



IBM Software Group

WebSphere High Availability based on System Automation for OS/390



Jean-Gabriel WEYER

jg_weyer@fr.ibm.com



Agenda

- **Pourquoi automatiser WebSphere Application Server ?**
- **« overview » des composants WebSphere**
- **« overview » de System Automation for OS/390**
- **Haute disponibilité de WebSphere à l'aide de System Automation**
- **Scenarios de recovery WebSphere**
- **Synthèse**



Pourquoi automatiser (1/2) ?

❑ Simplifier l'administration :

Regrouper les composants et gérer les dépendances

Automatiser le démarrage, les reprises, l'arrêt
Monitorer les messages et y répondre

Afficher l'état des composants et des groupes

➤ réduire les coûts

❑ Réduire l'indisponibilité

Reprise automatique

Gestion du sysplex

Déplacement des composants et des groupes

Assurer les objectifs de disponibilité

➤ Accroître la disponibilité des applications



Pourquoi automatiser WAS ?

❑ Simplifier la gestion

WAS est construit sur les bases éprouvées de z/OS, mais relativement complexes en terme de composants requis

WAS apporte un haut niveau de performance et d'évolutivité (scalability) avec toutefois:

une infrastructure complexe

plusieurs régions de contrôle et de serveur

❑ Démarrer, monitoring, les reprises, l'arrêt, de nombreux composants, automatiquement

❑ Réduire l'indisponibilité

Le e-business accroît la visibilité

Internet permet l'ouverture au public et à la demande sur une base de disponibilité 24x365

❑ z/OS & WAS recovery de base

Sysplex Distributor

IP backup IP avec DVIPA

❑ **L'automatisation est toujours nécessaire :**

- **Redémarrage automatique rapide des composants défectueux**
- **Déplacement automatique dans le sysplex si le recovery n'est pas possible**
- **Gestion des dépendances dans le recovery**
- **Nettoyage après les pannes**



WebSphere Application Server: z/OS Components

- ❑ z/OS apporte une plate forme stable, avec des fonctions de recovery et de sécurité intégrées pour WebSphere Application Server
 - TCP/IP pour la communication
 - Support Unix Systems Services avec HFS pour les fichiers
 - LDAP et DB2 pour les services de directory et de database
 - WLM pour l'équilibrage de charge
 - RACF pour la sécurité
 - RRS pour le recovery
- ❑ WebSphere Application Server for OS/390 and z/OS:
 - L'environnement de base apporte l'infrastructure
 - Les serveurs J2EE et MOFW contrôlent et font tourner les applications J2EE web ou les applications CORBA.



WebSphere Application Server: Server Components

- **Base Environment**

- Daemon**

- Initialise l'environnement WAS.

- Région de contrôle uniquement

- Systems Management Server (SMS)**

- Contrôle de la configuration WAS

- Contrôle de la région ou des régions à serveurs multiples démarrés par WLM (si nécessaire)

- Naming Server**

- Identifie les objets Java, tels que les EJBs

- Apporte le support opérationnel pour les serveurs J2EE et Corba (MOFW)

- Contrôle de la région ou des régions à serveurs multiples démarrés par WLM (si nécessaire)

- Interface Repository Server (IR)**

- Gère les interfaces avec les objets Corba

- Contrôle de la région ou des régions à serveurs multiples démarrés par WLM (si nécessaire)

- **Application Server Environment**

- J2EE servers**

- Environnement « run-time » pour les composants J2EE, tels que servlets ou EJBs

- Supporte Java uniquement

- Contrôle de la région ou des régions à serveurs multiples démarrés par WLM (si nécessaire)

- CORBA (MOFW) servers**

- Gestion des objets java

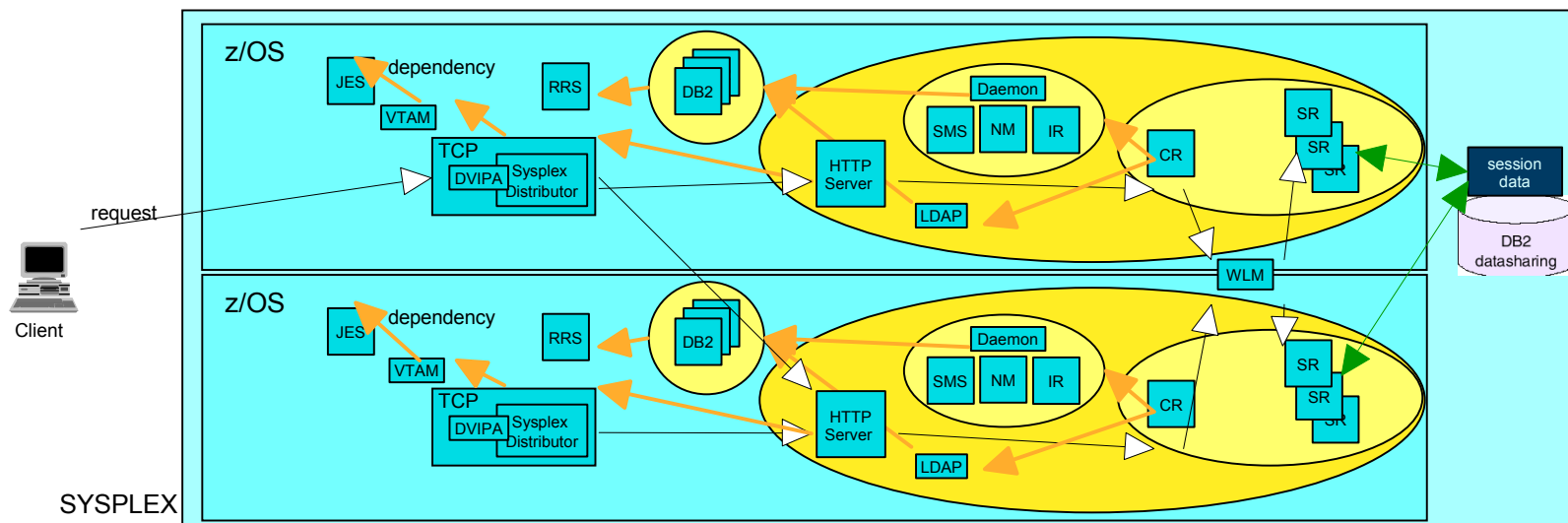
- Support de différents langages

- Contrôle de la région ou des régions à serveurs multiples démarrés par WLM (si nécessaire)

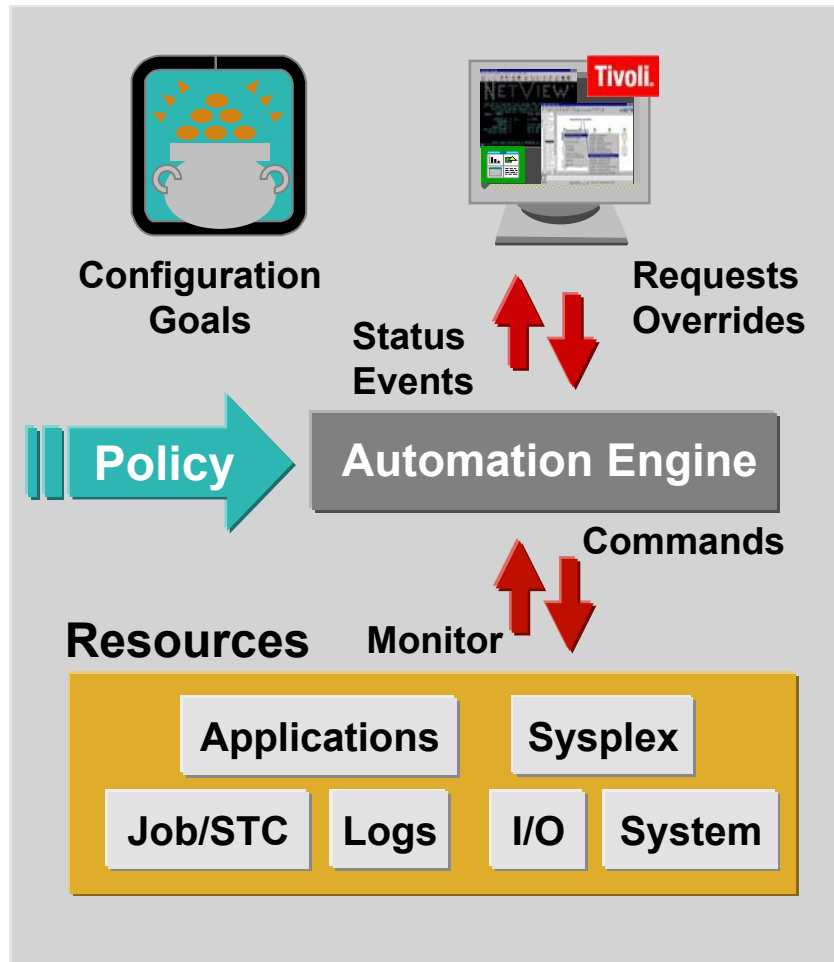


Avant de démarrer WAS...

