Industry Trends and Technology Perspective White Paper

Business and Application Aware DPM

A look at the evolving landscape of data protection management (DPM) and infrastructure resource management (IRM)

By Greg Schulz

Founder and Senior Analyst, the StorageIO Group

May 15, 2007

Data protection management (DPM) provides timely insight and analysis of data protection activities, including cross-technology domain event correlation analysis for problem resolution and planning purposes to optimize IT resource management. This paper looks at the shifting landscape of DPM and its role in enabling cost-effective IT infrastructure resource management (IRM) and service enhancement.

www.storageio.com P.O. Box 2026 Stillwater, MN 55082 651-275-1563 info@storageio.com

Introduction

The data protection management (DPM) landscape continues to evolve to support timely and cost effective data protection so information is safe, secure and accessible when and where it's needed. First generation DPM approaches were centered on basic backup reporting. The next generation focused on timely and insightful root cause and effect analysis across multiple technology domains to support business and application aware data protection. Data protection today is as much about ensuring that data protection meets compliance and coverage requirements for service level objectives as it is about optimizing the use of IT resources to contain and reduce cost expenditures.

IT infrastructure resource management (IRM)

DPM is part of a broader IT infrastructure resource management (IRM) focus (Figure 1) spanning multiple technology domains (applications, servers, connectivity, and storage devices). While IRM can have a storage centric view, the focus is on management of IT resources to deliver application services and information to meet business service requirement objectives addressing performance, availability, capacity and energy consumption (PACE) among other disciplines. As shown in Figure 1, IRM functions and activities include:

- Asset and facilities management including security
- Change and configuration management
- Performance analysis and capacity planning
- Chargeback and service level agreement (SLA) management
- Data protection and media management
- Infrastructure resource service and maintenance task scheduling

For example, traditional backup reporting has had a storage centric view focused on tape and media management, utilization, and backup success or failure information. In larger organizations, IRM functions and activities may be delegated to various groups addressing specific technology domains, such as server and storage teams, or cross-technology domain performance and capacity management groups.

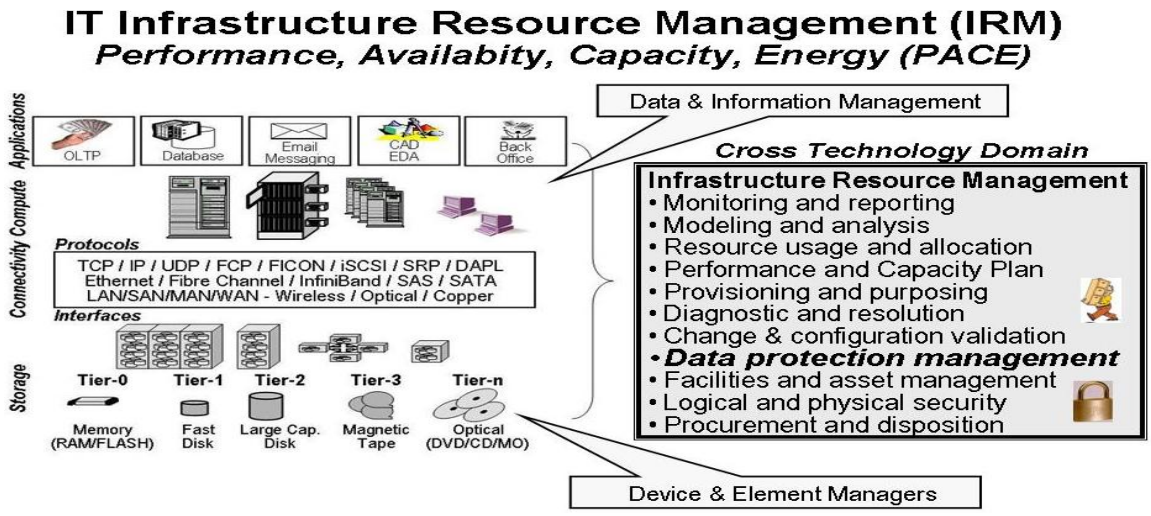

Figure 1: IT IRM and DPM relationship

Some IT infrastructure organizations may be more integrated than other environments with different individuals performing specific functions in a given technology domain or working across various IT resource and service delivery domains. That is, a storage administrator may also be a server admin or a

performance and capacity planning analyst. Given the complexities and interdependencies of applications that rely on multiple servers, operating systems, databases and IT resources, backup and DPM tasks similar to performance and capacity planning or change control and configuration validation need to look across the different technology domains that are used to support various business applications.

