IBM Systems and Technology Group

Experiences of Applying Samba in Enterprise NAS Products

SambaXP 2014

Dr. Ingo Meents
Göttingen, 2014-05-15

Contents

- **SONAS and V7000 Unified**
- **Customer use cases**
- **Role of Samba and CTDB in the (IBM) NAS Stack**
- **Selected contributions**
  - Support of Microsoft Management Console
  - Access control list improvements
  - Performance
  - Deterministic idenity mapping
  - Signing
  - SMB3
- **Next Steps**

NAS Products Employing Samba

Release 1.4.x
Samba 3.6

SONAS
- Enterprise NAS
- Scale out up to 21.6 PetaBytes

Release 1.5
GA June 2014
Samba 4

SONAS V7000 Unified
- Midrange NAS
- Unified block and file service
- Clustered up to 4PetaBytes

Both support
- Active-active clustering
- Cross-protocol interoperability with NFS, http, ftp, scp, CIFS/SMB
- Based on GPFS

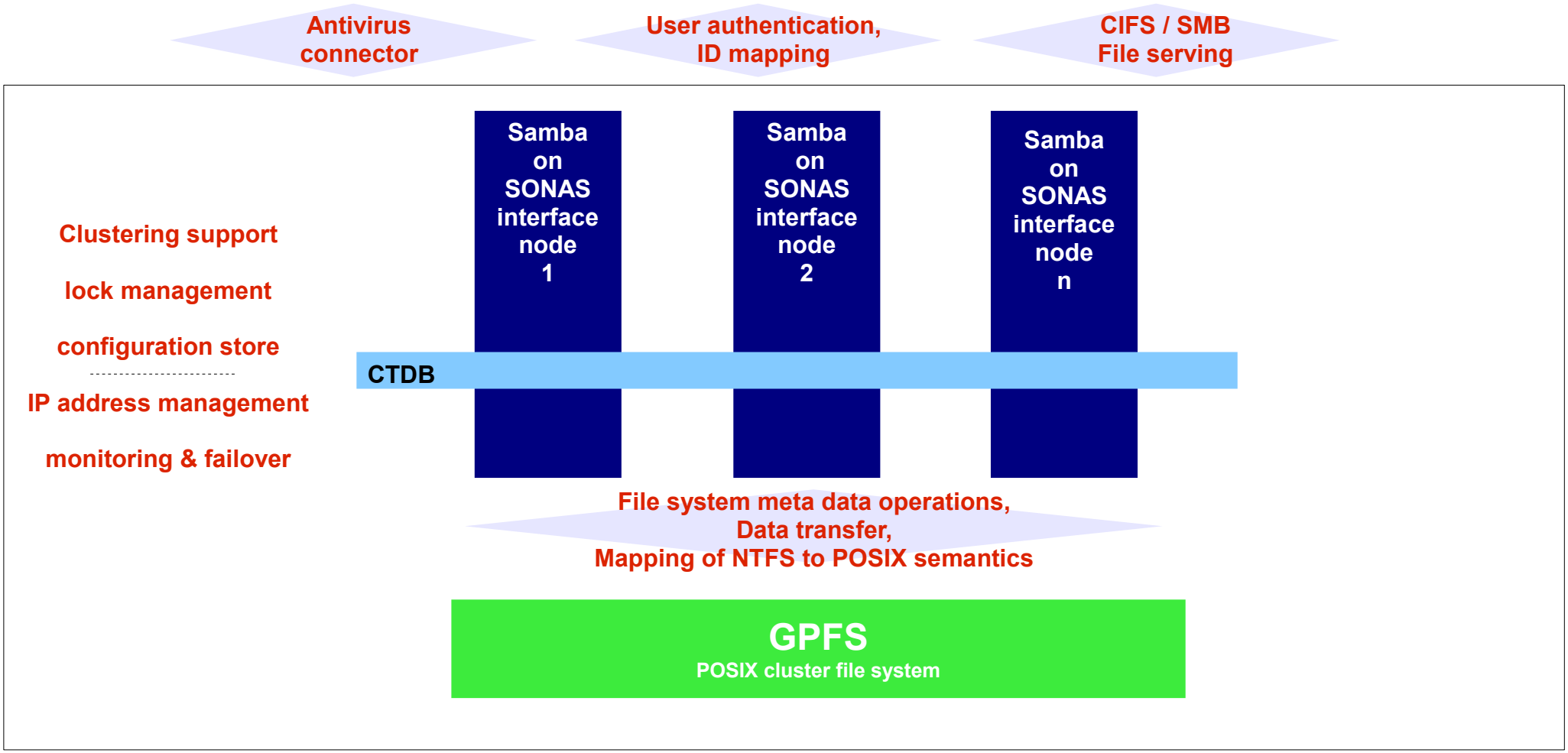
SMB Plays An Important Role For Our Customers

