Seagate Enterprise Data Warehouse Realstor Story

Amanda Strassle | August 8, 2017

Introduction

- **Amanda Strassle, IT Senior Director**

- **Global Business Data Center (sites in U.S. and Asia)**

- **Providing variety of hosting services for corporate systems both on-premise and in the public cloud**

- **Thousands of virtual machines, hundreds of databases and applications**

- **amanda.a.strassle@seagate.com**

Seagate's Enterprise Data Warehouse (EDW)

The subject installation at the root of our challenges

Critically important system for our engineering, quality and finance teams, driving business decisions on a daily basis.

Data loaded constantly (24x7) averaging 9 billion records each day. Equal amount purged each day.

Database Size: ~170TB
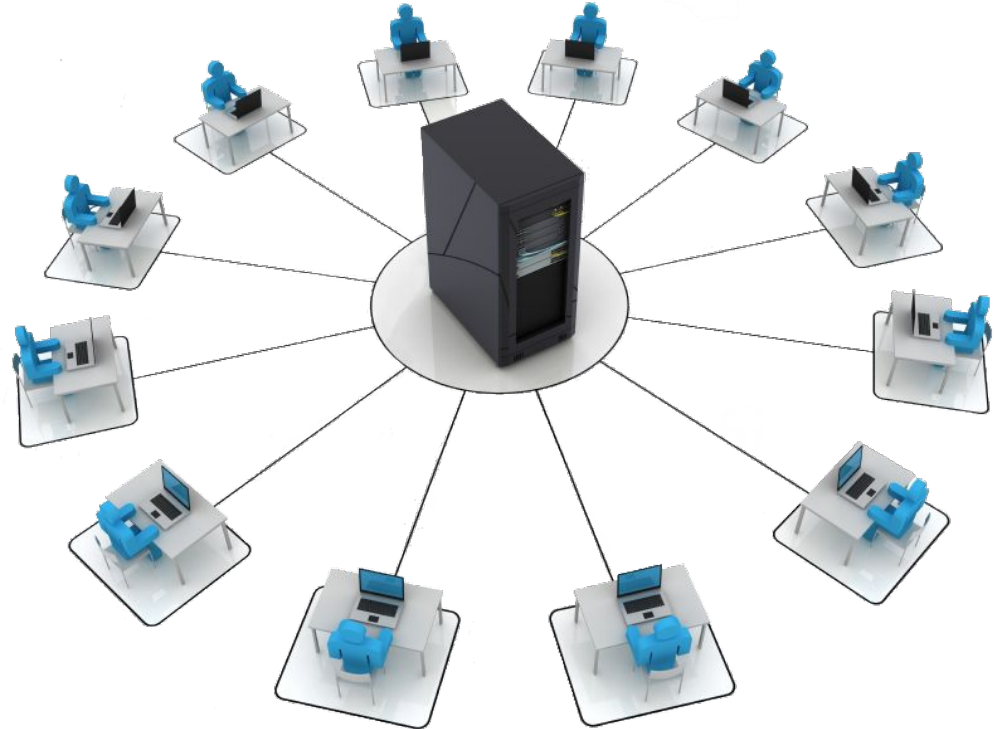
Our Challenges

Reduce Cost
Hardware was monolithic and costly. Desire to move to less expensive hardware.

