

A Centralized Data Protection Application for Cross Vendor Storage Systems

**Prateek Sinha & Nishi Gupta
Tata Consultancy Services**

Data centers consist of Storage Products from different vendors to meet the storage needs of multiple servers hosting various applications like: Databases, Mail Servers. These mixed environments pose a challenge for administrators as they have to use different Data Protection Applications (DPAs) for different combinations of Applications, Operating Systems and Storage Systems in an enterprise deployment.

Here we are presenting an approach to resolve the mentioned issue. A centralized DPA can be developed with a single user interface, which uses storage API's in the backend to perform data protection operations on cross-vendor storage systems. It consists of pluggable modules for different functionalities in areas example Backup/Recovery, Archival, Compliance, Deduplication.

We will be sharing the results of the PoC and the challenges that were faced and proposing the need of Standards in Storage Data Protection APIs.

We welcome any suggestions/inputs as this is an ongoing work.

- ❑ Trends in Data Protection
- ❑ Scope of Solutioning in Data Protection
- ❑ Data Protection Challenges
- ❑ Problem Description
- ❑ Proposed Solution
 - ❑ Benefits
 - ❑ Challenges
- ❑ Conclusion

Trends in Data Protection

- Despite the tightened IT budgets of the past few years, disaster recovery, business continuity, and data protection have emerged as top storage spending priorities for 2010. IDC's recent study on IT budgets and spending patterns has discovered a renewed emphasis on data protection solutions that address evolving data storage requirements of organizations in the coming years

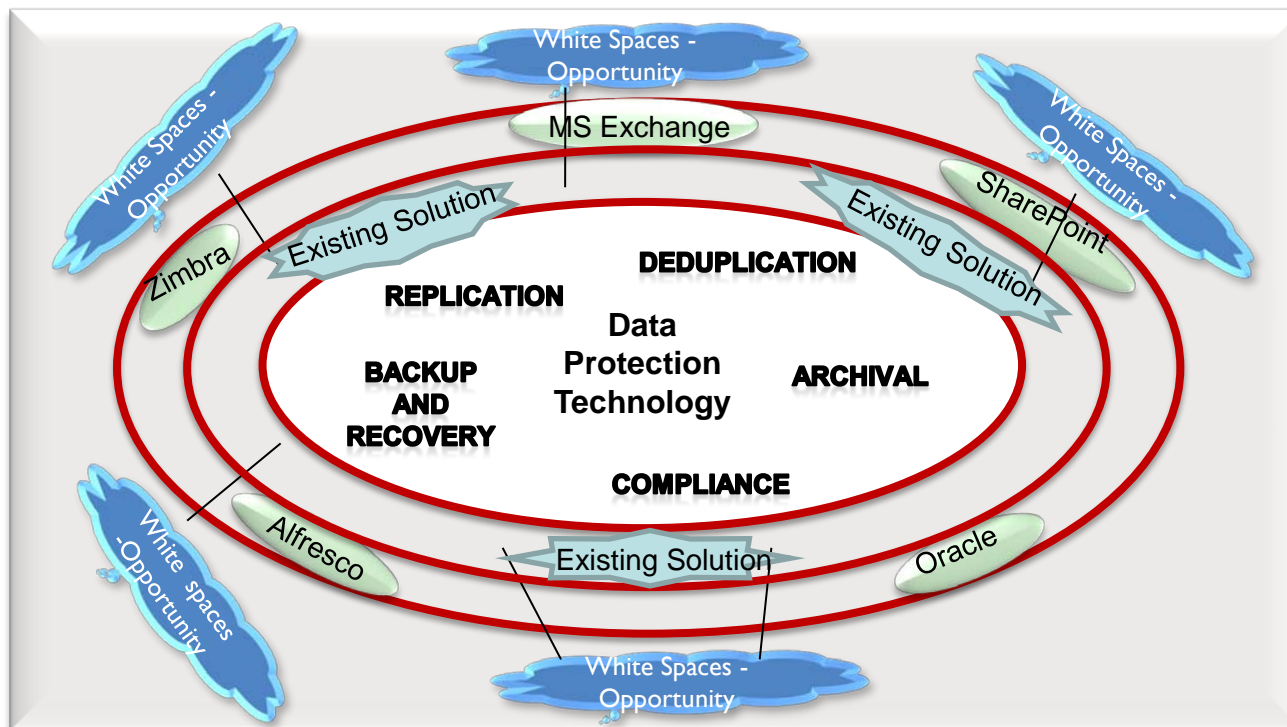
Source: IDC Report- Backup and Recovery: New Strategies Drive Disk-Based Solutions, December 2010

