2018 Predictive Analytics Symposium

Session 09: A/P - Sharing Shiny Applications

SOA Antitrust Compliance Guidelines
SOA Presentation Disclaimer

Sharing Shiny Applications

Guy Yollin

2019 SOA PREDICTIVE ANALYTICS SYMPOSIUM

Milliman

Important notices

- The views expressed in this presentation are those of the presenter, and are not necessarily shared by Milliman, Inc., RStudio, Inc., or the Society of Actuaries

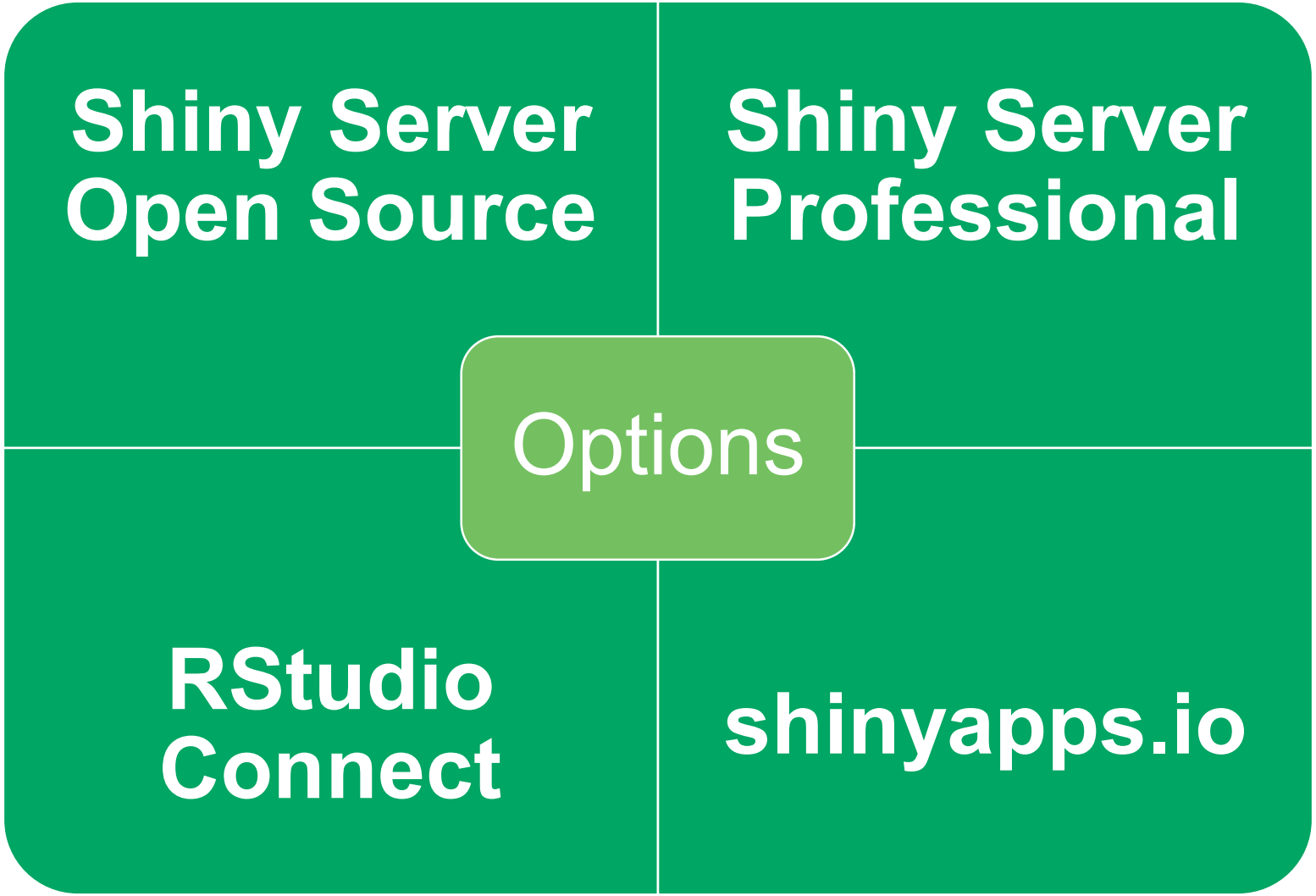
- This presentation contains a number of graphics that were produced by RStudio, Inc. and found on their website rstudio.com; attribution is noted on the slides where they appear

- This presentation is not sponsored or approved by RStudio, Inc.

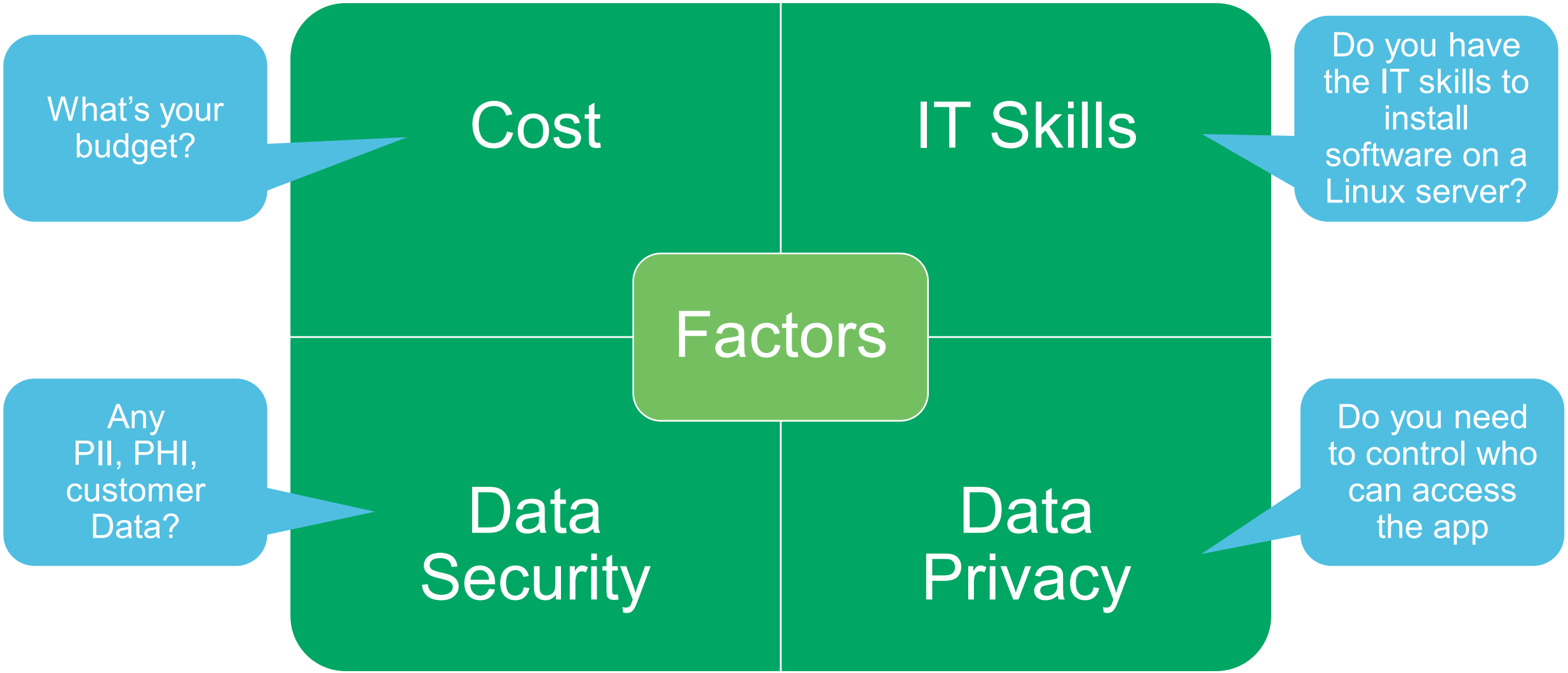
Agenda

- Introduction
- Quick overview of Shiny architecture
- Detailed look at Shiny hosting options
- Cost and analysis
- Real deployment examples
- Q & A

Shiny hosting options



Primary factors to consider when hosting Shiny apps



More comprehensive list of factors to consider

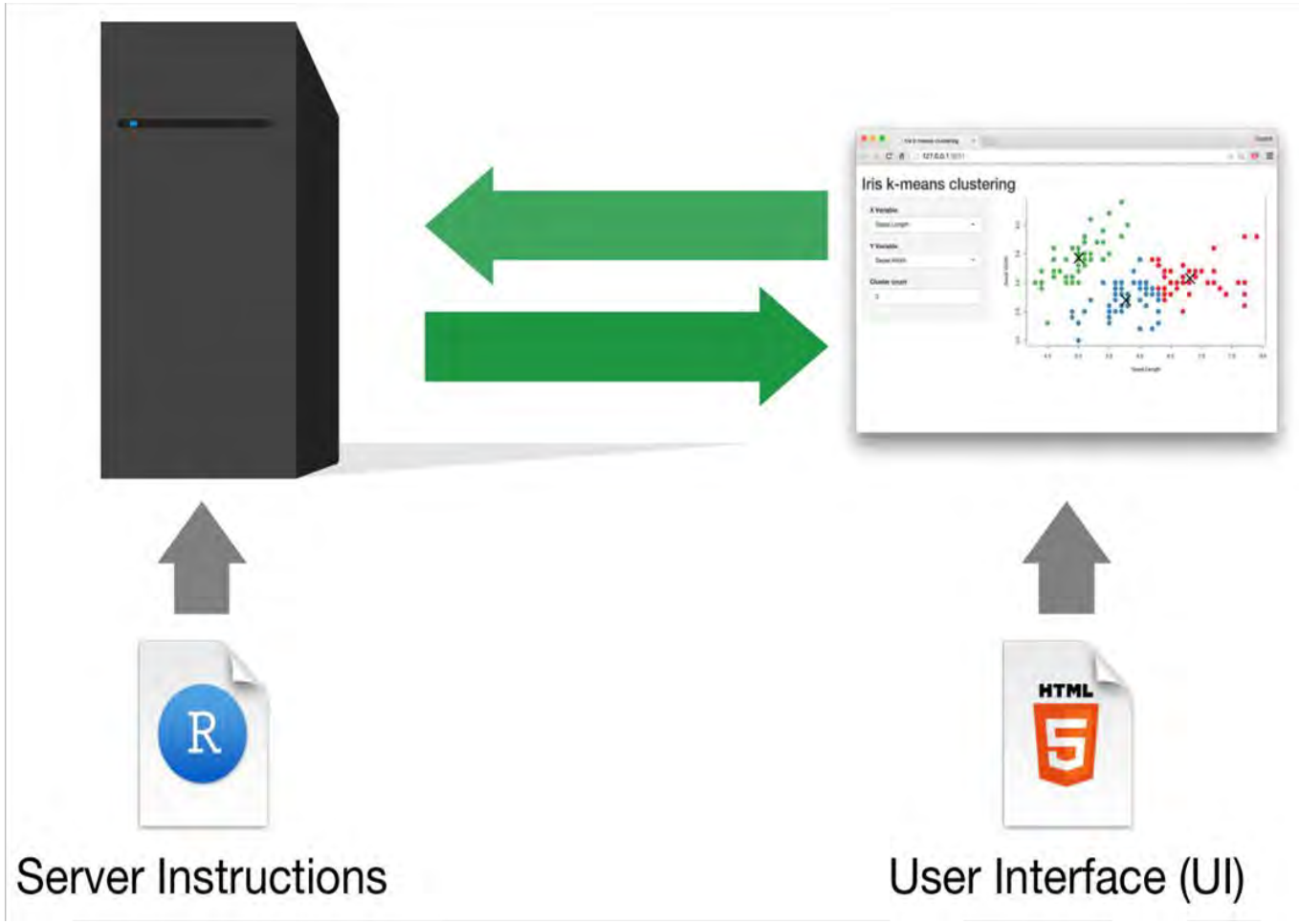
- Primary factors
  - Cost
  - IT skills
  - Data privacy
  - Data security

- Additional factors
  - Application performance
    - Startup time and application responsiveness
  - How many users do you need to support
  - How many apps do you need to host
  - Do you need to manage users
    - Control their app usage and resource utilization

Overview of Shiny architecture

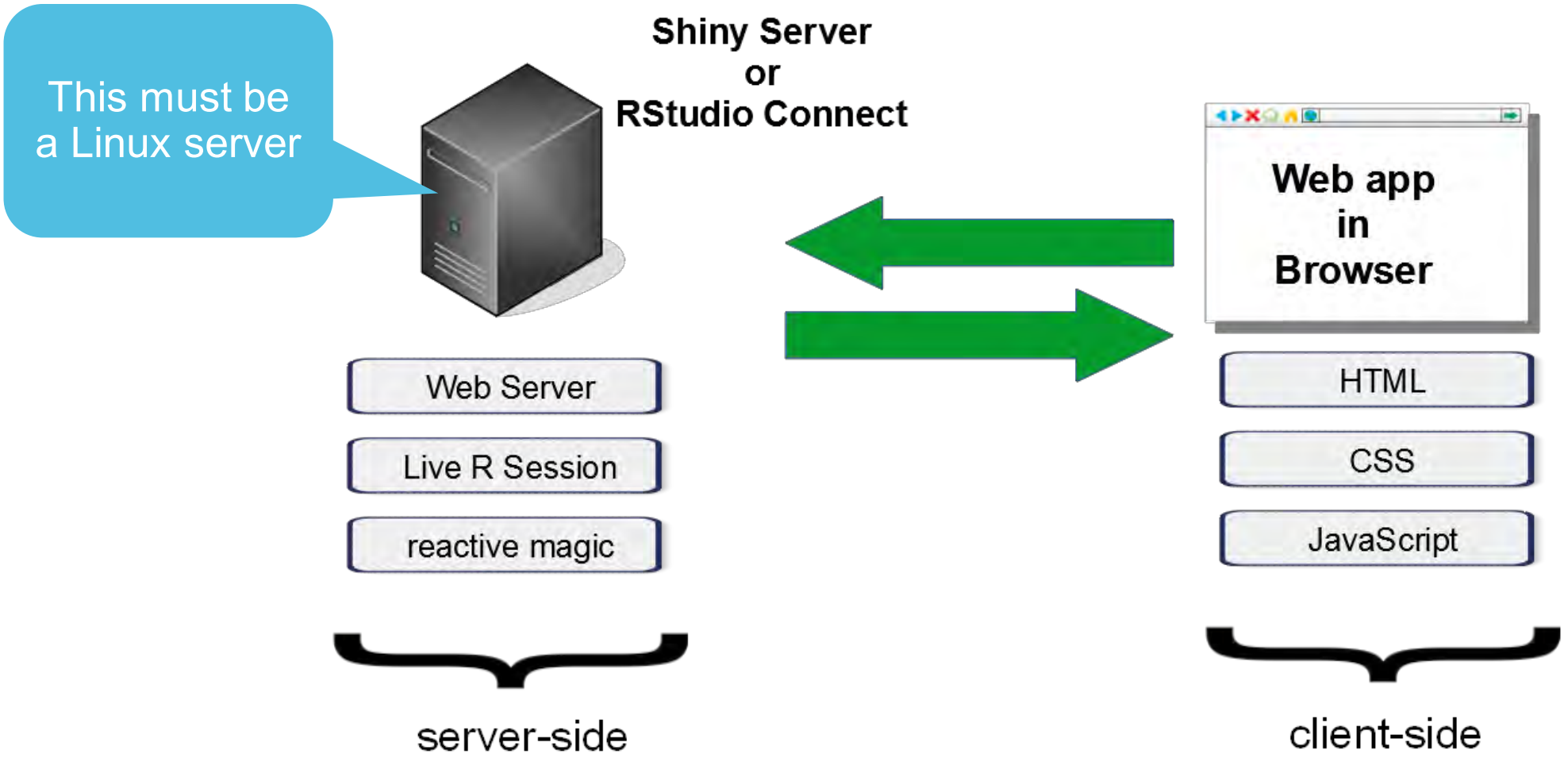
Logistics of shiny app hosting

Shiny client-server architecture (very high level)