The increase of regulatory requirements combined with pressure to meet service levels along with 24x7 data availability has resulted in data protection interdependencies across different business, application and IT entities. Consequently timely and effective DPM requires business and application awareness to correlate and analyze events that impact service and IT resource usage. *Business awareness* is the ability to collect and correlate IT assets to application interdependencies and resource usage with specific business owners or functions for reporting and analysis. *Application awareness* is the ability to relate IT resources to specific applications within the data protection environment to enable analysis and reporting.

A challenge with business and application aware DPM has been the fact that many organizations maintain information about business units, applications, IT resource, or asset ownership and usage in disparate database and repository formats. For example, information is kept in configuration management databases (CMDB), performance management databases (PMDB) and metadata repositories among other locations. To support business and application aware data protection, DPM tools need to support access of external sources including SQL databases, LDAP source code, and XML structured data files as well as flat files including CSV worksheets.

Effective data protection management includes knowing who, what, where, when, and why resources are being used to deliver service and how effectively service is being delivered with a business and application awareness. To enable timely IRM across different technology domains and disciplines including DPM, automated data collection and correlation of event and activity information is needed. Event correlation from different sources facilitates root cause analysis to find the real source of problems so that service levels and compliance objectives can be meet. With a focus on reducing electrical energy consumption and associated environmental impact, DPM can be used to insure that data protection resources are being optimally used to avoid costly upgrades.

IRM works in conjunction with data lifecycle management tasks that focus on organizing and managing data. Data protection and DPM are functions of IRM that support data and infrastructure management, including monitoring resource usage and activity as well as making sure that data is being protected to ensure recoverability and compliance.

The changing landscape of DPM

Backup reporting, the predecessor of DPM, has evolved from basic vendor supplied utilities for backup status and media usage reporting to third party heterogeneous backup utility reporting. DPM continues to evolve from collecting status information on multiple vendor backup software products and media utilization with interpretive report analysis to cross-technology domain event correlation and root cause analysis for a more in-depth active view of data protection effectiveness. DPM solutions today are expanding their focus and areas of coverage from a single dimension focus on backup to a multi-dimensional focus across different data protection techniques and technologies including replication.

Although an environment may have multiple tools and technologies to support IRM activities, DPM tools are evolving to support or co-exist with management of multiple data protection techniques including backup (to disk or tape), local and remote mirroring or replication, snapshots, and point-in-time (PIT) copy and file systems. Key to supporting multiple data protection approaches and technologies is the ability to scale and process in a timely manner, rapidly increasing large amounts of event and activity log

information. At the heart of a new breed of IRM tools, including DPM solutions, are robust cross-technology resource analysis and correlation engines to sift disparate data protection activity and event logs for interrelated information. Scalability is another attribute of next generation DPM requiring solutions to support event correlation automatically across 1,000s of local and distributed servers to enable timely insight and decision making.

In Figure 2, information is collected from across different technology domains, including backup applications, servers, applications such as databases, networks (LANs and SANs), and storage devices (SAN, NAS, disk, and tape). Leveraging a robust repository of correlated events and combining powerful analytics, DPM tools enable timely and descriptive alerts to be sent that provide insight into how well backup and data protection tasks are performing, to address errors and problems in real-time.

The power of a robust analysis and event correlation engine as seen in Figure 2 enables automation of the time consuming task of connecting the dots between thousands of various events and allows timely, more accurate analysis of events associated with DPM.

