international computing centre

International Computing Centre

Welcome

Summary of Performance Tests with Solid State Disks

IT Solutions for the UN Family

Today's Agenda

- Introduction to ICC
- Brief Considerations for I/O Performance Testing
- Description of Test Plan
- Test Results
- Conclusions

Introduction to ICC

IT Solutions for the UN Family

Intro – ICC History

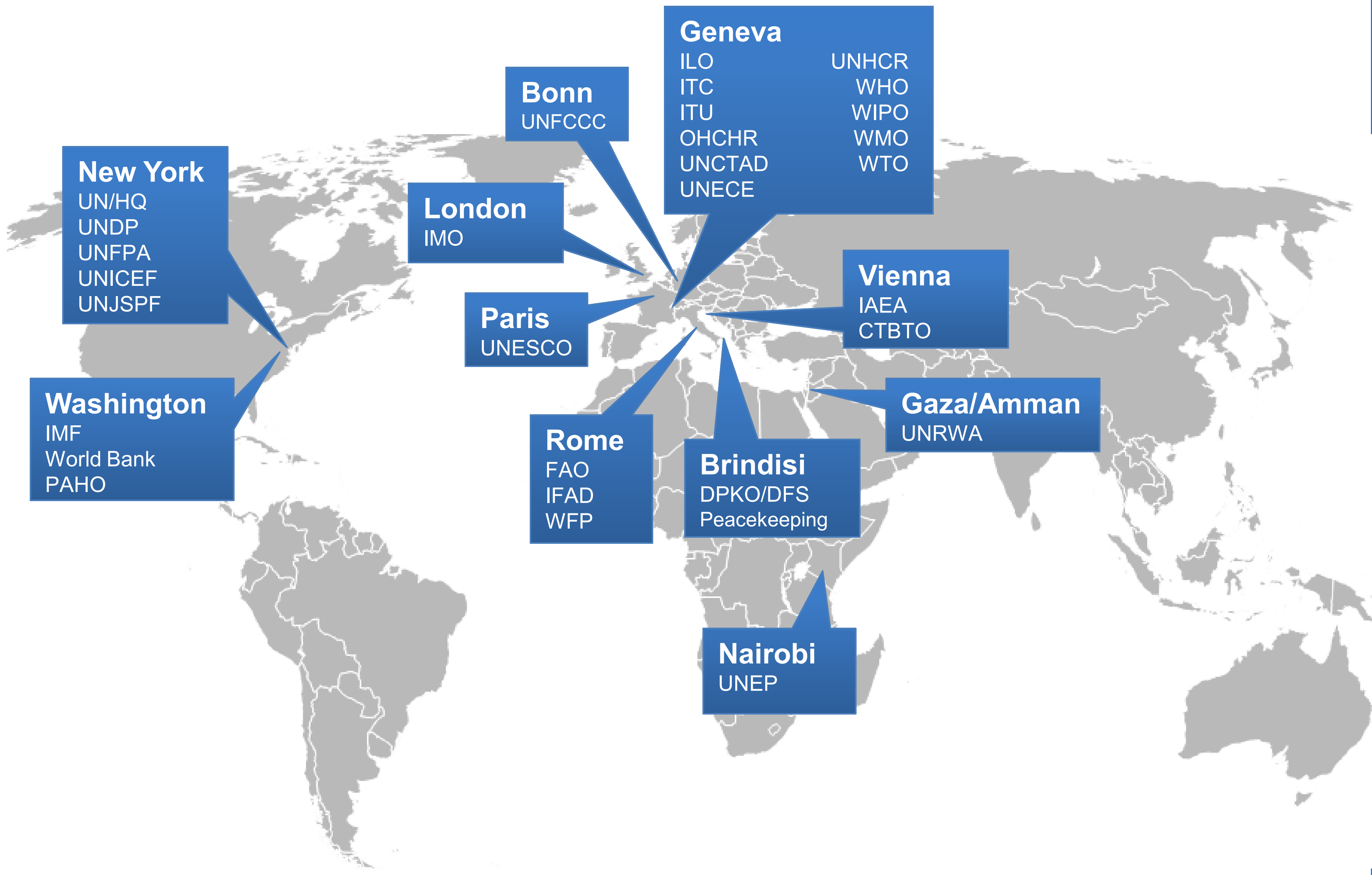
- Created by a resolution of the U.N. General Assembly in 1970
- Initially, ICC provided services to UN, UNDP, WHO for:
  - Mainframe Processing
  - Basic Data Communications

Intro – ICC Today

- Largest ICT Services entity in the UN System
- Deals with a broad spectrum of ICT services
- More than 35 UN entities using ICC services

- Currently, ICC offices (and Data Centers) are located in:
  - Geneva, Switzerland
  - New York, USA
  - Rome & Brindisi, Italy

