

Report of Activities for the Development Infrastructure WG in Japan

Co-operative benchmark evaluation
and development of tools for improving OSS reliability

Project Summary

December 3, 2004

Tomomi Suzuki (Hitachi, Ltd.)

Chairman of Development Infrastructure WG
Japan OSS Promotion Forum

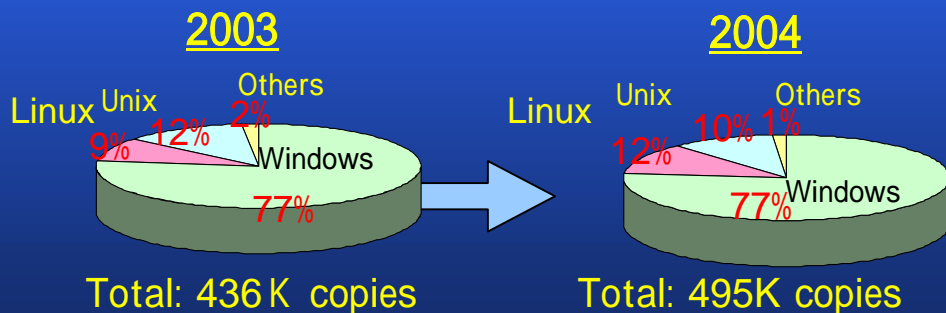


1. Server Linux Market in Japan (1)



Linux is spreading and expanding in the Japanese server market

Server OS Shipments in Japan

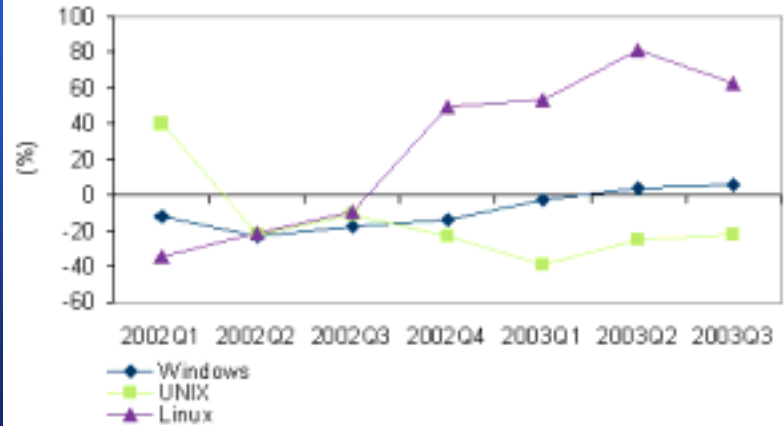


- 2003 39K copies (9%)
2004 59K copies (12%)
- **Will surpass Unix in 2004**
- A five-fold or greater increase in the base value of the Linux market share is expected, for the 5-year period from 2002 to 2007
- **Linux is the only server OS that has maintained a double-digit growth rate since 2002 Q4**