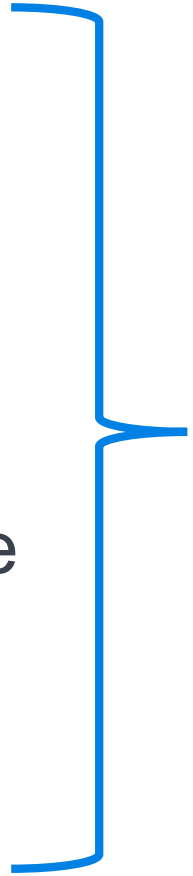
Graphic found on rstudio.com

Shiny client-server architecture (high level)



Logistics of Shiny app hosting

1. Setup and configure a Linux server
   - Ubuntu, Redhat/CentOS, SLES
   - On-premises or cloud-based
2. Install R
   - CRAN, Microsoft, Tibco, Oracle, etc.
3. Install and configure Shiny hosting software
   - **Shiny Server Open Source**
   - **Shiny Server Pro**
   - **RStudio Connect**
4. Install your shiny application



shinyapps.io simply provides an infrastructure with the first 3 steps already completed

Deep dive into hosting options

Comparison

Distribution/Cost/Support

	Shiny Server Open Source	Shiny Server Professional	RStudio Connect	shinyapps.io
Hosting Shiny apps with a web interface	✓	✓	✓	✓
Product or Service	**Product**	**Product**	**Product**	**Service**
Distribution Method	**Installed**	**Installed**	**Installed**	**Hosted**
Free	✓	✗	✗	✗ - paid tier ✓ - free tier
RStudio Support	✗	✓	✓	✓ - paid tier ✗ - free tier

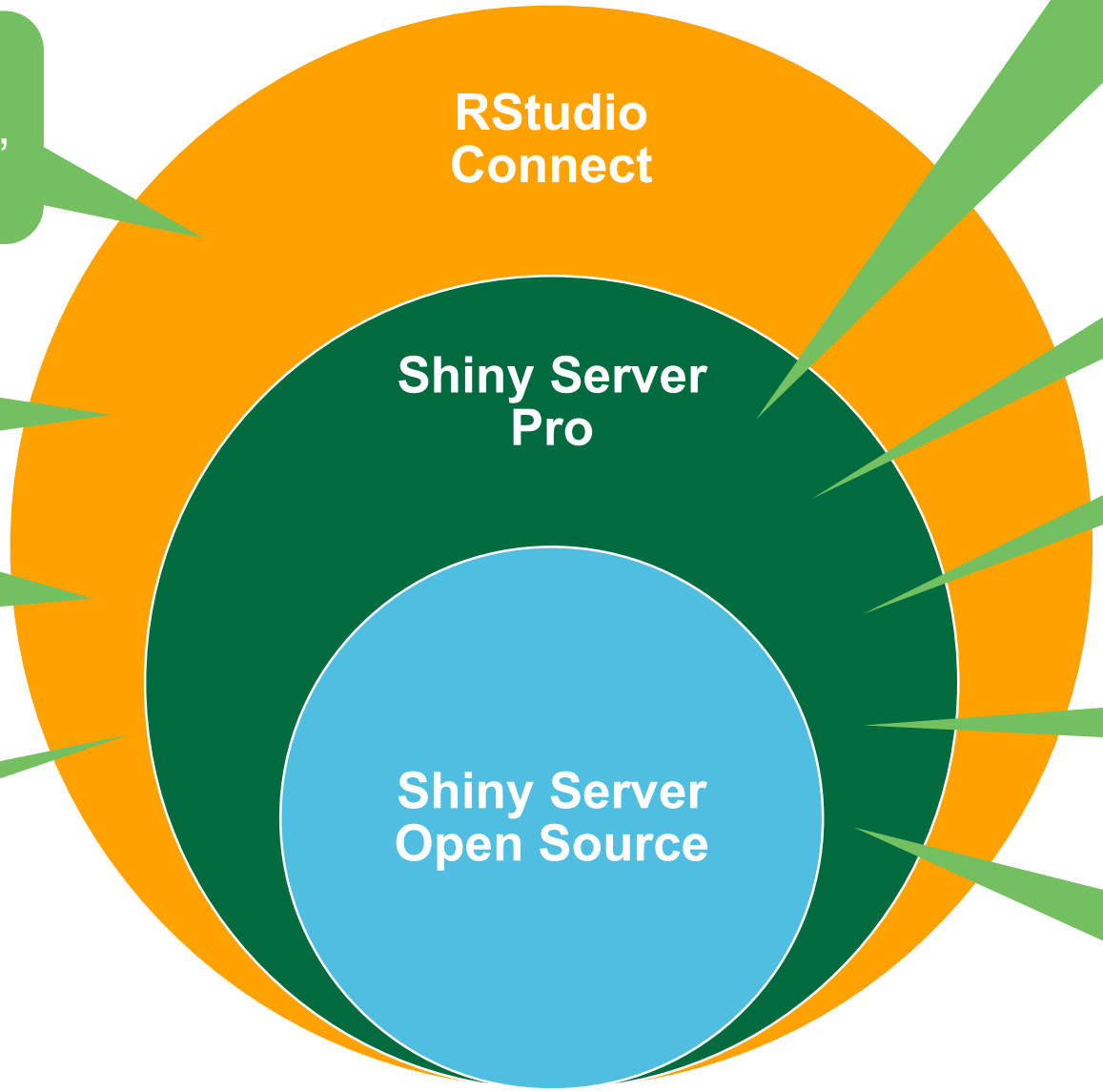
Comparison

Authentication/Security/Scaling/GUI

	Shiny Server Open Source	Shiny Server Professional	RStudio Connect	shinyapps.io
Password protected (i.e. data privacy)	✗*	✓	✓	✓ - paid tier (2) ✗ - free tier
Secure website access via HTTPS (i.e. data security)	✗*	✓	✓	✓
Scale applications across multiple R processes	✗	✓	✓	✓ - paid tier (3) ✗ - free tier
Push-button publishing and GUI interface for configuration	✗	✗	✓	✓

* Can be accomplished using additional open source software

RStudio product hierarchy



- Publish R Markdown docs, Jupyter notebooks, Plumber API's
- User self-management
- GUI for configuration and administration
- Push-button publishing
- Professional drivers to connect Shiny to data
- Administrative dashboard
- Application scaling
- Authentication (data privacy)
- Encrypted access via HTTPS (data security)

Shiny Server Pro (and open source) description

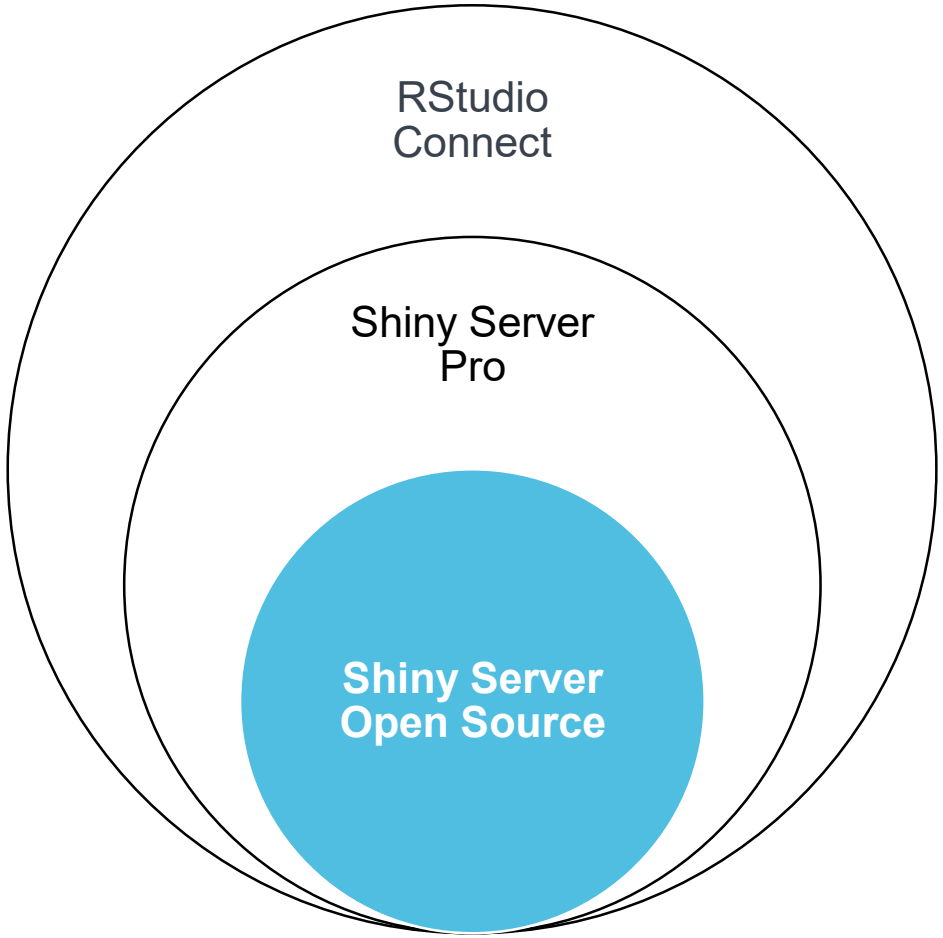
"Shiny Server Pro (and open source Shiny Server) is designed to only host shiny applications. The lack of push button publishing and a user interface typically requires IT administration for publishing and maintaining apps."

What is the difference between RStudio Connect, Shiny Server Pro, and Shinyapps.io?
https://support.rstudio.com/hc/en-us/articles/217240558-What-is-the-difference-between-shinyapps-io-and-Shiny-Server-Pro-

Shiny Server Open Source

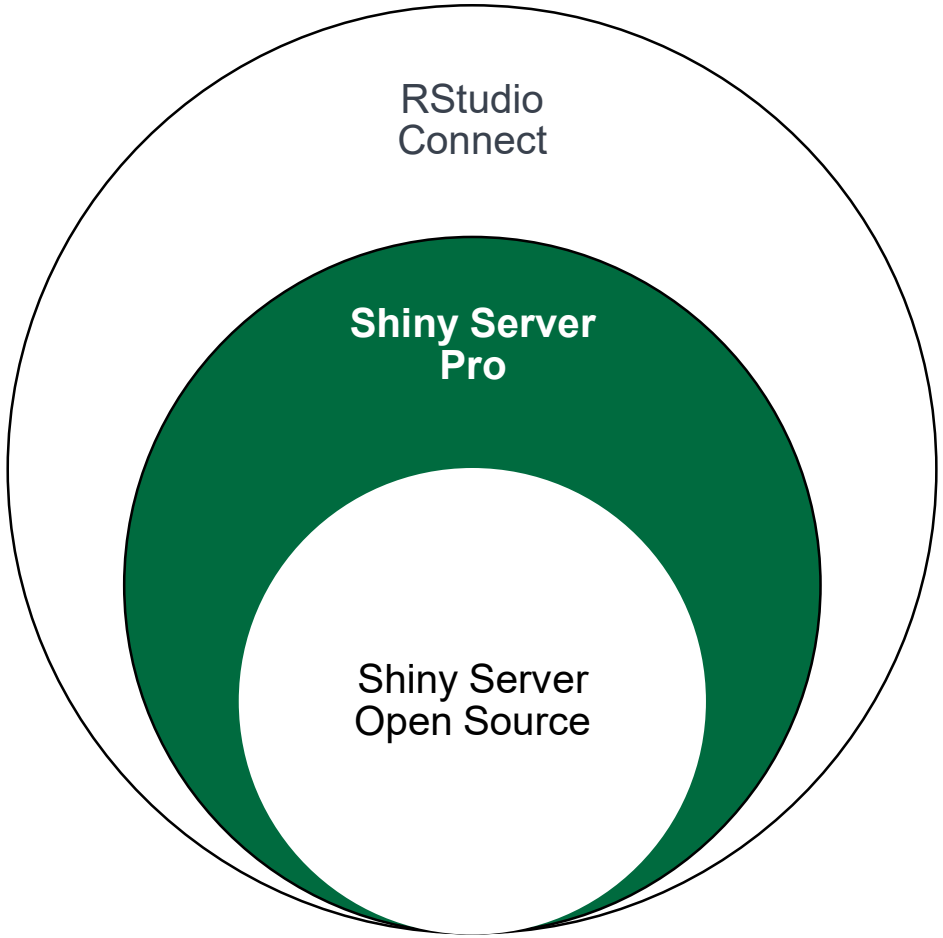
- Server with an HTTP API (web server)
- R execution engine
- Start & stop Shiny processes as needed

- Super popular!!!
  - Probably the most common Shiny hosting platform
  - It's free open source software
- Large volume of community contributed guidance on the web
  - How to add basic data security (HTTPS)
  - How to add basic password protection



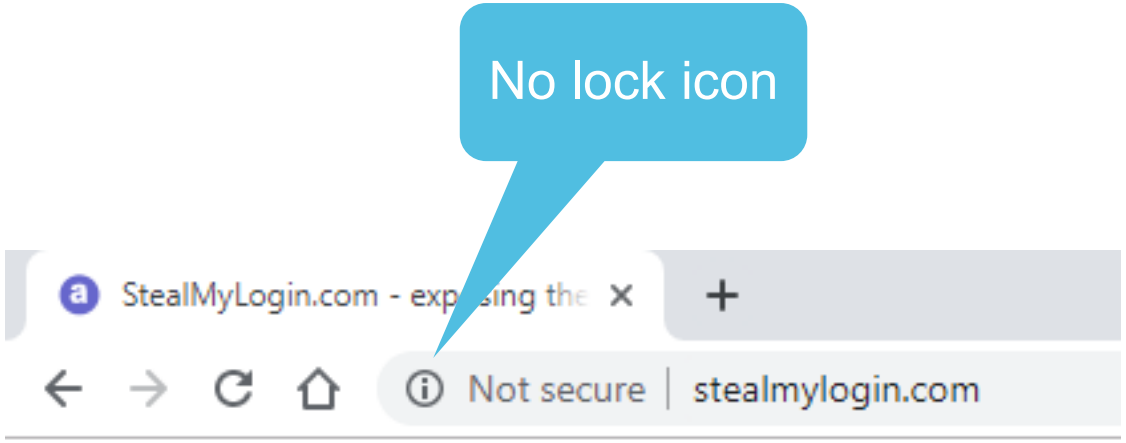
Shiny Server Pro

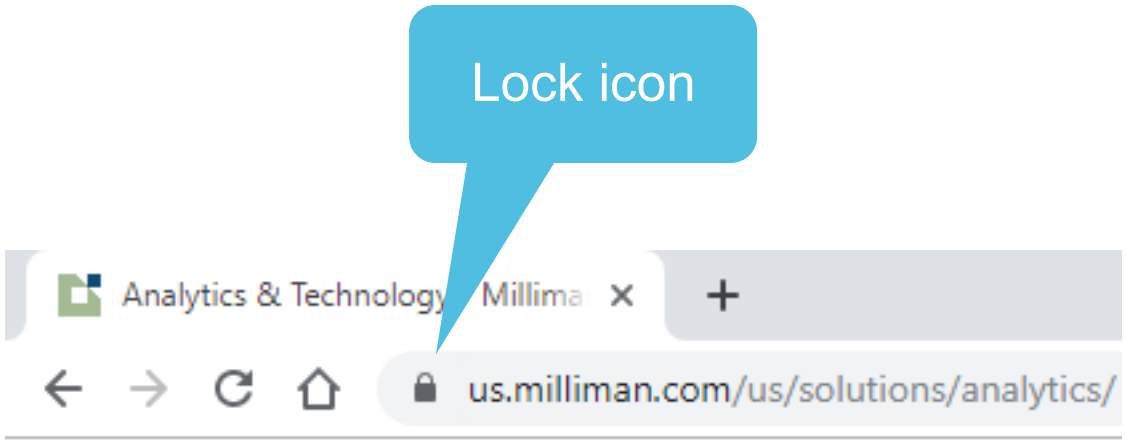
1. Server with an HTTP API (web server)
   - Support for HTTPS (secure web access)
   - Requires an SSL certificate
2. Authentication and identity awareness
3. Application scaling
   - For higher performance and more responsive applications
4. Administrative dashboard
5. Professional drivers (no extra charge)



