

INTEROP

NEW YORK CITY
SEPT 30 - OCT 4, 2013 | EXPO: OCT 2-3

IT Leads, Business Succeeds

Save an Extra \$200
Priority Code: MPIWK

REGISTER NOW

SPONSORED BY:



MOBILE

BIG DATA REPUBLIC

Transform Your Business With Data

ABOUT US REGISTER LOGIN

HOME BLOGGERS MESSAGES POLLS WEBINARS RESOURCES VIDEOS | STRATEGY ANALYTICS TECHNOLOGY FB TW LI G+ RSS

FINANCIAL SERVICES HEALTHCARE GOVERNMENT EDUCATION RETAIL

BIG DATA STORAGE

Confront 4 Key Big Data Storage Challenges



John Edwards, Technology Journalist & Author
10/9/2013
Comment
4 comments

Login
50% 50%

Like 2 Tweet 0 Share 2 +1 2

If you're planning to get the most out of big data, it makes sense to have a secure and reliable way of storing and accessing large amounts of data. Yet many enterprises fail to create a storage platform that adequately meets all of their big data needs.

Fortunately, formulating a big data storage strategy that will satisfy both current and future needs can be accomplished by following just a few simple steps.

Start with the platform

Designing and deploying the wrong type of storage platform is a mistake that's certain to lead to missed big data goals. It's also a mistake that's often difficult to rectify. Points to consider when designing a storage platform include estimating current and future utilization levels, specifying data-control mechanisms, and identifying the best data-migration approaches. It's also important to understand how well the planned platform will mesh with existing internal systems and whether it will be able to meet your organization's storage policies now and in the future.

Address replication

Replication is a potential stumbling point in any big data storage platform project. It's important to understand that some applications and databases have their own unique replication resource needs. To prevent the possibility of future headaches, it's important to recognize and accommodate these requirements during project planning.

Manage bandwidth and data flow

Underestimating bandwidth sizing is a pitfall that often leads to degraded performance. Likewise, overestimating bandwidth sizing can result in paying more for services than would otherwise be necessary.



Accurately calculating current and future bandwidth requires the careful study of several factors, including fallover requirements, the distance and number of hops the data must traverse, the amount of data being transmitted, and the number of users concurrently accessing the data. When calculating and balancing these numbers, it's better to slightly err on the side of specifying too much bandwidth and spending a little extra money, than to allocate too little bandwidth and potentially frustrate future big data users.

Paying careful attention to quality of service (QoS) and creating dedicated links between data centers enables better control over the data flow process while still reserving plenty of capacity for accommodating peak loads. Maintaining constant and consistent visibility of storage traffic flow and efficiency is another important key to maintaining consistently high big data performance.

Dedup intelligently

An intelligent deduplication strategy can cut waste and save money by unlocking significant amounts of valuable storage space. Ensuring that controllers are virtualization-ready is also important, as are technologies such as application virtualization, server virtualization, and virtual desktop infrastructure.

MORE BLOGS FROM JOHN EDWARDS

How Energy Companies Are Missing Big Data Opportunities

6 comments
Poor information management is crippling energy firms' ability to utilize big data to its full potential.

How Big Data Can Support Workforce Optimization

22 comments
Analytics allows businesses to scrutinize real-time employee data.

Boosting Fraud Detection Accuracy With Big Data

5 comments
Financial firms turn to big data tools to root out fraudsters.

Big Data Targets Alzheimer's Disease

9 comments
Researchers comb data to find possible causes and develop treatment paths.

More from John Edwards



FLASH POLL

In which vertical market do you work?

- Automotive
- Banking
- Consumer
- Education
- Energy
- Engineering
- Financial services
- Food and beverage
- Government
- Healthcare
- Insurance
- Manufacturing
- Media
- Real estate
- Religion
- Retail
- Technology

tools like application virtualization, server virtualization, and virtual desktop infrastructure (VDI). All of these tools help managers intelligently provision space and avoid the creation of capacity- and performance-robbing duplicates.

Final thought

It takes a considerable amount of time, thought, effort, and money to design and deploy an exceptional big data storage platform. The rewards, however, are potentially immense: fast performance, agility, and maximum uptime.

Related posts:

- [Dealing With the Dilemmas of the Data Warehouse](#)
- [Go Back to the Future With the Big Data Big Warehouse](#)
- [Can We Find Value in a Torrent of Data?](#)

[Email This](#) [Print](#) [Comment](#)

COMMENTS

[Newest First](#) | [Oldest First](#) | [Threaded View](#)



kiran, User Rank: Megabyte Messenger
10/10/2013 | 1:02:26 PM

Re: Flow hard

Yes thats understandable because the data keeps growing, the data we collect, on which we analyse, the data we use to make predicitions, they keep increasing in numbers and we might end up running out of space and our solution being a fail.

