IBM eServer™

Automated Domain Name Registration: Concepts

e business on demand software

Automated DNS Registration of Application-specific Hostnames

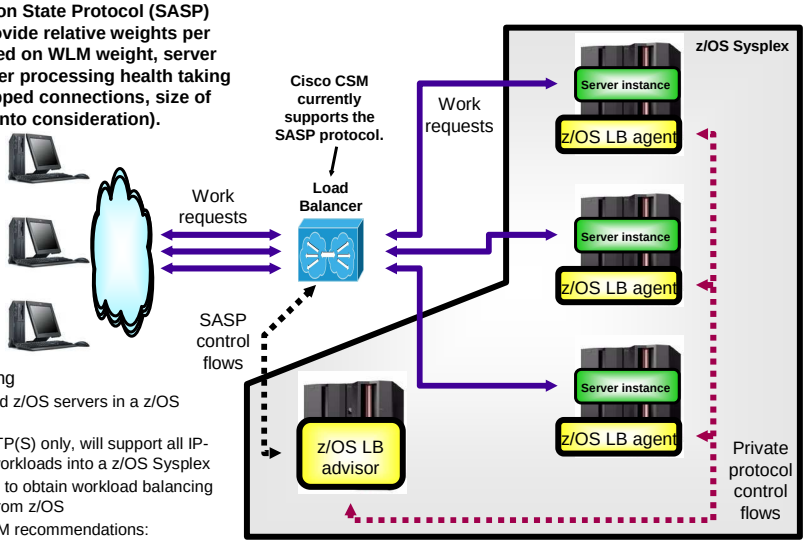
- **Automated DNS Registration of Application-specific Hostnames**
  - New z/OS® application, also known as Automated Domain Name Registration (ADNR)
  - Dynamically add/delete application-specific hostnames and the addresses of those applications in name servers according to application availability
  - Dynamically add/delete hostnames and their addresses representing a TCP/IP stack, MVS™ system, or subset of IP addresses
    - Use these traditional hostnames for occasions when you are not connecting to a server (like ping or traceroute)
  - Provide a migration path for DNS/WLM users
    - DNS/WLM and BIND 4.9.3 support to be removed in a future release
    - Remove restriction of 15 IPv4 addresses
    - Add support for IPv6 addresses
    - Participating applications no longer required to register with WLM
    - Allows name server to exist anywhere in the network, including z/OS
    - Capable of round-robin distribution only. Consider migrating to Sysplex Distributor or z/OS Load Balancing Advisor if load balancing is your primary goal and round-robin is not sufficient.

  - Leverage the z/OS Load Balancing Advisor
    - Required for ADNR
    - ADNR appears as a load balancer to the Advisor application
    - Load balancing with external load balancers can coexist with ADNR

Review: z/OS Load Balancing Advisor (LBA)

The Server Application State Protocol (SASP) control flows will provide relative weights per server instance (based on WLM weight, server availability, and server processing health taking such metrics as dropped connections, size of backlog queue, etc. into consideration).

SASP is also used by the new eWLM infrastructure.