- **Some typical use cases**
  - Home folders in enterprises, education, public
  - Project folders for engineering teams, access to result data of HPC computations
  - Genomic sequencing: most sequencers are based on Windows and deliver the raw data for further processing (NFS, compute cluster)
  - File server consolidation, multi-protocol
  - Private clouds with multi-protocl access (SMB, NFS, Rest, ftp)
- **Expectations are ranging**

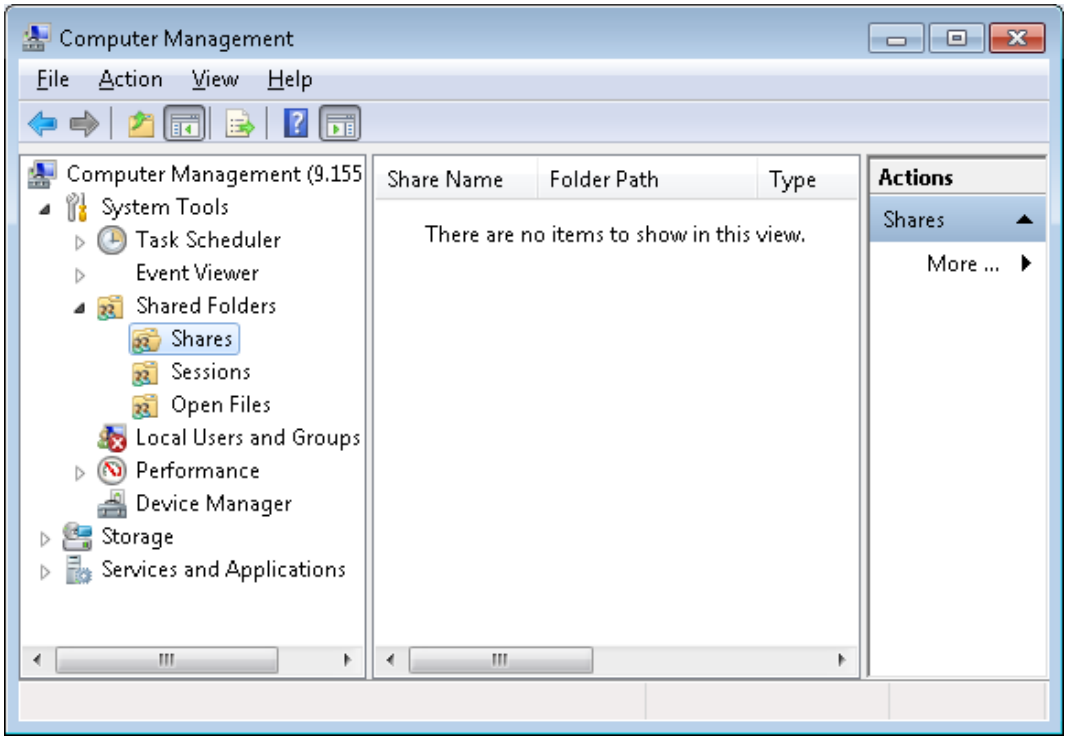
e.g. ACL Improvements

e.g. Unix Extensions

Samba And CTDB Play An Important Role In The NAS Stack

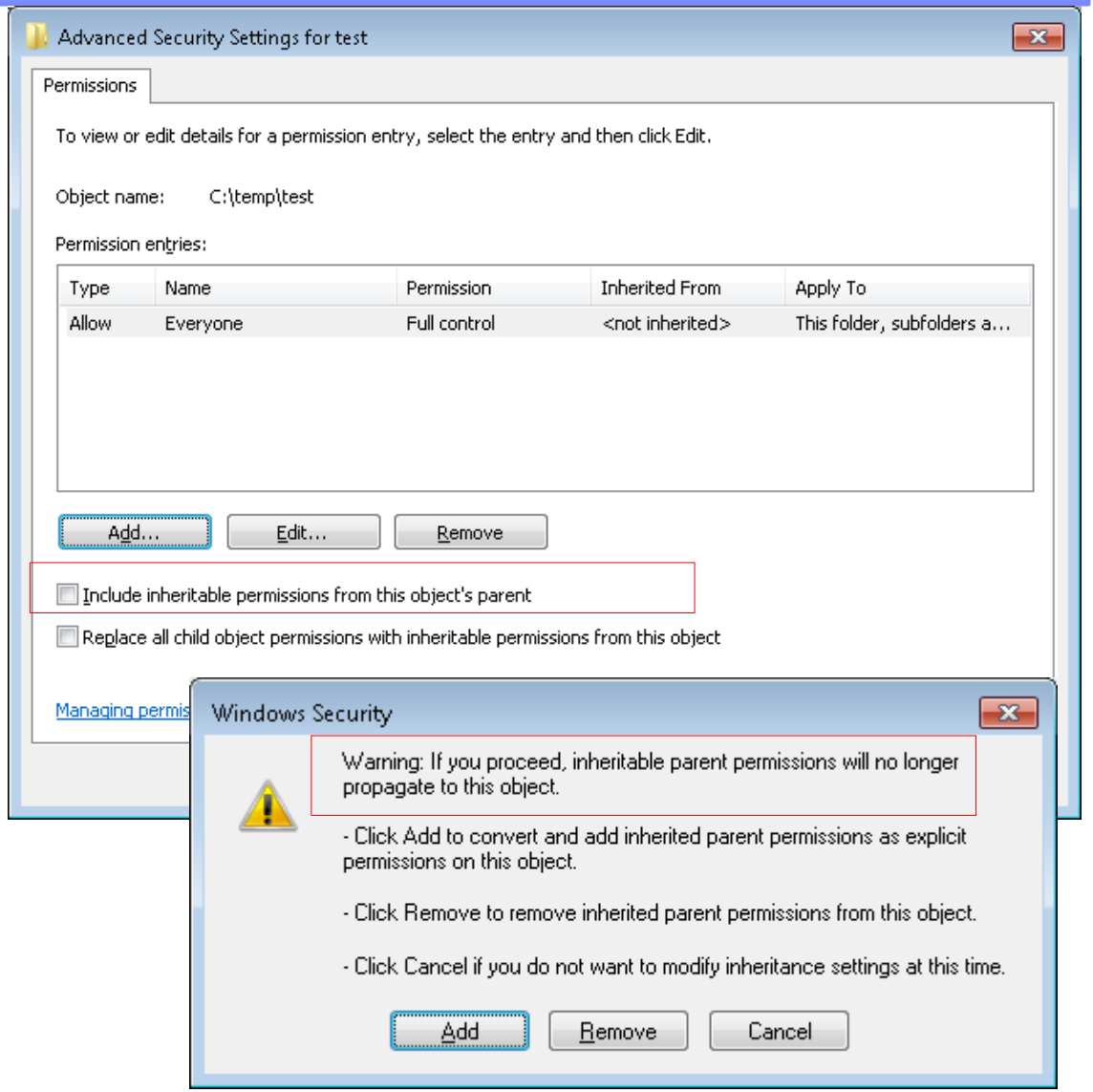


Support Of The Microsoft Management Console



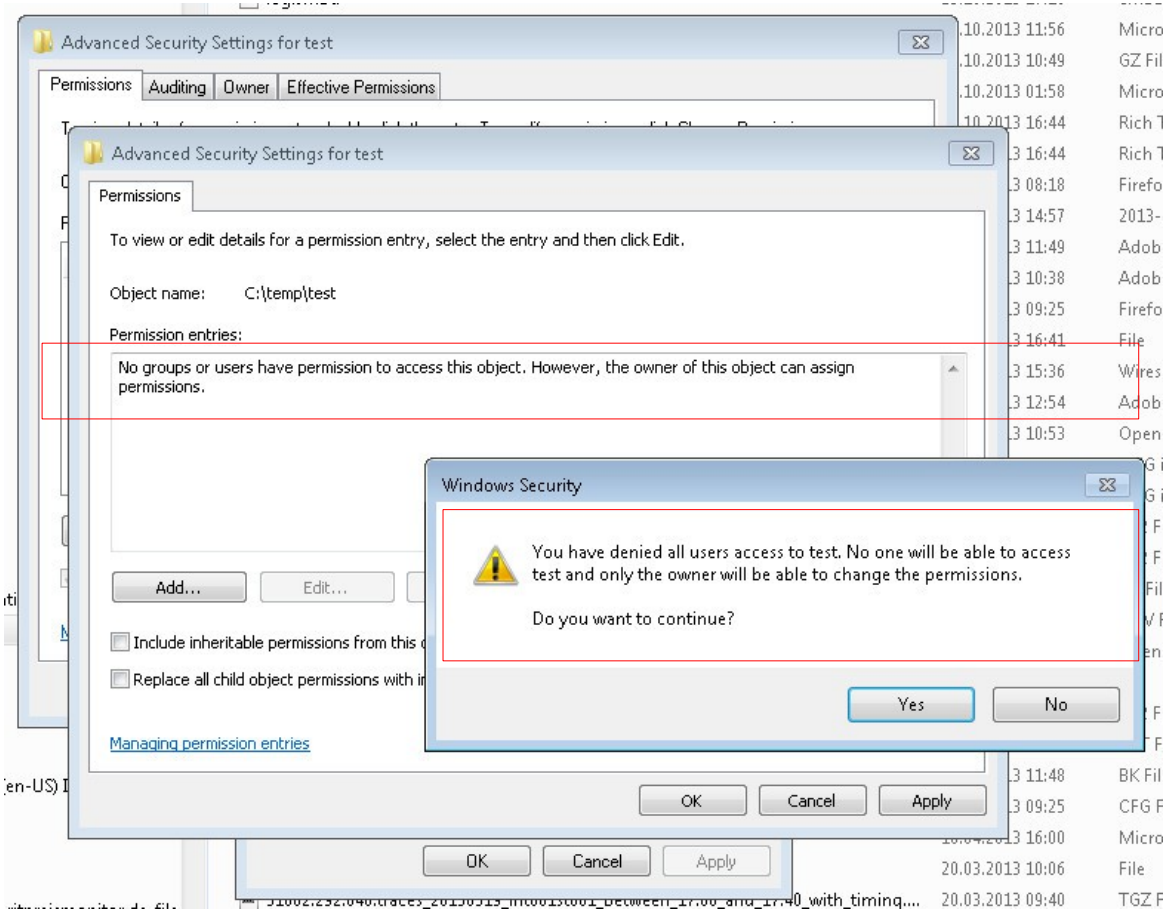
- **SONAS / V7000 Unified currently support**
  - List shares and exports
  - Change export permissions
- **Improvements made**
  - Scaling for large number of connections, counting open files