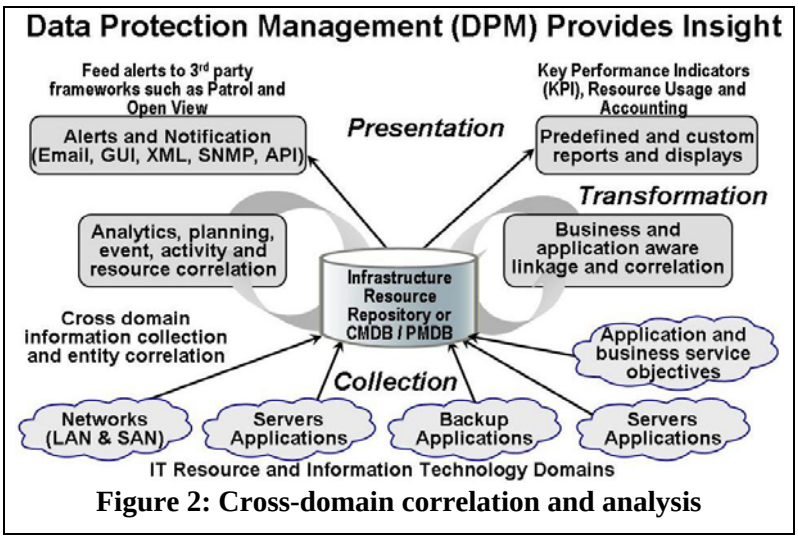

Figure 2: Cross-domain correlation and analysis

For example, instead of assuming that the root cause of a rash of backup job failures is a tape drive (a false positive) based on basic reporting results, a more in-depth analysis combing thorough event and activity logs from servers, backup logs, and storage devices could reveal that the tape drive going off-line is due to a server involved in the backup task intermittently going off-line as a result of a network congestion issue. However, in-depth analysis also reveals that certain tape cartridges being used for disk to disk to tape (D2D2T) replication and backup are being excessively reused, indicating a potential future problem.

In addition to eliminating false positives to insure effective data protection coverage, DPM can work with other IRM tools to maximize resource usage and application service delivery with current IT resources. The benefit of maximizing usage of existing IT resources including servers, storage, networks and software licenses without sacrificing performance or availability is the containment or reduction of upgrade cost as well as enhancing environmental impacts including power and cooling requirements.

Next generation DPM solution criteria

DPM continues to evolve from basic backup success and failure reporting to active management providing timely insight into why and how data protection tasks are performing, and their effectiveness for applications and systems that are growing in complexity and size. This includes looking across different technology domains to support heterogeneous backup and data protection tasks.

An object based data protection approach encompasses all of the resources and applications that comprise a business object, such as an online transaction processing (OLTP) system to support retail sales. In the OLTP scenario, a data protection object view involves local and remote data replication with snapshots (disk to disk, or D2D) combined with backing up to off-line medium (D2D2T) encompassing different

servers, operating systems, databases, and individual applications. The result is the ability to support restoration of a complete application or business function over multiple servers and data sources as well as restoration of a single file as needed.

Basic functionality for DPM includes backup activity and media usage reporting with a shift towards multi-tenant and cross-technology analysis and automated event correlation for complex data protection tasks. Next generation DPM capabilities include management of multiple types of data protection tasks in various combinations along with integration of different vendors' IRM and data protection technologies for event analysis and planning. Figure 3 shows the relative positioning and characteristics of various backup reporting and DPM centric tools and the progression from basic reporting to more in-depth application and business function aware event and activity analysis.

