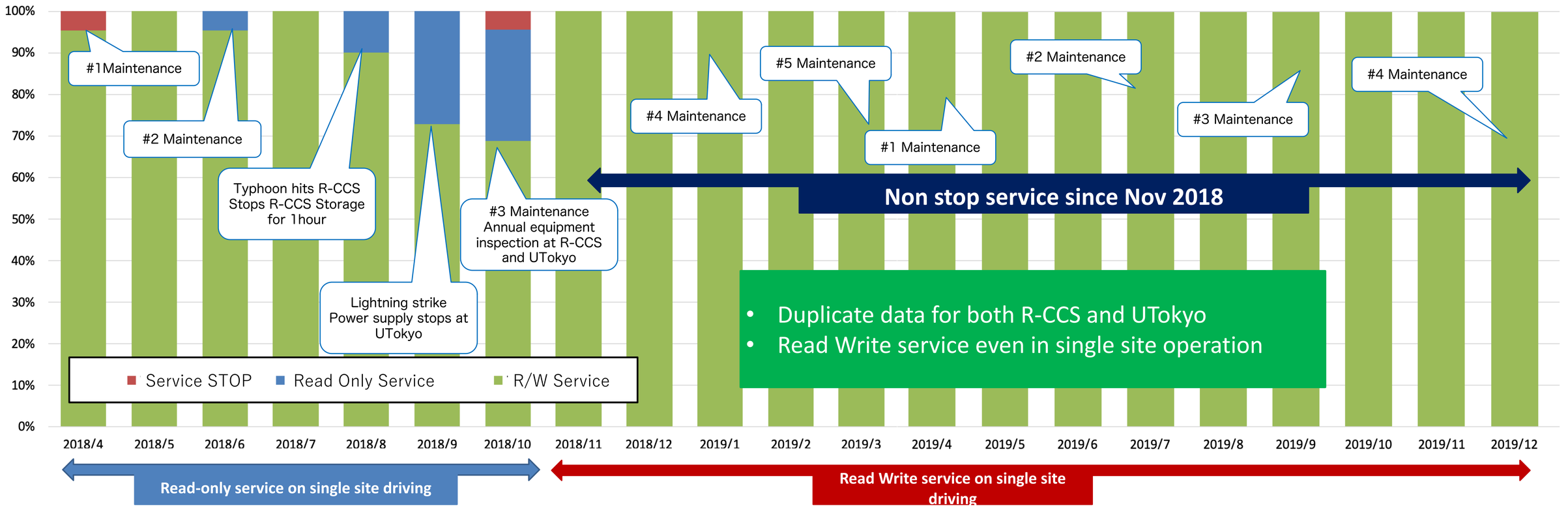
Introduction of HPCI shared storage that has achieved year-round non-stop operation

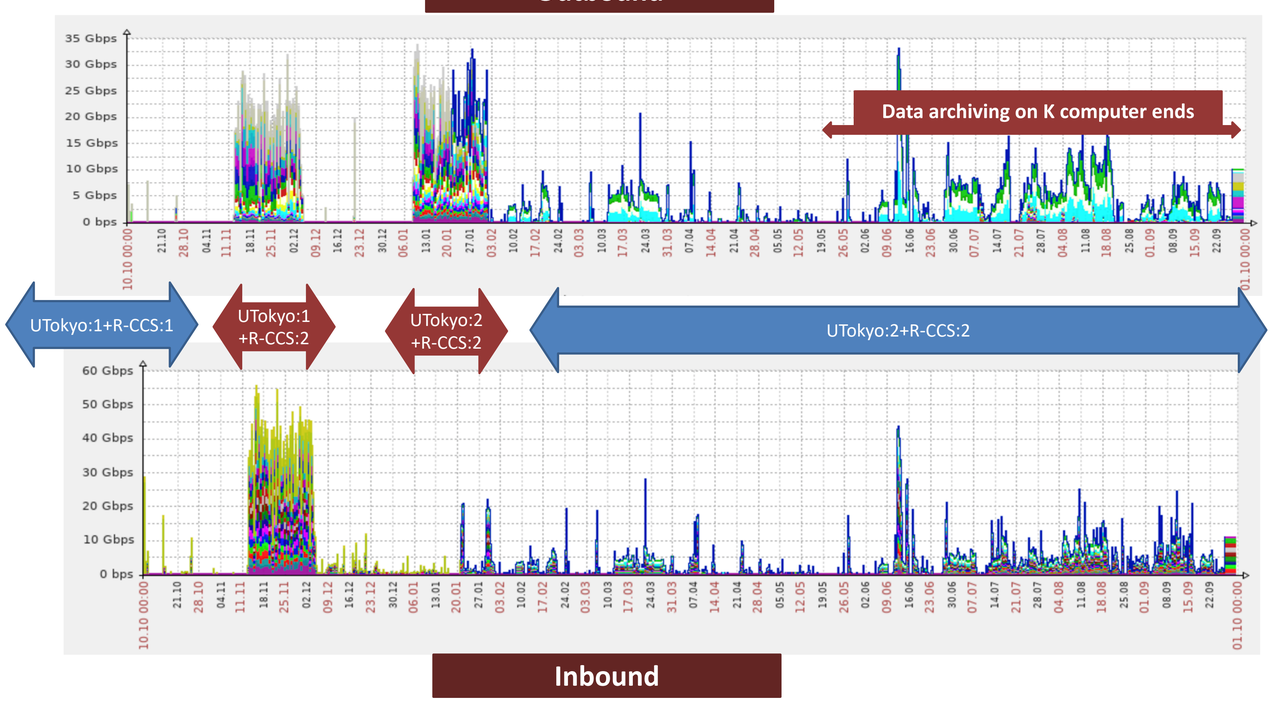
Hiroshi Harada(RIKEN), Osamu Tatebe(University of Tsukuba, Center for Computational Sciences), Toshihiro Hanawa(The University of Tokyo, Information Technology Center), Isamu Koseda(The University of Tokyo, Information Technology Center), Hidetomo Kaneyama(RIKEN), Noriyuki Soda(SRA), Akira Kondo(SOUM), Takahiro Yugawa(SOUM)