  - Functions added to filter sessions for more efficient counting
  - Support to set client side caching (csc) policy

ACL Improvements - DACL Protected



- **Customer missed Windows ACL compatibility**
- **Support for the "DACL Protected" ACL flag has been added**
- **Requires support in GPFS**
- **Used to break existing ACL inheritance in a particular sub folder**

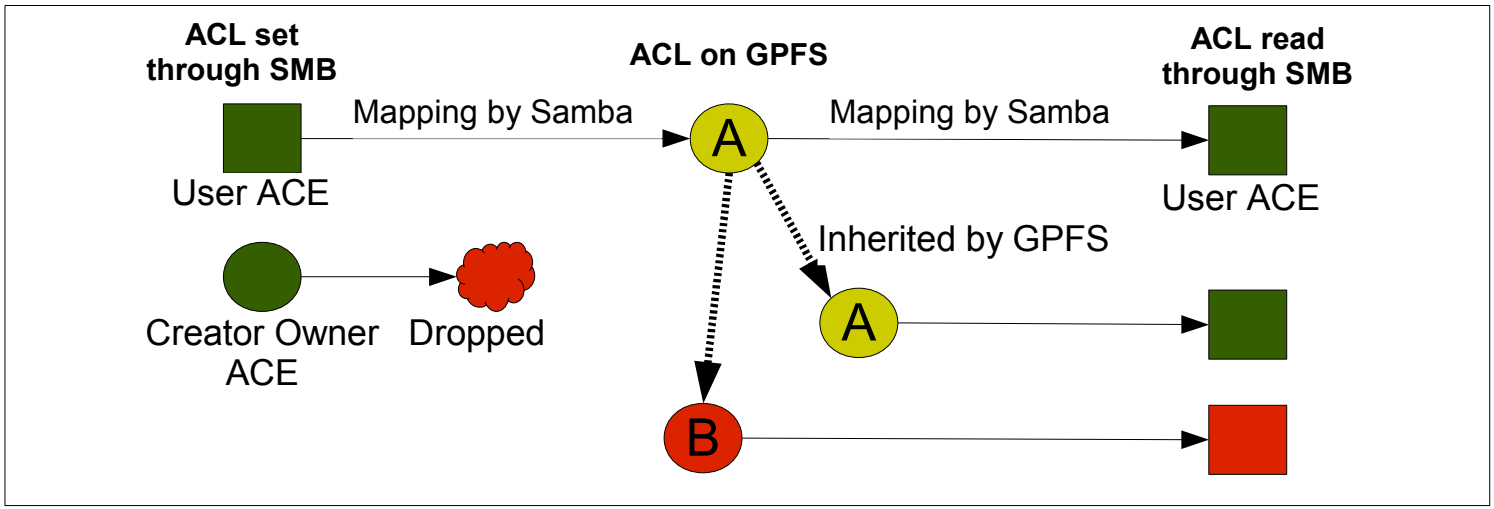
ACL Improvements - Owner can always change ACL