- Organizations of all sizes now realize that the loss of access to data directly affects their ability to operate. The recommended allocation of IT budget to this critical function ranges from 5% to 10%, depending on the type of business. Worldwide, SMB spending on data protection is estimated at \$30 billion to \$60 billion in 2010.

Source: Outlook Report- Data Protection and Recovery in the Small and Mid-sized Business (SMB), October 2010

Solution Scope : Data Protection

- ❑ Data Protection technologies are available through Storage Vendors. Using integration, these technologies can be leveraged to develop solutions that address the Data Protection needs of any Business Application.
- ❑ Some integrated solutions exist, but still there can be whitespaces/gaps that need to be filled.



Data Protection Challenges

□ Data Center Challenges

- Exponential growth of data
- Protecting the data of mission critical applications
- Management of complex and heterogeneous environments
- How to meet challenging service level agreements (SLAs) and compliance related requirements ?

□ Data Protection Solution Challenges

- Complex Enterprise Computing Environment
 - Platform, Network , Storage, Applications
- Heterogeneous mix of multiple Vendor Products
- Multiple Integration Points
 - Storage, Server, Applications, etc
- Interoperability and Compliance Requirements
- Growing demand for End to End Solutions

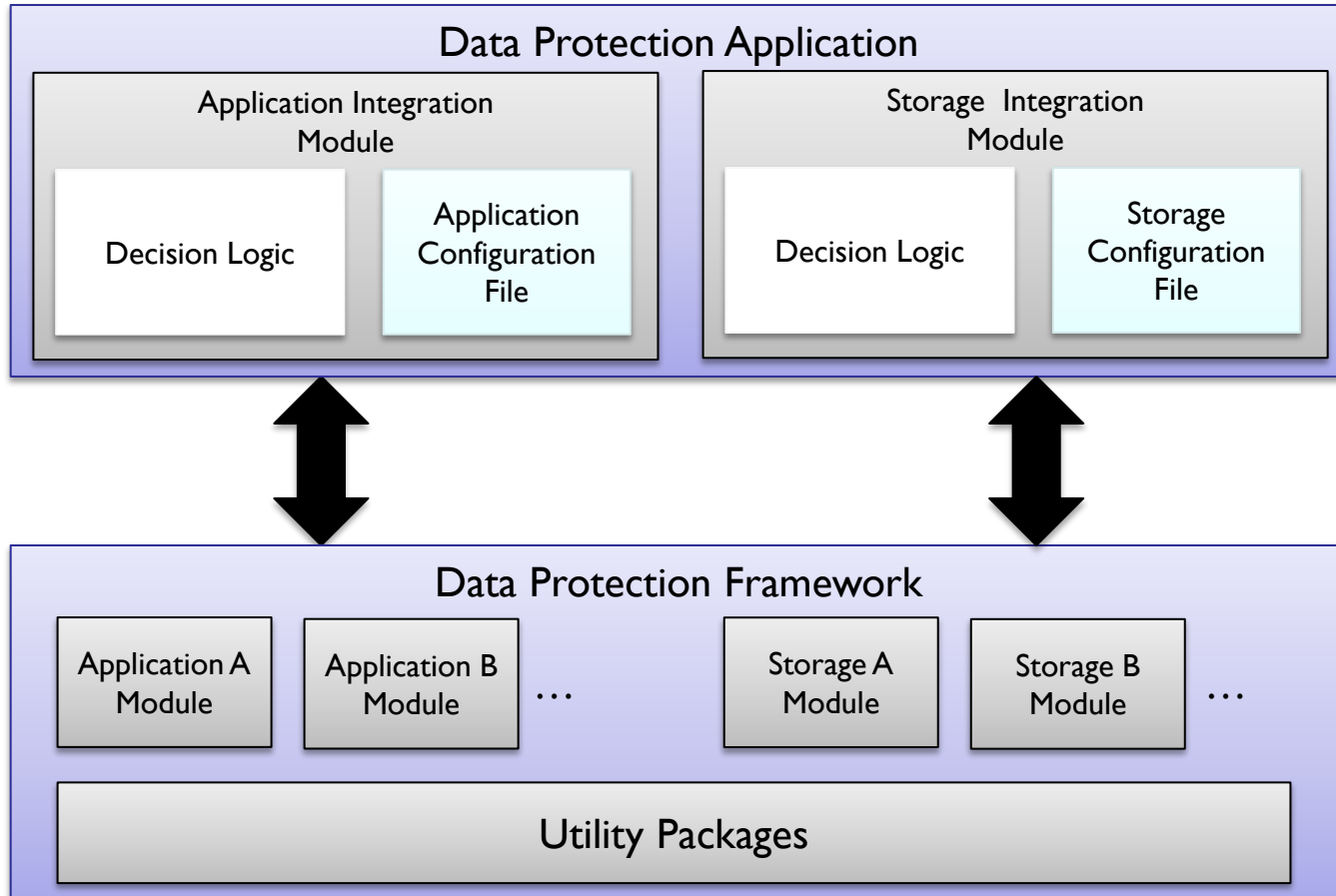
APPLICATIONS	Email Servers, ERP's, Database Servers, Content Management Systems, Industry Specific Applications, others
OPERATING SYSTEMS	Linux, Windows, Solaris, OS X, others
STORAGE NETWORK	iSCSI, FCP, NFS, CIFS, FCoE
STORAGE SYSTEM	EMC, NetApp, HDS, HP, Sepaton, others