Service Availability of HPCI Shared Storage 2018 ～2019

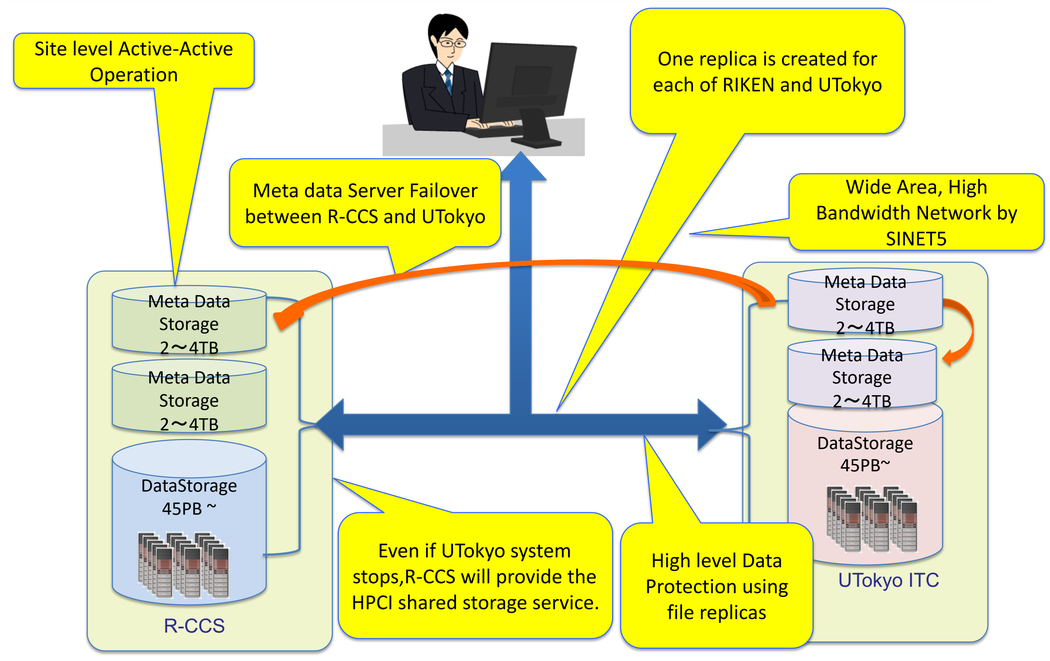


- Duplicate data for both R-CCS and UTokyo
- Read Write service even in single site operation

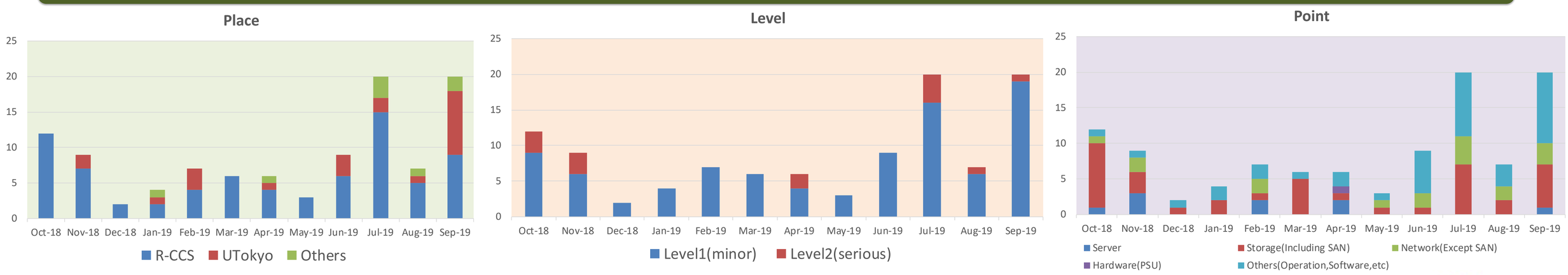
R-CCS network traffic on data multiplexing operation



HPCI shared storage system



Incident Analysis



HPCI shared storage Overview

- Provide 45PBytes single view file space for all the HPCI users.
- Available from all the HPCI super computer resources in Japan.
- HPCI single sign on authentication, if a user wants to transfer the computational results to the HPCI shared storage after the computational session, the user does not have to log in again to the HPCI shared storage.
- Automatic file replication, HPCI shared storage adopts 2 file replicas for each in default.
- File replicas improve access performance from distant clients, and fault tolerance.
- Support Secure network communication, data encryption method such as “gsi” can be specified.
- Support Data Integrity by calculating digest like md5 when accessing files automatically.
- The file digest is calculated at a storage node before writing to a storage, and managed in file system metadata.
- The file digest is also calculated when reading entire data of file.
- The read system call returns input/output error when the digest mismatches.
- The digest is calculated when accessing files sequentially to reduce additional overhead for calculation digest.
- For files which is created by random access write, the digest is calculated when creating replicas.
- The file digest check is also performed when creating a file replicas.
- The data corruption can be detected by comparing metadata digest with file data digest.
- To cope with data corruption during the network transfer from a client, the digest calculation is also supported by client side to ensure end-to-end data integrity.
- Minimum additional overhead and enough data integrity support.