Intro – ICC Clients



The largest ICT Service Provider in the United Nations System

Intro – Principal Services

- Enterprise Application Hosting
- Mainframe Services
- Internet, Messaging and Networking Services
- Managed Server Hosting
- Managed Storage Hosting
- Training Services
- IT Advisory Services

ICC Managed Data Centers

- **Hosting**
  - > 3,000 Servers/Appliances managed/hosted
  - > 350 Racks
  - > 2.5 PB of Storage and Backup
  - > 13,000 Network and SAN Ports
- **Shared Data Centers**
  Geneva and New York
  - Tier III Facilities
  - Power Utilization Effectiveness < 1.7
  - Metropolitan Area Network HUBs
  - Secure and under UN Jurisdiction
- **Managed On-Site Data Centers**
  - UNLB/DPKO (Brindisi)
  - UN (New York)
  - WFP (Rome)
  - PAHO (Washington DC)

IT Solutions for the UN Family

Oracle @ ICC

- ERPs*
  - EBusiness Suite – 3 orgs
  - PeopleSoft – 3 orgs
  - Siebel – 2 org (not managed hosting)
  - SAP – 1 org (runs on Oracle DBs)

- Key technologies
  - EBusiness, PeopleSoft, Siebel, SAP
  - 10g, 11g databases (10.2.0.4 and 11.2.0.1)
  - Oracle RAC + ASM
  - Grid Control + management packs
  - Solaris and Oracle Enterprise Linux

** ICC also manages Oracle DB instances supporting custom apps.*

International Computing Centre

ICC's ERP Spectrum

- Approx 90 Oracle instances for PeopleSoft
- Largest is 1.7 TB
- 48 CPU-core Sun Solaris server

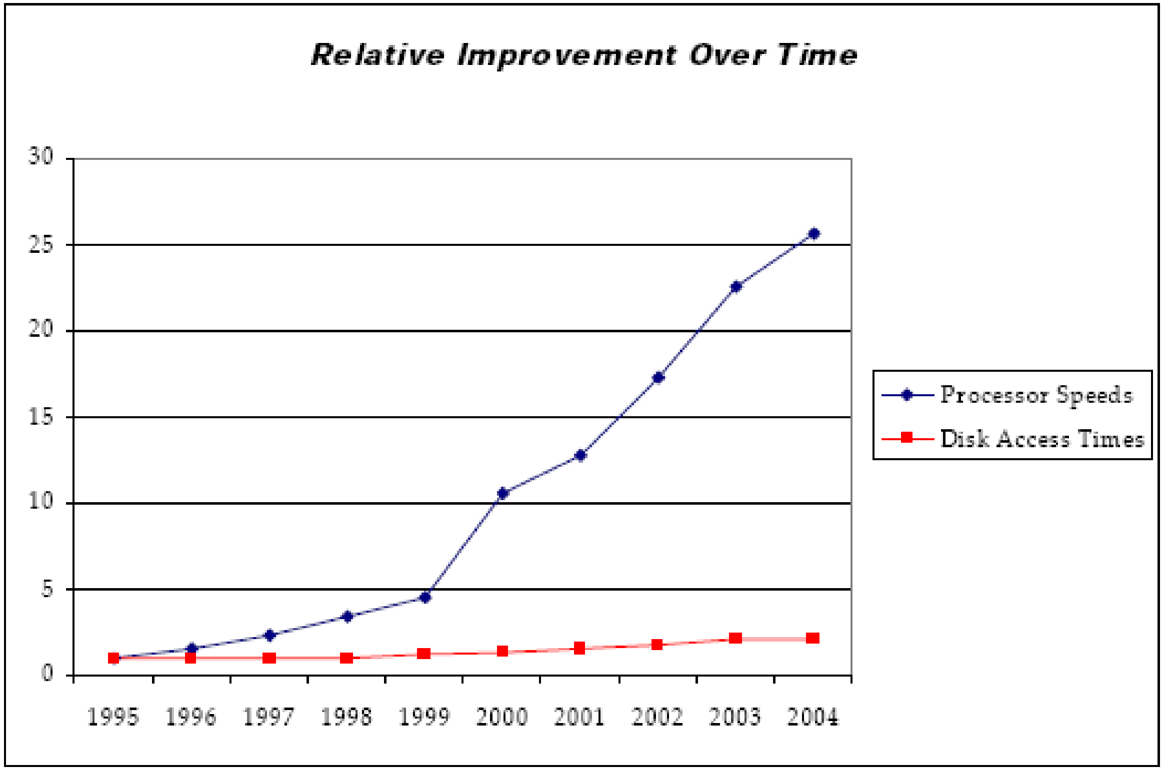
- Approx 30 Oracle instances for EBusiness
- Largest is 550 GB
- 32 CPU-core Oracle RAC (4 nodes)

Looking at Solid State Disks

IT Solutions for the UN Family

Testing Objectives

To measure differences in performance for GSM and Atlas-specific work when traditional disk storage is replaced by solid state disk (SSD).



