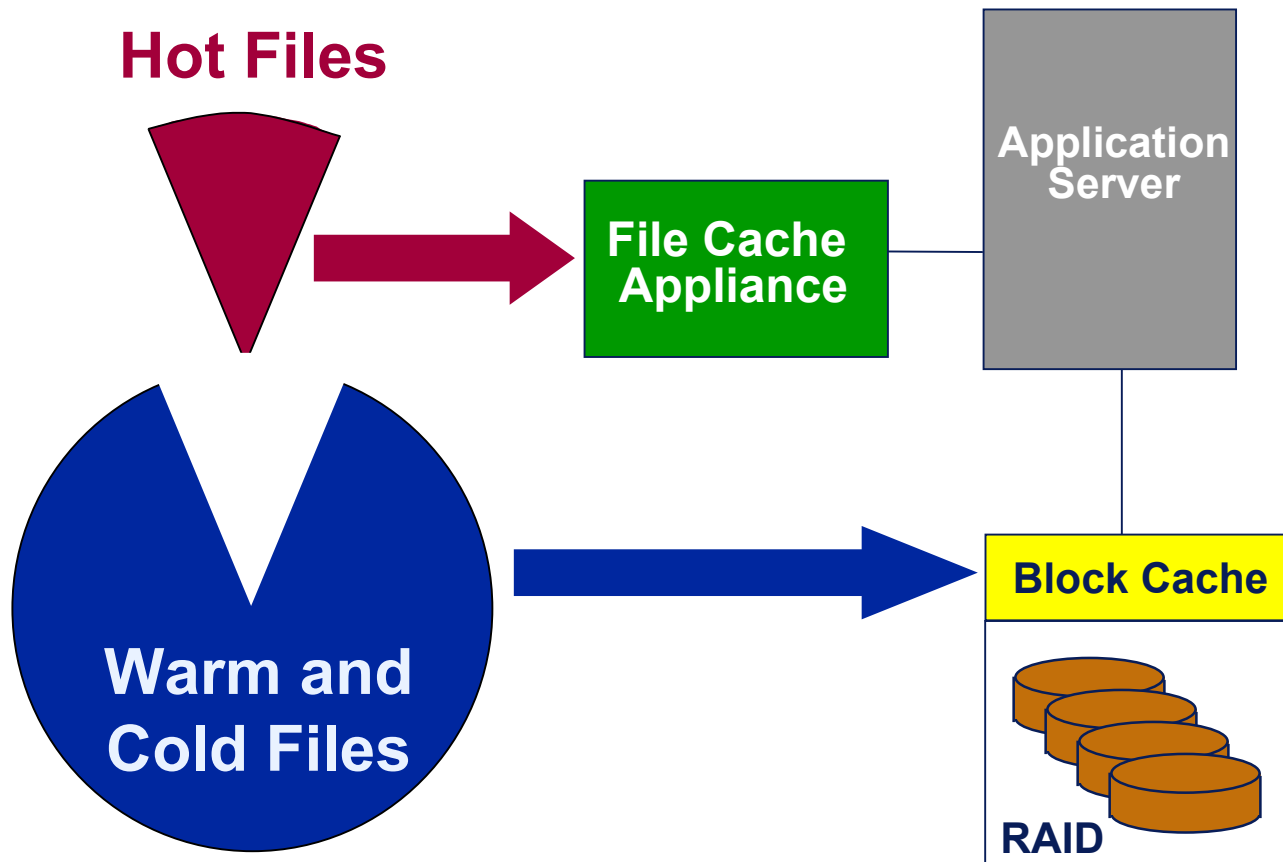
Less than 5% of the data can drive more than 50% of the I/O



Hot-File Examples:

- E-mail
+ Message queues
- Databases
+ Re-do logs
+ Index files
+ Temp spaces
- Messaging
+ Buffer queues