Linux Market in Japan



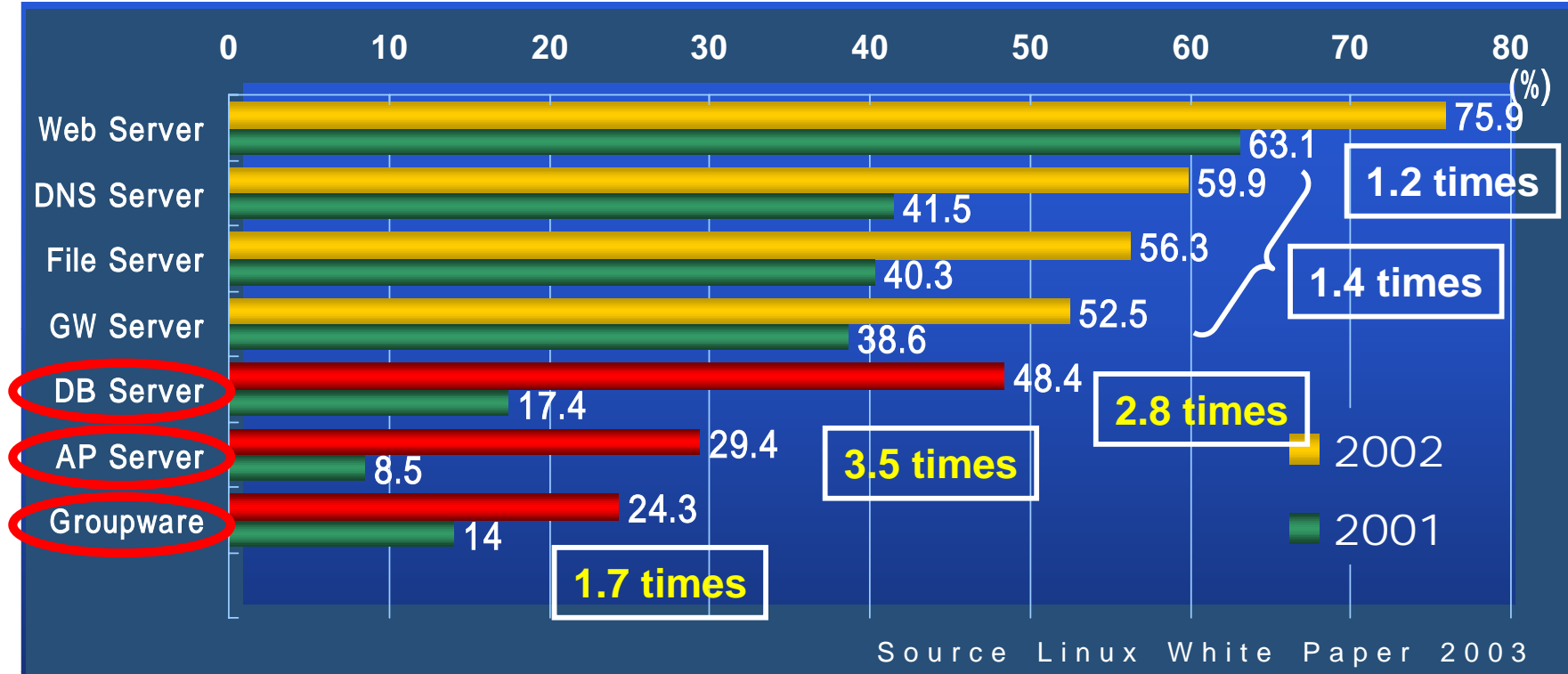
Shipment Value Growth Rate for Each OS



1. Server Linux Market in Japan (2)



What Do Companies Use Linux For?



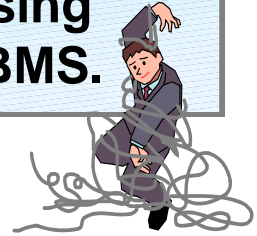
The increase in DB and AP servers is remarkable, while the Internet servers are still the most common.

2. OSS Server Problems



Problems from the Vendor's View

Systems that OSS will support have become complex because more and more user systems are constructed using Linux or middleware, such as application Servers and DBMS.



Moreover...

- Information necessary for the design and construction of efficient and reliable systems is insufficient, leading to similar assessments on the part of each vendor.
- Analysis tools for inspection of system problems or crashes (such as memory dump tools, kernel trace tools, and so on) are insufficient. As a result, it takes time for the vendor to determine the cause of a system problem.



3. Project Mission



Project Mission

**Resolution of vendor-side issues,
so as to encourage wider adoption of Server Linux and OSS
as well as the sharing of OSS knowledge among vendors.**

1. Sharing of knowledge about OSS system design and construction, through co-operative evaluation of OSS reliability, scalability, availability, and performance.

- Sharing of not only results but also processes, tools, and data
- Contribution to wider OSS adoption by making these results available to the community
- Savings on evaluation costs for vendors (especially the 2.6 kernel and new OSS versions)
- Augmenting of system reliability through system construction based on a broader knowledge base

2. Sharing of knowledge about OSS system troubleshooting, through the development of troubleshooting tools such as the memory dump tool and kernel trace tool.

- Sharing of knowledge on how to use these tools and improve skills
- Community contribution through the development and sharing of the necessary tools
- Reduction of troubleshooting analysis time
- Response to the needs of mission critical systems that are using OSS

4. Project Roadmap



Moving forward with the formation, fostering and expansion of the Enterprise Community for spreading the use of Server Linux, in accordance with the following roadmap.

~ 2005/3

Phase 1 Enterprise Community Formation

1. Concentration and friend Structure of a domestic vendor

Concrete project matter :

- Co-operative benchmark evaluation and sharing the results
- Development of tools for improving OSS reliability and sharing the knowledge

2. Starting co-operation with CK

~ 2006/3

Phase 2 Enterprise Community Fostering

1. Expansion of the project scope for Improvement of OSS reliability

-experimental user system
construction by our project

2. Participation of new members

3. Joint development With CK

~ 2007/3

Phase 3 Enterprise Community Expansion

1. Release the results
of joint development
and expansion of user

2. Public presentation
of experimental user
project

3. Expansion of joint
development with CK

5. Project Members and Situation



The OSS development and evaluation consortium was formed with 8 companies, and work has started as of Oct. 2004 as the IPA “Infrastructure Building for OSS” project.

■ WG members

-Chair

-Hitachi, Ltd.

-Member Companies (Consortium Companies)

-Software Research Associates, Inc.