IT Solutions for the UN Family

What We Tested

GSM (World Health Organization)

- EBusiness 11i on an Intel/OEL 5.x Platform
  - DB on 4-node RAC: Intel and OEL 5.x (Intel x5500 series)
  - Web/App on 4 node shared APPL_TOP (Intel x5300 series)
- RamSan 620 – 2 TB capacity

Atlas (UN Development Programme)

- PeopleSoft Fin/SCM 8.9, PT 8.48 on Solaris
  - DB on single instance on SPARC VII (48 cores, 96 threads)
  - App tier 4-nodes SPARC T2
  - Web tier 4-nodes SPARC T2
- RamSan 620 – 2 TB
- Sun F5100 – 2 TB

International Computing Centre

Principal Characteristics of Tested SSDs

RamSan 620 from Texas Memory Systems

- Up to 5 TB SLC NAND flash
- Multi-host-attach via 4 Gbps
- 2 RU chassis
- Low power (230 W)

Sun 9940V (HDS)

- 4 trays with 8 disks/tray - GSM
- 9 trays with 8 disks/tray - Atlas
- HW RAID 5
- ASM-disks, UFS slices allocated vertically across trays

Considerations

Fully simulating Prod is an expensive challenge

- Test vs Production environments:
  - Concurrency (random I/O) is hard to simulate.
  - Large jobs can be run in Test, but what *else* is running?

- Instead of attempting to predict how our ERPs would perform with SSD, we compared transactions on like-systems.

How *Fast* is the Disk?

- For disk the basic question is: "how fast can I access my data?"
- Access Time ~ seek time + latency
- Seek time is the time for the heads to identify and move to the right area on the platter.
- Latency is the time waited for the platter to spin to the correct spot [of data extraction or insertion].
- For fastest rotational disks, access time during low activity is ~5 ms. However, IO queues significantly impact this.
- For SSD's at peak times, we measured about 4 ms.

international computing centre

ICC

IT Solutions for the UN Family

Average Service Time in GSM Prod

```
Time: 10:22:48 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00  7.80 99.60      29.88     0.95    0.44   4.74
dm-4              0.00     0.00  8.80 69.40      33.55     0.43    0.73   5.70
dm-23             0.00     0.00  5.00 88.60      29.17     1.00    0.60   5.60
dm-27             0.00     0.00  4.80 88.60      29.31     0.47    0.57   5.28

Time: 10:22:53 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00 19.40 44.60      39.35     0.88    2.45  15.70
dm-4              0.00     0.00 21.40 52.60      37.74     0.52    2.20  16.30
dm-23             0.00     0.00 21.80 72.20      30.43     1.52    2.03  19.12
dm-27             0.00     0.00 21.60 54.20      32.93     0.38    2.05  15.56

Time: 10:22:58 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00  7.40 50.00      45.71     0.30    1.26   7.24
dm-4              0.00     0.00  7.60 56.60      38.05     0.48    0.86   5.50
dm-23             0.00     0.00  8.20 56.40      40.34     0.38    0.98   6.32
dm-27             0.00     0.00  7.60 72.40      33.52     0.43    1.04   8.36

Time: 10:23:03 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00  8.00 99.20      29.21     1.21    0.90   9.66
dm-4              0.00     0.00  7.20 87.00      33.32     0.70    0.55   5.22
dm-23             0.00     0.00  8.80 101.60      29.06     1.09    0.86   9.50
dm-27             0.00     0.00 11.00 84.00      28.56     0.44    1.00   9.48

Time: 10:23:08 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00  9.20  6.40      85.85     0.07    4.03   6.28
dm-4              0.00     0.00 11.60  8.20      69.33     0.10    4.78   9.46
dm-23             0.00     0.00  7.00  7.00      96.23     0.05    3.46   4.84
dm-27             0.00     0.00  8.00  7.20      83.01     0.08    5.24   7.96

Time: 10:23:13 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00  7.40 46.60      38.87     0.58    1.31   7.08
dm-4              0.00     0.00 12.80 26.40      50.86     0.31    2.25   8.82
dm-23             0.00     0.00  8.00 47.00      36.19     0.47    1.25   6.88
dm-27             0.00     0.00  6.40 66.20      37.05     0.37    0.77   5.56

Time: 10:23:18 AM
Device:         rrqm/s   wrqm/s   r/s   w/s   avgrq-sz avgqu-sz   svctm  %util
dm-1              0.00     0.00 28.80 134.60      31.59     2.29    1.58  25.84
dm-4              0.00     0.00 40.00 148.20      29.99     1.58    1.52  28.64
dm-23             0.00     0.00 31.80 132.80      31.18     2.31    1.62  26.68
dm-27             0.00     0.00 36.00 153.20      27.65     1.27    1.53  28.90
```

International Computing Centre

What *Throughput* Can the Disk Sustain?