- Avant de démarrer WAS, les sous systèmes suivants doivent être démarrés:
 - TCP/IP
 - Resource Recovery Services (RRS)
 - DB2
 - Lightweight Directory Access Protocol Server (LDAP)
- Les environnements applicatifs WLM doivent apparaître en état AVAILABLE.
- Avant de démarrer votre WAS, vous devez vous assurer que tous les sous-systèmes pré-requis sont disponibles



System Automation for OS/390 (SA OS/390)



- **Automation**

- Démarrage, reprises, arrêt

- Gère les applicatifs et les sous-systèmes

- Automatisation des tâches opérateurs

- Surveillance des messages et réponse

- Prévention des défaillances de ressources critiques (WTO buffers, spool)

- Automatisation CICS, IMS, DB2 & TWS

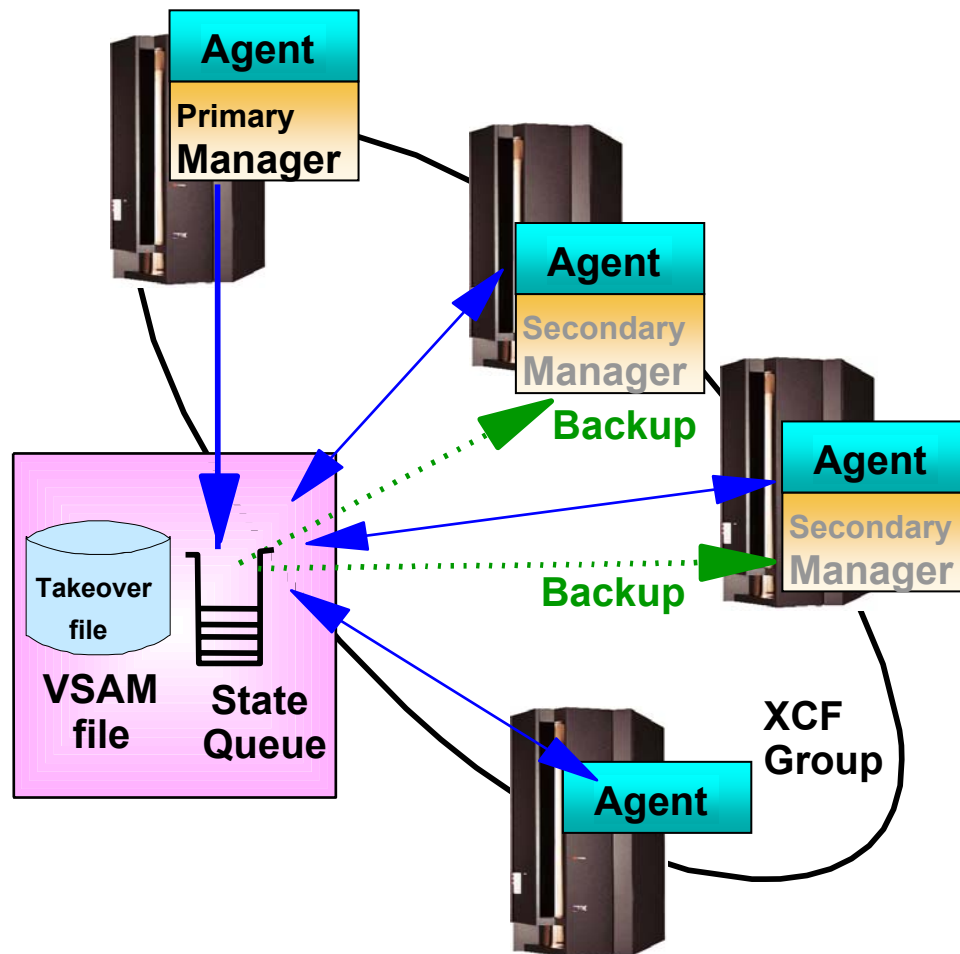
- **Interface graphique**

- Applications, systèmes, événements, ressources critiques du Sysplex

- Interface de commandes



SA OS/390 V2 Overview: concept Manager/Agent



- **Sysplex-wide automation**

- Distribué autant que possible, centralisé autant que nécessaire
 - Applicable aussi à un système unique
 - Regroupement de ressources en applications
 - Gestion par objectifs
 - Capacité de gestion des relations inter-systèmes

- **Centre de décisions automatisé**

- Connaissance centralisée des ressources et de leurs dépendances
 - Chaque ressource a un identifiant unique
 - eg: TSO/APL/SYS1 or CICS/APG
 - 5 états pour chaque ressources
 - observed, desired, automation, startability, compound
 - Prise en compte de l'état observé/demandé pour gérer les actions automatiques

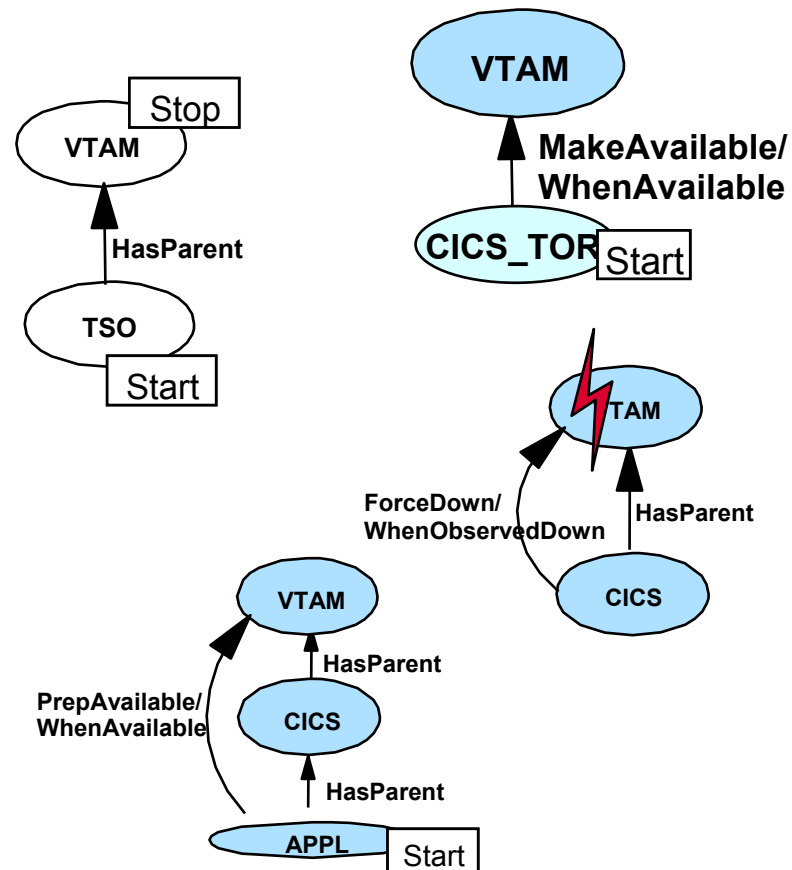
- **Automatisation locale par agents**

- Infrastructure NetView
 - Traitement des messages

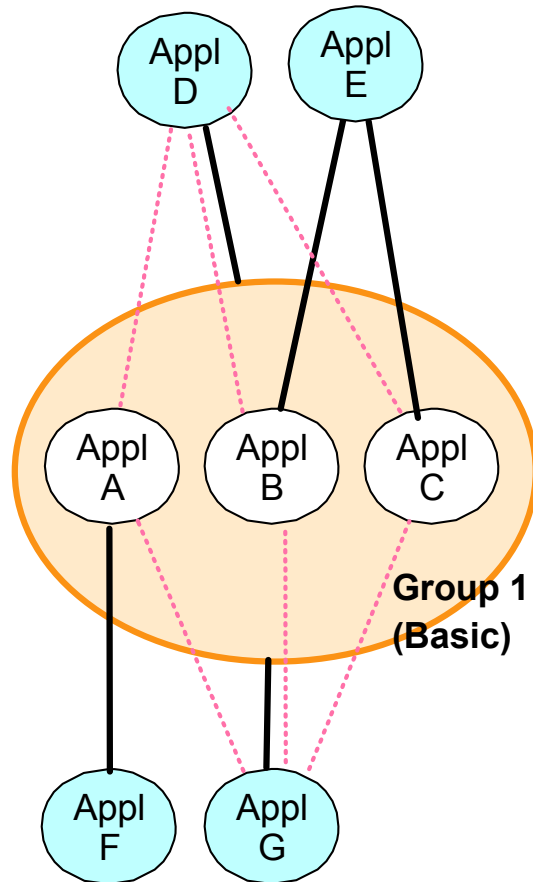


SA OS/390 gestion élaborée des relations

- séquences de démarrage et d'arrêt
- la relation définit comment une ressource est reliée à une autre
 - HasParent
 - MakeAvailable, MakeUnavailable
 - PrepAvailable, PrepUnavailable
 - ForceDown
- Conditions associées à une relation
 - WhenAvailable, WhenUnavailable
 - WhenAvailableOrStarting
 - WhenUnavailableOrStopping
 - WhenObservedDown
- options d'automatisation
 - Active vs Passive
- Les relations sont unidirectionnelles
- Les relations peuvent être définies inter systèmes



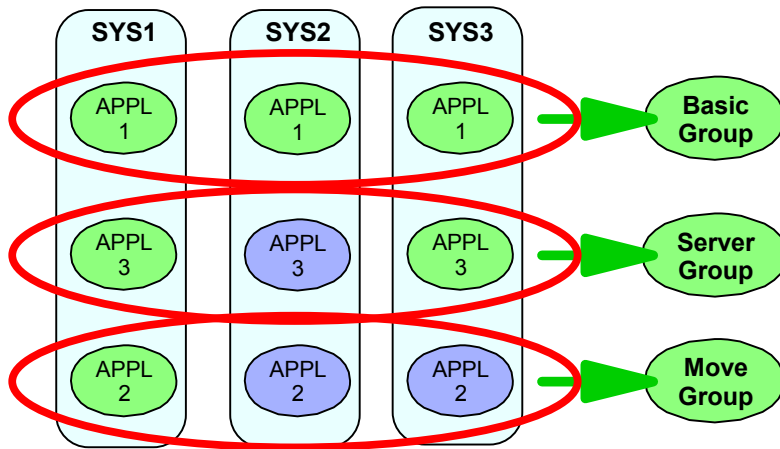
Réduire la complexité avec les regroupements



- Un groupe est un ensemble de plusieurs ressources
- L'état ("status") du groupe est déduit de celui des membres du groupe
- Un groupe peut être un sous ensemble, ou composé d'autres groupes ou dépendances
- L'appartenance à plusieurs groupes est possible
- Un groupe est identifié par un nom unique au niveau système ou sysplex
- Les composants du groupe peuvent être répartis dans le sysplex
- Permet d'automatiser et de contrôler une application (au sens métier) complète
- **Libère l'opérateur de la connaissance de chaque élément de l'application**



3 types de regroupements



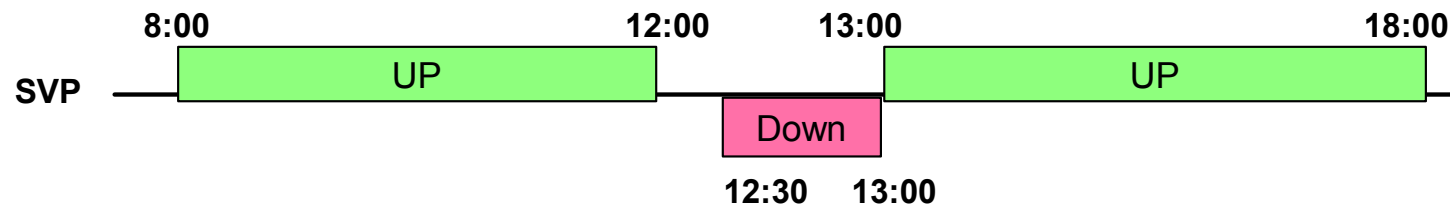
Vous passez une commande simple, l'automate prend en charge les détails