RStudio's original Shiny hosting product

HTTP versus HTTPS





HTTP (Hyper Text Transfer Protocol)

HTTPS (Hyper Text Transfer Protocol Secure)

HTTPS = HTTP + SSL

SSL (Secure Sockets Layer)

Note: The SSL protocol has evolved into the TLS (Transport Layer Security) protocol but the term 'SSL' and especially the term 'SSL certificate' are still widely used

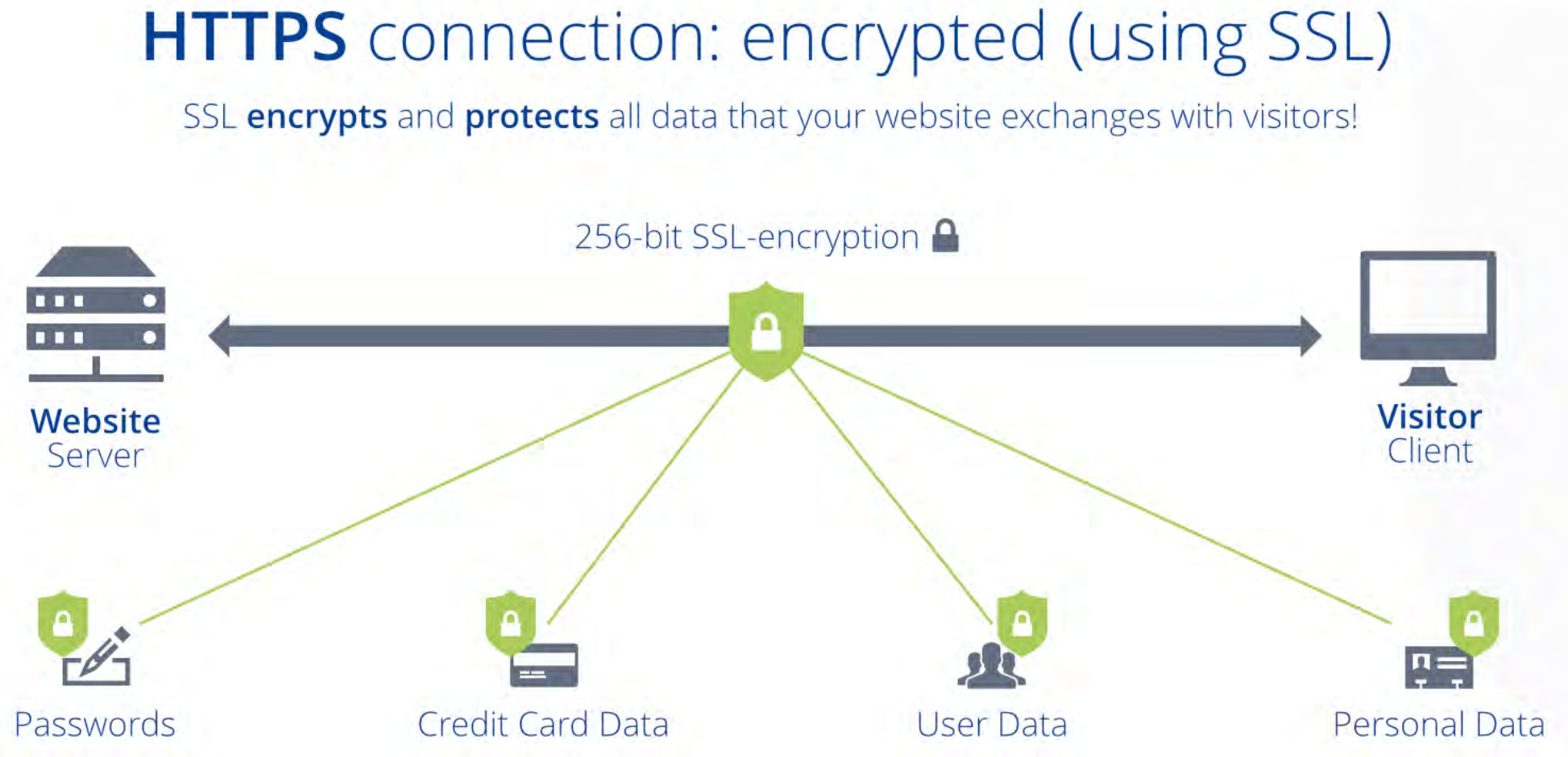
HTTP

https://seopressor.com/blog/http-vs-https/

HTTP

https://seopressor.com/blog/http-vs-https/

HTTPS



https://seopressor.com/blog/http-vs-https/

HTTPS

https://seopressor.com/blog/http-vs-https/

SSL certificate

- Digital certificate used to verify the identity of the website
- The certificate also contains the website's public encryption key used to encrypt website traffic to the website
- The SSL certificate must be obtained from a Certificate Authority (CA) and installed on the web server

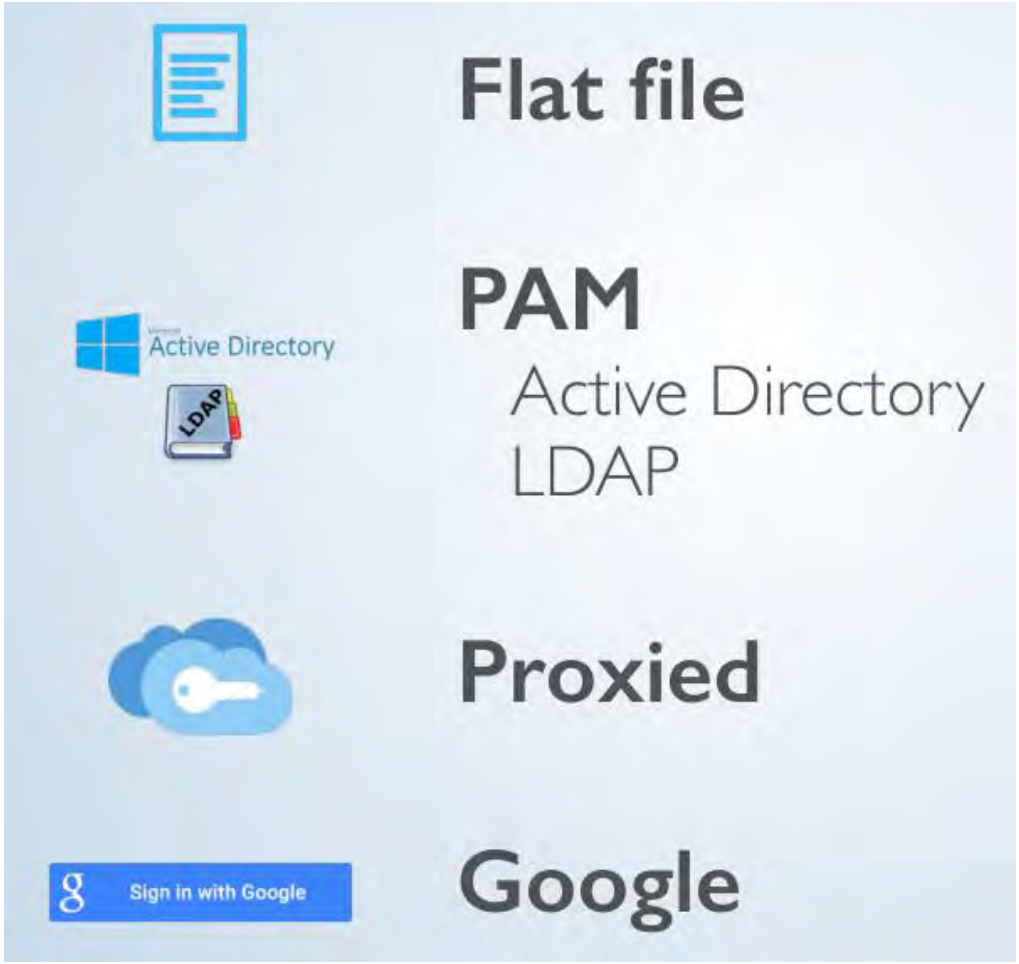
Section of configuration file telling Shiny Server Pro to use HTTPS and where the SSL certificate and public key is located

```
server {
  # Instruct this server to listen on port 443, the default port for HTTPS
  # traffic
  listen 443;
  ssl /etc/shiny-server/ssl-key.pem /etc/shiny-server/ssl.cert;

  ...
}
```

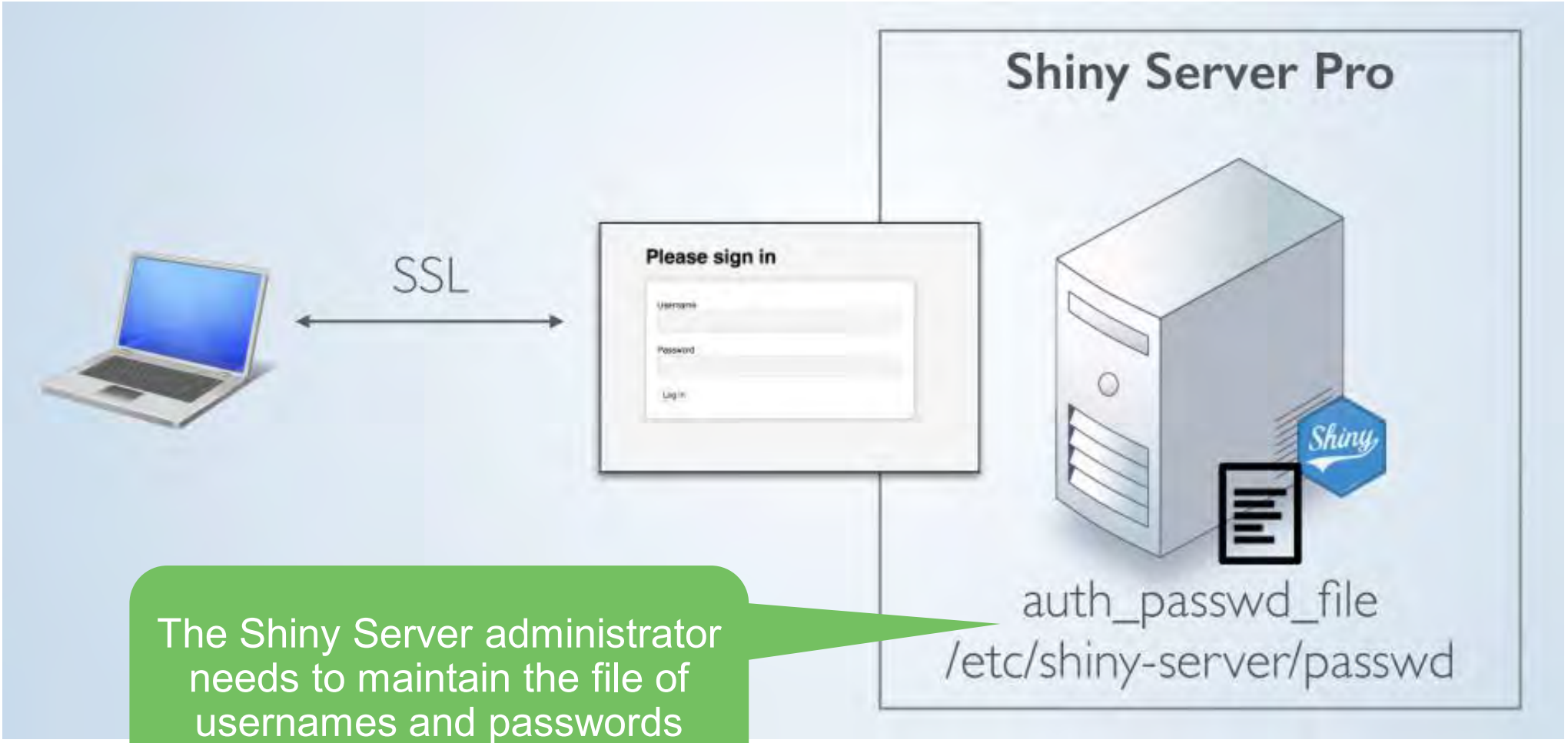
Obtaining and installing an SSL certificate is a common IT process but it does require basic web site system administration skills

Shiny Serve Pro authentication capabilities



Graphic found on rstudio.com

Flat-file authentication (built-in username/password system)