File Cache - a New I/O Architecture



Separate file cache is independently scalable & manageable

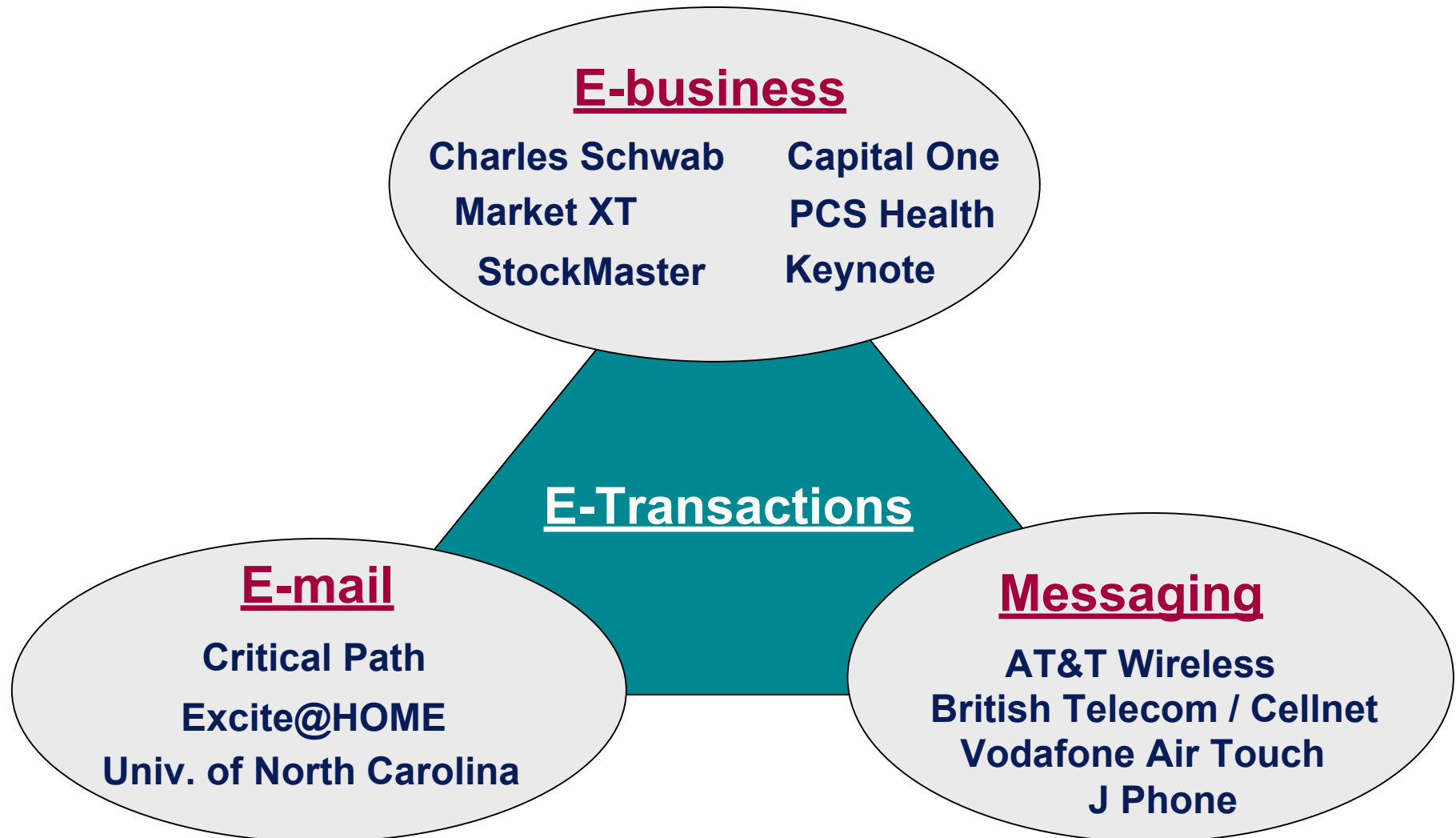
Architectural Adoption

Charles Schwab

File caching architecture replicated in multiple applications:

- ◆ Online Stock Trading
- ◆ Transaction Clearing
- ◆ MQ Series Messaging
- ◆ Siebel CRM Application

Rapidly Growing Applications



When Does File Caching Make Sense?

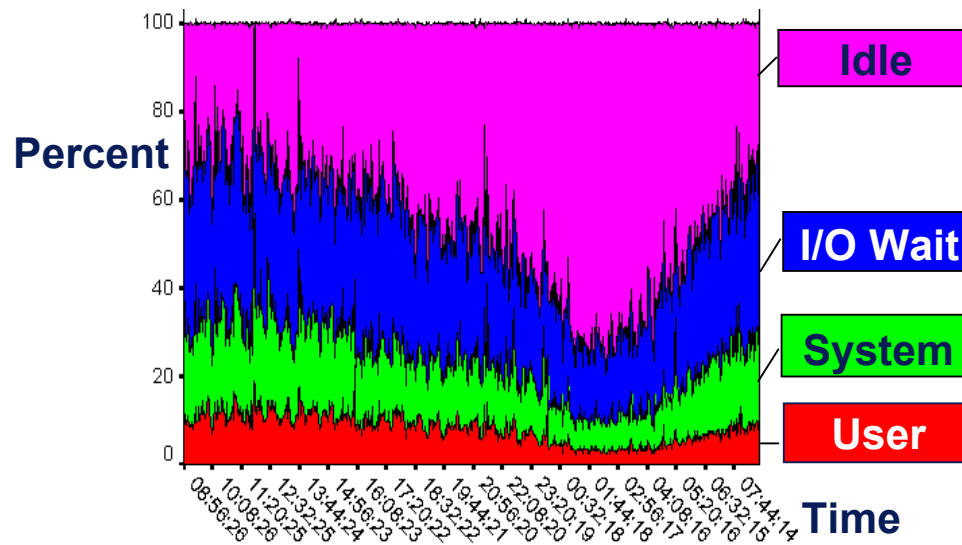
- **The application server is I/O bound**
- **The I/Os are skewed.**

A small percentage of the files
drive a large percentage of the I/O activity.