- **Basic group**
L'objectif de disponibilité est m sur m
- **Server group**
L'objectif de disponibilité est n sur m
Un ordre de préférence détermine la séquence de démarrage des composants
- **Move group**
L'objectif de disponibilité est 1 sur m
C'est un cas particulier du "server group"



Automatisation par objectifs

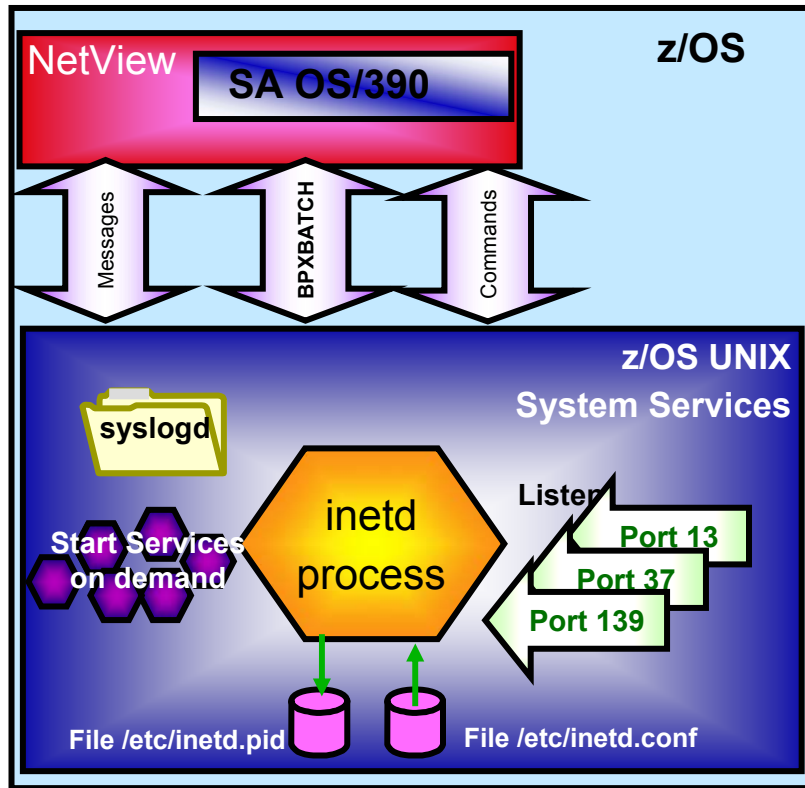
- L'administrateur définit les "objectifs" de disponibilité des applications en fonction des besoins "métier"
 - L'objectif consiste en état désiré, horaires de disponibilité et système cible
 - Relations entre ressources et groupes
 - Périodes de disponibilité:



- System Automation Manager s'efforce de maintenir le système en ligne avec les objectifs
- Les actions éventuelles sont simples et portent sur les exceptions
 - L'opérateur peut "prendre la main" sur les règles d'objectifs par des commandes "d'override ou start/stop"
- Mais la responsabilité des objectifs passe des opérations à l'administrateur



Unix System Services Automation



Exemple: Internet 'Super-Server' - inetd

- Automatisation intégrée
 - ressources et applications USS
 - Agent SA OS/390
 - Définition de règles en mode “SA V2”
 - Relations et groupes avec des ressources “MVS”
- Nouvelle infrastructure
 - USS application type
 - USS Control policy item
PATH, PORT, FILE and USERID
 - Méthodes de monitoring
 - API pour commandes USS



System Automation for OS/390: automatisation WebSphere

Opérations automatisées:

- Démarrage et arrêt de tous les pré-requis et de tous les composants de WebSphere dans un "cluster" Parallel Sysplex, dans le bon ordre
- Règles SA de définition des groupes et des relations de dépendance

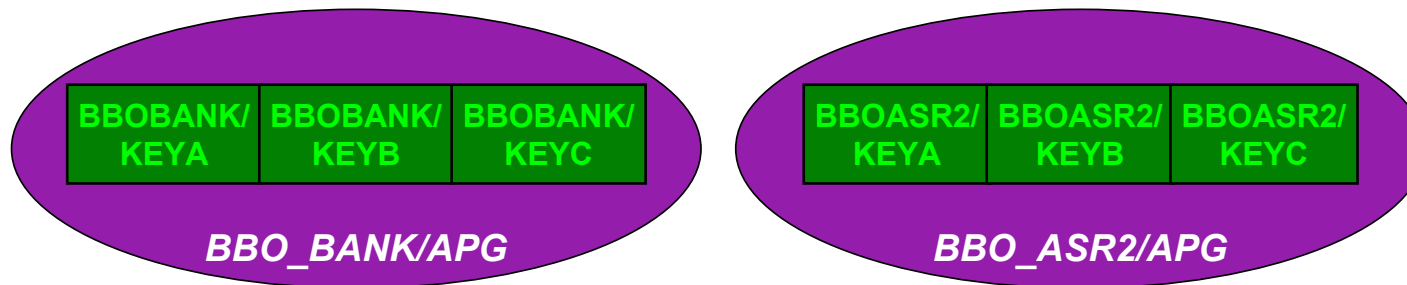
Disponibilité continue

- Monitoring des pré-requis et des composants
- redémarrage automatique des composants et des pré-requis défectueux
redémarrage des ressources dépendantes
- Déplacement de WebSphere vers une autre partition quand le redémarrage d'un composant vital échoue
- "Nettoyage" du système défectueux



SA OS/390 Groupes: J2EE Servers

Définition d'une application (APL) et de ses composants: 2 Instances de serveurs J2EE avec les applications „BANK“ et „ASR2“ qui peuvent tourner sur chacune des images du sysplex KEYA, KEYB, et KEYC



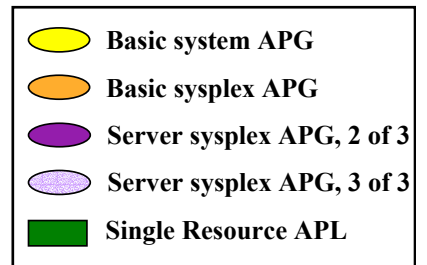
Regroupons les instances serveur en 2 applications **Server Sysplex Application Groups (APG)** avec un objectif de disponibilité de 2 sur 3

Application Group Type: SYSPLEX

Les ressources peuvent résider sur différents systèmes dans le sysplex

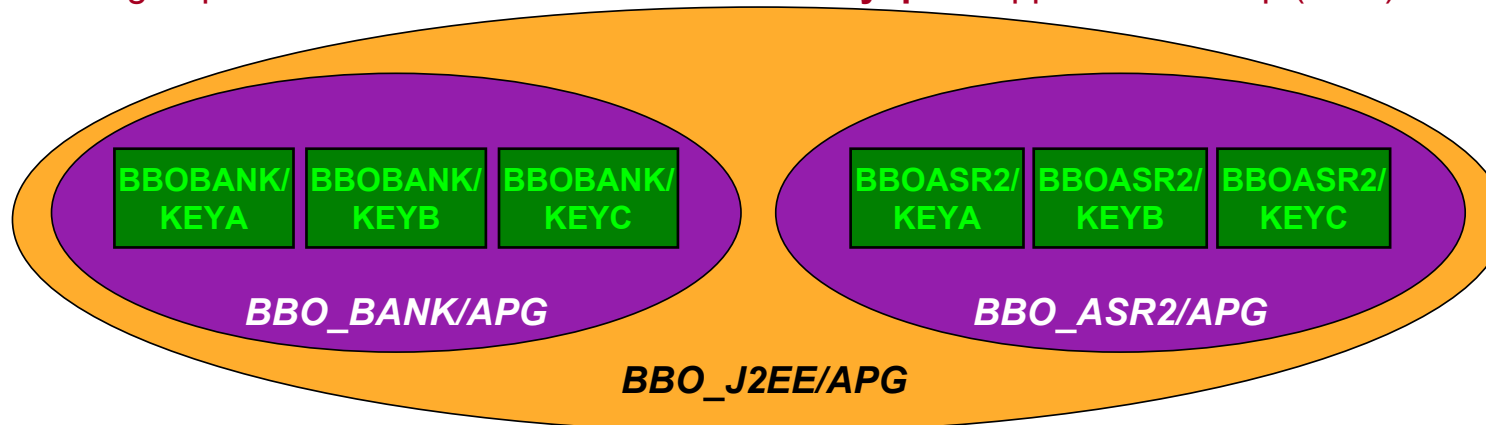
Application Group Nature: SERVER

L'objectif de disponibilité définit combien de membres du groupe doivent être disponibles pour que le groupe soit considéré comme disponible



SA OS/390 Groups : J2EE Servers

Regroupons les J2EE servers dans un **Basic Sysplex** Application Group (APG).

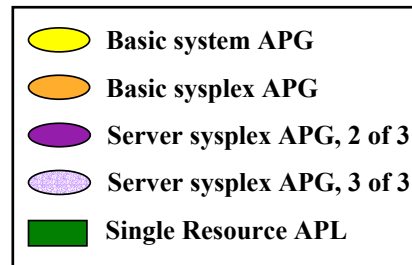


Application Group Type: SYSPLEX

Les ressources peuvent résider sur différents systèmes dans le groupe sysplex

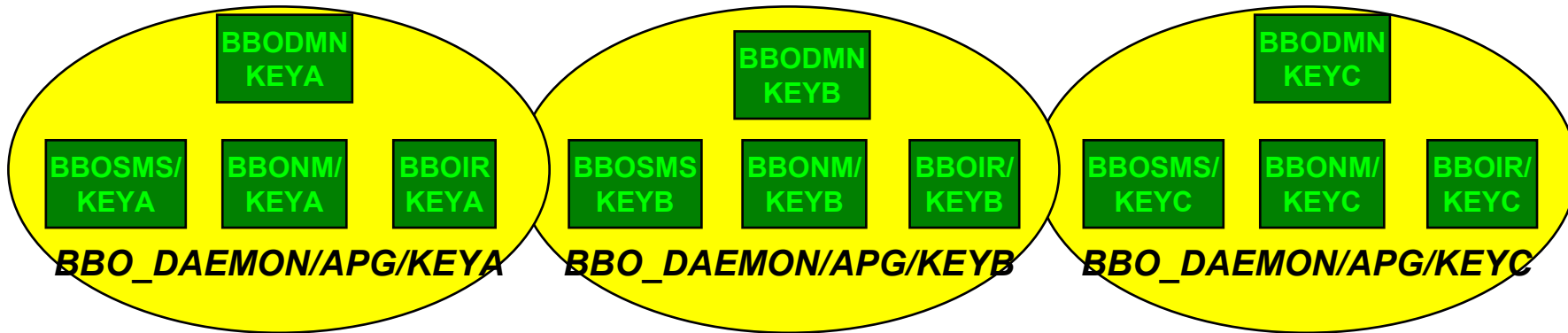
Application Group Nature: BASIC

Le groupe est disponible quand tous ses composants le sont.