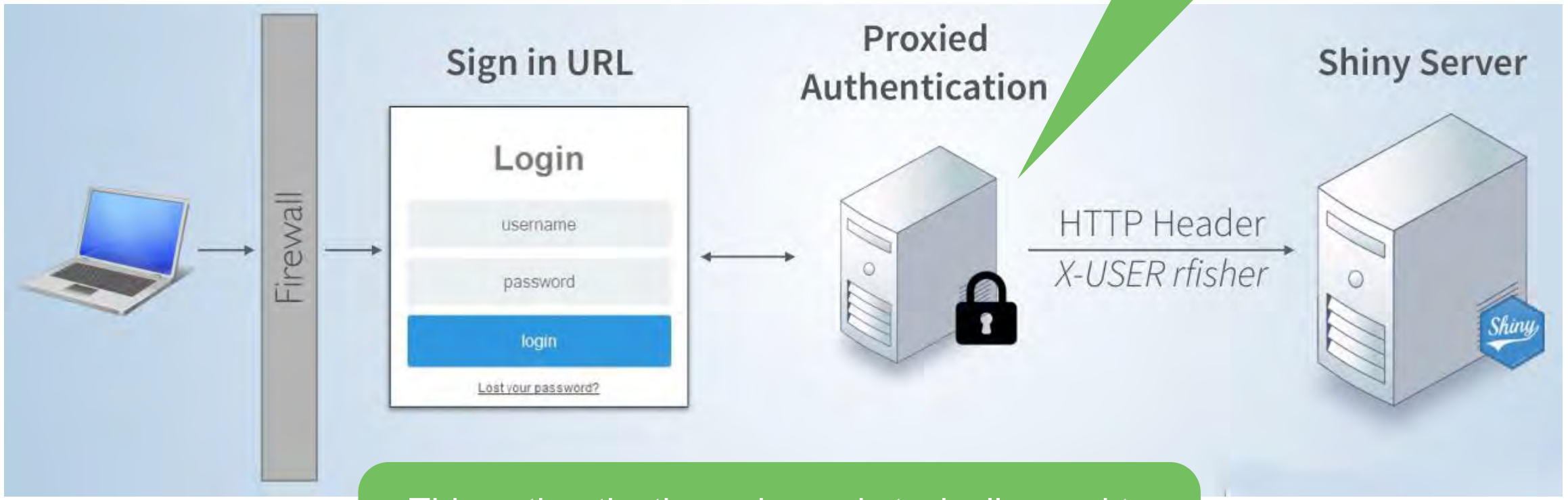
Graphic found on rstudio.com

Google authentication

The Shiny Server administrator needs to register the application with Google

Graphic found on rstudio.com

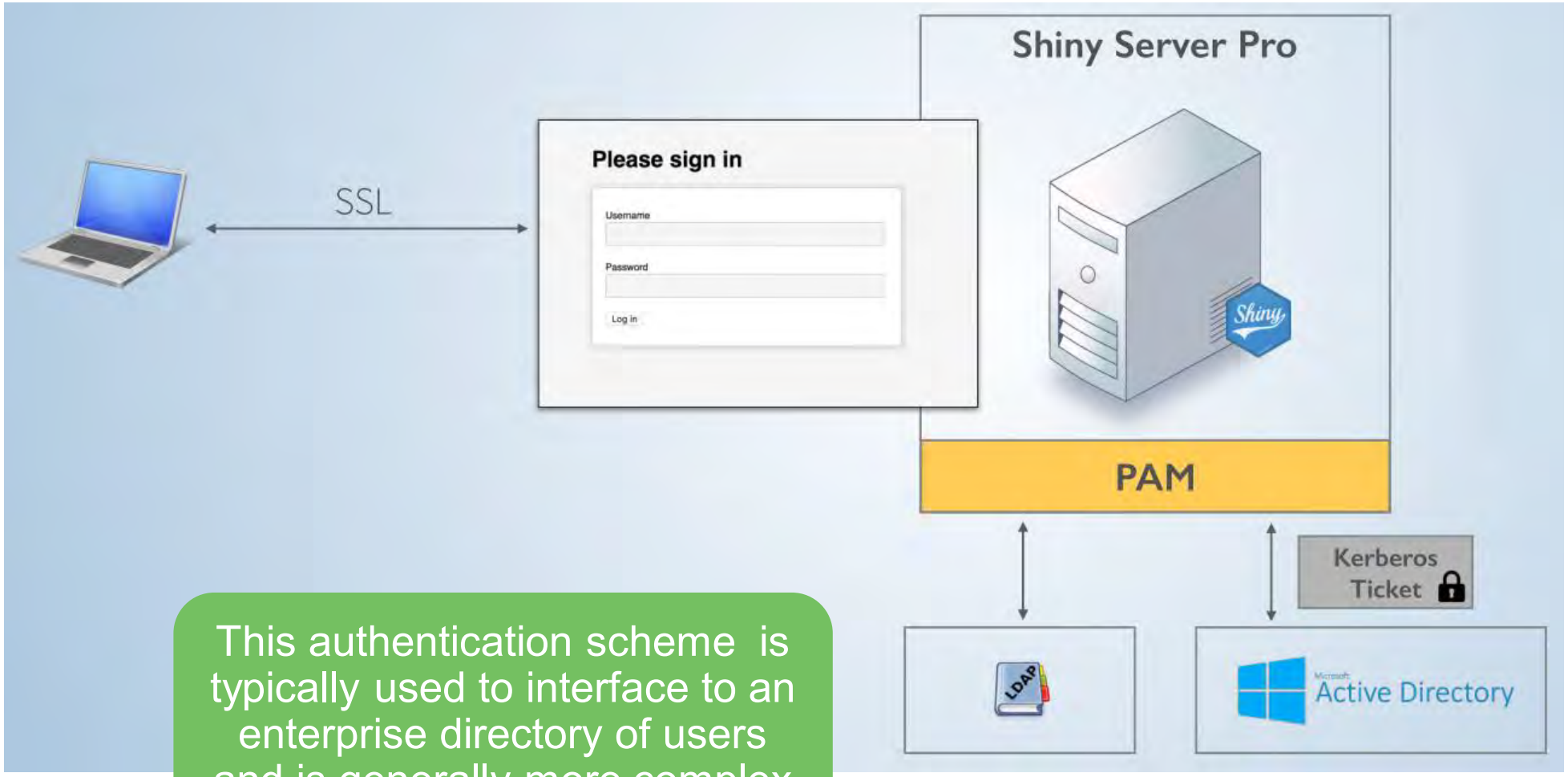
Proxied authentication



This authentication scheme is typically used to interface to a 3rd party authentication service (e.g. Auth0, Okta, etc.) and is generally more complex

Graphic found on rstudio.com

PAM+AD or PAM+LDAP authentication



This authentication scheme is typically used to interface to an enterprise directory of users and is generally more complex

Graphic found on rstudio.com

User identity accessible from the Shiny application

- RStudio professional products pass the username and user group to the Shiny application
- Facilitates application personalization
  - The application 'knows' who the user is and content can be dynamically modified for that user

Shiny app code has access the user's username and group

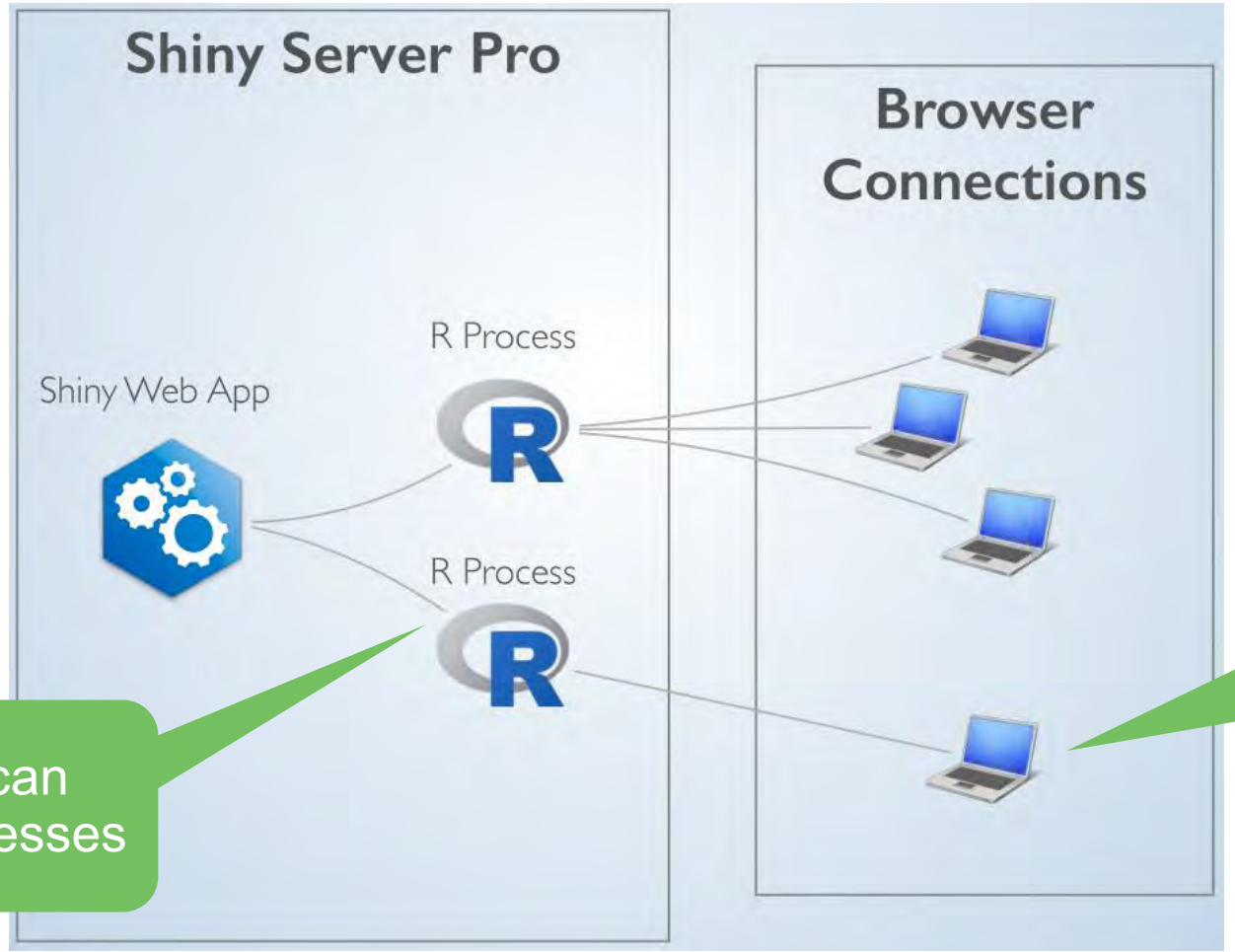
Graphic found on rstudio.com

```
shinyServer(function(input, output, session) {
  output$username <- reactive({
    session$user
  })

  output$groups <- reactive({
    session$groups
  })
})
```

Note: While it is relatively straight forward to add basic password protection to Shiny Server Open Source, it is not straight forward to have the Shiny app access the user identity without one of RStudio's professional products.

