

**Bad Homburg
19. März 2009
VMS Frühjahrstreffen**



**PRIMO·S,
Synchronisation und
Global Sections**



Werner Stegbauer
Manager PRIMO·S-Center, ABB Automation



Agenda

- About PRIMO·S
- PRIMO·S References
- PRIMO·S Highlights
- Synchronization and Global Sections

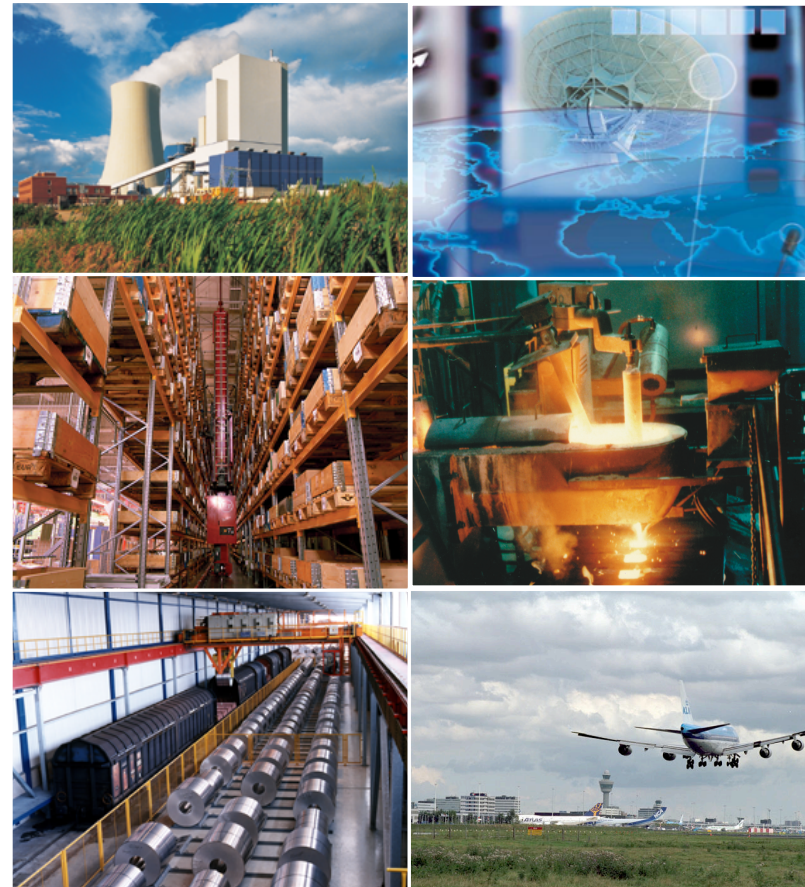
About PRIMO-S

PRIMO-S – ABB's real-time database

- Established product
 - Has been used successfully for many years
 - Over 4000 proven installations worldwide
- Compact, simple system
 - Little need for training
 - Installed and ready for use in just two minutes
 - Requires less than 6MB disk space under NT
- Available on a number of platforms
 - VMS (VAX, Alpha, Itanium – “award winning” in 2005)
 - UNIX (Tru64, Linux, SunOS, HP-UX, AIX)
 - Windows (from NT 3.51 to 2008)
- Low price

The PRIMO-S Application Fields

- Power stations
- Rolling mills
- Continuous casting plants
- Manufacturing process control
- Cargo management
- Telecommunications
- Baggage and cargo handling systems
- Simulation systems
- Diagnosis Systems