SA OS/390 Groupes : WAS environnement de base

Définissons les ressources applicatives (APL) requises par l'infrastructure, sur chaque partition du Sysplex:
 WAS Daemon, System Management, Naming, Interface Repository



Regroupons ces composants de base en un **Basic System** Application Group (APG).

Application Group Type: SYSTEM
 Les ressources résident sur un système.
Application Group Nature: BASIC
 Le groupe est disponible quand tous ses composants le sont.

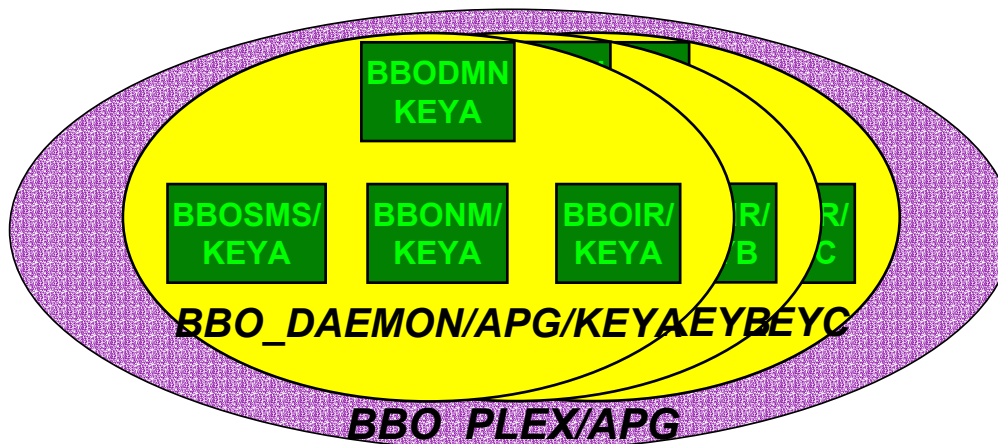
- Basic system APG
- Basic sysplex APG
- Server sysplex APG, 2 of 3
- Server sysplex APG, 3 of 3
- Single Resource APL

BBO_JZEE/APG



SA OS/390 Groupes : WAS Environnement de base

Regroupons tous les environnements de base dans un groupe **Server Sysplex** avec objectif de disponibilité de 3 sur 3



Application Group Type: SYSPLEX

Les ressources peuvent résider sur différents systèmes dans le groupe sysplex

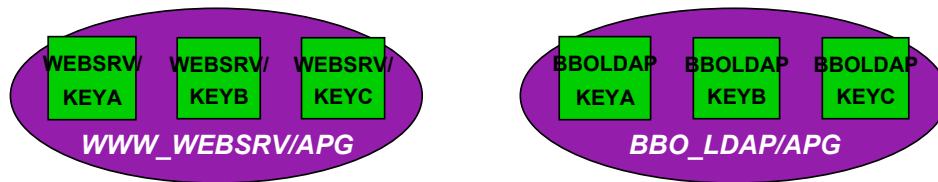
Application Group Nature: SERVER

L'objectif de disponibilité détermine combien d'éléments du groupe doivent être disponibles pour que le groupe le soit.

- Basic system APG
- Basic sysplex APG
- Server sysplex APG, 2 of 3
- Server sysplex APG, 3 of 3
- Single Resource APL

SA OS/390 Groupes : Web Server et LDAP

De la même façon définissons les **Server Sysplex Application Groups** avec objectif de disponibilité de 2 sur 3 pour le Web Server et LDAP

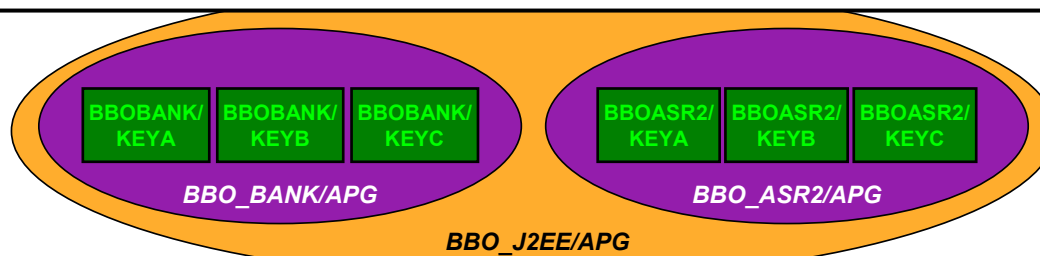


Application Group Type: SYSPLEX

Les ressources peuvent résider sur différents systèmes dans le groupe sysplex

Application Group Nature: SERVER

L'objectif de disponibilité détermine combien d'éléments du groupe doivent être disponibles pour que le groupe le soit.

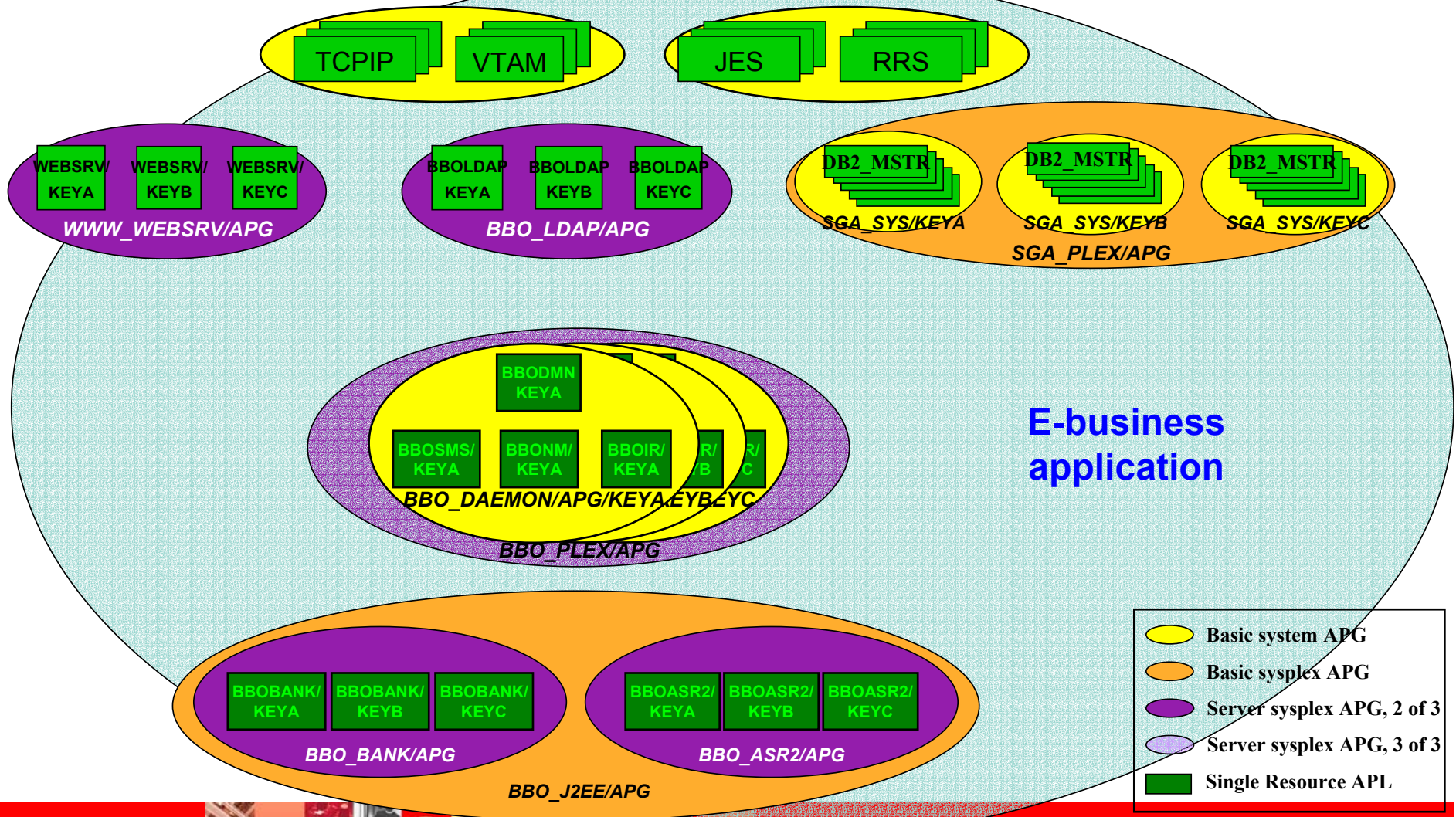


- Basic system APG
- Basic sysplex APG
- Server sysplex APG, 2 of 3
- Server sysplex APG, 3 of 3
- Single Resource APL

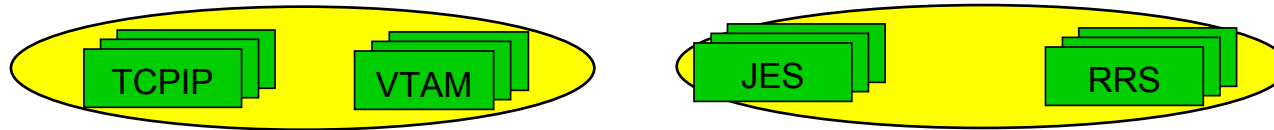
SA OS/390 Groupes: les pré-requis

Ajoutons les applications et groupes pré-requis pour WAS.