DB Management Burdens
Too much time spent continuously moving from FC to SATA in order to meet performance expectations.
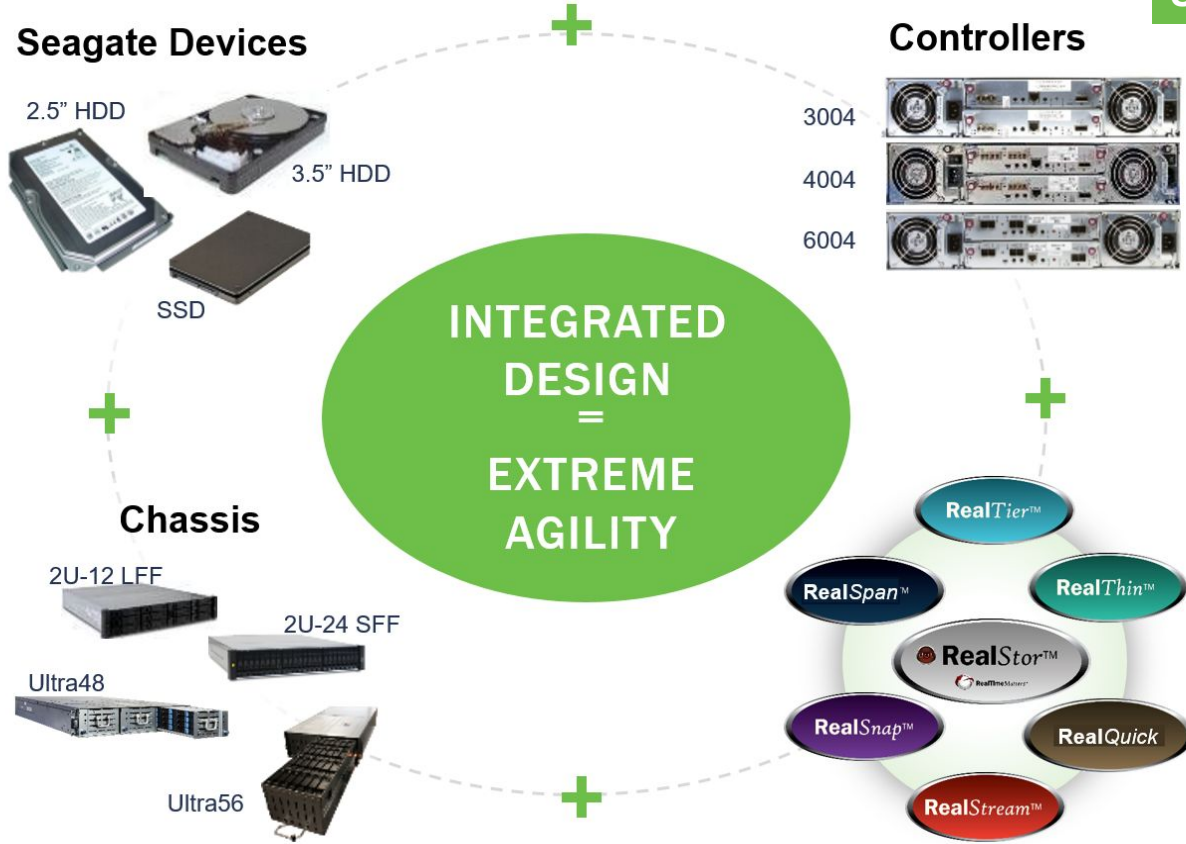
Space Constrained
Monolithic storage arrays were running out of capacity and physical footprint was limiting.

RealStor Array Modular Product Architecture

Seagate Devices

2.5" HDD

3.5" HDD

SSD

Controllers

3004

4004

6004

INTEGRATED DESIGN = EXTREME AGILITY

Chassis

2U-12 LFF

2U-24 SFF

Ultra48

Ultra56

RealTier™

RealSpan™

RealThin™

RealStor™

RealSnap™

RealQuick

RealStream™

Real Time Matters™

Real*Stor* Software Features Summary

- Real Time Storage Operating System – Foundation For Innovative Software Features

- Framework For Real Time Data Collection – Provides Information About Data Usage

	What is it?	Why is it Important
Real*Tier*** ✓	Tiering Software	Automatically move data where you need it, when you need it, in real-time
Real*Thin*** ✓	Thin Provisioning	Dedicate storage only when needed. Reclaim space for other applications with UNMAP and Zero-Block space reclamation
Real*Quick*** ✓	Rapid Raid Rebuild	Accelerated RAID rebuilds based on amount of used capacity
Real*Cache*** ✓	SSD Read Cache	SSD cacheing for read-intensive applications
Real*Snap*** ✓	Redirect-on-write Snapshots	Reduce RPO/RTO Mount snapshots as volumes for developing and testing
Real*Span***	Asynchronous Replication	Move volumes across from one array to another for off-site disaster recovery
Real*Stream*	Autonomic real-time ingest for IoT workloads	Provides consistent Quality-of-Service and Maximum performance for handling 1000's of concurrent incoming streams

RealStor™
RealTimeMatters®

* Dependent upon Virtual Pool

Storage Architectural Changes

Primary

Secondary

BEFORE
14 cabinets (7 in each CR)
380 TB usable capacity
2% FC SSD (7.6 TB) 62% 10K FC HDD (235.6 TB) 36% 7.2K SATA HDD (136.8 TB)
300 GB cache

Primary   **Secondary**

AFTER
2 cabinets (1 in each CR)
1,000 TB usable capacity
14% SSD (140 TB) 86% 10K Hybrid HDD (860 TB)
56 GB cache

Our Challenges Turned Into Improvements

Reduce Cost
Hardware was monolithic and costly. Desire to move to ... ha...