File Cache: Performance Multiplier

CPU utilization analysis -- E-mail server workload

(a) before file cache

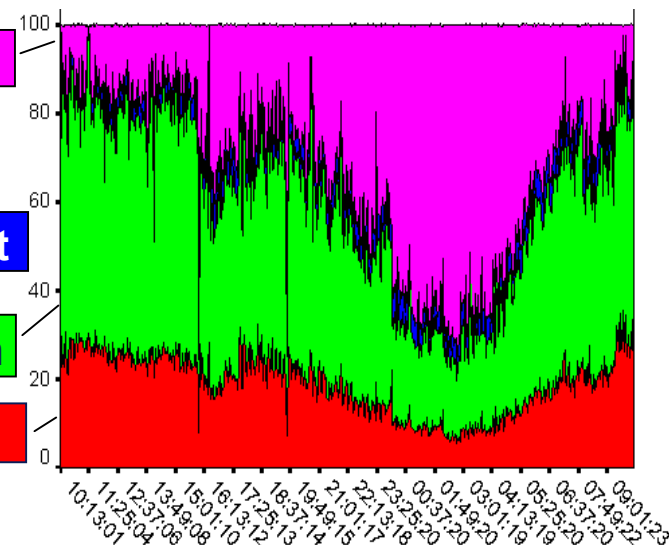


I/O Wait = 40% of CPU time
at peak load

13 messages/sec



(b) with file cache



I/O Wait = largely eliminated

55 messages/sec

Multiplier Effect

Critical Path

- Internet E-mail provider
- Message queues on file cache
- Documented 8x performance boost
- To date: 80 Sun Ultra 2 servers and 40 file cache units

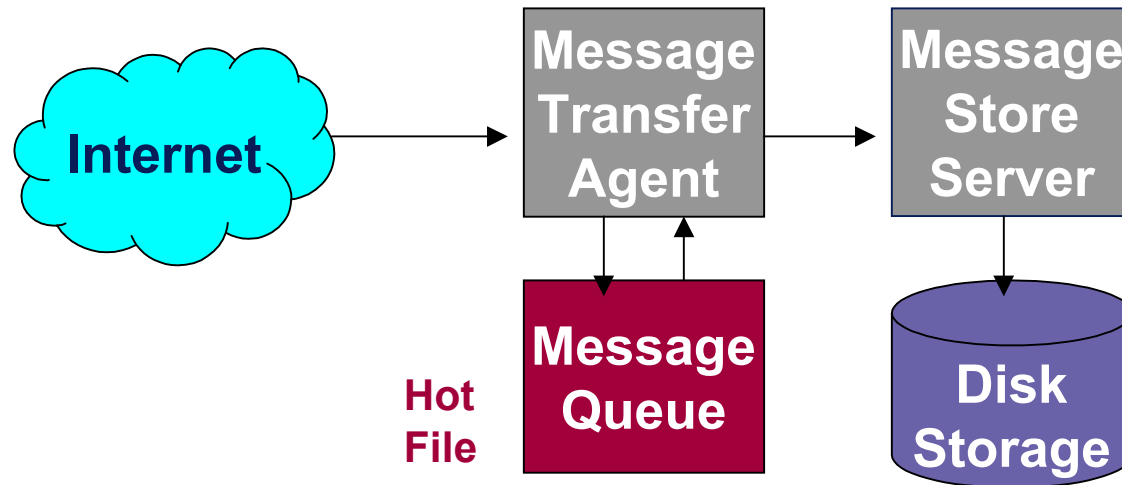
Critical Path's customer base:

3Com, Arthur Anderson, AT&T, Bell Atlantic, Bristol-Myers, Squibb, Chevron, Deloitte & Touche, DirecTV, DuPont, Merck, EDS, E*TRADE, FedEx, Glaxo Wellcome, Hewlett Packard, Hughes Aircraft, ICQ, Ingram Micro, Kraft, KPMG, Lockheed Martin, Lucent, Motorola, NASA, Network Solutions, Pfizer, Pixar, Sprint, StarMedia, Time, Universal, Studios, US West, Warner Bros.

Compelling Value Proposition

Dollars in 000s	Before File-cache	With File-cache	Net Savings
Relative Performance	1.00	4.00	
Required Servers	320	80	
Cost per Server	<u>\$ 70</u>	<u>\$ 80</u>	
Total Investment Cost	\$22,400	\$6,400	\$16,000
Annual Admin Cost per Server	\$ 14	\$ 16	
Annual Admin Costs	<u>\$ 4,480</u>	<u>\$1,280</u>	
3-Year Admin Costs	\$13,440	\$3,840	<u>\$ 9,600</u>
3-Year Life Cycle TCO	\$35,840	\$10,240	\$25,600

E-mail and Wireless Messaging



- ◆ **Message in queue is written 9 times before disposition**
- ◆ **Performance increases 300 - 500% with hot-file caching**
- ◆ **Proven application: replicated as standard architecture**