An Approach to address the challenges

Data Protection Solution Requirements

Business Problem	Current Scope	Requirements
Protection of mission critical /business critical application data	Yes	Backup <ul style="list-style-type: none"> • Online • Offline • Minimum RPO Restores/Recovery <ul style="list-style-type: none"> • Application Level • Granular : File Level, etc [Depending on the Application: Project Level, Database Level] Replication <ul style="list-style-type: none"> • Block level • Minimum RPO/RTO
Management of complex heterogeneous Environment	Yes	A Centralized User Interface for Administrator
Handling exponential data growth	Yes	Storage Optimization
Meeting challenging Service Level Agreements (SLA's) and Compliance requirements	Yes	<ul style="list-style-type: none"> • Automation • Data Archival • RPO/RTO

Solution Architecture



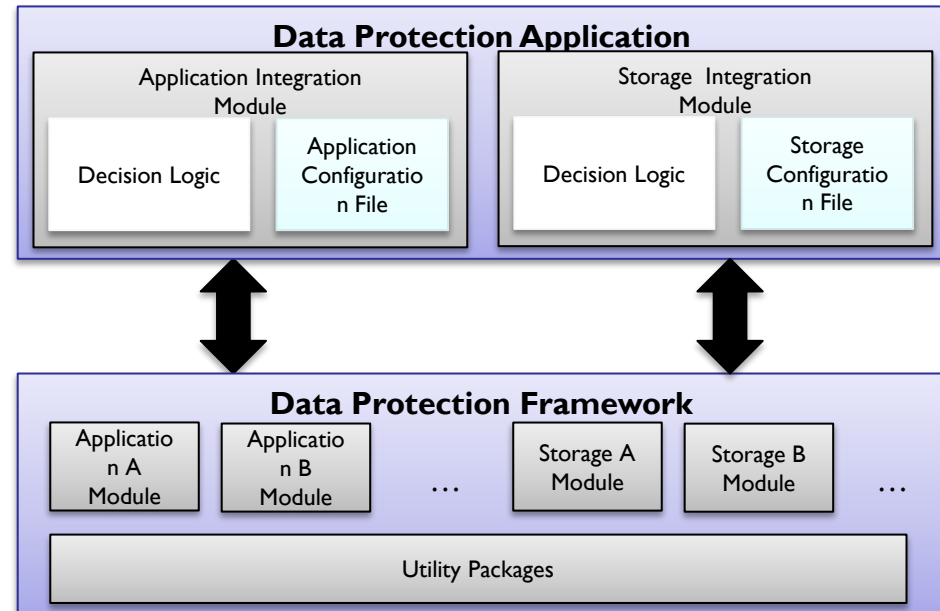
Solution Components

□ Data Protection Application

- GUI/CLI
- Decision making logic
- A dynamic configuration file
 - Application Configuration file
 - Storage Configuration file

