



NASA-GPT: Retrieval- Augmented Generation on NASA Technical Reports

Derek Dalle

ARC/TNA

November 19, 2024

Unclassified ▾

CUI

NASA files, contains CUI
🔒 usa_citizen

Unclassified

NASA files, Unclassified Unlimited

HECC

Supercomputing Knowledge Base



At high subsonic speeds, what is the approximate limit of chord-height ratio for small-perturbation theory?



SWOT satellite target accuracy RMS



Python script to generate Spice Toolkit CK file



What was JPL's Budget in 2017?


Message...



Retrieval-Augmented Generation is (roughly) Generative AI + citations

1. Search a database of documents
2. Find the best documents
3. Find the best “chunks” of those documents
4. Read those chunks into the context window
5. Then let the Large Language Model do its thing

Unclassified ▾

New Chat 

What is the size of the inlet bleed holes on the XB-70?

```
<|python_tag|>{"name": "search_nasa", "parameters": {"sentence": "XB-70 inlet bleed holes size"}}
```

1. (U) Smith, R. H. (1966) [Initial flight experience with the XB-70 air-induction system](#)
2. (U) Goecke, S. A. (1968) [Base pressure measurements on the XB-70 airplane at Mach numbers from 0.4 to 3.0](#)
3. (U) Putnam, T. W. (1970) [XB-70 compressor-noise reduction and propulsion-system performance for choked inlet flow](#)
4. (U) Fukuda, M. K. (1975) [Control of shock wave-boundary layer interactions by bleed in supersonic mixed compression inlets](#)
5. (U) Gallagher, R. J. (1970) [Investigation of a digital simulation of the XB-70 inlet and its application to flight-experienced free-stream disturbances at Mach numbers of 2.4 to 2.6](#)
6. (U) Davis, David O. (GRC) (2012) [Research on Supersonic Inlet Bleed](#)
7. (U) Sanders, Bobby W. (GRC) (2008) [Aerodynamic Design of a Dual-Flow Mach 7 Hypersonic Inlet System for a Turbine-Based Combined-Cycle Hypersonic Propulsion System](#)

The size of the inlet bleed holes on the XB-70 is as small as three-hundredths of an inch, according to [1].

Message...



Unclassified

New Chat

What is the size of the inlet bleed holes on the XB-70?

```
<|python_tag|>{"name": "search_nasa", "parameters": {"sentence": "XB-70 inlet bleed holes size"}}
```

1. (U) Smith, R. H. (1966) [Initial flight experience with the XB-70 air-induction system](#)
2. (U) Goecke, S. A. (1968) [Base pressure measurements on the XB-70 airplane at Mach numbers 2.4 to 2.6](#)
3. (U) Putnam, T. W. (1970) [XB-70 compressor-noise reduction and propulsion-system performance](#)
4. (U) Fukuda, M. K. (1975) [Control of shock wave-boundary layer interactions by bleed in supersonic flow](#)
5. (U) Gallagher, R. J. (1970) [Investigation of a digital simulation of the XB-70 inlet and its application to the inlet](#)
6. (U) Davis, David O. (GRC) (2012) [Research on Supersonic Inlet Bleed](#)
7. (U) Sanders, Bobby W. (GRC) (2008) [Aerodynamic Design of a Dual-Flow Mach 7 Hypersonic Inlet System for a Turbine-Based Combined-Cycle Hypersonic Propulsion System](#)

The size of the inlet bleed holes on the XB-70 is as small as **three-hundredths of an inch**, according to [1].

... on the throat ramps were found to be clogged. For polishing the inlet had filled portions of the bleed holes as small as **three-hundredths of an inch.** ... was a contributor to the instability of shock wave unstart at Mach number 3.

... at engine stalls have not been a problem with the



NASA-GPT has several document databases embedded

- NASA Technical Reports Server (NTRS) <https://ntrs.nasa.gov>
- JPL Technical Reports Server <https://trs.jpl.nasa.gov>
- High-End Computing Capability knowledge base <https://www.nas.nasa.gov/hecc/support/kb>
- Adding much more

AI turns the HECC “Knowledge Base” into an advanced “help” capability with NASA-specific knowledge.

Generative AI does a pretty good job of these tasks on its own, but this fills in some of the blanks, e.g. that NASA’s Cascade Lake nodes are called `cas_ait` and have 40 cores per node.

