



IBM Software Group – Enterprise Networking Solutions  
z/OS V1R13 Communications Server

# IBM Multi-site Workload Lifeline

## Overview

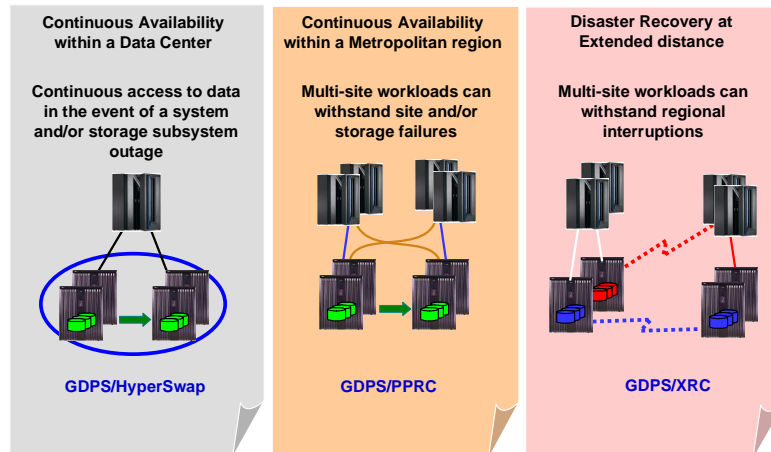


© Copyright International Business Machines Corporation 2011. All rights reserved.

This presentation provides an overview of the IBM Multi-site Workload Lifeline for z/OS® product.

## Continuous availability and disaster recovery

- What do z/OS customers do today?



Page 2

© Copyright International Business Machines Corporation 2011. All rights reserved.

Customers have three choices for recovering workloads in the event of a failure in their data center.

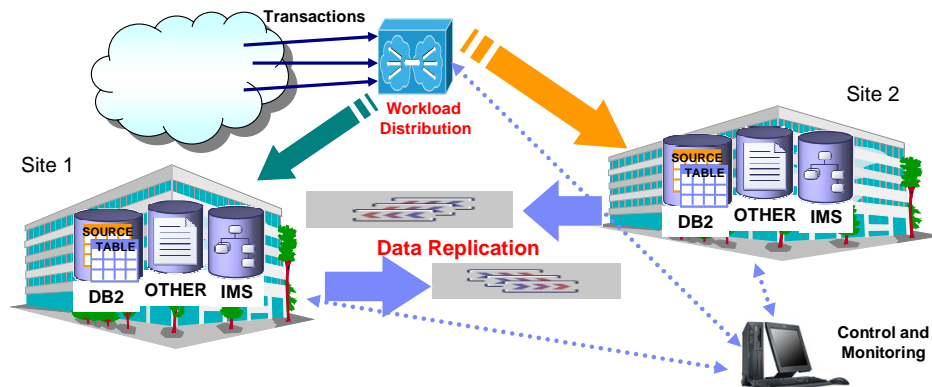
The first (leftmost) box represents a z/OS parallel sysplex that resides in a single site. It is using Geographically Dispersed Parallel Sysplex™ (GDPS®) along with HyperSwap® to provide recovery of storage subsystems using hardware disk mirroring. Continuous access to applications and data is possible, in the event of a system failure, application failure, or storage subsystem failure. However, access to the data is not possible in the event of a localized outage.

The second box represents a z/OS parallel sysplex that is split between two sites, a limited distance apart. It is using the Peer-to-Peer Remote Copy (PPRC) feature of GDPS to provide recovery of storage subsystems using hardware synchronous disk replication. Nearly continuous access to applications and data is possible in the event of a failure of one of the sites. Access to the data, after a site failure, can take an hour or more as applications are restarted on the remaining site. However, access to the data is not possible in the event of a regional outage.

The third box represents two z/OS parallel sysplexes that are located in sites that are an unlimited distance apart. It is using the Extended Remote Copy (XRC) feature of GDPS to provide recovery of storage subsystems using hardware asynchronous disk replication. Recovery of the applications and data is possible in the event of a regional outage. Access to the data, after a site failure, can take several hours as systems and applications are restarted on the recovery site. However, the data in the recovery site can only be used in the event of a failure in the primary site.