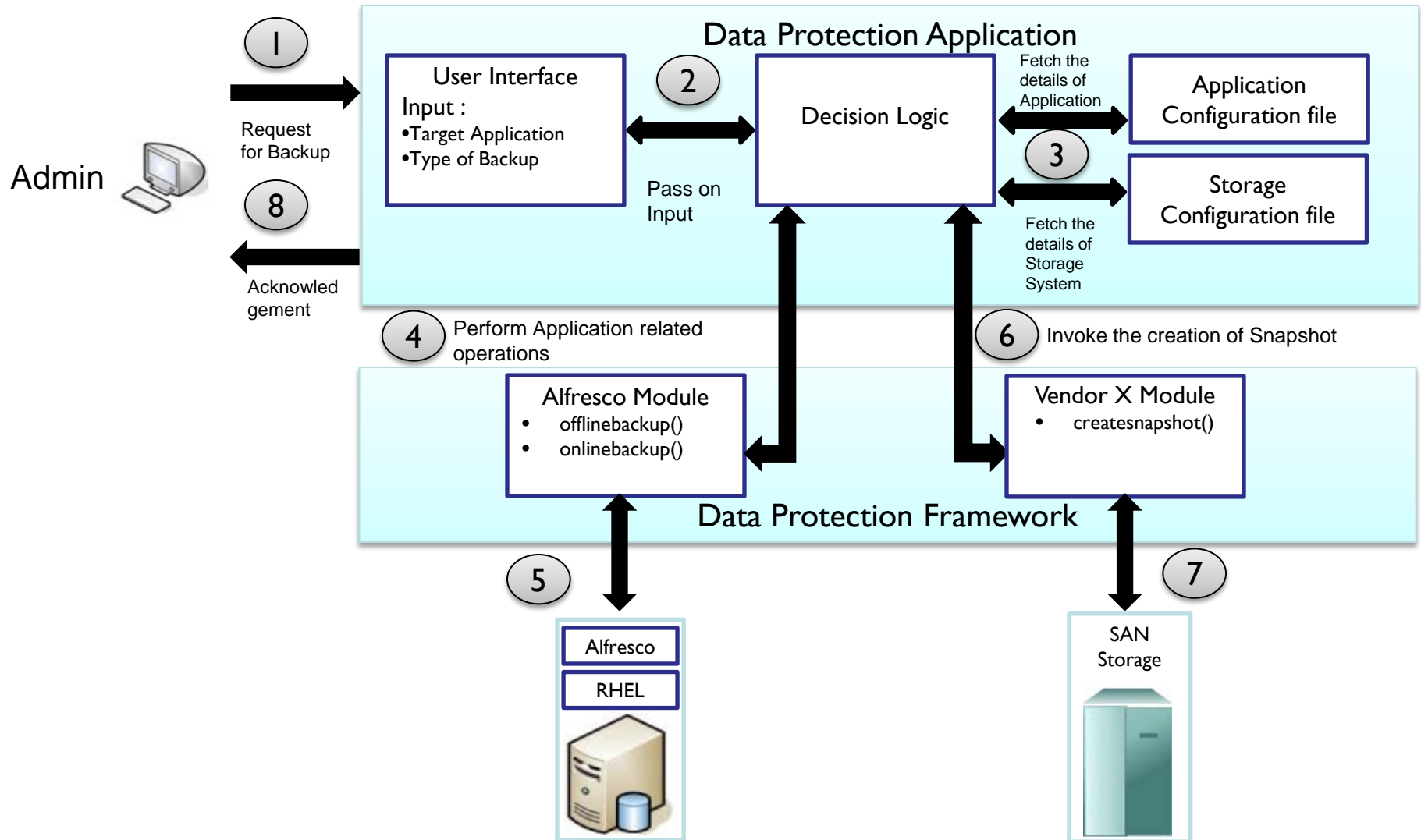
□ Data Protection Framework

- Application Modules: Business Application specific APIs
 - Automation of Application specific operations
- Storage Modules: Storage System Specific APIs
 - Snapshots
 - Cloning
 - Replication
 - Deduplication