The most common computer will be a cell phone - Michael Peterson

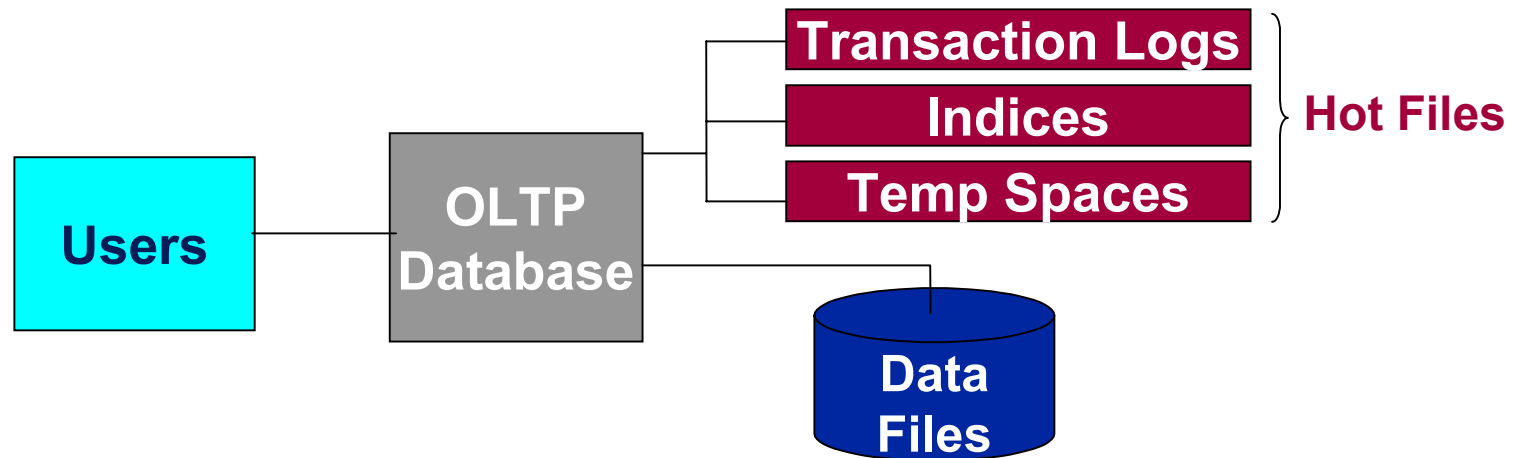
Transaction

- ◆ **Is an application specific unit of work**

- Dynamic Web page generation
- E-mail message
- Credit card electronic verification
- Database insertion
- Etc.

Each transaction may represent *dozens* to *hundreds* of disk I/Os

Transaction Databases



- ◆ Sybase, Oracle, Informix, Microsoft SQL Server
- ◆ Oracle is the 800 pound gorilla
 - ERP: Oracle Financials, SAP, Baan, PeopleSoft
 - CRM: Siebel, Vantive, Clarify, Remedy
- ◆ Key: hot-file identification for ease of use

Symptoms of Potential I/O Bottlenecks

- ◆ Performance drops radically as users are added to a system
- ◆ System no longer able to “keep up”
- ◆ Batch jobs don’t complete within the available processing time window
- ◆ Month-end close takes days instead of hours
- ◆ Extensive tuning has not solved performance issues
- ◆ SAR (Unix) or PerfMon (NT) indicate >50% I/O Wait

Case Study: Oracle Database

Large Cable Operator

Issue: Losing Revenue

Reason: Batch report results were delayed,
month-end close required 3 days,
nightly database refresh took 12 hours.

SSD Impact

Batch processing reduced from overnight to 30 minutes.

Month-end close reduced from 3 days to 3 hours.

Database refresh reduced from 12 to 4 hours.

Case Study: Peoplesoft Database

Solid Data Systems - IO Dynamics for Oracle - [finprd_20000126_050938]

File View Connection Help

Log File: \\Fiona\I-O\OHFM\Log Files\SCgroup\finprd.mdb Log Table: finprd_20000126_050938 Start: 1/26/00 5:11:57 AM End: 1/27/00 9:29:51 PM