Eliminated annual maintenance of $400,000

DB Management Burdens
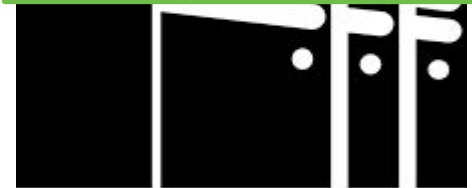Too much time spent continuously moving from FC... ...heet pe... ...s.
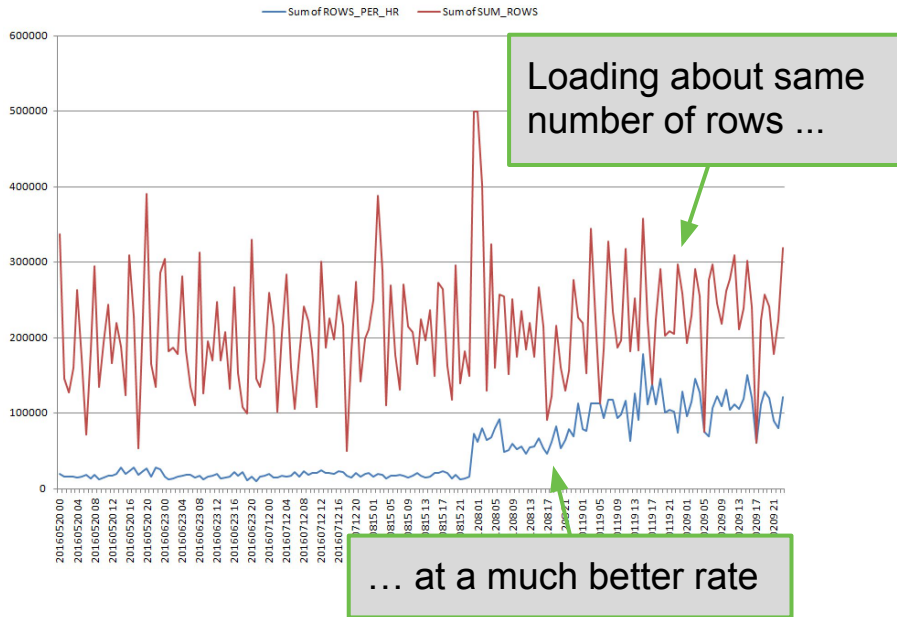
Only SSD and Hybrid

Space Constrained
Monolithic storage arrays were running out of ...

163% increase in usable storage & significant HW footprint reduction (2 cabinets from 14**)**

Benefits Realized

Data Delivered Faster To End User

Now loading data from drive, head and media factories up to 50% faster.

Engineers able to start analysis sooner and detect issues earlier.

Loading about same number of rows ...

… at a much better rate

Benefits Realized

Faster Bill of Materials (BOM) Maintenance

OVER 200% IMPROVEMENT

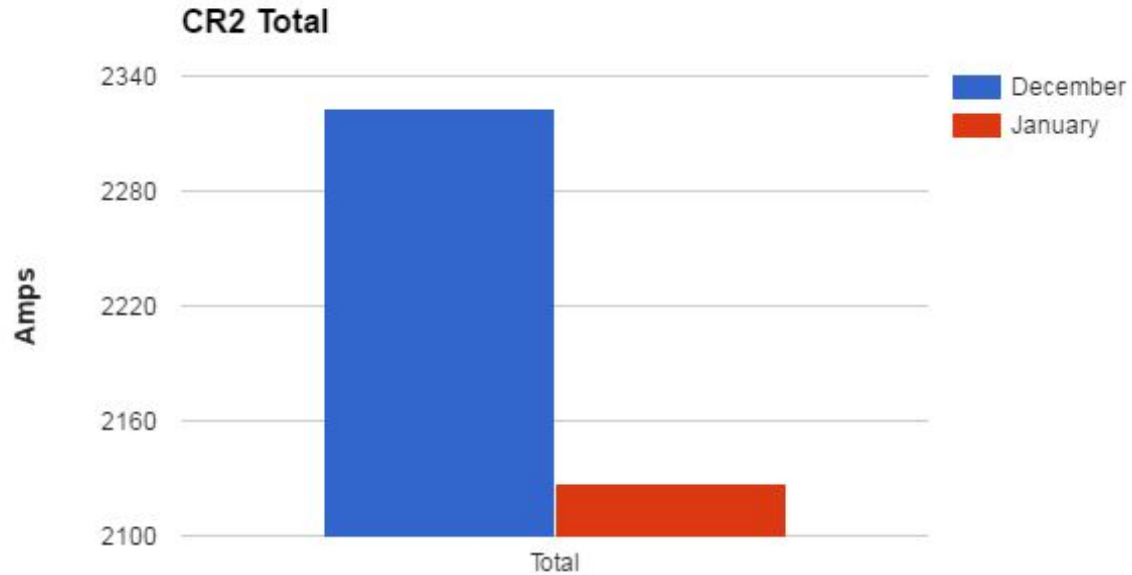| Before	After
Reload time required waiting until the weekend resulting in up to a week backlog of BOM updates not being loaded into the table. Potential for users to not be looking at the latest BOM data.	Today, we have the ability to reload the table in an overnight maintenance action, if necessary, during the workweek while minimizing any user impact.
Truncate/Reload in Dec 2015 Total # of rows in table: ~500 Million Total reload time: 19 hrs	**Truncate/Reload in Jan 2017** Total # of rows in table: 913 Million Total reload time: 15 hrs

Almost twice as many rows in less time.

Benefits Realized

Computer Room Power & Floor Space Savings

8.6% reduction in power consumption in one computer room (represents only 1 of 2 replaced arrays removed)

Translates into a total estimated annual power savings of $35,800, improved performance and 75% reclaimed computer room floor space.

CR2 Total

Power Measures (before & after original storage removed)

Seagate Realstor with SSD and Hybrid Drives

Outcome of our experience

FASTER ☑

CHEAPER ☑

BIGGER ☑

BETTER ☑