	Aptare	Bocada	Illuminator	Symantec	Tek-Tools	WysDM
Multi-vendor data protection reporting and activity dashboards	Yes	Yes	EMC & NetApp replication	Yes	Yes	Yes
Basic backup event and media usage reporting	Yes	Yes		Yes	Yes	Yes
Cross IT technology domain analysis engine						Yes
Business and application awareness						Yes
Replication management			Yes			Yes
Centricity	Backup event and activity reporting	Backup activity reporting	Replication -centric manager	Backup activity reporting	IT activity and profile reporting	DPM and IRM event correlation analysis

Figure 3: Relative positioning of various[1] IRM backup reporting and DPM centric tools

Next generation DPM solutions are evolving from basic backup status reporting including job step and backup stream activity (or media usage reporting) to more in-depth automated analysis. Functionalities to consider in a DPM solution include the ability to link data protection activities to application and business function and service requirements, including inter-system and application dependencies. DPM solutions need to be able to collect and analyze data from multiple IT resource for a multi-dimensional view, including different types of data protection technologies.

DPM solutions should be able to identify and assist with optimization of IT resource usage in terms of effective performance, availability, and capacity utilization to maximize current technology. This capability should work with and enhance other IRM activities enabling planning for future IT resource needs (servers, storage, networks, media, power and cooling). Additional DPM capabilities should include the ability to identify orphan storage and data that needs to be backed up and protected as well as support for audits for compliance to ensure that protected data is complete and recoverable in a timely manner to meet service objectives.

A common theme heard from IT customers is the need to do more with less, maximize existing resources, cost containment, reducing energy consumption, deferring technology upgrades and enabling IRM and

[1] Information based on publicly available material

www.storageio.com P.O. Box 2026 Stillwater, MN 55082 651-275-1563 info@storageio.com

DPM that scales to support thousands to tens of thousands of servers. Another theme commonly heard is to eliminate false positives when performing DPM and IRM functions to focus on real problems in order to optimize IT resource usage while identifying and isolating potential errors to avoid cascading chains of event failure scenarios. DPM solutions need to co-exist with other IRM tools and repositories including CMDB and PMDB via multiple interfaces and access methods. Ease of use and customization features, including wizard based user interfaces and report generation, along with standard reports and dashboards for real-time or historical perspective and event notification have become basic, must have functionality for effective DPM solutions.

How WysDM is prepared for next generation DPM

The WysDM solution is positioned to support next generation DPM along with complementing other IRM tools and technologies. At the heart of the WysDM technology is a robust analysis engine that uses information collected from various technology domains to enable application and business aware event correlation and analysis. Most backup reporting and DPM tools today provide basic reporting and interpretive analysis of report results in reports or dashboards, including multi-step backup job status. WysDM builds on this baseline functionality by leveraging its correlative analysis engine to connect the dots across different IT resources that are involved in application service delivery that need to be protected to remain compliant with regulation and service requirements.

Desirable DPM feature checklist

- ✓ Cross-technology domain event collection
- ✓ Scalable root cause analysis engine
- ✓ Data collection from diverse sources
- ✓ Business and application aware
- ✓ Multiple data protection technologies
- ✓ Heterogeneous across different vendors
- ✓ Real-time and historical reporting
- ✓ Support 1,000s to tens of 1,000s of servers
- ✓ Timely event notification mechanisms
- ✓ Co-existence with other IRM tools
- ✓ Configuration and compliance validation

While WysDM certainly provides backup data protection management, real-time analysis and event reporting, the technology can also be leveraged as part of a larger IRM approach to ensure effective business and application aware data protection and resource utilization while co-existing with other IRM technologies.

DPM has evolved from basic backup status and media usage reporting to encompass complex objective based data protection technologies and techniques in order to protect business functions defined as application objects. Consequently, given the complexity and intersystem application and technology dependencies, automated analysis and event correlation across various IT technology domains is imperative for timely and effective data protection management. DPM is part of IRM, thus solutions need to co-exist and complement other IRM enabling technologies and tools for timely and effective IT resource and data protection management. The StorageIO Group recommends that IT organizations not currently looking at DPM or for those with 1st generation backup reporting tools to explore how current DPM tools can co-exist and compliment other IRM tools to improve application service delivery while optimizing IT resource management.

About the author

Greg Schulz is founder and senior analyst of the StorageIO Group and author of the book *Resilient Storage Networks — Designing Flexible Scalable Data Infrastructures* (Elsevier Digital Press).