Database File	Reads/Sec (Avg)	Writes/Sec (Avg)	Reads/Sec (Peak)	Writes/Sec (Peak)	IO Density (Avg)	File Size (Peak)
/mnt/finprd/dat1/finprd_aplarge01	15.4	0.1	169.3	4.6	3.6	928,186,368
/mnt/finprd/dat2/finprd_gllarge01	11.2	0.1	125.3	8.3	2.7	939,524,096
/mnt/finprd/dat3/finprd_gllarge02	8.0	0.1	120.7	13.2	2.0	939,524,096
/mnt/finprd/dat4/finprd_aplarge04	4.5	0.0	39.4	4.5	1.7	536,870,912
/mnt/finprd/dat2/finprd_aplarge02	4.4	0.0	37.2	2.8	3.4	268,435,456
/mnt/finprd/dat3/finprd_aplarge03	4.3	0.0	43.7	2.9	3.4	268,435,456
/mnt/finprd/dat4/finprd_psindex10	4.1	0.1	50.7	10.3	1.0	1,073,242,112
/mnt/finprd/dat2/finprd_psindex09	3.6	0.1	48.3	14.1	0.9	1,073,717,248
/mnt/finprd/dat3/finprd_psapp02	3.3	0.1	31.0	3.7	1.7	915,136,512
/mnt/finprd/dat1/finprd_psindex01	2.9	0.3	31.4	24.3	0.9	1,073,741,824
/mnt/finprd/dat4/finprd_psindex04	2.9	0.3	23.3	16.5	0.9	1,073,741,824
/mnt/finprd/dat3/finprd_psindex11	2.9	0.3	28.1	27.7	0.9	1,072,381,952
/mnt/finprd/dat3/finprd_psindex07	2.8	0.2	17.4	27.2	0.8	1,073,741,824
/mnt/finprd/dat2/finprd_psapp01	2.7	0.1	51.2	6.2	1.2	848,740,352
/mnt/finprd/dat2/finprd_psindex06	2.7	0.2	17.5	15.6	0.9	1,073,741,824
/mnt/finprd/dat1/finprd_psindex05	2.5	0.2	26.0	25.1	0.8	1,073,741,824
/mnt/finprd/dat2/finprd_psindex02	2.4	0.3	33.5	28.0	0.7	1,073,741,824
/mnt/finprd/dat2/finprd_psindex08	2.4	0.2	39.6	14.4	0.6	1,073,741,824
/mnt/finprd/dat3/finprd_psindex03	2.0	0.3	20.9	20.3	0.6	1,073,741,824
/mnt/finprd/dat1/finprd_amlarge01	1.4	0.0	45.8	0.2	0.6	618,790,912
/mnt/finprd/dat1/finprd_system	1.3	0.0	22.4	0.6	5.1	157,286,400
/mnt/finprd/dat4/finprd_amlarge02	1.3	0.0	23.0	0.1	0.4	618,790,912
/mnt/finprd/dat1/finprd_pttbl01	0.7	0.0	9.5	0.4	0.9	327,827,456
/mnt/finprd/dat1/finprd_fslarge01	0.6	0.4	16.7	43.0	0.4	717,225,984
/mnt/finprd/dat3/finprd_pclarge01	0.5	0.0	26.4	0.6	1.0	136,314,880
/mnt/finprd/tmp2/finprd_temp2	0.5	1.2	22.7	22.4	0.8	2,095,759,360
/mnt/finprd/tmp1/finprd_temp1	0.4	0.4	21.5	22.3	0.2	2,096,390,144
/mnt/finprd/dat4/finprd_scg	0.3	0.0	9.6	0.2	0.6	209,715,200
/mnt/finprd/dat1/finprd_rollback01	0.1	0.8	13.4	89.3	0.1	1,610,612,736
/mnt/finprd/dat3/finprd_rollback02	0.1	1.7	14.9	127.4	0.2	1,610,612,736
/mnt/finprd/dat4/finprd_psimage	0.1	0.0	14.8	0.0	0.4	77,373,440
/mnt/finprd/dat4/finprd_rollback03	0.1	0.5	12.7	51.4	0.4	356,515,840
/mnt/finprd/dat3/finprd_tools	0.0	0.0	12.4	0.1	0.4	225,607,680
/mnt/finprd/dat2/finprd_users	0.0	0.0	0.1	0.0	0.0	15,728,640
/mnt/finprd/dat1/finprd_arlarge01	0.0	0.0	0.8	0.6	0.5	20,848,640