- **User is not able to re-gain control of file ACL if all entries are removed EVEN if owner**
- **Windows-POSIX compatibility issue**
- **Missing Windows ACL compatibility**
- **Use case: Home folder workload with thousands of users: the administrator cannot re-grant access for that many users**

ACL Improvements – Inheritance & Creator Owner

- **Issues**
  - under certain circumstances inheritance works unexpectedly
  - Creator/owner not supported (drop folder use case)



- **Introduction of new nfs:mode = simple cleans up both issues**

Performance And Stability

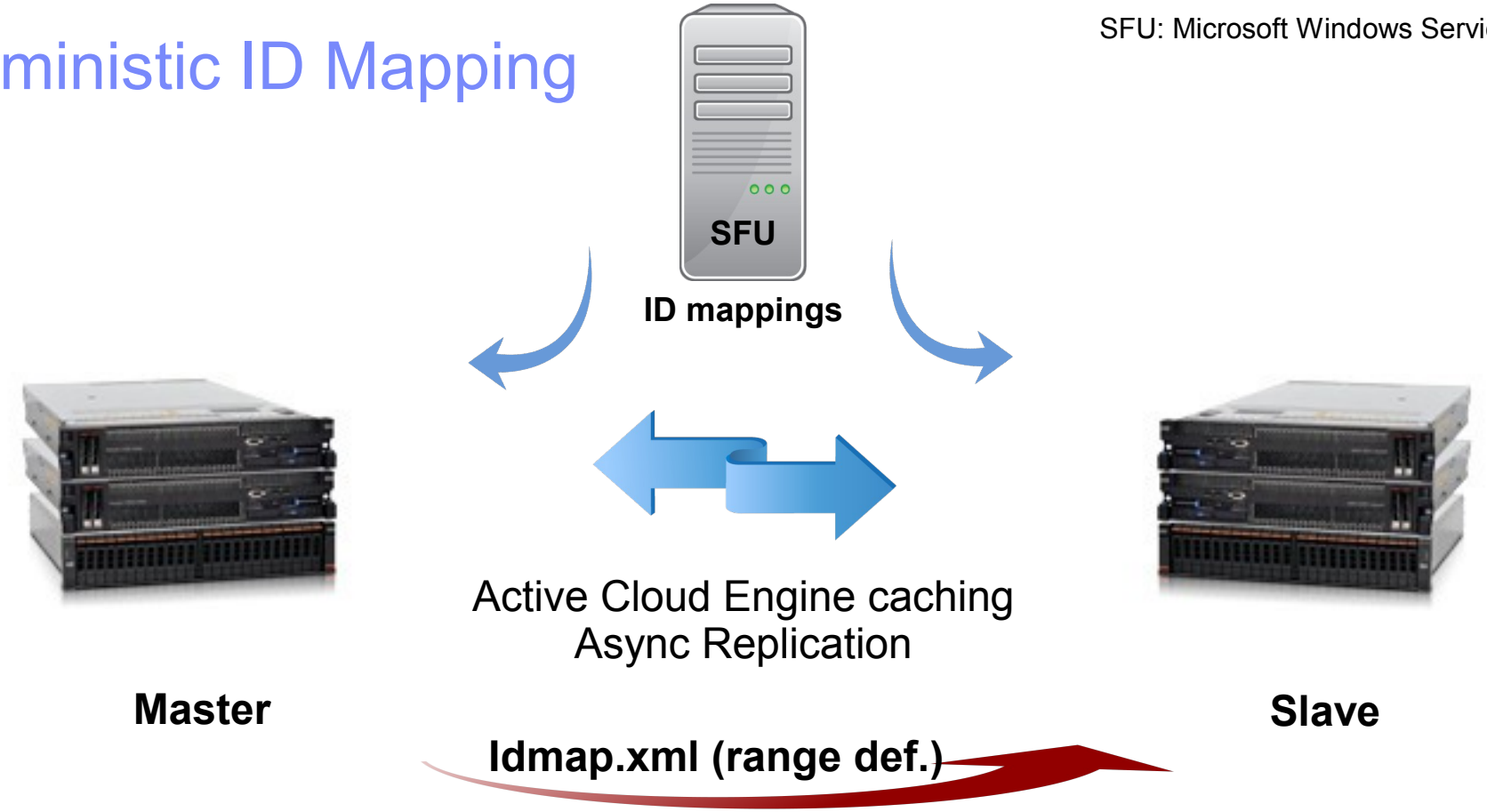
Performance
Typical Benchmarks Specsfs FSCT Etc.