[Login](#)



50% 50%

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



netcrawl, User Rank: Petabyte Pathfinder
10/10/2013 | 9:41:08 AM

Re: Flow hard

@kiran it's a daunting task, one of the key issues in most data solutions is the inability to predict storage capacity, how to meet growing data volume. In the long run storage capacity could transforms into serious problem- storage scalability.

[Login](#)



50% 50%

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



kiran, User Rank: Megabyte Messenger
10/10/2013 | 7:08:18 AM

Re: Flow hard

There has to be thorough planning and understanding of what specifications we need of our big data platform. And early decisions are better with more reasoning and understandability rather than making wrong decisions that will make the storage difficult and of no use.

[Login](#)



50% 50%

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)



Saul Sherry, User Rank: Blogger
10/10/2013 | 3:52:55 AM

Flow hard

Managing that flow is the most difficult element - it almost takes a predictive analytics platform to imagine where you will be in 18 months time and how much bandwidth will be needed.

[Login](#)



50% 50%

That said, sensible measures and plans for growth should suffice unless you get a massive spike in activity because your brand or service suddenly becomes hugely popular... and who doesn't want a problem like that?

[Reply](#) | [Post Message](#) | [Messages List](#) | [Start a Board](#)

Technology

Telecommunications

Transportation

Other

[All Polls](#)

DIGITAL AUDIO



Big Data Storage Radio

Latest Archived Broadcast

Locking Up Your Cloud: Security in the Ether

Data security is precious enough when it's sitting in your office, so what's the best way to maintain security standards once you migrate to the cloud?

[Listen](#)

[Full schedule](#) [Archived shows](#) [Download to calendar](#)

BDR IN YOUR INBOX

InformationWeek reports
THE ULTIMATE RESOURCE FOR IT, BY IT
Unlimited access to comprehensive research and analysis reports

Top 5 Mobile Security Threats for 2012
State of Unified Communications

FEATURED VIDEO

Big Data Explained: What is a Data Mart?

1



Use this video to explain quickly to someone the difference between a data warehouse and a data mart.

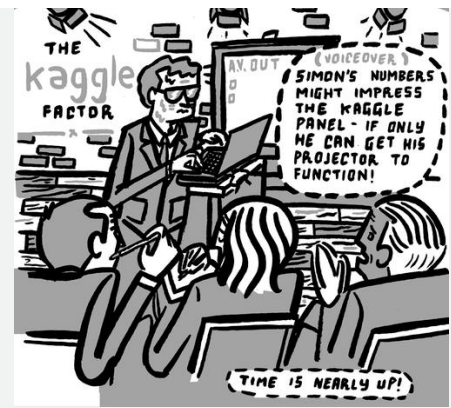
[Watch This Video](#)

[More Video Blogs](#)

CAPTION CONTEST

Write a Caption, Win a Gift Card

[Click Here](#)



Latest Comment: "The Kaggle Factor set a new viewing record of nearly 23.7 viewers. With ..."

19

Cartoon Archive

FOLLOW US ON TWITTER

Edu•Tek Ltd. @edutek_ltd 21m
Big Data in the Classroom
buff.ly/18AY4BV via
[@BigDataRepublic](https://twitter.com/BigDataRepublic) #edtech

TIBCO Software @TIBCO 47m
Is #BigData #analysis art or science?
Nick Evans of @jaywingsays
discusses in @BigDataRepublic:
bit.ly/18Srqn

Big Data Republic @BigDataRepublic2h
RT @data_nerd: Data Mining is about
explaining the past and predicting
the future. How can #analytics help
you? dld.bz/aqgf8...
Retweeted by Adma Raia
Expand

LIKE US ON FACEBOOK

Big Data Republic
Like You like this.

You and 3,257 others like Big Data Republic.

Facebook social plugin

ACCOLADES



Media Pioneer Award Winner



UBM TECH

OUR MARKETS: [Business Technology](#) | [Channel](#) | [Electronics](#) | [Game & App Development](#)

Working With Us: [Advertising Contacts](#) | [Event Calendar](#) | [Tech Marketing Solutions](#) | [Corporate Site](#) | [Contact Us / Feedback](#)

[Terms of Service](#) | [Privacy Statement](#) | Copyright © 2013 UBM Tech, All rights reserved