The Geodetic Cloud Computing Service: a new paradigm in GNSS analysis

Henno Boomkamp

IGS workshop 2016
Sydney 8-12 February
GCCS output: ITRF network polyhedrons @ 30 min
orbits, satellite clocks, Earth rotation parameters are by-products

The GCCS is the operational service of the GPSdancer project
• Routine global network solutions based on GPSdancer instances in the cloud
• Target accuracy < 3 mm RMS w.r.t. formal ITRF for any static receiver
• Guaranteed consistency of models and standards (…IERS 2010, IGS repro 2)
• Operator privacy: input data & station products protected by RSA keys and NDA
• Paid service: participating network operators carry the cost of their own analysis
• Under consideration: guarantee of service covered by GCCS liability insurance

Delaunay triangulation N = 10,000
Ambiguities resolved along Minimum Spanning Tree
Global network analysis v. Precise Point Positioning

**Global network solutions**
- Station geometry: GDOP 1..2
- Satellite geometry: GDOP 10..12
- Global network solutions
  - Station geometry: $\sigma_{\text{sta}} = 2.8\text{ mm}$
    - (IGS SINEX v. formal ITRF)
  - Satellite geometry: $\sigma_{\text{sat}} = 20\text{ mm}$
    - (IGS ACs v. combination)
- Regional PPP solutions
  - Station geometry: $\sigma_{\text{sta}} = 10\text{ mm H, 20 mm V}$
    - (typical RTK networks, therefore: all end-users)
- ~400 IGS stations
- ~30,000 regional CORS stations
- “Noisy” global satellite outputs become rigid PPP inputs

**Why global solutions?**
- Observability of global products is inferior to that of station products: avoid station dependency on abs. orbit accuracy
- Feedback between station products and the global ITRF network realization: **guaranteed analysis consistency**
GPS dancer / GCCS global batch least squares

**Global parameters:** orbits, satellite clocks, Earth rotation parameters

**Station parameters:** position, receiver clocks, troposphere, float ambiguities if relevant

Each GPSdancer instance \( j \) contributes:

\[
\begin{align*}
\{ & \quad G\downarrow j \ x \downarrow g \ & + \ C\downarrow j \ x \downarrow s\downarrow j \ & = \ y\downarrow g\downarrow j \\
& \quad G\downarrow j \ x \downarrow g \ & + \ S\downarrow j \ x \downarrow s\downarrow j \ & = \ y\downarrow s\downarrow j \\
\end{align*}
\]

All station parameters can be pre-eliminated

\[
\begin{align*}
\sum_{j=1}^{N} (G\downarrow j - C\downarrow j S\downarrow j \uparrow -1 C\downarrow j \uparrow t) x \downarrow g & = (y\downarrow g\downarrow j - C\downarrow j S\downarrow j \uparrow -1 y\downarrow s\downarrow j )
\end{align*}
\]

**GPS dancer initial design:** (2 Mb/s)
- “Perfect constraint” solution method requiring square dance accumulation of **vectors**

**GCCS operational approach:** (1 Gb/s)
- *First* accumulate \( G \) on servers per operator
- *Then* accumulate \( G \) among all servers via a square dance accumulation of **matrices**

IGS Workshop 2016 Sydney
Initial GPSdancer design: public P2P process
one process per receiver, on public internet (~2 Mb/s design speed)

GPSdancer design ensures:
• Data and products remain private to the station operators
• Processing capacity can grow along with network size

• Each GPSdancer instance communicates with $\log_2 N$ others
• Connections may change whenever the network changes

No practical solution was found (so far) for maintaining all firewalls!
GCCS operational layout: clustering of GPSdancer instances
still one process per receiver, but **private** cloud servers per operator

- Internet security issues solved by double firewall logic
- Significantly faster communication links (> 1 Gb/s) in cloud

Each server has two IP addresses with separate firewall rules
- Operator has a private IP for input data & output products
- GCCS has second IP for communication between servers
Main evolutions since late 2013:

- GPSdancer instances run in “gated communities” i.e. firewalled cloud servers per network operator
- Cost reduction: GCCS no longer needs investors from industry start-up funding is now 50% ESA (BIC DA) and 50% EU (Hessen/EFRE)
1st Oct 2015: formal creation of the GCCS

How to create a web service
- 400+ pages of forms & documents
- 30+ meetings
- 15+ offices / agencies involved
- 15+ incompatible procedures
- 20+ signatures GCCS
- 20+ signatures 3rd parties
- 10+ months start to finish
- 100+ phone calls
- 150+ e-mails

Geodetic Cloud Computing Service
UG (Haftungsbeschränkt)
Amtsgericht Darmstadt HRB 94625
Ust-Idnr DE 302 314 101
DE70 5085 0150 0000 7556 80
Sparkasse BIC HELADEF1DAS
Step 1: deployment
ICDs with manufacturers
Certificate of cloud service
Step 2: CalVal
mid 2015
Step 3: Pilot project
end 2015

...REFAG2014 plan

...IGS2016 status

10 months of paperwork

- Deployment: Oct 2015:
  3x server @ 8 CPU & 32 GB RAM until 10/2018
  1x server @ 2 CPU & 8 GB RAM indefinitely
  1x virtual server @ 1 CPU indefinitely
  + incidental test capacity @ DARZ Darmstadt

- CalVal to be reduced from 6 months to < 4 months
- Pilot project to be replaced by 3-month free trials for any network operator
Sequential matrix accumulation of $M$ instances per CPU:

- Amount of RAM per CPU is now independent of $M$.
- Cloud computing cost per instance decrease by $1/M$.
- Product latency increases as $L = A + B.M$.
- GCCS baseline: $M = 25$, $L = 30$ minutes.
NEQ accumulation (2): square dance process among servers

“Square-dance” accumulation algorithm between \( n \) servers
- Number of sequential exchange cycles is \( \log_2 n \)
- GCCS home: DARZ (www.da-rz.de) + AWS Frankfurt
- GCCS servers must have 1Gb/s to Frankfurt Ring

\( n = 100 \): 7 cycles \( \times 250 \text{ MB} \times 8 \text{ bits} = 14 \text{ Gb} \leq 14 \text{ seconds} \)

DE-CIX Frankfurt has reached the new record of 5 terabits per second in December 2015
GCCS routine analysis @ 30 min intervals

- Conventional batch LSQ
- 24 hr arc length
- 30 sec output clocks

• Data cleaning & initialization
• LSQ fit “float solution” (zero diff.)
• EMST ambig resolution (double diff.)
• LSQ fit “precise run” (double diff.)
• LSQ fit (zero diff.)

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Arc decoupling</th>
<th>30 min of new data per arc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initialization</td>
<td>Estimated</td>
<td>only 30 min of new data per arc</td>
</tr>
<tr>
<td>New data only</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
<tr>
<td>Fixed</td>
<td>Estimated</td>
<td>station positions &amp; troposphere</td>
</tr>
<tr>
<td>Fixed</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
<tr>
<td>New data only</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
<tr>
<td>Fixed</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
<tr>
<td>Fixed</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
<tr>
<td>New data only</td>
<td>Estimated</td>
<td>clocks</td>
</tr>
</tbody>
</table>

**Ambiguity resolution based on AmbiZap (Blewitt & al.) but adapted to parallel processing, and to single diff. exchange along true EMST baselines.**

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Accurate geometry</th>
<th>Clocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbits</td>
<td>Estimated</td>
<td>Fixed</td>
</tr>
<tr>
<td>Sat clocks</td>
<td>Estimated</td>
<td>Fixed</td>
</tr>
<tr>
<td>ERP</td>
<td>Estimated</td>
<td>Estimated</td>
</tr>
<tr>
<td>Sta pos</td>
<td>Estimated</td>
<td>Fixed</td>
</tr>
<tr>
<td>Sta clocks</td>
<td>Estimated</td>
<td>Estimated</td>
</tr>
<tr>
<td>Troposphere</td>
<td>Estimated</td>
<td>Fixed</td>
</tr>
<tr>
<td>Ambiguities</td>
<td>Estimated</td>
<td>Estimated</td>
</tr>
<tr>
<td>Unresolved amb.</td>
<td>Estimated</td>
<td>Fixed</td>
</tr>
<tr>
<td>Phase L3</td>
<td>2 mm</td>
<td>7 mm</td>
</tr>
<tr>
<td>Code P3</td>
<td>80 cm</td>
<td>80 cm</td>
</tr>
</tbody>
</table>