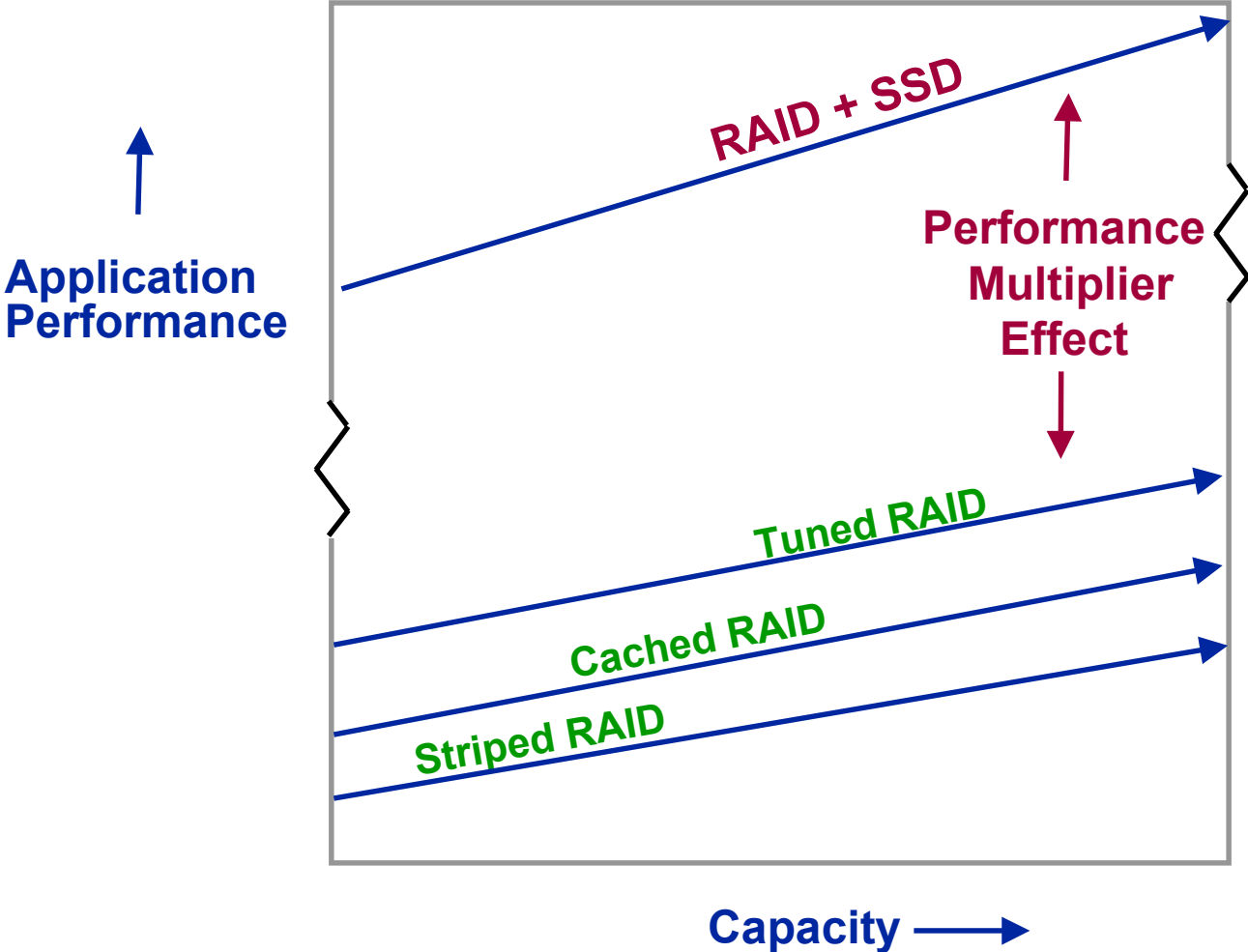
Office

Start Office Distiller Vi for W... Comma... Calend... Explorin... Acrobat... Microso... Solid ... 5:42 PM

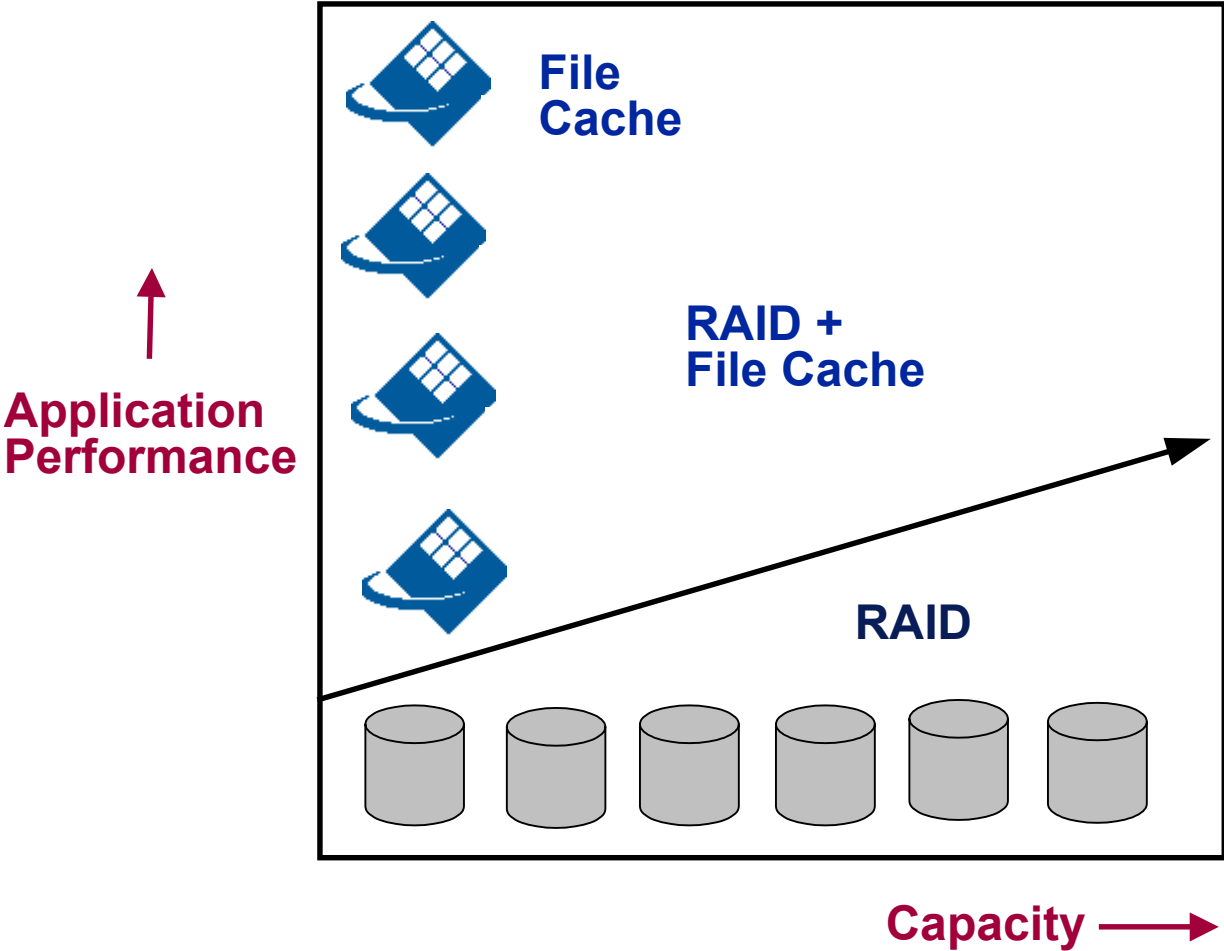
Free download:
http://www.soliddata.com/products/iodynamics_oracle.html