Plaçons les dans un groupe SYSPLEX



Relations: J2EE Servers et WAS Base APG

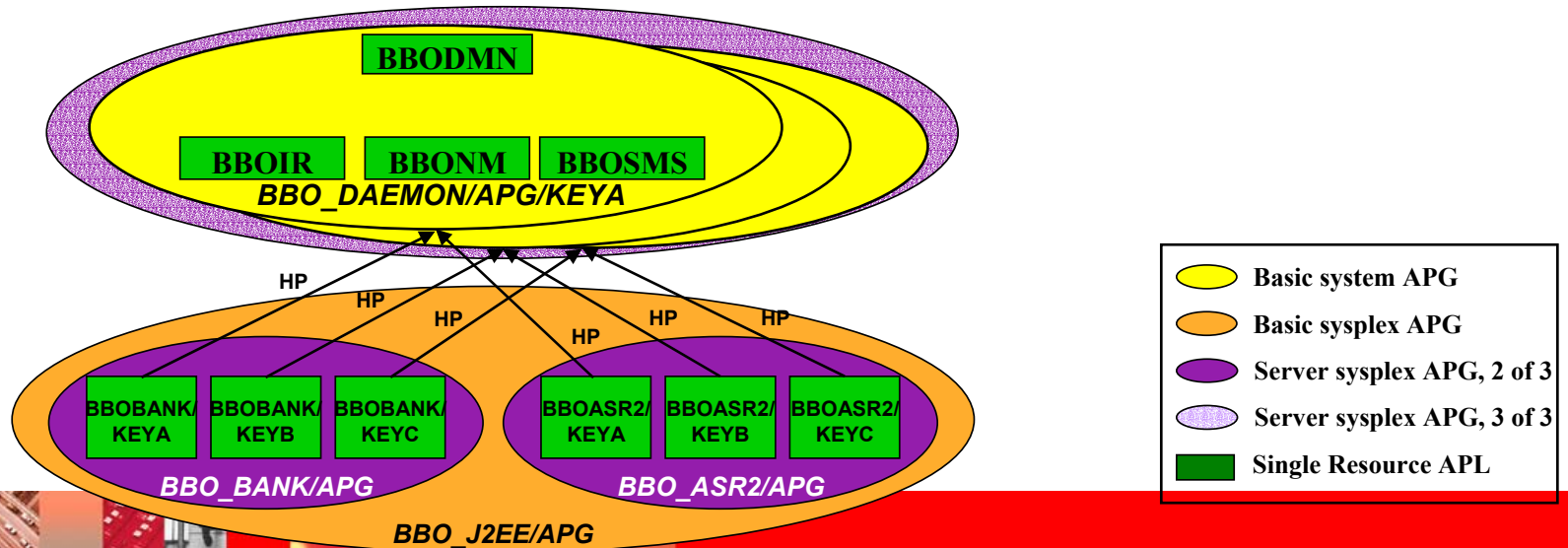


HasParent (HP) Relationship:

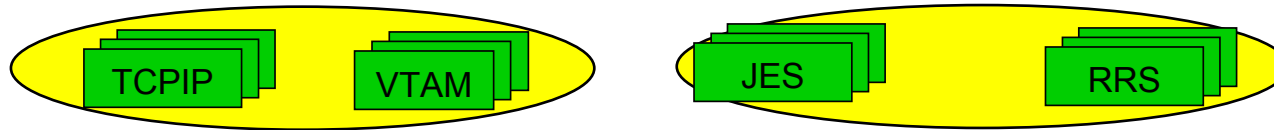
Une ressource dépendante ne peut être démarrée tant qu'une ressource pré-requise ne l'est pas.

Une ressource pré-requise ne peut être stoppée avant que la ressource dépendante ne soit arrêtée.

Une demande de démarrage d'une ressource dépendante provoquera une demande de démarrage d'une ressource pré-requise si celle-ci est "down".



Relations: WAS Base Environment



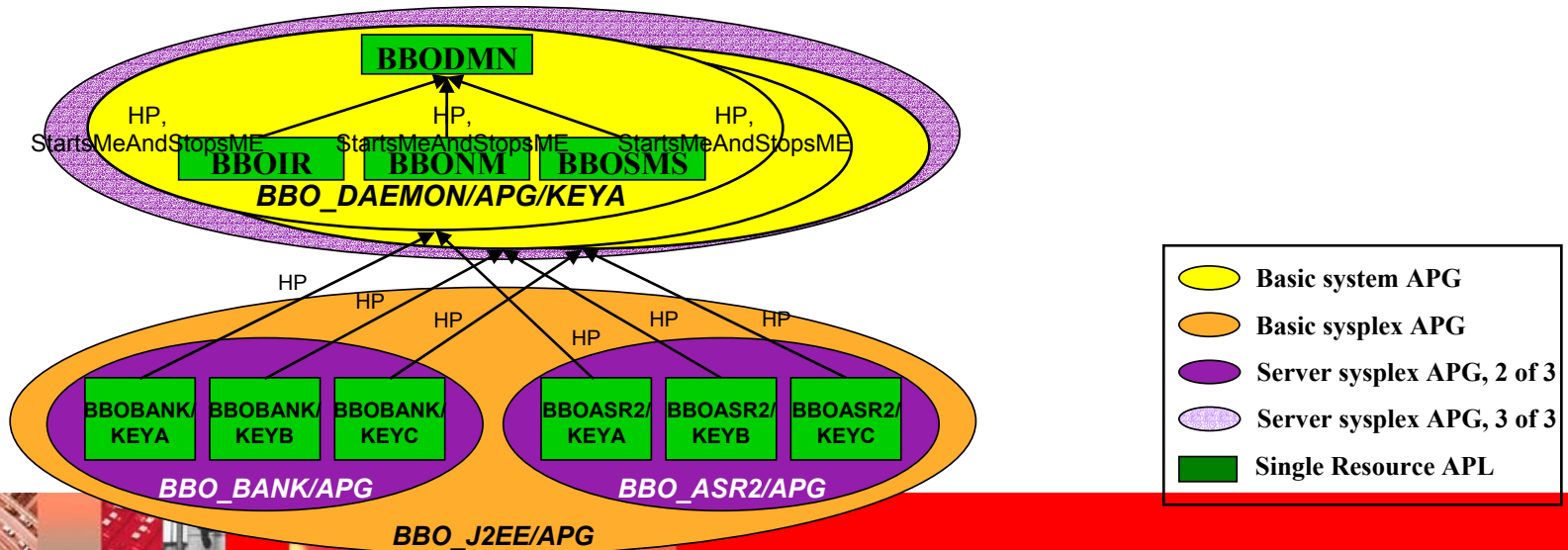
HasParent (HP) Relationship:

Une ressource dépendante ne peut être démarrée tant qu'une ressource pré-requise ne l'est pas.

Une ressource pré-requise ne peut être stoppée avant que la ressource dépendante ne soit arrêtée.

Relationship condition StartsMeAndStopsMe:

La ressource pré-requise démarrera (et arrêtera) la ressource dépendante lorsqu'elle-même sera démarrée (et arrêtée).