- Just because a disk can get to the data very quickly, does not automatically imply that the disk can read/write large *volumes* of data.
- Throughput depends on a lot of things, but eliminating common factors such as interface to disk and OS config, it is a function of the "type of spindle".
- Typical values for SATA and SAS disks are 100 to 200 MB/s (for sequential I/O).
- SSD values are typically >600 MB/s (for random! I/O).

Comparing cherries to lemons here, but bottom line is: cherries are sweeter.

Benefits of SSD vs Traditional Disk

- SSD have:
  - About 3 times the read throughput of disk.
  - Since they are non-rotational, there is little contention in an OLTP-type environment.
  - About 10 times lower latency.

- There are no performance down-sides to SSD, however its cost is relatively high.

The GSM Test Plan

Testing Philosophy

The idea was to remain close to reality in order to make the tests as relevant as possible:

- Databases were created on same HW with same config parameters.
- Same web/app tier was used.
- Tests run for the following areas:
  - Standard reports, queries and batch jobs.
  - Standard scripts such as post-refresh.
  - Administration jobs such as backups, stats-gathering.

Test Infrastructure

- Databases were set up with identical sizing configuration on same 4-node RAC:
  - GSMSSD used SSD for all storage: data, index, redos, FRA
  - PAYUAT was set up on all traditional disk (15K RPM SAS).
- Connectivity between DB nodes and storage is via same Fibre Channel switches, using same 4 Gbps HBAs.
- Traditional disk was protected via HW-implemented RAID 5.
- SSD was twice-protected* via HW RAID 5.
- No additional ASM redundancy was configured in either case.

* See RamSan dual-protection design.

IT Solutions for the UN Family

Tests run in both GSMSSD & PAYUAT

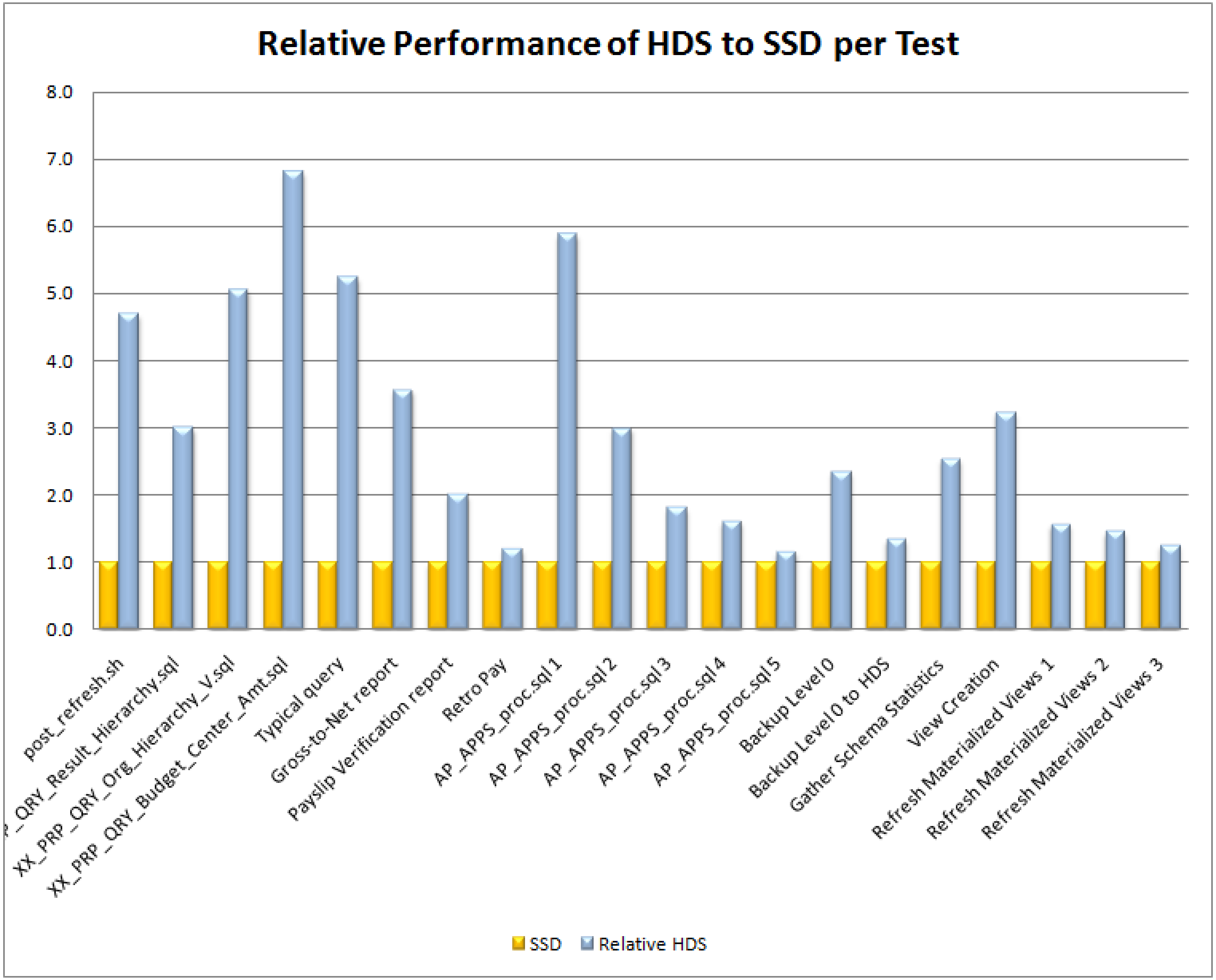
- Application workload:
  - post_refresh.sh
  - Reporting-type script
  - Typical query
  - WHO Gross-to-Net Report
  - WHO Payslip Verification Report
  - Retro-Pay
  - OSW reports
- Administrative workload:
  - Backup/restore
  - Gather schema stats
  - Create views
  - Refresh for materialized views