Stability
Scale out test cases Light load (rw, lock) across thousands of connections

- **Replacement of fcntl locks by user space robust mutxes for non-persistent custered tdbs**
- **Mutexes for the (local) gencache_notrans.tdb**
- **Make use of dead records to avoid contention on the lock for the freelist**
- **Add memory cache to gencache tdb**
- **Optimization of create path: reduction of the number of accesses to the locking tdb**

Deterministic ID Mapping

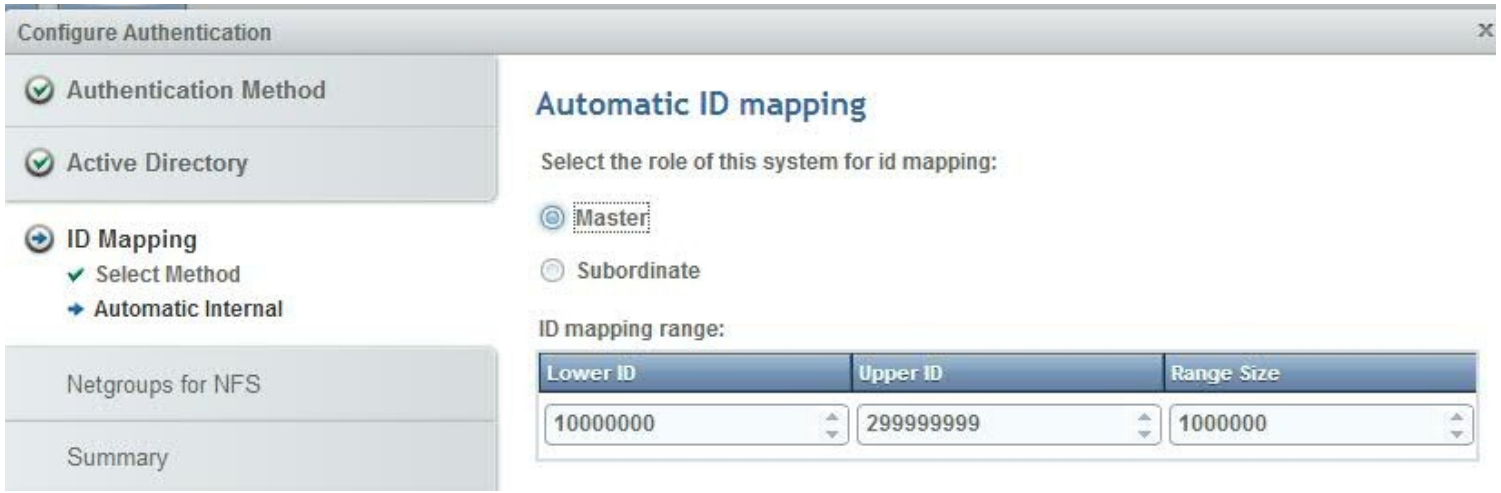
SFU: Microsoft Windows Services for UNIX