-NTT DATA CORPORATION

-NS Solutions Corporation

-Sumitomo Computer Systems Corporation

-Nomura Research Institute, Ltd.

-MIRACLE LINUX CORPORATION

-UNIADEX, Ltd.

-Member Companies (not Consortium Companies)

-NTT COMWARE CORPORATION

-Nihon Unisys, Ltd.

-Observers

-Novell, Inc., OSDL, Red Hat KK

6. Project Outline (1)



There are 2 sets of tasks in this project

1. Co-operative benchmark evaluation

- (i) Evaluation of the **Java Application layer**
- (ii) Evaluation of the **DB layer**
- (iii) Evaluation of the **OS layer**

Through co-operative benchmark evaluation:



- **share tools and procedures**
- **share results and knowledge**

2. Development of tools for improving Linux and OSS reliability.

- (i) Memory dump inspection tool
- (ii) Kernel trace tool
- (iii) Visual tool for disk allocation assessment



Through co-operative tool development:

- contribute to the improvement of Linux and OSS reliability

6. Project Outline (2)



Evaluation Considerations

1. The target of benchmark evaluation should be **OSS**
2. Benchmark evaluations of **commercial products** will be carried out as necessary, in order to provide a comparison with respect to reliability, efficiency, and scalability, as an enterprise system.
3. Benchmark evaluations for **3-tier Web systems** will be carried out, as these packages are a major part of enterprise systems.

(i) OS layer

- RedHat Linux AS 2.1, 3.0 - The #1 distribution in the Japanese market at present
- SuSE 9.0 - The only distribution based on the 2.6 kernel
- Miracle Linux 3.0 - Equivalent to Redflag (Asianux), and thus useful for CJK co-operation in WG1

(ii) Java Application layer

- Tomcat 5.0 - The #1 web container in the Japanese market at present
- JBoss 4.0 - The only OSS EJB container that is J2EE 1.4 Compatible

(iii) DB layer

- PostgreSQL 7.4 - The #1 RDBMS in the Japanese market at present
- MySQL 5.0 - The #1 RDBMS in the Worldwide market at present