International Computing Centre

Results

No.	Test	SSD [sec]	NonSSD [sec]	SSD Faster
1	post_refresh.sh	197	925	470%
2	XX_PRP_QRY_Result_Hierarchy.sql	1	3	300%
3	XX_PRP_QRY_Org_Hierarchy_V.sql	1.6	8.1	506%
4	XX_PRP_QRY_Budget_Center_Amt.sql	162	1104	681%
6	Typical query	37	194	524%
7	Gross-to-Net report	152	542	357%
8	Payslip Verification report	60	120	200%
9	Retro Pay	66655	79579	119%
10	AP_APPS_proc.sql 1	17	100	588%
11	AP_APPS_proc.sql 2	1866	5592	300%
12	AP_APPS_proc.sql 3	1252	2268	181%
13	AP_APPS_proc.sql 4	8889	14278	161%
14	AP_APPS_proc.sql 5	507	590	116%
15	Backup Level 0	2026	4755	235%
16	Backup Level 0 to HDS	2100	2820	134%
17	Gather Schema Statistics	50	127	254%
18	View Creation	2612	8448	323%
19	Refresh Materialized Views 1	151	237	157%
20	Refresh Materialized Views 2	191	278	146%
21	Refresh Materialized Views 3	23	29	126%

Results


Relative Performance of HDS to SSD per Test
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
post_refresh.sh
_QRY_Result_Hierarchy.sql
XX_PRP_QRY_Org_Hierarchy_V.sql
XX_PRP_QRY_Budget_Center_Amt.sql
Typical query
Gross-to-Net report
Payslip Verification report
Retro Pay
AP_APPS_proc.sql 1
AP_APPS_proc.sql 2
AP_APPS_proc.sql 3
AP_APPS_proc.sql 4
AP_APPS_proc.sql 5
Backup Level 0
Backup Level 0 to HDS
Gather Schema Statistics
View Creation
Refresh Materialized Views 1
Refresh Materialized Views 2
Refresh Materialized Views 3
SSD
Relative HDS

