

CersatCloud: stockage et traitements massivement distribué (MooseFS et job distribués)

jeudi 19 novembre 2015 11:40 (40 minutes)

Summary

La présentation portera sur la plateforme big data & cloud que nous avons montée au Laboratoire d'Océanographie Spatiale de l'Ifremer, pour répondre aux besoins de stockage et d'analyse d'archives de données provenant de satellite océanographiques depuis plusieurs décennies. Initialement expérimentale, cette plateforme héberge aujourd'hui plusieurs petaoctets de données et est utilisée pour les travaux quotidiens des scientifiques ainsi que pour des campagnes de reprocessing massives pour les agences spatiales (ESA, CNES). Je présenterai les défis métiers que nous avons relevé ces 5 dernières années en matière de gestion des données, et les solutions techniques déployées qui s'appuient sur les concepts ou technologies du Big-Data (horizontal scalability, distributed filesystems, Hadoop, Spark...) et du Cloud (plateforme mutualisée, environnements virtuels dédiés à la demande).

Auteur principal: M. PAUL, Frédéric (CERSAT IFREMER)

Orateur: M. PAUL, Frédéric (CERSAT IFREMER)