## GDPS Active-Active sites – what is it?

- Two or more sites, separated by *unlimited* distances, running the same applications
  - Having the same data to provide cross-site workload balancing and Continuous Availability / Disaster Recovery
- Paradigm shift: failover model => near continuous availability model



Page 3

© Copyright International Business Machines Corporation 2011. All rights reserved.

Although the current GDPS solutions provide good protection in the event of a site failure, there are limitations. Data can be accessed from only one site at a time, and all data written to disks is mirrored to the remote site. IBM Multi-site Workload Lifeline plays an integral role in the new disaster recovery solution called GDPS Active-Active Sites. The solution represents a shift from a failover model to a near continuous availability model.

In the diagram, the Workload Distribution box makes up the software and hardware that is required to perform load balancing of workloads. Multi-site Workload Lifeline determines the availability of the server applications that make up a workload, the capacity of the systems where the server applications reside, and other metrics. Multi-site Workload Lifeline uses this information to provide recommendations to two tiers of load balancers. The first tier of load balancers must be hardware appliances that support a protocol called SASP (which stands for Server/Application State Protocol that is described in RFC 4678). Multi-site Workload Lifeline provides recommendations to this first tier on which site new workload transactions should be directed. The second tier of load balancers can either be hardware appliances supporting SASP or the z/OS sysplex distributor. The second tier load balancers receive the workload transactions from the first tier load balancers and are then responsible for directing the transaction to one instance of the server application within the selected site. If the second tier load balancer is a hardware appliance, then Multi-site Workload Lifeline provides recommendations to this tier on which server application within the site is best able to handle to new workload transaction.

Referring back to the diagram, the Multi-Site Workload Lifeline (as part of the Workload Distribution box) provides distribution of workload requests between the sites. It can route around a failed site or around failed workloads on a site. Multi-site Workload Lifeline dynamically selects a site based on ability of site to handle additional workload, or based on custom configuration for the workload. The Data Replication box implements software-based replication to keep data sources in sync between sites, which allows the data to be accessed from either site, at unlimited distance between the sites. The Control and Monitoring box manages and monitors each workload defined in the configuration.

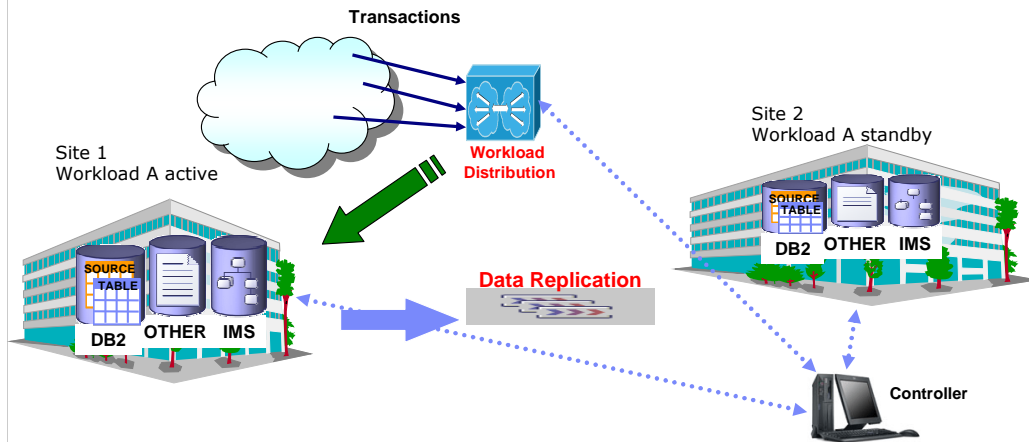
A workload is the aggregation of the these three components:

First, the software, or user written applications and the middleware run time environment (for example, a CICS® region and DB2® subsystem).

Second, the data sources, or related set of objects that must preserve transactional consistency (for example, DB2 Tables).

Third, the network connectivity, or the one or more TCP/IP addresses and ports that map the applications or middleware.