Shiny Server Pro application scaling and performance tuning

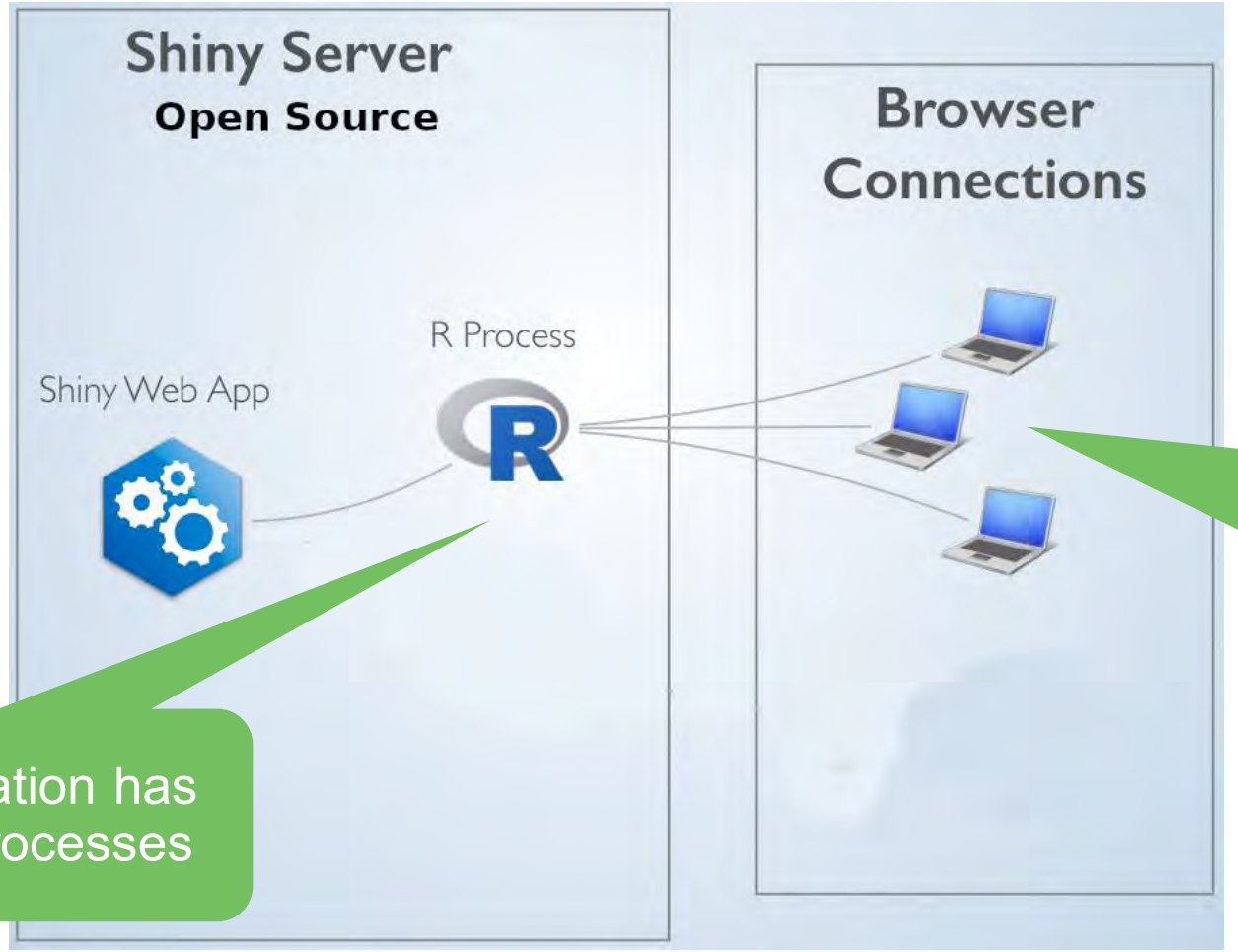


Every application can have multiple R processes

The number of connections can be limited down to one

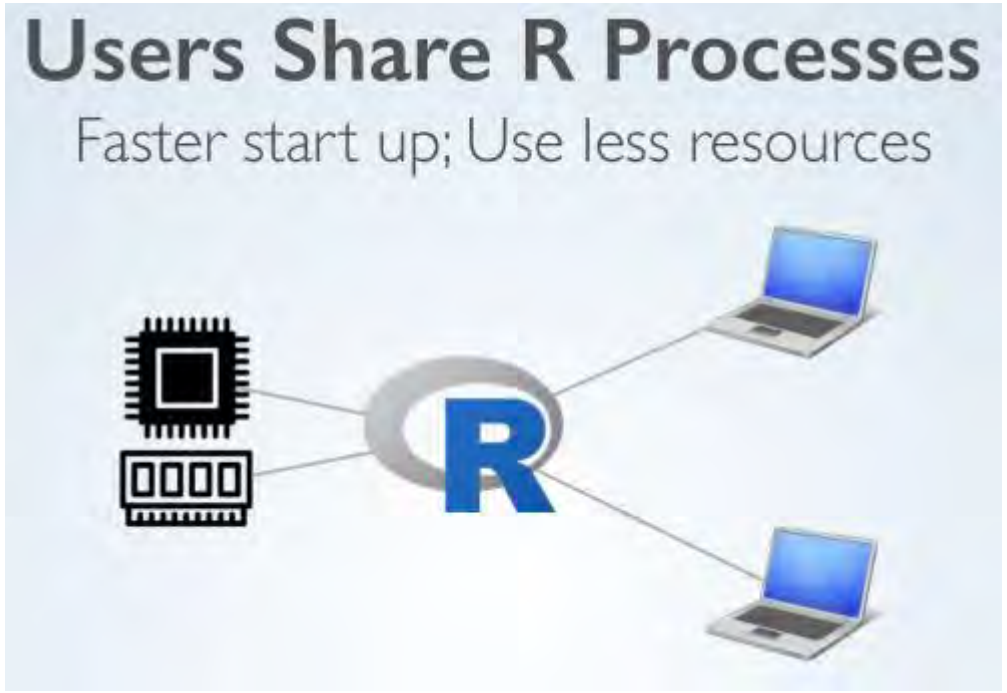
Graphic found on rstudio.com

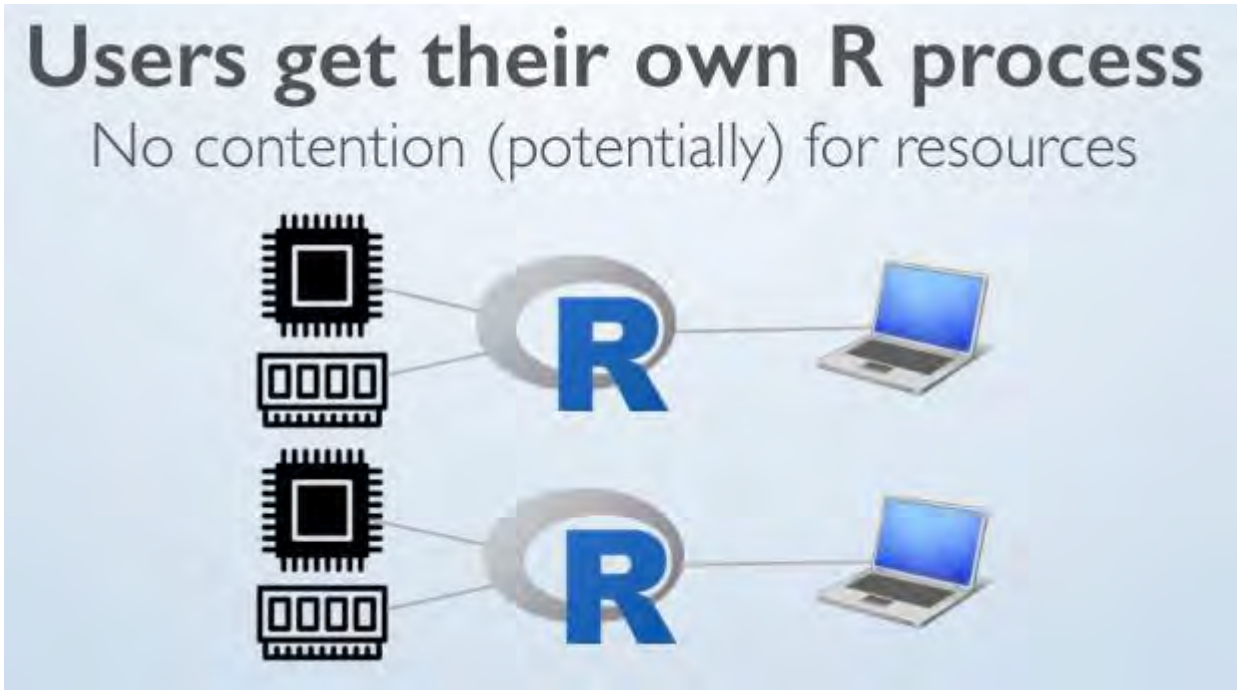
Shiny Server Open Source – no scaling



Original graphic found on rstudio.com

Shiny Server Pro application scaling and performance tuning

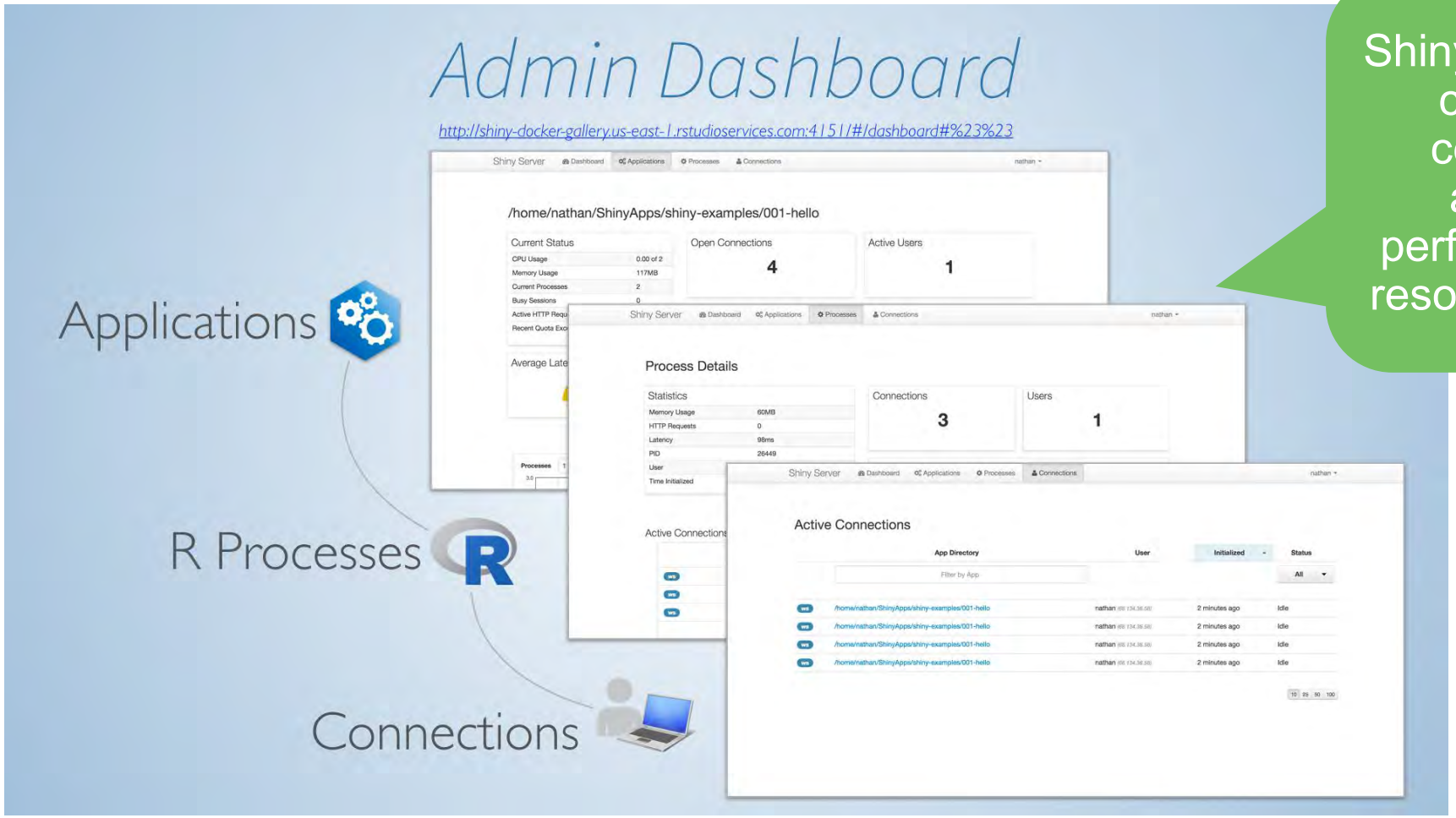




Graphics found on rstudio.com

Shiny Server Pro allows tuning to balance app performance and resource utilization

Shiny Server Pro administrative dashboard



Shiny administrator can monitor connections, application performance, and resource utilization

Graphic found on rstudio.com

Professional drivers to connect Shiny to data

Available data sources

Connect to some of the most popular databases available today. We support several relational databases, cloud data warehouses, Hadoop, and NoSQL data sources.

- Microsoft SQL Server
- Oracle
- Teradata
- PostgreSQL
- Apache Hive
- Apache Impala
- Apache Cassandra
- Amazon Athena
- Amazon Redshift
- MongoDB
- Google BigQuery
- IBM Netezza
- Salesforce
- MySQL

Graphic found on rstudio.com

Professional drivers are available to RStudio professional product customers at no additional charge

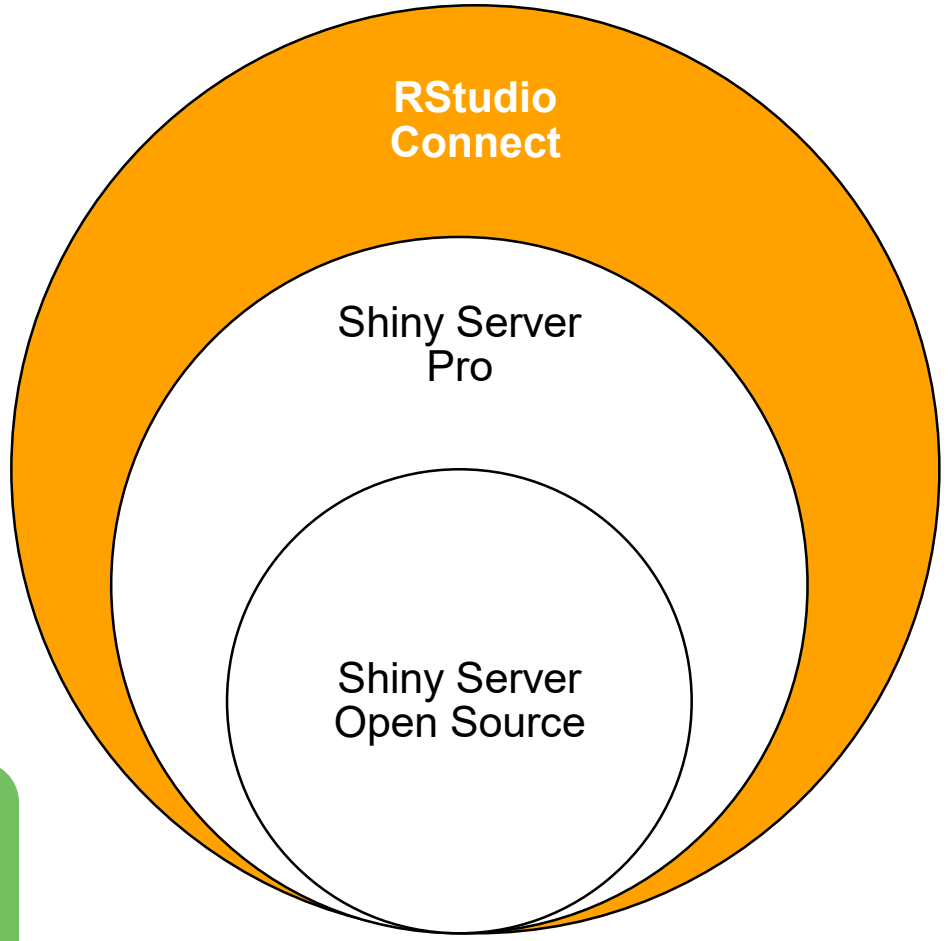
RStudio Connect product description

"RStudio Connect is a new publishing platform for all the work your teams create in R. Share Shiny applications, R Markdown reports, dashboards, plots, APIs, and more in one convenient place. Use push-button publishing from the RStudio IDE, scheduled execution of reports, and flexible security policies to bring the power of data science to your entire enterprise."

What is the difference between RStudio Connect, Shiny Server Pro, and Shinyapps.io?
https://support.rstudio.com/hc/en-us/articles/217240558-What-is-the-difference-between-shinyapps-io-and-Shiny-Server-Pro-

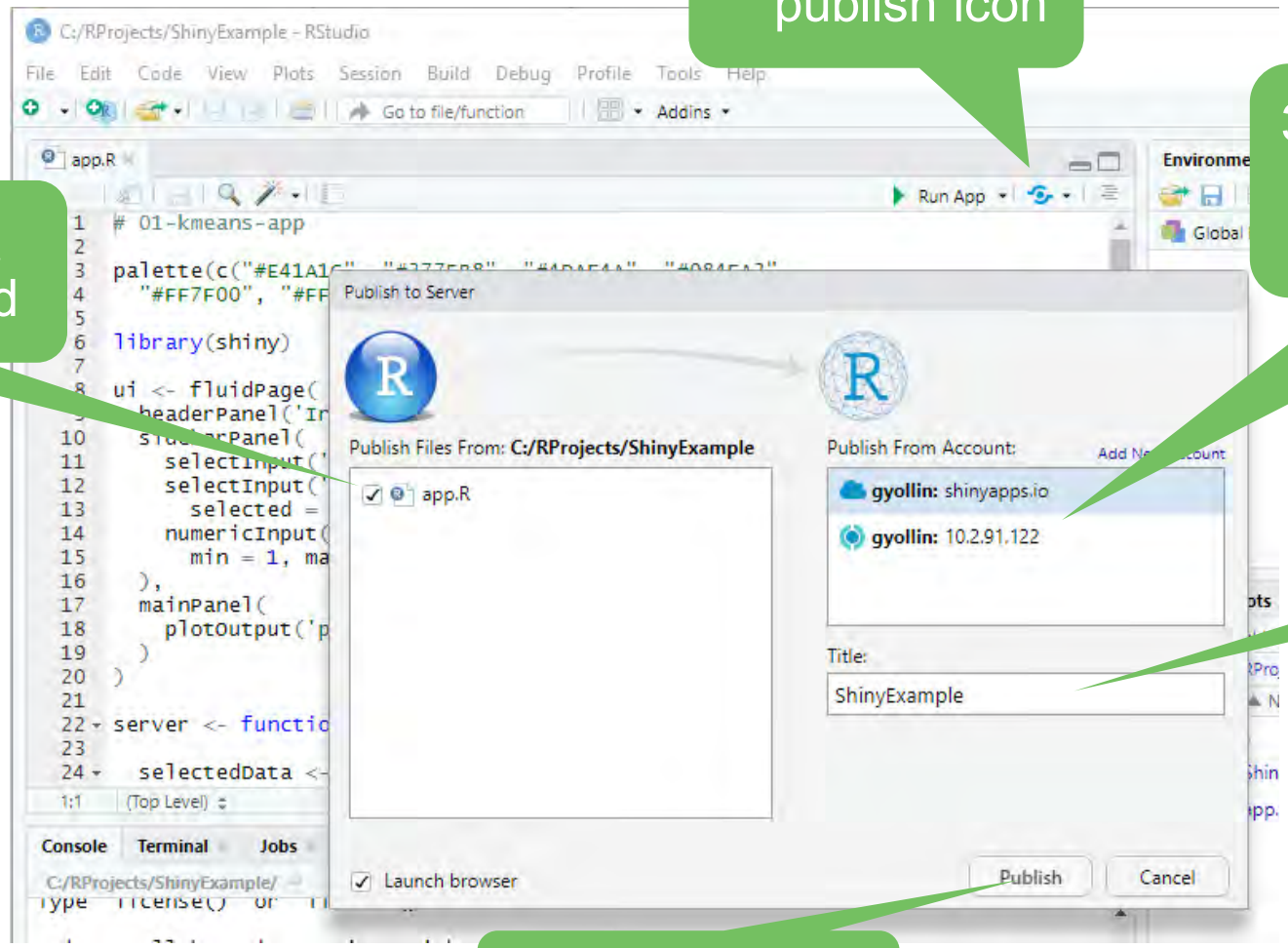
RStudio Connect

1. Push-button publishing from the RStudio IDE
2. GUI for some configuration/administration
   - Once the product is installed on a Linux server
3. User self-management
4. Publish R Markdown documents, Jupyter notebooks and Plumber APIs

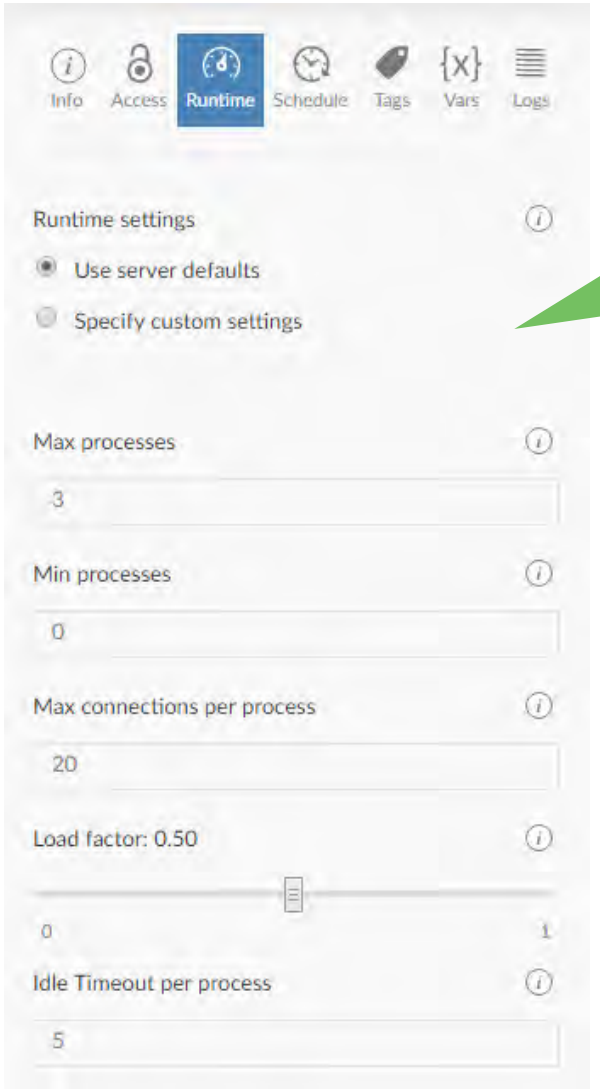


Note: The primary use-case for RStudio Connect is for **internal collaboration within an enterprise**, not hosting Shiny apps for external use on the public internet