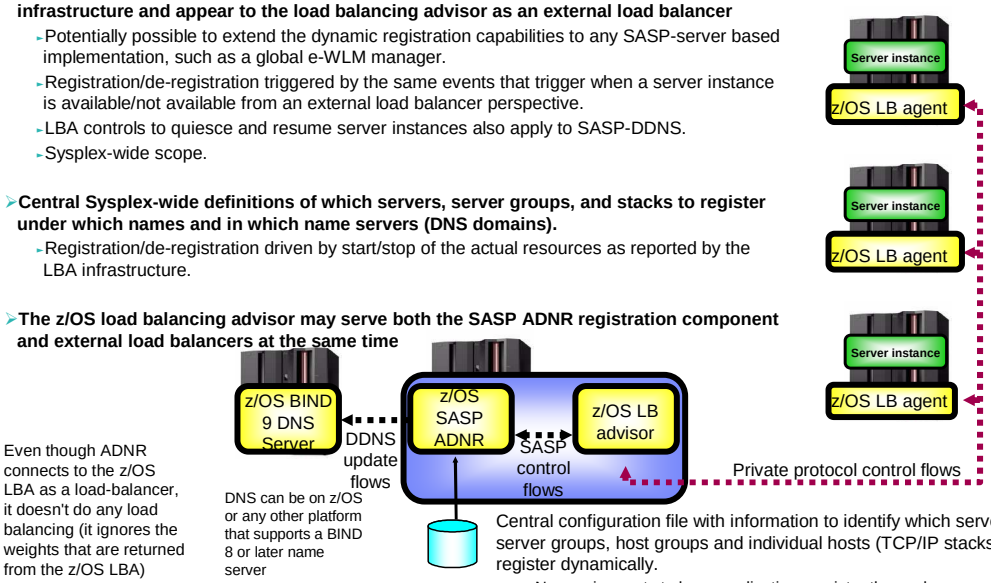
z/OS workload balancing

- Support for clustered z/OS servers in a z/OS Sysplex
- Not focused on HTTP(S) only, will support all IP-based application workloads into a z/OS Sysplex
- Allows external LBs to obtain workload balancing recommendations from z/OS
- Based on z/OS WLM recommendations:
  - System level recommendations based on displaceable capacity
  - Or Server specific recommendations (based on how well individual servers are performing)

The z/OS Load Balancing Advisor technology is a z/OS Communications Server technology that was made generally available in 4Q2004 via APARs PQ90032 (V1R4) and PQ96293 (V1R5/V1R6). It is fully integrated into z/OS V1R7.

Replacing the dynamic DNS registration part of the DNS/WLM component with a dynamic DNS solution

- **ADNR registration component will use existing z/OS load balancing advisor infrastructure and appear to the load balancing advisor as an external load balancer**
  - Potentially possible to extend the dynamic registration capabilities to any SASP-server based implementation, such as a global e-WLM manager.
  - Registration/de-registration triggered by the same events that trigger when a server instance is available/not available from an external load balancer perspective.
  - LBA controls to quiesce and resume server instances also apply to SASP-DDNS.
  - Sysplex-wide scope.
- **Central Sysplex-wide definitions of which servers, server groups, and stacks to register under which names and in which name servers (DNS domains).**
  - Registration/de-registration driven by start/stop of the actual resources as reported by the LBA infrastructure.
- **The z/OS load balancing advisor may serve both the SASP ADNR registration component and external load balancers at the same time**



Even though ADNR connects to the z/OS LBA as a load-balancer, it doesn't do any load balancing (it ignores the weights that are returned from the z/OS LBA)

DNS can be on z/OS or any other platform that supports a BIND 8 or later name server

Central configuration file with information to identify which servers, server groups, host groups and individual hosts (TCP/IP stacks) to register dynamically.

- No requirements to have applications register themselves

How sysplex resources are represented in DNS

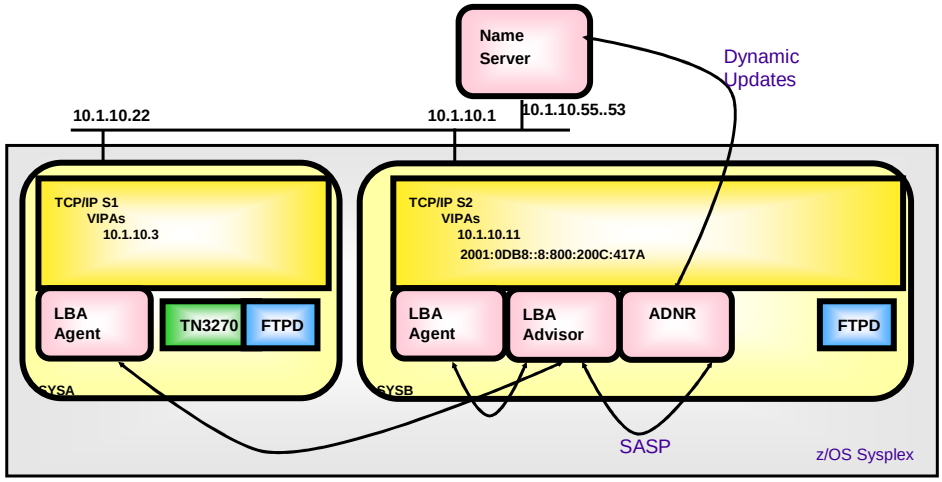
- **Each sysplex resource is represented by one or more Resource Records (RRs) in the name server**
- **Most RRs map a resource name to an IP address, or an IP address to a resource name**
  - ADNR only adds the type of RRs that map a resource name to an IP address
  - ADNR does NOT add the type of RRs that map an IP address to a resource name
    - DNS/WLM didn't do so either
  - For RRs added by ADNR a resource name can be:
    - The name of a group of equivalent server applications or a single server application (such as tn3270)
    - The name of a particular instance of one of these server applications (such as sysa.tn3270)
    - A name representing any TCP/IP stack or system in the sysplex (such as prodplex)
    - A name representing a particular TCP/IP stack or system in the sysplex (such as sysa)
    - A name representing any desired set of IP addresses (such as sysplexxcf)
- **Resources mapped to IPv4 addresses are represented in DNS with an 'A' record**
- **Resources mapped to IPv6 addresses are represented in DNS with an 'AAAA' (quad-A) record**
- **Some example RRs added by ADNR follow on the next page**