6. Project Outline (3)



Outline of output

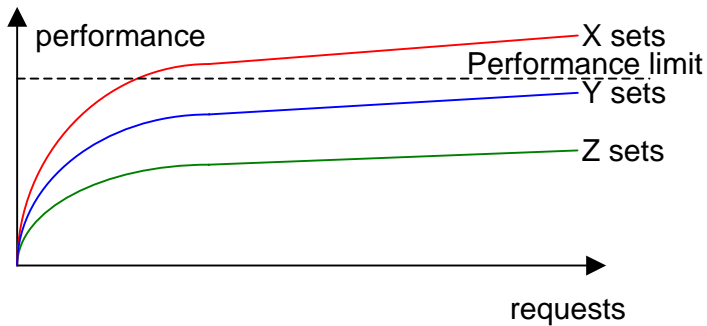
Evaluation environment definition

- construction of hardware & software
- benchmark tools

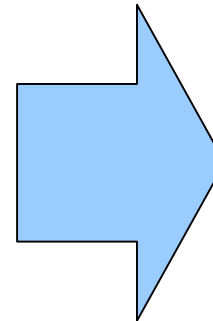
Evaluation procedure report

- installer and parameter information
- benchmark items
- procedure

Evaluation results



Create a sheet for each benchmark evaluation



OSS
Community

Public and
Shared

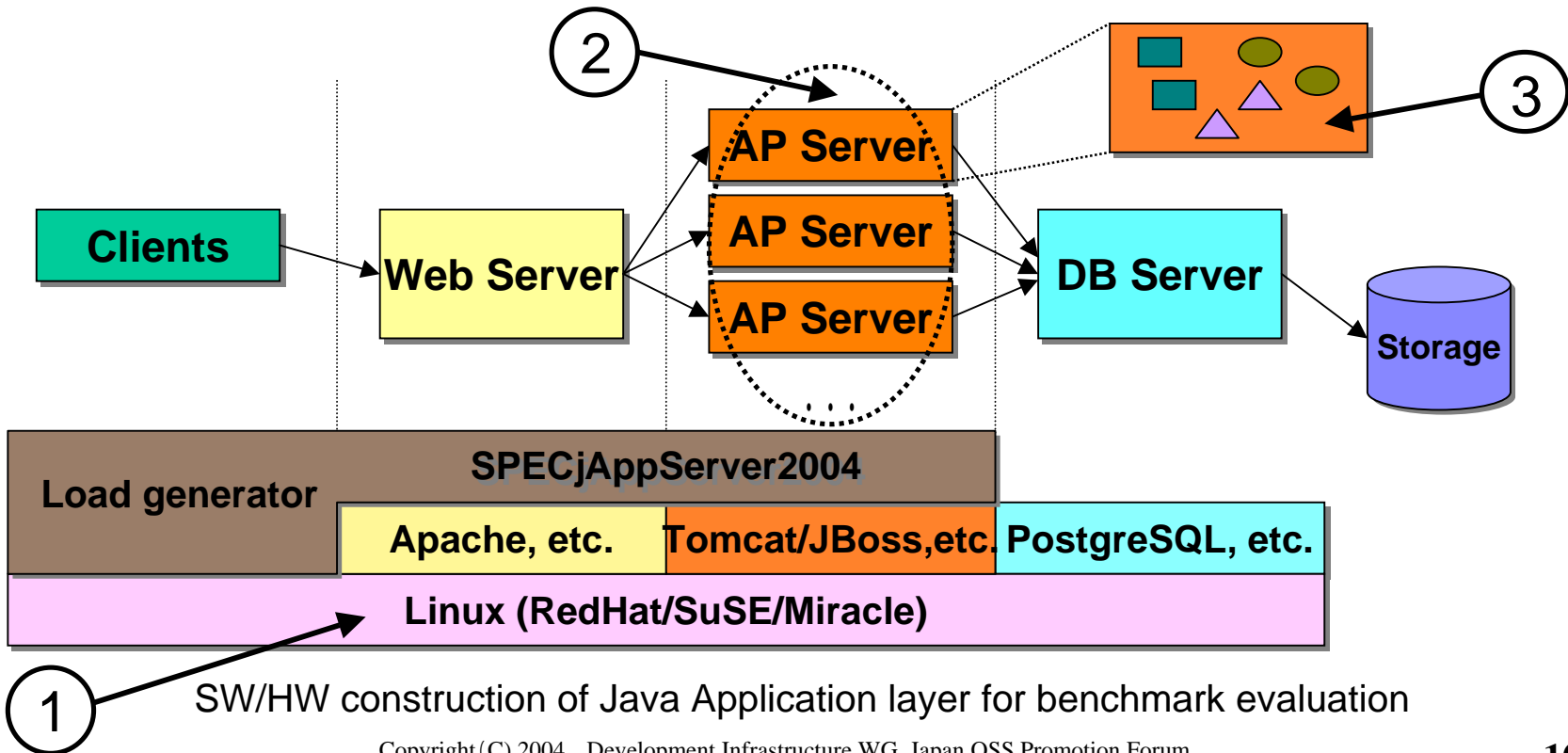


7. Benchmark Evaluation Plan (1) Java Application Layer



Benchmark evaluation items

1. Compare the performance between the 2.4 and 2.6 kernels
2. Evaluation when the application server is clustered
3. EJB Container level analysis
4. Compare OSS (JBoss) and another product (WebLogic)