HECC ▾

New Chat ↗

How do I start an interactive PBS job on 4 Cascade Lake nodes for 6 hours using the group 'e0847'. Use the queue 'normal' and explain the different parts of the command.

```
<|python_tag|>{"name": "search_nas_hecc_kb", "parameters": {"sentence": "starting interactive PBS job on Cascade Lake nodes"}}
```

1. (U) NAS Support (2024) [Running Jobs with PBS / PBS on Pleiades / PBS Job Queue Structure](#)
2. (U) NAS Support (2021) [Systems Reference / Aitken / Preparing to Run on Aitken Cascade Lake Nodes](#)
3. (U) NAS Support (2021) [Running Jobs with PBS / PBS on Pleiades / Pleiades devel Queue](#)
4. (U) NAS Support (2023) [New User Orientation / Basic Tasks](#)
5. (U) NAS Support (2024) [Systems Reference / GPU Nodes / PBS Job Sharing on GPU Nodes](#)
6. (U) NAS Support (2021) [Running Jobs with PBS / PBS on Pleiades / Releasing Idle Nodes from Running Jobs](#)
7. (U) NAS Support (2020) [Running Jobs with PBS / PBS on Electra / Running "At-Scale" Jobs in the sky_wide Queue](#)
8. (U) NAS Support (2020) [News / 2020 / Reminder: Changes to PBS Job Requests for Pleiades V100 GPU Resources on July 16](#)
9. (U) NAS Support (2021) [Running Jobs with PBS / Optimizing/Troubleshooting / Common Reasons Why Jobs Won't Start](#)
10. (U) NAS Support (2017) [News / 2017 / PBS Scheduling Updated for Broadwell Jobs](#)

To start an interactive PBS job on 4 Cascade Lake nodes for 6 hours using the group `e0847` and the queue `normal`, you can use the following command:

```
qsub -I -l select=4:ncpus=40:mpiprocs=40:model=cas_ait -l walltime=6:00:00 -q normal -W group_list=e0847
```

Here's a breakdown of the different parts of the command:

- `qsub` : This is the command to submit a job to PBS.
- `-I` : This option tells PBS to start an interactive job.
- `-l select=4:ncpus=40:mpiprocs=40:model=cas_ait` : This option specifies the resources required for the job. In this case, we're requesting:
 - `select=4` : 4 nodes
 - `ncpus=40` : 40 CPUs per node
 - `mpiprocs=40` : 40 MPI processes per node
 - `model=cas_ait` : Cascade Lake nodes (specifically, the `cas_ait` model)
- `-l walltime=6:00:00` : This option specifies the maximum wall-clock time for the job (in this case, 6 hours).

Message...



LFC: Large File Control helps manage versions of large files

Open-source package at:

<https://github.com/nasa/lfc>

Documentation at:

<https://nasa.github.io/lfc-doc>

nasa / lfc

Type to search

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

lfc Public

Edit Pins Watch 2 Fork 0 Star 0

main 1 Branch 0 Tags

Go to file Add file Code

nasa-ddalle Add support for mmap on large files f567ad0 · last month 123 Commits

doc	Modify conf.py title, version, copyright	2 months ago
lfc	Add support for mmap on large files	last month
test	Add support for mmap on large files	last month
.aerohub.json	Auto-commit repo options in '.aerohub.json'	last year
.flake8	Add flake8 settings to template	last year
.gitignore	Update doc template to include test results	2 years ago
LICENSE	Add text version of full license	3 months ago
README.rst	Add link to GitHub pages in README	2 months ago
pdb_test_py3.sh	Update test scripts w/ report gen and final exit statys	2 years ago
run_test_py3.sh	Test 1.0.0 to 100%	last year
setup.py	Add support for mmap on large files	last month
tox.ini	Add coverage rc using tox.ini	last month

About

LFC is an extension to git for managing large files

nasa.github.io/lfc-doc/

Readme View license Activity Custom properties

0 stars 2 watching 0 forks

Report repository

Releases

No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

LFC works like `git`, with `git`, and is parallel to `git`

- Suppose you have a repo `myrepo/`
- There is an `lfc init` command that looks like `git init`.

```
$ lfc init
```

```
$ lfc remote add pfe23:/nobackup/user/myrepo.git/cache -d
```

- Now suppose you have a large database file `database.xlsx`

```
$ lfc add database.xlsx
```

```
$ lfc push
```

```
$ git push
```

- LFC creates `database.xlsx.lfc`, and git tracks that file ... that's it!



Derek Dalle <derek.j.dalle@nasa.gov>