## Active/Standby configuration (1 of 2)

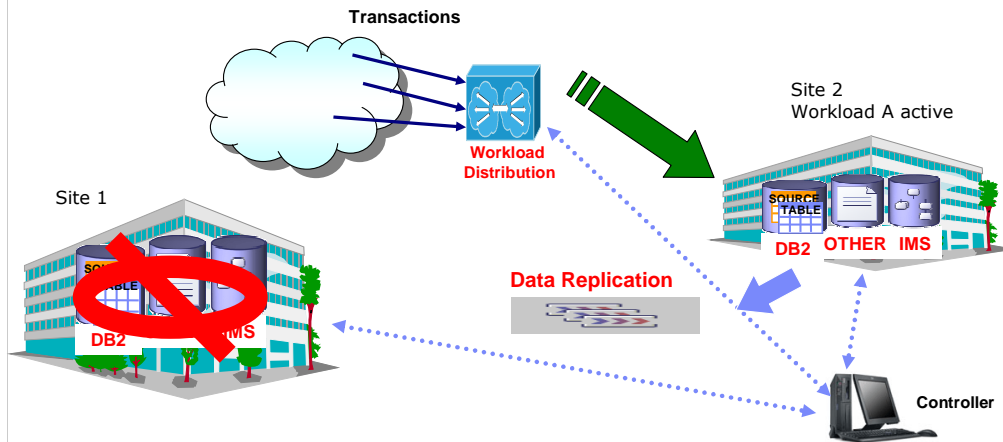


Page 4

© Copyright International Business Machines Corporation 2011. All rights reserved.

The IBM Multi-site Workload Lifeline product, as part of the Workload Distribution box, is an integral part of the initial GDPS Active-Active Sites solution, supporting the Active/Standby Configuration. In the diagram, a single workload is defined called workload A and is initially active on site 1. Multi-site Workload Lifeline is responsible for directing all incoming transactions for workload A to the server applications that reside in site 1. Software-based replication is responsible for keeping the data sources used by workload A in sync between site 1 and site 2.

## Active/Standby configuration (2 of 2)

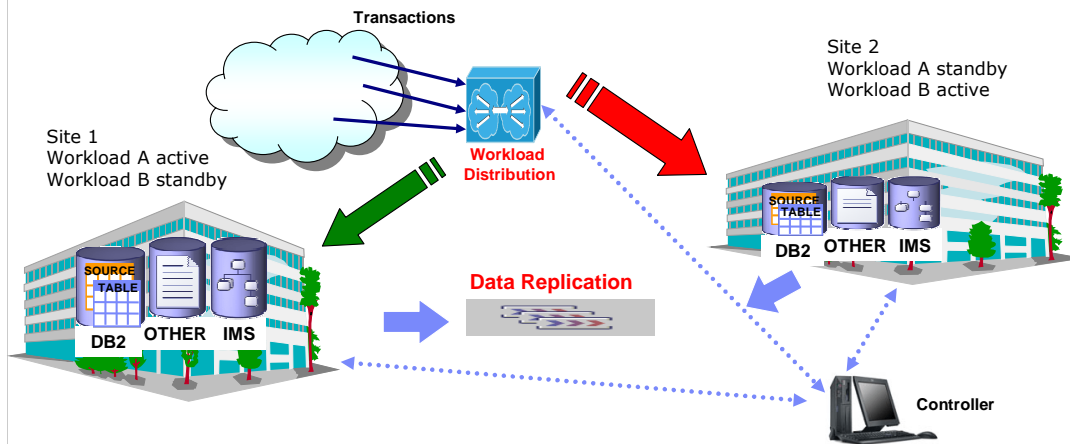


Page 5

© Copyright International Business Machines Corporation 2011. All rights reserved.

Continuing with the configuration on the prior slide, assume that site 1 encounters an unexpected outage. The IBM Multi-site Workload Lifeline product, as part of the Workload Distribution box, detects the site outage. Working with the controller, Multi-site Workload Lifeline will redirect all new transactions for workload A to site 2, which is now the active site for the workload. Software-based replication will queue any updates to the data sources on site 2 on behalf of workload A and forward them to site 1 once the site becomes available again.

## Active/Standby configuration (multiple workloads)



Page 6

© Copyright International Business Machines Corporation 2011. All rights reserved.

In the diagram, the IBM Multi-site Workload Lifeline product is able to manage multiple workloads. A workload called workload A is defined and active on site 1. A second workload called workload B is defined and active on site 2. Multi-site Workload Lifeline is responsible for directing all incoming transactions for workload A to the server applications that reside in site 1 and all incoming transactions for workload B to the server applications that reside in site 2. Software-based replication is responsible for keeping the data sources used by workload A and workload B in sync between site 1 and site 2. Multi-site Workload Lifeline can redirect either workload to its standby site in the event of a site outage or a manual operation to gracefully move a workload between sites.