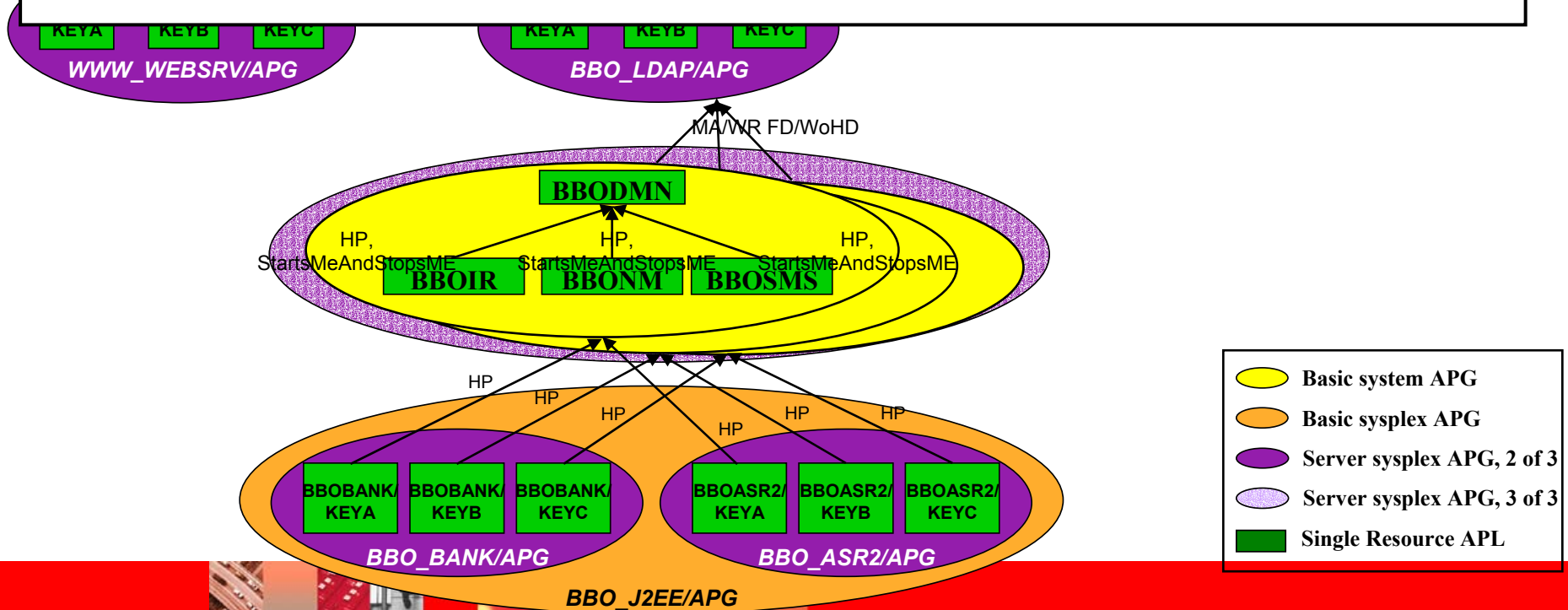


Relations: WAS Base et LDAP

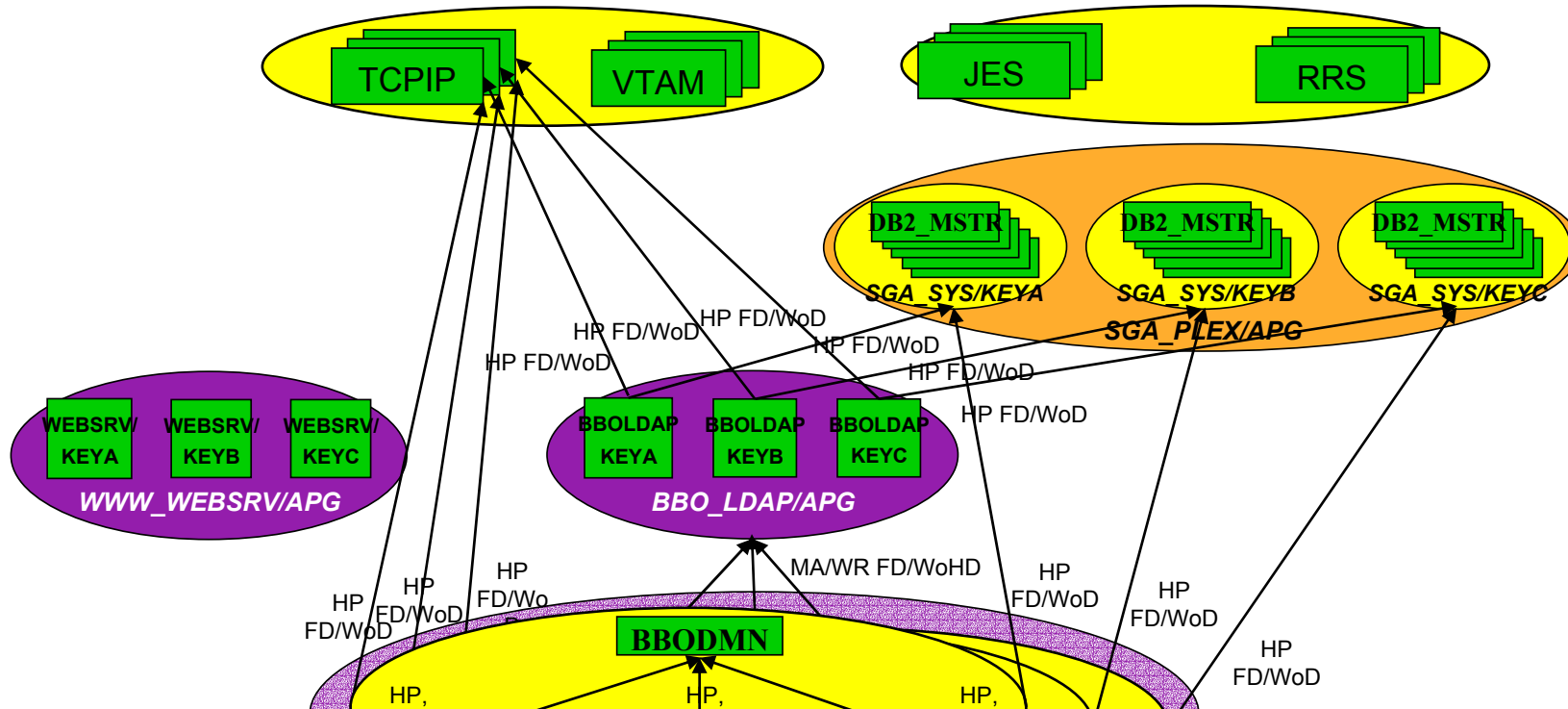
MakeAvailable/WhenRunning (MA/WR):

La ressource dépendante ne peut être démarrée tant que la condition requise n'a pas été remplie au niveau de la ressource pré-requise. SA OS/390 démarre la ressource dépendante dès que la condition est satisfaite. "WhenRunning" signifie que l'état ("status") observé est "Available" ou "Degraded".

Forcedown/WhenobservedHardDown (FD/WoHD): Force un « shutdown » si la ressource pré-requise est vue en « hardown », c'est-à-dire que SA ne peut la redémarrer.



Relations: WAS Base et LDAP avec DB2 et TCPIP



HasParent (HP) Relationship:

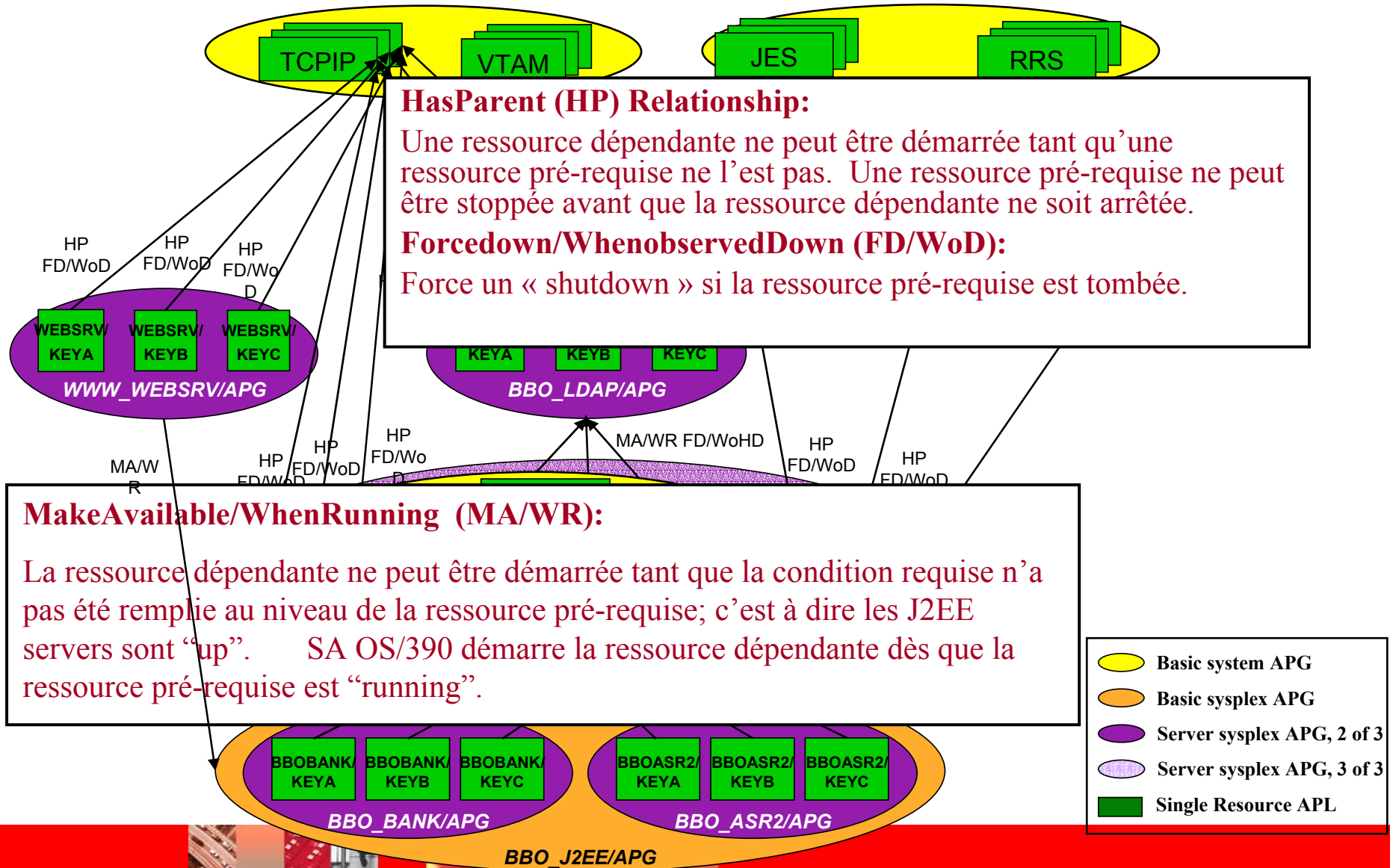
Une ressource dépendante ne peut être démarrée tant qu'une ressource pré-requise ne l'est pas. Une ressource pré-requise ne peut être stoppée avant que la ressource dépendante ne soit arrêtée.

Forcedown/WhenobservedDown (FD/WoD):

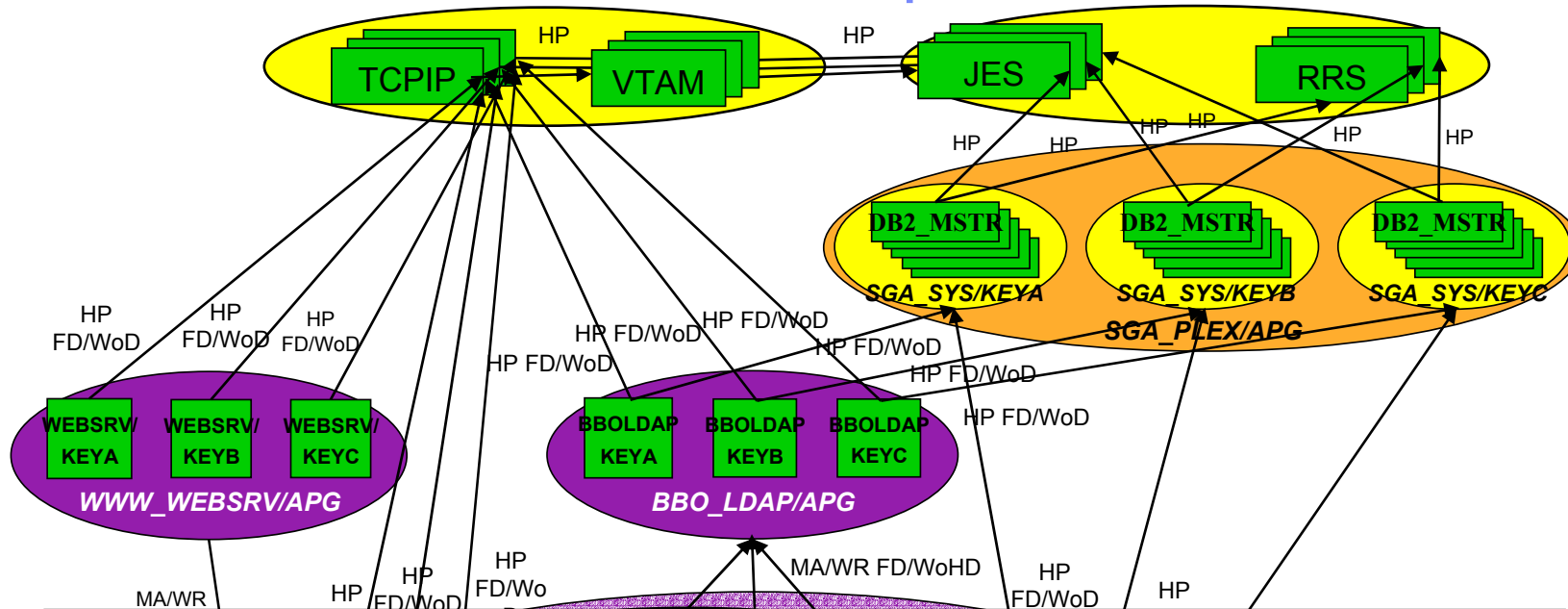
Force un « shutdown » si la ressource pré-requise est tombée.