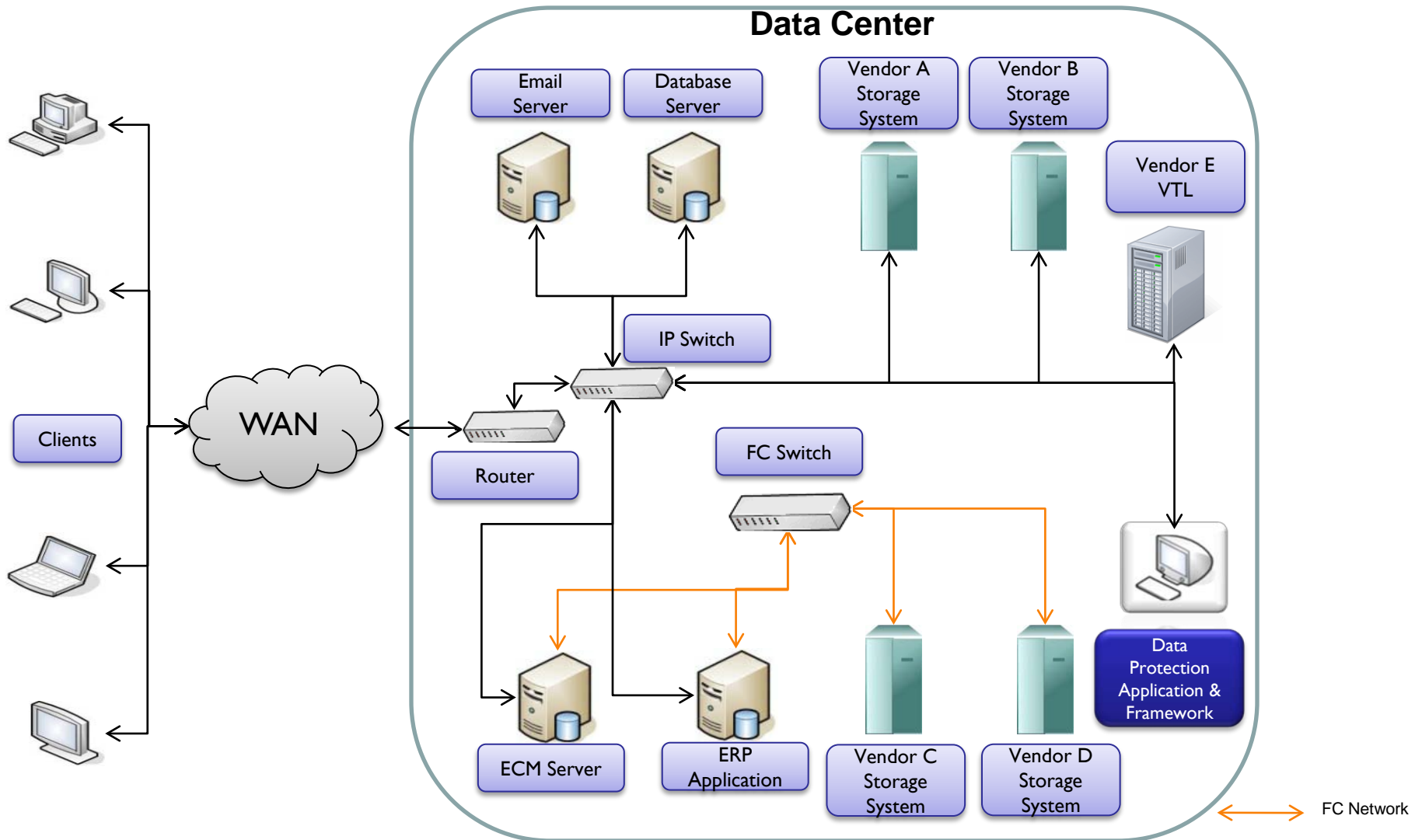


- ❑ Backup of Application data
 - ❑ Offline D2D Backup
 - ❑ Online D2D Backup
- ❑ Tiering of backups
 - ❑ Movement of Disk based backups data to tapes
- ❑ Replication of Application data
- ❑ Deduplication of Application data
 - ❑ Production data
 - ❑ Backed up Data
- ❑ Restore/Recover data from
 - ❑ Tape
 - ❑ Snapshots
- ❑ Compliance
 - ❑ Retaining data for specific time duration
- ❑ Backend for any data protection software application
 - ❑ The framework can be reused by any existing DPA

Sample Case : Alfresco Hosted on Proprietary Storage System of Vendor X



Deployment Diagram



Solution Summary

Business Problem	Can Be Addressed	How?	Description
Protection of mission critical application data	Yes	<ul style="list-style-type: none"> Integration of Applications and Backup/Recovery technologies of Storage Systems. Application Consistent Online Backups. Granular Restores 	Application Specific Modules of the framework would automate all the steps that are required to bring the application in consistent state before snapshots are taken.
Management of Complex Heterogeneous Environment	Yes	<ul style="list-style-type: none"> A Centralized User Interface for Administrator 	A web-based centralized user interface can be developed. This will help the administrator to perform all the Data Protection Related operations on all Applications and Storage System.
Handling exponential data growth	Yes	<ul style="list-style-type: none"> Leveraging the deduplication technology of storage systems 	The Storage System Module will invoke the specific API for Deduplication.