ABB

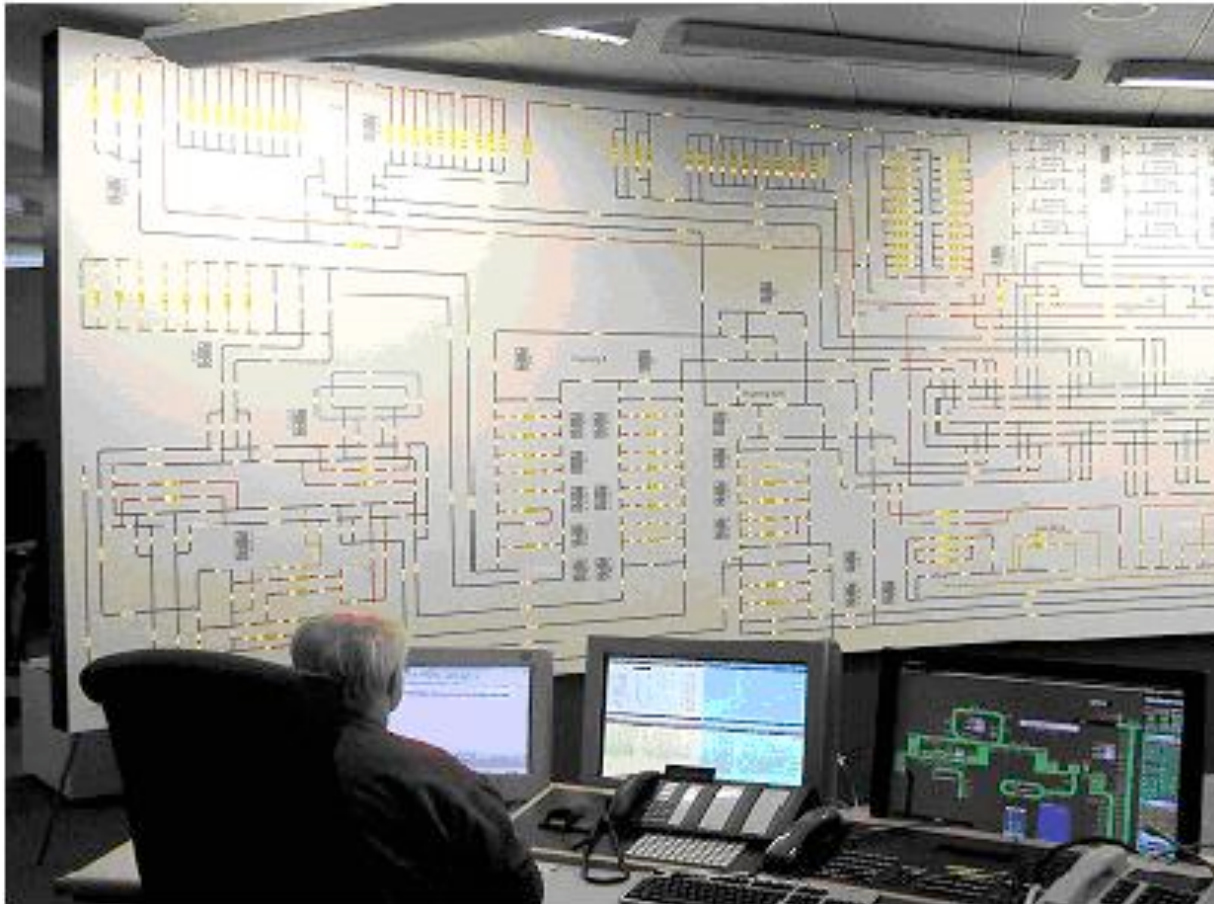
PRIMO-S References

PRIMO·S Reference (1) - FRAPORT

- Frankfurt Airport
 - Baggage Handling System called GFA (google for it)
 - Up to 100,000 pieces of luggage per day
 - 4 times 2 Alphas (each pair in a cluster) – now in Itanium
 - Replace fault tolerant VAX with PRIMO·S redundancy concept Replication
 - 45 minutes guaranteed transfer time between flights
- Reasons for PRIMO·S
 - Performance
 - Replication with HeatUp and fast switch over in failure case

PRIMO-S Reference - FRAPORT

- Frankfurt Airport – Control room



Picture: FRAPORT

FRAPORT moves to Itanium (2008/2009)

■ Fraport Performance Comparison

Mit meinem Messprogramm habe ich den CPU Faktor zwischen Alpha und Itanium ermittelt. - Dieser liegt bei 2,62.

Für die Verarbeitung der 100.000 Telegramme benötigt der PRIMOS-Telver auf der Itanium 9,85 sec CPU gegenüber 94,91 sec CPU auf der Alpha, das ist Faktor 9,64 schneller, also **ca. 4 * schneller als zu erwarten gewesen wäre.**

Jedenfalls habe ich dann den Original Telver genommen .. und habe darin die PRIMOS durch ORACLE Funktionen ersetzt.

Mit diesem Telver wurden die folgenden Zeiten gemessen: **1938,5 sec**

Damit ergibt sich der direkte **Vergleichsfaktor zu 196,80**

- **PRIMO-S is 196 times faster than Oracle!!**
 - How long is 1938 seconds?