Example DNS resource records added by ADNR

Comments (delimited by a semicolon) are not added by ADNR. These were added here only for illustration.

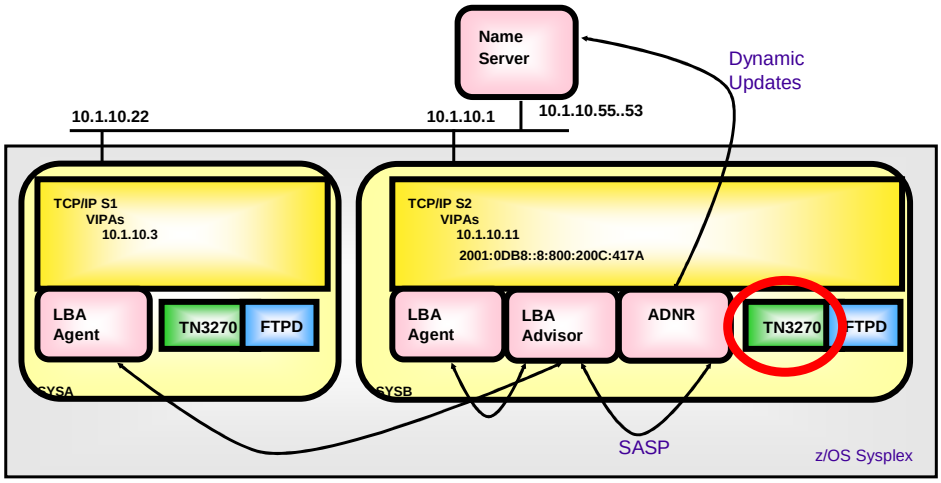
```
; Added due to a host_group statement
; NAME                          CLASS  TYPE    RECORD DATA      COMMENTS
; ------------------------------------------------------------------------------------------
prodplex.mvsplex.mycorp.com     IN     A       10.1.10.3        ;name for any IP address in sysplex
prodplex.mvsplex.mycorp.com     IN     A       10.1.10.11       ;name for any IP address in sysplex
prodplex.mvsplex.mycorp.com     IN     AAAA    2001:0DB8::8:800:200C:417A ;name for any IP addr in sysplex
sysa.mvsplex.mycorp.com         IN     A       10.1.10.3        ;name for an IP address on 'sysa'
sysb.mvsplex.mycorp.com         IN     A       10.1.10.11       ;name for an IP address on 'sysb'

;Added due to a server_group statement
; NAME                          CLASS TYPE   RECORD DATA   COMMENTS
; ------------------------------------------------------------------------------------------
tn3270.mvsplex.mycorp.com       IN     A    10.1.10.3     ;name for any instance of tn3270 in sysplex
tn3270.mvsplex.mycorp.com       IN     A    10.1.10.11    ;name for any instance of tn3270 in sysplex
sysa.tn3270.mvsplex.mycorp.com  IN     A    10.1.10.3     ;name for tn3270 instance on 'sysa'
sysb.tn3270.mvsplex.mycorp.com  IN     A    10.1.10.11    ;name for tn3270 instance on 'sysb'
```