## GDPS Active-Active sites workload distribution requirements (1 of 2)

- **Ability to distribute workloads between sites (and route around failed sites)**
  - Based on capacity/health of sites and server application instances within a site
- **Ability to detect workload or site failures**
- **Ability to switch workloads from one site to another site**
  - Perform “graceful” takeover for site maintenance

There are multiple requirements for workload distribution in a GDPS Active/Active Sites solution.

Transactions for a workload must be able to be directed to either site, based on the health and availability of the site and the server applications on that site.

If the server applications for a workload are no longer available on the active site, it must be detected and surfaced as a workload failure. If the active site for a workload is no longer available, it must also be detected and surfaced as a site failure.

Workloads must be able to be moved gracefully between sites in order to facilitate site-wide maintenance.

## GDPS Active-Active sites workload distribution requirements (2 of 2)

- **Ability to maintain workload configuration states in event of workload manager failure**
  - Keep a peer workload manager in sync with workload states
- **Ability to dynamically add/modify workloads**
- **Ability to surface distribution recommendations to network management applications**

Continuing with the requirements for workload distribution in a GDPS Active/Active Sites solution:

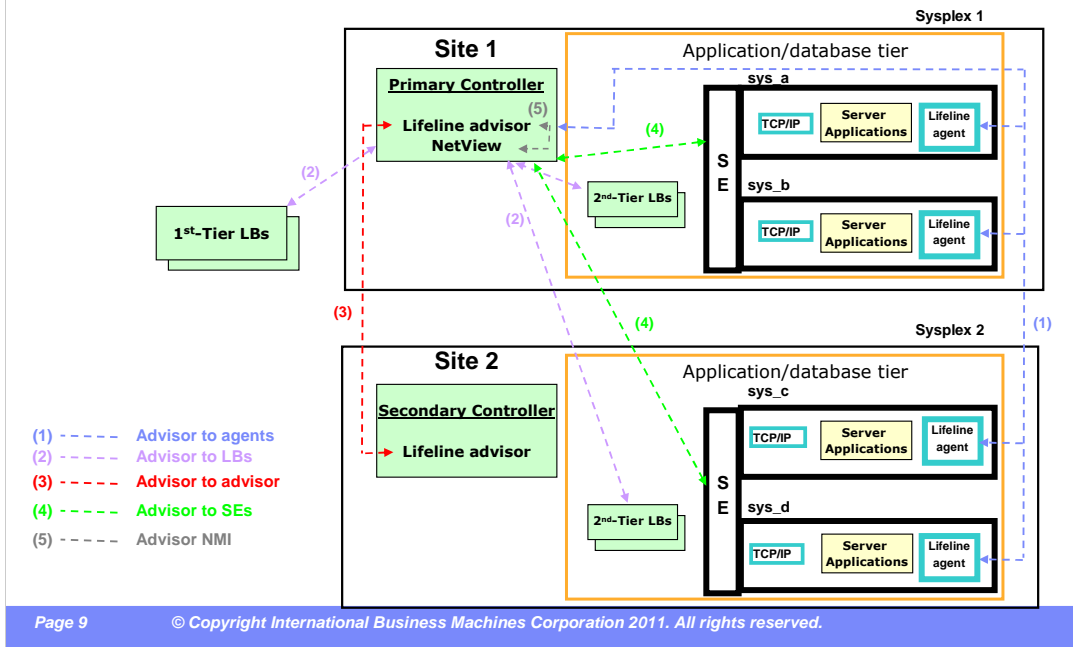
A failure in any component making up the workload distribution function must be recoverable by a backup component. This means maintaining state information about all workloads between the primary and backup components.

The workload distribution function must be able to handle workload additions dynamically without requiring a recycling of the workload distribution function.

All metrics used by the workload distribution function for determining how workload transactions are directed to a site must be made available to network management applications.



## Multi-site Workload Lifeline structure



Page 9

© Copyright International Business Machines Corporation 2011. All rights reserved.

The diagram describes the flows between the components and the functions for the IBM Multi-site Workload Lifeline product.