PRIMO·S Reference (2) - STATOIL

- STATOIL - Snøwhit
 - Norwegian company the builds a huge liquefied natural gas (LNG) plant at the Arctic Circle close to Hammerfest
 - See www.statoil.com
 - PRIMO·S used in emulation system for the target plant and also for training of operators
 - Simulates 25 controllers in one PC
- Reasons for using PRIMO·S
 - Performance
 - Message Arrays for inter-task communication
 - Direct Access to data to fulfill extreme performance requirements

PRIMO-S Reference - STATOIL

- Snøwhit – Hammerfest - Melkøya



Picture: STATOIL

PRIMO-S Reference - STATOIL

- Melkøya



Picture: STATOIL

PRIMO-S Highlights

PRIMO-S Highlights

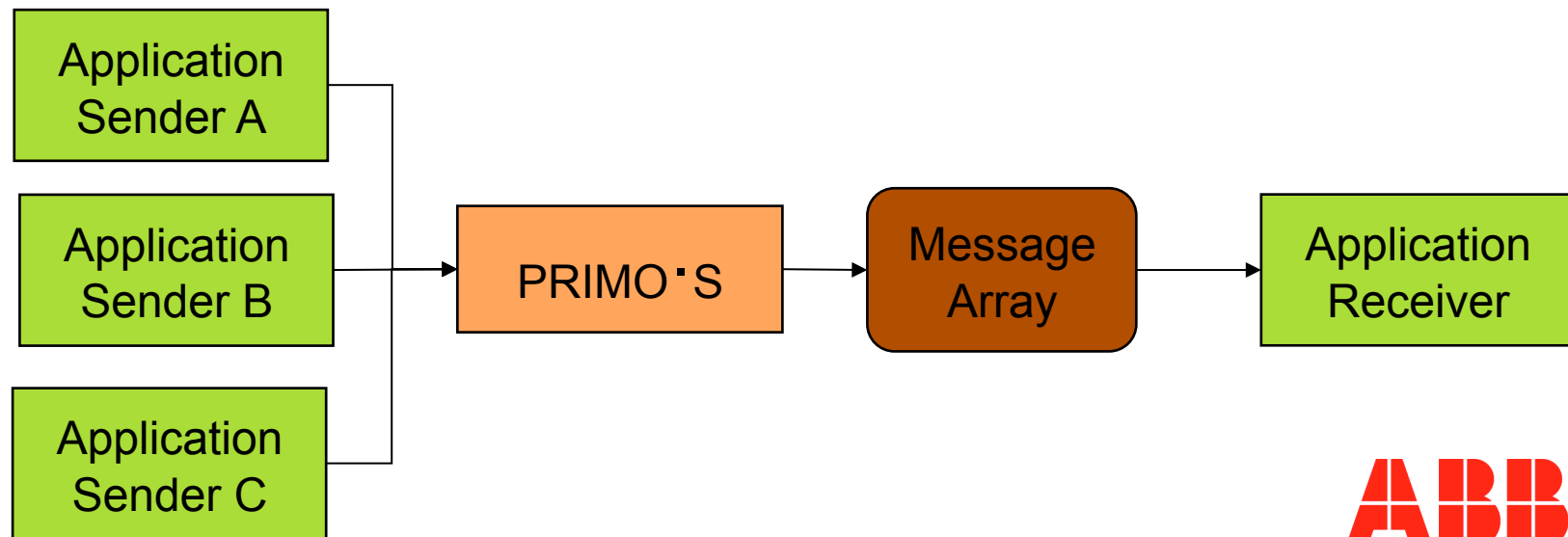
- Performance
- Message Arrays
- Replication
- Logging
- Standards
- Code Generator

Performance

- PRIMO·S is very fast
 - More than 300,000 accesses per second
- Up to 35 times faster than conventional SQL data base
 - Measured by Institute in Cologne, Germany
- Data is accessed in memory
 - Global section on VMS
 - Memory mapped files on UNIX and Windows
- PRIMO·S itself is a shared image (DLL)
- Direct Access (via address) to data provided
 - For experienced user
 - And extrem performance requirements
- A Demo ..