ICC international computing centre

IT Solutions for the UN Family

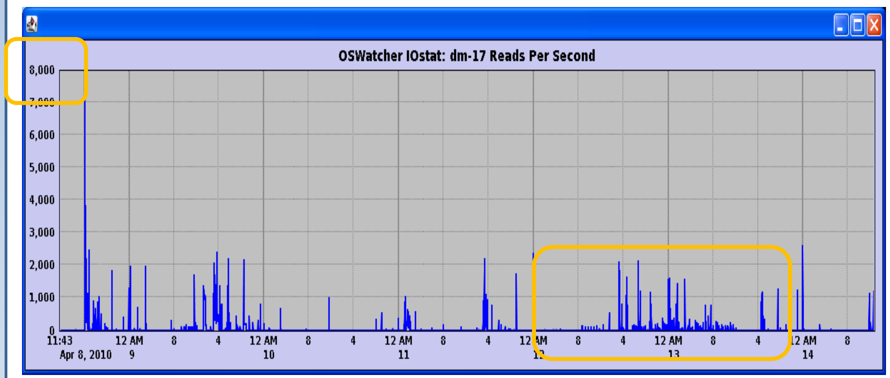
Disk Performance from OS Watcher

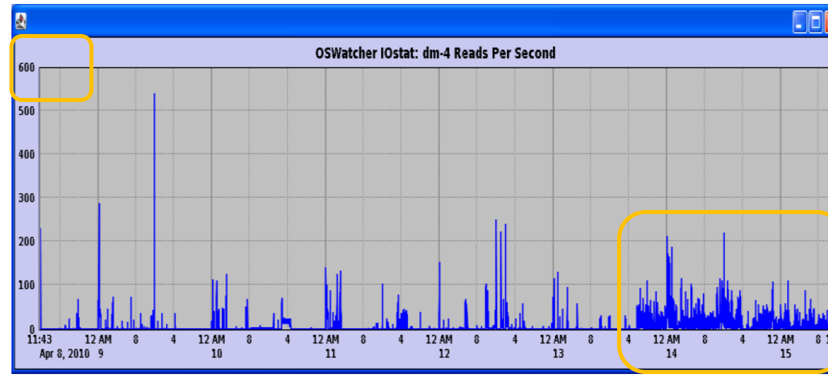
International Computing Centre

SSD

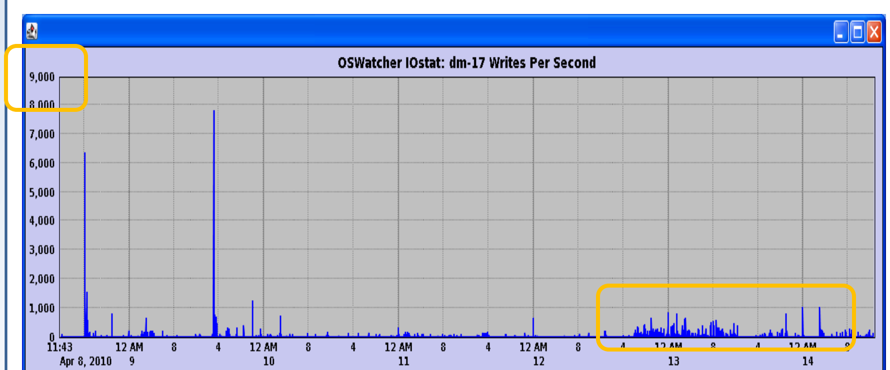
HDS

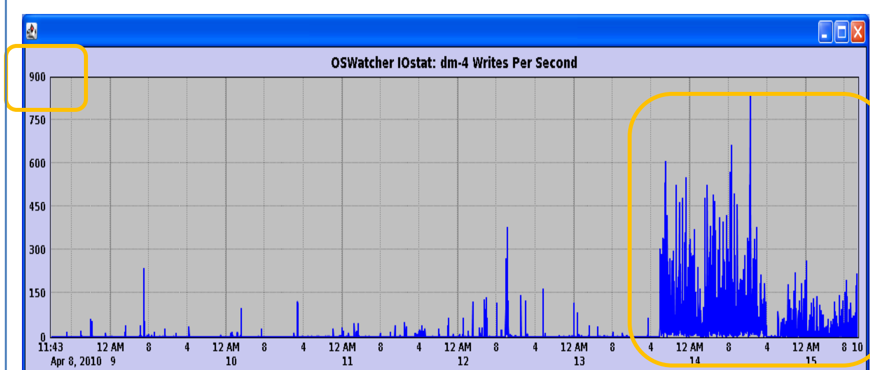
READS





WRITES

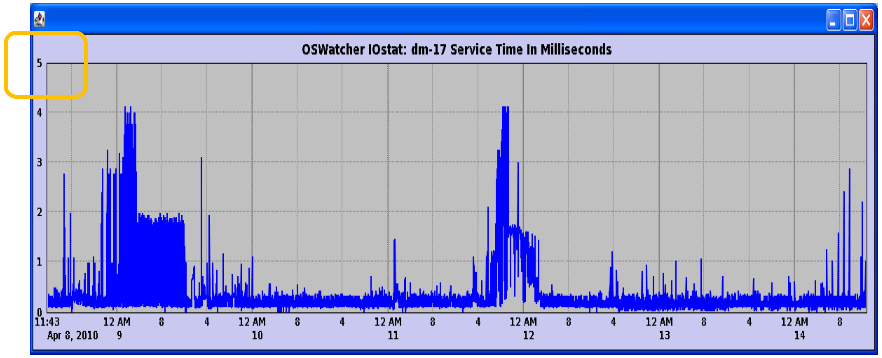




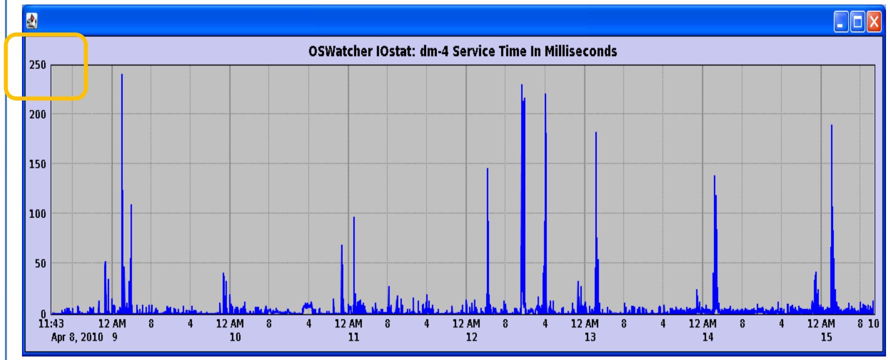
OS Watcher cont'd

SSD

Service Time



HDS



ICC international computing centre

IT Solutions for the UN Family

Database Wait Events for Payroll Run

International Computing Centre

SSD

Wait Class

- s - second
- cs - centisecond - 100th of a second
- ms - millisecond - 1000th of a second
- us - microsecond - 1000000th of a second
- ordered by wait time desc, waits desc

Wait Class	Waits	%Time -outs	Total Wait Time (s)	Avg wait (ms)	Waits /txn
Cluster	44,764,330	0.00	19,756	0	16.78
System I/O	3,922,831	0.00	2,938	1	1.47
Other	13,380,071	67.29	2,596	0	5.02
User I/O	2,767,876	0.00	1,120	0	1.04
Concurrency	134,993	0.66	105	1	0.05
Network	14,837,778	0.00	81	0	5.56
Commit	8,090	0.00	15	2	0.00
Application	28,146	0.01	12	0	0.01
Configuration	200,069	93.85	5	0	0.07

- Total I/O wait time:
  - System I/O 2,938 s
  - User I/O 1,120
  - **Total = 4,850 s**

HDS

Wait Class

- s - second
- cs - centisecond - 100th of a second
- ms - millisecond - 1000th of a second
- us - microsecond - 1000000th of a second
- ordered by wait time desc, waits desc

Wait Class	Waits	%Time -outs	Total Wait Time (s)	Avg wait (ms)	Waits /txn
System I/O	5,552,600	0.00	117,748	21	28.90
Cluster	23,465,974	0.01	14,363	1	122.13