The first flow shows the communication between the lifeline advisor running on the primary controller and each of the lifeline agents running on each of the LPARs within both sites. The agents gather information about the LPAR where they are active and the server applications for workloads that reside on that LPAR. This information is communicated to the advisor which uses the information to generate distribution recommendations. The advisor also uses this information to determine when all server applications for a workload are no longer available and surface a workload failure.

The second flow shows the communication between the lifeline advisor and first and second tier load balancers. The advisor provides recommendations to the first tier load balancers about which site to route new transactions for a workload. The advisor can also provide recommendations to the second tier load balancers about which server application within the site to route a new transaction.

The third flow shows the communication between the lifeline advisor, called the primary advisor, on the primary controller and a backup lifeline advisor, called the secondary advisor, on the secondary controller. The primary advisor provides workload state information to the secondary advisor. In the event of a failure of the primary advisor, the secondary advisor will assume the role of primary advisor and maintain the current state of all workloads.

The fourth flow shows the communication between the lifeline advisor and service element on the central processor complexes (CPCs). The advisor uses this communication to determine the status of each LPAR where a workload can be active. If all LPARs that comprise a site are no longer available, the advisor can surface a site failure.

The fifth flow shows the communication between the lifeline advisor and a network management application, in this case, NetView®. The information the advisor maintains, such as LPAR health, server application availability, and workload states is available to network management applications using a documented network management interface.

## Workload Lifeline role in Active/Standby environment

- **Advisor provides distribution recommendations to multiple tiers of load balancers**
  - Server application health metrics and LPAR capacity information provided by agents running in all LPARs across both sites are used to build recommendations
- **Site recommendations to 1<sup>st</sup>-tier load balancers**
  - Direct 1<sup>st</sup>-tier load balancers to route new connections for a workload to a 2<sup>nd</sup>-tier load balancer within a site
  - Site selection determined by where the workload is currently active

The IBM Multi-site Workload Lifeline provides these functions in an Active/Standby configuration.

The advisor provides distribution recommendations to two tiers of load balancers. These recommendations are derived from information received from communication with agents active on LPARs that make up both sites. Information provided by the agents include the health and availability of the server applications for the workload, and the available capacity of the LPAR where the agent is running.

The advisor provides distribution recommendations to first tier load balancers about which site to direct new transactions. A given workload can only be active on one site, so the advisor will direct the first tier load balancer on which site a specific workload is active. The first tier load balancer will send the transactions for a workload to a second tier load balancer within the active site for that workload.

## Workload Lifeline in Active/Standby environment (1 of 4)

- **Server application recommendations to 2<sup>nd</sup>-tier load balancers**
  - Direct 2<sup>nd</sup>-tier load balancers to route new connections for a workload to specific server applications within the site
  - Server application selection determined by recommendations provided by the agents within the site
  - Sysplex Distributor can assume role of 2<sup>nd</sup>-tier load balancer
    - No server application recommendations provided by advisor in this case
- **Advisor provides ability to group different server applications into a workload**
  - Distinguish different workloads and perform different distribution decisions based on the workload (direct each workload to its Active site)

The IBM Multi-site Workload Lifeline provides these additional functions in an Active/Standby configuration.

The advisor provides distribution recommendations to second tier load balancers, provided they are hardware appliances supporting SASP. The advisor will provide recommendations about which server application instance to direct a new transaction. The advisor uses the information sent by the agents residing on the LPARs where the server applications are active. If the second tier load balancer is the z/OS sysplex distributor, the advisor does not provide distribution recommendations, as the sysplex distributor is able to derive these recommendations on its own.

Hardware appliances typically distribute transactions across multiple instances of the same server application. However, a workload can consist of different types of server applications that access the same data sources. The advisor can map these different types of server applications into a single workload, and provide matching distribution recommendations for each of the different server application types.

## Workload Lifeline in Active/Standby environment (2 of 4)

- **Advisor responsible for detecting workload failures**
  - Monitor the capacity of LPARs within a workload's Active site and availability/health of the server applications that make up the workload
  - Ability to dynamically switch a workload to an alternate site after detecting a failure
- **Advisor responsible for detecting site failures**
  - Monitor the availability and reachability of the LPARs that make up the site
  - Communication with agents active on the LPARs verifies IP network connectivity to the site
  - Communication with Service Elements (SE) over HMC network verifies LPAR status
  - Ability to dynamically switch all workloads to an alternate site after detecting a failure

Continuing the functions provided by the IBM Multi-site Workload Lifeline:

The advisor can detect when a workload fails. Using the information provided by agents, the advisor can determine the availability of the server applications within the site. If no server applications can be used on the site where the workload is active, the advisor will surface a workload failure. The advisor can be configured to automatically redirect the failed workload to the standby site.