ADNR system example



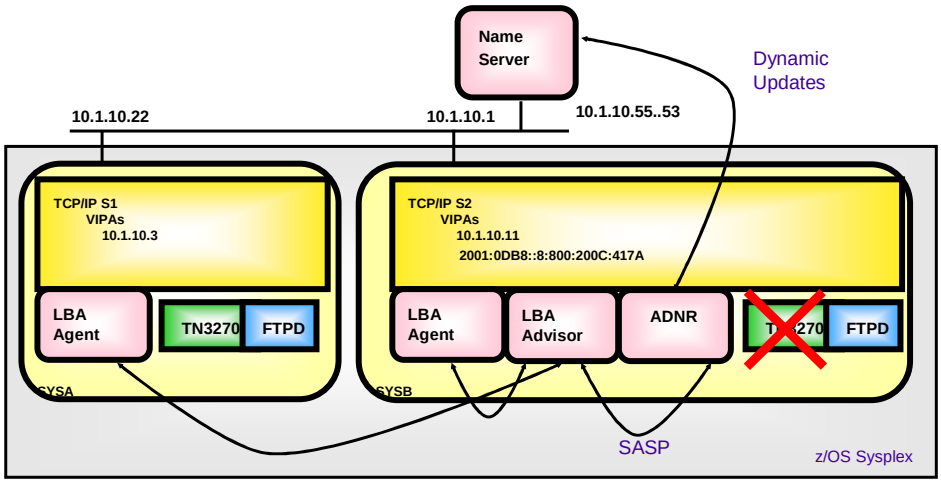

```
Selected name server contents:
   tn3270.mvsplex.mycorp.com                    10.1.10.3
   sysa.tn3270.mvsplex.mycorp.com               10.1.10.3
   ftp.mvsplex.mycorp.com                                  10.1.10.3,
 10.1.10.11
   sysa.ftp.mvsplex.mycorp.com                  10.1.10.3
   sysb.ftp.mvsplex.mycorp.com                  10.1.10.11
```

ADNR system example (cont.) - Server instance becomes available




```
Selected name server contents:
   tn3270.mvsplex.mycorp.com                10.1.10.3, 10.1.10.11
   sysa.tn3270.mvsplex.mycorp.com           10.1.10.3
   sysb.tn3270.mvsplex.mycorp.com           10.1.10.11
   ftp.mvsplex.mycorp.com                   10.1.10.3, 10.1.10.11
   sysa.ftp.mvsplex.mycorp.com              10.1.10.3
   sysb.ftp.mvsplex.mycorp.com              10.1.10.11
```

ADNR system example (cont.) - Server instance becomes unavailable




```
Selected name server contents:
   tn3270.mvsplex.mycorp.com                10.1.10.3, 10.1.10.11
   sysa.tn3270.mvsplex.mycorp.com           10.1.10.3
   sysb.tn3270.mvsplex.mycorp.com           10.1.10.11
   ftp.mvsplex.mycorp.com                   10.1.10.3, 10.1.10.11
   sysa.ftp.mvsplex.mycorp.com              10.1.10.3
   sysb.ftp.mvsplex.mycorp.com              10.1.10.11
```

How sysplex resources are represented in ADNR configuration

- **Sysplex resources in the ADNR configuration file are represented by groups**
  - Two types of groups
    - Groups which represent equivalent server applications (server_group) or a single server
    - Groups which represent a TCP/IP stack, MVS system, or subset of sysplex addresses (host_group)
- **Groups designate a port and protocol (e.g. TCP, UDP) and contain:**
  - Members which specify one or more IP addresses
- **Groups representing a TCP/IP stack, MVS system or subset of sysplex addresses (host_group) always have an implied port of zero and an implied protocol of zero**
- **Groups representing server applications specify the port the servers listen on and the TCP/IP protocol used to connect to that port**
- **Therefore, a sysplex resource is ultimately represented by the triplet of:**
  - Port +
  - Protocol +
  - IP address
- **Conceptually, ADNR has this type of organization:**
  - Group1 (implicitly or explicitly defines port and protocol)
    - Member1 = IP Address1, IP Address2
    - Member2 = IP Address3, IP Address4, IP Address5
  - Group2 (implicitly or explicitly defines port and protocol)
    - Member3 = IP Address6

How sysplex resources are represented by the z/OS LBA Advisor/Agent

- **"GWM" = Global Workload Manager = SASP terminology for the LBA Advisor**

- **Group**
  - LBA Advisor/Agent keep the same group concept of the ADNR group
  - All ADNR resources configured within a group are kept within the same group in LBA
    - Members are subordinate to a group

- **Member**
  - LBA member is slightly different than an ADNR member
  - LBA member represents one sysplex resource and is the triplet of:
    - Port
    - Protocol
    - IP address
  - ADNR member is a collection of one or more sysplex resources

- **Conceptually, LBA Advisor/Agent has this type of organization:**
  - Group1
    - Member1 = port1, prot1, IP Address1
    - Member2 = port1, prot1, IP Address2
  - Group2
    - Member3 = port2, prot1, IP Address3
    - and so on.