View file and object-level I/O data with free software tools (I/O Dynamics for Oracle) or commercial packages -- e.g., Quest Software, Precise Software, BMC, Oracle ...

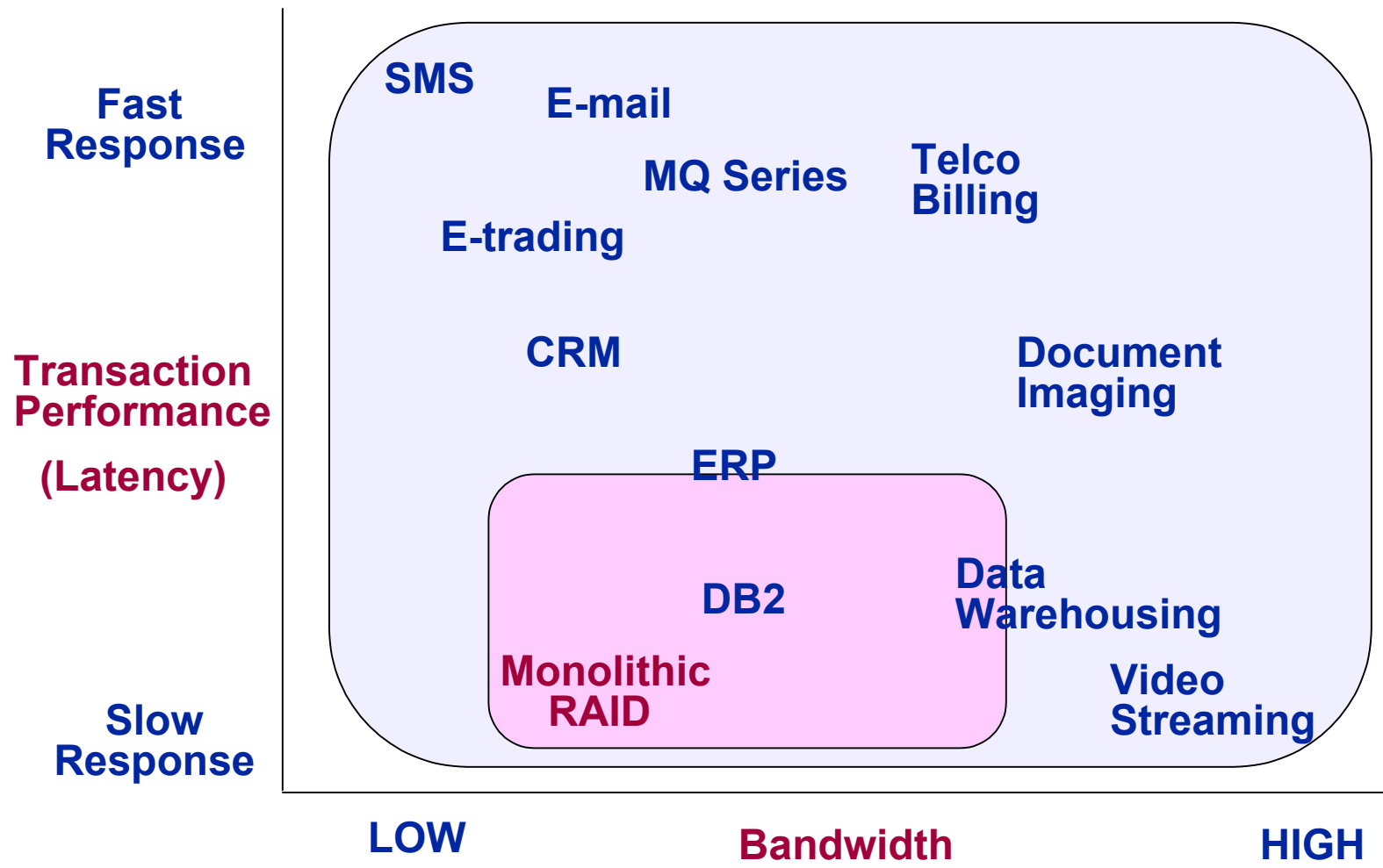
E-Transaction Performance Architecture



2-Dimensional Modular Scalability



Application I/O Requirements

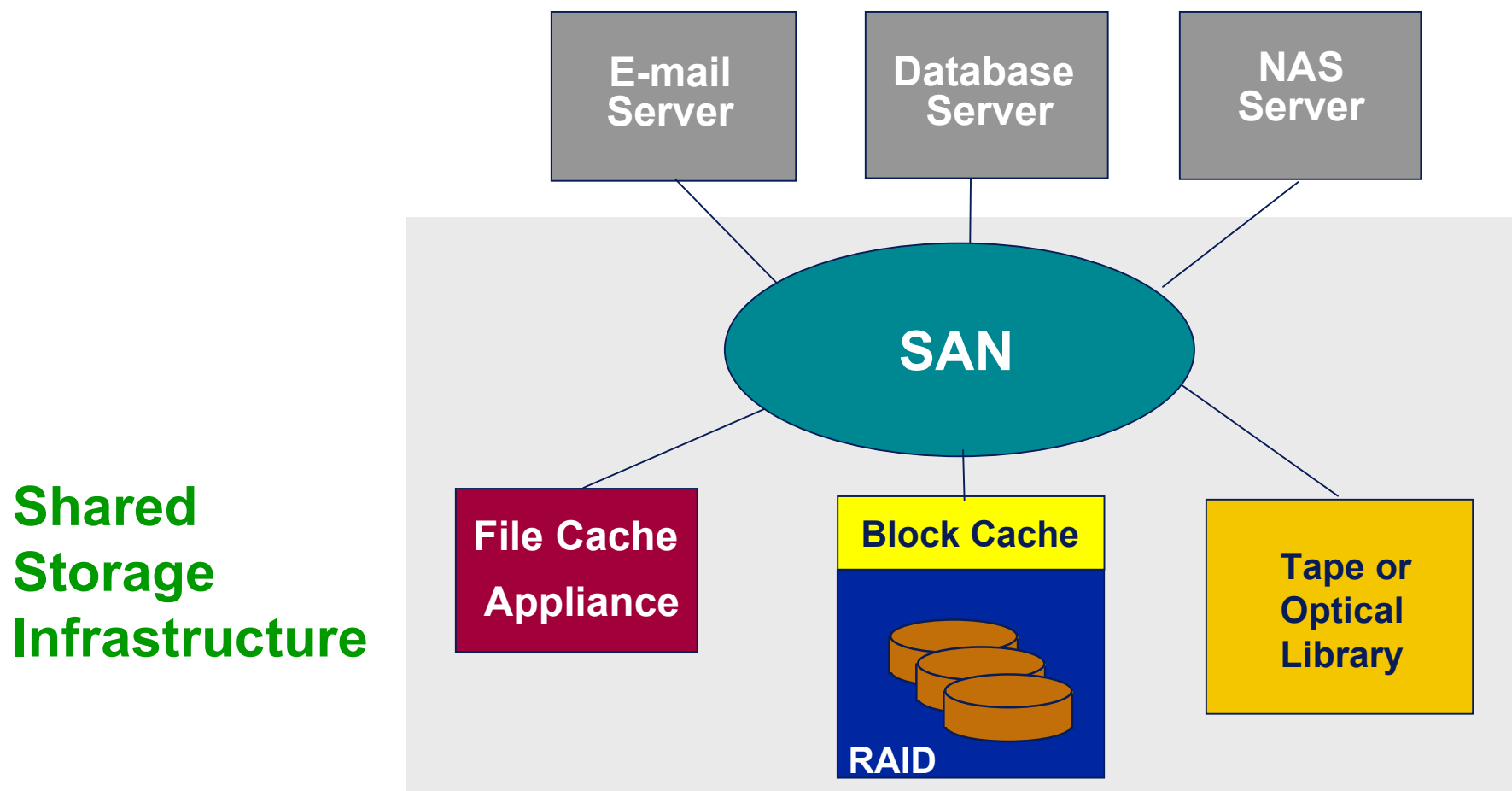


Ease of Use

PCS Health Systems

- Application: Online pharmacy services for 100+ insurance companies and HMOs
- Problem: Continual tuning effort and \$\$\$ required to meet SLA
- Solution: Placed database temp spaces on solid-state file cache
- Result: 10x response time improvement and reduced management effort

File Cache for the Managed SAN



File Cache as Shared Infrastructure

SHARED

- **Affordable in a wider range of applications**

INFRASTRUCTURE

- **Extends scalability and functionality of servers & SANs**
 - **SAN file systems and storage services**
 - **Address lookup tables in virtual storage architectures**
 - **Write-intensive storage management functions (e.g., snapshot copy)**

Conclusions

- Growing applications require scalable transaction performance
- Bandwidth investments will uncover transaction processing as the next e-infrastructure bottleneck
- Access density of mechanical disk keeps falling, and servers need scalable architectural solutions
- Proven applications are accelerating architectural adoption of solid-state file caching
- New software tools are making file cache easier to use
- New SAN infrastructures and virtual storage software will facilitate wider use of file caching as an architectural solution

Questions

For a copy of this presentation with full speaker's notes or answers to other questions please send e-mail request to:

mcasey@soliddata.com