Message Arrays

- Has built-in inter-task communication
 - Technology called Message Arrays
 - Three times faster than VMS mailboxes
 - More than 100,000 send/receive per second
 - Powerful Message Monitor
 - Sender and receiver may be on a remote computer



Synchronization and Global Sections

PRIMO-S Implementation Architecture

- Map data base (= one single File) in address space of each application
 - VMS – Global Sections, UNIX: mmap, Windows: memory based files
- Advantages
 - Extrem fast access via address calculation
 - No management of data buffers – the pager does it for you
 - No explicit write back to disk
- Disadvantages
 - Size of virtual address space limits size of data base
 - Must solve problem to save data on disk
 - Concurrent Accesses must be synchronized

Three Synchronization Problems

- A: Concurrent Accesses
- B: Save data on disk
- C: Keep two Global Sections during Replication in synch

A: Concurrent Accesses

- Should be completely inline code in non-conflict case
- Needs exclusive lock (Masterlock)
 - VAX/Alpha/Itanium: BBSSI
 - Windows: WaitForSingleObject
- In case of access conflicts we need support from operating system for stop/restart of waiters
- Demo: Two profil with counter 50 on perfdb
 - See code native.h on Vms and windows and usage in psv.c

B: Save data on disk

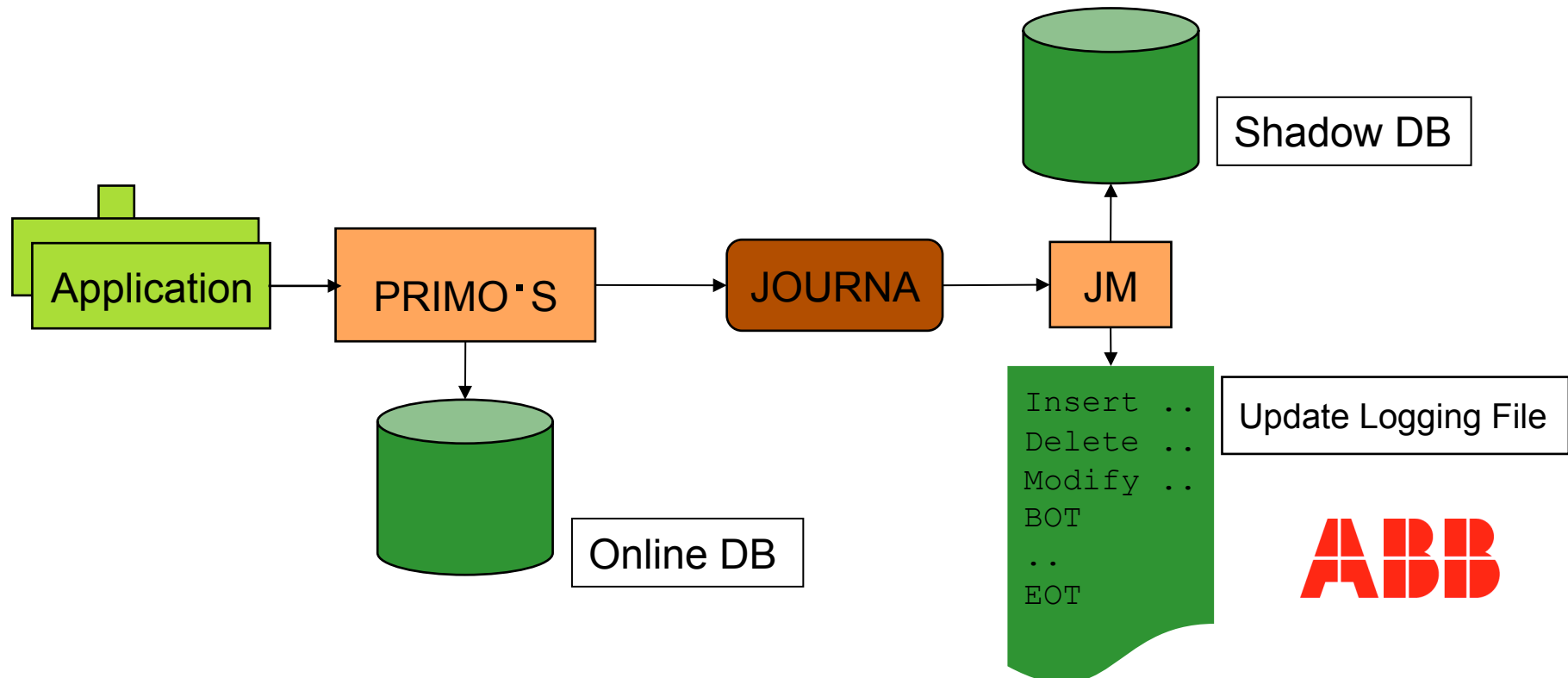
- If system is stopped normally it's no problem
- Is only a problem if system fails:
 - So what we need is:
 - If system fails at any time during runtime
 - .. Data has to be reconstructed as accurate as possible
 - .. in a consistent manner
- First try: Update global section
 - Not reliable
- Better approach: Logging

B: Logging Concept

- Save complete data base at startup in shadow data base
- Force all changes on disk in asynchronuos manner in logfile
- If system fails use shadow and logfile to recover
 - Watch open transactions
- Conclusion: Logging solves problem B
- Demo?