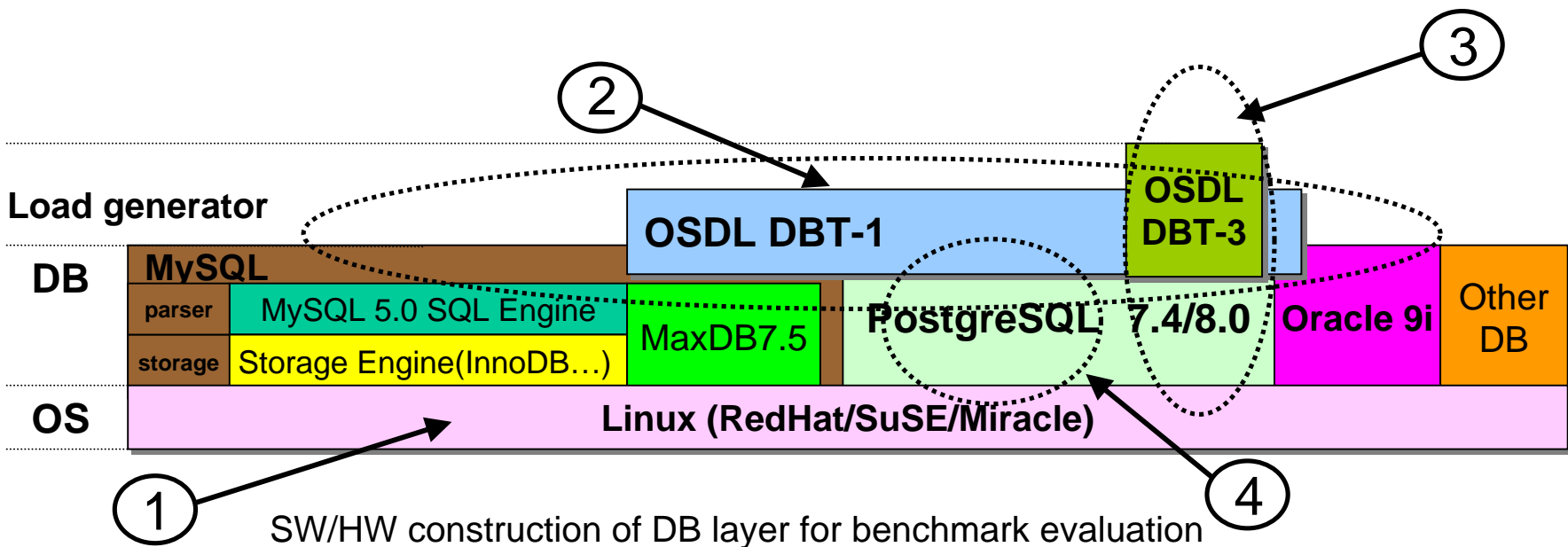


7. Benchmark Evaluation Plan (2) DB Layer



Benchmark evaluation items

1. Compare the performance between the 2.4 and 2.6 kernels
2. Performance evaluation of the Web system
3. Performance evaluation of DSS (Decision Support System)
4. Performance evaluation of the large DBMS
(backup, load, reconstruction of index, etc.)

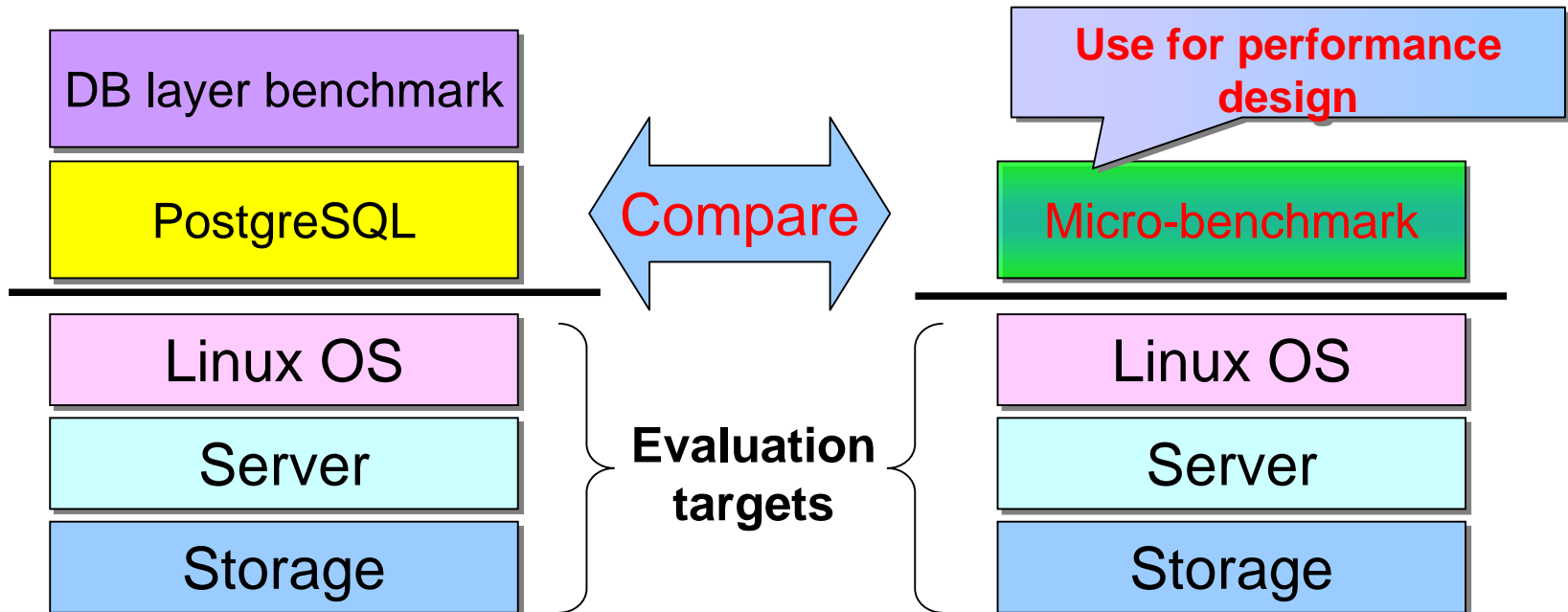


7. Benchmark Evaluation Plan (3) OS Layer



Benchmark evaluation items

1. Compare performance for PostgreSQL I/Os and original micro-benchmark programs that follow the same procedure as PostgreSQL I/Os
2. Evaluation of correlation with PostgreSQL's performance and the above program by using profiles



8. Development of tools for improving Linux and OSS reliability (1)



Enterprise mission critical systems require more powerful RAS

(Minimization of downtime, early failure notification)