IGS Workshop 2016 Sydney
The first GCCS solutions: Nov/Dec 2015

Number of receivers processed

<table>
<thead>
<tr>
<th></th>
<th>GCCS</th>
<th>IGS AC</th>
<th>factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 hours of processing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>solution arcs</td>
<td>48 x 24 hr</td>
<td>4 x 30 hr</td>
<td>9.6</td>
</tr>
<tr>
<td>global parameters estimated</td>
<td>450,000</td>
<td>50,000</td>
<td>9.0</td>
</tr>
<tr>
<td>station parameters estimated</td>
<td>16,500,000</td>
<td>650,000</td>
<td>25.4</td>
</tr>
<tr>
<td>observations processed</td>
<td>90,000,000</td>
<td>3,400,000</td>
<td>26.5</td>
</tr>
</tbody>
</table>

Mean station baseline error vs ITRF 2014 (mm), day 335-348 RMS 9.3 mm
**Cost breakdown**

GCCS is a *paid service*: network operators must carry the cost for what they use

<table>
<thead>
<tr>
<th>GCCS</th>
<th>For every 1000 stations, GCCS can fund 1 full time job or equiv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liability insurance</td>
<td><strong>Optional</strong>: collective insurance offers GCCS Guarantee of Service</td>
</tr>
<tr>
<td>ITRF backbone</td>
<td><strong>Should phase out</strong> if IGS operators are willing to pay their own cost</td>
</tr>
<tr>
<td>Cloud computing</td>
<td><strong>Variable</strong> - depends on operator requirements, e.g.:</td>
</tr>
<tr>
<td></td>
<td>- Redundant processing (geographically separated servers)</td>
</tr>
<tr>
<td></td>
<td>- Physical security level</td>
</tr>
<tr>
<td></td>
<td>- Optional cloud archiving of data and/or products</td>
</tr>
</tbody>
</table>

$\text{GCCS is a paid service}$: network operators must carry the cost for what they use

Current GCCS server capacity is available until at least October 2018
- 650 stations = **400 IGS + 250 free trials**, 25 stations @ 3 months per operator

\[ \text{~$500/yr} \]

\[ \ldots \text{first come, first served – requests to info@ITRF.online} \]

IGS Workshop 2016 Sydney
Do we still need the IGS/ITRF?

...yes! The GCCS merely complements the IGS

- The **IGS** defines an accurate reference frame for just a few hundred points on Earth
- The **GCCS** transfers the ITRF to all regional stations (and therefore to end-users)

What would happen if IGS disappears?

- The IGS tends to emphasize its **products**, but probably more important are:
  - Continuous feedback among state-of-art GNSS analysis from all agencies
  - International cooperation among many participating organisations
Summary

• GCCS places all stations in **global** analysis along with the ITRF stations
  – *as opposed to:* feeding regional stations with orbits and clocks as ITRF reference
  – All stations get direct baselines to the ITRF stations, i.e. ITRF position time series
  – IGS ACs routinely demonstrate that accuracy of this analysis *can* be 2.8 mm RMS

• GCCS cloud deployment (…3 large servers) started Oct 2015
  – Generous computing capacity secured until at least Oct 2018

• Formal CalVal campaign against IGS expected from May to Sep 2016
  – Publication of CalVal report marks the formal start of GCCS operations

• **Free trials will be available** for all interested network operators
  – 64 trial slots, up to 3 months of GCCS analysis for 25 sites (economic value ~$3000)
  – Questions & requests for trial slots: info@itrf.online

**Next IGS workshop: routine ITRF realizations for 10,000+ sites?**