Logging

- No data loss in failure situations
 - Recovery uses Shadow DB and Update Logging File
 - Logging may be local or remote, single or multiple
- Synchronization on-line of partner host after failure (HeatUp)

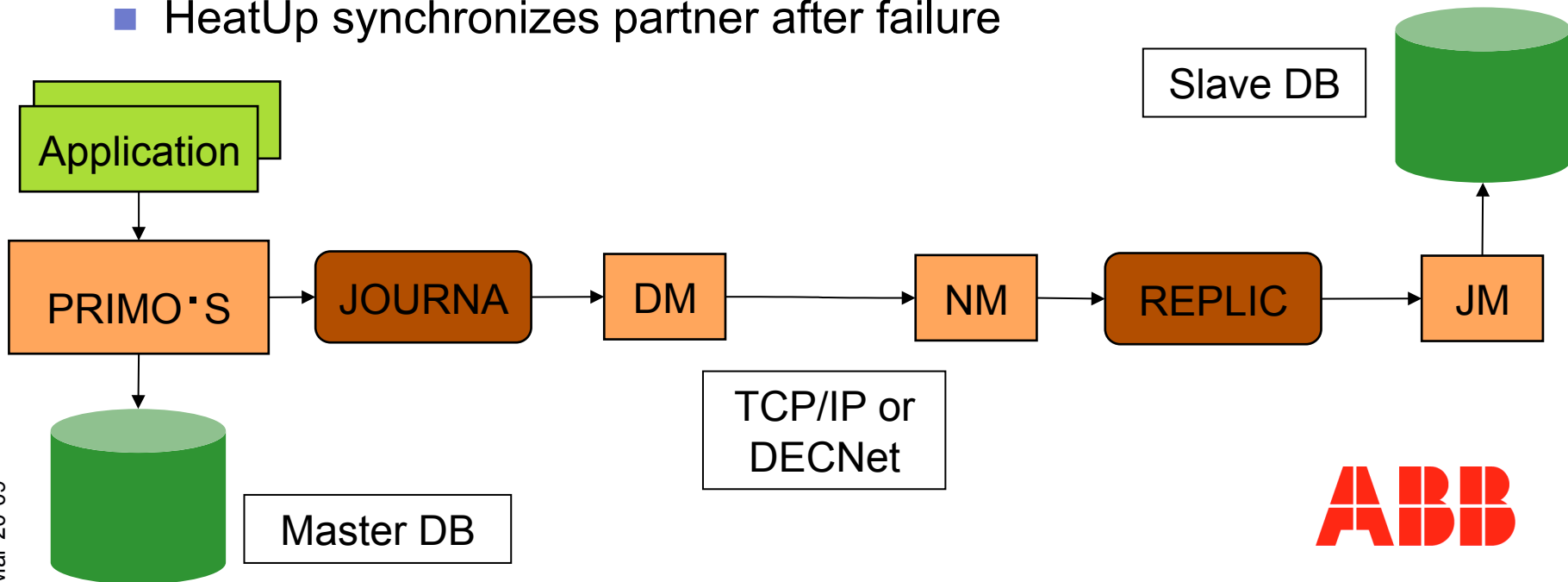


C: Sync two Global Sections during Replication

- Move all updates from master to slave and update remote global section
- Heatup brings partner after restart online in sync
- Optimization for FRAPORT:
 - Block processing is about 3 times faster than single processing
- Demo?
 - Perfdb on demo\tmp

Replication

- Master/Slave principle to replicate data to partner host (may be foreign SQL db)
 - Replicates update event-driven – as soon as possible
 - Several thousand per second
 - Signals errors by callback to user
 - HeatUp synchronizes partner after failure



About me

- Werner Stegbauer
- ABB Automation GmbH
PRIMOS-Center, DEATG/LEMA
Kallstadter Str. 1
68309 Mannheim, Germany
- Tel: +49(0)621/381-1720
- E-mail : werner.stegbauer@de.abb.com
- Fax: +49(0)621/381-1610

AABB