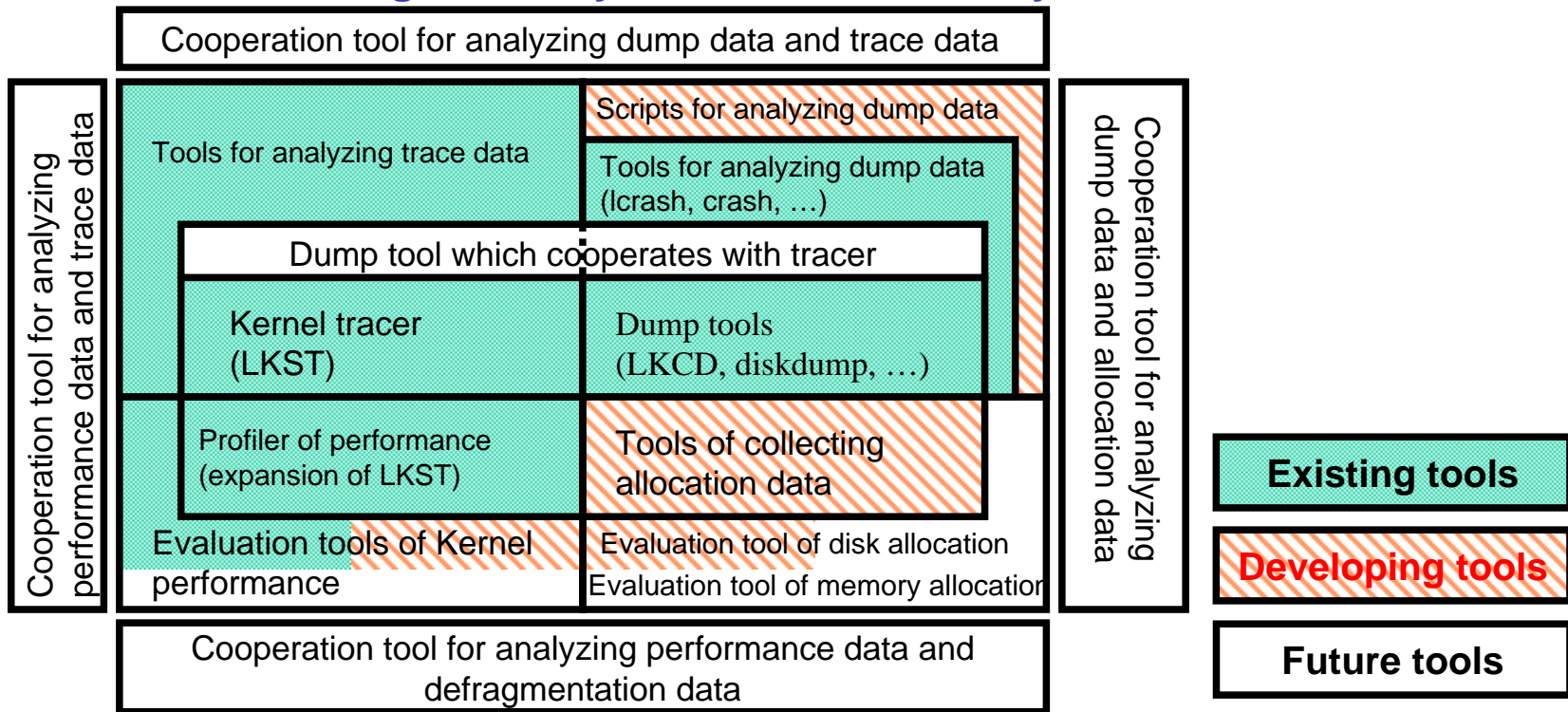
Classify tools for reliability into the following three types:

trace tools, dump tools and disk allocation evaluation tools.

Develop tools that are lacking

(scripts for analyzing dump data, disk allocation evaluation tools)

Tools for realizing reliability in mission critical system



8. Development of tools for improving Linux and OSS reliability (2)



Linux Dump Analysis Tool

LDAI* (Wrapper module)