The advisor can detect when a site fails. Using the information provided by communicating with service elements, the advisor can determine the availability of the LPARs within the site. If no LPAR is available on the site where the workload is active, the advisor will surface a site failure. The advisor can be configured to automatically redirect the active workloads on the failed site to the standby site.

## Workload Lifeline in Active/Standby environment (3 of 4)

- **Advisor communicates with a peer advisor**
  - Shares workload state information
    - A workload can be inactive
    - A workload can be active to a specific site
  - Peer advisor takes over responsibilities in the event the primary advisor fails
- **Advisor provides graceful movement of a workload to an alternate site (a 'planned' failure)**
  - Prevents new connections for the workload from being distributed to the Active site
  - Terminates any existing connections being distributed to the Active site
  - Reroutes new transactions to the alternate site

5 Additional functions provided by Multi-site Workload Lifeline in an Active/Standby configuration:

The advisor communicates with a peer advisor, called the secondary advisor, to share workload state information. The secondary advisor uses this communication to determine the availability of the primary advisor. If the primary advisor is not available, the secondary advisor can be configured to automatically become the primary advisor.

The advisor can gracefully move workloads from the active site to the standby site. It does this in several steps. First, it directs first tier load balancers to stop distributing new transactions for the workload. Second, it terminates existing connections for the workload. Finally, it directs first tier load balancers to start distributing new transactions for the workload to the standby site.

## Workload Lifeline in Active/Standby environment (4 of 4)

- **Advisor has ability to dynamically add or modify existing workloads to an active configuration**
  - Allows changes without recycling the advisor
- **Advisor provides Network Management Interface (NMI)**
  - Surface workload states, distribution recommendations, and component information to network management agents
- **Agents communicate with a Communications Server TCPIP stack**
  - Extracts information about available server applications and server application health on the LPAR where agent is active

Remaining functions provided by Multi-site Workload Lifeline in an Active/Standby configuration:

The advisor can be dynamically updated to learn about updates to existing workload configuration or new workload configuration. This ensures the advisor continues to provide distribution recommendations without interruption by having to recycle the advisor.

The advisor also will surface information the advisor it maintains. For example, information such as LPAR health, server application availability, and workload states is made available to network management applications using a documented network management interface.

Finally, agents on each LPAR communicate with the Communications Server TCP/IP stack to determine whether server applications are available and their health. The agents send this information to the advisor to be used for generating distribution recommendations.

## Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?
- Did it help you solve a problem or answer a question?
- Do you have suggestions for improvements?

Click to send email feedback:

[mailto:iea@us.ibm.com?subject=Feedback\\_about\\_overview.ppt](mailto:iea@us.ibm.com?subject=Feedback_about_overview.ppt)

This module is also available in PDF format at: [../overview.pdf](#)

You can help improve the quality of IBM Education Assistant content by providing feedback.

## Trademarks, copyrights, and disclaimers

IBM, the IBM logo, ibm.com, CICS, DB2, GDPS, Geographically Dispersed Parallel Sysplex, HyperSwap, NetView, Parallel Sysplex, and z/OS are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of other IBM trademarks is available on the web at "[Copyright and trademark information](http://www.ibm.com/legal/copytrade.shtml)" at <http://www.ibm.com/legal/copytrade.shtml>

Other company, product, or service names may be trademarks or service marks of others.

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. WHILE EFFORTS WERE MADE TO VERIFY THE COMPLETENESS AND ACCURACY OF THE INFORMATION CONTAINED IN THIS PRESENTATION, IT IS PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. IN ADDITION, THIS INFORMATION IS BASED ON IBM'S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. NOTHING CONTAINED IN THIS PRESENTATION IS INTENDED TO, NOR SHALL HAVE THE EFFECT OF, CREATING ANY WARRANTIES OR REPRESENTATIONS FROM IBM (OR ITS SUPPLIERS OR LICENSORS), OR ALTERING THE TERMS AND CONDITIONS OF ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS OR SOFTWARE.

© Copyright International Business Machines Corporation 2011. All rights reserved.