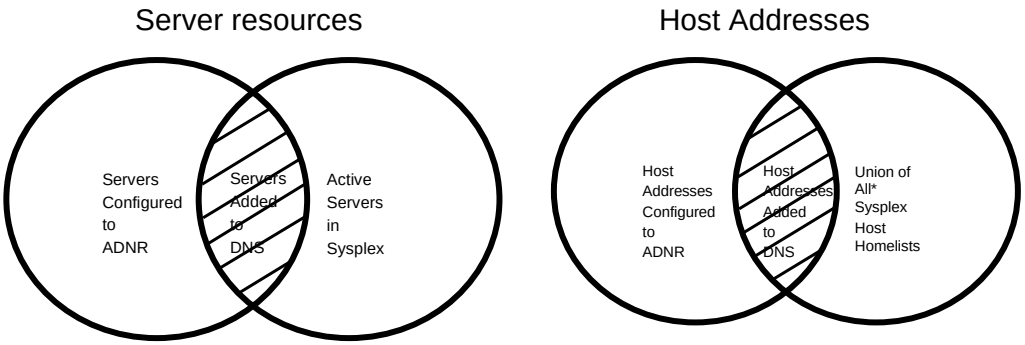
Which resources are dynamically added to DNS?

- **Resources configured to ADNR and which are available in the sysplex**
- **Two sysplex resource types**
  - Hostnames representing the host itself or a TCP/IP stack (host_group)
  - Hostnames representing a group of equivalent server applications (server_group) or a single server
- **Resources to be dynamically added to DNS must be configured to ADNR**
  - ADNR registers resources to LBA Advisor
  - LBA Agent must be running on each sysplex host
  - LBA Agents/Advisor inform ADNR of resource availability changes
  - ADNR update-adds/update-deletes resources according to current availability
- **Only available resources are update-added to the name server**
  - Active server applications (server_group)
    - Listening TCP server applications
    - UDP server applications
  - IP addresses in the sysplex host TCP/IP stack Home lists (host_group)

When are resources dynamically removed from DNS?

- **Resources removed from a name server (update-deleted)**
  - Whenever resource is unavailable from the LBA Advisor's perspective
  - Resource Record (RR) update-deleted from name server
  - RR with same hostname may still exist with a different IP address
- **For both sysplex resource types (host_group and server_group), RRs removed when...**
  - Configuration changes
    - IP address removed from Home list (VARY OBEY)
    - ADNR configuration changed to remove group/member/IP address/dns/zone
  - Operational changes
    - Target server application's TCP/IP stack stopped
    - LBA Agent on MVS system stopped
    - Member quiesced by LBA Agent operator command
- **For server names (server_group)...**
  - All of the above, plus when...
  - Target server application is stopped, crashes, and so on.

Servers and host addresses added to name server



*Certain types of addresses are restricted. See the following page.

Restricted addresses

- **Addresses which cannot be used by ADNR**
  - Distributed Dynamic VIPAs
  - Loopback
  - Unavailable IPv6 addresses
  - Deprecated IPv6 addresses
- **Addresses which should not be used by ADNR**
  - Autoconfigured IPv6 addresses, including link-local
    - Lack of predictability for ADNR configuration
    - Interface failure makes endpoint unreachable
    - Clients will not be able to use link-local addresses to connect to endpoints

VIPA types and uses

- **Static VIPAs vs. dynamic and distributed dynamic VIPAs**
  - Static VIPAs
    - Work with ADNR
    - ADNR offers a unique type of virtualization access to applications using static VIPAs
    - Static VIPAs provide fault tolerence for interface failures
    - ADNR extends this fault tolerence for server, TCP/IP stack, or MVS system failure
  - Dynamic VIPAs
    - Work with ADNR
    - Better suited for alternative virtualization approaches like VIPA takeover, unique application-instance DVIPA
  - Distributed dynamic VIPAs
    - Cannot work with ADNR
    - Used with sysplex distributor
- **Summary**
  - Static VIPAs
    - ADNR can offer unique virtualization advantages
  - Dynamic VIPAs and distributed dynamic VIPAs
    - Use for alternative virtualization approaches

Name server location

- **The name server or servers which ADNR updates may be located anywhere in the network, including on z/OS**
- **The local name server which a client queries to resolve the names of sysplex resources can be anywhere in the network.**
  - Assumes the domain containing the sysplex resources has been properly delegated from its parent domain.
    - e.g. The name server that is authoritative for the mycorp.com domain needs to delegate the mvsplex.mycorp.com domain to the name server ADNR is updating.