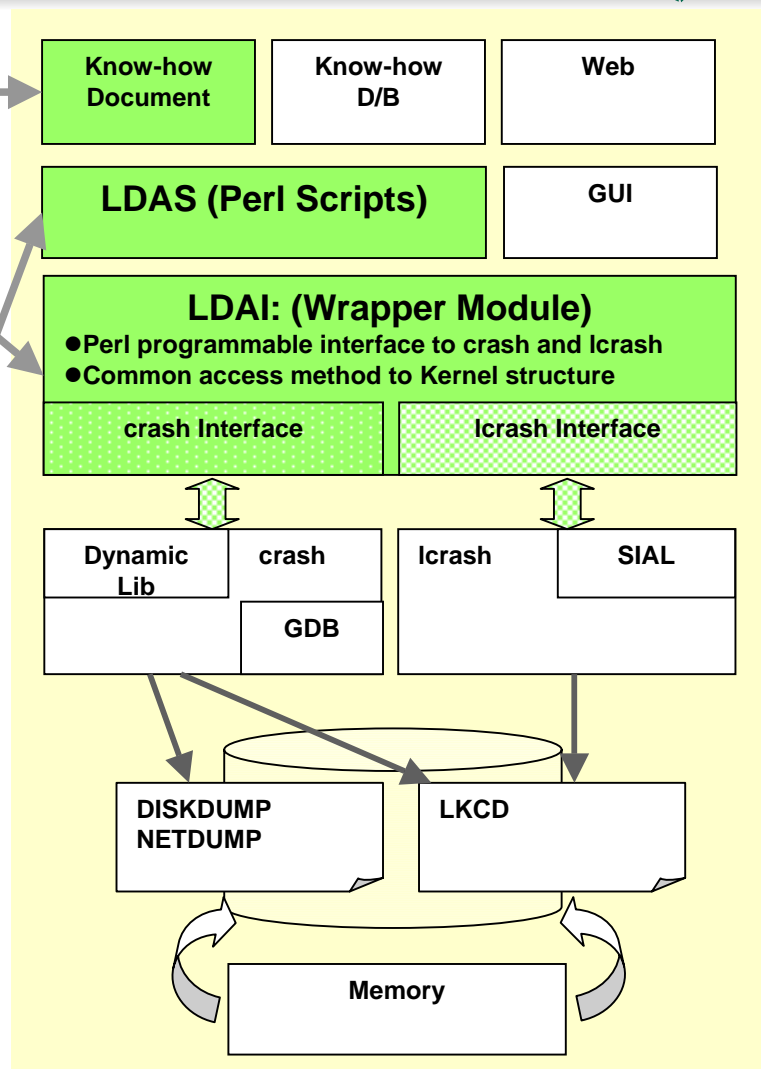
- Provides common access method to kernel by wrapping the existing tool crash/lcrash.
 - Provides new interface that can handle the programmable language Perl.
- (*LDAI: Linux Dump Analysis Interface)

LDAS* (Perl scripts)

- Provides new dump analysis commands using LDAI (Wrapper module).
 - Provides quick dump analysis environment using scripted dump analysis procedures.
- (*LDAS: Linux Dump Analysis Scripts)

Know-how sharing

- LDAS can be a know-how D/B and good educational texts for kernel dump analysis.
- Perl scripts on LDAS can be reused and updated easily and new commands can be generated without reloading LDAI.



8. Development of tools for improving Linux and OSS reliability (3)



LKST (Linux Kernel State Tracer) for evaluating performance of Kernel

Know-how sharing for using the existing tools and new tools properly

- oprofile (<http://oprofile.sourceforge.net/>)
- hardmeter (<http://sourceforge.jp/projects/hardmeter/>)
- Lockmeter (<http://oss.sgi.com/projects/lockmeter>)

Evaluation of kernel performance

- (1) Execution time of system call
- (2) Execution time of IO
- (3) Execution time of user process
- (4) Details of kernel processing time
 - Execution time and execution count of memory allocation
 - Execution time of lock, etc.

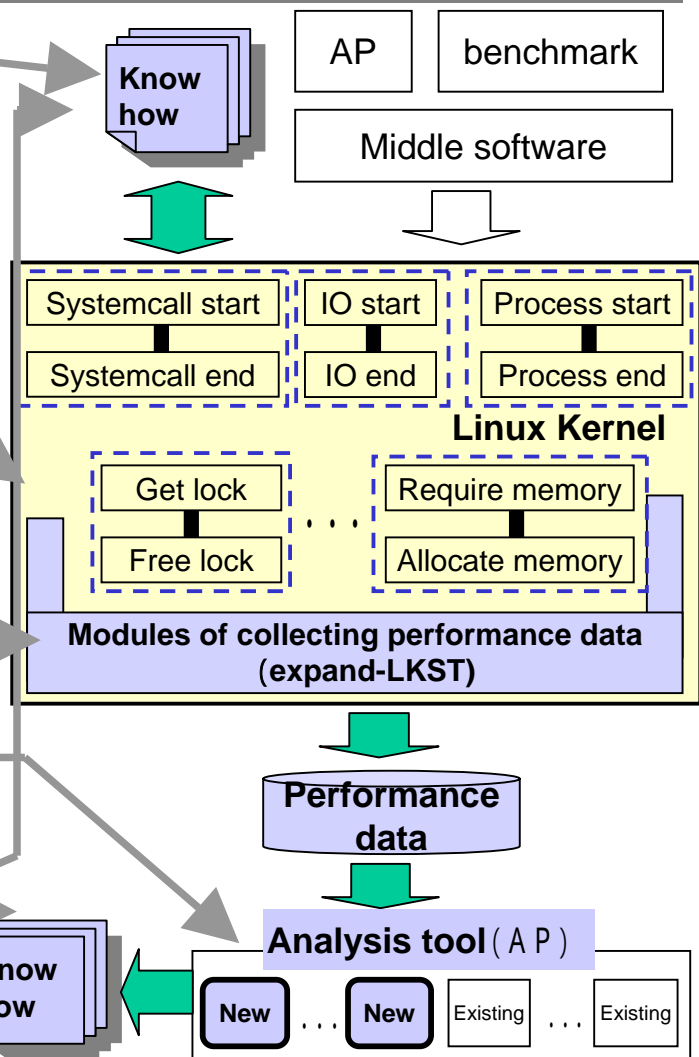
Expand LKST modules for collecting performance data