User I/O	969,769	0.00	8,907	9	5.05
Other	7,141,540	58.93	6,433	1	37.17
Concurrency	144,239	1.35	171	1	0.75
Commit	12,951	0.24	144	11	0.07
Network	7,325,200	0.00	41	0	38.13
Configuration	106,556	97.00	30	0	0.55
Application	64,720	0.00	19	0	0.34
Administrative	136	0.00	15	108	0.00

- Total I/O wait time:
  - System I/O 117,748 s
  - User I/O 8,907
  - **Total = 126,655 s**

The Atlas Test Plan

IT Solutions for the UN Family

Atlas - *select count(*)*

- Migrated from ZFS to UFS on the same storage system.
- A series of processes were selected for testing performance differences.
- Same processes were later repeated with SSDs.

Process Name	Process Type	Average Response time Before UFS (ZFS)	Average Response Time After UFS Conversion (29 Nov)	UFS on SSD 2-Feb-2010
Select Count(*) from PS_DISTRIB_LINE;	SQL	7 min 43 sec	2 min 16 sec	18 sec
Select Count(*) from PS_KK_ACTIVITY_LOG;	SQL	Never Return	31 min 35 sec	3 min 52 sec
Select Count(*) from PS_KK_EXCPTN_TBL;	SQL	15 min 34 sec	6 min 21 sec	60 sec
Select Count(*) from PS_KK_SOURCE_HDR;	SQL	5 min 3 sec	2 min 9 sec	22 sec
Select Count(*) from PS_PC_RES_PA_TA041;	SQL	26 min 5 sec	8 min 12 sec	35 sec
Select Count(*) from PS_PO_LINE_DISTRIB;	SQL	2 min 8 sec	30 sec	4 sec
Select Count(*) from PS_PO_LINE_SHIP;	SQL	1 min 59 sec	20 sec	8 sec
Select Count(*) from PS_PYCYCL_RST_TBL;	SQL	18 min 1 sec	5 min 12 sec	24 sec
Select Count(*) from PS_VCHR_ACCTG_LINE;	SQL	42 min 33 sec	11 min 26 sec	1 min 35 sec
Select Count(*) from PS_VOUCHER;	SQL	2 min 55 sec	1 min 9 sec	10 sec
Select Count(*) from PS_VOUCHER_LINE;	SQL	4 min 17 sec	2 min 5 sec	14 sec