Meeting challenging Service Level Agreements (SLA's) and Compliance requirements	Yes	<ul style="list-style-type: none"> Automation Archival 	The Data Protection Application will use the Framework to automate the Backup, Recovery, Replication process and also use the archival feature of the storage system to meet compliance related requirement

❑ Administrators

- ❑ Can use Single Application/Script to manage Backups over a combination of Application and Storage Systems
- ❑ Application administrators need not focus on the Data Protection of application data.

❑ End Users

- ❑ Backups at Storage Layer are quick and independent of the size of data being backed up.

❑ Data Protection Software Vendors

- ❑ Existing Data Protection Technologies of Storage Systems can be leveraged
- ❑ Can use the Framework to support unsupported Applications and Storage Systems.

❑ Storage Vendors

- ❑ Need not bother about integration of their products with business application

Challenges in developing such a Framework

- ❑ Most of the Business applications are closed source
- ❑ Understanding of Business Application for
 - ❑ Application Consistent Backups
 - ❑ Recovery at different granular levels
- ❑ There are few Storage Vendors that expose the Storage API's to a third party.
- ❑ Analysis of APIs
 - ❑ Differences in signature of methods/functions, etc

- ❑ Industry is really lacking Consolidated Backup Solutions which can support application consistent backups across different storage systems.
- ❑ There is a need for Standardization in Data Protection area as it will reduce the time needed in integration of Business Applications and Storage Systems.

Questions ?

Mail us @

prateek.sin@tcs.com

nishi.g@tcs.com

Special Thanks to Reena Dayal and Udayan Singh

We would also like to thank Amit Shukla, Girish Agnihotri, Brijesh Das MK and Ankur Srivastava of Storage Center of Excellence for making this demonstration possible