Push-button publishing from RStudio IDE

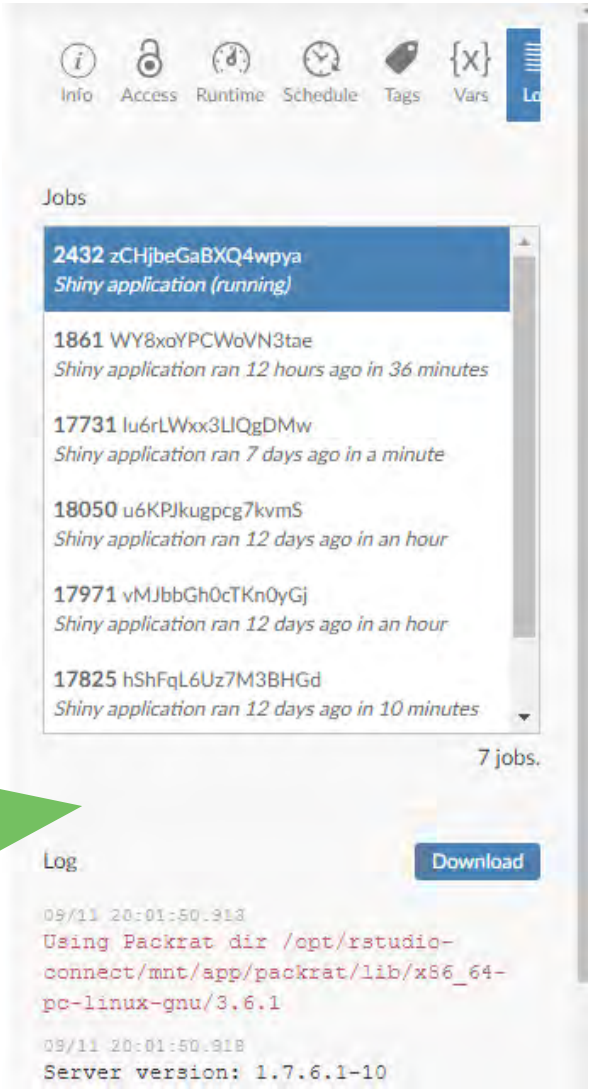


GUI for some configuration/administration tasks

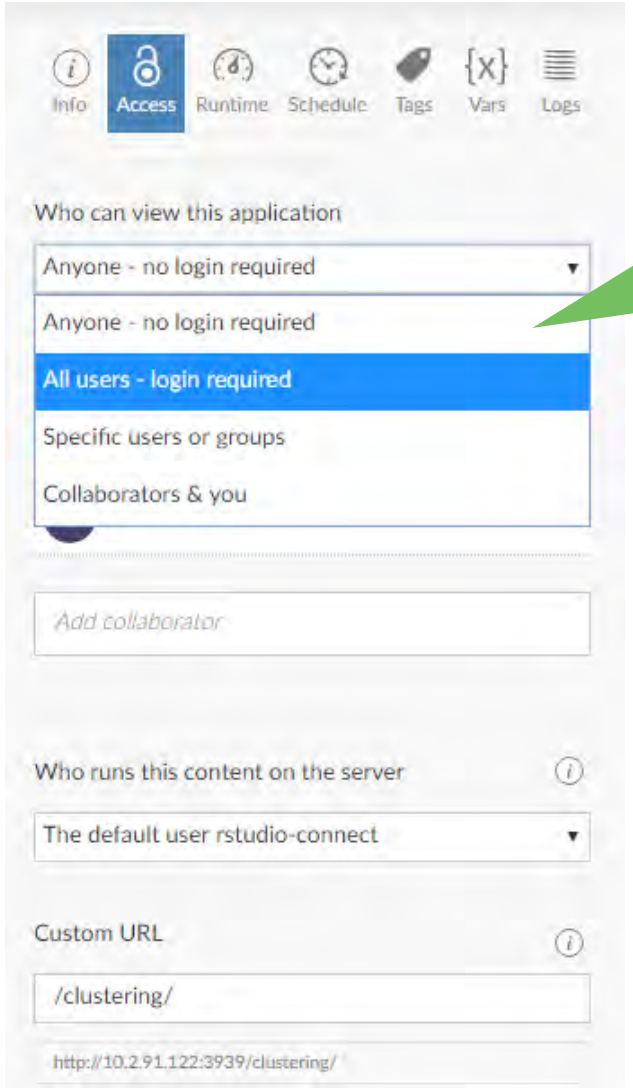


Adjust app performance tuning parameters

Review and download logs

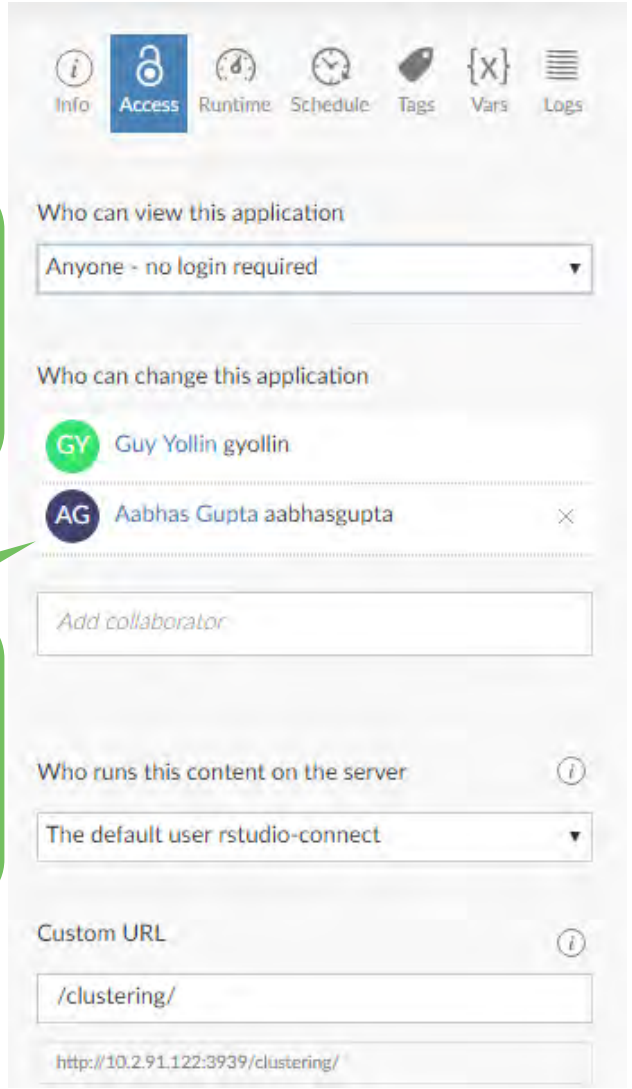


User self-management by publishers



Choose who can access the app

Add collaborators who can change the app



Additional RStudio Connect hosting capabilities

Product	Installed or Hosted	Push Button Publishing	Shiny Apps	R Markdown Documents	Jupyter Notebooks	Plumber API's
RStudio Connect	Installed	●	●	● *(1)*	● *(1)*	●
Shiny Server Pro	Installed		●	●		
Shinyapps.io	Hosted	●	●	● *(2)*		

Graphic found on rstudio.com

(1) Supports advanced features for refreshing, scheduling, and distributing documents

(2) Only when using runtime: shiny in the YAML header

https://support.rstudio.com/hc/en-us/articles/217240558-What-is-the-difference-between-RStudio-Connect-Shiny-Server-Pro-and-Shinyapps-io-

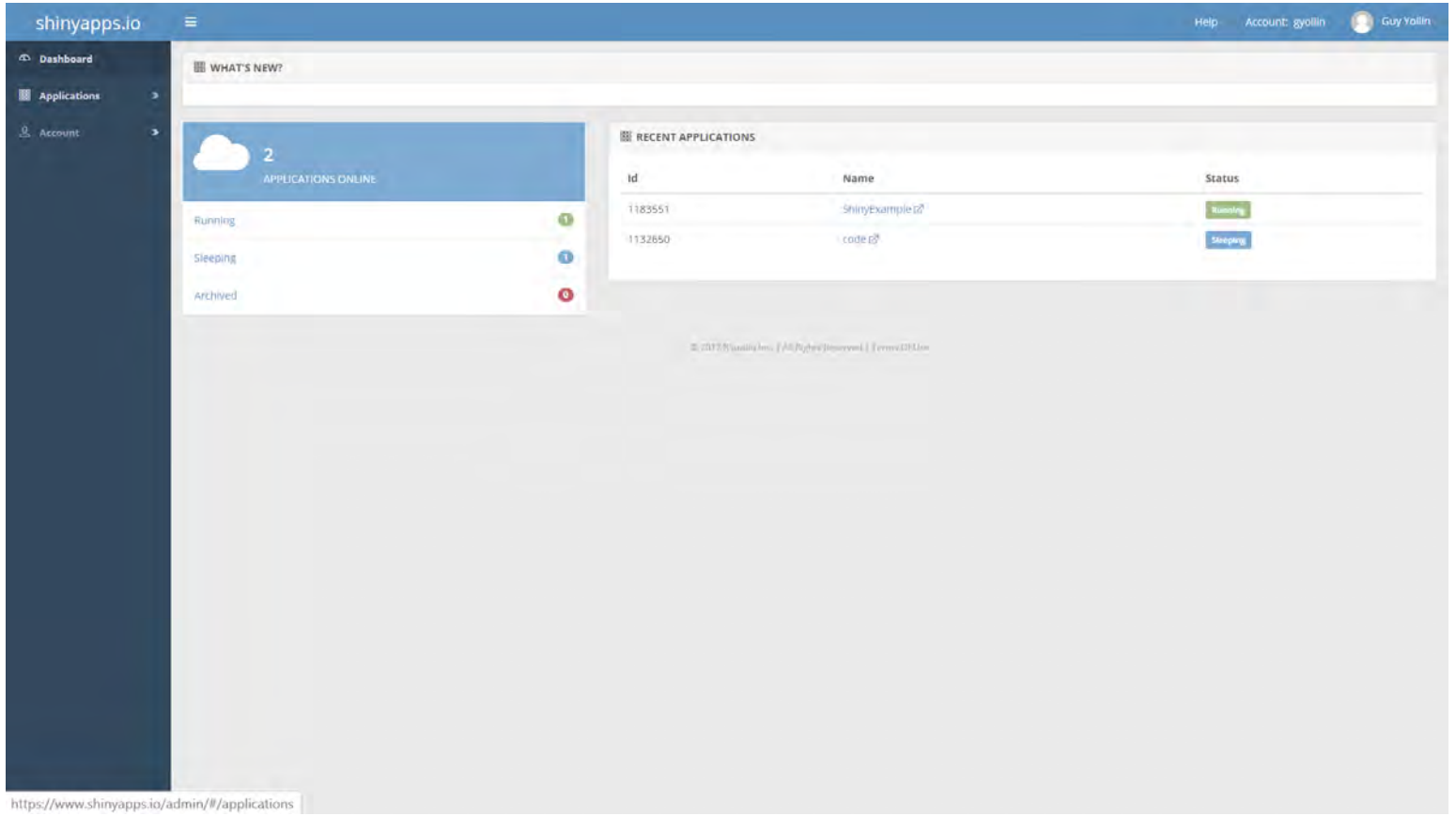
Shinyapps.io product description

"Shinyapps.io is software as a service (SaaS) hosted in the cloud by RStudio. It has both free and paid plans. Anyone can publish their Shiny apps to shinyapps.io with the push of a button. You don't need to own a server or know how to configure a firewall to deploy and manage your applications in the cloud. No hardware, installation, or annual purchase contract required."

What is the difference between RStudio Connect, Shiny Server Pro, and Shinyapps.io?
https://support.rstudio.com/hc/en-us/articles/217240558-What-is-the-difference-between-shinyapps-io-and-Shiny-Server-Pro-

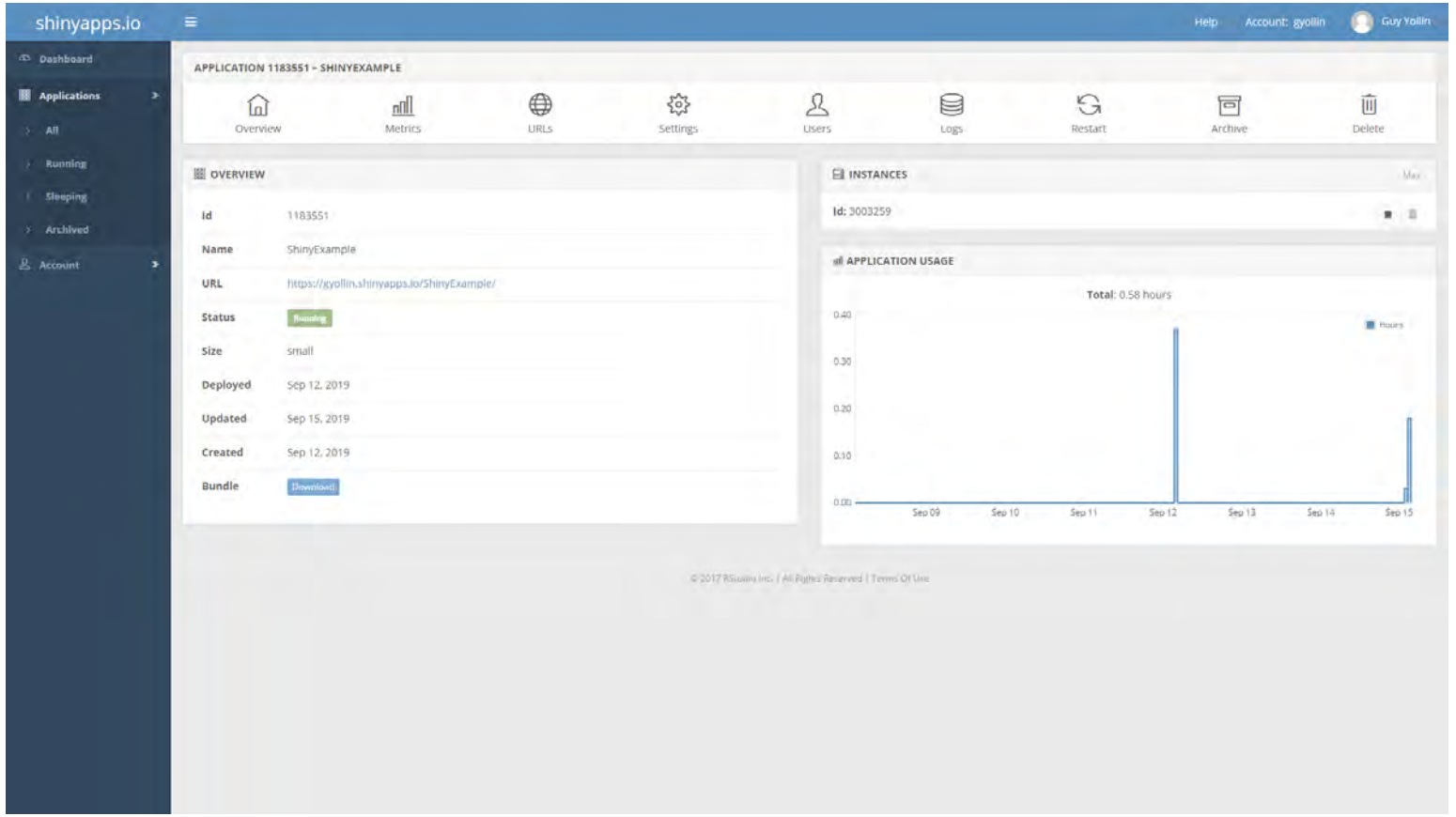
Shinyapps.io

- Shiny app hosting server
  - No software installation required
  - All management and configuration via GUI
- Free and paid tiers include:
  - Secure website access via HTTPS
  - Push-button publishing



Shinyapps.io paid features

- Authentication
  - Standard tier and higher
    - Google
    - Github
    - shinyapps.io
- Application scaling
  - Basic tier and higher
- Custom domain name
  - Professional tier



Shinyapps.io challenges/shortcomings

- Data hosting challenge
  - Persistent storage of data requires accessing a data store external to the shinyapps.io environment; i.e. files created and stored by the app locally are not guaranteed to persist after app resets
- Currently no Service Level Agreements (SLA) regarding performance
- No PII/PHI data security compliance
  - The infrastructure used is not a HIPAA-compliant stack
  - RStudio recommends RStudio Connect or Shiny Server Pro for applications requiring security audits

Note: shinyapps.io is not really suitable for hosting commercial applications

Cost and analysis

Professional products license limitations

- Concurrent User limit
  - Maximum number of users who can be logged into a server at any one time
  - Defined as one human on one browser connecting to one server
  - Limit applies to:
    - Shiny Server Pro
    - RStudio Connect
- Named User limit
  - Maximum number of different uses who can log into a server over a 1 year period of time
  - Limit applies to:
    - RStudio Connect

Professional products license limitations

- Active Hours limit
  - Maximum number of hours that your Shiny application can be non “idle”
  - Limit applies to:
    - shinyapps.io
- Number of applications that can be hosted
  - Limit applies to:
    - shinyapps.io