ICC international computing centre

IT Solutions for the UN Family

Atlas – *SQR Reports*

Process Name	Process Type	Average Response time Before UFS (ZFS)	Average Response Time in Production (18 Feb 2010)	Average Response Time for SSD (18 Feb 2010)	Improvement (%)
1. Project Budget Balance - PBB UNGM556 (.pdf)	SQR Report	15 Sec	14 Secs	12 Secs	117%
2. CDR without Encumbrance - CDR UNGL143 (.htm)	SQR Report	2 min 13 sec	1 min 33 secs	59 Secs	158%
3. CDR with Encumbrance - UNGL143A (.pdf)	SQR Report	2 min 1 sec	1hr 30 mins	16 mins	563%
4. CDR by Project - UNGL143P (.pdf)	SQR Report	Error	Running Since 4 hrs still not finished.	22 mins	1091%
5. CDR by Activity - UNGL143B (.cvs)	SQR Report	>10 min	1 hr 27 mins	15 mins	580%
6. Project Transaction Detail RPT. UNEXPDET (.cvs)	SQR Report	>10 min	18 Secs	14 Secs	129%
7. Budget Transaction Detail Rep - GLS8005 (.pdf)	SQR Report	>1 h	18 secs	13 Secs	138%

Atlas – *PSoft Query*

Process Name	Process Type	Parameters used	Average Response time Before UFS (ZFS)	Average Response Time in Production (18 Feb 2010)	Average Response Time for SSD (18 Feb 2010)	Improvement (%)
1. UN_EXP Detail (online html, xls)	Query	BU:UNDP1, Budget Dept: B0064, Project:00069345, From/to Acctg Dt: 01/11/2009- 30/11/2009	Did not run	10 Sec	4 Sec	250%
2. UNCDF_AWD_BUD_EXPENSE (online, excel)	Query	Award ID (% for All) = 00051440,Project ID (% for All)=%	>5 min	3.5 Sec	2 Sec	175%
OFA_AP_SCA_MISSING_INFO	Query	period from o1 jun 09 to 18 nov 09 and donor 0006	Timed out	2 Sec	1 Sec	200%
OFA_AR_SCA_12000_FOR_DONOR	Query	2009, month =8 , donor %	Timed out	28 Sec	30 Sec	93%
OFA_SCA_TB_AP	Query	jan 1 2009 to nov 30 2009 for 0006 donor	>5 min	2 Sec	3 Sec	67%
OFA_SCA_TB_AR	Query	period from o1 jan 09 to 18 nov 09 and donor 0006	>3 min	17 Sec	28 Sec	64%
UN_AP_HQ_PCYCL_SIMULATOR	Query	NA (SQLPLUS)	>5 min			
UN_AP_SCA_APJV_REVIEW	Query	jan 1 2009 to nov 18 , donor % BU UNDP!	Error	8 Sec	2 Sec	400%
UN_FUND_STATUS	Query	yr 2009, BU UNDP1, Budget Dept B0150, period 1 to 12	Did not complete	8 min 9 sec	2 min 10 sec (used % for fund,project and donor field)	376%
UN_SCA_GLOBAL_FUND_12000_UNDP	Query	NA	>5 min	3 min 35 sec	25 Sec	860%

Atlas – *Administrative + Other*

Process Name	Process Type	Average Response Time in Production (18 Feb 2010)	Average Response Time for SSD (18 Feb 2010)	Improvement (%)
Runstats on entire database	oracle	avg 13 hours	4.25 hours	306%
Purchase Item Inquiry Performance	Sql	3 Secs	1 Secs	300%
Semi Manual Reconcilliation	Page	2 mins	30 Secs	400%
Sample processes(XRFWIN)	SQR Report	15 Secs	14 Secs	107%
Create index for Purchase Item Inquiry	Sql	15 Secs	8 Secs	188%
Update statistics on PS_VENDOR_LOC	Sql	19 Secs	10 Secs	190%
select count(*) from PS_VENDOR_LOC;	Sql	23 Secs	5 Secons	460%

Conclusions

- Where I/O is a significant part of the workload, SSDs perform 3-5 times faster than rotational disks.
- The performance *difference* makes SSDs an *IT solution,* not just a storage choice
- From both user and administrator perspectives, the difference would be significant.
- From an organizational perspective, the increase in speed would represent considerable savings.

So, then what happened?

So, what happened?

- RFO was issued, only two vendors replied.
- We chose RamSan vendor Texas Memory Systems, but had questions on availability.
- Virgin Blue reservation system 21-hour outage
- …other things…money…
- We still want it, but are looking at more selectively using flash instead of putting whole DB on it.
- Exadata is being talked about, too.