- Basic system APG
- Basic sysplex APG
- Server sysplex APG, 2 of 3
- Server sysplex APG, 3 of 3
- Single Resource APL

Relations: WebServer avec J2EE Servers et TCPIP



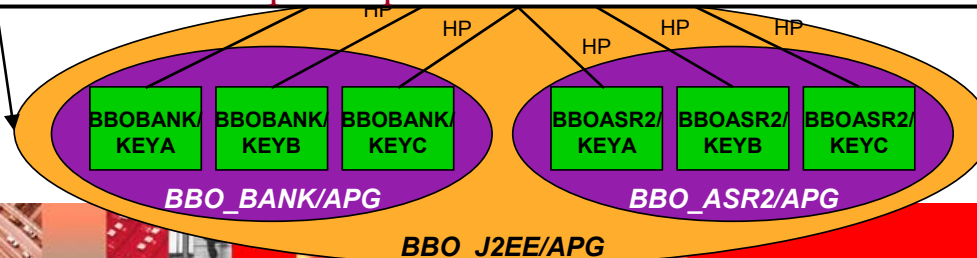
Relations: Base Components



HasParent (HP) Relationship:

Une ressource dépendante ne peut être démarrée tant qu'une ressource pré-requise ne l'est pas. Une ressource pré-requise ne peut être stoppée avant que la ressource dépendante ne soit arrêtée.

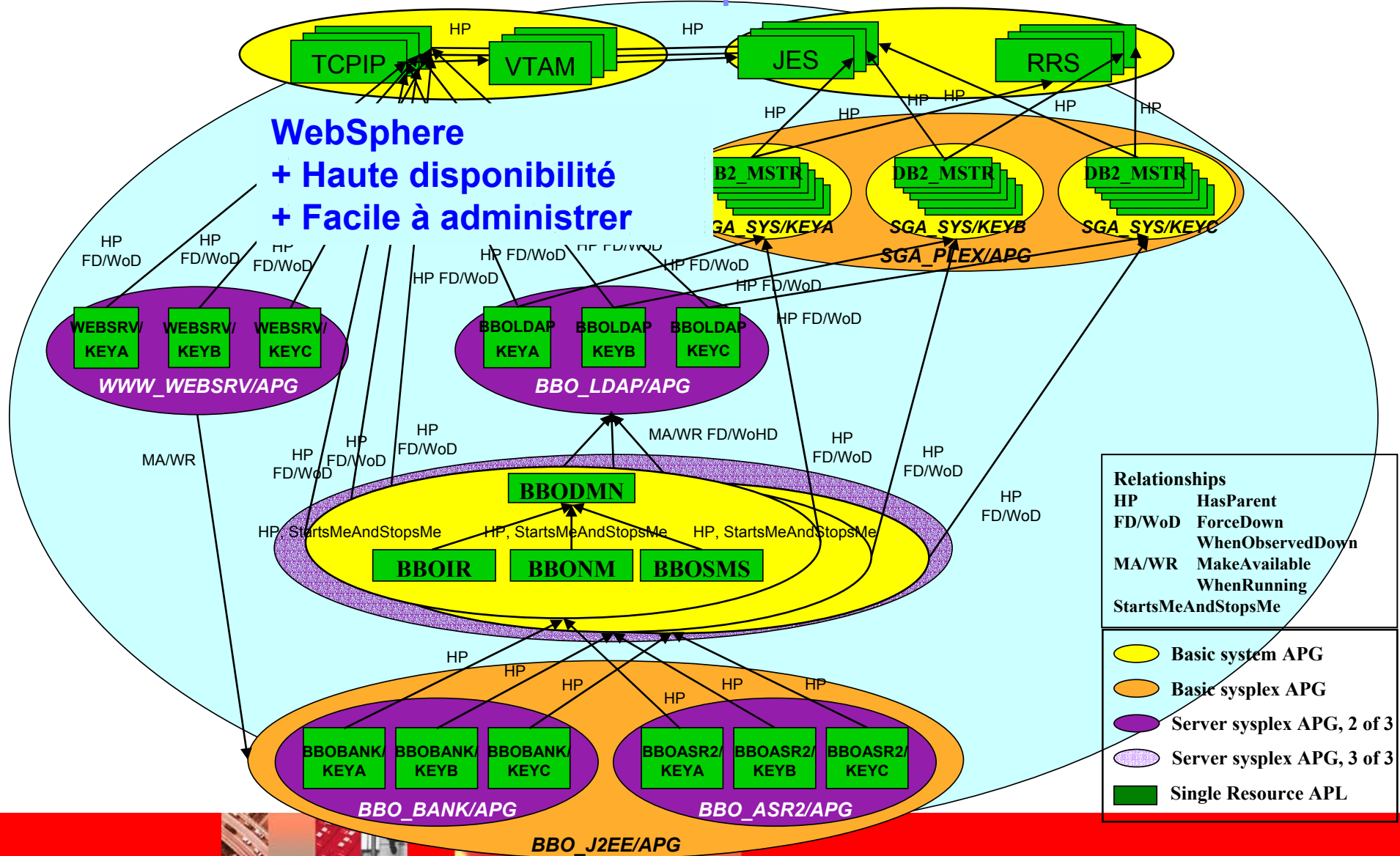
Une demande de démarrage d'une ressource dépendante provoquera une demande de démarrage d'une ressource pré-requise si celle-ci est "down".



System APG

- Basic sysplex APG
- Server sysplex APG, 2 of 3
- Server sysplex APG, 3 of 3
- Single Resource APL

Relations: Tous les composants...



Ce qu'apporte SA OS/390 :

- une table de messages enrichie:
messages émis par WebSphere et les produits qui lui sont associés,
durant le démarrage et l'arrêt
- une procédure commune:
pour "nettoyer" les espaces adresses serveurs que WebSphere n'a pu
arrêter correctement
- un "white paper":
concepts and set-up guidelines
- une "Policy Database":
Des définitions que vous pouvez utiliser pour paramétrer SA OS/390
pour une automatisation personnalisée de votre WAS
- des „sample tools“ :
jobs, OMVS shell scripts, REXX execs que vous pouvez utiliser pour
automatiser les différentes étapes de l'administration de WAS



Scenarios de recovery WebSphere

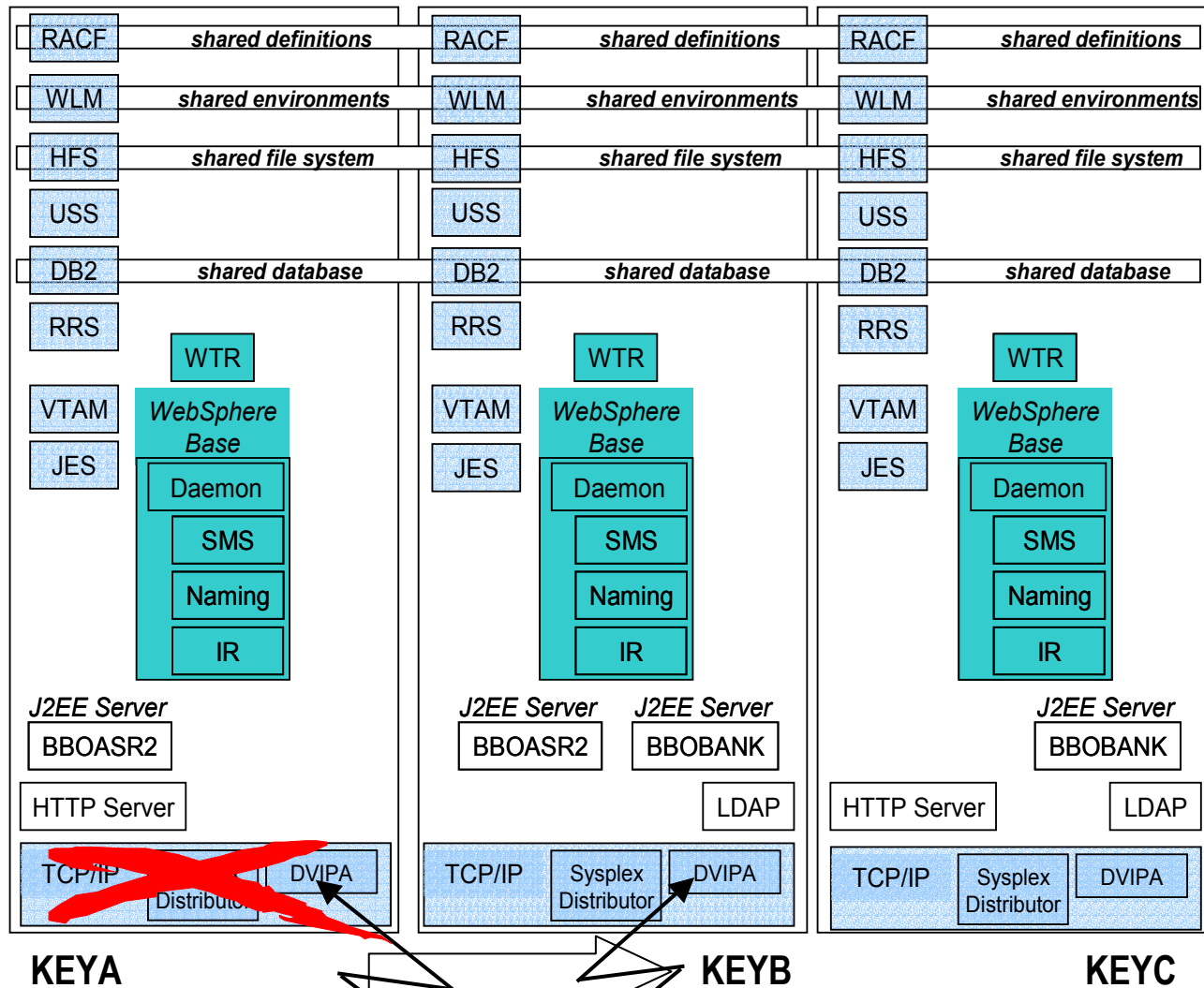
- Que faire si TCPIP tombe sur le système “KEYA” ?
- Que faire si un HTTP server tombe ?
- Que faire si certains composants du WAS Server tombent ?