Shiny Server Pro annual cost

	Shiny Server Professional	20-user License pack	150-user License pack
Price	**$9,995**	**$4,995**	**$14,995**
Concurrent users	**20**	**20**	**150**
Named users	**unlimited**		
Number of applications	**unlimited**		

Consult with RStudio for official pricing information

RStudio Connect pricing

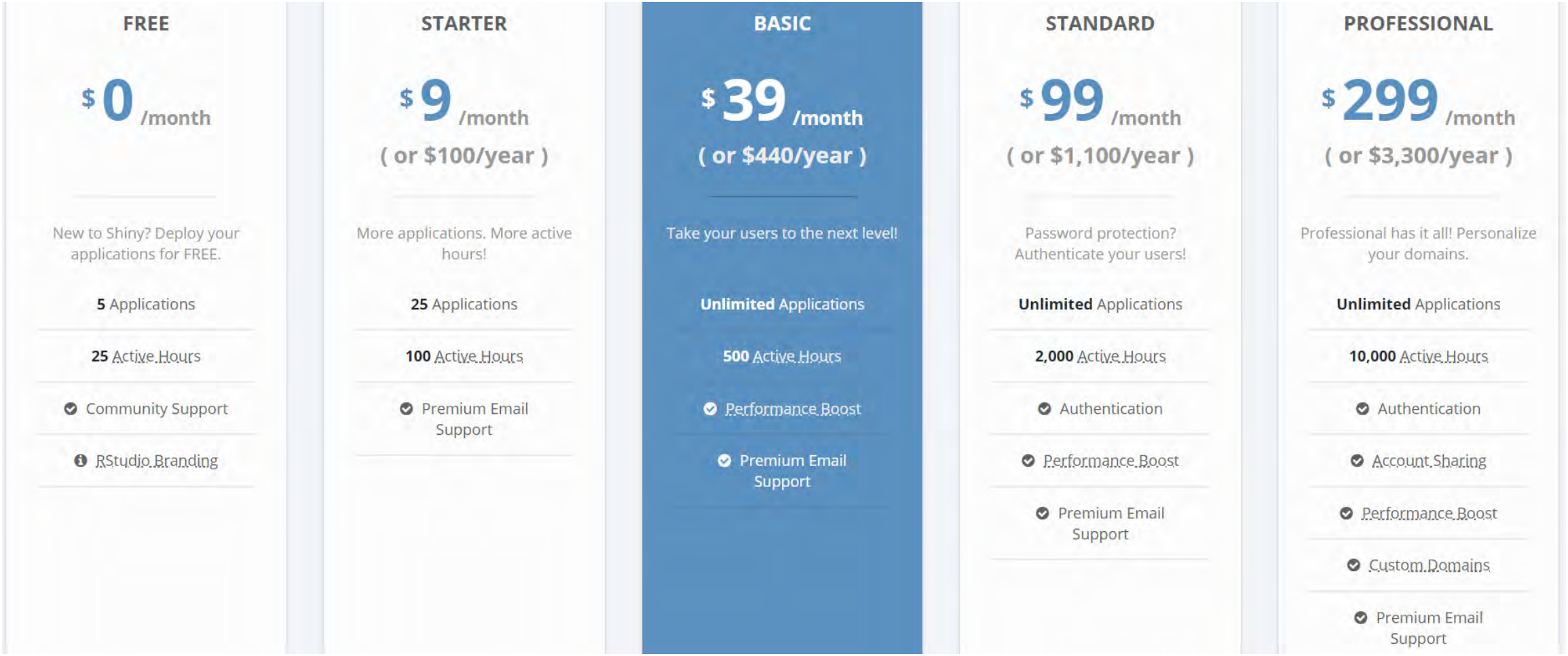
Concurrent User limit is the Named User limit

	RStudio Connect Base	RStudio Connect Standard	RStudio Connect Enterprise*
	$14,995 per year	$24,995 per year	$74,995 per year
Named Users Named Users are publishers of content or have authenticated access to published content	20	100	1,000
Upgrade Path	Upgrade to Standard	Add Named User Packs of 50 for $4,995 and 250 for $14,995	Add Named User Packs of 250 for $14,995
Scaling Run across a multi-node cluster, providing an option to scale compute horizontally and ensure high availability	N/A	Add an Execution Server for $9,995	Please contact sales

Graphic found on rstudio.com

Consult with RStudio for official pricing information

Shinyapps.io pricing



Graphic found on rstudio.com

Consult with RStudio for official pricing information

RStudio Connect rubric

From RStudio:

Consider RStudio Connect if you can answer yes to these questions:

1. Do you want push button publishing?
2. Do you want to publish R Markdown documents, Plumber API's, and Jupyter Notebooks in addition to Shiny applications?
3. Do you want a user interface so that content creators can manage their own data products?

I would add:

- Is the deployment for internal collaboration within your enterprise?
- Do you have a budget of $15K+ per year to support this internal collaboration?

Shiny Sever Pro rubric

From RStudio:

Consider Shiny Server Pro if you can answer yes to all of these questions:

1. Do you want to deploy apps without push button publishing?
2. Do you want to manage apps with configuration files rather than a UI?
3. Are you only interested in sharing Shiny and apps (and not other data products)?
4. Do you want concurrent user licensing?

I would add:

- Does your software release process require testing and formal approval prior to release?
- Do you want an unlimited number of potential users and only be restricted by concurrent users?
- Do you have a budget of $10K+ per year to support Shiny app hosting?

Shinyapps.io rubric

From RStudio:

Use Shinyapps.io if you can answer yes to all of these questions:

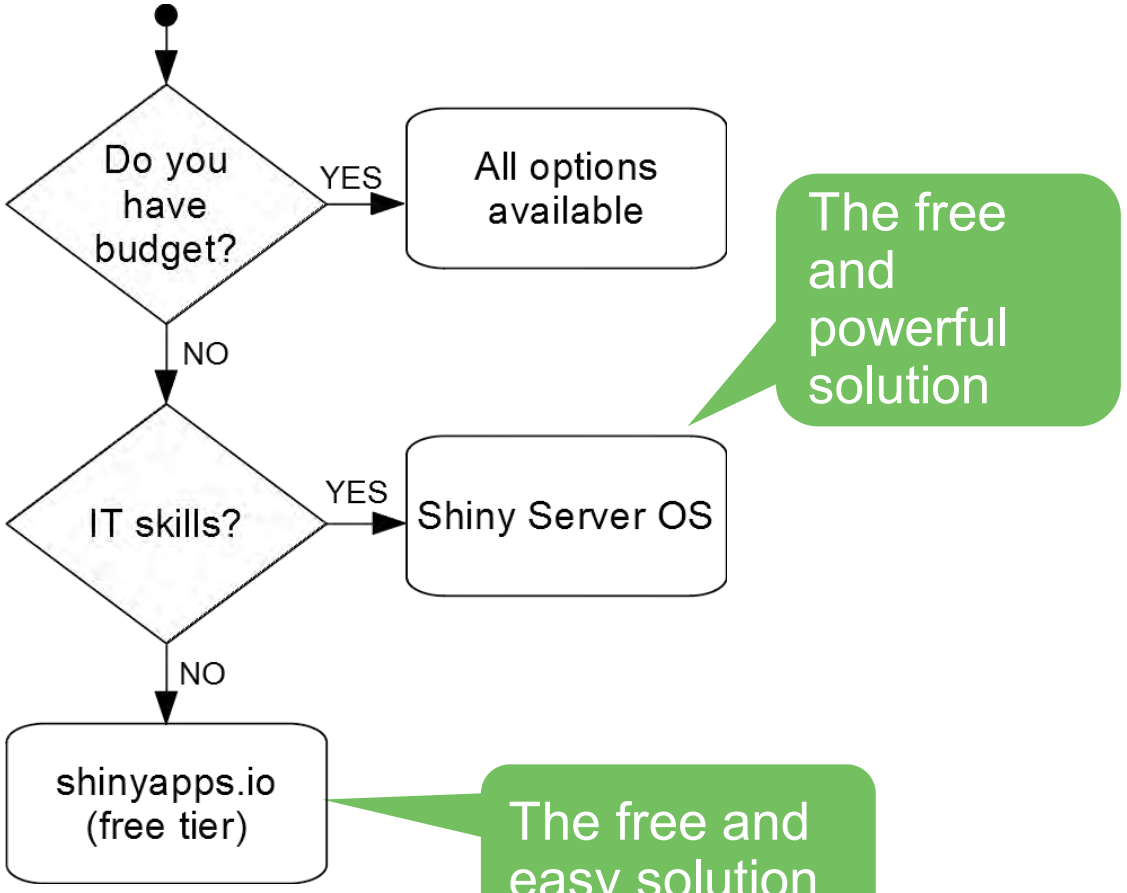
1. Are you okay with your application being outside your firewall?
2. Are you okay with the data that the application is pulling from being accessible to our cloud? (You have to open up a hole in your firewall if the data is behind the firewall today.)
3. Are you okay with your end client creating a user account on shinyapps.io? (if you are looking to use authentication).
4. Are you okay with a shared computation platform for your analyses? (for example, we don't have any SLAs today on performance)

I would add:

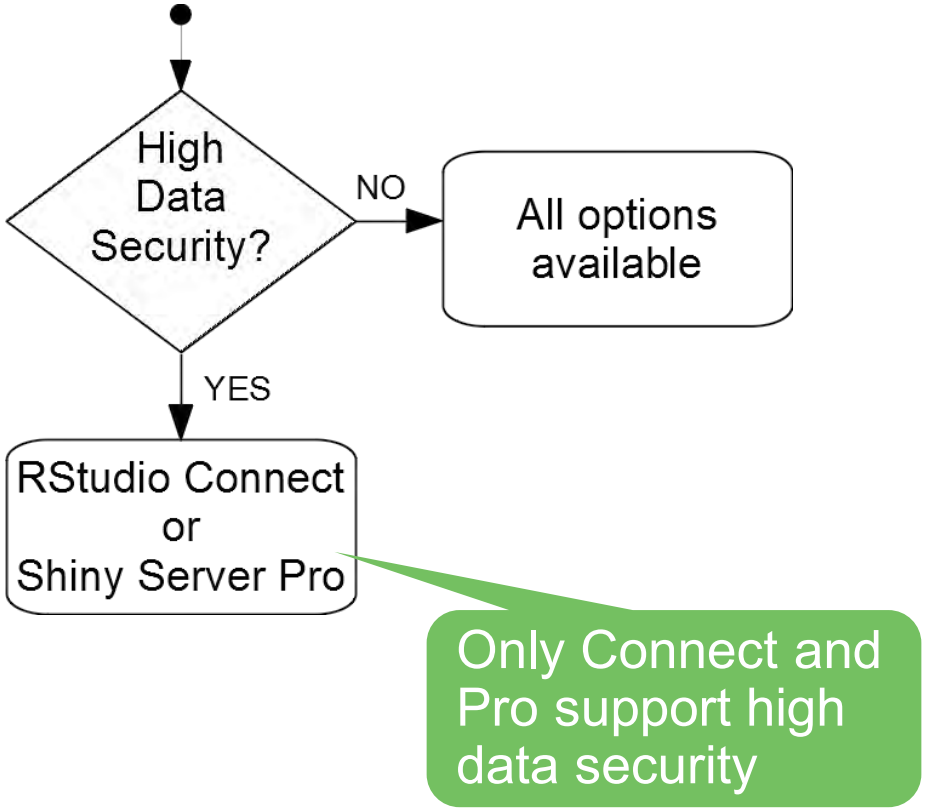
- Do you have budget for the desired features within the shinyapps.io pricing plans?

Simple decisions given constraints

Cost constrained



Security constrained



General guidance

Use-case: Quick and easy sharing - think **shinyapps.io** first

- Confirm there is no rigorous data security required
- No authentication required
  - Shinyapps.io free or basic
- Authentication required
  - Shinyapps.io standard

Use-case: Internal collaboration - think **RStudio Connect** first

- Fully-featured data science collaboration tool

Use-case: commercial app on the public internet - think **Shiny Server Pro** first

- All of the capabilities to support a production quality commercial application

Real deployment examples

Internal & External

Real internal/external authentication

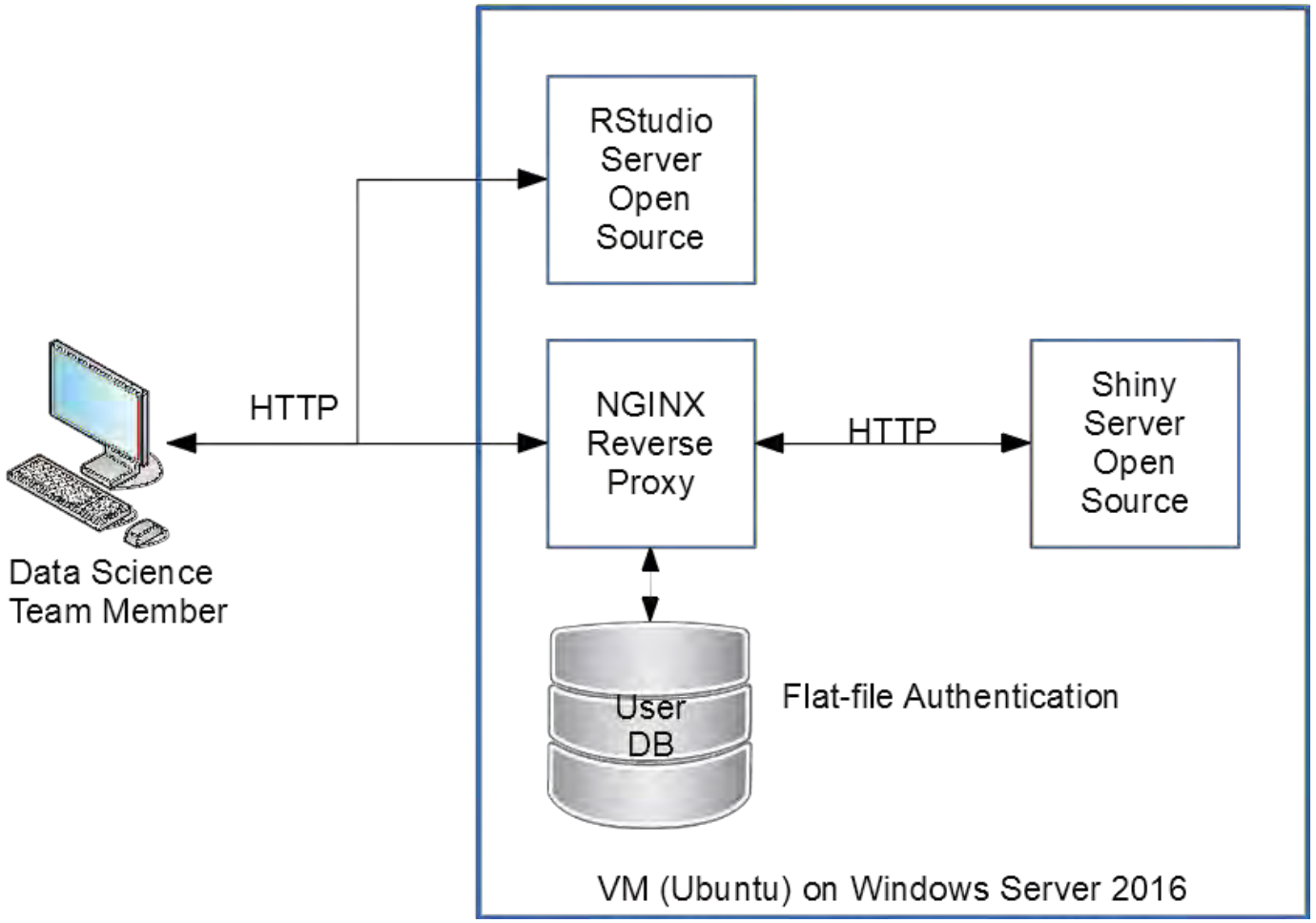
Use-case: Internal collaboration development / testing for data science team