Update frequency

- **How closely the name servers reflect the true availability of status of sysplex resources depends largely upon LBA Advisor configuration**
  - update_interval statement
  - Range is 10 seconds to 10 minutes. Default is 60 seconds.
  - Changes in sysplex resource availability status pushed to ADNR at least every "update_interval" (also see Convergence Period, below)
  - Changes from available to unavailable may occur more quickly
  - Convergence Period: During ADNR initialization or after a F REFRESH, name servers are not updated until 2 X the update_interval, then they occur at least every update_interval.

- **Tradeoffs of higher vs. lower update frequency.**

With more frequent updates...

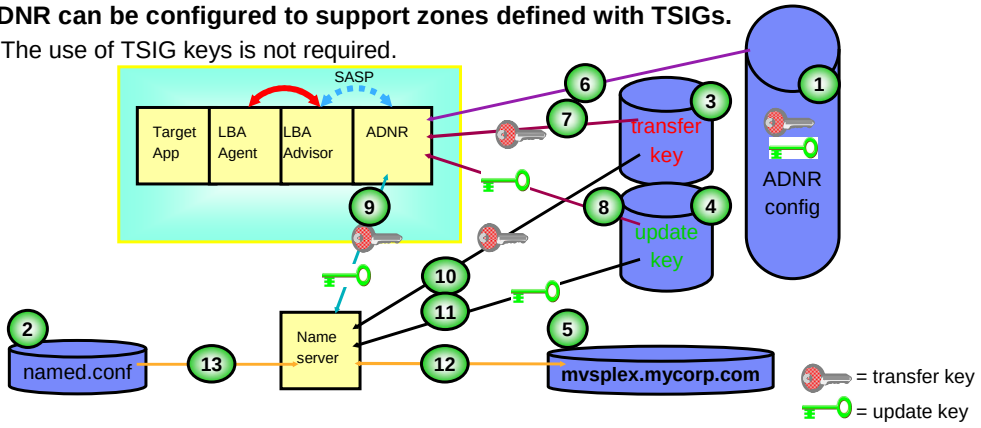
  - Benefits
    - True application availability status approaches real-time, reducing latency between when changes appear on the target application and when the name server is updated.
  - Costs
    - More frequent network flows
      - Between the LBA Advisor and the LBA Agents
      - Between the LBA Advisor and ADNR
    - Additional cycles consumed by LBA Advisor and ADNR

Caching issues

- **Name servers and some client resolvers may cache DNS responses for a period of time**
- **Near real-time availability information provided by ADNR can be defeated by prolonged caching of DNS responses**
  - e.g. A sysplex server becomes unavailable and the ADNR name server is updated indicating such. However, a client resolver has already cached the address of the unavailable server. If a connection attempt is made to the equivalent group of servers before the cache expires, the client connection will fail.
- **Caching time is usually governed by the Time To Live (TTL)**
- **Each resource record (RR) in a name server has a TTL**
- **TTL of RRs added by ADNR is determined either by ADNR configuration, or by LBA Advisor configuration**
  - TTL can be configured on a zone basis by ADNR configuration, or
  - TTL will be determined by the LBA Advisor's update_interval which is the default
  - Allowing the TTL to be determined by the update_interval is recommended

ADNR does support DNS transaction signature

- **A transaction signature (TSIG) is used to prevent an unauthorized zone transfer and dynamic update.**
- **A transaction signature is based on an IETF standard (RFC 2845).**
- **ADNR can be configured to support zones defined with TSIGs.**
  - The use of TSIG keys is not required.



Trademarks, copyrights, and disclaimers

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

MVS z/OS

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements or changes in the products or programs described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

Information is provided "AS IS" without warranty of any kind. THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (for example, IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

IBM makes no representations or warranties, express or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.