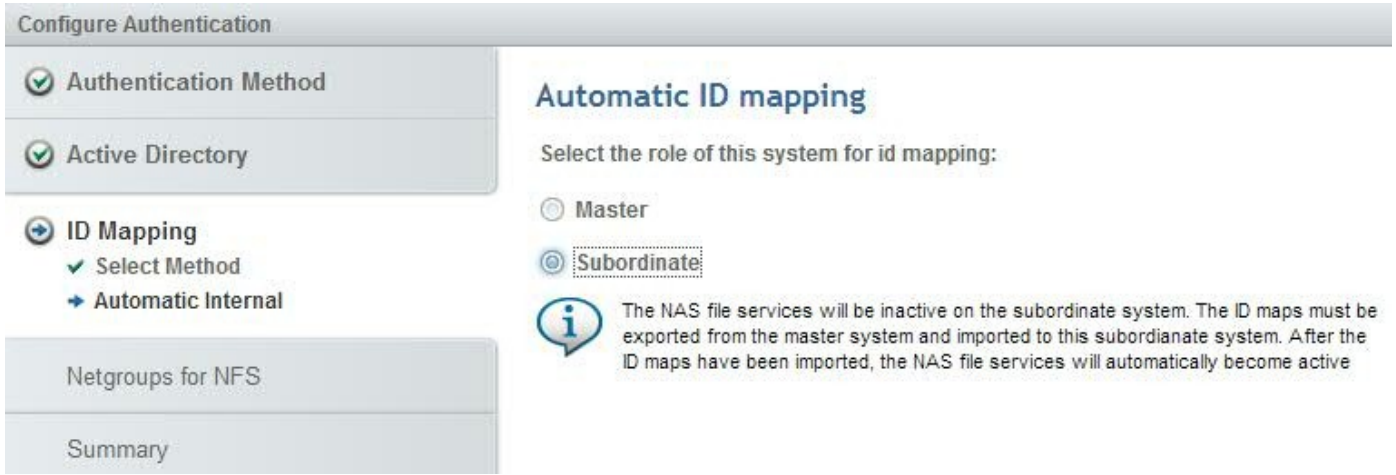
- **Mapping of Windows SIDs to UIDs/GIDs for Linux/GPFS**
- **Goal is to get rid of external ID mapping but maintain consistent mapping across separate systems for all protocols (NFS, SMB, etc.)**
- **Approach: autorid, dedicated range for well-knowns and built-ins, master/slave (read only) concept**