- Requirements
  - Must host multiple applications
  - Minimum data security but restrict access to data science team members
- Solution
  - Shiny Server Open Source deployed on VM internally with nginx authentication

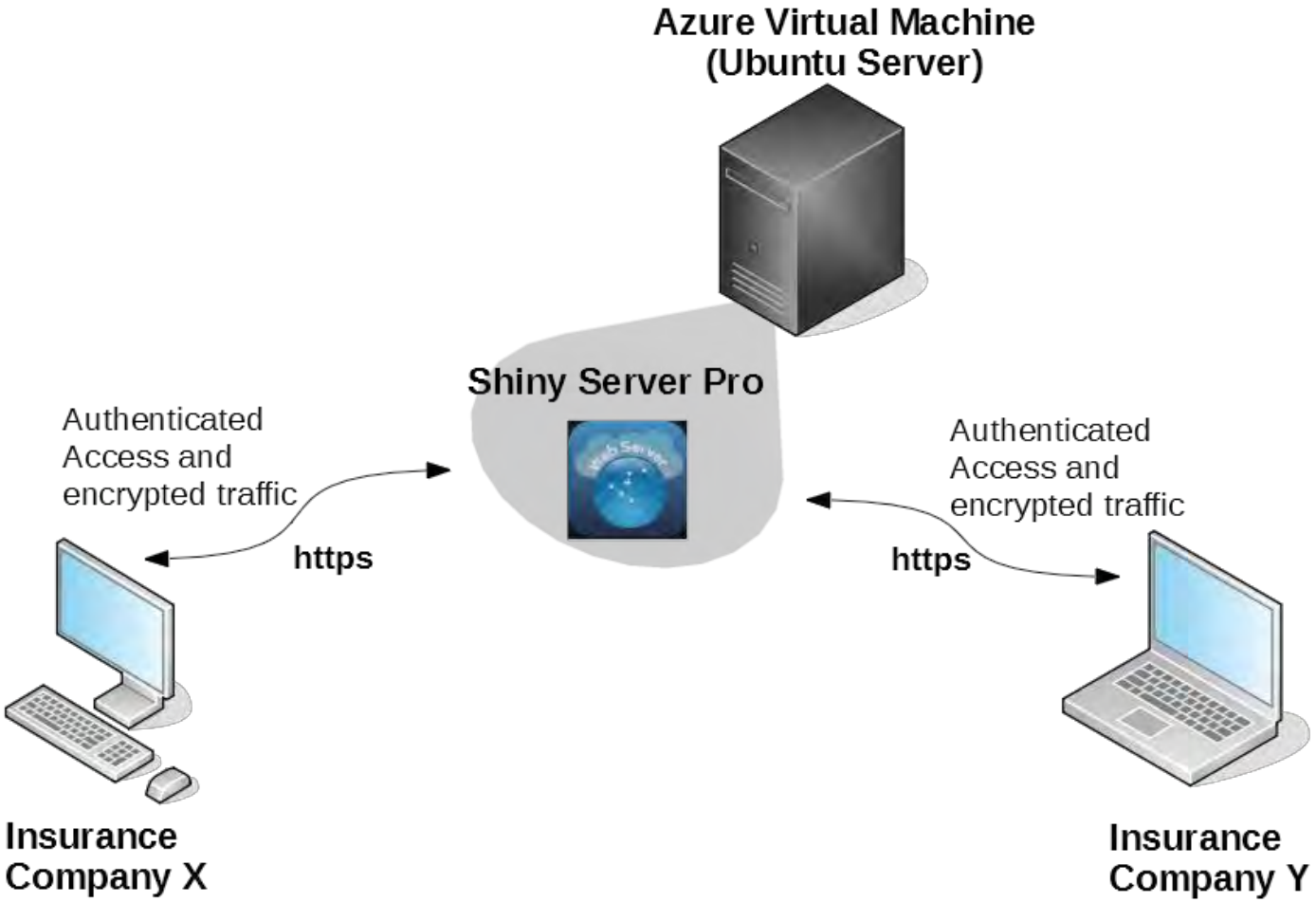
Use-case: Production applications for customer access on public internet

- Requirements
  - Rigorous data security requirement (despite no PII/PHI)
  - Content personalized to client
- Solution
  - Shiny Serve Pro deployed on a VM on Microsoft Azure with Proxied authentication

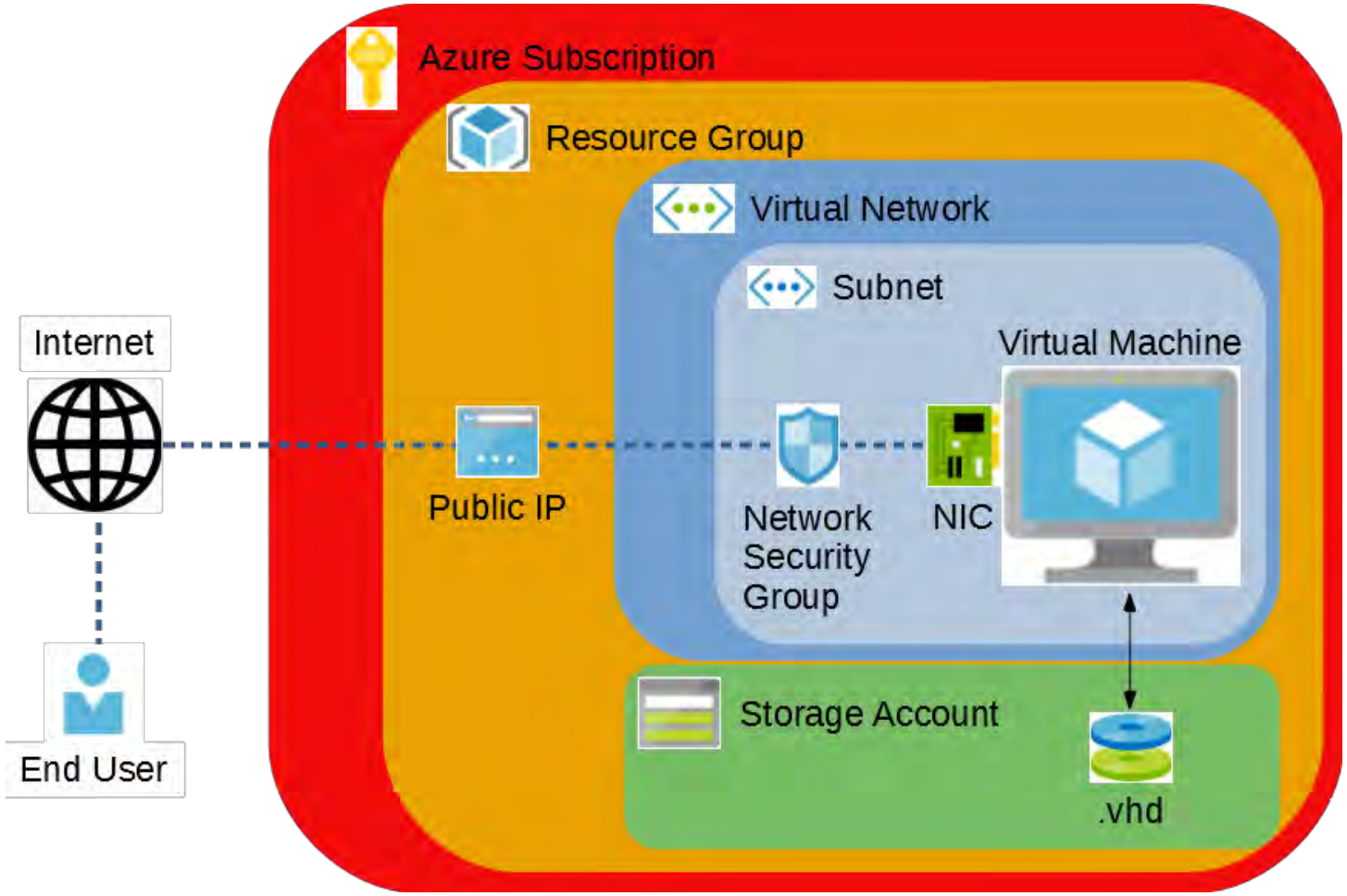
Internal Network – Dev/Test – Minimum Data Security



Public Internet – Commercial Application – High Data Security

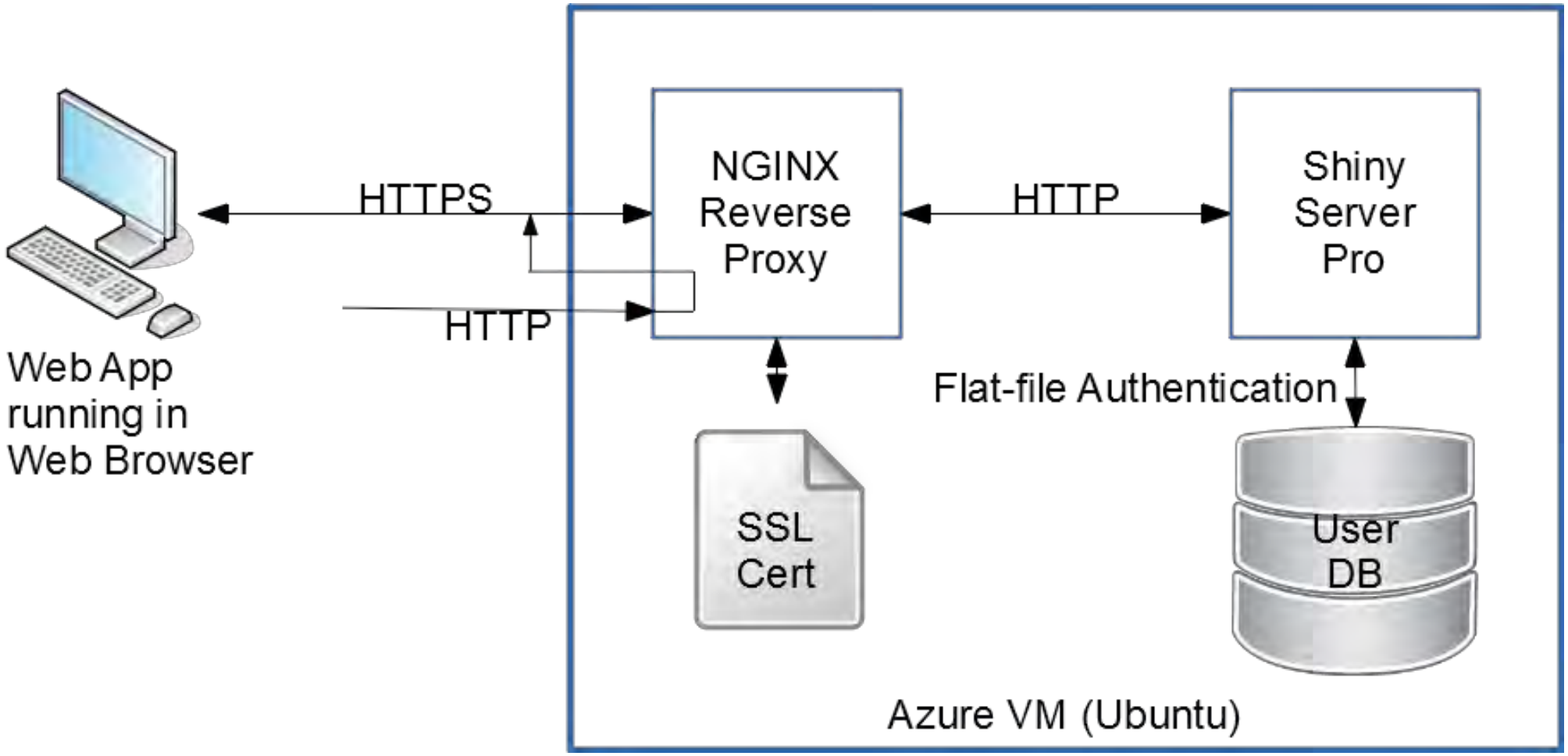


Public Internet – Commercial Application – High Data Security



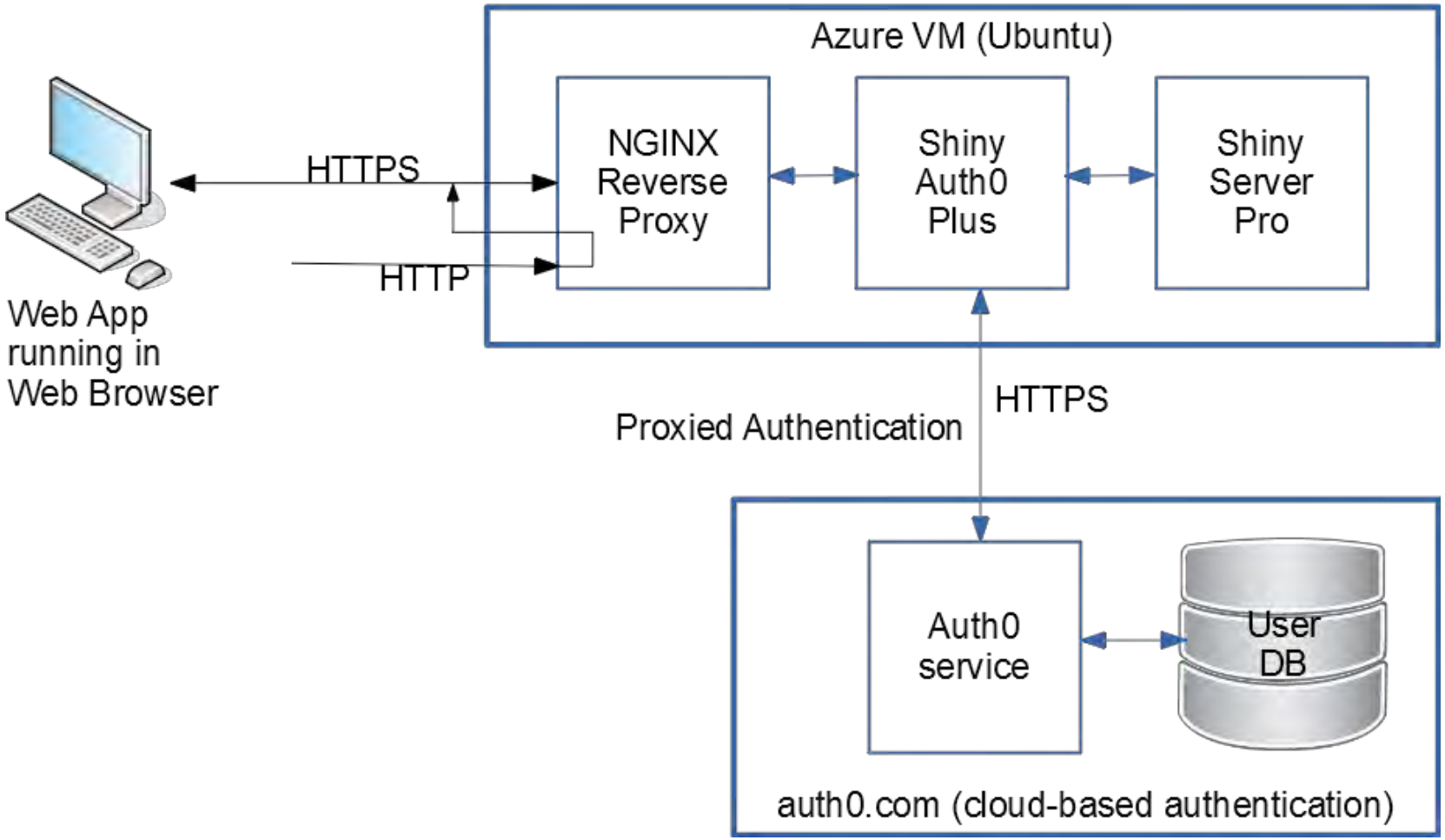
Public Internet – Commercial Application – High Data Security

Gen 1 deployment



Public Internet – Commercial Application – High Data Security

Gen 2 deployment



Questions & Answers

Milliman

Thank you

Guy Yollin
guy.yollin@milliman.com