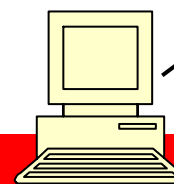
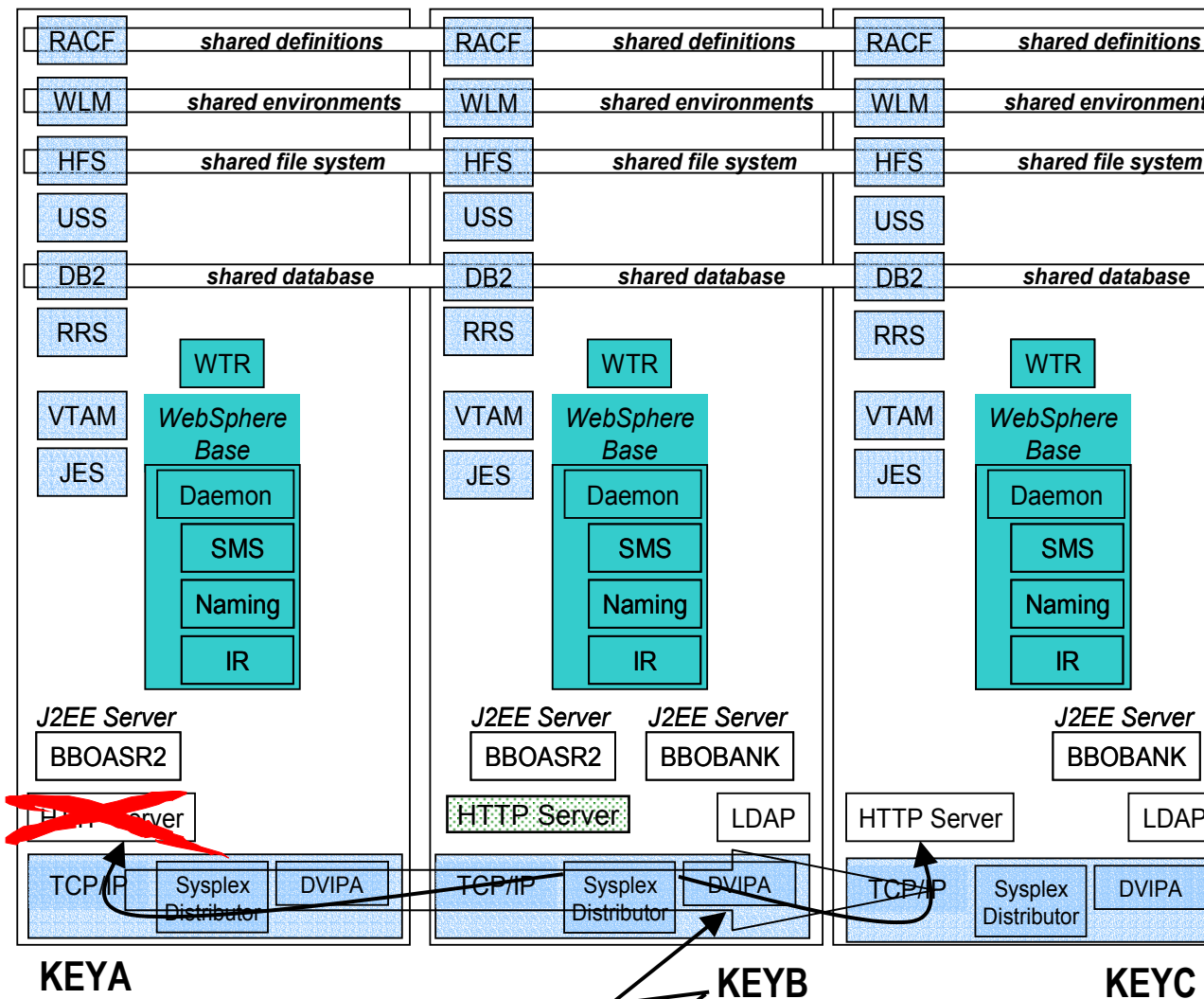
- Un utilisateur web se connecte en DVIPA
- DVIPA tourne sur le système „KEYA“
- Le Sysplex Distributor de KEYA répartit les requêtes sur les 3 systèmes.

TCPIP sur KEYA tombe !

- La connexion à KEYA est perdue
- Le TCPIP de backup sur KEYB, reprend
- L'utilisateur web est automatiquement connecté à „KEYB“
- Le Sysplex Distributor de KEYB répartit les requêtes sur les 3 systèmes



- Le Sysplex Distributor sélectionne le HTTP Server KEYA pour traiter la requête HTTP de l'utilisateur
- **HTTP Server KEYA tombe !**
- la connexion HTTP et les réponses HTTP en attente sont perdues
- Le Sysplex Distributor automatiquement route la requête HTTP suivante vers le HTTP Server KEYC
- SA redémarre le HTTP server KEYA
- Si cela échoue, SA démarre le 2ème HTTP Server KEYB
- Le Sysplex Distributor distribue les requêtes HTTP entre KEYB et KEYC



LDAP ou un J2EE Server tombe !

L'approche est très comparable à la défaillance du HTTP Server:

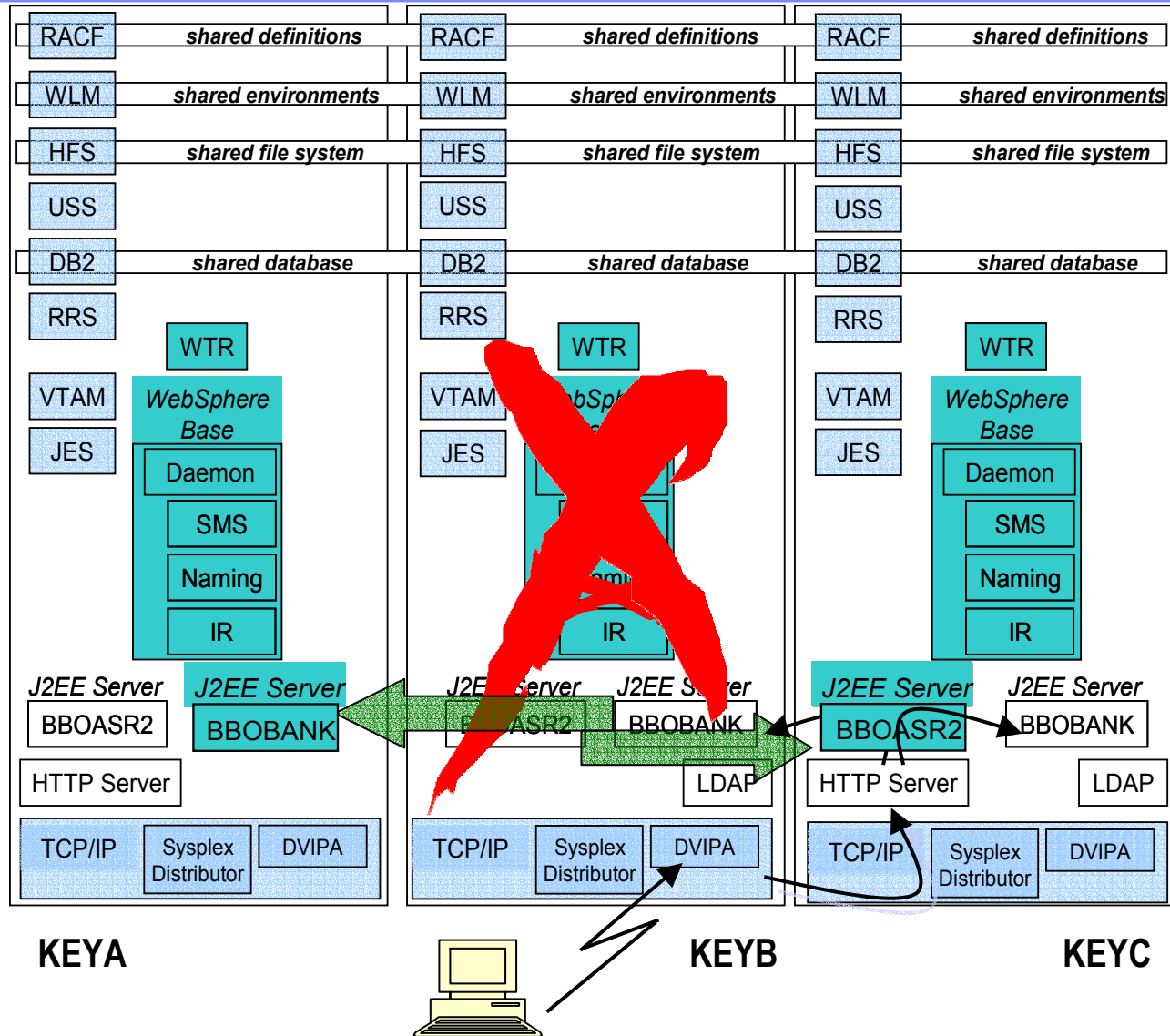
- Le Sysplex Distributor sélectionne un LDAP server et un J2EE server pour satisfaire la requête
- Si un LDAP server ou un J2EE server tombe, le Sysplex Distributor va „router“ la requête suivante vers une instance LDAP ou J2EE disponible
- System Automation va relancer le composant défaillant sur le même système
- Si le redémarrage échoue, System Automation va démarrer un a LDAP server ou un J2EE server (instances) sur le 3ème système
- System Automation va „nettoyer“ le système défaillant si nécessaire



- Sysplex Distributor sélectionne le J2EE server BBOBANK sur KEYB

Des composants du J2EE server KEYB tombent !

- WAS server et J2EE server sur KEYB s'arrêtent
- Les réponses des applications en cours sont perdues
- Le Sysplex Distributor „route“ les requêtes suivantes vers BBOBANK sur KEYC
- SA redémarre les serveurs WAS et J2EE sur KEYB
- Si cela échoue, SA démarre BBOBANK sur KEYA et BBOASR2 sur KEYC
- Le Sysplex Distributor répartit les requêtes pour BBOBANK entre KEYA et KEYC



Un autre composant pré-requis pour WAS tombe

L'approche est très comparable à celle adoptée pour le problème du WAS Server :

- Le WAS server et le J2EE server s'arrêtent
- Le Sysplex Distributor sélectionne un autre J2EE server pour satisfaire les nouvelles requêtes
- System Automation relance les composants défectueux, le WAS server, et le J2EE server sur le même système
- Si le redémarrage échoue, System Automation démarre un J2EE server sur un autre système (où le WAS server est déjà „running“ !)
- System Automation va „nettoyer“ le système défectueux si nécessaire



Synthèse: l'automatisation de WebSphere permet de réduire les coûts et d'améliorer la disponibilité des applicatifs

❑ Simplifier l'administration

- Démarrer tous les composants et les pré-requis de WAS dans un sysplex, dans la bonne séquence
- Arrêter tous les composants de WAS dans un sysplex, dans la bonne séquence
- Afficher l'état (status) de tous les composants et pré-requis
- Utiliser les règles SA pour définir des groupes et des relations

❑ Réduire l'indisponibilité

- Monitorer les composants et leurs pré-requis
- Redémarrer automatiquement les composants défectueux
- Déplacer WebSphere vers une autre partition quand le redémarrage de composants vitaux échoue
- "Nettoyer" les systèmes défectueux
- Automatiser la maintenance de WAS