GUI for Deterministic ID Mapping

Master



Slave



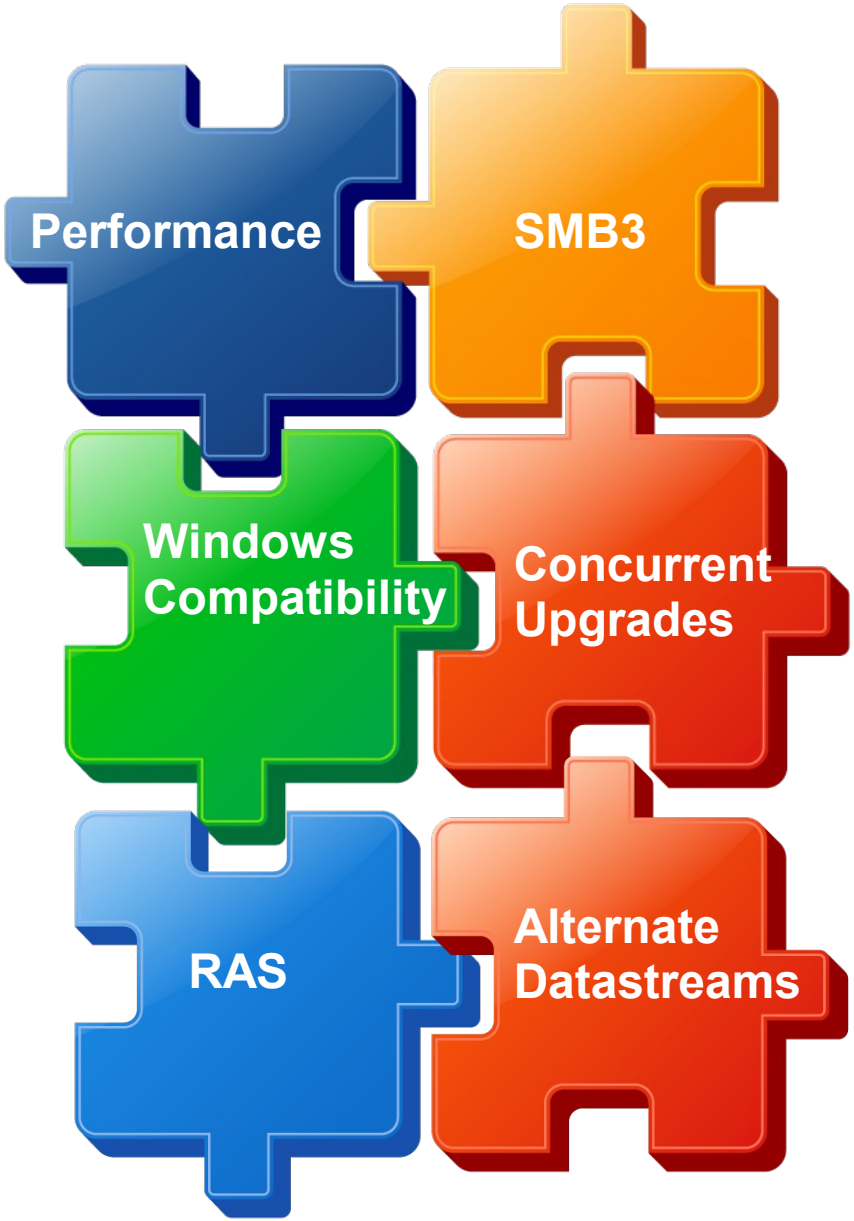
SMB3

- **Requests for SMB 3 support starts rising, especially of the optional features**
  - Multi-channel
  - RDMA
  - Witness protocol
  - Multichannel
  - Persistent File Handles
  - Directory Leasing
  - Support for Storage Features (TRIM, etc.)
- **Samba already supports**
  - SMB3 secure negotiation, improved packet signing, encryption
  - Starting with Samba 4.0

→ see talks of Michael Adam and Stefan Metzmacher

Items for the future

- **Samba**
- **Integration with GPFS**
- **Integration with NFS**
- **Management functions**



Further Details in the Following Talks

- **Wed 14, 2014**
  - 2:15-3:00pm, Stefan Metzmacher (SerNet), A new DCERPC infrastructure for Samba
  - 5:00-5:45pm, Michael Adam (SerNet), Samba, SMB3, Clustering - The Road To Hyper-V!?
  - 5:45-6:30pm, Mathias Dietz (IBM), Samba in a cross protocol environment
- **Thu 15, 2014**
  - 10:45-11:30am, Amitay Isaacs (IBM), Sustaining CTDB Development
  - 11:30-12:15pm, Martin Schwenke (IBM), Scaling IP address handling in CTDB
  - 2:00-2:45pm, Alexander Werth (IBM), Recent improvements in using NFS4 ACLs with Samba

Thank you!

Ingo Meents
meents @ de.ibm.com
06131 / 84-2534
IBM Research & Development
Mainz

Many thanks to
Volker Lendecke, Michael Adam, Stefan Metzmacher, Gregor Beck, Björn Baumbach, Ralph Würthner, Alexander Werth, Christof Schmitt, Peter Somogy, Shekhar Amlekar, Axel Magard, Steven French, Amitay Isaacs, Martin Schwenke, and many others in support, test and sales

Legal Information and Trademarks

The following terms are trademarks of International Business Machines Corporation in the United States, other countries, or both:
IBM, IBM Logo, on demand business logo, Enterprise Storage Server, xSeries, BladeCenter, eServer, ServeRAID and FlashCopy, System Storage, Tivoli, Easy Tier, Active Cloud Engine, General Parallel File System and GPFS.
The following are trademarks or registered trademarks of other companies.
Intel is a trademark of the Intel Corporation in the United States and other countries.
Java and all Java-related trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries.
Lotus, Notes, and Domino are trademarks or registered trademarks of Lotus Development Corporation.
Linux is a registered trademark of Linus Torvalds.
Microsoft, Windows, Windows NT, Windows Vista, Windows Server 2008, Windows 7, Windows Server 2008R2, Windows 2008, Windows Server 8 are trademarks or registered trademarks of Microsoft Corporation.
POSIX is a registered trademark of the IEEE.
UNIX is a registered trademark of The Open Group in the United States and other countries.
Storwize and the Storwize logo are trademarks or registered trademarks of Storwize Inc., an IBM Company.
McAfee and the McAfee logo are trademarks or registered trademarks of McAfee, Inc. in the United States and other countries.
* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.
All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.
The information on the new products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information on the new products is for informational purposes only and may not be incorporated into any contract. The information on the new products is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of any features or functionality described for our products remains at our sole discretion.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.
Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.