



- (51) International Patent Classification:
G06F 30/20 (2020.01)
- (21) International Application Number:
PCT/IB2021/053514
- (22) International Filing Date:
28 April 2021 (28.04.2021)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
15/930,258 12 May 2020 (12.05.2020) US
- (71) Applicant: **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, New York, 10504 (US).
- (71) Applicants (for MG only): **IBM UNITED KINGDOM LIMITED** [GB/GB]; PO Box 41, North Harbour, Portsmouth Hampshire PO6 3AU (GB). **IBM (CHINA) INVESTMENT COMPANY LIMITED** [CN/CN]; 25/F, Pangu Plaza, No. 27, Central North 4th Ring Road, Chaoyang District, Beijing 100101 (CN).
- (72) Inventors: **PINEL, Florian**; IBM Corporation, c/o Intellectual Property Law, 650 Harry Road, Almaden Research Center, San Jose, California 95120-6099 (US). **BOBBITT, Russell, Patrick**; IBM Corporation, 1101 Kitchawan Road, PO Box 218, Yorktown Heights, New York 10598-0218 (US). **BYRON, Donna**; IBM Corporation, 550 King Street, Littleton, Massachusetts 01460-1250 (US).
- (74) Agent: **LITHERLAND, David**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Hursley, Winchester Hampshire SO21 2JN (GB).

(54) Title: OPTIMIZED DEPLOYMENT OF ANALYTIC MODELS IN AN EDGE TOPOLOGY

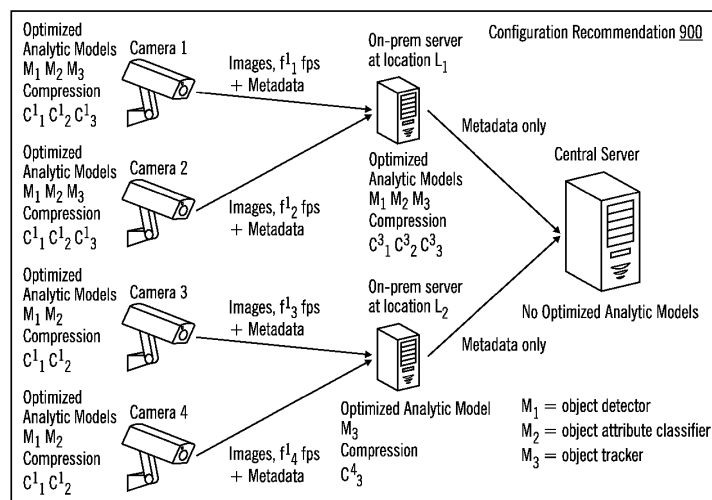


FIG. 9

(57) Abstract: Provided are techniques for optimized deployment of analytic models in an edge topology. A description of a multi-tiered edge topology with a plurality of nodes, a multimedia stream analytics composition, and performance objectives are received, where the multimedia stream analytics composition includes tasks that use analytic models. The analytic models are optimized and clustered to form clusters of optimized analytic models. A representative optimized analytic model is selected from each of the clusters. A configuration recommendation is determined that indicates deployment of the tasks and of each selected representative optimized analytic model on the plurality of nodes to meet the performance objectives. One or more workflows are generated from the configuration recommendation and executed on the plurality of nodes to generate output for the multimedia stream analytics composition.

[Continued on next page]

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))