Develop performance analysis tools

- (1) Extract and visualize of performance data
- (2) Extract and visualize of state transition data

Know-how sharing for analyzing performance

- (1) How to collect data suitable for the evaluation purpose
- (2) How to use the analysis tool of performance





Evaluation tool of disk allocation

Filesystem information collection tools

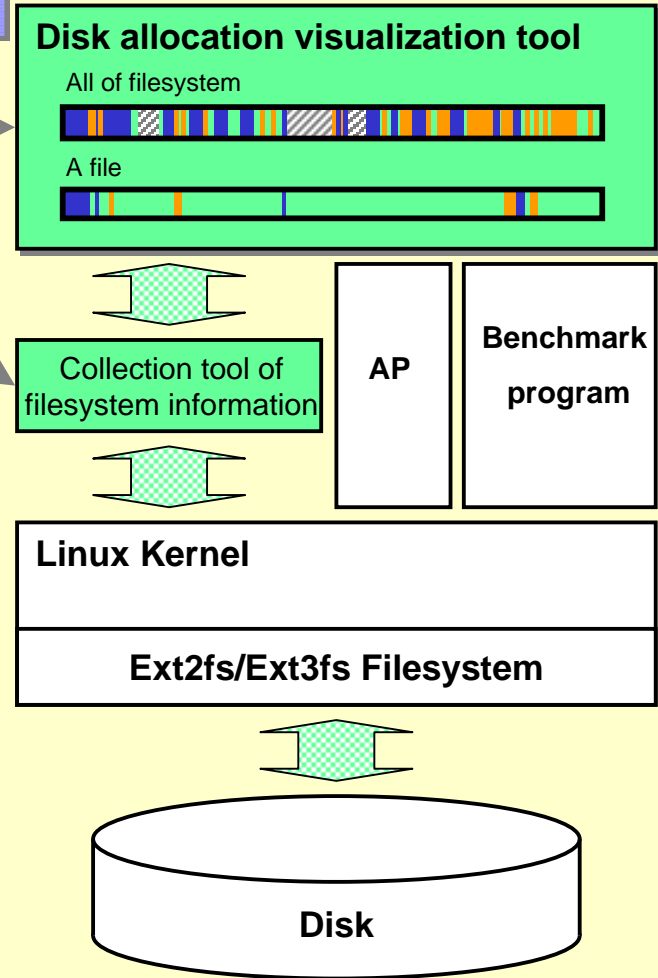
- Research of management information (ext2fs, ext3fs)
- Collect the structure data of filesystem
 - Collect the kind of filesystem
 - Collect the management information, etc. (Ex: partition, mount status)
- Collect the disk allocation management data of each partition

Evaluation of disk allocation

- visualize the disk allocation status of a partition
- visualize the disk allocation status of a file

Know-how sharing

- Evaluation of fragmentation (Influence on performance degradation)
- Evaluation of defragmentation tools
- How to use disk for improving performance



9. Proposal for CJK Co-operation



■ Opinions of Japanese WG members

- There are many combinations of hardware, software, and distributions, so a co-operative benchmark project among CJK will be beneficial for attaining performance data and information on delicate points in system design.
- These activities will contribute to reliable OSS systems in CJK.

■ Our schedule:

- Presentation of part of the benchmark results in Dec. 2004 at the North East Asia Promotion Forum
- Presentation of all benchmark results in this paper in Mar. 2005.



- The Development Infrastructure WG of the Japan OSS Promotion Forum will encourage wider adoption of Server Linux and OSS for enterprise systems.
- We will:
 - Evaluate benchmarks for performance and reliability
 - Develop tools for improving Linux and OSS reliability, and share knowledge and results.
- We will make these results available to the CJK and worldwide communities, and improve them.
- Please work with us on this project through WG1.

